Macquarie University Mahidol University

# Three Essays on Investor Confidence: A New Measure, Empirical Application and Experimental Evidence

*by*

Christoph Meier

*Supervisors:*

Dr. Roy Kouwenberg†

Dr. Lurion DeMello‡
Dr. Andreas Hellmann‡

*A thesis submitted in partial fulfilment of the requirements for the degree of*

Doctor of Philosophy

*in a cotutelle agreement between*

†CMMU, College of Management
Mahidol University
Bangkok, Thailand

‡Faculty of Business and Economics
Macquarie University
Sydney, Australia

27 March 2017

# Declaration of the candidate

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person, nor material that to a substantial extent has been accepted for the award of any other degree or diploma at any other educational institution, except where due acknowledgement is made in the thesis. Any contribution made to the research by colleagues with whom I have worked elsewhere during my candidature is fully acknowledged.

The work presented in this thesis has been submitted to Macquarie University in Sydney, Australia, and Mahidol University in Bangkok, Thailand, and supported by the joint PhD supervision (cotutelle) agreement.

I agree that this thesis be accessible for the purpose of study and research in accordance with the normal conditions established by the Executive Director, Library Services or nominee, for the care, loan and reproduction of theses.

Sydney, March 2017

Christoph Meier

# Acknowledgements

Firstly, I would like to express my sincere gratitude to my advisors. I thank Assoc. Prof. Roy Kouwenberg and Dr. Lurion DeMello for their continuous support of my PhD studies and related research, and for their patience, motivation and immense knowledge. I could not have imagined having better advisors and mentors for my PhD journey.

My sincere thanks also go to Prof. George Moschis for introducing me to the craft of inquiry and for adding enlightenment to my first research endeavour and to Prof. Stefan Trueck for much helpful advice. I further thank Assoc. Prof. Sue Wright, Assoc. Prof. Elaine Evans and Dr. Karen Handley who provided me with the opportunity to join their team for interesting research projects. I also thank all my other colleagues from Mahidol University and Macquarie University for their assistance and support even though there is insufficient room to add their names to this list. I thank Macquarie University for its financial support by awarding me a Macquarie University Research Excellence Scholarship (iMQRES) and a Postgraduate Research Fund (PGRF) grant.

I further thank my fellow PhD candidates, any of whom have become close friends over the years (special thanks to Adam, Tem, Martin and Schnacki), for all the fun we have had.

Last but not least, I would like to thank my family: my wife (+1) who has supported me immeasurably throughout this journey and who gave me strength and meaning to succeed both in my studies and in life; my parents and sister for supporting me in spirit throughout the writing this thesis; and my grandparents who have always shown interest in my research and have patiently listened to my subsequent monologues. My cats can't read: nevertheless, I thank them for waking me up at 5.00 a.m. every day which added extra productive hours to the day.

# Abstract

This PhD research is committed to contributing to the literature on investor overconfidence, one of the most robust findings in the field of behavioural finance. Overconfidence, a cognitive bias where decision makers tend to be overly optimistic not only about their aptitudes and skills, but also about the precision of their forecasts and information, is associated with poor decision making. Individuals suffering from overconfidence tend to be excessive stock traders, Chief Executive Officers (CEOs) who rush into mergers and acquisitions, risky drivers, naïve entrepreneurs and sloppy retirement planners.

The literature yields the many attempts to link stock market phenomena to overconfidence. However, existing measures that have been used to test these hypotheses are typically only loosely related to the overconfidence of investors in their own abilities, or use proxies that lack a formal model of cognitive psychology.

In the first of three research projects, I propose a measure of aggregate investor confidence that is based on a cross-disciplinary model containing determinants of confidence. The measure captures major economic events intuitively, and is statistically distinct from existing proxies. Using a 1926–2014 United States (US) sample, I find that the new measure is a better predictor of aggregate trading activity than past stock returns, which have been used in prior studies.

The second research project explores the role of aggregate investor confidence in asset pricing factors. Empirical tests reveal interesting patterns. Firstly, and in line with a behavioural model by Daniel, Hirshleifer, and Subrahmanyam (1998), aggregate investor confidence partially explains variations in the profitability of momentum strategies. Additionally, aggregate investor confidence appears to play a key role in the size factor, complementing an early hypothesis by Roll (1981). Indeed, investors seem to systematically change their risk perceptions which ultimately impacts on market outcomes.

The third research project takes a qualitative stance. Using a new methodology proposed by Glaser, Langer, and Weber (2013), we utilise the ability to

assess time series variations of individual overconfidence levels in an experimental asset market. We find that arriving signals that strongly support prior decisions cause overconfidence to prevail, while strongly opposing signals cause the effect to vanish 'overconfidence crashes'. However, previously lost overconfidence can re-emerge when these opposing signals reverse.

Additionally, we find strong evidence in favour of the hypothesis by Hong and Stein (2007) which states that investors interpret arriving information differently with opposing feedback having particularly strong effects. We also find measurement bias in the methodology proposed by Glaser et al. (2013). This is consistent with methodological concerns documented by Langnickel and Zeisberger (2016) and Biais, Hilton, Mazurier, and Pouget (2005) who report that assessment tasks using confidence intervals typically yield inflated overconfidence scores, as individuals tend to be insensitive to confidence levels in their estimations.

# Contents

# List of Tables

# List of Figures

# Chapter 1

# Introduction

## 1.1 A brief evolution of behavioural finance

One of the most fundamental assumptions in neoclassical economics is that individuals are rational agents who make informed and utility-maximising decisions.

In the field of finance, investors are assumed to behave similarly. When in possession of relevant information, they are able to determine fair security prices depending on their future cash flows discounted by a rate according to an appropriate level of underlying risk. Consequently, those securities have to be fairly priced, as each utility-maximising agent would immediately identify any mispricing and increase demand for a security that is undervalued given its level of risk or, conversely, decrease demand for a security that it overvalued also given its level of risk.

Therefore, markets where those securities are traded have to be efficient, as first postulated by Fama (1970). In other words, security prices should not only be in accordance with underlying risk, but also be unpredictable and follow no pattern or 'random walk' (Samuelson, 1965; Fama, 1965).

An early attempt to determine appropriate security prices is the capital asset pricing model (CAPM) which was initially proposed by Sharpe (1964) and Lintner (1965). This simple model suggests a linear relationship between a stock's level of risk relative to the market (the stock's $\beta$) and its appropriate expected return.

In the two and a half decades since then, a multitude of "anomalies" has been reported in which particular stock trading strategies yield higher returns than CAPM would suggest. A vastly incomplete list of those anomalies includes the 'small-firm effect' (Banz, 1981; Keim, 1983), where small market-capitalisation stocks yield higher risk-adjusted returns, and the weekend-effect (French, 1980), with frequent negative stock returns at the beginning of the week and after holidays.

Consequently, the debate in the literature yielded three possible explanations for the existence of anomalies. Firstly, those strategies are indeed profitable but only retrospectively (French, 1980). Once those anomalies are discovered and known to a broader audience, they will disappear through the price-correcting trades of rational market participants. Secondly, existing models designed to capture the cross-section of stock returns fail to appropriately consider underlying risk and, therefore, 'abnormal' returns are simply due to risk not being accounted for in those models (Fama and French, 2004, 1996a). Thirdly, some of those anomalies may have alternative explanations, based on systematic flaws in the fundamental assumption of agents that are fully rational.

The first possible explanation was quickly confirmed for some of those anomalies. Soon after their discovery, some seasonal patterns seem to have vanished (Agrawal and Tandon, 1994; Schwert, 2003). An attempted remedy for the second possible explanation is the development of more sophisticated models to evaluate securities with these models intended to capture a security's risk to a fuller extent. Fama and French (1993) propose a simple three-factor model which increases the ability to explain cross-sectional variations in stock returns. The third potential explanation has laid the groundwork for nonrational perspectives incorporating concepts of bounded rationality.

Research following the spirit of the third notion, that is, that agents are perhaps not fully rational, has borrowed concepts from the field of cognitive psychology where the pioneers Tversky and Kahneman (1974) and Kahneman and Tversky (1979) provided early frameworks explaining human decision making under risk.

In subsequent years, much research was produced that applied concepts of bounded rationality to known phenomena. For instance, Shefrin and Statman (1985) demonstrate that investors tend to sell their well-performing stocks too early and their poorly-performing stocks too late which the authors attribute to regret avoidance. That is, investors strongly dislike realising losses as they would then trigger the sensation of regret.

However, one of the most prominent biases in human judgement is overconfidence which has been applied in many domains in economics. The following section discusses the relevance of the concept, and embeds the three research projects of this study conceptually.

## 1.2 The role of investor confidence

Overconfidence bias is one of the most prominent judgement biases with broad application in many disciplines. A rich body of literature suggests that individuals suffering from overconfidence tend to make decisions. For instance, Camerer and Lovallo (1999) and Koellinger, Minniti, and Schade (2007) find that on average, individuals are overly optimistic about the chance of success of their new business venture despite the knowledge that most entrepreneurs fail within the first few years. Similarly, managers tend to be too optimistic about the likelihood of the success of mergers and acquisitions, as well as the profitability of investments in general (Malmendier and Tate, 2005, 2008).

Barber and Odean (2001) and Odean (1998) find that traders are too optimistic about their abilities to pick stocks and, therefore, they trade too much.

Overconfidence consistently leads to adverse decision making which eventually causes significant damage. Not surprisingly, DeBondt and Thaler (1994a) (1994) conclude that overconfidence is "perhaps the most robust finding in the psychology of judgement [...]" (p.6).

Despite the huge body of literature, little evidence exists on how individuals calibrate their confidence after feedback arrives on their potentially poor decision making, especially after the extreme misalignment of confidence. In other words, does the overconfident entrepreneur return to an "appropriate" level of confidence after the failure of his business, and how does the overconfident CEO adjust his level of confidence subsequent to unsuccessful takeovers? Naturally, these questions are difficult to test as testing would require the measurement of confidence before and after such events. Furthermore, the "real-world" environment rules out alternative determinants of confidence.

This notion motivates the third paper presented in this thesis. Utilising the properties of an experimental asset market, we explore the process of calibration that investors undergo after receiving either highly reinforcing or contradicting feedback about prior decisions. We find that initially overconfident investors radically calibrate their confidence levels after receiving strongly conflicting feedback which causes their overconfidence to disappear.

This finding intuitively aligns with the possibility that overconfidence may co-exist with another common behavioural bias, the self-attribution bias. In their studies over four decades ago, Langer and Roth (1975) and Bem (1965) find that individuals tend to update their level of confidence differently depending on whether they receive positive or negative feedback. Typically, they credit success to their own aptitude and skill but blame external factors or bad luck for failure.

A combination of self-attribution bias and overconfidence among investors is intuitively appealing. In their behavioural model, Daniel et al. (1998) propose this notion as a possible explanation of the price momentum effect. If investors who actively trade stocks in a market interpret portfolio performance as feedback on their investment decisions, it is likely that they will interpret signals differently. When achieving high returns, they may credit their own abilities and conclude they are investment geniuses, even if those positive returns are shared among the entire market. If their portfolios bear losses, however, they are likely to blame external and unforeseeable circumstances for failure.

If investors follow this pattern, they should necessarily become increasingly overconfident over time which appears rather counter-intuitive A more intuitively appealing remedy may be 'confidence crashes' which cause overconfident investors to calibrate their level of confidence after receiving highly contradicting feedback. Chapter 4 addresses this notion.

Overconfident investors are associated with excessive trading (Odean, 1999, 1998) which has been documented both, not only in experimental settings (Glaser and Weber, 2007, 2009), but also on a macro-level (Statman, Thorley, and Vorkink, 2006).

Overconfidence is not only associated with trading activity, but also with other phenomena, such as price momentum. Daniel et al. (1998) propose a behavioural model to explain the effect. Testing such models is challenging as this requires an appropriate measure of aggregate investor confidence.

Many studies apply a range of 'proxies' of aggregate investor confidence, with these proxies summarised in table 1.1. However, it must be noted that these proxies capture different concepts of investor confidence. For instance, the American Association of Individual Investors (AAII) surveys their members' confidence that the stock market will perform well or poorly in the future on a weekly basis. Similarly, the University of Michigan and the Confidence Board both survey consumer confidence: these are targeted at capturing the degree of optimism or pessimism among consumers in their outlook on the economic environment.

**Table 1.1: Alternative measures of investor confidence and sentiment**

This table summarises proxies of investor confidence and sentiment used in the prior literature.

| Index | Name | Publication | Data | Data range | Data frequency |
|---|---|---|---|---|---|
| PRSTAT | Past returns | Statman et al. (2006) | | historical data | |
| BWSENT | Baker and Wurgler Investor Sentiment Index | Baker and Wurgler (2007) | NYSE data | 07/1965 – 12/2010 | Monthly |
| SHIL$_{1y}$ | Shiller One-Year Confidence Index | Shiller (2000a) | Survey data | 01/2001 – 12/2014 | Monthly |
| SHIL$_{cr}$ | Shiller Crash Confidence Index | Shiller (2000a) | Survey data | 01/2001 – 12/2014 | Monthly |
| SHIL$_{val}$ | Shiller Valuation Confidence Index | Shiller (2000a) | Survey data | 01/2001 – 12/2014 | Monthly |
| AAII | American Association of Individual Investors Investor Sentiment Survey | American Association of Individual Investors | Survey data | 01/2001 – 12/2014 | Weekly |
| MICHIGAN | Consumer Sentiment Index (University of Michigan) | Lemmon and Portniaguina (2006) | Survey data | 01/1978 – 12/2014 | Monthly |
| CONCONC | Conference Board Consumer Confidence Survey® | Lemmon and Portniaguina (2006) | Survey data | not included in this research | |

Shiller (2000b) surveys both individual and institutional investors about their confidence in the stock market: this includes surveying their confidence that the market has a bullish outlook, is fairly valued, will not crash or is currently not in a bubble. Similarly, Baker and Wurgler (2007) construct an investor sentiment index which includes the dividend premium; first-day initial public offering (IPO) returns and volume; closed-end fund discounts; and the equity share in new issues. The aim of the index is to capture how euphoric or sorrowful investors are about the current state of the stock market.

Many of the conceptual models that propose the role of investor confidence in explaining stock market phenomena require an interpretation of the investor as "one who overestimates the precision of his private information signal" (Daniel et al., 1998, p.1841). However, I argue that existing proxies do not accurately capture this notion, as they typically capture investors' confidence in their outlook on the economic environment, but not their ability to judge the precision of their private information. This distinction is critical, as existing measures aim to captures individuals' optimism or pessimism in changes in their environment, but not their *own* ability to assess the accuracy of information. In other words, I follow definitions of 'investor confidence' by Daniel et al. (1998); Odean (1998, 1999) and Statman et al. (2006) with the 'ability to estimate the precision of information' as a central assumption. In contrast, confidence in the changes in one's environment is referred to as 'optimism', 'pessimism' or 'sentiment' throughout this thesis.

Using the notion of confidence in on's *own* ability to assess information as a starting point, I propose a new measure of aggregate investor confidence in chapter 2. Based on a cognitive model of confidence formation proposed by Griffin and Tversky (1992), I construct a parsimonious measure that assumes that investors interpret stock returns as feedback, with this being similar to the recommendations by Odean (1999) and Gervais and Odean (2001). I find that the new measure is statistically distinct from the related confidence proxies summarised above and that it is a better predictor of trading activity than using past returns as a dependent variable (Statman et al., 2006).

Chapter 3 extends this notion and applies the measure to explore a range of hypotheses regarding the role of investor confidence and asset pricing. In line with the hypothesis of Daniel et al. (1998), I find that the new measure partially predicts variations in the profitability of momentum returns, and is a better predictor than lagged market states (Cooper, Gutierrez, and Hameed, 2004). Furthermore, the new proxy links investor confidence to the size premium, complementing a hypothesis by Roll (1981).

## 1.3 Structure of the thesis

This thesis is organised based on three research projects, with each project representing a chapter in the thesis, while also able to be individually used for publication in academic journals. However, all three projects in this thesis fall under a common theme: Investor confidence and overconfidence. Figure 1.1 illustrates the structure of this thesis.

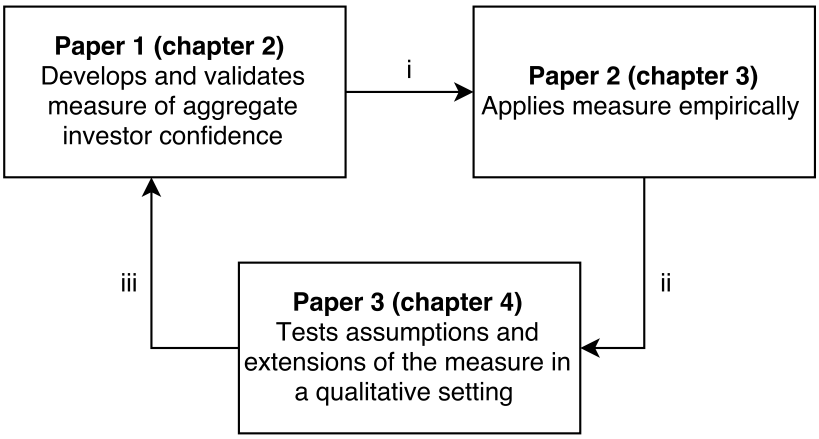



**Figure 1.1: Structure of the thesis**
Paper 1 develops and validates a new measure of aggregate investor confidence which is then used in a range of applications (Paper 2). Paper 3 extends the notion of Paper 1 and Paper 2, complementing prior quantitative findings with a qualitative perspective.

In the next chapter (chapter 2), I propose a new measure of aggregate investor confidence and test it against conceptually related proxies. In addition, I investigate its ability to predict trading activity and find that it is a better predictor than past returns, as used in prior studies (e.g. Statman et al., 2006).

Chapter 3 extends the analysis by investigating the role of aggregate investor confidence in a range of asset pricing applications. I find an association between lagged aggregate investor confidence and price momentum, supporting a behavioural model proposed by Daniel et al. (1998). Again, the new measure proposed in the current research is a better predictor of variation in momentum returns than those used in prior studies (e.g. Cooper et al., 2004). Additionally, and complementing other findings documented in chapter 2, the strong ability to predict size returns supports a hypothesis by Roll (1981).

Chapter 4 takes a qualitative stance. The motivation of the study is to explore the *actual* confidence calibration process of individuals after receiving feedback. Empirical studies using aggregate data fail to appropriately explain variations of *over*confidence, due to empirically measure the lack of an approach to measure *appropriate* level of confidence. We find that the severity of feedback is a relevant determinant of overconfidence. Strongly conflicting information can cause overconfidence to vanish, whereas strongly supportive information can boost overconfidence. We also find supportive evidence for hypotheses of Hong and Stein (2007) and Scheinkman and Xiong (2003) who suggest that *strong* information signals are associated with *high* investor disagreement.

The motivation of the current research reported in this thesis was to triangulate the theme—investor confidence—from different quantitative and qualitative perspectives. The following section summarises the key findings and contributions to the literature.

## 1.4 Contributions

### 1.4.1 Aggregate investor confidence in the stock market

The main contribution of the first project is that it provides a new measure of aggregate investor confidence that is:

1. Conceptually and statistically distinct to related proxies of investor confidence, such as
   (a) Index of Investor Sentiment Survey by the American Association of Individual Investors (AAII)
   (b) University of Michigan Consumer Sentiment Index
   (c) Robert Shiller's series of investor confidence indices, and
   (d) Investor Sentiment Index proposed by Baker and Wurgler (2007)
2. Based on a cognitive model by Griffin and Tversky (1992) using a cross-disciplinary approach.
3. Parsimonious and can be constructed from both current and historical data.
4. A better predictor of stock trading activity than past returns (Statman et al., 2006).

### 1.4.2 Aggregate investor confidence, price momentum and asset pricing

The second project contributes to the literature as follows:

1. Empirical findings support the behavioural model of Daniel et al. (1998), in which it is proposed that overconfidence in the accuracy of private information signals at least partially explains price momentum. Aggregate

investor confidence explains approximately 3–5% of the variation in the profitability of momentum returns.

2. Furthermore, the findings suggest a strong link between overconfidence and the size factor. I find that aggregate investor confidence accounts for approximately 5% of the variation of returns in the Fama-French size portfolio, supporting a hypothesis by Roll (1981).

3. Lagged investor confidence impulses provide insight to the anatomy of the relationship between investor confidence and pricing factors. Aggregate investor confidence affects momentum returns approximately 2–15 months after an impulse, with this being 1–12 months for the size factor, respectively.

4. The new index is a better predictor of variations in momentum returns than lagged market state, as used in a prior study (Cooper et al., 2004).

### 1.4.3 Investor overconfidence in experimental asset markets across market states

In contrast to the first and second projects, the third project contributes to the literature by producing the following qualitative insights:

1. We utilise the time-series properties of a methodology recently proposed by Glaser et al. (2013), finding that the arrival of extreme evidence has a strong impact on the within-subject variation of overconfidence:

   (a) Strong feedback opposing one's prior decision making causes overconfidence to vanish ('overconfidence crashes').

   (b) Supportive feedback, however, causes existing overconfidence to persist.

(c) Respondents who previously lost their overconfidence due to strongly opposing feedback became overconfident again after the arrival of feedback that strongly supported their initial decisions.

2. We find evidence complementing hypotheses by Hong and Stein (2007) and Scheinkman and Xiong (2003) who propose that time-series variance of stock prices can be interpreted as investor disagreement. They argue that, when strong information signals arrive, investors will interpret these signals differently, causing heterogeneous stock valuations and, hence, high volatility. When our experimental asset market experiences a crash, we find that variance in stock price predictions increases dramatically. The effect is particularly pronounced when these signals are in opposition to one's prior decision making.

3. In line with recommendations by Langnickel and Zeisberger (2016) and Biais et al. (2005), we find that the methodology by Glaser et al. (2013) conveys the same weakness as other methodologies that use confidence intervals to assess overconfidence. As respondents appear to fail to appropriately incorporate the stated confidence levels, those methods typically yield inflated overconfidence scores.

# Chapter 2

# Aggregate investor confidence in the stock market

**ABSTRACT**

Overconfidence is one of the most robust findings in the field of behavioural finance, and is associated with excessive trading and risk taking among market participants. Assessment of the level of confidence of individuals in their abilities and skills is well documented. However, the literature lacks an aggregate measure of investor confidence, with this required in order to test its implications on a macro-level. This paper introduces a simple measure of aggregate investor confidence by adopting a formal model of overconfidence. The applications of the measure suggest that, in aggregate, higher trading activity occurs when investor confidence soars, particularly for smaller stocks. Subsequently, the effect partially reverses, implying a correction to an initial overreaction. The newly introduced investor confidence index possesses better ability to predict trading activity than past returns, as used in prior studies. Additionally, investors tend to have a higher risk appetite when confident, as shown by increased investment in small stocks with higher risk.

*This paper was presented at the 2016 Financial Econometrics and Empirical Asset Pricing Conference in Lancaster, UK, and at the 2016 Macquarie University Faculty of Business and Economics (FBE) PhD workshop. I thank Roy Kouwenberg, Lurion DeMello, two anonymous referees, and participants of the 2016 Multinational Finance Society (MFS) conference conference in Stockholm, Sweden, the 2016 Macquarie University FBE PhD workshop seminar, and the 2016 Financial Econometrics and Empirical Asset Pricing Conference in Lancaster, UK for their valuable comments and insights.*

## 2.1 Introduction

This paper proposes a new measure of aggregate investor confidence. Starting with a simple theoretical model on the formation of confidence, a proxy of aggregate investor confidence in their *own* investment abilities is developed, based on feedback information available to all market participants. The measure adopts suggestions of Griffin and Tversky (1992) who identify two primary drivers of confidence: strength and weight of evidence. The strength of evidence is the magnitude of arriving feedback, whereas the weight of evidence is the reliability thereof.

The measure follows intuition that investors form beliefs about a 'typical' stock return given a market state, and interpret subsequent portfolio returns as feedback on prior investment decisions. For instance, if investors enjoy portfolio returns that are much higher than a typical return, the strength of evidence is high. As a consequence, the level of confidence in their own ability to pick stocks increases. If the distribution of strength signals is wide, some investors will have high gains, while others will incur losses. As investors suffer from self-attribution bias, they tend to credit success to their own aptitude and blame bad luck for failure (Miller and Ross, 1975). Consequently, winning investors increase their level of confidence, while those who are losing only lower their confidence partially, if at all. As a result, aggregate investor confidence is high when the strength of evidence is high and the weight of evidence is low, a condition also associated with overconfidence (Griffin and Tversky, 1992).

Several tests suggest that the investor confidence index introduced in this study is statistically distinct from existing measures, such as the Investor Sentiment Survey of the American Association of Individual Investors (AAII), the Baker and Wurgler (2007) Investor Sentiment Index, the University of Michigan Consumer Sentiment Index and a series of investor confidence indices proposed by Shiller (2000b). While those measures capture investor confidence and sentiment in an investment environment, the new measure proposed in this study aims to capture investor confidence in their *own* ability to invest in stocks.

Anecdotally, the new measure captures major booms and busts in the stock market in the last 90 years, with stark spikes around major stock market crashes and high confidence levels in bullish times.

The results of empirical applications show that aggregate investor confidence is positively associated with trading activity and aggregate risk appetite. In other words, high (low) aggregate investor confidence precedes high (low) trading activity and a larger (smaller) proportion of small stocks traded. The new measure is a better predictor of trading activity than past returns which were used as a proxy for investor overconfidence in prior studies (e.g. Statman et al., 2006).

The remainder of this paper is organised as follows. Section 2.2 summarises the related literature, while section 2.3 outlines and describes the motivation for the empirical approach. Section 2.4 summarises the data and intuitively inspects the newly introduced investor confidence index in its ability to reflect major historical events. Section 2.5 validates the measure and tests its applicability. Section 2.6 presents the conclusion.

## 2.2 Literature review

### 2.2.1 Confidence and overconfidence in finance

Investor overconfidence is one of the most prominent concepts in the field of behavioural finance (DeBondt and Thaler, 1994a), plausibly explaining many asset pricing patterns (Daniel and Hirshleifer, 2015). Lichtenstein and Fischhoff (1977) define an individual's confidence as his ability to assess the probability of a future event. In other words, individuals are able to make predictions about future events with some accuracy. However, many individuals tend to be overconfident which leads to an overly optimistic precision of their forecasts.

Overconfidence is applied in a multitude of domains in the field. (Malmendier and Tate, 2005) find that overconfident CEOs tend to invest too much and are overly optimistic about the success of mergers and acquisitions, which eventually harms firm value Malmendier and Tate (2008). Scheinkman and Xiong (2003) suggest that overconfident investors are at least partially responsible for speculative bubbles, and Daniel et al. (1998) suggest that investors are overconfident about the accuracy of private information which, at least to a certain degree, explains price momentum.

Overconfidence appears to be a shortcoming in the ability to assess one's own skill and the precision of knowledge, with this systematically leading to adverse outcomes. Therefore, it seems plausible that individuals would learn from prior mistakes and thus, with experience, the magnitude of the effect would be reduced (Gervais and Odean, 2001). However, the literature yields mixed evidence. Einhorn and Hogarth (1978); McKenzie, Liersch, and Yaniv (2008); Menkhoff, Schmidt, and Brozynski (2006); and Wagenaar and Keren (1986) suggest that overconfidence is pronounced in both experienced and inexperienced individuals and, thus, it is not necessarily cured by experience and self-reflection.

As a consequence, if overconfidence is a psychological human trait with no simple cure, it can be assumed that, in aggregate, some degree of overconfidence has to be present through time. Even if experience had a diminishing effect on the bias, a heterogeneous population of inexperienced and experienced individuals should preserve the effect.

### 2.2.2 Overconfidence in investor behaviour

Overconfidence has a well-documented effect on investor behaviour. Grinblatt and Keloharju (2009); Odean (1998, 1999); Statman et al. (2006); and Deaves, Lueders, and Luo (2008) find that overconfident investors tend to trade excessively which eventually hurts their portfolio performance.

Griffin, Nardari, and Stulz (2007) deliver robust evidence from international markets that trading activity, in aggregate, tends to increase when past market returns are high. Using a comprehensive sample of online brokerage data, Glaser and Weber (2007) find that overconfident investors not only tend to trade more, but also increase their risk appetite. In line with the notion of this study, Glaser and Weber (2009) report that increased trading activity is mainly due to recent high individual portfolio returns, rather than returns of the market. Deaves et al. (2008) produce evidence that is consistent with the studies above from their experimental study.

Despite rich evidence of individual investor overconfidence and trading behaviour, the literature lacks a macro-approach to quantify investor confidence in their *own* abilities to predict future prices. I argue that such an approach is compelling, as the literature yields much empirical evidence on the effect of investor sentiment on trading activity and asset pricing (e.g. Cooper et al. (2004), Statman et al. (2006) or Daniel et al. (1998)).

This study intends to fill this gap. An aggregate measure of investor confidence is introduced in order to discover behavioural patterns at a macro-level. The following section summarises existing models of investor sentiment and confidence and arguing that these measures are conceptually distinct from the measure introduced in this study.

### 2.2.3 Models of investor sentiment

Many attempts, using a wide range of confidence proxies, have been undertaken to explain phenomena in security markets. However, some studies use different types of confidence and assume them to be equivalent[1], or use proxies of confidence that are conceptually only vaguely related. Daniel et al. (1998) propose that price momentum occurs due to investors' overconfidence in the precision

[1] For a good discussion, see Moore and Healy (2008) and Glaser et al. (2013).

of their private information. Subsequently, Lemmon and Portniaguina (2006) use consumer confidence indices as a proxy of aggregate confidence but find no association with price momentum. One reason for this finding could be the inappropriate assumptions that unrelated concepts are, in fact, the same. Investors are unlikely to become more confident about the precision of their private stock market information if consumers, on average, become more optimistic about the state of the economy.

Similarly, a range of proxies and indices have been developed to capture investor confidence and sentiment, including the indices of Baker and Wurgler (2007), which measure the general degree of optimism and pessimism in the stock market. Shiller (2000a) develops a range of measures of investor attitudes: speculative bubble expectations and investor confidence; the perception among investors that 'nothing can go wrong'; as well as investor confidence that 'securities are fairly priced' at a given point in time.

However, these approaches aim to capture investor confidence and the outlook of an investment environment including the expected prosperity of the economy and the stock market. These approaches are difficult to align with formal overconfidence models, where overconfidence affects an individual's perceived ability to make forecasts based on his private information. These proxies may affect an investor's general degree of optimism in the economy, but not his investment skills.

More closely related to the rationale of this paper is the study by Statman et al. (2006) who investigate the relationship between past stock returns and current trading volume. They suggest that, due to biased self-attribution, some investors erroneously attribute returns to their own abilities to predict future prices, even though those returns are shared across the entire market. Cooper et al. (2004); Gervais and Odean (2001); and Odean (1998) conclude that average overconfidence across investors should be higher subsequent to market gains, as most investors hold long positions. Given that this assumption holds true, the aggregate level of investor confidence in their own skills to predict security prices should be higher in bull markets, but lower in bear markets.

The methodology of Statman et al. (2006) and Cooper et al. (2004) addresses the impact of past market returns, but does not formally measure confidence. That is, past returns seem to be related, but only vaguely consider the confidence formation process. This paper extends on this view. Similar to Statman et al. (2006) and Cooper et al. (2004), I assume that investors interpret portfolio returns as feedback on their prior investment decisions and, accordingly, update their level of confidence.

The purpose of this study is to provide a measure of aggregate investor confidence that is more closely related to the notion of investors' *own* abilities to predict future security prices. This allows more stringent application of theoretical models of investor confidence, including price momentum (Daniel et al., 1998), trading activity (Statman et al., 2006), or the size effect (Banz, 1981).

The following section summarises the empirical approach and describes the process used to develop the introduced measure of aggregate investor confidence.

## 2.3 Empirical approach

In the current study, the construction of a measure of aggregate investor confidence starts with a cognitive model of the determinants of confidence proposed by Griffin and Tversky (1992). According to the model, one's confidence is determined by arriving evidence which either supports or contradicts prior decisions or beliefs. This feedback has two components: strength and weight. Strength is the vividness of evidence, while weight is the reliability of evidence. In other words, investors should become more confident in their ability to trade stocks if their portfolios recently performed very well (high strength), but more so if those gains are hard to attribute to either their own abilities or simply to luck (low weight).

Griffin and Tversky (1992) note that individuals tend to be "highly sensitive to variations in the extremeness of evidence and not sufficiently sensitive to variations in its credence or predictive validity" (p.413). Analogously, investors should experience sharp changes in their levels of confidence after experiencing extreme changes in their portfolio returns. Griffin and Tversky (1992) quantify both effects: the size of an effect can be measured as the difference between means,with the weight being the standard error of evidence.

When applied to this domain, strength is measured by an investor's recent performance versus their long-term average performance. This approach follows the rationale that investors, on average, form expectations of a 'typical return' in accordance with the current market state. Thus, they 'anchor'[2] their expectations to a metric and accordingly form beliefs about a typical return. If new evidence (the most recent portfolio return) arrives, the strength impulse for that point in time is positive (negative) if that latest evidence is larger (smaller) than the 'typical return'.

Therefore, the strength variable $v$ can be interpreted as the extremeness of evidence and, hence, is defined as the difference in value-weighted market return in period $t$ (the latest month) and the average returns in a baseline period $u$, which an investor uses as an anchor in a given market state to define the 'typical' return. The baseline period return is computed as the average value-weighted market return over a given number of months.

To determine the length of the baseline period, durations of portfolio formation periods used in prominent underreaction and overreaction literature are adopted (Asness, Moskowitz, and Pedersen, 2013; Conrad and Kaul, 1993; Daniel and Moskowitz, 2013; DeBondt and Thaler, 1985, 1987; Fama and French, 2012; Jegadeesh and Titman, 1993, 2001). Consequently, baseline periods of 6, 12, 24 and 36 months are selected for further analysis.[3]

[2] In the spirit of Kahneman and Tversky (1974).

[3] In order to mitigate potential criticism of data mining, results for different baseline period

Thus, the strength variable of a stock in an index can be expressed as follows:

$$v_{i,t} = r_{i,t} - \overline{r}_{i,u,t}, \tag{2.1}$$

where $v_{i,t}$ is the strength impulse of stock $i$ in period $t$; $r_{i,t}$ is the return of stock $i$ in period $t$; and $\overline{r}_{i,u,t}$ is the simple moving average of stock $i$ in baseline period $u$, the look-back period used to compute the moving average at time $t$.

Griffin and Tversky (1992) suggest the standard error of the strength variable as a measure for the weight variable. The rationale is that if the level of strength impulses is high, more investors will either outperform or underperform the market. As investors suffer from self-attribution bias[4], outperforming investors will attribute recent strong performance to their own abilities, whereas underperforming investors will make external factors responsible for their failure. As a consequence, the confidence of outperforming investors increases, whereas underperforming investors only partially lower their level of confidence.

Following this rationale, the weight variable $w$ is computed as a function of the cross-sectional standard deviation of strength impulses for each security $i$ in an index $j$.

Therefore,

$$W_t = \frac{1}{\sqrt{\frac{\sum_{i=1}^{n}(v_{i,t}-\overline{v}_t)^2}{n-1}}}, \tag{2.2}$$

where $W$ is the weight of evidence in period $t$; $v$ is the strength of stock $i$ in the index; and $\overline{v}_t$ is the mean strength in the index in month $t$[5].

Subsequent to this step, a market equity-weighted mean score for the strength variable is computed for each month $t$:

durations are contrasted.

[4] In the spirit of Bem (1965), Langer and Roth (1975) and Miller and Ross (1975).

[5] It has to be noted that weight is defined as the reciprocal of cross-sectional standard deviation, as *high* standard deviation corresponds with *low* weight.

$$V_t = \sum_{i=1}^{n} \frac{m_{i,t} v_{i,t}}{\sum_{i=1}^{n} m_{i,t}}, \tag{2.3}$$

where $V_t$ is the value-weighted strength of all stocks in the index; and $m_i$ is the respective market value of a stock.

It must be noted that a *high* value of overconfidence corresponds with a *low* weight of evidence.[6] Therefore, the weight variable $w_t$ is the inverse of the stock strength dispersion. Consequently, aggregate overconfidence in a stock market in a given period of time should be *high* when $V_t$ is *high* and $W_t$ is *low*.

As all prior variables are computed with simple returns, natural logarithms of $(1 + V_t)$ and $(1 + W_t)$ are computed, that

$$ln(INVCON_t) = ln(1 + V_t) - ln(1 + W_t). \tag{2.4}$$

## 2.4 Data description

The data to compute the strength and weight proxy are from the monthly Center for Research in Security Prices (CRSP)-Compustat database. All common stocks from NYSE, AMEX and NASDAQ are included (share code 10 and 11) between July 1927 and December 2014. Holding period returns are provided by CRSP and defined as the change in total value of an investment in a common stock over a given period of time per dollar of initial investment, which includes dividend returns.

### 2.4.1 Control variables

Each proxy first undergoes stationarity tests which are typically passed after the first difference of the natural logarithm of a respective confidence/sentiment proxy or control variable. The detailed stationarity tests, including the Kwiatkowski–Philips-

[6] Griffin and Tversky (1992) state that overconfidence occurs "when strength is high and weight is low" (p. 414).

Schmidt–Shin (KPSS) test of trend-stationarity and the augmented Dickey–Fuller (ADF) test of a unit root, as well as time-series plots of all stationary variables are available upon request.

To isolate the confidence and sentiment components of the measures used in this study from business cycle components, I regress each stationary proxy on a range of macroeconomic variables, following the recommendations of Baker and Wurgler (2007). More specifically, these variables include growth in the industrial production index[7] and growth in consumer durables, non-durables, and services[8], as well as a dummy variable for National Bureau of Economic Research (NBER) recessions. I assume that the residuals of these regressions are largely free of major business cycle effects, as previously suggested by Baker and Wurgler (2007).

### 2.4.2 Alternative measures of investor confidence/ sentiment

In order to test the validity of the newly constructed investor confidence index (INVCON), it is examined for associations with other existing indices. The purpose of this test is to verify that, if controlled for macroeconomic variables that may influence investor sentiment in general, the index is conceptually and statistically distinct from related indices. These include the American Association of Individual Investors (AAII) individual investor confidence indices, consumer confidence indices used by Lemmon and Portniaguina (2006), the Baker and Wurgler (2007) investor sentiment index, as well as the Shiller (2000b) 1-Year, Valuation and Crash Confidence indices.

Since July 1987, the American Association of Individual Investors (AAII) has surveyed its members on a weekly basis in regard to their stock market

[7] The data on growth in the industrial production index is available in the Federal Reserve Statistical Release G.17.

[8] Data for consumer durables, non-durables, and services are available from the US Bureau of Economic Analysis (BEA) National Income Accounts Table 2.10.

expectations for the next six months. If a member expects a positive (negative) trend, he would respond ‘bullish’ (‘bearish’), and otherwise ’is neutral’. For the reason of parsimony, I only use the bullish component of the index, as it reflects the *proportion* of all surveyed investors with positive expectations about future stock market performance. As the survey explores AAII members’ expectations about the direction of the market in the upcoming months, it captures investors’ perception of the direction of the stock market, and not the individual’s confidence in their ability to identify profitable stocks.

Lemmon and Portniaguina (2006) explore the relationship between investor sentiment and the small stock premium, with consumer confidence as a proxy of investor optimism. These researchers analyse the University of Michigan Consumer Sentiment Index, as well as the Conference Board Index of Consumer Confidence, and treat the residual in the variation of the indices, which cannot be explained by a set of macroeconomic control variables, as consumer sentiment (optimism and pessimism). Due to the similar nature of the indices, the current study solely uses the University of Michigan Consumer Sentiment Index. As the index captures consumer expectations about present and future environmental conditions, it is conceptually distinct from the investor confidence index (INVCON) developed in this study.

Baker and Wurgler (2007) develop an investor sentiment index which reflects investor optimism and pessimism in the stock market, using a variety of proxies that intuitively relate to investor moods. These include trading volume, dividend premia, closed-end fund discounts or first-day IPO returns. For instance, if a firm’s first-day IPO return is high, it is possible that, on average, investors are very optimistic about its future outlook. The causal relationship points from confidence to first-day IPO returns. The same applies for trading volume, closed-end fund discounts or dividend premia. Similar to the AAII and University of Michigan Consumer Sentiment indices, the Baker and Wurgler (2007) investor sentiment index captures the mood or excitement of investors about the state of the stock market, and not their own investing abilities.

Shiller (2000b) develops a series of measures of investor attitudes among institutional and individual investors at six-monthly (and later monthly) intervals based on the results of questionnaires. In this study, the three most closely related indices are covered which are the One-Year Confidence Index, the Valuation Confidence Index and the Crash Confidence Index. The One-Year Confidence Index surveys respondents about their beliefs in developments in the Dow Jones in the coming year. Respective index scores reflect the percentage of the population expecting an increase. The rationale behind the investor confidence index proposed by Baker and Wurgler (2007) and the Shiller (2000b) One-Year Confidence Index is similar as they both measure the expectations and emotions of market participants, but neither measure investor confidence in their own ability to predict future security prices and, thus, they are conceptually different from the measure proposed in the current study.

The second index from Shiller (2000b) is the Valuation Confidence Index which captures the proportion of investors who think that current stock prices do not exceed their fundamental value. Thus, it reflects the proportion of the population who believe that the stock market is not overvalued which is closely related to bubble expectations. The third index borrowed from Shiller (2000b) is the Crash Confidence Index which measures the percentage of the population who attach little probability (less than 10%) to the occurrence of a catastrophic stock market crash in the next six months. Both the Valuation Confidence Index and the Crash Confidence Index follow a different rationale to the measure proposed in the current study.

Table 2.1 summarises alternative measures used for the validation process of the INVCON indices. As all alternative measures are conceptually different from the measure proposed in this study, null-hypotheses of zero correlation serve as a basis for index validation.

**Table 2.1: Alternative confidence measures**

The bullish component of the investor sentiment survey of the American Association of Individual Investors (AAII), the University of Michigan Consumer Confidence Index (Michigan), the Baker and Wurgler (2007) investor sentiment index (BWSENT), as well as the Shiller (2000b) 1-Year Confidence ($\text{Shil}_{1y}$), Valuation ($\text{Shil}_{val}$), and Crash Confidence indices ($\text{Shil}_{cr}$) are selected as alternative investor confidence measures in this study. This table summarises key index features, including data source, data range and data frequency.

| Index | Name | Publication | Data type | Data range | Availability |
|---|---|---|---|---|---|
| AAII | AAII Sentiment Survey | AAII | Survey data | 06/1987–12/2014 | Weekly |
| Michigan | UMich Consumer Sentiment Index | Lemmon and Portniaguina (2006) | Survey data | 01/1978–12/2014 | Monthly |
| BWSENT | Investor Sentiment Index | Baker and Wurgler (2007) | NYSE data | 07/1965–12/2010 | Monthly |
| $\text{Shil}_{1y}$ | Shiller One-Year Confidence Index | Shiller (2000b) | Survey data | 01/2001–12/2014 | Monthly |
| $\text{Shil}_{val}$ | Shiller Valuation Confidence Index | Shiller (2000b) | Survey data | 01/2001–12/2014 | Monthly |
| $\text{Shil}_{cr}$ | Shiller Crash Confidence Index | Shiller (2000b) | Survey data | 01/2001–12/2014 | Monthly |

### 2.4.3 Investor confidence index

Figure 2.1 below illustrates INVCON12 scores with a baseline period of 12 months, cleaned for macroeconomic control variables. In order to smooth the process, I apply a Hodrick-Prescott filter. However, raw scores are used for all empirical tests. The trend is intuitively appealing and matches anecdotal accounts of investor sentiment summarised by Baker and Wurgler (2007), Kindleberger (2000) and Malkiel (1999).

Consistent with the market gains attribution theory of Gervais and Odean (2001) and Odean (1999) that investors suffering from self-attribution bias tend to attribute market gains to their own investment talent, even if gains are shared with the entire market. That is, investors who observe *high* portfolio returns tend to erroneously credit their own ability to pick stocks, a phenomenon also known as the ego-preserving self-serving bias (Langer and Roth, 1975). Consistently, INVCON12 tends to be high during periods with high market returns, and low during periods with low market returns.

During speculative bubbles, INVCON12 scores consistently rise to relatively high levels. This applies for the months immediately before the Wall Street Crash of 1929, the Great Depression starting 1937, the ‘Tronics Boom in the early 1960s, the High-Tech New-Issues Bubble in the early 1980s, Black Monday in 1987, as well as the months before the burst of the Dotcom-Bubble around the turn of the millennium and the Global Financial Crisis (GFC)[9].

INVCON12 is thus in line with prior findings. Cooper et al. (2004), Gervais and Odean (2001) and Odean (1998) suggest that average (over)confidence across investors should be higher subsequent to market gains, as most investors hold long positions[10].

[9] It has to be noted that macroeconomic accounts discussed in this section have an illustrative purpose only. I do not argue causality and acknowledge that further steps could improve construct validity.

[10] It has to be noted that in aggregate, high confidence is only difficult to distinguish from overconfidence. That is, in order measure the level of overconfidence, one must first know an “appropriate” level of confidence that can be compared with an individual’s *actual* level of confidence. The difference would be considered over- or underconfidence. Glaser

Given that this assumption holds true, the aggregate level of investor confidence in their own skills to predict security prices should be higher in bull markets, and lower in bear markets.

## 2.5 Empirical tests

### 2.5.1 Validation

The first stage of the validation process is a series of correlation analyses of each alternative measure of investor confidence or sentiment on the investor confidence measure proposed in this paper. To control for macroeconomic effects and to allow comparison, raw scores of all indices are regressed on macroeconomic control variables.[11] The residuals of these regressions are assumed to be free of macroeconomic influences.

Figure 2.2 illustrates the correlation coefficients, as well as a scatter plot of INVCON12 and all alternative measures of investor confidence and sentiment. The investor confidence index with a baseline period of 12 months proposed in this study shows a weak positive correlation with the bullish component of the AAII index ($\rho$ =.06), thus rejecting the null hypothesis of no correlation.

The results of a weak positive association are therefore sound, as recent performance impulses may drive both investors' optimism about the direction of the stock market and changes in INVCON12. Correlation analyses between INVCON12 and the University of Michigan Consumer Sentiment Index, and the Baker and Wurgler (2007) Investor Sentiment Index, as well as all Shiller indices show zero correlation with INVCON12. Consequently, I assume that INVCON12 is both conceptually and statistically distinct from these alternative proxies.

et al. (2013) propose a methodology ("true overconfidence") as a remedy for this issue.

[11] These include an industrial growth index (Federal Reserve Statistical Release G.17), as well as growth of consumer durables, non-durables, and services (BEA National Income Accounts Table 2.10).

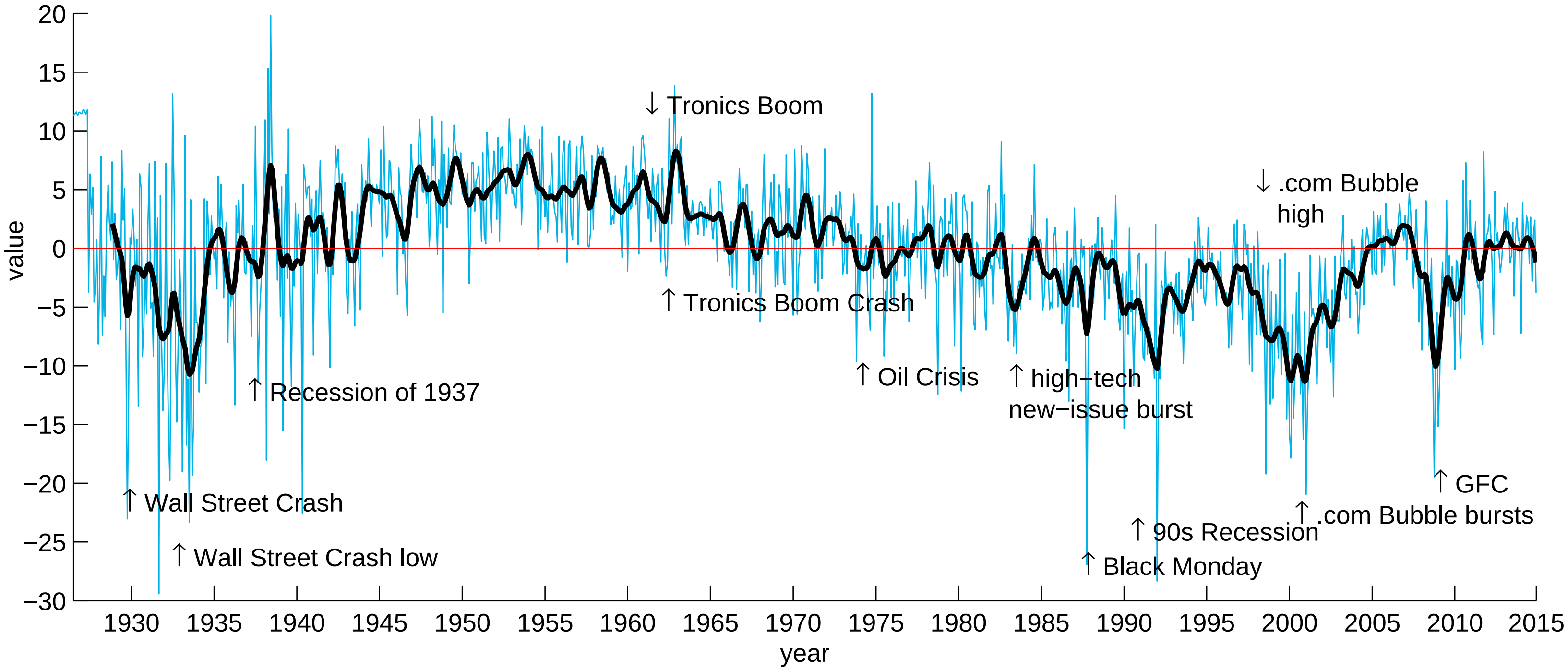



**Figure 2.1: Monthly values of the INVCON12 index**

This figure shows monthly orthogonalized scores of the INVCON12 index from July 1927 to December 2014. Raw values of the index are orthogonalised to a number of macroeconomic control variables. Following Baker and Wurgler (2007) I regress raw scores of the INVCON index on growth in the industrial production index (Federal Reserve Statistical Release G.17), growth in consumer durables, non-durables and services (all BEA National Income Accounts Table 2.8.5), growth in employment, changes in Consumer Price Index (CPI) (both US Bureau of Labor Statistics) as well as a NBER recession dummy. For the sole purpose of producing this diagram, raw scores are re-scaled (thin line) and a Hodrick-Prescott filter is applied to eliminate short-term fluctuations (thick line).

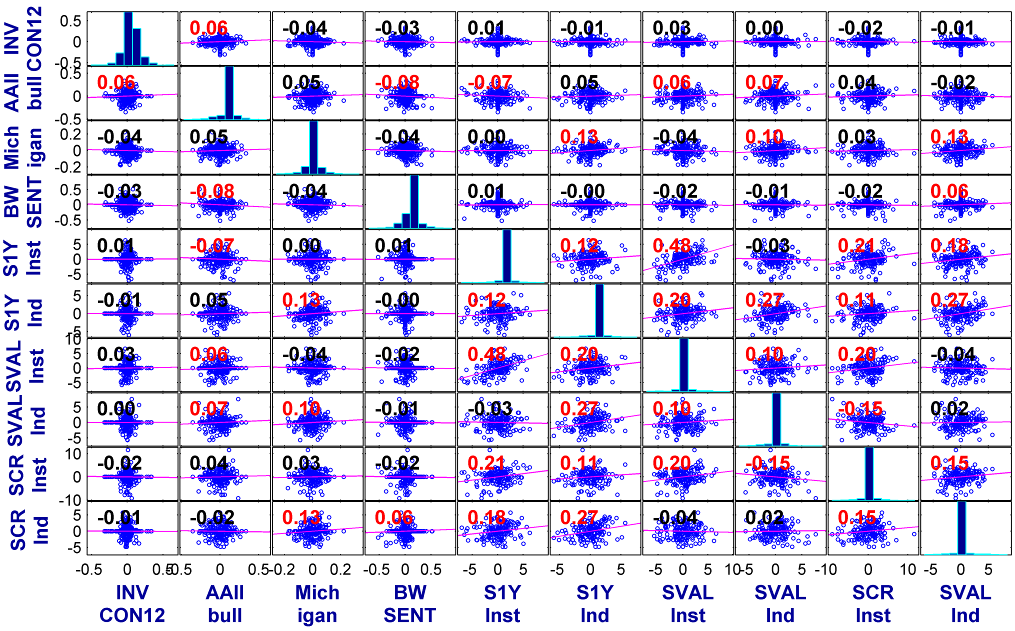



**Figure 2.2: Correlations between changes in INVCON12 and changes in all alternative confidence/sentiment proxies investigated in this study**

The figure shows the correlations between all investor confidence or sentiment proxies investigated in this study. Histograms of the variables appear along the matrix diagonal and scatter plots of variable pairs appear off-diagonal, respectively. The slopes of the least-squares reference lines in the scatter plots are equal to the displayed correlation coefficients, which appear red if significantly different from zero. The investor confidence index with a baseline period of 12 months (INVCON12) proposed in this study shows no significant correlation with most alternative proxies, except the bullish component of the AAII index. It must be noted that correlation coefficient values reported in the body of the text slightly deviate from those in this figure. The reason behind this deviation is the restriction of sample size in the individual analysis of an alternative investor sentiment proxy and INVCON. This plot uses a uniform sample length in order to allow full analysis.

In addition, I explore the relationships between all alternative proxies of investor sentiment/confidence using factor analysis. In line with the results of the correlation analyses, I find that INVCON12 shares a common factor with the bullish component of the AAII Investor Confidence Index.[12]

### 2.5.2 Empirical application

#### Investor confidence and trading activity

The second series of this empirical investigation are tests of the forecast ability of share turnover.[13] Prior literature suggests that, in aggregate, investors should increase trading activity when confidence is high and reduce trading activity when confidence is low. As the literature does not recommend a specific time frame for lead–lag effects, Statman et al. (2006) apply vector autoregression (VAR) impulse response functions. That is, they trace the effect of shocking one endogenous variable by one standard deviation on the current and future values of another endogenous variable, while controlling for exogenous variables. Positive responses of security turnover to market return shocks are found for lags of 1–4 months.

Thus, in order to compare the forecast ability of Statman et al. (2006), the same lags must be applied. Market return lags $l$ of 1–4 months are selected for both market and individual security turnover.

Lagged confidence measures are used to explain turnover in time period $t$ as follows:

$$Y_t = a + bX_{t-l} + e_t, \tag{2.5}$$

where $Y_t$ is the value-weighted turnover of an index; and $X_{t-l}$ is the value of the first difference of the logarithm of INVCON indices with baseline periods

[12] A rotated component matrix is presented in Table 2.C.1 in Appendix 2.C.

[13] Turnover is computed as the cross-sectional value-weighted ratio of shares of a company divided by the respective number of shares outstanding in a month.

of 6, 12, 24 and 36 months, controlled for macroeconomic variables. Respective INVCON values are lagged by $l = 1$–$4$ months; $a$ is a constant; $b$ is the respective coefficient; and $e_t$ is the random error term.

Baker and Wurgler (2007) suggest that small stocks are more affected by sentiment. Therefore, I further sort the sample by firm-size quintiles, measured by a firm's end-of-the-month market capitalization. Table 2.3 summarises a series of time series models for the first difference of the natural logarithm of value-weighted security turnover of all shares listed in AMEX, NASDAQ and NYSE between July 1927 and December 2014. In order to explore a size effect, the sample is sorted into month-end market equity firm-size quintiles. 'Smallest' refers to the bottom 20% of all firms listed in the three indices in a given month, and 'largest' is the top quintile, respectively. Independent variables are the log-differences of four lags of the Investor Confidence Index (INVCON) proposed in this study.

Investor confidence indices with baseline periods of 6 (six) and 12 months are generally better predictors of share turnover. However, all four versions of the INVCON index are successful in explaining up to 3% of monthly share turnover with the expected positive sign, with the exception of the largest firm size quintile. The effect is more pronounced for smaller firms, and is statistically significant with positive coefficients for up to two months after measuring investor confidence. The effect reverses after three months, being statistically significant for the smallest firm size quintile only, thus suggesting a market correction of initial overreaction for these small stocks which is a well-documented phenomenon (e.g. DeBondt and Thaler (1985) and Lo and MacKinlay (1990)). In other words, high aggregate investor confidence is associated with higher trading activity in the following months, when controlled for macroeconomic variables. The effect monotonously decreases with the number of lags and tends to be smaller for large firms, and reverses after the initial overreaction for those stocks where the effect was initially pronounced strongest.

**Table 2.2: Investor confidence and security turnover**

This table reports beta coefficient estimates, $t$-statistic significance values and adjusted $R^2$ values for time-series models with the first difference of logged market equity-weighted security turnover of all shares listed in AMEX, NASDAQ and NYSE between July 1927 and December 2014, sorted in month-end market equity firm-size quintiles. 'Smallest' refers to the bottom 20% of all firms listed in the three indices in a given month, and 'largest' to the top quintile, respectively. Independent variables are the first differences of logged four versions of the Investor Confidence Index proposed in this study. INVCON6 (IC6) is based on a baseline period of 6 months and INVCON36 (IC36) uses assumptions of a time period of 36 months for investors to form beliefs about 'typical' market returns at a given point in time, respectively.

| | | $1^{st}$ lag | | | $2^{nd}$ lag | | | $3^{rd}$ lag | | | $4^{th}$ lag | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Size | Coeff. | $t$-stat. | Adj. $R^2$ | Coeff. | $t$-stat. | Adj. $R^2$ | Coeff. | $t$-stat. | Adj. $R^2$ | Coeff. | $t$-stat. | Adj. $R^2$ |
| IC6 | smallest | 0.76*** | 4.61 | 1.89% | 0.74*** | 4.53 | 1.82% | -0.21 | -1.28 | 0.06% | -0.01 | -0.08 | 0.07% |
| | 2 | 0.86*** | 5.81 | 3.01% | 0.43*** | 2.91 | 0.70% | 0.09 | 0.59 | 0.00% | -0.20 | -1.32 | 0.03% |
| | 3 | 0.85*** | 6.04 | 3.26% | 0.33*** | 2.27 | 0.39% | 0.19 | 1.35 | 0.08% | -0.17 | -1.15 | 0.00% |
| | 4 | 0.47*** | 3.89 | 1.32% | 0.31*** | 2.56 | 0.53% | 0.00 | 0.00 | 0.00% | -0.05 | -0.39 | 0.00% |
| | largest | 0.18*** | 1.59 | 0.14% | 0.34*** | 3.06 | 0.79% | 0.01 | 0.07 | 0.00% | 0.01 | 0.05 | 0.00% |
| IC12 | smallest | 0.72*** | 4.58 | 1.87% | 0.65*** | 4.13 | 1.51% | -0.41*** | -2.61 | 0.55% | -0.01 | -0.07 | 0.09% |
| | 2 | 0.80*** | 5.73 | 2.94% | 0.36** | 2.56 | 0.53% | -0.07 | -0.47 | 0.00% | -0.20 | -1.41 | 0.05% |
| | 3 | 0.81*** | 6.06 | 3.29% | 0.23* | 1.68 | 0.17% | 0.06 | 0.41 | 0.00% | -0.17 | -1.24 | 0.00% |
| | 4 | 044*** | 3.76 | 1.24% | 0.26** | 2.20 | 0.37% | -0.10 | -0.85 | 0.00% | -0.07 | -0.63 | 0.00% |
| | largest | 0.16 | 1.52 | 0.12% | 0.33*** | 3.09 | 0.81% | -0.08 | -0.72 | 0.00% | 0.00 | -0.01 | 0.00% |
| IC24 | smallest | 0.68*** | 4.43 | 1.77% | 0.60*** | 3.91 | 1.36% | -0.36** | -2.33 | 0.42% | 0.00 | -0.01 | 0.04% |
| | 2 | 0.78*** | 5.62 | 2.86% | 0.33** | 2.40 | 0.46% | -0.07 | -0.47 | 0.00% | -0.17 | -1.19 | 0.04% |
| | 3 | 0.77*** | 5.78 | 3.03% | 0.22* | 1.67 | 0.17% | 0.07 | 0.50 | 0.00% | -0.16 | -1.19 | 0.00% |
| | 4 | 0.41*** | 3.58 | 1.13% | 0.26** | 2.27 | 1.13% | -0.11 | -0.93 | 0.00% | -0.02 | -0.19 | 0.00% |
| | largest | 0.12 | 1.17 | 0.04% | 0.32*** | 3.14 | 0.85% | -0.09 | -0.84 | 0.00% | 0.04 | 0.36 | 0.00% |
| IC36 | smallest | 0.67*** | 4.41 | 1.77% | 0.62*** | 4.05 | 1.48% | -0.36** | -2.36 | 0.45% | -0.09 | -0.61 | 0.14% |
| | 2 | 0.76*** | 5.58 | 2.86% | 0.35** | 2.56 | 0.54% | -0.10 | -0.75 | 0.00% | -0.22 | -1.56 | 0.13% |
| | 3 | 0.74*** | 5.64 | 2.92% | 0.24* | 1.76 | 0.21% | 0.04 | 0.31 | 0.00% | -0.21 | -1.54 | 0.00% |
| | 4 | 0.38*** | 3.41 | 1.02% | 0.26** | 2.30 | 0.42% | -0.12 | -1.08 | 0.02% | -0.07 | -0.62 | 0.00% |
| | largest | 0.10 | 1.00 | 0.00% | 0.33*** | 3.28 | 0.94% | -0.10 | -0.96 | 0.00% | -0.01 | -0.07 | 0.00% |

***, ** and * Statistically significant at a 99%, 95% and 90% confidence level, respectively

Unreported regressions show similar results when using raw INVCON scores which are not controlled for macroeconomic variables. This serves as a robustness check against potential look-ahead biases which could be caused by controlling for macroeconomic variables.

In order to test the persistence of the effect over time, I apply a two-way sort where the firm-size quintiles, based on end-of-the-month market equity breaks, are further divided into time sub-samples. Table 2.3 summarises the results.

The first time period spans the period between June 1927 and December 1949. One-month lagged INVCON indices significantly predict value-weighted security turnover, except for the largest firms.

Although all INVCON variations deliver significant results, indices with baseline periods of 6 to 24 months appear to capture aggregate investor perception of a 'typical' return better than longer time horizons, such as 36 months. This notion is intuitively appealing, as it seems likely that more recent events play a larger role than those in 'distant memory'. However, to the best of my knowledge, the literature does not yield clear recommendations for an approximate time frame for 'active investor memory'.

The second sub-sample begins in January 1950. It should be noted that while all other control variables are available from the beginning of the sample in July 1927, BEA accounts only capture monthly data from January 1959. Therefore, if the observed effect was due to the lack of control variables before 1959, it should appear in the second sub-sample. Quite strikingly, this is not the case. In the 1950s and 1960s, variations in the INVCON indices explain up to 13.5% of turnover for the smallest firms, and around 3–6% for the following three size quintiles. Again, the effect seems to be stronger for smaller firms, but remains highly significant for the second-largest size quintile.

Between January 1970 and December 1989, the explanatory power of INVCON indices to explain variations in stock turnover vanishes among the smallest stocks. This is quite surprising, as the effect was quite strong in the prior two decades. However, although weaker across all size quintiles, INVCON indices persistently explain some variation of share turnover for the three medium size quintiles.

However, in the subsequent two and a half decades, the explanatory power of variation of share turnover resumes to a level around 3%, with the smallest and largest size quintiles as exceptions. Therefore, the ability of aggregate investor confidence to partially explain trading activity is regarded as stable over time.

I further investigate the relative ability to forecast variations in value-weighted stock turnover, in comparison with the approach of Statman et al. (2006) who use past returns as a proxy for aggregate investor confidence with this notion conceptually related to the current study's rationale. Statman et al. (2006) argue that past returns relate to trading activity as investors who recently experienced high portfolio gains are more likely to develop a high degree of aggregate confidence. In order to test the lagged effects of past returns on trading activity, Statman et al. (2006) use impulse-response functions, finding significant results for up to one year.

As the effects reported by Statman et al. (2006) appear to be strongest for approximately four months, I construct a series of multiple linear regression models, with four lags of the first difference of the logarithm of INVCON12 as independent variables, and the first difference of the logarithm of value-weighted security turnover as a dependent variable, double sorted by firm size quintiles and four time horizons. In order to allow direct comparison, I adapt the methodology of Statman et al. (2006) by using the first four lags of value-weighted security returns of all stocks listed in AMEX, NASDAQ and NYSE as independent variables. In addition, I control for macroeconomic variables, as summarised earlier, in order to maintain consistent variable treatment.

Table 2.4 summarises the regression results. Based on their respective ad-

**Table 2.3: Investor confidence and security turnover, double sorted by time and firm-size**

This table reports beta coefficient estimates, $t$-statistic significance values and adjusted $R^2$ values for a series of time-series models, with the one-month lagged first difference of logged market equity-weighted security turnover of all shares listed on AMEX, NASDAQ and NYSE between July 1927 and December 2014, double sorted into month-end market equity firm-size quintiles and four time breaks. 'Smallest' refers to the bottom 20% of all firms listed in the three indices in a given month and 'largest' to the top quintile, respectively. Independent variables are the first differences of the four logged versions of the Investor Confidence Index (INVCON) proposed in this study.

| | | Jul. 1927–Dec. 1949 | | | Jan. 1950–Dec. 1969 | | | Jan. 1970–Dec. 1989 | | | Jan. 1990–Dec. 2014 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Size | Coeff. | $t$-stat. | Adj. $R^2$ | Coeff. | $t$-stat. | Adj. $R^2$ | Coeff. | $t$-stat. | Adj. $R^2$ | Coeff. | $t$-stat. | Adj. $R^2$ |
| IC6 | smallest | 0.82** | 2.42 | 1.74% | 2.43*** | 5.81 | 12.10% | 0.18 | 0.48 | 0.00% | 0.37* | 1.80 | 0.74% |
| | 2 | 0.97*** | 3.16 | 3.18% | 1.50*** | 3.75 | 5.20% | 0.62** | 2.06 | 1.35% | 0.53*** | 2.66 | 2.00% |
| | 3 | 1.03*** | 3.38 | 3.67% | 1.25*** | 3.15 | 3.60% | 0.54** | 2.13n | 1.47% | 0.56*** | 3.16 | 2.93% |
| | 4 | 0.40 | 1.56 | 0.52% | 0.96*** | 3.01 | 3.28% | 0.48** | 2.03 | 1.30% | 0.42** | 2.45 | 1.65% |
| | largest | 0.12 | 0.53 | 0.00% | 0.42 | 1.39 | 0.39% | 0.11 | 0.53 | 0.00% | 0.25 | 1.42 | 0.34% |
| IC12 | smallest | 0.74** | 2.30 | 1.57% | 2.42*** | 6.18 | 13.50% | 0.19 | 0.53 | 0.00% | 0.34* | 1.75 | 0.68% |
| | 2 | 0.87*** | 2.95 | 2.79% | 1.54*** | 4.09 | 6.20% | 0.62** | 2.17 | 1.53% | 0.50*** | 2.65 | 1.98% |
| | 3 | 0.96*** | 3.28 | 3.52% | 1.30*** | 3.47 | 4.42% | 0.52** | 2.18 | 1.55% | 0.53*** | 3.16 | 2.94% |
| | 4 | 0.33 | 1.36 | 0.31% | 0.90*** | 3.00 | 3.26% | 0.47** | 2.08 | 1.37% | 0.43*** | 2.67 | 2.01% |
| | largest | 0.10 | 0.46 | 0.00% | 0.40 | 1.40 | 0.41% | 0.07 | 0.36 | 0.00% | 0.25 | 1.54 | 0.46% |
| IC24 | smallest | 0.75** | 2.32 | 1.68% | 2.30*** | 5.95 | 12.64% | 0.07 | 0.21 | 0.00% | 0.33* | 1.73 | 0.67% |
| | 2 | 0.86*** | 2.89 | 2.78% | 1.50*** | 4.05 | 6.07% | 0.58** | 2.08 | 1.39% | 0.47** | 2.57 | 1.85% |
| | 3 | 0.92*** | 3.10 | 3.25% | 1.22*** | 3.30 | 3.98% | 0.49** | 2.10 | 1.41% | 0.49*** | 3.03 | 2.67% |
| | 4 | 0.30 | 1.21 | 0.18% | 0.88*** | 2.97 | 3.17% | 0.43* | 1.95 | 1.16% | 0.42*** | 2.67 | 2.02% |
| | largest | 0.05 | 0.24 | 0.00% | 0.37 | 1.30 | 0.29% | 0.06 | 0.30 | 0.00% | 0.21 | 1.31 | 0.24% |
| IC36 | smallest | 0.75** | 2.32 | 1.76% | 2.12*** | 5.48 | 10.88% | 0.06 | 0.17 | 0.00% | 0.36* | 1.93 | 0.91% |
| | 2 | 0.88*** | 2.95 | 3.05% | 1.30*** | 3.53 | 4.60% | 0.57** | 2.07 | 1.36% | 0.46** | 2.57 | 1.84% |
| | 3 | 0.90*** | 3.03 | 3.25% | 1.02*** | 2.76 | 2.71% | 0.47** | 2.02 | 1.28% | 0.51*** | 3.22 | 3.04% |
| | 4 | 0.28 | 1.14 | 0.12% | 0.70** | 2.39 | 1.94% | 0.40* | 1.86 | 1.02% | 0.43*** | 2.82 | 2.28% |
| | largest | 0.05 | 0.23 | 0.00% | 0.24 | 0.86 | 0.00% | 0.02 | 0.08 | 0.00% | 0.21 | 1.38 | 0.30% |

***, ** and * Statistically significant at a 99%, 95% and 90% confidence level, respectively

justed $R^2$ values, INVCON12 appears to have better ability to forecast trading activity than both raw and cleaned past returns, with very few exceptions. As a consequence, I conclude that INVCON12 is a better proxy for aggregate investor confidence than past returns.

Furthermore, I construct an encompassing model to explain variations in trading activity by adding both the first four lags of the first difference of the logarithm of INVCON12, as well as the first four lags of value-weighted security returns of all stocks listed in AMEX, NASDAQ and NYSE as independent variables. Consequently, I use eight independent variables for an encompassing model. The purpose of this analysis is to identify the forecast ability of related proxies.

Table 2.5 summarises the output of the encompassing regression which suggests that lags two, three and four of the past returns variable become insignificant when including lags of INVCON12. Quite interestingly, the estimated coefficient of the first lag of the past returns variable changes its sign to negative. In line with the findings reported in Table 2.4, I assume that INVCON12 is a better proxy for aggregate investor confidence than past returns, at least for the case in aggregate trading activity.

### Investor confidence and risk attitude

Prior findings suggest that when individuals become overconfident, they tend to underestimate risk which leads to increased trading activity (Barber and Odean, 2001; Odean, 1998, 1999), too many new business ventures (Koellinger et al., 2007) and overly optimistic managerial decision making (Malmendier and Tate, 2005, 2008). In other words, individuals seem to systematically underestimate risk when overconfident.

This study proposes a time-series measure of aggregate investor confidence. However, it only tracks variations of aggregate investor confidence *without* specifically contrasting a high degree of confidence from overconfidence. Nev-

**Table 2.4: Investor confidence, past returns and security turnover**

This table reports $F$-statistic significance values, $p$-values and adjusted $R^2$ values for a series of multiple time-series models, with the first difference of logged market equity-weighted security turnover of all shares listed on AMEX, NASDAQ and NYSE between July 1927 and December 2014, double sorted in month-end market equity firm-size quintiles and four time breaks. 'Smallest' refers to the bottom 20% of all firms listed in the three indices in a given month, and 'largest' to the top quintile, respectively. Independent variables are the first four lags of the first differences of the logged Investor Confidence Index (INVCON) with a baseline period of 12 months. The second set of independent variables are raw values of market equity-weighted past returns (PR) of AMEX, NASDAQ and NYSE, following the rationale of Statman et al. (2006). The third set of independent variables are the same past returns (PR) as in set two, cleaned for macroeconomic control variables.

| | | $1^{st}$ lag | | | $2^{nd}$ lag | | | $3^{rd}$ lag | | | $4^{th}$ lag | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Size | $F$-stat. | $p$-value | Adj. $R^2$ | $F$-stat. | $p$-value | Adj. $R^2$ | $F$-stat. | $p$-value | Adj. $R^2$ | $F$-stat. | $p$-value | Adj. $R^2$ |
| IC12 | smallest | 8.43*** | 0.00 | 10.09% | 11.64*** | 0.00 | 15.17% | 2.97** | 0.02 | 3.20% | 2.83** | 0.02 | 2.40% |
| | 2 | 8.23*** | 0.00 | 9.83% | 5.38*** | 0.00 | 6.86% | 4.54*** | 0.00 | 5.61% | 4.44*** | 0.00 | 4.41% |
| | 3 | 7.82*** | 0.00 | 9.33% | 3.78** | 0.01 | 4.46% | 4.67*** | 0.00 | 5.80% | 5.34*** | 0.00 | 5.50% |
| | 4 | 3.50** | 0.01 | 3.63% | 2.55** | 0.04 | 2.53% | 3.40** | 0.01 | 3.88% | 3.05** | 0.02 | 2.67% |
| | largest | 4.46*** | 0.00 | 4.97% | 1.08 | 0.37 | 0.14% | 1.15 | 0.33 | 0.26% | 0.96 | 0.43 | 0.00% |
| PR | smallest | 7.83*** | 0.00 | 8.89% | 9.89*** | 0.00 | 13.00% | 3.75** | 0.01 | 4.41% | 4.32*** | 0.00 | 4.27% |
| raw | 2 | 7.76*** | 0.00 | 8.80% | 3.19** | 0.01 | 3.56% | 4.23*** | 0.00 | 5.14% | 3.76** | 0.01 | 3.57% |
| | 3 | 8.62*** | 0.00 | 9.82% | 2.16* | 0.07 | 1.92% | 3.03** | 0.02 | 3.30% | 3.97*** | 0.00 | 3.83% |
| | 4 | 2.83** | 0.03 | 2.55% | 1.27 | 0.28 | 0.45% | 1.61 | 0.17 | 1.02% | 1.60 | 0.18 | 0.79% |
| | largest | 2.09* | 0.08 | 1.53% | 0.31 | 0.87 | 0.00% | 1.16 | 0.33 | 0.27% | 1.27 | 0.28 | 0.36% |
| PR | smallest | 7.77*** | 0.00 | 8.81% | 11.32*** | 0.00 | 14.79% | 3.27** | 0.01 | 3.67% | 4.70*** | 0.00 | 4.73% |
| cleaned | 2 | 7.76*** | 0.00 | 8.81% | 3.83*** | 0.00 | 4.54% | 3.90*** | 0.00 | 4.65% | 4.02*** | 0.00 | 3.89% |
| | 3 | 8.60*** | 0.00 | 9.79% | 2.50** | 0.04 | 2.46% | 2.84** | 0.02 | 3.01% | 3.85*** | 0.00 | 3.68% |
| | 4 | 2.77** | 0.03 | 2.46% | 1.48 | 0.21 | 0.80% | 1.41 | 0.23 | 0.69% | 1.71 | 0.15 | 0.94% |
| | largest | 2.07* | 0.09* | 1.50% | 0.27 | 0.90 | 0.00% | 0.95 | 0.44 | 0.00% | 1.32 | 0.26 | 0.43% |

***, ** and * Statistically significant at a 99%, 95% and 90% confidence level, respectively

**Table 2.5: Encompassing regression: Investor confidence and past returns**

This table illustrates the output of an encompassing regression with the first difference of the logarithm of value-weighted security turnover as a dependent variable and four lags of both INVCON12 and past returns (PR) as independent variables.

| | Lags | Coeff. | $t$-stat. |
|---|---|---|---|
| IC12 | 1 | 1.10*** | 4.62 |
| | 2 | 1.03*** | 3.82 |
| | 3 | 0.44* | 1.70 |
| | 4 | 0.24* | 1.70 |
| PR | 1 | -0.46** | -2.2 |
| | 2 | 0.24 | 1.06 |
| | 3 | 0.10 | 0.43 |
| | 4 | -0.12 | -0.57 |

***, ** and * Statistically significant at a 99%, 95% and 90% confidence level, respectively.

ertheless, the hypothesis of high confidence being associated with altered risk perception is yet to be verified. Therefore, I test if high levels of confidence translate into changed risk perception. In other words, in aggregate, does investor risk appetite change with variations in confidence?

In order to test this notion, I identify stocks that are considered more risky. Fama and French (1992) introduce a size factor to explain cross-sectional variation in security returns. If small firms are considered more risky and highly confident investors have higher risk appetite, then aggregate preference towards small stocks should be higher. Consequently, I construct a SMALLS ratio which is defined as the proportion of trading volume of small stocks as a share of trading volume of the entire market. Small firms are defined as those in the two bottom market equity quintiles at the end of a month. Therefore,

$$SMALLS = \frac{\sum_{i=1}^{n} volume_{small,i,t}}{\sum_{i=1}^{n} volume_{i,t}} \tag{2.6}$$

where $volume_{small}$ is the trading volume of all small stocks $i$ in month $t$; and $volume$ is the trading volume of all stocks $i$ in month $t$. The first difference of SMALLS meets stationarity assumptions.

**Table 2.6: SMALLS time-series regression**

This table reports beta coefficient estimates, $t$-statistic significance values and adjusted $R^2$ values for a series of time-series models, with the first difference of the SMALLS ratio, which captures the proportion of the volume traded in small stocks over the volume traded in all stocks listed on AMEX, NASDAQ and NYSE between July 1927 and December 2014. 'Small' stocks refers to the bottom 40% of all firms listed in the three indices in a given month, sorted on month-end market equity. Independent variables are the first differences of the four logged versions of the investor confidence index (INVCON) proposed in this study.

| | $1^{st}$ lag | | | $2^{nd}$ lag | | | $3^{rd}$ lag | | | $4^{th}$ lag | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Coeff. | $t$-stat. | Adj. $R^2$ | Coeff. | $t$-stat. | Adj. $R^2$ | Coeff. | $t$-stat. | Adj. $R^2$ | Coeff. | $t$-stat. | Adj. $R^2$ |
| INVCON6 | 0.068*** | 4.68 | 1.95% | 0.010 | 0.69 | 0.00% | -0.002 | -0.14 | 0.00% | -0.005 | -0.33 | 0.00% |
| INVCON12 | 0.065*** | 4.75 | 2.01% | 0.008 | 0.58 | 0.00% | -0.007 | -0.49 | 0.00% | -0.006 | -0.41 | 0.00% |
| INVCON24 | 0.067*** | 4.97 | 2.24% | 0.005 | 0.38 | 0.00% | -0.004 | -0.30 | 0.00% | -0.006 | -0.45 | 0.00% |
| INVCON36 | 0.066*** | 4.92 | 2.21% | 0.006 | 0.44 | 0.00% | -0.004 | -0.30 | 0.00% | -0.009 | -0.67 | 0.00% |

***, **, * Statistically significant at a 99%, 95% and 90% confidence level, respectively.

Table 2.6 reports the output for time-series regression, with the SMALLS ratio as a dependent variable and changes in the four INVCON variables as independent variables for up to four lags. Quite strikingly, the month following high investor confidence impulses is associated with higher relative volume of small stocks, which this robust across all four versions of INVCON indices. However, the effect disappears in the second month and the months that follow.

The results support the notion that higher investor confidence is associated with higher risk appetite (Glaser and Weber, 2007). In other words, if aggregate investor confidence is high, investors seem to develop a higher risk appetite, and vice versa.

## 2.6 Conclusion

The notion that investor overconfidence affects market outcomes is a robust finding in prior research. However, this paper specifically quantifies aggregate investor confidence which is a crucial condition for its applicability in formal overconfidence models.

In my first key finding, I document that the model of aggregate investor confidence proposed in this study, which is based on an adaptation of a formal overconfidence model by Griffin and Tversky (1992), is conceptually and statistically distinct from existing investor confidence and sentiment measures. Essentially, while existing measures of confidence and sentiment capture investors' excitement, optimism or comfort about their investment environment, the new measure introduced in this study intends to capture impulses affecting investors' beliefs about their ability to predict security prices or, in other words, the ability to accurately process and forecast information.

The second key finding is that, in aggregate, higher levels of investor confidence are associated with higher security turnover for around two months. The effect appears to be more pronounced for small stocks and partially reverses in the third month. This finding is therefore sound as an initial 'overreaction' is likely to be eventually corrected by rational traders.

In prior studies, past returns (PR) are used as a proxy for aggregate investor overconfidence (e.g. Statman et al., 2006). I show in empirical tests that the new index of aggregate investor confidence is a better predictor of trading activity than lagged past returns (PR).

The third key finding is that, in aggregate, confident investors not only tend to trade more, they also appear to have a higher risk appetite, with this documented by the increase in the relative volume of small stocks traded when confidence is high. In line with prior findings, investors seem to systematically alter their risk perception in tandem with their level of confidence.

The new measure of aggregate investor confidence offers a broad range of testable implications, including further exploration of the effect of investor confidence on risk appetite, as well as applications to explain price momentum and tests of self-attribution bias (i.e. isolating positive from negative confidence impulses).

# Appendix

## 2.A Stationarity tests

### 2.A.1 Control variables

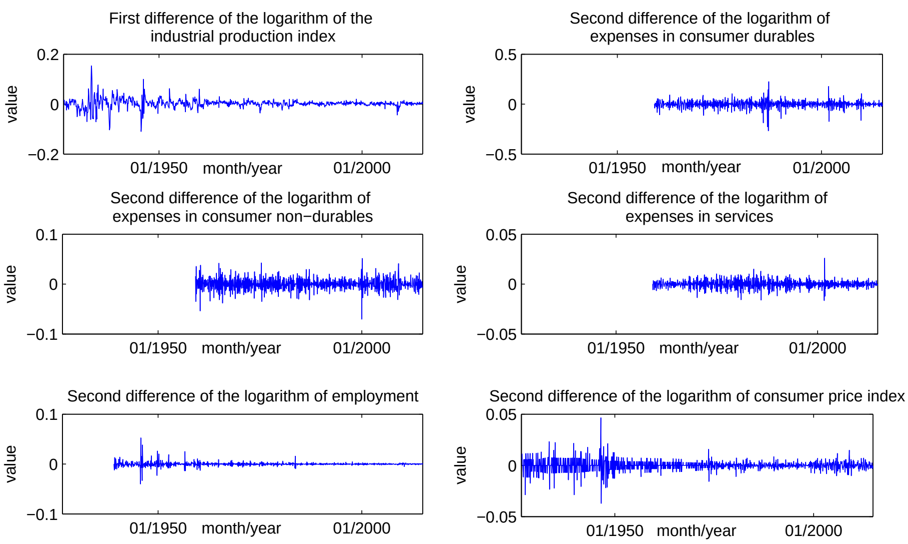



**Figure 2.A.1: First/second difference of control variables**

## 2.A.2 INVCON index with baseline periods of 6, 12, 24 and 36 months

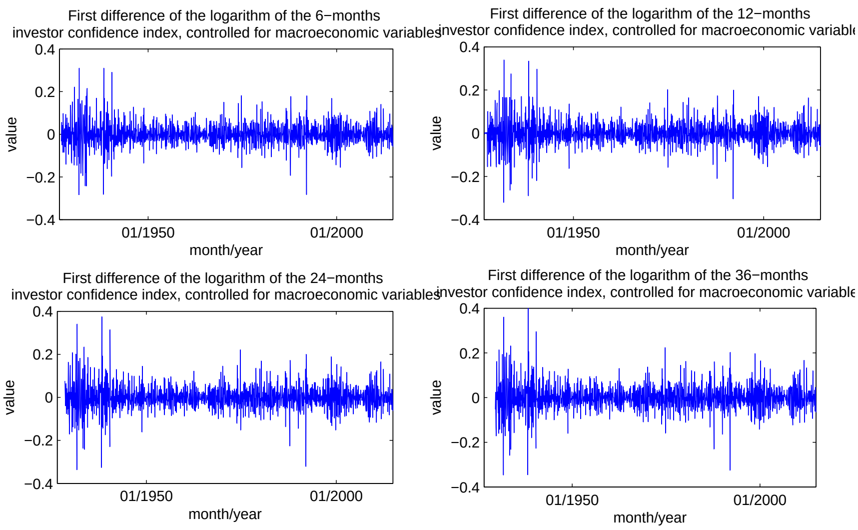



**Figure 2.A.2: First difference of aggregate investor confidence measures**

### 2.A.3 Alternative measures of investor confidence/ sentiment

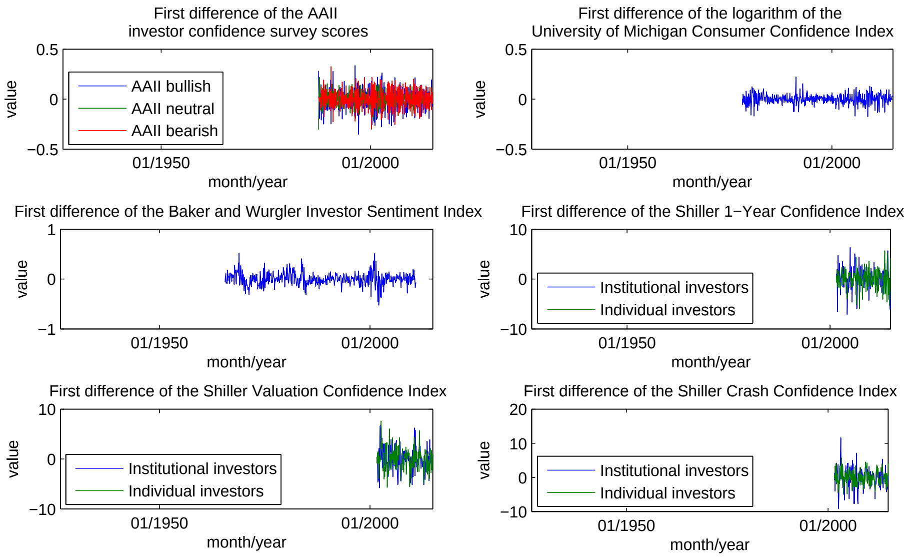



**Figure 2.A.3: First difference of alternative measures of investor confidence/ sentiment**

### 2.A.4 SMALLS ratio

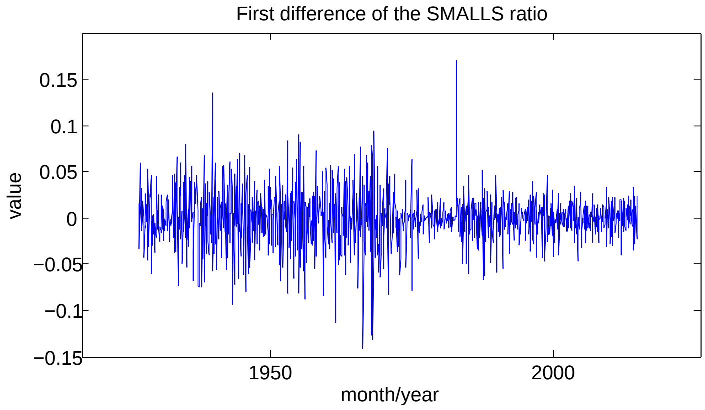



**Figure 2.A.4: First difference of the SMALLS ratio**

## 2.B Construction of the INVCON index

This list can be used as a step-by-step guide to the construction of the investor confidence (INVCON) index

1. I compute simple time-series moving average of past returns from $t - n$ periods to $t - 1$ period, where $n$ is the lookback period. This paper uses lookback periods of 6 to 36 months.

2. I then isolate the most recent return for period $t$.

3. I take the difference between the most recent return in period $t$ and the respective moving average computed in step 1, which results in the item-specific strength impulse $v_t$.

4. I compute the standard deviation of item-specific strength impulses, described in step 3, which results in the weight variable, for month $t$.

5. I compute a weighted average (based on market equity weights in period $t$) for item-specific strength impulses $v_t$ described in step 3, which results in the strength variable $V$ for period $t$.

6. I take the natural logarithm and compute the ratio as follows:
   $ln(INVCON_t) = ln(1 + V_t) - ln(1 + W_t)$.

7. To achieve stationarity, I take the first difference of INVCON indices.

8. To clean the index from macroeconomic factors, I run time-series regressions, with INVCON indices as a dependent variable, and stationary control variables as independent variables.

   I store the residuals of these regressions which are assumed to be free from macroeconomic effects.

   A full list of control variables, including their respective sources, as well as the process of achieving stationarity, is summarised above. Typically, the first difference of the natural logarithm leads to stationarity.

The Kwiatkowski–Philips–Schmidt-Shin (KPSS) test and the augmented Dickey–Fuller (ADF) test are examples of formal stationarity tests.

Please note: All alternative measures of investor sentiment/confidence are collected in their raw form and subsequently undergo the same cleaning procedure.

## 2.C Factor analysis: INVCON index and related measures

**Table 2.C.1: Rotated component matrix**

This table reports factor loading scores after performing factor analysis with all of the alternative investor sentiment/confidence proxies after achieving stationarity. I use the VARIMAX rotation method with Kaiser normalisation. The Shiller (2000b) One-Year Valuation Index for institutional investors seems to share a common factor with the One-Year Valuation Index for institutional investors. All individual Shiller Investor Confidence Measures, as well as the University of Michigan Consumer Sentiment Index appear to share a common factor. However, Shiller's Valuation Index for individual investors seems to be inversely related to the Shiller Crash Confidence Index for institutional investors. The Baker and Wurgler (2007) Investor Sentiment Index shares some communalities with Shiller's Crash Confidence Index for individual investors. INVCON12 loads on a common factor with the bullish component of the AAII Investor Confidence index.

| | Component | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 |
| INVCON12 | | | .817 | | |
| $AAII_{bull}$ | | | .849 | | |
| Michigan | | .671 | | | |
| BWSENT | | | | | .790 |
| $Shiller_{1Y}$ (inst.) | .859 | | | | |
| $Shiller_{1Y}$ (ind.) | | .749 | | | |
| $Shiller_{val}$ (inst.) | .868 | | | | |
| $Shiller_{val}$ (ind.) | | .554 | | -.600 | |
| $Shiller_{cr}$ (inst.) | | | | .840 | |
| $Shiller_{cr}$ (ind.) | | .366 | | | .678 |

# Chapter 3

# Aggregate investor confidence, price momentum and asset pricing


**ABSTRACT**

This paper applies a new measure of aggregate investor confidence which extracts feedback impulses from stock market data. According to the measure, aggregate investor confidence is positively associated with the profitability of momentum strategies. In a 1927–2014 US sample, aggregate investor confidence requires around three months to notably affect market outcomes and remains statistically significant for up to 16 months. Aggregate investor confidence can also partially explain the size premium, in line with conceptual accounts from prior literature. In aggregate, investors tilt their preference toward small market capitalisation and growth stocks when confidence is high. In contrast to price momentum, aggregate investor confidence immediately affects the size premium.



*This paper was presented at the 2016 Annual Meeting of the Academy of Behavioral Finance and Economics in Las Vegas, USA, and the 2016 Conference of the Multinational Finance Society (MFS) at the Stockholm University Business School, Stockholm,. I thank Roy Kouwenberg, Lurion DeMello, three anonymous referees, and participants of the 2016 MFS conference in Stockholm of the 2016 Annual Meeting of the Academy of Behavioral Finance and Economics in Las Vegas for their valuable comments and insights. This paper was awarded a grant from the 2016 Macquarie University Postgraduate Research Fund (PGRF).*

## 3.1 Introduction

Momentum strategies buy past winners and sell past losers, and persistently generate statistically and economically significant returns in international asset markets. The phenomenon is strong and persistent not only for equity, but also for other for asset classes, including bonds, commodities, currencies or exchange-traded futures.

The literature yields various attempts to explain the phenomenon from a behavioural perspective.[1] One testable model was proposed by Daniel et al. (1998) in which momentum forms through the overconfidence of investors. However, little empirical evidence exists that links price momentum to aggregate investor overconfidence, largely due to the lack of an appropriate measure thereof.[2]

In order to fill this gap, I apply the new measure of aggregate investor confidence proposed in chapter 2, showing that the measure is statistically distinct from existing measures of investor sentiment and partially explaining variations in aggregate trading activity[3] and risk appetite.

Applying the new measure, I investigate the role of aggregate investor confidence in the profitability of price momentum. Lagged aggregate investor confidence, controlled for macroeconomic variables, explains variations in the profitability of momentum strategies for up to 16 months. Consistent with the overconfidence model of Daniel et al. (1998), momentum returns tend to be higher (lower) following time periods with high (low) aggregate investor confidence. While following a similar motivation, the new measure of aggregate investor confidence is a better predictor of variations in the profitability of momentum returns than market dummies, as used by Cooper et al. (2004).

[1] Section 2 gives a more detailed review of the literature.

[2] Cooper et al. (2004) define UP and DOWN market states and find that momentum returns are positive subsequent to UP markets, and negative after DOWN markets.

[3] Similarly, Statman et al. (2006) use impulse–response functions to identify lead–lag effects for investor overconfidence and stock trading activity

The investigation of lagged aggregate investor confidence on the profitability of momentum strategies reveals interesting patterns. Beta coefficients are initially close to zero or negative, and turn significant and positive only after several months. An in-depth understanding of this initial underreaction is an area for future research.

Other applications of aggregate investor confidence in asset pricing factors expose further patterns. In line with an early interpretation of the small-firm effect by Roll (1981) that the small-firm premium is possibly caused by investor mis-assessment of risk, I find significantly positive predictability of the small-minus-big (SMB) factor for approximately one year ($F$ = 3.28, adj. $R^2$ = 5.12%). One potential explanation is that overconfident investors tend to systematically underestimate firm-specific risk associated with size. This interpretation is supported by findings, reported in chapter 2, which indicated that the proportion of small stocks traded is larger after periods of high aggregate investor confidence.

The layout of this paper is as follows: Section 3.2 summarises the literature that is used to motivate this study. section 3.3 describes the rationale and construction of the aggregate investor confidence index and summarises the hypotheses. In Section 3.4, I briefly describe the data used for the analysis. Section 3.5 performs analyses on the predictability of momentum, size and value returns and the anatomy of lead–lag effects. Section 3.6 discussed the findings, outlines potential avenues for future research and concludes.

## 3.2 Literature review

### 3.2.1 Price momentum

Price momentum is a pervasive phenomenon in the field of finance. That is, stocks that have performed particularly well in the past 3–12 months also tend to perform well in subsequent time periods. In contrast, stocks that have performed particularly poorly over the past 3–12 months also tend to perform poorly in the near future. The effect, however, tends to revert over 3–5 years, as past winners then turn into losers, and vice versa (Jegadeesh, 1990; Jegadeesh and Titman, 1993; Lehmann, 1990).

Price momentum is not uniquely found in US stocks. Among others, Chan, Hameed, and Tong (2000); Fama and French (2012); and Rouwenhorst (1998) find momentum in international stocks. Annual excess returns typically range between 8% and 27%, with weaker effects among large firms. Japanese stocks consistently generate close to zero or negative momentum returns.

Asness et al. (2013) find positive cross-country price momentum returns not only for stocks, but also for other asset classes. Asness et al. (2013) find momentum in bonds; Okunev and White (2003) and Asness et al. (2013) in currencies; Erb and Harvey (2006) and Asness et al. (2013) in commodities; and Miffre and Rallis (2007) in commodity futures markets.

### 3.2.2 Behavioural models of price momentum

The literature yields a multitude of attempts to explain price momentum. For instance, Liew and Vassalou (2000) attempt to explain momentum with future nominal gross domestic product (GDP) growth, and Maio and Santa-Clara (2012) use an inter-temporal 'capital asset pricing model (CAPM) to explain the phenomenon. However, the current study takes a behavioural perspective. The following section briefly summarises some proposed models.

#### News watchers and momentum traders

Hong and Stein (1999) propose a behavioural model that links initial underreaction with long-term overreaction (DeBondt and Thaler, 1985, 1987). In their model, investors can be categorised into "news watchers" and "momentum traders", with the former group of investors closely following news announcements to update their information which is then translated into a security's fundamentals, while the latter group observes occurrences of price drift and members of this group act as arbitrageurs.

Despite the straightforward nature of this conceptual model, its empirical applicability is limited. How can we identify if an investor is a newsreader or a momentum trader, and could she perhaps be a combination of the two? Furthermore, how can we explain a variation of momentum returns over time, and why are momentum returns sometimes negative (e.g. Cooper et al. (2004) and Daniel and Moskowitz (2013))?

If the model holds true and momentum traders identify any price drift as momentum, the resulting price correction therefore should be identified as a trend, continuously causing overshooting and undershooting of fundamental values. Furthermore, assuming that price drifts are at least partially due to gradually diffusing information, one may assume that these drifts disappear if information diffusion accelerates. Modern technology allows almost instant access to information which should at least alleviate this effect.

Despite acknowledging that the model developed by Hong and Stein (1999) is conceptually sound, it remains a challenge to find empirical support due to its lack of applicability.

#### An investor sentiment model of conservatism and representativeness bias

Barberis, Shleifer, and Vishny (1998) propose a model of investor sentiment. The model borrows behavioural biases to explain how investors form biased beliefs which are responsible for underreaction and overreaction and, thus, momentum. On the one hand, conservatism leads to the slow updating of investors' beliefs, as existing beliefs are not instantly altered upon the arrival of new evidence (Edwards, 1968). For instance, the investor of asset $i$ may initially hold a positive belief about that asset. After the arrival of negative news about the future outlook of the asset, he refuses to alter his previously formed belief to the full extent, but gradually becomes more rational once further evidence arrives. Eventually, his belief is consistent with his asset's fundamental price.

Representative bias is borrowed from Tversky and Kahneman (1974). Investors suffering from this bias may erroneously project the recent growth of a company that has performed particularly well in the last few years too far into the future, ignoring regression to the mean.

Despite the intuitive nature of this model, it seems very difficult to test the model empirically. That is, measuring variations in conservatism and representative bias among investors appears to be a challenge. Furthermore, justifying the absence of momentum in some markets and temporal variations in others seems counter-intuitive.

#### Self-attribution bias and overconfidence

Daniel et al. (1998) and Daniel, Hirshleifer, and Subrahmanyam (2001) propose a model of investor psychology as a unified approach to address security price underreaction and overreaction. According to the model, these two effects can be explained with aggregate self-attribution bias and investor overconfidence due to private information. Self-attribution bias in the spirit of Bem (1965); Miller and Ross (1975); and Langer and Roth (1975) is responsible for (erroneously) updating an investor's level of confidence.

In other words, investors tend to overestimate the precision of private information signals. If information arrives subsequent to the formation of an (overconfident) investor's belief, his confidence will be reinforced by the arrival of evidence, and he will chase the trend of the underlying security.

Daniel et al. (1998) and Gervais and Odean (2001) suggest that investors may erroneously attribute market returns to their own skill and develop overconfidence when market returns are bullish.

Consequently, if the model of Daniel et al. (1998) holds true, momentum returns should be higher subsequent to periods of market gains and less pronounced after bearish markets. Indeed, Johnson (2002); Sagi and Seasholes (2007); and Stivers and Sun (2010) find that momentum returns tend to behave pro-cyclically.

Cooper et al. (2004) suggest, in regard to the hypothesis of Daniel et al. (1998), that investors are overconfident due to private information and build up unreasonable confidence until they face discrediting news and adjust their level of confidence to the appropriate level.

Consequently, momentum returns should be higher (lower) after "up-market" (down-market) conditions. Cooper et al. (2004) define "up-markets" ("down-markets") as market conditions where lagged three-year market returns are non-negative ("negative") and find that momentum returns after up-markets average 0.93% per month but, after down-markets, the average momentum return is only -0.37% per month.

However, the literature lacks evidence of a detailed understanding of the relationship between investor confidence and momentum returns. Cooper et al. (2004); Daniel et al. (1998); and Gervais and Odean (2001) suggest that aggregate investor confidence should be higher subsequent to market gains and lower after market losses.

Nevertheless, these studies do not quantify the degree of investor confidence. As a result, the relationship between the degree of investor confidence and momentum returns has not been tested.

The aim of this paper is to provide an empirical link between the conceptual model of Daniel et al. (1998) and price momentum, which exceeds the work of Cooper et al. (2004). I apply a proxy for aggregate investor confidence, as proposed in chapter 2. The measure borrows from a model of the determinants of confidence proposed by Griffin and Tversky (1992).

Prior findings reported in chapter 2 suggest that, in aggregate, trading activity is higher after months with high levels of investor confidence. Consistently, time periods subsequent to high aggregate investor confidence appear to be associated with increased investor risk appetite, proxied by a larger proportion of investment in small stocks.

## 3.3 Empirical approach

### 3.3.1 Measuring aggregate investor confidence

In chapter 2, I propose a measure for aggregate investor confidence in the stock market which borrows its rationale from Griffin and Tversky (1992) who define two primary components of confidence: strength and weight of evidence. Strength is the extremeness of evidence, such as the number of goals scored by a football player, a recent exam result of a student, or the recent portfolio performance of a stock trader. Weight, on the other hand, is the credibility of evidence, for instance, the performance of the football player over the season, the performance of other students during the exam, or the performance of the stock trader over time and in comparison with the market. This differentiation plays a key role in understanding the rationale for overconfidence in the spirit of Griffin and Tversky (1992).

The first component, strength of evidence, and its impact on judgement, can frequently be observed in everyday situations. Football players who recently scored many goals are quickly considered great talents, students who recently performed well are undoubtedly hard-working and highly intelligent and investors who recently enjoyed an outstanding portfolio performance form the belief that they perhaps are investment geniuses. In other words, individuals tend to form biased beliefs about the probability of a future event, as they tend to ignore shortcomings in sample size and attribute favourable outcomes preferably to their own aptitude and skill. These heuristics are well documented in the literature as representativeness bias (Tversky and Kahneman, 1974) and self-attribution bias (Miller and Ross, 1975).

These heuristics-driven biases largely depend on the reliability of evidence. Suppose you observe the aforementioned football player over the entire season. If he continuously scores goals, he is increasingly likely to be highly talented. Otherwise, the observer may alter his initial belief and acknowledge that the player may have been lucky or the opponent simply had a bad day. Likewise, an investor may appear highly talented when observing high portfolio performance in isolation. However, his judgement may differ if market returns are equal to the investor's performance, especially when the standard deviation of such impulses is high. That is, the higher the standard deviation of strength impulses, the more market participants may believe themselves to be superior investors.

I follow this rationale when developing an aggregate measure of investor confidence in the stock market. The measure predicts trading volume and provides implications for aggregate risk behaviour when confidence is high. More specifically, aggregate trading volume, with small stocks measured by stock turnover, tends to be higher when aggregate investor confidence is high. Furthermore, the proportion of small stocks traded, measured as stocks in the bottom 40% of market capitalisation firms listed on AMEX, NASDAQ and NYSE, is higher after months with high aggregate investor confidence.

According to the model proposed in chapter 2, stock traders' confidence in their ability to pick stocks is primarily driven by feedback, That is: how well have they performed recently in respect to a 'typical' return given the current market state? Thus, a recent performance impulse is defined as the deviation of a recent return from the typical return, computed as a simple average of returns over time windows of 6, 12, 24 and 36 months, each of which, hereafter, is referred to as a 'baseline period'. The lengths of the baseline periods are common time horizons used in the underreaction and overreaction literature (see, e.g. DeBondt and Thaler (1985) and DeBondt and Thaler (1987)) Thus, the strength variable of a stock in an index can be expressed as follows:

$$v_{i,t} = r_{i,t} - \overline{r}_{i,u,t}, \tag{3.1}$$

where $v_{i,t}$ is the strength of evidence in period $t$; $r_{i,t}$ is the return of stock $i$ in month $t$; and $\overline{r}_{i,u,t}$ is the simple moving average of stock $i$ in baseline period $u$.

For the weight of evidence variable, I use the standard deviation of strength impulses. The rationale behind this proxy is as follows. If the standard deviation of strength impulses is high, at least some of these results is due to chance. As a result, some investors will erroneously attribute these gains to their own abilities and become overconfident. Therefore,

$$W_t = \frac{1}{\sqrt{\frac{\sum_{i=1}^{n}(v_{i,t}-\overline{v}_t)^2}{n-1}}}, \tag{3.2}$$

where $W$ is the weight of evidence in period $t$; $v$ is the strength of stock $i$ in the index; and $\overline{v}_t$ is the mean strength in the index in month $t$.

The weight of evidence, $W$, is expressed as the reciprocal of the standard deviation of the strength of evidence variable, as Griffin and Tversky (1992) suggest that overconfidence is high when strength is *high* and weight is *low*.

As the goal is to compute an aggregate measure of investor confidence, a market equity-weighted mean score for the strength variable is computed:

$$V_t = \sum_{i=1}^{n} \frac{m_{i,t} v_{i,t}}{\sum_{i=1}^{n} m_{i,t}}, \tag{3.3}$$

where $V_t$ is the value-weighted strength of all stocks in the index; and $m_i$ is the respective market equity weight of a stock.

As all prior variables are computed with simple returns, natural logarithms of $(1 + V_t)$ and $(1 + W_t)$ are computed, that

$$ln(INVCON_t) = ln(1 + V_t) - ln(1 + W_t). \tag{3.4}$$

In order to control for possible macroeconomic factors, I take the first difference of the index to achieve stationarity, and subsequently regress the index values on a range of macroeconomic variables. These variables include growth in the industrial production index[4]; as well as a dummy variable for NBER recessions; and following Baker and Wurgler (2007), growth in consumer durables, non-durables and services[5], with these authors assuming that the residuals of such regressions are free of major business cycle effects. All alternative measures of confidence and sentiment applied in this paper undergo the same treatment. To avoid possible look-ahead bias, I use raw variables to undergo tests as a robustness check.

### 3.3.2 Hypotheses development

This section describes the motivation for the expected hypotheses derived from the literature. Hypotheses 1 and 2 aim to meet the main objectives of this paper, namely, exploring the relationship between aggregate investor confidence and the profitability of momentum strategies. Hypothesis 3 extends this by attempting to isolate overconfidence components and their ability to explain variations in the profitability of momentum strategies. Hypotheses 4 and 5 address the impact on high-minus-low (HML) and small-minus-big (SMB) factors.

#### Aggregate investor confidence and price momentum

One of the primary objectives of this paper is to test the overconfidence hypothesis of Daniel et al. (1998) which postulates that price momentum is largely due to investor overconfidence in the accuracy of their private information. Testing this hypothesis empirically is challenging for two reasons. Firstly, few closely related measures of aggregate investor confidence in their own abilities to pick stocks exist. Lemmon and Portniaguina (2006) use the University

[4] The data for growth in the industrial production index are available at: Federal Reserve Statistical Release G.17

[5] Data for consumer durables, non-durables and services are available at BEA National Income Accounts Table 2.10

of Michigan's Consumer Sentiment Index as a proxy and find no significant relationship with book-to-market and momentum factors.

Secondly, isolating overconfidence from appropriate levels of confidence is challenging, as both 'appropriate' levels of confidence and observed levels of confidence have to be assessed separately in order to identify true overconfidence. Glaser et al. (2013) discuss this issue in detail and propose a remedy for experimental studies.

In chapter 2, I propose a measure of aggregate investor confidence which uses stock market data to compute feedback impulses through market returns which, in turn, affect the population of investors in aggregate. According to the model, confidence increases when recent returns exceed a typical return given a particular market, especially when the origin of those returns is difficult to attribute. The proxy possesses predictive ability for trading activity, measured by stock turnover, as well as the proportion of risky stocks traded, after being orthogonalised against major business cycle components.

Consequently, I assume that the measure at least partially captures variations in aggregate investor confidence, and I apply it to test the hypothesis of Daniel et al. (1998). Following this notion, momentum returns should be more pronounced when aggregate investor confidence is high.

**Hypothesis 1:** *Price momentum returns are higher in periods following high levels of aggregate investor confidence, and vice versa.*

I acknowledge that the measure does not directly isolate overconfidence from confidence. However, given that the measure captures all components of confidence, it should also capture the overconfidence component. Thus, I expect a significant and positive association between the aggregate investor confidence proxy and Fama-French's[6] momentum factor (up-minus-down [UMD]),

[6] I refer to the UMD factor as the Fama-French UMD factor as I use data from Kenneth French's website which follows Fama-French methodology. The factor was first introduced by Carhart (1997).

which reflects the time-series profitability of momentum strategies. However, in order to rule out spurious correlations and reverse causality issues, the aggregate investor confidence index should be applied as a lagged variable.

The second hypothesis addresses the length of the lead–lag effect of aggregate investor confidence and momentum. Statman et al. (2006) test the suggestions of Gervais and Odean (2001) and Odean (1998) that aggregate confidence should be higher subsequent to market gains in the case of trading volume using impulse response functions. They find that a significant positive effect persists for several months. Consequently, I expect a similar effect for the case of price momentum. In other words, I apply 1–24 lags to a selection of investor confidence and sentiment proxies. In particular, I expect significant and positive associations for multiple lags of INVCON indices and the Fama-French UMD factor.

**Hypothesis 2:** *Price momentum returns remain significant for several time periods after following high levels of aggregate investor confidence.*

Alternative investor confidence and sentiment proxies are included in this analysis, in order to explore their ability to explain variations in the profitability of momentum strategies. As I argue that other proxies are not closely related to the notion of overconfidence in private information in the spirit of Daniel et al. (1998), I do not expect significant relationships. This is particularly the case for the University of Michigan Consumer Sentiment Index, the Investor Sentiment Index by Baker and Wurgler (2007), as well as a series of investor confidence indices proposed by Shiller (2000b). To an extent, the American Association of Individual Investors (AAII)'s Investor Sentiment Index may capture overconfidence, as the survey questionnaire is aimed at investors' beliefs about market prospects in the near future.

#### Aggregate investor confidence and asset pricing factors

A secondary objective of this study is to explore the effect of aggregate investor confidence on asset pricing factors. Therefore, I include the Fama-French HML and SMB factors as dependent variables, with these undergoing the same tests as the UMD factor.

Fama and French (1993) introduce a small-minus-big (SMB) factor[7] that captures the small-firm effect. Banz (1981) and Fama and MacBeth (1973) document that, on average, firms with small market capitalisation earn higher than normal risk-adjusted returns than simple asset pricing models suggest. Roll (1981) suggests that “the mis-assessment of risk has the potential to explain why small firms, low price/earnings ratio firms, and possibly high dividend yield firms display large excess returns” (p.887). A large body of literature documents the systematic understating of risk by overconfident decision makers (e.g. Barber and Odean (2001); Gervais and Odean (2001); Grinblatt and Keloharju (2009); or Malmendier and Tate (2008)).

Consequently, risk overall should, in aggregate, be understated when aggregate investor confidence is high. In other words, higher aggregate confidence among investors leads to stronger mis-assessment of risk in the spirit of Roll (1981) and, ultimately, to a stronger pronounced SMB effect. As a result, I expect a significant and positive association between lagged INVCON impulses and the SMB factor.

**Hypothesis 3:** *The small-firm effect is higher in periods following high levels of aggregate investor confidence, and vice versa.*

Roll (1981) furthermore proposes that the same mis-assessment of risk potentially explains abnormal returns of low price/earnings ratio firms. Fama (1995) suggest that high book-to-market ratios, which reflect low stock prices

[7] The factor is the return of a value-weighted portfolio of small stocks minus the return of a value-weighted portfolio of large market capitalisation stocks. The data are available on Kenneth French’s website.

relative to their book value, indicate low earnings on book equity. Consequently, I expect a negative relationship, if any relationship at all, between lagged aggregate investor confidence impulses and the Fama-French HML factor[8] (Hypothesis 4). However, the conceptual relationship is not as intuitively striking, especially from a behavioural perspective.

## 3.4 Data description

This section summarises the data used for this study, as well as its sources. The data validation tests, especially regarding stationarity tests of respective investor confidence and sentiment proxies, are reported in chapter 2.

### 3.4.1 Security data

In chapter 2, I compute an aggregate confidence index for NYSE AMEX and NASDAQ. The data to compute the strength and weight of evidence proxies are collected from the monthly CRSP-Compustat database. All common stocks with share codes 10 and 11 between July 1927 and December 2014 are included. The CRSP-Compustat database provides holding period returns which is the change in the total value of an investment in a common stock per dollar of initial investment, including dividend returns. Asset pricing model data, including returns of the four Fama-French factors, are from Kenneth French's website.[9]

### 3.4.2 Control variables

I follow the recommendations of Baker and Wurgler (2007) to isolate raw confidence index scores from major business cycle components. Each stationary proxy is regressed with respect to a range of macroeconomic variables, in-

[8] Factor values reflect the return of a value-weighted portfolio of high book-to-market stocks minus a value-weighted portfolio of low book-to-market stocks.

[9] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

cluding the industrial production index[10]; growth in consumer durables, non-durables and services[11]; as well as a dummy variable for NBER recessions. Residuals of these regressions are assumed to be largely free of major business cycle effects, as previously suggested by Baker and Wurgler (2007). This step is essential, as the literature suggests that some momentum returns are driven by macroeconomic variables.[12]

### 3.4.3 Alternative measures of investor confidence/ sentiment

In this study, I adapt a range of alternative measures of investor confidence and sentiment, as identified in chapter 2. The purpose is to explore the potential implications for asset pricing, as well as the shared and distinct properties they may possess. These indices include the American Association of Individual Investors (AAII) Investor Sentiment Survey[13]; the University of Michigan Consumer Sentiment Index[14]; the investor sentiment index by Baker and Wurgler (2007)[15]; and a range of investor confidence indices by Shiller (2000b).[16] It must be noted that monthly data of Shiller's confidence indices are available only from 2001 which results in a relatively small sample of 161 observations.

The American Association of Individual Investors (AAII) surveys its members in regard to their future expectations of developments of the stock market. Respondents report either bullish ($\text{AAII}_{bull}$), neutral ($\text{AAII}_{neutral}$) or bearish ($\text{AAII}_{bear}$) expectations. Baker and Wurgler (2007) develop an index of in-

[10] The data on growth in the industrial production index are available at: Federal Reserve Statistical Release G.17

[11] Data for consumer durables, non-durables and services are available from BEA National Income Accounts Table 2.10

[12] See, for example, Chordia and Shivakumar (2002), who find that industry-based momentum returns are captured by macroeconomic variables.

[13] The data are available at http://www.aaii.com/sentimentsurvey/

[14] The data are available at http://www.sca.isr.umich.edu/

[15] The data are available at http://people.stern.nyu.edu/jwurgler/

[16] The data are available upon request at http://som.yale.edu/faculty-research/our-centers-initiatives/international-center-finance/data/stock-market-confidence-indices/ I express my gratitude to the International Center of Finance at the Yale School of Management for providing the data.

vestor sentiment (BWSENT) which is intended to capture investors' optimism and pessimism, by including trading volume, dividend premiums, closed-fund discounts or first-day IPO returns as components. The University of Michigan surveys consumers in regard to their expectations of the economy, as well as their investment and saving behaviour (Michigan). Shiller (2000b) introduces a family of investor attitude measures among institutional and individual investors at six-monthly (and later monthly) intervals. This study includes the three Shiller indices that are the most closely related, namely, the One-Year Confidence Index, the Valuation Confidence Index and the Crash Confidence Index.

I show in chapter 2 that aggregate investor confidence indices are largely unrelated with these alternative measures of investor confidence and sentiment, except for the bullish component of the AAII Investor Sentiment Index.

## 3.5 Empirical tests

### 3.5.1 Aggregate investor confidence and the Fama-French momentum factor

Hypothesis 1 tests the ability of aggregate investor confidence indices to predict the profitability of momentum strategies. I propose four alterations of the index which are based on the length of the baseline period. That is, it is assumed that investors form beliefs about a 'typical' return according to current market states. The length of these baseline periods varies between six and 24 months, hence the numeric component in the index name. To capture the impact of a confidence impulse for several months thereafter, I use 1–24 lags of the respective aggregate investor confidence index as an independent variable, and the Fama–French momentum factor as a dependent variable in a series of time-series regressions in the following form:

$$Y_t = a + bX_{t-1} + bX_{t-2} + \ldots + bX_{t-24} + e_t, \quad (3.5)$$

where $Y_t$ is the value of the UMD factor in month $t$; $X_{t-l}$ is the value of the first difference of the logarithm of 24 lags $l$ of aggregate investor confidence indices with baseline periods of 6, 12, 24 and 36 months, controlled for macroeconomic variables. Respective aggregate investor confidence values are lagged by $l$ 1–24 months; $a$ is a constant; $b$ is the respective coefficient; and $e_t$ is the random error term.

Table 3.1 below summarises the test output of Hypothesis 1. All 24 lags of the first differences of the logarithm of all aggregate investor confidence modifications show significant ($F$ = 1.73, 2.19, 2.31 and 2.08, respectively) and positive associations with the Fama-French UMD factor. However, the reported adjusted $R^2$ values are relatively low (adj. $R^2$ = 1.67%, 2.68%, 3.01% and 2.86%, respectively), suggesting co-existing alternative explanations. Hypothesis 3 further investigates this notion.

Similar to the findings of Statman et al. (2006) who find that investor confidence affects trading activity for several months, Hypothesis 2 states that an aggregate investor confidence impulse entails several monthly lags of significant and positive predictive association.

Table 3.1 summarises the findings of this hypothesis test. All four modifications of aggregate investor confidence show somewhat similar patterns. Generally, an aggregate investor confidence impulse affects the profitability of momentum strategies for many lags, which typically range from 12–15 months. In line with the hypothesis, associations are consistently positive, with the exception of the first lag of aggregate investor confidence with baseline periods of 12 and 24 months, which are negative. The negative component was not hypothesised earlier, and thus may be a possible recommendation for future research.

Figure 3.1 below illustrates beta coefficient estimates of the four aggregate

**Table 3.1: Investor confidence and the Fama-French UMD factor**

This table summarises test output of four time-series regression models, with the four versions of the 24 lags of first difference of the logarithm of INVCON6, INVCON12, INVCON24 and INVCON36 as independent variables. INVCON6 is based on a baseline period of 6 months, which refers to assumptions of a time period of 6 months for an investor to form beliefs about a 'typical' market return at a given point in time. I use monthly data ranging from July 1927 to December 2014, resulting in 1062 monthly observations.

| | INVCON6 | | INVCON12 | | INVCON24 | | INVCON36 | |
|---|---|---|---|---|---|---|---|---|
| lags | Coeff. | $t$-stat. | Coeff. | $t$-stat. | Coeff. | $t$-stat. | Coeff. | $t$-stat. |
| 1 | 0.01 | 0.28 | -0.09*** | -2.99 | -0.10*** | -3.30 | 0.03 | 0.89 |
| 2 | 0.09** | 2.08 | -0.06 | -1.33 | -0.07 | -1.64 | 0.13*** | 3.21 |
| 3 | 0.11** | 2.36 | 0.05 | 0.96 | 0.04 | 0.71 | 0.16*** | 3.25 |
| 4 | 0.13** | 2.52 | 0.06 | 1.17 | 0.06 | 1.01 | 0.16*** | 3.02 |
| 5 | 0.12** | 2.13 | 0.08 | 1.41 | 0.08 | 1.44 | 0.14** | 2.41 |
| 6 | 0.16*** | 2.70 | 0.07 | 1.18 | 0.07 | 1.17 | 0.17*** | 2.85 |
| 7 | 0.23*** | 3.77 | 0.09 | 1.52 | 0.10 | 1.58 | 0.21*** | 3.46 |
| 8 | 0.21*** | 3.32 | 0.15** | 2.45 | 0.16** | 2.47 | 0.18*** | 2.91 |
| 9 | 0.23*** | 3.61 | 0.14** | 2.24 | 0.15** | 2.25 | 0.23*** | 3.72 |
| 10 | 0.22*** | 3.28 | 0.17*** | 2.69 | 0.17*** | 2.65 | 0.26*** | 4.15 |
| 11 | 0.22*** | 3.28 | 0.16*** | 2.59 | 0.16** | 2.52 | 0.24*** | 3.81 |
| 12 | 0.15** | 2.30 | 0.15** | 2.38 | 0.15** | 2.25 | 0.20*** | 3.08 |
| 13 | 0.14** | 2.05 | 0.10 | 1.61 | 0.09 | 1.42 | 0.19*** | 2.98 |
| 14 | 0.12* | 1.84 | 0.10 | 1.60 | 0.09 | 1.40 | 0.15** | 2.34 |
| 15 | 0.16** | 2.44 | 0.08 | 1.29 | 0.06 | 0.98 | 0.17*** | 2.75 |
| 16 | 0.07 | 1.11 | 0.11* | 1.91 | 0.11 | 1.65 | 0.10 | 1.63 |
| 17 | 0.01 | 0.12 | 0.03 | 0.55 | 0.02 | 0.31 | 0.07 | 1.19 |
| 18 | 0.00 | 0.07 | 0.00 | -0.05 | -0.02 | -0.34 | 0.06 | 1.04 |
| 19 | 0.01 | 0.19 | 0.00 | -0.08 | -0.03 | -0.54 | 0.07 | 1.23 |
| 20 | 0.04 | 0.70 | 0.01 | 0.19 | -0.01 | -0.24 | 0.09* | 1.67 |
| 21 | 0.07 | 1.47 | 0.04 | 0.88 | 0.01 | 0.22 | 0.14*** | 2.66 |
| 22 | 0.05 | 1.12 | 0.09** | 1.97 | 0.06 | 1.28 | 0.11** | 2.26 |
| 23 | 0.02 | 0.38 | 0.06 | 1.54 | 0.05 | 1.16 | 0.07* | 1.65 |
| 24 | 0.00 | 0.00 | 0.01 | 0.41 | 0.01 | 0.35 | 0.02 | 0.61 |
| | $F$-Stat. | Adj. $R^2$ | $F$-Stat. | Adj. $R^2$ | $F$-Stat. | Adj. $R^2$ | $F$-Stat. | Adj. $R^2$ |
| Mdl | 1.73** | 1.67% | 2.19*** | 2.68% | 2.31*** | 3.01% | 2.08*** | 2.86% |

***, ** and * Statistically significant at a 99%, 95% and 90% confidence level, respectively (two-tailed).

investor confidence measures. A positive predictive association is maintained for several months, and reaches its pinnacle at around month 10. A similar pattern can be observed for aggregate investor confidence measures with different baseline periods. Quite surprisingly, coefficients of measures with baseline periods of 12 and 24 months are significantly negative for the first lag.

A negative first lag could be related to the short-term reversal effect, which is a common statistical artefact in stock pricing (e.g. DeBondt and Thaler (1987) or Fama and French (1996b)) . Typically, the first lag is skipped when forming performance-based portfolios in order to mitigate this effect (e.g. Jegadeesh and Titman (1993) or Moskowitz, Ooi, and Pedersen (2012)).

The pattern in figure 3.1 potentially sheds light on the anatomy of confidence formation and its subsequent translation into behavioural patterns. That is, a confidence impulse through feedback is slowly incorporated into an investor's confidence, reaching a peak in magnitude after approximately 10 months. In subsequent months, the effect diminishes. This interpretation seems sound and is in line with the findings of Statman et al. (2006) who use impulse–response functions to explore the lead–lag effect of investor confidence on trading activity.

The following section includes a selection of alternative investor confidence and sentiment measures in the analysis, in order to explore potential abilities to predict the profitability of momentum strategies, as well as potential common patterns in lead–lag effects. Additionally, I add Fama–French HML and SMB factors to the analysis as dependent variables.

**Alternative proxies of aggregate investor confidence**

This section tests Hypotheses 4 and 5 empirically, and explores the effect of alternative measures of investor confidence and sentiment on asset pricing factors, as well as potential common features with aggregate investor confidence measures.

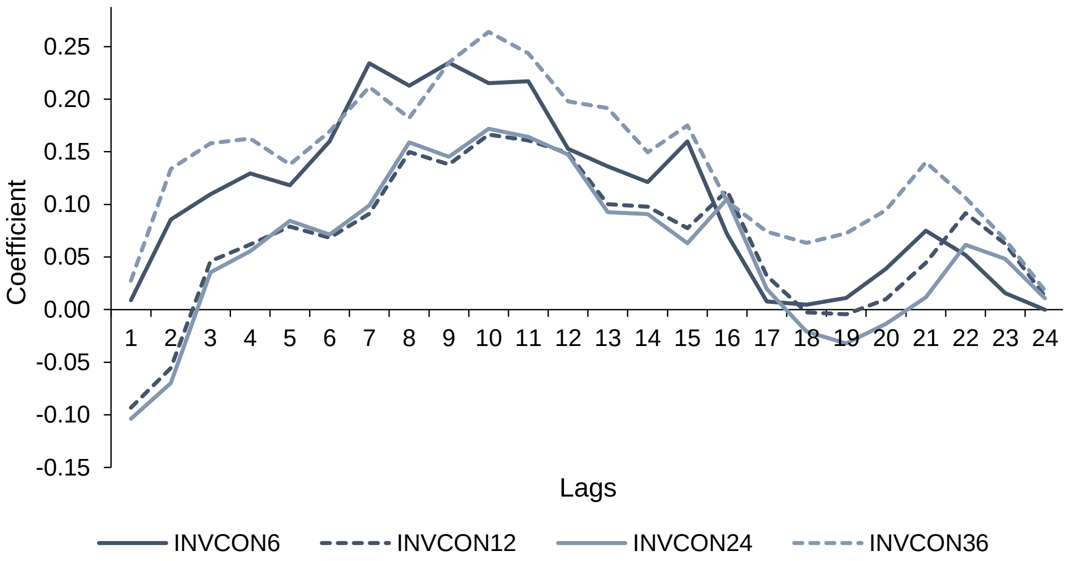



**Figure 3.1: Beta coefficients of 24 lags of the first difference of INVCON indices and the Fama-French UMD factor**

This figure illustrates beta coefficients of four time-series regression models with 24 lags of each variation of INVCON indices as independent variables and the Fama-French UMD factor as a dependent variable. Predictable association is initially negative at 1 lag for INVCON12 and INVCON24. All modifications of the INVCON index show a similar pattern. Positive predictability sustains between month 2 and approximately month 15 subsequent to a confidence impulse. The magnitude reaches its pinnacle approximately 10 months after a confidence impulse.

Table 3.2 reports the test output of time-series regression models with 24 lags of the first difference of the logarithm of a respective investor confidence or sentiment measure as the independent variables, and the Fama-French UMD factor as a dependent variable.

The analysis is structured as follows. I identify the superior version of aggregate investor confidence for each asset pricing factor based on its significance and adjusted $R^2$ value and subsequently I explore significant models in greater detail, in order to identify potential common patterns.

Although all versions of lagged aggregate investor confidence measures have statistically significant positive relationships with the Fama-French UMD factor, the version with a baseline period of 24 months has the greatest explanatory power, with an $F$-statistic value of 3.21 and an adjusted $R^2$ of 3.01%.

Quite strikingly, Shiller's 1-Year Confidence Index is a highly significant predictor of UMD (F = 2.5) with high predictability. The high adjusted $R^2$ value of 20.92% is largely due to lag 24, as reported in table 3.3 and figure 3.2 below. However, the effect may be due to a relatively small sample of monthly observations of the Shiller confidence indices. The only other significant predictors of the UMD factors are $AAII_{bull}$ ($F = 1.83$, adj. $R^2 = 6.18\%$) and $BWSENT$ ($F = 1.43$, adj. $R^2 = 1.95\%$). $AAII_{bull}$ is investigated in greater detail below.

Table 3.3 below reports the coefficients and $t$-statistic values of 24 lags of three selected investor confidence proxies. This finding raises some questions. Firstly, no obvious pattern exists and, secondly, the same index is insignificant for institutional investors. Unreported regressions with up to 48 lags suggest that the strong association in lag 24 is unique.

$AAII_{bull}$ shows a similar pattern with INVCON24. Initial low and insignificant coefficients turn significant and positive after four lags, and remain significant till lag 11, supporting Hypothesis 2 of the multi-period impact of confidence impulses.

**Table 3.2: Investor confidence/sentiment proxies and Fama-French UMD factor**

This table reports the $F$-statistic significance values and adjusted $R^2$ values for a series of time-series models, with the Fama-French UMD factor as a dependent variable, and 24-months lags of INVCON6, INVCON12, INVCON24, INVCON36 as well as the selected alternative investor confidence or sentiment proxies ($AAII_{bull}$, $BWSENT$, $Michigan$, as well as three Shiller confidence indexes) as independent variables. $INVCON6$ is based on a baseline period of 6 months, which refers to the assumption of a time period of 6 months for an investor to form beliefs about 'typical' market return at a given point in time. All alterations of the INVCON indices possess predictability of the UMD factor Generally, 24 lags of selected alternative proxies appear not to possess predictability of the Fama–French UMD factor, with some notable exceptions. Shiller's 1-Year Confidence Index has high explanatory power of variations in the UMD factor with an adjusted $R^2$ of 20.92%. Interestingly, the effect is insignificant for institutional investors. The bullish component of the AAII index is another notable exception. For each dependent variable, a superior INVCON proxy, as well as significant alternative proxies, are selected for further analysis.

| Proxy | $F$-stat. | Adj. $R^2$ | N |
|---|---|---|---|
| INVCON6 | 1.73** | 1.67% | 1062 |
| INVCON12 | 2.19*** | 2.68% | 1062 |
| INVCON24 | 2.31*** | 3.01% | 1062 |
| INVCON36 | 2.08*** | 2.86% | 1062 |
| $AAII_{bull}$ | 1.83** | 6.18% | 329 |
| $BWSENT$ | 1.43* | 1.95% | 545 |
| $Michigan$ | 1.12 | 0.71% | 443 |
| $Shil_{1Y}$ (ind.) | 2.50*** | 20.92% | 161 |
| $Shil_{1Y}$ (inst.) | 0.58 | 0.00% | 161 |
| $Shil_{val}$ (ind.) | 0.89 | 0.00% | 161 |
| $Shil_{val}$ (inst.) | 0.72 | 0.00% | 161 |
| $Shil_{cr}$ (ind.) | 0.81 | 0.00% | 161 |
| $Shil_{cr}$ (inst.) | 1.06 | 1.07% | 161 |

***, **, *. Statistically significant at a 99%, 95% and 90% confidence level, respectively (two-tailed).

**Table 3.3: Selected investor confidence proxies and the Fama-French UMD factor**

This table reports t-statistic significance values and beta coefficient estimates for a series of time series models, with the Fama-French UMD factor as a dependent variable, and 24-months lags of the first differences of INVCON24, the bullish component of the AAII Investor Confidence Index, as well as the individual investor version of Shiller's 1-Year Confidence Index. INVCON12 is selected, as it possesses superior predictability over other modifications of INVCON indices. $AAII_{bull}$ and individual investor version of Shiller's 1-Year Confidence Index ($Shil_{1Y}$) are selected due to overall model significance, as reported in Table 3.2 above. Unreported analyses of time series regression models with up to 48 lags show no repetition of the phenomenon. Figure 3.2 below illustrates the pattern reported in this table.

| | INVCON24 | | $AAII_{bullish}$ | | $Shil_{1Y}$ (ind.) | |
|---|---|---|---|---|---|---|
| lags | Coeff. | t-stat | Coeff. | t-stat | Coeff. | t-stat |
| 1 | -0.10*** | -3.30 | -0.02 | -0.50 | 0.004 | 1.56 |
| 2 | -0.07 | -1.64 | 0.04 | 1.06 | -0.001 | -0.28 |
| 3 | 0.04 | 0.71 | 0.04 | 0.98 | -0.005** | -2.10 |
| 4 | 0.06 | 1.01 | 0.09*** | 2.33 | -0.002 | -0.71 |
| 5 | 0.08 | 1.44 | 0.13*** | 3.34 | 0.001 | 0.57 |
| 6 | 0.07 | 1.17 | 0.05 | 1.28 | 0.000 | -0.04 |
| 7 | 0.10 | 1.58 | 0.13*** | 3.04 | 0.000 | 0.05 |
| 8 | 0.16** | 2.47 | 0.13*** | 3.19 | -0.004 | -1.54 |
| 9 | 0.15** | 2.25 | 0.08* | 1.89 | -0.001 | -0.53 |
| 10 | 0.17*** | 2.65 | 0.12*** | 2.95 | -0.003 | -1.06 |
| 11 | 0.16** | 2.52 | 0.14*** | 3.42 | 0.001 | 0.19 |
| 12 | 0.15** | 2.25 | 0.06 | 1.39 | 0.002 | 0.91 |
| 13 | 0.09 | 1.42 | 0.04 | 0.92 | 0.000 | -0.04 |
| 14 | 0.09 | 1.40 | 0.02 | 0.44 | -0.001 | -0.26 |
| 15 | 0.06 | 0.98 | 0.01 | 0.27 | 0.006** | 2.19 |
| 16 | 0.11 | 1.65 | 0.02 | 0.45 | 0.003 | 1.23 |
| 17 | 0.02 | 0.31 | 0.06 | 1.47 | 0.002 | 0.73 |
| 18 | -0.02 | -0.34 | -0.02 | -0.42 | 0.006** | 2.13 |
| 19 | -0.03 | -0.54 | 0.00 | 0.12 | -0.002 | -0.81 |
| 20 | -0.01 | -0.24 | 0.04 | 1.03 | 0.004 | 1.59 |
| 21 | 0.01 | 0.22 | 0.03 | 0.76 | 0.006** | 2.59 |
| 22 | 0.06 | 1.28 | 0.03 | 0.96 | 0.000 | -0.01 |
| 23 | 0.05 | 1.16 | 0.05 | 1.48 | 0.000 | -0.01 |
| 24 | 0.01 | 0.35 | 0.02 | 0.67 | 0.011*** | 4.55 |

***, **, *. Statistically significant at a 99%, 95%, or 90% confidence level, respectively (two-tailed).

Figure 3.2 below illustrates the patterns of predictability of INVCON24 and AAII$_{bull}$ for the UMD factor. Predictability is initially low and significantly inverse for INVCON24, and reaches its pinnacle for both proxies at around 10 months. No obvious pattern is observable for Shiller's 1-Year Confidence Index for individual investors.

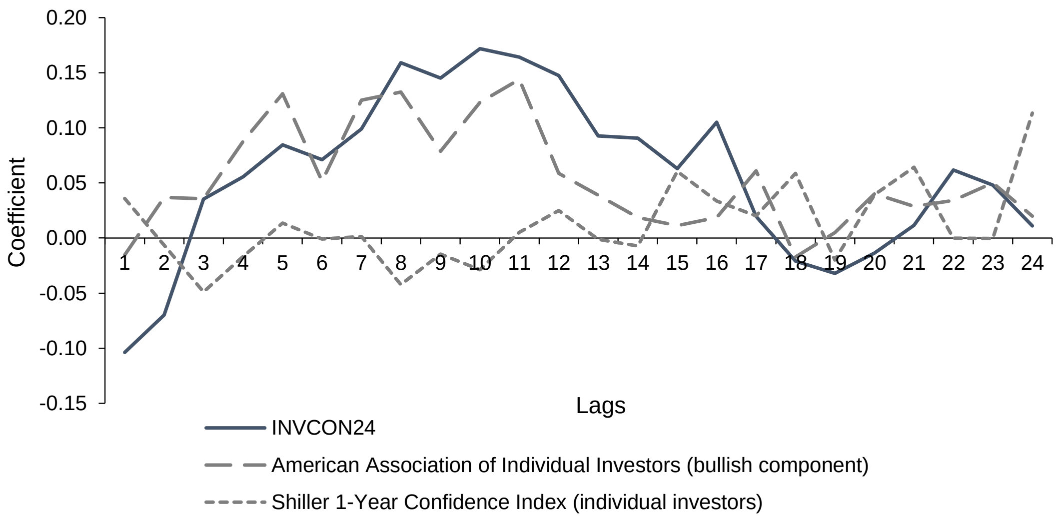



**Figure 3.2: Beta coefficients of selected time-series regression models and the Fama-French UMD factor**
This figure illustrates beta coefficients of three time-series regression models with 24 lags of the first difference of INVCON24 and AAII$_{bull}$, as well as the individual investor version of Shiller's 1-Year Confidence Index as independent variables and the Fama-French UMD factor as a dependent variable. The predictable association is initially negative at 1 lag for INVCON24, with a subsequent positive association for both INVCON12 and the bullish component of the AAII index. Positive predictability is sustained between month 3 and approximately month 16 subsequent to a confidence impulse and reaches its pinnacle at approximately 10 months after a confidence impulse. The individual investor version of Shiller's 1-Year Confidence Index shows a strong positive association at 24 lags. Unreported regressions suggest that the effect is unique for lag 24.

The pattern is suggestive that a confidence impulse affects aggregate investor behaviour over several time periods. Further research is needed to explore the formation of investor confidence in greater detail. This research could potentially shed light on the initial low or negative impact.

### 3.5.2 Aggregate investor confidence and size and value returns

This section explores the relationship predictability of Fama-French HML and UMD factors using investor confidence and sentiment proxies. Table 3.4 below summarises the test output of time-series regression models, with 24 lags of each stationary proxy, controlled for macroeconomic variables, included in this study as independent variables, and Fama-French HML and SMB factors as dependent variables.

All versions of the INVCON measure significantly predict variations in SMB factor loadings. The first column reports test statistic values of these time-series regression models. Shiller's Crash Confidence Index shows high predictability ($F = 1.65$, adj. $R^2 = 10.35\%$).

The relatively low $F$-statistic value is likely to be the result of low test power which is due to a relatively small sample size. Nevertheless, the relationship complements the explanation of the small-firm effect by Roll (1981). When investors are confident that 'nothing can go wrong', they should be more likely to make errors in assessing risk.

However, significant results for institutional investors and insignificant results for individual investors raise some questions. Given that the interpretation of Roll (1981) is valid, institutional investors should shift their preference towards small stocks when confidence is high.

Gompers and Metrick (1998) report that institutional investors tend to have a strong preference for large market capitalisation stocks. Barber and Odean (2001) find that overconfident investors tend to tilt their portfolio preferences toward small stocks. One possible interpretation could be that institutional investors may shift their portfolio preferences, which are typically dominated by large stocks, towards small ones, especially when they are confident that no crash would occur in the near future. The same rationale is a sensible

**Table 3.4: Investor confidence/sentiment proxies and Fama-French SMB and HML factors**

This table reports $F$-statistic significance values and adjusted $R^2$ values for a series of time-series models, with the Fama-French SMB and HML factors as dependent variables, and 24-month lags of INVCON6, INVCON12, INVCON24 and INVCON36, as well as the selected alternative investor confidence or sentiment proxies (AAII$_{bull}$, BWSENT, Michigan, as well as three Shiller confidence indices) as independent variables. INVCON6 is based on a baseline period of 6 months, which refers to assumptions of a time period of 6 months being needed for an investor to form beliefs about 'typical' market return at a given point in time.

| | SMB factor | | | HML factor | | |
|---|---|---|---|---|---|---|
| Proxy | $F$-stat. | Adj. $R^2$ | N | $F$-stat. | Adj. $R^2$ | N |
| INVCON6 | 2.30*** | 2.95% | 1062 | 1.68** | 1.58% | 1062 |
| INVCON12 | 2.88*** | 4.17% | 1062 | 1.30 | 0.70% | 1062 |
| INVCON24 | 3.28*** | 5.12% | 1062 | 1.40* | 0.94% | 1062 |
| INVCON36 | 2.23*** | 2.86% | 1062 | 1.47* | 1.12% | 1062 |
| AAII$_{bull}$ | 1.18 | 1.38% | 329 | 0.81 | 0.00% | 329 |
| BWSENT | 1.20 | 0.91% | 545 | 1.13 | 3.21% | 545 |
| Michigan | 0.74 | 0.00% | 443 | 1.05 | 0.13% | 443 |
| Shil$_{1Y}$ (ind.) | 0.87 | 0.00% | 161 | 0.92 | 0.11% | 161 |
| Shil$_{1Y}$ (inst.) | 1.03 | 0.60% | 161 | 0.53 | 0.00% | 161 |
| Shil$_{val}$ (ind.) | 1.51* | 8.24% | 161 | 0.95 | 0.00% | 161 |
| Shil$_{val}$ (inst.) | 0.59 | 0.00% | 161 | 0.70 | 0.00% | 161 |
| Shil$_{cr}$ (ind.) | 1.07 | 1.28% | 161 | 1.02 | 0.00% | 161 |
| Shil$_{cr}$ (inst.) | 1.65** | 10.35% | 161 | 1.47* | 7.73% | 161 |

***, ** and * Statistically significant at a 99%, 95% or 90% confidence level, respectively (two-tailed).

interpretation of the relationship found between the INVCON measure and the SMB factor. When confident, investors tend to shift their preferences toward small stocks, which is the result of systematic underestimation of associated risk.

**Table 3.5: Investor confidence and the Fama-French SMB factor**

This table reports *t*-statistic significance values and beta coefficient estimates for a series of time-series models, with the Fama-French SMB factor as a dependent variable, and 24-month lags of INVCON6, INVCON12, INVCON24 and INVCON36 as independent variables.

| | INVCON6 | | INVCON12 | | INVCON24 | | INVCON36 | |
|---|---|---|---|---|---|---|---|---|
| Lags | Coeff. | *t*-stat. | Coeff. | *t*-stat. | Coeff. | *t*-stat. | Coeff. | *t*-stat. |
| 1 | 0.06*** | 2.94 | 0.11*** | 5.06 | 0.11*** | 5.15 | 0.05** | 2.54 |
| 2 | 0.09*** | 3.36 | 0.14*** | 5.02 | 0.15*** | 5.38 | 0.07** | 2.30 |
| 3 | 0.05 | 1.54 | 0.14*** | 4.25 | 0.17*** | 4.98 | 0.03 | 0.92 |
| 4 | 0.04 | 1.09 | 0.08** | 2.35 | 0.13*** | 3.50 | 0.04 | 1.12 |
| 5 | 0.02 | 0.59 | 0.07** | 2.02 | 0.12*** | 3.19 | 0.05 | 1.37 |
| 6 | -0.04 | -1.00 | 0.06 | 1.52 | 0.12*** | 3.06 | 0.01 | 0.34 |
| 7 | -0.06 | -1.38 | 0.02 | 0.56 | 0.09** | 2.04 | 0.02 | 0.59 |
| 8 | -0.02 | -0.46 | 0.03 | 0.52 | 0.08* | 1.94 | 0.09** | 2.15 |
| 9 | 0.01 | 0.18 | 0.08* | 1.65 | 0.13*** | 3.11 | 0.12*** | 2.65 |
| 10 | 0.00 | -0.11 | 0.11** | 2.39 | 0.16*** | 3.72 | 0.10** | 2.18 |
| 11 | -0.01 | -0.31 | 0.10** | 2.27 | 0.15*** | 3.60 | 0.07* | 1.67 |
| 12 | -0.04 | -0.94 | 0.10** | 2.30 | 0.13*** | 3.02 | 0.02 | 0.40 |
| 13 | -0.07 | -1.59 | 0.04 | 1.04 | 0.07* | 1.65* | -0.03 | -0.75 |
| 14 | -0.04 | -0.86 | 0.00 | 0.04 | 0.02 | 0.43 | 0.00 | -0.05 |
| 15 | -0.06 | -1.27 | 0.03 | 0.69 | 0.04 | 0.83 | -0.03 | -0.59 |
| 16 | -0.04 | -0.88 | 0.01 | 0.24 | 0.00 | 0.00 | -0.01 | -0.30 |
| 17 | -0.06 | -1.50 | 0.02 | 0.56 | 0.01 | 0.34 | -0.05 | -1.24 |
| 18 | 0.00 | 0.03 | -0.02 | -0.30 | -0.03 | -0.80 | 0.01 | 0.15 |
| 19 | 0.01 | 0.13 | 0.04 | 1.31 | 0.02 | 0.48 | 0.00 | -0.07 |
| 20 | 0.03 | 0.78 | 0.04 | 1.22 | 0.01 | 0.33 | 0.01 | 0.15 |
| 21 | 0.02 | 0.70 | 0.06** | 1.90 | 0.03 | 0.78 | -0.01 | -0.21 |
| 22 | 0.03 | 0.97 | 0.04* | 1.70 | 0.02 | 0.57 | 0.00 | 0.00 |
| 23 | 0.03 | 0.96 | 0.04** | 1.84 | 0.03 | 1.01 | -0.01 | -0.53 |
| 24 | 0.01 | 0.31 | 0.02 | 1.23 | 0.02 | 0.84 | -0.02 | -0.73 |

***, **, * Statistically significant at a 99%, 95% and 90% confidence level, respectively (two-tailed).

Table 3.5 reports beta coefficient estimates and t-statistic values for four time-series regressions with 24 lags of aggregate investor confidence proxies as independent variables and the Fama-French SMB factor as dependent variables.

While all the modifications of the INVCON measure show significant relationships over several months, INVCON24 seems to best capture the hypothesised effect. We can observe a strong and positive relationship for around one year after a confidence impulse. In line with the interpretation of Hypothesis 2, as well as the findings of Statman et al. (2006) in the case of trading volume, aggregate investor confidence impulses seem to impact on market outcomes over several months.

Table 3.5 reports the beta coefficient estimates and *t*-statistic values for three time-series regressions with 24 lags of investor confidence proxies as independent variable, and the Fama-French SMB factor as a dependent variable. Investor confidence proxies are selected based on overall model significance as reported in Table 3.4.

The reported coefficients of the two Shiller indices do not suggest obvious patterns; that is, coefficients neither follow a particular trend, nor seasonality. This finding challenges earlier interpretations. However, the pattern produced by INVCON24 follows the rationale. That is, high levels of aggregate investor confidence are positively associated with SMB loading over several periods, before the effect vanishes after approximately one year.

Figure 3.3 illustrates this pattern. INVCON24 impulses are initially strong and positive, which persists for approximately one year. Beta coefficient estimates of 24 lags of both Shiller indexes appear to follow no particular pattern.

I repeat the analysis with the Fama-French HML factor as a dependent variable. Table 3.7 reports beta coefficient estimates and t-statistic values for 24 lags of all four modifications of the aggregate investor confidence index as independent variables and the Fama-French HML factor as dependent variables. As previously stated, most measures do not show a significant relationship, with INVCON6 being the exception.

The negative sign of the relationship, as anticipated, is, however, intuitively sound. Barber and Odean (2000) find, in their sample, that overconfident traders tend to have a slight preference for growth stocks. Thus, and

**Table 3.6: Alternative investor confidence proxies and the Fama-French SMB factor**

This table reports $t$-statistic significance values and beta coefficient estimates for three time-series regression models, with the Fama-French SMB factor as the dependent variable, and 24-month lags of the first differences of INVCON24, as well as the institutional investor version of Shiller's Crash Confidence Index ($Shil_{cr}$) as independent variables. The INVCON24 measure is selected as it possesses superior predictability over other modifications of the INVCON indices. The institutional investor version of Shiller's Crash Confidence index, as well as the individual version of Shiller's Valuation Confidence Index are selected due to their overall model significance, as reported in Table 3.4 above.

| | INVCON24 | | $Shil_{val}$ (ind.) | | $Shil_{cr}$ (inst.) | |
|---|---|---|---|---|---|---|
| Lags | Coeff. | $t$-stat. | Coeff. | $t$-stat. | Coeff. | $t$-stat. |
| 1 | 0.11*** | 5.15 | 0.00 | -0.08 | 0.01* | 1.73 |
| 2 | 0.15*** | 5.38 | 0.01 | 0.47 | 0.00 | 0.30 |
| 3 | 0.17*** | 4.98 | 0.00 | 0.15 | -0.02** | -2.34 |
| 4 | 0.13*** | 3.50 | -0.01 | -0.50 | 0.01 | 1.27 |
| 5 | 0.12*** | 3.19 | 0.02* | 1.82 | -0.01 | -1.39 |
| 6 | 0.12*** | 3.06 | -0.01 | -1.18 | 0.02* | 1.88 |
| 7 | 0.09** | 2.04 | 0.00 | -0.14 | 0.01 | 1.37 |
| 8 | 0.08* | 1.94 | -0.02 | -1.47 | 0.01 | 0.97 |
| 9 | 0.13*** | 3.11 | 0.00 | -0.40 | -0.02*** | -2.65 |
| 10 | 0.16*** | 3.72 | 0.01 | 0.65 | 0.02** | 2.06 |
| 11 | 0.15*** | 3.60 | 0.02* | 1.77 | -0.01 | -1.02 |
| 12 | 0.13*** | 3.02 | 0.00 | -0.09 | 0.00 | -0.15 |
| 13 | 0.07* | 1.65 | -0.01 | -0.94 | 0.01 | 0.74 |
| 14 | 0.02 | 0.43 | 0.00 | 0.31 | 0.00 | -0.17 |
| 15 | 0.04 | 0.83 | -0.02* | -1.71 | 0.00 | 0.27 |
| 16 | 0.00 | 0.00 | -0.01 | -0.93 | -0.01 | -0.86 |
| 17 | 0.01 | 0.34 | 0.02 | 1.58 | -0.01 | -0.66 |
| 18 | -0.03 | -0.80 | 0.01 | 0.53 | -0.01 | -1.22 |
| 19 | 0.02 | 0.48 | 0.01 | 1.00 | 0.02* | 1.98 |
| 20 | 0.01 | 0.33 | -0.02* | -1.70 | 0.00 | -0.14 |
| 21 | 0.03 | 0.78 | 0.00 | -0.29 | 0.00 | -0.10 |
| 22 | 0.02 | 0.57 | 0.00 | 0.25 | -0.01 | -1.34 |
| 23 | 0.03 | 1.01 | 0.01 | 0.63 | -0.01 | -0.75 |
| 24 | 0.02 | 0.84 | 0.00 | 0.41 | -0.02** | -2.34 |

***, ** and * Statistically significant at a 99%, 95% and 90% confidence level, respectively (two-tailed).

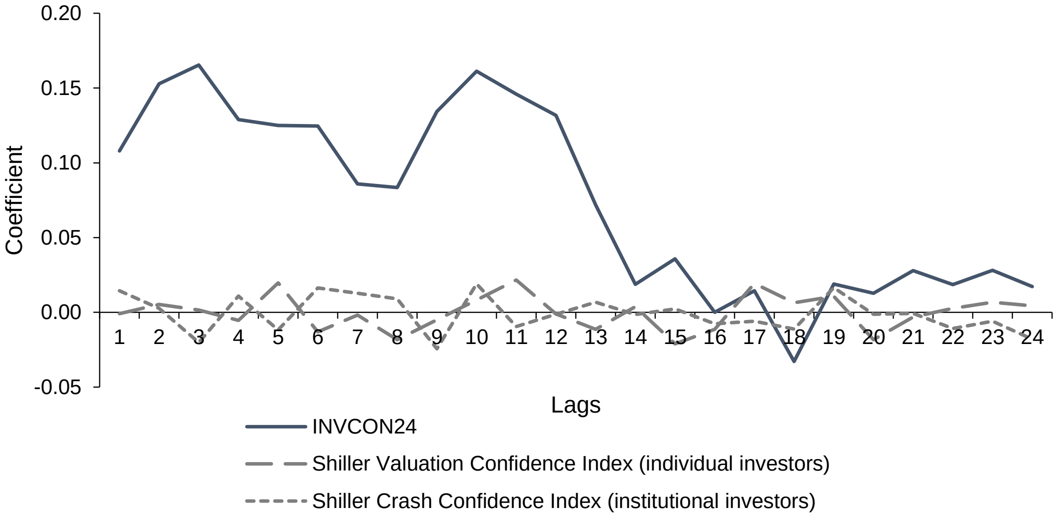



**Figure 3.3: Beta coefficients of selected time-series regression models and the Fama-French SMB factor**
This figure illustrates the beta coefficients of three time-series regression models with 24 lags of the first difference of INVCON24, the individual investor version of Shiller's Valuation Confidence Index, as well as the institutional investor version of Shiller's Crash Confidence Index as independent variables and the Fama-French SMB factor as a dependent variable. The INVCON24 measure possesses a strong and positive predictable association by the SMB factor for approximately one year subsequent to a confidence impulse. The predictable association of Shiller's indices does not follow a particular pattern. It must be noted that the beta coefficient estimates of the Shiller indices were multiplied with a factor of 10 for rescaling purposes.

loosely consistent, the negative sign of INVCON6 beta coefficients is intuitively sound. When aggregate investor confidence is high, investors tend to tilt toward growth stocks and away from value stocks.

Table 3.8 reports beta coefficient estimates and $t$-statistic values for two time-series regressions with 24 lags of investor confidence proxies as independent variables and the Fama-French HML factor as dependent variables. Investor confidence and proxies are selected based on their overall model significance reported in Table 3.4 above.

Shiller's Crash Confidence Index (Shil$_{cr}$) for institutional investors, which captures the belief that 'nothing can go wrong' in the near future, shows a positive relationship for the first three lags. However, an interpretation of the result is not anticipated to be straightforward. One may expect a negative association in that investors generally prefer growth stocks when they perceive

**Table 3.7: Investor confidence and the Fama-French HML factor**

This table reports $t$-statistic significance values and beta coefficient estimates for a series of time-series models, with the Fama-French HML factor as a dependent variable, and 24-month lags of INVCON6, INVCON12, INVCON24 and INVCON36 as independent variables.

| | INVCON6 | | INVCON12 | | INVCON24 | | INVCON36 | |
|---|---|---|---|---|---|---|---|---|
| Lags | Coeff. | $t$-stat. | Coeff. | $t$-stat. | Coeff. | $t$-stat. | Coeff. | $t$-stat. |
| 1 | 0.02 | 0.69 | 0.01 | 0.24 | 0.03 | 1.26 | 0.01 | 0.27 |
| 2 | -0.02 | -0.64 | 0.01 | 0.19 | 0.04 | 1.27 | -0.03 | -0.93 |
| 3 | -0.04 | -1.25 | -0.02 | -0.64 | 0.01 | 0.14 | -0.02 | -0.54 |
| 4 | -0.02 | -0.43 | -0.02 | -0.44 | 0.02 | 0.36 | 0.04 | 0.91 |
| 5 | -0.08* | -1.96 | 0.03 | 0.64 | 0.07 | 1.57 | 0.01 | 0.24 |
| 6 | -0.10** | -2.33 | -0.02 | -0.42 | 0.03 | 0.76 | 0.02 | 0.43 |
| 7 | -0.12** | -2.56 | -0.02 | -0.54 | 0.04 | 0.92 | 0.02 | 0.43 |
| 8 | -0.08 | -1.59 | -0.03 | -0.76 | 0.04 | 0.85 | 0.05 | 0.99 |
| 9 | -0.09* | -1.90 | -0.01 | -0.15 | 0.06 | 1.29 | 0.01 | 0.22 |
| 10 | -0.10** | -2.12 | -0.02 | -0.46 | 0.04 | 0.88 | -0.03 | -0.63 |
| 11 | -0.08 | -1.60 | -0.02 | -0.52 | 0.03 | 0.58 | -0.03 | -0.53 |
| 12 | -0.07 | -1.35 | 0.00 | -0.02 | 0.04 | 0.77 | -0.04 | -0.91 |
| 13 | -0.05 | -1.07 | 0.00 | 0.08 | 0.04 | 0.73 | -0.04 | -0.90 |
| 14 | -0.05 | -0.94 | 0.02 | 0.37 | 0.04 | 0.80 | -0.03 | -0.68 |
| 15 | -0.07 | -1.50 | 0.02 | 0.39 | 0.04 | 0.81 | -0.07 | -1.43 |
| 16 | -0.01 | -0.19 | -0.02 | -0.44 | 0.00 | -0.06 | -0.03 | -0.61 |
| 17 | 0.04 | 0.81 | 0.02 | 0.44 | 0.03 | 0.70 | -0.02 | -0.31 |
| 18 | 0.06 | 1.35 | 0.04 | 0.99 | 0.05 | 1.00 | -0.02 | -0.47 |
| 19 | 0.00 | 0.11 | 0.05 | 1.20 | 0.04 | 0.91 | -0.07 | -1.53 |
| 20 | 0.03 | 0.61 | -0.01 | -0.29 | -0.02 | -0.53 | -0.04 | -0.86 |
| 21 | -0.01 | -0.38 | 0.00 | 0.05 | 0.01 | 0.13 | -0.08** | -1.98 |
| 22 | -0.01 | -0.34 | -0.05 | -1.30 | -0.04 | -1.12 | -0.07 | -1.82 |
| 23 | 0.01 | 0.45 | -0.04 | -1.25 | -0.04 | -1.24 | -0.03 | -0.97 |
| 24 | 0.01 | 0.29 | -0.01 | -0.23 | -0.01 | -0.33 | -0.01 | -0.36 |

***, ** and * Statistically significant at a 99%, 95% and 90% confidence level, respectively (two-tailed).

the stock market to be safe. Instead, when institutional investors are certain that no stock market crash will occur, their portfolio preference tends to shift toward value stocks. This counterintuitive relationship may be a possible field for future research.

Figure 3.4 below plots the beta coefficient estimates of 24 lags of INVCON6 and the institutional investor version of Shiller's Crash Confidence Index. While the beta coefficient estimates of Shiller's Crash Confidence In-

**Table 3.8: Selected investor confidence proxies and the Fama-French HML factor**

This table reports $t$-statistic significance values and beta coefficient estimates for two time-series regression models, with the Fama-French HML factor as a dependent variable, and 24-month lags of the first differences of INVCON6 and as well as the institutional investor version of Shiller's Crash Confidence Index ($\text{Shil}_{cr}$) as independent variables. The INVCON6 measure is selected as it possesses superior predictability over other modifications of the INVCON indices. The institutional investor version of Shiller's Crash Confidence is selected due to its overall model significance, as reported in Table 3.4 above.

| | INVCON6 | | $\text{Shil}_{cr}$ (inst.) | |
|---|---|---|---|---|
| Lags | Coeff. | $t$-stat. | Coeff. | $t$-stat. |
| 1 | 0.02 | 0.69 | 0.02** | 2.08 |
| 2 | -0.02 | -0.64 | 0.02* | 1.80 |
| 3 | -0.04 | -1.25 | 0.02** | 1.99 |
| 4 | -0.02 | -0.43 | 0.00 | 0.18 |
| 5 | -0.08* | -1.96 | -0.01 | -0.74 |
| 6 | -0.10** | -2.33 | -0.01 | -1.41 |
| 7 | -0.12** | -2.56 | 0.01 | 1.33 |
| 8 | -0.08 | -1.59 | 0.01 | 1.09 |
| 9 | -0.09* | -1.90 | 0.00 | -0.43 |
| 10 | -0.10** | -2.12 | 0.01 | 1.01 |
| 11 | -0.08 | -1.60 | 0.01 | 0.69 |
| 12 | -0.07 | -1.35 | -0.01 | -0.94 |
| 13 | -0.05 | -1.07 | 0.00 | -0.13 |
| 14 | -0.05 | -0.94 | 0.01 | 1.46 |
| 15 | -0.07 | -1.50 | -0.01 | -0.62 |
| 16 | -0.01 | -0.19 | 0.01 | -0.90 |
| 17 | 0.04 | 0.81 | 0.01 | 0.80 |
| 18 | 0.06 | 1.35 | 0.00 | 0.07 |
| 19 | 0.00 | 0.11 | 0.00 | -0.25 |
| 20 | 0.03 | 0.61 | 0.01 | 1.42 |
| 21 | -0.01 | -0.38 | 0.01 | 1.28 |
| 22 | -0.01 | -0.34 | 0.00 | -0.07 |
| 23 | 0.01 | 0.45 | 0.00 | -0.61 |
| 24 | 0.01 | 0.29 | 0.00 | -0.14 |

***, ** and * Statistically significant at a 99%, 95% and 90% confidence level, respectively (two-tailed).

dex show no obvious pattern, INVCON6 impulses are inversely related to the HML factor for several months which is in line with Hypothesis 4.

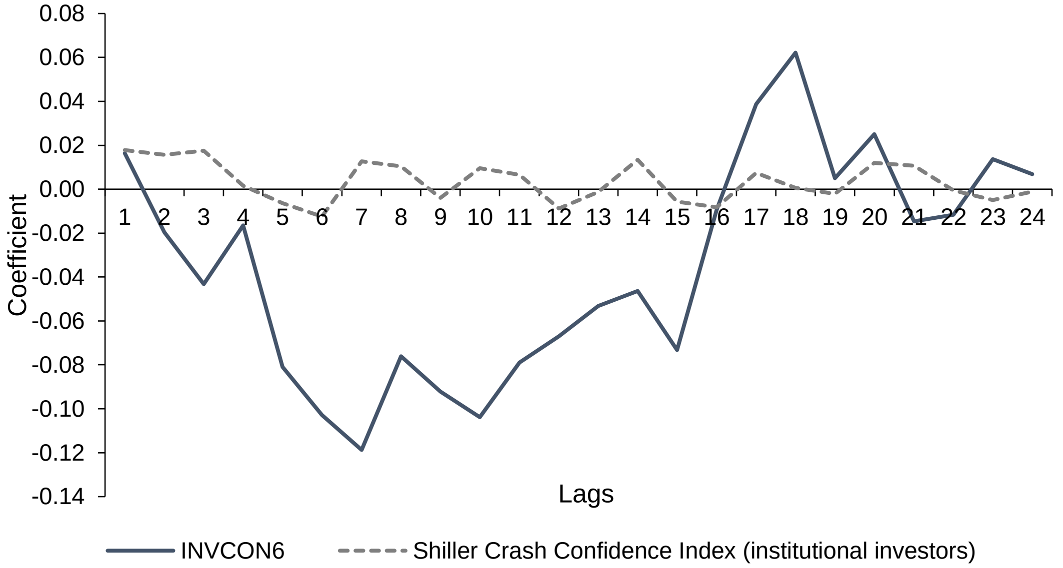



**Figure 3.4: Beta coefficients of selected time-series regression models and the Fama-French HML factor**
This figure illustrates the beta coefficients of three time-series regression models with 24 lags of the first difference of INVCON6, as well as the institutional investor version of Shiller's Crash Confidence Index as independent variables and the Fama-French HML factor as a dependent variable. The predictable association is positive for lags 1–3 for the institutional investor version of Shiller's Crash Confidence Index. The INVCON6 measure has no significant association with the HML factor for the first three lags, with this association turning to significant and negative between 5 and 10 lags. It must be noted that the beta coefficient estimates of the Shiller indexes were multiplied by a factor of 10 for rescaling purposes.

### 3.5.3 Robustness checks and encompassing tests

In order to test for potential look-ahead biases, the time-series regressions described above are repeated with raw INVCON scores which are not controlled for macroeconomic variables. Tables reporting these results can be found in the appendices for this chapter.

Generally, macroeconomic control variables appear to account for some of the explanatory power of these regression models. However, all models show similar patterns for both raw and cleaned index scores as independent variables, thus suggesting robust findings.

As an additional robustness check, I run encompassing regressions to compare the predictability of INVCON indices with the measures proposed in the prior literature for explaining momentum returns. Cooper et al. (2004) find that market states, defined as UP markets and DOWN markets, can partially explain variations in momentum returns. They define UP markets as a 3-year lagged non-negative market return. Consequently, I compute UP dummies at time $t$ as cumulative value-weighted market returns from $t - 36$ to $t - 1$ exceeding 0.

I first test the ability of the UP market dummy variable to predict variations in momentum profits, and subsequently perform an encompassing regression in order to to simultaneously test the UP market dummy variable and the INVCON24 measure.

**Table 3.9: Encompassing regressions: Investor confidence, market states and the Fama-French UMD factor**

This table reports the test results of three time-series regressions, with a market state dummy (UP) in the spirit of Cooper et al. (2004) INVCON24, as well as an encompassing combination of these as independent variables, and the Fama-French UMD factor as a dependent variable.

| Model | | 1 | 2 | 3 |
|---|---|---|---|---|
| INVCON24 | Coeff. | -0.10*** | | -0.10*** |
| | $t$-stat. | -3.30 | | -3.26 |
| UP | Coeff. | | 0.037*** | 0.04*** |
| | t-stat. | | 4.58 | 4.41 |
| | $F$-stat. | 2.31*** | 20.98*** | 3.04*** |
| | Adj. $R^2$ | 3.01% | 1.91% | 4.79% |

***, ** and * Statistically significant at a 99%, 95% and 90% confidence level, respectively (two-tailed).

Table 3.9 reports the output of these tests. Model 1 uses 24 lags of INVCON24 as independent variables. The model is significant at the 99% confidence level and an $F$-statistic of 2.31. Model 2 is a time-series regression with an UP market dummy variable as an independent variable. The first lag of the dummy variable is statistically significant ($F$=20.98), with a positive coefficient sign. The result meets the expectation that, in aggregate, UP markets are associated with higher degrees of overconfidence, as most investors hold long positions (Cooper et al., 2004).

In contrast to the UP dummy, the first lag of INVCON24 is negative, as discussed above. However, INVCON24 appears to do a better job of explaining variations in momentum profits, represented by a higher adjusted $R^2$.

Model 3 combines models 1 and 2 in an encompassing regression. Quite strikingly, both proxies remain statistically significant predictors of variations in momentum profits, suggesting that they capture different components. A combined model further increases the adjusted $R^2$ value to almost 5%.

The unreported factor tests with UP market dummy variables as independent variables and the Fama-French HML and SMB factors yield sound results. The UP dummy variable is significantly positive in its association with the SMB factor, and significantly negative in its association with the HML factor, which is in line with discussion above. However, the adjusted $R^2$ values are much lower compared to models using INVCON proxies as independent variables.

## 3.6 Conclusion

This study analyses the impact of aggregate investor confidence on price momentum, as well as the size and value premium, using AMEX, NYSE and NASDAQ data between 1927 and 2014. I document several key findings. Firstly, aggregate investor confidence is positively associated with the profitability of momentum strategies, which is consistent with the overconfidence hypothesis by Daniel et al. (1998) and Daniel et al. (2001). Although being better predictors of variation in momentum returns than proxies from prior literature, relatively low adjusted $R^2$ values suggest that price momentum may have several sources. Further research is required to shed light on this issue.

Secondly, more detailed analysis reveals that a confidence impulse has, on average, significantly positive impact on price momentum for up to 16 months.

This finding is consistent with the findings of Statman et al. (2006) who show that overconfidence drives trading activity for several periods. These new findings complement the literature by providing not only an empirical link between the overconfidence hypothesis by Daniel et al. (1998) and price momentum, it also provides insight into the anatomy of confidence formation and its impact on market outcomes. A confidence impulse is statistically notable after approximately three months in the case of price momentum, and reaches its pinnacle at around 10 months.

The third key finding is a positive association between aggregate investor confidence and the size premium. This finding is in line with the interpretations by Roll (1981) which propose that mis-assessment of risk is a possible source of the size effect, and with the suggestion of Barber and Odean (2001) that overconfident investors tend to tilt their portfolio preferences toward smaller, more risky stocks. In aggregate, investors tend to increase the proportion of small stocks in their portfolio when confidence is high, as reported in chapter 2. In contrast to price momentum,', a positive effect is notable immediately after a confidence impulse and is sustained for approximately one year.

This difference could be explained as follows. If confident investors systematically underestimate risk, they could immediately begin trading riskier stocks once the level of confidence increases. The case for price momentum is different. Given that the model of Daniel et al. (1998) holds with investors becoming overconfident about private information, the chain of events would be as follows. First, an investor becomes overconfident. Second, he receives pieces of private information which he then evaluates. Third, he 'chases' those prior decisions that turned out to be profitable, in retrospect. As this mechanism is more complex, it is likely that more time passes before the impact is notable.

Fourthly, I document an inverse relationship between aggregate investor confidence and the HML factor. Although the relationship is rather weak, the finding is sound and loosely consistent with Barber and Odean (2001) who suggest that overconfident investors have a slight preference for growth stocks.

This study provides empirical evidence in many domains. In aggregate, investor confidence is associated with the profitability of price momentum strategies. Confident investors tend to have a preference for stocks with low market capitalisation, as well as for growth stocks. The time between confidence formation and its effect on market outcomes tends to depend on the complexity of the phenomenon. Simple risk assessment tasks are influenced almost immediately, whereas more complex mechanisms, such as momentum formation, take longer. However, further research is necessary to verify this relatively speculative account.

Most alternative proxies of investor confidence or sentiment have little impact on the asset pricing factors investigated in this study, with some notable exceptions. First, the bullish component of the Investor Confidence Index of the American Association of Individual Investors (AAII) shows similar time-series patterns with the aggregate investor confidence index used in this study in their predictive behaviour toward the UMD factor. Second, the individual investor version of Shiller's 1-Year Confidence Index is positively associated with the profitability of momentum returns, but this is not the case with the institutional investor version.

Likewise, although Shiller's Crash Confidence Index for institutional investors is positively associated with size and value premiums, the individual version is not. These questions, as well as further investigation on the anatomy of confidence formation and subsequent trading behaviour, may be avenues for future research.

# Appendix

## 3.A Raw time-series regression outputs

**Table 3.A.1: Raw INVCON scores and the Fama-French UMD factor**

This table reports significance values and beta coefficient estimates of four time-series regressions, with 24 lags of raw scores of each INVCON version as independent variables and the Fama-French UMD factor as a dependent variable, as well as overall model significance and adjusted $R^2$ scores. In order to test for potential look-ahead bias, I use raw scores which are not filtered for macroeconomic variables. Similar model specifications suggest no presence of look-ahead biases.

| | INVCON6 | | INVCON12 | | INVCON24 | | INVCON36 | |
|---|---|---|---|---|---|---|---|---|
| Lags | Coeff. | $t$-stat. | Coeff. | $t$-stat. | Coeff. | $t$-stat. | Coeff. | $t$-stat. |
| 1 | -0.11 | -3.42*** | -0.10 | -3.01*** | -0.10 | -3.32*** | -0.12 | -3.83*** |
| 2 | -0.08 | -1.99** | -0.06 | -1.51 | -0.07 | -1.65* | -0.09 | -2.06** |
| 3 | 0.01 | 0.16 | 0.04 | 0.80 | 0.04 | 0.71 | 0.03 | 0.67 |
| 4 | 0.03 | 0.51 | 0.06 | 1.15 | 0.06 | 1.05 | 0.06 | 1.12 |
| 5 | 0.05 | 0.99 | 0.09 | 1.48 | 0.09 | 1.46 | 0.09 | 1.51 |
| 6 | 0.04 | 0.76 | 0.08 | 1.30 | 0.07 | 1.22 | 0.08 | 1.33 |
| 7 | 0.09 | 1.41 | 0.10 | 1.65 | 0.10 | 1.63 | 0.11 | 1.77* |
| 8 | 0.17 | 2.67*** | 0.16 | 2.57*** | 0.16 | 2.51*** | 0.16 | 2.52** |
| 9 | 0.17 | 2.60*** | 0.15 | 2.41*** | 0.15 | 2.29** | 0.14 | 2.17** |
| 10 | 0.19 | 3.02*** | 0.18 | 2.83*** | 0.17 | 2.70*** | 0.19 | 2.98*** |
| 11 | 0.17 | 2.68*** | 0.17 | 2.70*** | 0.17 | 2.56** | 0.22 | 3.50*** |
| 12 | 0.19 | 2.87*** | 0.17 | 2.63*** | 0.16 | 2.41** | 0.21 | 3.28*** |
| 13 | 0.13 | 2.04** | 0.12 | 1.85* | 0.10 | 1.57 | 0.17 | 2.62*** |
| 14 | 0.12 | 1.83* | 0.12 | 1.91* | 0.10 | 1.57 | 0.16 | 2.57** |
| 15 | 0.10 | 1.62 | 0.11 | 1.67* | 0.08 | 1.25 | 0.12 | 1.86* |
| 16 | 0.15 | 2.27** | 0.14 | 2.267** | 0.12 | 1.89* | 0.15 | 2.36** |
| 17 | 0.06 | 0.95 | 0.06 | 0.93 | 0.03 | 0.55 | 0.08 | 1.23 |
| 18 | 0.01 | 0.16 | 0.03 | 0.44 | 0.00 | -0.06 | 0.05 | 0.89 |
| 19 | 0.01 | 0.23 | 0.02 | 0.36 | -0.02 | -0.28 | 0.05 | 0.90 |
| 20 | 0.01 | 0.21 | 0.04 | 0.62 | 0.00 | -0.02 | 0.05 | 0.93 |
| 21 | 0.03 | 0.68 | 0.06 | 1.16 | 0.03 | 0.51 | 0.06 | 1.18 |
| 22 | 0.06 | 1.37 | 0.11 | 2.20** | 0.07 | 1.53 | 0.10 | 2.15** |
| 23 | 0.05 | 1.12 | 0.07 | 1.69* | 0.06 | 1.39 | 0.07 | 1.86* |
| 24 | 0.01 | 0.32 | 0.02 | 0.57 | 0.02 | 0.55 | 0.04 | 1.19 |
| | F-stat. | Adj. $R^2$ | $F$-stat. | Adj. $R^2$ | $F$-stat. | Adj. $R^2$ | $F$-stat. | Adj. $R^2$ |
| Mdl | 3.67*** | 5.90% | 2.80*** | 4.07% | 3.30*** | 5.17% | 3.54*** | 5.73% |

***, ** and * Statistically significant at a 99%, 95% and 90% confidence level, respectively (two-tailed).

**Table 3.A.2: Raw INVCON scores and the Fama-French SMB factor**

This table reports significance values and beta coefficient estimates of four time-series regressions, with 24 lags of raw scores of each INVCON version as independent variables and the Fama-French SMB factor as a dependent variable, as well as overall model significance and adjusted $R^2$ scores. In order to test for potential look-ahead bias, I use raw scores which are not filtered for macroeconomic variables. Similar model specifications suggest no presence of look-ahead biases.

| | INVCON6 | | INVCON12 | | INVCON24 | | INVCON36 | |
|---|---|---|---|---|---|---|---|---|
| Lags | Coeff. | $t$-stat. | Coeff. | $t$-stat. | Coeff. | $t$-stat. | Coeff. | $t$-stat. |
| 1 | 0.12 | 5.54*** | 0.10 | 4.73*** | 0.11 | 5.17*** | 0.12 | 5.514*** |
| 2 | 0.16 | 5.86*** | 0.14 | 4.92*** | 0.15 | 5.41*** | 0.16 | 5.66*** |
| 3 | 0.18 | 5.68*** | 0.14 | 4.26*** | 0.16 | 4.91*** | 0.16 | 4.87*** |
| 4 | 0.14 | 4.00*** | 0.10 | 2.67*** | 0.13 | 3.40*** | 0.12 | 3.27*** |
| 5 | 0.12 | 3.23*** | 0.08 | 2.16** | 0.12 | 2.98*** | 0.12 | 2.96*** |
| 6 | 0.10 | 2.63*** | 0.07 | 1.63 | 0.11 | 2.81*** | 0.11 | 2.80*** |
| 7 | 0.03 | 0.83 | 0.02 | 0.42 | 0.07 | 1.73* | 0.07 | 1.72* |
| 8 | -0.01 | -0.12 | 0.01 | 0.35 | 0.07 | 1.64 | 0.08 | 1.83* |
| 9 | 0.01 | 0.32 | 0.07 | 1.61 | 0.12 | 2.90*** | 0.14 | 3.15*** |
| 10 | 0.03 | 0.78 | 0.10 | 2.39** | 0.15 | 3.46*** | 0.15 | 3.55*** |
| 11 | 0.02 | 0.44 | 0.10 | 2.31** | 0.13 | 3.10*** | 0.13 | 3.03*** |
| 12 | 0.00 | 0.07 | 0.10 | 2.24** | 0.12 | 2.69*** | 0.11 | 2.45** |
| 13 | -0.04 | -0.87 | 0.04 | 0.99 | 0.05 | 1.25 | 0.04 | 0.92 |
| 14 | -0.06 | -1.49 | 0.00 | -0.04 | 0.00 | 0.09 | -0.01 | -0.25 |
| 15 | -0.03 | -0.65 | 0.02 | 0.49 | 0.02 | 0.52 | 0.02 | 0.55 |
| 16 | -0.05 | -1.20 | 0.00 | -0.10 | -0.01 | -0.27 | 0.00 | -0.04 |
| 17 | -0.03 | -0.75 | 0.01 | 0.23 | 0.01 | 0.12 | 0.01 | 0.31 |
| 18 | -0.07 | -1.72 | -0.03 | -0.73 | -0.04 | -0.96 | -0.03 | -0.79 |
| 19 | -0.01 | -0.35 | 0.03 | 0.76 | 0.01 | 0.36 | 0.02 | 0.51 |
| 20 | 0.00 | 0.04 | 0.03 | 0.78 | 0.01 | 0.33 | 0.01 | 0.33 |
| 21 | 0.03 | 0.86 | 0.05 | 1.31 | 0.03 | 0.77 | 0.02 | 0.65 |
| 22 | 0.03 | 0.91 | 0.03 | 1.05 | 0.02 | 0.55 | 0.01 | 0.46 |
| 23 | 0.03 | 1.17 | 0.04 | 1.30 | 0.03 | 1.00 | 0.02 | 0.69 |
| 24 | 0.03 | 1.21 | 0.02 | 0.75 | 0.01 | 0.69 | 0.01 | 0.29 |
| | $F$-stat. | Adj. $R^2$ | $F$-stat. | Adj. $R^2$ | $F$-stat. | Adj. $R^2$ | $F$-stat. | Adj. $R^2$ |
| Mdl | 3.67*** | 5.90% | 2.80*** | 4.07% | 3.30*** | 5.17% | 3.54*** | 5.73% |

***, ** and * Statistically significant at a 99%, 95% and 90% confidence level, respectively (two-tailed).

**Table 3.A.3: Raw INVCON scores and the Fama-French HML factor**

This table reports significance values and beta coefficient estimates of four time-series regressions, with 24 lags of raw scores of each INVCON version as independent variables and the Fama-French HML factor as a dependent variable, as well as overall model significance and adjusted $R^2$ scores. In order to test for potential look-ahead bias, I use raw scores which are not filtered for macroeconomic variables. Similar model specifications suggest no presence of look-ahead biases.

| | INVCON6 | | INVCON24 | | INVCON24 | | INVCON36 | |
|---|---|---|---|---|---|---|---|---|
| Lags | Coeff. | $t$-stat. | Coeff. | $t$-stat. | Coeff. | $t$-stat. | Coeff. | $t$-stat. |
| 1 | 0.03 | 1.28 | 0.01 | 0.46 | 0.03 | 1.23 | 0.03 | 1.44 |
| 2 | 0.04 | 1.46 | 0.02 | 0.47 | 0.04 | 1.32 | 0.04 | 1.24 |
| 3 | 0.01 | 0.20 | -0.02 | -0.57 | 0.01 | 0.21 | 0.00 | 0.01 |
| 4 | -0.01 | -0.39 | -0.02 | -0.56 | 0.02 | 0.42 | 0.01 | 0.23 |
| 5 | 0.01 | 0.27 | 0.02 | 0.56 | 0.08 | 1.72 | 0.06 | 1.37 |
| 6 | -0.06 | -1.29 | -0.02 | -0.53 | 0.05 | 0.99 | 0.03 | 0.62 |
| 7 | -0.08 | -1.82* | -0.03 | -0.63 | 0.06 | 1.19 | 0.04 | 0.81 |
| 8 | -0.11 | -2.25** | -0.04 | -0.76 | 0.06 | 1.22 | 0.04 | 0.76 |
| 9 | -0.07 | -1.42 | 0.00 | -0.09 | 0.08 | 1.72* | 0.06 | 1.28 |
| 10 | -0.09 | -1.79* | -0.02 | -0.39 | 0.06 | 1.22 | 0.02 | 0.51 |
| 11 | -0.10 | -2.03** | -0.03 | -0.54 | 0.04 | 0.85 | -0.02 | -0.36 |
| 12 | -0.07 | -1.42 | 0.00 | -0.08 | 0.05 | 1.04 | -0.02 | -0.32 |
| 13 | -0.06 | -1.23 | 0.00 | -0.03 | 0.04 | 0.91 | -0.04 | -0.71 |
| 14 | -0.05 | -0.99 | 0.01 | 0.15 | 0.04 | 0.90 | -0.03 | -0.71 |
| 15 | -0.04 | -0.83 | 0.01 | 0.28 | 0.04 | 0.88 | -0.02 | -0.46 |
| 16 | -0.07 | -1.55 | -0.03 | -0.59 | 0.00 | -0.10 | -0.06 | -1.26 |
| 17 | -0.01 | -0.26 | 0.02 | 0.33 | 0.03 | 0.65 | -0.02 | -0.44 |
| 18 | 0.03 | 0.72 | 0.04 | 0.94 | 0.04 | 0.94 | -0.01 | -0.17 |
| 19 | 0.05 | 1.26 | 0.04 | 1.00 | 0.04 | 0.80 | -0.02 | -0.38 |
| 20 | 0.00 | 0.04 | -0.02 | -0.46 | -0.03 | -0.64 | -0.06 | -1.42 |
| 21 | 0.03 | 0.68 | 0.00 | 0.11 | 0.00 | 0.06 | -0.02 | -0.61 |
| 22 | -0.01 | -0.33 | -0.04 | -1.23 | -0.04 | -1.19 | -0.06 | -1.80* |
| 23 | -0.01 | -0.27 | -0.04 | -1.31 | -0.04 | -1.30 | -0.05 | -1.73* |
| 24 | 0.01 | 0.57 | -0.01 | -0.31 | -0.01 | -0.35 | -0.02 | -0.77 |
| | $F$-stat. | Adj. $R^2$ | $F$-stat. | Adj. $R^2$ | $F$-stat. | Adj. $R^2$ | $F$-stat. | Adj. $R^2$ |
| Mdl | 1.80** | 1.83% | 1.45* | 1.05% | 1.48* | 1.13% | 1.53** | 1.25% |

***, ** and * Statistically significant at a 99%, 95% and 90% confidence level, respectively (two-tailed).

# Chapter 4

# Investor overconfidence in experimental asset markets across market states

Christoph Meier (contribution: 80%)
Lurion DeMello (contribution: 20%)

**ABSTRACT**

In this study, we explore how individual overconfidence adjusts after receiving extreme feedback that either supports or contradicts prior decision making. We find that highly contradicting feedback causes overconfidence to vanish as 'confidence crashes', while supportive signals cause overconfidence to increase. Further evidence suggests that strong feedback impulses are associated with higher investor disagreement, supporting prior hypotheses that investors interpret such impulses differently. We also find that methodologies that measure overconfidence in prediction tasks systematically overstate confidence scores, as respondents tend to fail to internalise stated confidence intervals appropriately.

*Ethics Committee approval has been obtained for the collection of the data used in this research (approval number: 5201600242).*

## 4.1 Introduction

Overconfidence is one of the most robust findings in the field of behavioural finance (DeBondt and Thaler, 1994a; Daniel and Hirshleifer, 2015). A multitude of empirical studies document that overconfidence is associated with poor financial decision making (e.g. Odean, 1999, 1998; Malmendier and Tate, 2008, 2005). Experimental studies yield evidence that is consistent with this notion.[1]

In contrast to the traditional assessment of individual overconfidence in experiments[2], a new methodology proposed by Glaser et al. (2013) allows the continuous assessment of individual overconfidence in the *same* task. This property provides an opportunity to measure *changes* in overconfidence after the arrival of feedback that either reinforces or contradicts prior decisions.

The aim of this study is to investigate the adjustment process of investor overconfidence after the arrival of extreme feedback signals. We find that, on average, overconfidence vanishes when highly contradicting signals about a prior decision arrives, resulting in 'confidence crashes'. In contrast, new evidence that supports a decision causes overconfidence to persist. However, if the nature of these evidence signals reverses, lost overconfidence can re-emerge.

In line with the hypotheses of Hong and Stein (2007) and Scheinkman and Xiong (2003), we find that, after the arrival of strong information signals, disagreement among market participants increases, particularly when such information contradicts one's prior decisions. We interpret the pattern as reflective of the heterogeneity of individual personality traits through which some agents disregard information that conflicts with their prior beliefs and consequently fail to update their stock valuations appropriately. As a consequence, higher

[1] See, for example, Glaser, Langer, and Weber (2010); Cesarini, Sandewall, and Johannesson (2006); Hilton, Regner, Cabantous, Charalambides, and Vautier (2011); Kirchler and Maciejovsky (2002) and Soll and Klayman (2004).

[2] The method compares the number of confidence intervals produced by a participant where the boundaries *include* the true value of a knowledge question with the number of intervals that *should* be stated correctly given a particular confidence level.

time-series variance can be observed after such events.

Complementing the findings of Langnickel and Zeisberger (2016); Teigen and Jørgensen (2005); and Cesarini et al. (2006) who document that respondents of knowledge questions are insensitive to confidence levels, we find that the methodology of Glaser et al. (2013) is likely to produce inflated overconfidence scores for stock price prediction tasks. Self-reported confidence levels tend to be significantly *lower* than those asked from respondents during assessment tasks. As a result, individuals are assumed to be more overconfident than they actually are.

The remainder of this paper is organised as follows. Section 2 briefly summarises the relevant literature and outlines the research gaps. Section 3 develops the hypotheses, while Section 4 outlines the research design. Section 5 reports and discusses the results, while Section 6 presents the conclusion.

## 4.2 Literature review

The literature yields rich evidence on how overconfidence affects individual decision making. Individuals tend to be too optimistic about their estimations, aptitude, attributes and forecasts. That is, people perceive themselves as better than they actually are, with this commonly called overconfidence (Tversky and Kahneman, 1974).

Moore and Healy (2008) argue that three distinct types of overconfidence are typically treated as the same concept: overestimation, overplacement and overprecision.

Overestimation is defined as being overly optimistic about one's own abilities. For instance, students overestimate their performance in exams (Clayson, 2005); young drivers overestimate their driving skills (Gregersen, 1996); children overestimate their physical abilities (Plumert and Schwebel, 1997); physicians overestimate their patients' medical literacy (Kelly and Haidet, 2007); and public speakers overestimate the effectiveness of their communication

(Keysar and Henly, 2002).

Overplacement refers to the 'better-than-average' effect. That is, individuals tend to estimate their own abilities and attributes to be better than the median of those abilities and attributes in the population (Larrick, Burson, and Soll, 2007). College students believe they are, for instance, more polite, intelligent and reliable, but that they are less of a liar, less disrespectful and less unpleasant than the average college student population (Alicke and Govorun, 2005; Alicke, Klotz, Breitenbecher, Yurak, and Vredenburg, 1995). Among American drivers, 93% believe that they possess driving skills that are above the average level of skills of the driving population, as do 69% of Swedish drivers (Svenson, 1981).

The third type of overconfidence, and the one most relevant to this study, is overprecision. That is, individuals tend to be too optimistic about the accuracy of their beliefs. Typically, overprecision is measured by asking participants to state a confidence interval around a point estimate answer to a numerical question, or a future estimate of a value. These intervals are subsequently shown to be too narrow. For instance, news vendors tend to be too certain about the precision of their sales forecasts (Ren and Croson, 2013).

Likewise, participants of experiments tend to be overly confident about the precision of their forecasts, such as the price of a car, box office gross of a movie or the overall quality score of a college (Soll and Klayman, 2004). Due to their simplicity, confidence interval estimates are a common tool to assess the overconfidence of agents in experimental asset markets. However, these questions only allow a single assessment per task, as once the true answer of a numerical knowledge question is known, confidence interval estimation loses its purpose. As a result, the design of this 'classic' approach to overconfidence assessment allows between-subject variation of overconfidence, but prohibits the exploration of the process of individuals' confidence calibration over time in an ongoing task.

As a remedy, Glaser et al. (2013) design a methodology of "true overconfidence" assessment, utilising the properties of an experimental asset market. In their methodology, participants observe the chart of a stock price, and, subsequently predict confidence intervals of the stock price at a point in the future. As incremental stock price changes follow one of two possible distributions with given probabilities, we can compute the exact confidence intervals of future stock prices. However, as this computation is fairly complex (a more detailed description is given below), participants are unable to compute these confidence intervals during the experiment.

The main objectives of this paper are threefold. Firstly, we investigate within-subject variation in overconfidence over time. This domain is relevant, as theory suggests an increase in overconfidence after recurring positive feedback due to self-attribution bias, and a decrease in the level of confidence after conflicting feedback (Daniel et al., 1998).

Secondly, we test the hypotheses of Hong and Stein (2007) and Scheinkman and Xiong (2003) who suggest that strong news signals cause investor disagreement about stock valuation and, hence, high time-series variance.

Thirdly, this paper tests the potential shortcomings of the methodology of Glaser et al. (2013), following the suggestions of Langnickel and Zeisberger (2016) and Teigen and Jørgensen (2005) who find systematic overestimation of overconfidence levels in general knowledge question tasks due to the insensitivity of respondents to different confidence levels.

## 4.3 Hypotheses development

The motivation of this study is grounded in the hypotheses of Gervais and Odean (2001); Odean (1999); and Cooper et al. (2004), that overconfidence among investors should be higher after bullish markets, as most hold long positions and interpret recent portfolio gains as positive feedback crediting their ability to pick stocks. Therefore,

**Hypothesis 1a:** *Investors who hold long positions become overconfident in market booms,*

and

**Hypothesis 1b:** *Investors who hold short positions lose their overconfidence in market booms.*

Figure 4.1 below summarises the calibration process of confidence after the arrival of new information. The arrival of supporting evidence leads to the reinforcement of one's confidence (as illustrated in area a, see Koriat, Lichtenstein, and Fischhoff (1980)).

Conversely, conflicting evidence should result in a decrease in the level of confidence. However, due to self-attribution bias, the reduction of confidence is disproportionately lower, as the individual seeks external reasons for failure and therefore, fails to appropriately calibrate his confidence (as illustrated in area b) (Bem, 1965; Miller and Ross, 1975). When this is applied to stock markets, Odean (1998) and Gervais and Odean (2001) suggest that market losses reduce aggregate investor confidence, but possibly in an asymmetrical fashion.

The novelty of this paper is adding a third condition to the calibration of confidence: confidence crashes (illustrated in area c). Sunstein and Zeckhauser (2011) and Loewenstein, Weber, Hsee, and Welch (2001) suggest that individuals tend to overreact when being confronted with fearsome risk, especially in highly vivid situations, such as realising extreme losses subsequent to market crashes. The rationale of these 'confidence crashes' follows from the coexistence of self-attribution bias and overconfidence (e.g. Daniel et al. (1998)). If investors suffer from self-attribution bias, they will credit their own aptitude and skill for gains, but blame external factors for failure. Consequently, they will become increasingly overconfident over time. It is plausible that the effect may persist for longer time periods but that, eventually, re-calibration of one's

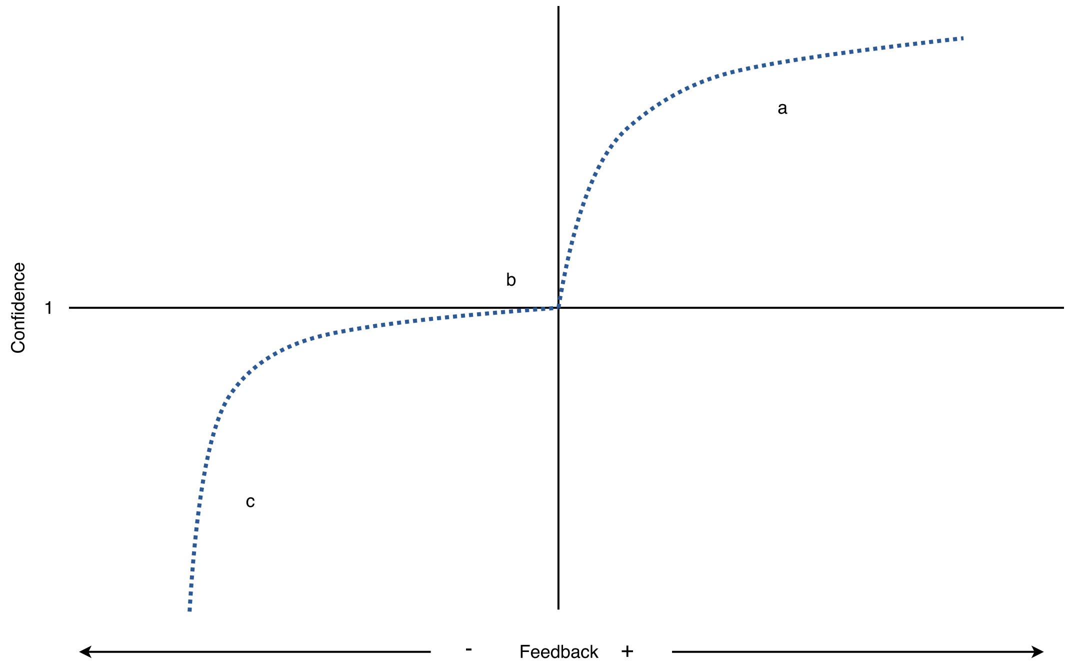



**Figure 4.1: Confidence adjustment to feedback.**
Confirming evidence of prior decisions leads to an increase in one's level of confidence (area a). This effect leads to overconfidence over time. Due to self-attribution bias, an individual will disproportionately adjust her confidence downwards (area b) when feedback is controversial, thus a kink appears in the curve. However, extremely poor feedback will lead to a harsh drop in confidence (area c).

confidence occurs. Therefore, we propose 'confidence crashes' as one potential source of such re-calibration.

We interpret the process of confidence crashes as follows. After receiving extremely conflicting evidence, investors are temporarily 'stunned' and—due to extreme losses—become insecure about their own abilities in estimating future prices in asset markets. In other words, the confidence intervals of their forecasts widen dramatically upon arrival of alarmingly dis-confirming evidence.

As the interpretation of arriving signals depends on prior decisions, investors who hold long positions should lose their overconfidence, and those who hold short positions should become overconfident during stock market crashes.

Therefore,

**Hypothesis 2a:** *Stock market crashes cause overconfident investors who hold long positions to lose their overconfidence,*

and

**Hypothesis 2b:** *Stock market crashes cause investors who hold short positions to become overconfident.*

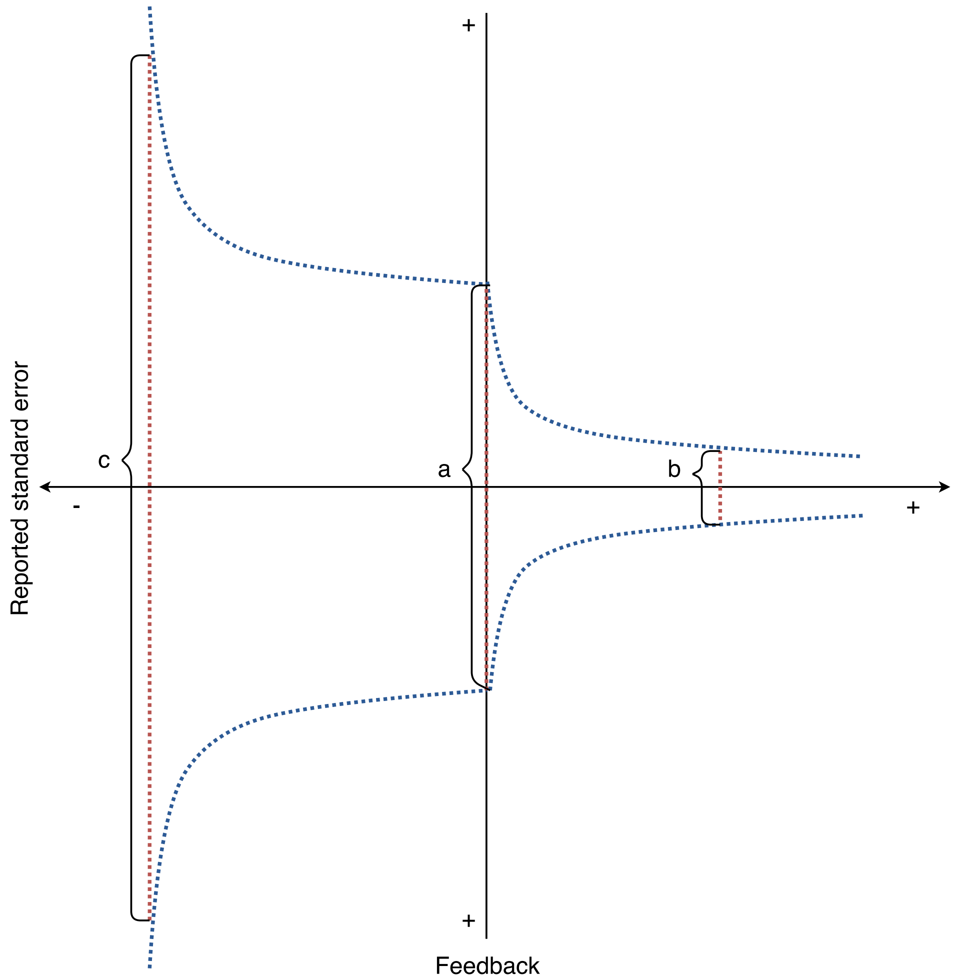



**Figure 4.2: Confidence interval adjustment to feedback.**
This diagram illustrates the proposed adjustment of the width of an individual's confidence interval of their belief about future stock prices.

Figure 4.2 translates the calibration process from Figure 4.1 and describes the domain of this paper. Initially, the individual will start with a moderately wide confidence interval (interval a), which disproportionately narrows down after the arrival of confirming evidence, thus the overly optimistic and narrow confidence interval (interval b). After the arrival of mildly dis-confirming information, confidence intervals widen at a lower pace than is the case for confirming evidence of a similar magnitude, thus the kink that occurs at neutral feedback. However, the individual dramatically widens her confidence interval after the arrival of extremely dis-confirming evidence (interval c), due to fear and anxiety.

We further test a hypothesis by Hong and Stein (2007) and Scheinkman and Xiong (2003) who propose increased time-series variance in stock prices after the arrival of strong news impulses. They argue that investors interpret such impulses differently, resulting in investor disagreement. In other words, investors have different opinions what the true value fo a stock would be. Consequently, we should expect high variation of stock price valuations subsequent to the arrival of strong information signals. Analogously, we expect that the variance in stock price valuation should be *high* when feedback impulses are *strong*.

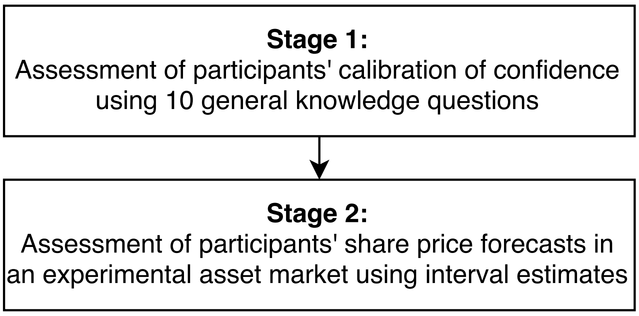
## 4.4 Experimental design

### 4.4.1 Experimental procedure

This study designs an experimental asset market where participants can buy or sell an asset, with the asset following a given distribution of returns. Furthermore, participants are asked to provide confidence interval estimates of their price forecasts. The experiment consists of two stages, as illustrated in Figure 4.1 below. The first stage is an initial assessment of confidence. The purpose is to assess if individuals who are diagnosed as overconfident during the first stage of the experiment also begin the second stage with overconfidence. The second stage assesses the within-subject adjustment of confidence interval forecasts in different market states of an experimental asset market, as well as testing the hypotheses.

**Figure 4.1: Structure of the experiment**

The experiment consists of two stages. The first stage assesses the participants' initial calibration of confidence, using 10 general knowledge questions. Participants have to estimate confidence intervals that include the true value 90% of the time. The second stage assesses the participants' accuracy of forecasting future stock prices in an experimental asset market through several market states.




The two-stage design of the experiment is adapted from Glaser and Weber (2007) and Glaser et al. (2013) who initially assess participants' level of overconfidence using general knowledge questions before applying the new methodology for assessment of overconfidence among investors in an experimental asset market. The authors use general knowledge questions where respondents are asked to state a range (upper and lower bound), within which they believe the true value of a numerical knowledge question lies with 90% confidence. Table 4.1 lists the general knowledge questions used in this study, with these having been adapted from Glaser et al. (2010).[3] As the overconfidence assessment questions by Glaser et al. (2010) are phrased in a European context, we make some adjustments to align with local context.

Subsequent to Stage 1, participants are asked to provide confidence interval estimates of future share prices. While the methodology is largely adapted from Glaser et al. (2013), we add some features in order to test our hypotheses. For instance, participants can choose to either 'buy' or 'sell' the stock initially. The purpose of this decision is to mirror 'positive' and 'negative' feedback

[3] A complete catalogue of these questions is available upon request. We thank Markus Glaser for sharing these questions and for general helpful advice in regard to the design of this experiment.

**Table 4.1: General knowledge questions**

During Stage 1 of the experiment, participants are asked to estimate 90% confidence intervals for numeric answers of 10 general knowledge questions. An appropriately confident participant should define nine confidence intervals that capture the correct answer.

We are interested in your judgement in regard to different figures (sizes, lengths, intervals, . . . ).
We ask you to provide your judgement by specifying an upper bound and a lower bound. You should specify these boundaries in such a way, that the true answer appears as in the example of the first question below:

The age of William Shakespeare at his death is *in your opinion*:
- almost certainly (i.e. with 95% probability) above the lower bound, and,
- almost certainly (i.e. with 95% probability) below the upper bound.

Differently stated, we ask you to provide an interval that contains the correct answers with 90% probability:

| | Lower bound | Upper bound |
|---|---|---|
| Age of William Shakespeare at his death: | | |
| Length of the Mississippi River (in km): | | |
| Total number of medals awarded to all participants during the Winter Olympic Games in Sochi 2014: | | |
| Average number of rainy days per year in Bergen (Norway): | | |
| Weight (in kg) of an empty Airbus A380: | | |
| Height of the Eiffel Tower (in m): | | |
| Duration of the pregnancy (in days) of a koala: | | |
| Diameter of the moon (in km): | | |
| Total number of Premier League goals scored by David Beckham: | | |
| Grams of sugar in a 1.5 litre bottle of Coca Cola: | | |

impulses if the price of an asset increases or decreases.

Figure 4.2 illustrates the interface for the first round of the forecasting stage of the experiment. Participants are asked to observe a price chart and are given information about the return distributions. After deciding to either 'buy' or 'sell' the stock, participants continue to the actual forecasting rounds, as illustrated in Figure 4.3. During each of the 12 forecasting rounds, participants are asked to define a confidence interval for $t + 9$, within which they believe the prices of the stock will lie 90% of the time.

The experimental market undergoes a range of market states which are shown in Figure 4.4. Initially, the market is relatively stable. Subsequently, the markets will turn bullish, with participants who initially decided to long (short) the share are confronted with confirming (dis-confirming) feedback about their prior decisions. Analogously, we can observe if individuals who decided to long or short the stock will remain overconfident. The third state consists of a market crash where the positive trend abruptly reverses. Mirroring stimuli across buyers and sellers allows us to explore the confidence adjustment process for both sudden crash and boom situations which addresses the confidence crash hypothesis.

### 4.4.2 Computing overconfidence

#### Knowledge questions

The methodology to compute overconfidence in stage 1 of the experiment is simple. We ask 10 general knowledge questions where participants provide confidence intervals as answers. We then count the number of stated confidence intervals $C$ that contain the correct value, and can compare these to the number of confidence intervals that should be stated correctly based on a given confidence level $K$ (for instance, 9 out of 10 for a confidence level of 90%). Thus, the level of overconfidence $OCCI$ can be expressed as

**Figure 4.2: Interface of initial participant decision**

Participants are presented with a complex distribution of price changes, as well as a price chart. In the first round, they are asked to either buy or sell the security. In all subsequent rounds, the buy and sell buttons, as well as the respective text, will disappear. Adopted from Glaser et al. (2013).

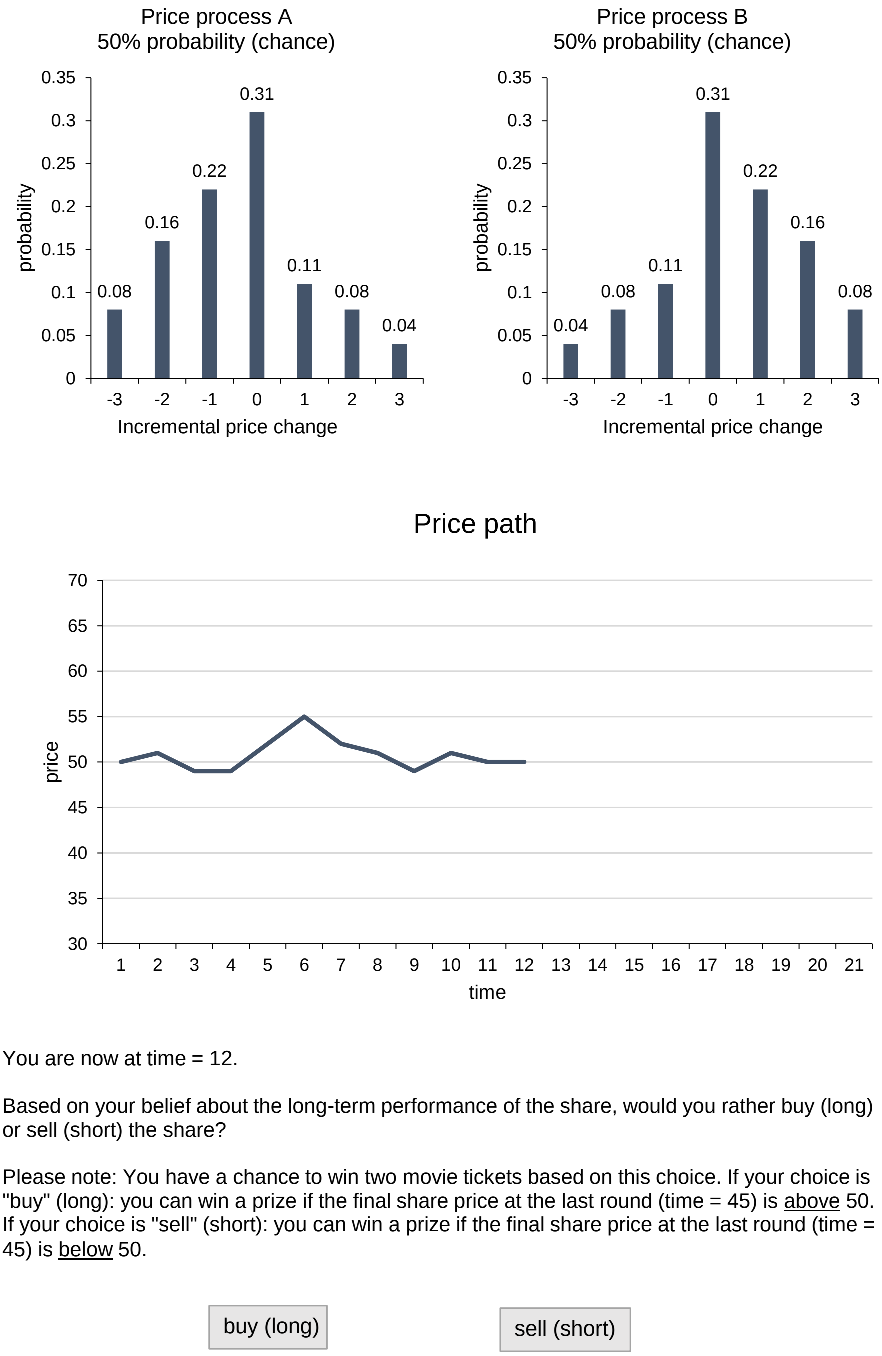


**Figure 4.3: Interface of price forecast assessment**

After the participant decided to either 'buy' or 'sell' the asset, he proceeds to the forecasting stage of the experiment. In each round, the participant observes a price change and subsequently estimates a 90% confidence interval for the stock's price in $t + 9$. Adapted from Glaser et al. (2013).

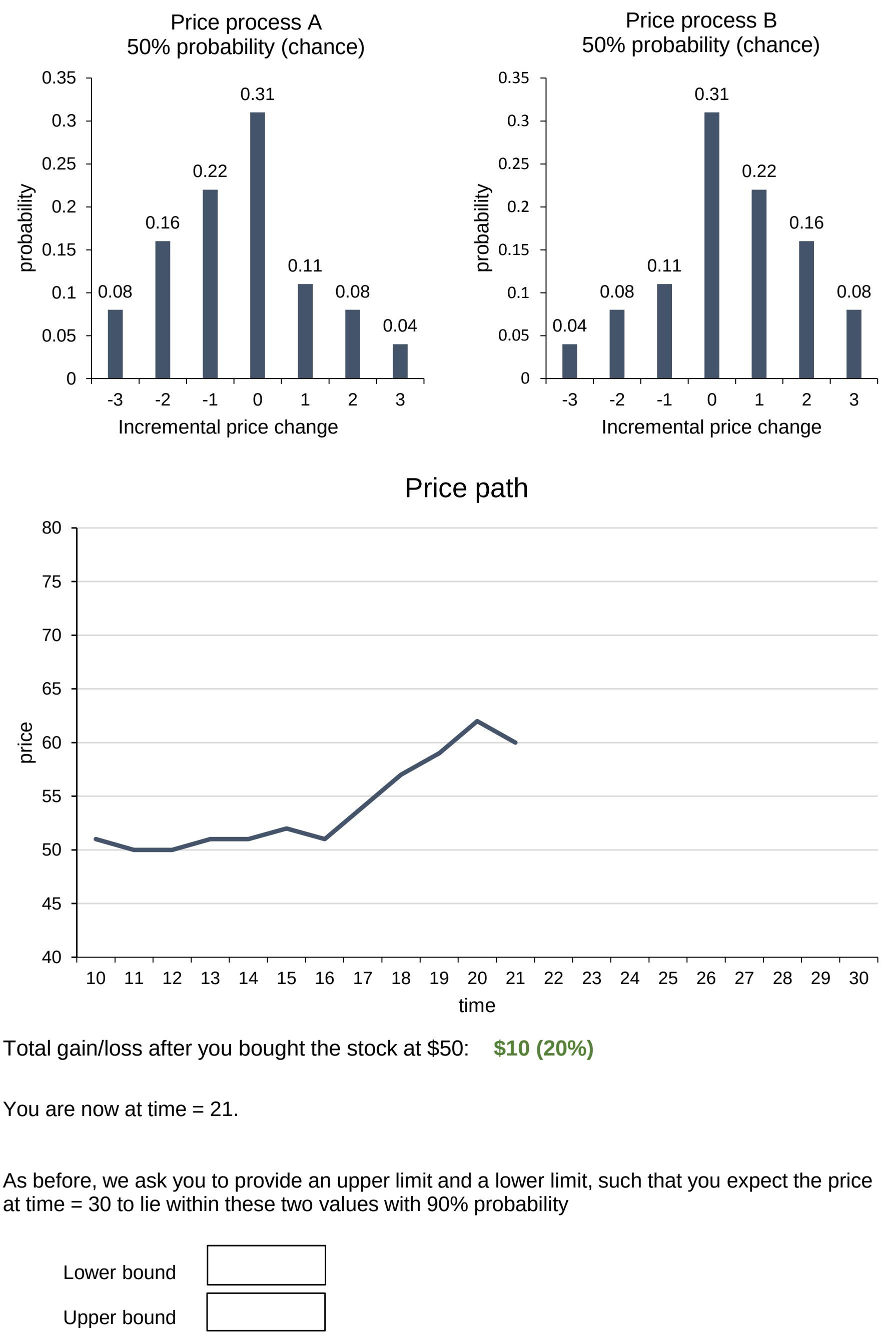
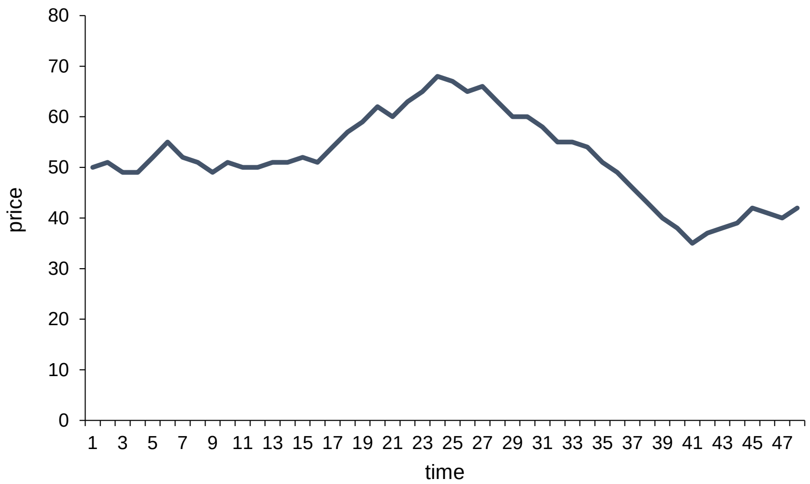


**Figure 4.4: Price path throughout the experiment**

Throughout the experiment, the experimental asset market undergoes various states. Initially, the market is relatively stable before experiencing a boom. Subsequently, the market crashes, before it recovers near the end of the experiment.




$$OCCI = K - C. \tag{4.1}$$

That is, if a participant defines only five of the 10 confidence intervals correctly, his level of overconfidence ($OCCI$) is four (10% error margin, or one wrong confidence interval out of 10 questions is allowed at a 90% confidence level). If he defines seven intervals correctly, his level of overconfidence ($OCCI$) is two, respectively. Consequently, *positive* (*negative*) scores reflect *overconfidence* (*underconfidence*).

However, Glaser et al. (2013) acknowledge that this measure has a number of weaknesses. Firstly, very narrow confidence intervals can be due to two conditions. The participant can either be rightfully confident, as he knows the correct answer with high certainty, or he does not know the answer, but is simply lucky. Thus, this method fails to distinguish between knowledgeable (thus confident) and ignorant yet lucky participants.

Furthermore, this method produces an average confidence score for several unrelated general knowledge questions. That is, it is possible to be overconfident for some estimates, but underconfident for others. As a result, overall confidence may appear to be well calibrated which, obviously, is not the case (Glaser et al., 2013).

Langnickel and Zeisberger (2016) express the criticism that participants are typically insensitive to changes in the confidence levels requested. In particular, participants tend to fail to adjust confidence intervals when the requested confidence levels change. Following Langnickel and Zeisberger (2016), we include a self-assessment question which asks the participants to state the expected number of correctly defined confidence intervals $S$ that he or she provided. We calculate an overconfidence score based on self-reported confidence level $OCSR$ as

$$OCSR = S - C. \tag{4.2}$$

**Artificial charts**

Measuring overconfidence in stock price forecasts using artificial charts is not as straightforward as the use of knowledge questions for numerical answers. Glaser et al. (2013) develop a new method to measure overconfidence in this domain which is adopted in this study. Participants initially observe an artificial stock price chart over 12 periods. The distribution $D$ of possible future stock returns is either positively ($k+$) or negatively ($k-$) skewed. In other words, the price of the stock has either a positive or negative long-term trend, thus a positive or negative expected value. Participants are informed that one of the two distributions is randomly picked. Furthermore, the two distributions are both numerically and graphically presented to the participants. The possible price changes are -3, -2, -1, 0, +1, +2 and +3.

The distributions $D_{k-}$ and $D_{k+}$ of the two price paths are illustrated in table 4.2. $D_{k-}$ has a negative trend and $D_{k+}$ has a positive trend, respectively.

The subscript of $D_k$ reflects the odds that an outcome of +1 is from distri-

**Table 4.2: Incremental price changes**

This table summarises a distribution of incremental price changes per round. The future price process either follows a negative trend $D_{k-}$ or a positive trend $D_{k+}$.

| Price change | -3 | -2 | -1 | 0 | +1 | +2 | +3 |
|---|---|---|---|---|---|---|---|
| $D_{2-}$ | 8% | 16% | 22% | 31% | 11% | 8% | 4% |
| $D_{2+}$ | 4% | 8% | 11% | 31% | 22% | 16% | 8% |

bution $D_{k+}$ rather than $D_{k-}$. Glaser et al. (2013) further assume that incremental price changes are independent from prior price changes.

In the beginning of Stage 2 of the experiment, participants observe price movements from $t = 0$ until $t = 12$. Subsequently, they are asked to state their estimated confidence interval for the share price at $t = 21$ with 90% probability. In the following round at $t = 15$, participants observe another price change and are asked to provide a confidence interval for $t = 24$ and so on, until the final round for $t = 48$, resulting in 12 prediction rounds.

The correct distribution for a price in period 24 ($\pi_{24}$) can be derived with the distribution pair [$D_{k-}$;$D_{k+}$] from price $\pi_{15}$. Glaser et al. (2013) design this methodology so the relevant probabilities do not depend on the price path, but only the total price change since $c_{15} = (\pi_{15} - 50)$. Therefore,

$$p_k(pos.|c_{15}) = \frac{k^{c_{15}}}{1 + k^{c_{15}}} \tag{4.3}$$

and

$$p_k(neg.|c_{15}) = 1 - (pos.|c_{15}) = \frac{1}{1 + k^{c_{15}}}. \tag{4.4}$$

For instance, given that the chart in figure 4.3 above was generated with $D_{2-}$ or $D_{2+}$, and has a price of $\pi = 52$, we can compute a total price change of $c_{15} = (\pi_{52} - 50) = (52 - 50) = 2$. Consequently, using equation 4.3, we can compute the likelihood that $D_{2+}$ is the underlying distribution of the price path instead of $D_{2-}$ with $p_2(pos.|2) = \frac{2^2}{(1+2^2)} = 80\%$. Therefore, the likelihood that the price path was generated by $D_{2-}$ is $1 - (pos.|2) = 20\%$.

Subsequently, Glaser et al. (2013) determine the distribution of price from $t_{20}$ to $t_{40}$ following $D_{2+}$, which they define with

$$p_k(pos.|c_{24}) \times D_{k-}^{24} + p_k(neg.|c_{24}) \times D_{k-}^{24}. \tag{4.5}$$

Figure 4.5 illustrates the distributions $D_{k-}^{24}$ and $D_{k+}^{24}$ for $k = 2$. The right-hand panel is a mixed distribution of $D_{k-}^{24}$ and $D_{k+}^{24}$.

The left-hand side of the figure illustrates two possible distributions of cumulative future price changes, following either a positive or negative trend. Price changes since the beginning of the experiment allow us to determine the probability that the true trend is positive or negative. Applying these probability weigths produces a mixed distribution of possible cumulative price changes, which then can be used to determine desired confidence intervals by assessing the area under the distribution curve. The complexity of this method prevents participants from computing optimal confidence intervals during the experiment.

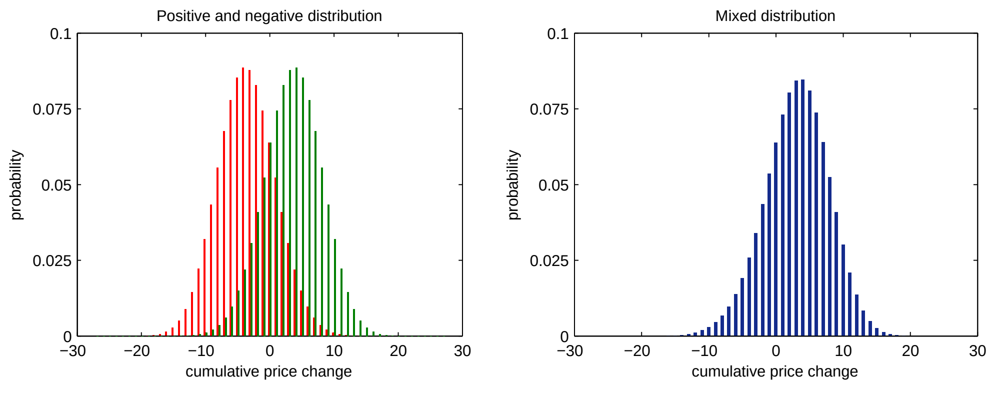



**Figure 4.5: Distribution of possible cumulative price changes**
The left-hand side of this figure illustrates the probability distributions of a stock that either follows a negative (red) or positive (green) long-term path. The cumulative price change since the beginning of the experiment allows assessment of the probability that the path is positive or negative. We can use this piece of information to determine a mixed distribution (right-hand side of the figure), which can be used to compute upper and lower bounds of desired confidence intervals through the area under the mixed distribution curve.

Glaser et al. (2013) compute individual overconfidence levels by taking an individual's provided upper and lower bound in a price prediction round to calculate the area under the respective mixed distribution curve. The computed area can then be subtracted from the given confidence level, which yields an individual's overconfidence level for a prediction round.

To implement the original methodology of Glaser et al. (2013), we take each participant's produced lower bound $l$ and upper bound $u$ for the respective distribution $D$ in round $t$ which we have previously computed in equation 4.5. As these distributions are not expressed as functions, we use the trapezoidal method to calculate the captured area under the curve $A$ for each participant $i$. Therefore,

$$A_{i,t} = \int_l^u f(x)dx = \frac{(l-u)}{2N}\left[f(x_1) + 2f(x_2) + ... + 2f(x_N) + f(x_{N+1})\right]. \tag{4.6}$$

Eventually, we can calculate a respondent's overconfidence score $OC_{area}$ following the area under the curve method for participant $i$ in round $t$ as

$$OC_{area,i,t} = C - A_{i,t}. \tag{4.7}$$

We can then compute a participant's mean overconfidence score over all rounds $n$ based on the area under the curve method as

$$TOC_{area,i} = \frac{\sum_{i=1}^{n} OC_{area,i,t}}{n}. \tag{4.8}$$

Following the recommendations of Langnickel and Zeisberger (2016), that individuals tend to be insensitive to different confidence levels, we also calculate overconfidence scores based on self-reported confidence levels $CSR$, which represents the self-assessed proportion of correctly defined confidence levels for each participant. Therefore,

$$OCSR_{area,i,t} = CSR - A_{i,t} \tag{4.9}$$

and

$$TOCSR_{area,i} = \frac{\sum_{i=1}^{n} OCSR_{area,i,t}}{n}. \tag{4.10}$$

However, the method has two potential shortcomings. First, the highest possible *overconfidence* score is equal to the stated confidence level $C$. If, however, a respondent produces extremely wide confidence intervals, the highest possible level of $underconfidence$ (which is equivalent to the $lowest$ possible level of overconfidence) would be $1 - C$. As a result, if we select a confidence levels of *i*) 90% and *ii*) 60%, the highest level of *underconfidence* in *i*) can be 10%, compared to 40% in scenatio *ii*). As a consequence, respondents who produce extremely wide confidence intervals will be diagnosed with a $higher$ level of underconfidence in the given confidence level is $lower$.

Second, if a provided confidence interval 'misses' the mixed distribution entirely, a participant's overconfidence would be close to $C$. Figure 4.6 illustrates this condition. A participant producing confidence interval A will appear highly overconfident, as the defined interval only captures a small proportion of the area under the curve, despite the high width of the interval. The respondent of confidence interval B, however, would be considered far *less* overconfident, despite the narrow interval produced.

To address our concerns about the possibility of labelling a participant that produced extremely wide confidence intervals that missed the optimal distribution entirely as extremely overconfident, we propose an alternative methodology that uses the widths of optimal confidence intervals $CII_t$ instead of the respective area under the curve. By doing that, we create an alternative measure robust against the bias mentioned above.

We begin with the mixed distributions summarised in equation 4.5, where

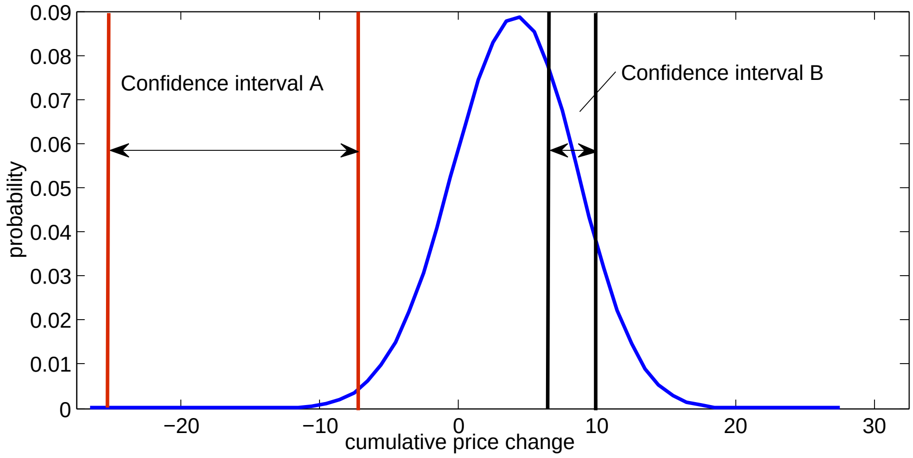



**Figure 4.6: Stated confidence intervals and resulting overconfidence scores**
This figure illustrates how a very *broad* confidence interval (confidence interval A) can result in a *high* overconfidence score, using the area under the curve method proposed by Glaser et al. (2013). According to the method, a participant who produced confidence interval B will be *less* overconfident than participant A.

we can compute the area under the curve given a stated confidence level (in our case 90%). The corresponding values on the x-axis represent price changes in $t + 9$ rounds, and the distance between these two values define the widths of ideal confidence intervals $CII_t$ for each round. These widths are the basis for computing overconfidence scores $OC_{width}$ for a participant $i$ at time $t$. $OC_{width,i,t}$ at time $t$ is computed by calculating the difference between the widths of the ideal confidence interval $CII_t$ and the width of the participant's provided confidence interval $WCI_t$. Therefore,

$$OC_{width,i,t} = CII_t - WCI_{i,t}. \tag{4.11}$$

In order to compare overconfidence scores from the stock price prediction task with $OCCI$ and $OCSR$ scores, we calculate a participant's mean overconfidence $TOC_i$ for this task as the sum of his overconfidence scores for each round:

$$TOC_{width,i} = \frac{\sum_{i=1}^{n} OC_{width,i,t}}{n}. \tag{4.12}$$

### 4.4.3 Participants and incentives

We conducted the online experiment between August 2016 and September 2016 with a total of 123 participants. Of these, 33 participants were from the finance industry, and 90 were undergraduate students from Macquarie University's Faculty of Business and Economics. Depending on their answers, participants could win up to four of a total 20 movie tickets, with an approximate value of \$18 per ticket, totalling up to \$72. Half of these tickets were randomly drafted among all participants who initially decided to 'short' the stock (the random component of the incentive depending on the share performance), while the remainder of the tickets were awarded to those participants who stated confidence intervals closest to the ideal values based on the methodology of Glaser et al. (2013).

Invitations for the survey were announced in several Finance and Accounting undergraduate courses, and a link was provided in an online announcement forum accessible to all enrolled students.

Following the recommendations of Langnickel and Zeisberger (2016), we only included complete sets or responses that took between five and 60 minutes for the survey to be completed. This criterion was added to ensure thorough responses without random guessing. We also excluded excessively long responses, as breaks may affect participants' memory and, thus, the effect of feedback on prior choices on their cognitive systems.

The average response time was 21.51 minutes. Of the participants, 42.3% were female, while 30.09% were older than 25 years, resulting in a relatively young group of participants.

## 4.5 Results

### 4.5.1 General knowledge questions

Similarly to Langnickel and Zeisberger (2016), we find that the proportion of appropriately stated confidence intervals for general knowledge questions are relatively low (31.5%), and Cronbach's alpha is moderately high (0.63)[4], suggesting a relatively reliable measure and high general overconfidence among participants.

Table 4.1 summarises the descriptive statistics of Stage 1 of the experiment. We find that overconfidence scores are generally high when using the traditional methodology of overconfidence assessment[5] ($OCCI$). As participants tend to be insensitive to the requested confidence levels (Langnickel and Zeisberger, 2016), we added a self-reported level of confidence ($OCSR$), by asking participants their own estimates of the number of correctly defined confidence intervals in this stage. On average, the student participants estimated that they answered 44.00% of these questions correctly. Participants who were finance professionals tended to be more confident, and reported, on average, that 52.27% of their confidence intervals captured the correct value.

Mean $OCCI$ scores are significantly higher among student participants (p = 0.01), but the difference is statistically insignificant for $OCSR$ scores. It must be noted that mean overconfidence scores based on self-reported confidence levels are lower than scores determined through the traditional methodology across students (1.64 and 6.24, respectively) and finance professionals (1.33 and 5.06, respectively), which is in line with the findings of Langnickel and Zeisberger (2016).

[4] Langnickel and Zeisberger (2016) report 35.3% appropriately stated confidence intervals and Cronbach's alpha of 0.67, respectively.

[5] For example, Biais et al. (2005) and Hilton et al. (2011).

**Table 4.1: Overconfidence scores among students and finance professionals**

This table summarises the descriptive statistics of general knowledge overconfidence scores among students and finance professionals using two alternative assessment methods. $OCCI$ refers to the conventional method, where the overconfidence score is computed by subtracting the number of correctly defined confidence intervals from the number based on confidence level specifications (i.e. 9). $OCSR$ refers to an alternative method which subtracts number of correctly stated confidence intervals from the self-assessed the number of correct intervals. Positive scores refer to overconfidence, and negative scores to underconfidence.

| Occupation | Score | N | Min. | Max. | Mean | Std. Dev. |
|---|---|---|---|---|---|---|
| Student | $OCCI$ | 90 | 2 | 9 | 6.24 | 2.01 |
| | $OCSR$ | 90 | -2 | 7 | 1.64 | 2.07 |
| Professionals | $OCCI$ | 33 | -1 | 9 | 5.06 | 2.55 |
| | $OCSR$ | 33 | -5 | 7 | 1.33 | 2.41 |

### 4.5.2 Stock price predictions

In Stage 2 of the experiment, participants were asked to make stock price predictions after observing the artificial chart of a stock. As summarised in Table 4.2, overconfidence scores are partially consistent across assessment tasks. Based on $OCCI$, participants who were overconfident in their answers to general knowledge questions also tended to be overly optimistic about the accuracy of their stock price predictions for both, $TOC_{area}$ and $TOC_{width}$ (Pearson's $\rho = 0.26, p < 0.01$). However, this was not the case for overconfidence scores using self-reported confidence levels as a base. While we find a relatively high correlation between $OCCI$ and $OCSR$ scores (Pearson's $\rho = 0.62, p < 0.01$), our tests suggest no association between $OCSR$ and the three $TOC$ measures.

The purpose of initially assessing overconfidence in Stage 1 of the experiment was to identify how participants enter Stage 2 of the experiment. It is reasonable to assume that participants do not begin such an experiment with zero overconfidence. Indeed, we found that participants in both samples tended to be overconfident at the beginning of the price prediction stage.

Table 4.3 summarises mean overconfidence scores $OC_{width}$ and $OC_{area}$ over 12 rounds of the experiment, split among participants who initially decided to

**Table 4.2: Correlation between overconfidence scores: General knowledge confidence interval method, self-assessment method and artificial chart method.**

This table presents Pearson correlation coefficients between assessed overconfidence scores over the two stages of the experiment. $TOC$ is the overconfidence score of a participant answering after 12 prediction rounds, the subscripts differentiate between the two methods (area under the curve and interval width). $OCCI$ is the overconfidence score of 10 general knowledge questions using the conventional method, and $OCSR$ and $TOCSR$ are overconfidence scores of a participant using self-reported levels of confidence.

| | $OCCI$ | $OCSR$ | $TOC_{area}$ | $TOCSR_{area}$ | $TOC_{width}$ |
|---|---|---|---|---|---|
| $OCCI$ | 1.00 | .62* | .27* | 0.10 | .26* |
| $OCSR$ | .62* | 1.00 | 0.05 | 0.15 | 0.01 |
| $TOC_{area}$ | .26* | 0.05 | 1.00 | .70* | .88* |
| $TOCSR_{area}$ | 0.10 | 0.15 | .70* | 1.00 | .58* |
| $TOC_{width}$ | .26* | 0.01 | .88* | .58* | 1.00 |

*. Correlation is significant at the 0.01 level (2-tailed).

long or short the stock. Column 2 indicates the price of the stock in a round, and column 3, 6, 9 ans 12 show the respective gain or loss with respect to round 1. It must be noted that gains.

Following $OC_{width}$, participants appear to show high levels of overconfidence in the first rounds which then quickly decreases after round two. As the stock price increases and reaches its pinnacle at round 5, overconfidence scores become negative among participants in short positions, while those participants in long positions are, on average, still overconfident.

As the trend reverses and the market crashes, mean $OC$ scores among participants in long positions become low, while mean scores for those in short positions recover.

The pattern is particularly pronounced for our student sub-sample and, to a lesser extent, among professionals in long positions where we fail to observe a clear pattern.

To test our hypotheses formally, we apply a series of non-parametric tests, due to the non-normality of our data. The results of these tests are presented in the following section.

**Table 4.3: Mean overconfidence scores over 12 price prediction rounds of the experiment**

This table reports the mean overconfidence scores over 12 rounds of the experiment using both methods of individual overconfidence assessment ($TOC_{width}$ and $TOC_{area}$). Column 2 shows the price of the stock in the respective round, and column 3, 6, 9 and 12 the respective gain or loss with respect to the beginning of the experiment.

| | | $OC_{width}$ | | | | | | $OC_{area}$ | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Long | | | Short | | | Long | | | Short | | |
| Round | Price | Gain/loss | Mean | $t$ | Gain/loss | Mean | $t$ | Gain/loss | Mean | $t$ | Gain/loss | Mean | $t$ |
| 1 | 50 | 0 | 5.93 | 6.85*** | 0 | 3.89 | 2.26*** | 0 | 0.39 | 14.01*** | 0 | 0.42 | 8.54*** |
| 2 | 52 | 2 | 5.84 | 6.67*** | -2 | 5.63 | 4.64*** | 2 | 0.40 | 13.73*** | -2 | 0.37 | 7.36*** |
| 3 | 57 | 7 | 1.81 | 1.95* | -7 | 2.57 | 2.06** | 7 | 0.35 | 10.34*** | -7 | 0.45 | 7.93*** |
| 4 | 60 | 10 | 1.88 | 2.07** | -10 | 2.37 | 1.81* | 10 | 0.35 | 10.77*** | -10 | 0.43 | 8.08*** |
| 5 | 68 | 18 | 0.66 | 0.65 | -18 | -0.37 | -0.19 | 18 | 0.63 | 18.16*** | -18 | 0.63 | 11.05*** |
| 6 | 66 | 16 | 1.49 | 1.59 | -16 | 0.37 | 0.24 | 16 | 0.64 | 18.82*** | -16 | 0.58 | 10.27*** |
| 7 | 60 | 10 | 2.08 | 2.19** | -10 | 1.86 | 1.34 | 10 | 0.48 | 15.22*** | -10 | 0.43 | 8.36*** |
| 8 | 55 | 5 | 1.03 | 1.02 | -5 | 2.83 | 2.16** | 5 | 0.34 | 9.70*** | -5 | 0.37 | 6.64*** |
| 9 | 49 | -1 | 0.63 | 0.61 | 1 | 2.31 | 1.49 | -1 | 0.49 | 12.96*** | 1 | 0.60 | 11.37*** |
| 10 | 40 | -10 | 2.17 | 1.90 | 10 | 4.11 | 2.36** | -10 | 0.55 | 16.08*** | 10 | 0.65 | 13.10*** |
| 11 | 37 | -13 | 0.41 | 0.39 | 13 | 0.66 | 0.38 | -13 | 0.45 | 13.41*** | 13 | 0.55 | 10.06*** |
| 12 | 42 | -8 | 1.34 | 1.15 | 8 | 1.54 | 0.87 | -8 | 0.40 | 12.68*** | 8 | 0.44 | 7.83*** |

***, ** and * Statistically significant at a 99%, 95% and 90% confidence level, respectively, (two-tailed).

### Tests of hypotheses

Hypothesis 1a postulates that investors who hold long positions become overconfident in market booms. The rationale is simple: participants who initially decide to long the stock should become overconfident over the first seven rounds of the experiment (the rounds that mimic a market boom). Columns 5 and 8 of table 4.3 report a series of one-sample t-tests with the null that mean $OC_{width}$ overconfidence scores for each round is equal to zero. In other words, a significant t-score suggests that the mean overconfidence score for that round is significantly different from zero.

The results are in line with Hypothesis 1a. Overconfidence is present throughout all 'boom' rounds for those participants that initially decided to long the stock, with one exception at round 5.

Hypothesis 1b claims that investors who hold short positions lose their overconfidence in market booms. Quite strikingly, while starting the experiment overconfident, overconfidence among participants who decided to short the stock disappears in round 5, the first boom phase of the experiment. This notion supports Hypothesis 1b. Indeed, strongly contrary evidence appears to cause overconfidence to vanish.

Hypothesis 2a postulates that stock market crashes cause overconfident investors who hold *long* positions to lose their overconfidence. The simulated stock market crash reaches its pinnacle in round 9 of the experiment. In line with the hypothesis, mean overconfidence among respondents who initially decided to long the stock vanishes.

Hypothesis 2b mirrors Hypothesis 2a; that is, stock market crashes cause investors who hold *short* positions to become overconfident. Quite strikingly, participants of our sample who initially decided to short the stock become overconfident, on average, during the simulated stock market crash in rounds 8–10, as illustrated in Table 4.3.

The combination of Hypotheses 1 and 2 provides interesting insights about individuals' confidence calibration process after the arrival of vivid feedback. Initially overconfident investors lose (maintain) their overconfidence when receiving feedback that strongly contradicts (supports) their prior decision making. However, if those feedback signals revert and strongly support (contradict) those prior decisions, overconfidence re-appears (disappears).

Hypotheses tests using $OC_{area}$, the measure originally proposed by Glaser et al. (2013), reveals interesting patterns. Interestingly, both groups are highly overconfident throughout all 12 rounds of the experiment. Furthermore, both groups (long/short) become particularly overconfident when markets are either very bullish or bearish. This finding is difficult to align with the intuition that on average, overconfidence increases after gains and decreases after losses (Odean, 1998). One possible explanation for this phenomenon could be the misattribution of those participants that produce very wide confidence intervals that miss the actual distribution curve due to miscalibration entirely, as illustrated in Figure 4.6 above.

#### Price stimuli and investor disagreement

A secondary aim of this study addresses how individuals form heterogeneous beliefs after the arrival of new information. Hong and Stein (2007) and Scheinkman and Xiong (2003) propose that if investors interpret information signals differently, a greater news impulse should be associated with higher time-series variance in confidence interval widths due to the heterogeneous valuation of investors.

Figure 4.1 illustrates the variance of stated confidence intervals over 12 rounds of the experiment. Quite strikingly, variance spikes when the experimental asset market crashes[6], which is in line with the hypothesis by Hong and Stein (2007) and Scheinkman and Xiong (2003).

[6] F-tests of differences in variance in widths of price prediction intervals between round 1 and round 5 (the boom) and round 1 and round 10 (the crash) yield significant results with $F_{long,t_2,t_5} = 1.35$, $p = 0.08$, $F_{long,t_2,t_10} = 1.72$, $p < 0.01$, $F_{short,t_2,t_5} = 2.51$, $p < 0.01$ and $F_{short,t_2,t_10} = 2.07$, $p = 0.02$, respectively.

The effect is particularly strong among investors who hold long positions (dashed green line). One possible explanation for this pattern could be an interaction between self-attribution bias and the strength of arriving signals. When strong news impulses arrive, investors interpret them differently, particularly, when those impulses conflict with their prior beliefs.

**Methodological concerns**

Recent studies have demonstrated some shortcomings of traditional methodological approaches to assessing overconfidence. For instance, Langnickel and Zeisberger (2016) and Biais et al. (2005) show that individuals tend to be insensitive to requested confidence levels when producing confidence intervals in assessment tasks (e.g. "What is the height of the Eiffel Tower (in m)?"). Our results support this view. In order to test the extent to which participants take stated confidence levels into consideration when producing confidence intervals, we added self-assessed confidence levels (i.e. "How many true values of the first 10 general knowledge questions on the previous page do you expect to lie within your provided ranges?" and "For the 12 share price ranges that you provided (upper bound and lower bound) on the previous pages, how often do you think that the actual share price was within the range?", respectively), which follows the rationale of Langnickel and Zeisberger (2016). As the confidence level was given at 90% and participants, in fact, constructed 90% confidence intervals, the answer should be on average 9 (out of 10) for general knowledge questions, and on average 10.8 (out of 12) for price prediction questions, respectively.

On average, self-assessed confidence levels were 46.42% for general knowledge questions and 49.32% for the price prediction task, which is consistently lower than the stated 90% confidence levels.

Test results reported in Table 4.4 suggest that self-reported confidence levels are significantly different from the stated ones. As a result, overconfidence

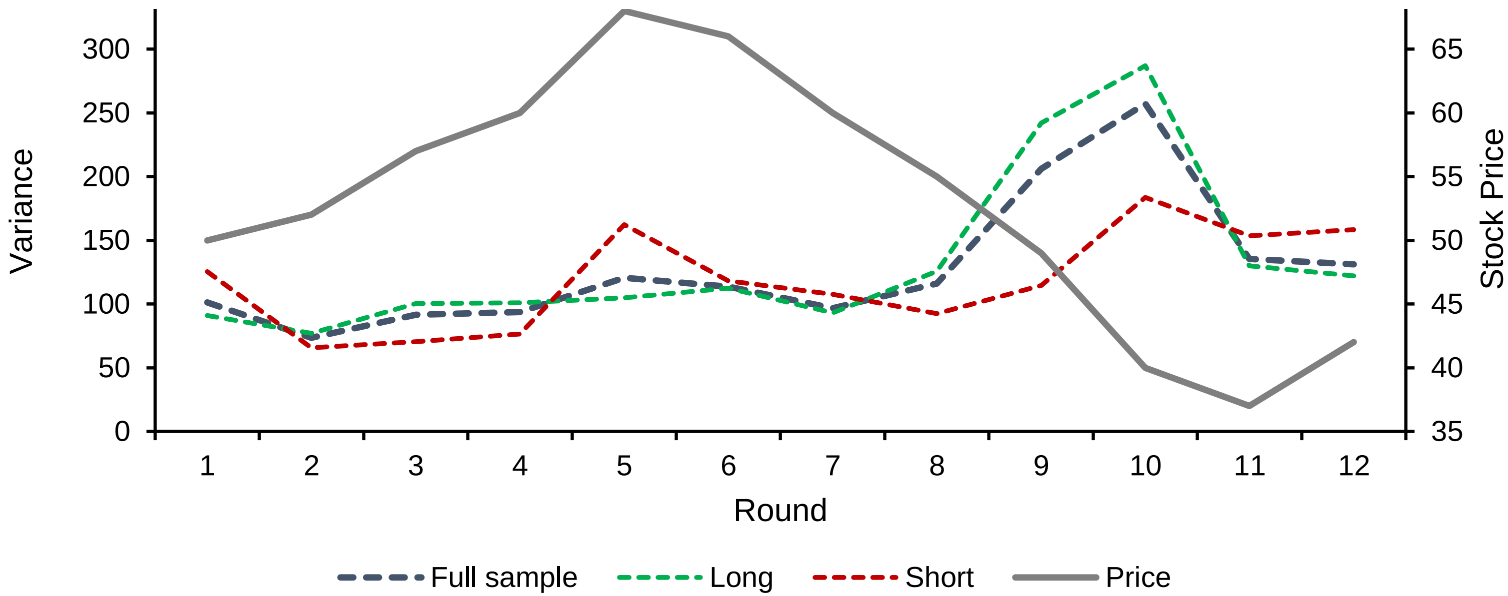



**Figure 4.1: Variance of confidence interval widths scores over 12 stock price prediction rounds**
This chart illustrates the variance of the width of the provided confidence intervals/overconfidence scores over 12 stock price prediction rounds among participants in long and short positions.

**Table 4.4: Self-reported confidence levels versus stated confidence levels**

This table reports a series of one-sample t-tests of mean self-reported versus confidence levels stated during two stages of assessment tasks among student and professional participants. Appropriately calibrated participants should report 9 correctly answered general knowledge questions and, on average, 10.8 correct stock price predictions.

| Measure | Sample | N | Mean | $t$ |
|---|---|---|---|---|
| General knowledge | Students | 90 | 44.00% | -24.36*** |
| | Professionals | 33 | 52.73% | -10.37*** |
| Price predictions | Students | 90 | 48.70% | -18.76*** |
| | Professionals | 33 | 51.01% | -10.22*** |

*** statistically significant at a 99% confidence level (two-tailed).

scores are systematically inflated for general knowledge questions which is in line with the findings of Langnickel and Zeisberger (2016). We find strikingly similar patterns across both participant groups and assessment tasks.

Participants appear to have difficulties incorporating a requested confidence level into their interval predictions. This challenges the methodology commonly used for overconfidence assessment, including the 'traditional' methodology for general knowledge questions (i.e. Russo and Schoemaker, 1992), but also the methodology recently proposed by Glaser et al. (2013) to assess overconfidence in stock price prediction tasks.

## 4.6 Conclusion and recommendations

The traditional assessment of overconfidence uses a number of knowledge questions, paired with a given confidence level. However, this methodology has two shortcomings. First, as demonstrated by Langnickel and Zeisberger (2016) and Biais et al. (2005), individuals typically tend to be insensitive to the requested confidence levels, with this commonly resulting in inflated levels of overconfidence. Second, repetitions of the questions after feedback are impossible. Once participants know the true value of a knowledge question, confidence intervals are no longer required. Glaser et al. (2013) propose a methodology to compute "true" overconfidence by using complex distribution pairs that can be applied for stock price predictions. The new methodology allows the assessment of changes in overconfidence after feedback as stock price predictions can be measured over several repetitions.

Utilising this new method, we test the calibration process of overconfident individuals after the arrival of strongly supporting or contradicting evidence to prior decisions. We find that overconfidence tends to disappear when arriving feedback strongly contradicts their prior decision making ('confidence crashes'). On the other hand, we find that subsequent to those confidence crashes, overconfidence can re-emerge when feedback signals revert and prior decisions are reinforced.

Our findings potentially contribute to the literature by providing a possible explanation of the re-calibration process of overconfidence. The assumption that investors suffer from both overconfidence and self-attribution bias (e.g. Daniel et al., 1998), raises the following obvious point. If investors become more (over)confident after the arrival of feedback supporting prior decisions, but do not fully adjust their level of confidence after the arrival of evidence that contradicts their decision making due to self-attribution bias, these individuals must become exceedingly overconfident as time passes. Our evidence suggests that, as overconfidence vanishes when extremely contradicting information arrives, there is indeed a process that restores mean confidence calibration during the said events.

Furthermore, we find supporting evidence for the hypotheses by Hong and Stein (2007) and Scheinkman and Xiong (2003) that time-series variance in security valuation increases when strong news impulses arrive. In line with these hypotheses, we find a stark increase of time-series variance in the provided confidence intervals following market crashes. This supports the notion that agents interpret these impulses differently.

We also find that the methodology proposed by Glaser et al. (2013) bears the same flaw as overconfidence assessment using stated confidence levels, as demonstrated by Langnickel and Zeisberger (2016) and Biais et al. (2005). As participants are not sensitive to confidence levels[7], overconfidence scores assessed using the 'classic' methodology are likely to be inflated.

Given the persistent importance of overconfidence in the field, as recently reviewed by Daniel and Hirshleifer (2015), a methodology insensitive to confidence level biases is an apparent gap in the literature. We are confident that the near future will yield promising approaches to fill this gap.

[7] We find that, on average, self-assessed confidence levels are significantly lower, suggesting that dictated confidence levels are mostly not incorporated into the cognitive process of producing confidence intervals.

# Appendix

## 4.A Robustness checks

### 4.A.1 Non-parametric correlations among overconfidence measures

**Table 4.A.1: Correlation between overconfidence scores: General knowledge confidence interval method, self-assessment method and artificial chart method.**

This table presents Spearman's correlation coefficients between assessed overconfidence scores over the two stages of the experiment. $TOC$ is the overconfidence score of a participant answering after 12 prediction rounds, the subscripts differentiate between the two methods (area under the curve and interval width). $OCCI$ is the overconfidence score of 10 general knowledge questions using the conventional method, and $OCSR$ and $TOCSR$ are overconfidence scores of a participant using self-reported levels of confidence.

| | $OCCI$ | $OCSR$ | $TOC_{area}$ | $TOCSR_{area}$ | $TOC_{width}$ |
|---|---|---|---|---|---|
| $OCCI$ | 1.00 | 0.59* | 0.25* | 0.09 | 0.28* |
| $OCSR$ | 0.59* | 1.00 | 0.07 | 0.13 | 0.06 |
| $TOC_{area}$ | 0.25* | 0.07 | 1.00 | 0.70* | 0.85* |
| $TOCSR_{area}$ | 0.10 | 0.13 | 0.70* | 1.00 | 0.61* |
| $TOC_{width}$ | 0.28* | 0.06 | 0.85* | 0.61* | 1.00 |

*. Correlation is significant at the 0.01 level (2-tailed).

### 4.A.2 Non-parametric tests for median overconfidence scores over 12 rounds of the experiment

**Table 4.A.2: Non-parametric tests for median overconfidence scores over 12 rounds of the experiment**

This table presents tests outputs for non-parametric Wilcoxon-signed rank tests (WC) for median $TOC_{width}$ and $TOC_{area}$ overconfidence scores. Consistent with mean scores, median $TOC_{area}$ scores remain positive throughout the 12 rounds of the experiment, while median $TOC_{width}$ suggest mixed results. According to the tests output, participants who initially decided to long the stock remain significantly overconfidence, while those that initially decided to short the stock are overconfident only in the beginning of the experiment and during rounds 8 and 9, which is consistent with prior expectations.

| | | $TOC_{width}$ | | | | | | $TOC_{area}$ | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Long | | | Short | | | Long | | | Short | | |
| Round | Price | Gain/ loss | Median | WC | Gain/ loss | Median | WC | Gain/ loss | Median | WC | Gain/ loss | Median | WC |
| 1 | 50 | 0 | 7 | reject | 0 | 7 | reject | 0 | 0.34 | reject | 0 | 0.54 | reject |
| 2 | 52 | 2 | 7 | reject | -2 | 7 | reject | 2 | 0.39 | reject | -2 | 0.37 | reject |
| 3 | 57 | 7 | 4 | reject | -7 | 4 | reject | 7 | 0.35 | reject | -7 | 0.47 | reject |
| 4 | 60 | 10 | 4 | reject | -10 | 4 | | 10 | 0.37 | reject | -10 | 0.45 | reject |
| 5 | 68 | 18 | 4 | reject | -18 | 4 | | 18 | 0.79 | reject | -18 | 0.80 | reject |
| 6 | 66 | 16 | 4 | reject | -16 | 4 | | 16 | 0.79 | reject | -16 | 0.74 | reject |
| 7 | 60 | 10 | 4 | reject | -10 | 4 | | 10 | 0.52 | reject | -10 | 0.45 | reject |
| 8 | 55 | 5 | 4 | reject | -5 | 4 | reject | 5 | 0.33 | reject | -5 | 0.37 | reject |
| 9 | 49 | -1 | 4 | reject | 1 | 5 | reject | -1 | 0.56 | reject | 1 | 0.74 | reject |
| 10 | 40 | -10 | 6.5 | reject | 10 | 4 | | -10 | 0.70 | reject | 10 | 0.75 | reject |
| 11 | 37 | -13 | 3.5 | reject | 13 | 5 | | -13 | 0.52 | reject | 13 | 0.61 | reject |
| 12 | 42 | -8 | 5 | reject | 8 | 5 | | -8 | 0.43 | reject | 8 | 0.52 | reject |

**Table 4.A.3: Descriptive statistics of overconfidence scores during Stage 2 of the experiment ($OC_{width}$ long)**

This table reports summary statistics of overconfidence scores of all participants that have initially decided to long the stock over 12 rounds of Stage 2 of the experiment following the $OC_{width}$ method.

| Round | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Mean | | 5.93 | 5.84 | 1.81 | 1.88 | 0.66 | 1.49 | 2.08 | 1.03 | 0.63 | 2.17 | 0.41 | 1.34 |
| 95% CI for Mean | Lower B. | 4.21 | 4.10 | -0.03 | 0.08 | -1.36 | -0.37 | 0.19 | -0.98 | -1.42 | -0.11 | -1.65 | -0.98 |
| | Upper B. | 7.65 | 7.57 | 3.65 | 3.67 | 2.68 | 3.34 | 3.97 | 3.05 | 2.67 | 4.45 | 2.47 | 3.66 |
| 5% Trimmed Mean | | 6.96 | 6.84 | 2.64 | 2.68 | 1.49 | 2.32 | 3.04 | 1.89 | 1.57 | 2.91 | 1.31 | 2.55 |
| Median | | 7.00 | 7.00 | 4.00 | 4.00 | 4.00 | 4.00 | 4.00 | 4.00 | 4.00 | 6.50 | 3.50 | 5.00 |
| Variance | | 66.02 | 67.26 | 75.28 | 72.04 | 90.69 | 76.71 | 79.20 | 90.29 | 93.46 | 115.38 | 94.66 | 119.97 |
| Std. Deviation | | 8.13 | 8.20 | 8.68 | 8.49 | 9.52 | 8.76 | 8.90 | 9.50 | 9.67 | 10.74 | 9.73 | 10.95 |
| Minimum | | -33.00 | -31.00 | -31.00 | -31.00 | -31.00 | -31.00 | -31.00 | -31.00 | -31.00 | -28.00 | -29.00 | -45.00 |
| Maximum | | 14.00 | 15.00 | 13.00 | 13.00 | 13.00 | 12.00 | 14.00 | 14.00 | 12.00 | 16.00 | 12.00 | 13.00 |
| Range | | 47.00 | 46.00 | 44.00 | 44.00 | 44.00 | 43.00 | 45.00 | 45.00 | 43.00 | 44.00 | 41.00 | 58.00 |
| Interquartile Range | | 9.00 | 8.00 | 10.00 | 10.00 | 14.75 | 9.00 | 10.00 | 12.00 | 13.50 | 15.00 | 9.00 | 12.25 |
| Skewness | | -2.47 | -2.21 | -1.58 | -1.57 | -1.32 | -1.64 | -1.71 | -1.33 | -1.40 | -0.98 | -1.45 | -1.85 |
| Kurtosis | | 8.22 | 5.92 | 2.61 | 2.99 | 1.59 | 3.04 | 3.22 | 1.68 | 1.84 | 0.20 | 1.57 | 3.93 |

**Table 4.A.4: Descriptive statistics of overconfidence scores during Stage 2 of the experiment ($OC_{width}$ short)**

This table reports summary statistics of overconfidence scores of all participants that have initially decided to short the stock over 12 rounds of Stage 2 of the experiment following the $OC_{width}$ method.

| Round | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Mean | | 3.89 | 5.63 | 2.57 | 2.37 | -0.37 | 0.37 | 1.86 | 2.83 | 2.31 | 4.11 | 0.66 | 1.54 |
| 95% CI for Mean | Lower B. | 0.39 | 3.16 | 0.04 | -0.29 | -4.27 | -2.80 | -0.95 | 0.17 | -0.84 | 0.57 | -2.86 | -2.08 |
| | Upper B. | 7.39 | 8.09 | 5.11 | 5.03 | 3.53 | 3.54 | 4.67 | 5.49 | 5.47 | 7.66 | 4.17 | 5.17 |
| 5% Trimmed Mean | | 5.04 | 6.12 | 3.01 | 2.76 | 0.37 | 1.08 | 2.44 | 3.52 | 3.36 | 5.03 | 1.62 | 2.48 |
| Median | | 7.00 | 7.00 | 4.00 | 4.00 | 4.00 | 4.00 | 4.00 | 4.00 | 5.00 | 7.00 | 4.00 | 5.00 |
| Variance | | 103.81 | 51.48 | 54.43 | 59.83 | 129.12 | 85.01 | 66.83 | 60.03 | 84.22 | 106.57 | 104.64 | 111.20 |
| Std. Deviation | | 10.19 | 7.17 | 7.38 | 7.73 | 11.36 | 9.22 | 8.18 | 7.75 | 9.18 | 10.32 | 10.23 | 10.54 |
| Minimum | | -33.00 | -13.00 | -16.00 | -16.00 | -26.00 | -30.00 | -21.00 | -21.00 | -26.00 | -32.00 | -35.00 | -34.00 |
| Maximum | | 14.00 | 15.00 | 12.00 | 12.00 | 12.00 | 12.00 | 12.00 | 12.00 | 11.00 | 16.00 | 13.00 | 12.00 |
| Range | | 47.00 | 28.00 | 28.00 | 28.00 | 38.00 | 42.00 | 33.00 | 33.00 | 37.00 | 48.00 | 48.00 | 46.00 |
| Interquartile Range | | 15.00 | 12.00 | 13.00 | 15.00 | 15.00 | 13.00 | 13.00 | 9.00 | 10.00 | 15.00 | 12.00 | 15.00 |
| Skewness | | -1.85 | -0.95 | -0.83 | -0.89 | -1.20 | -1.25 | -1.16 | -1.35 | -1.72 | -1.64 | -1.62 | -1.46 |
| Kurtosis | | 4.21 | 0.22 | -0.15 | -0.31 | 0.27 | 1.97 | 1.06 | 1.76 | 2.84 | 3.14 | 3.23 | 2.30 |

**Table 4.A.5: Descriptive statistics of overconfidence scores during Stage 2 of the experiment ($OC_{area}$ long)**

This table reports summary statistics of overconfidence scores of all participants that have initially decided to long the stock over 12 rounds of Stage 2 of the experiment following the $OC_{area}$ method.

| Round | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Mean | | 0.39 | 0.40 | 0.35 | 0.35 | 0.63 | 0.64 | 0.48 | 0.34 | 0.49 | 0.55 | 0.45 | 0.40 |
| 95% CI for Mean | Lower B. | 0.34 | 0.34 | 0.28 | 0.29 | 0.56 | 0.57 | 0.42 | 0.27 | 0.41 | 0.48 | 0.39 | 0.34 |
| | Upper B. | 0.45 | 0.46 | 0.42 | 0.42 | 0.70 | 0.71 | 0.54 | 0.41 | 0.56 | 0.62 | 0.52 | 0.46 |
| 5% Trimmed Mean | | 0.39 | 0.40 | 0.35 | 0.35 | 0.66 | 0.67 | 0.49 | 0.33 | 0.50 | 0.57 | 0.46 | 0.40 |
| Median | | 0.34 | 0.39 | 0.35 | 0.37 | 0.79 | 0.79 | 0.52 | 0.33 | 0.56 | 0.70 | 0.52 | 0.43 |
| Variance | | 0.07 | 0.07 | 0.10 | 0.10 | 0.11 | 0.10 | 0.09 | 0.11 | 0.12 | 0.10 | 0.10 | 0.09 |
| Std. Deviation | | 0.26 | 0.27 | 0.32 | 0.31 | 0.33 | 0.32 | 0.30 | 0.33 | 0.35 | 0.32 | 0.32 | 0.30 |
| Minimum | | -0.10 | -0.10 | -0.10 | -0.10 | -0.10 | -0.10 | -0.10 | -0.10 | -0.10 | -0.10 | -0.10 | -0.10 |
| Maximum | | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 |
| Range | | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Interquartile Range | | 0.36 | 0.40 | 0.54 | 0.50 | 0.40 | 0.43 | 0.45 | 0.58 | 0.67 | 0.55 | 0.53 | 0.46 |
| Skewness | | -0.10 | -0.09 | 0.21 | 0.24 | -1.19 | -1.19 | -0.53 | 0.11 | -0.40 | -0.65 | -0.30 | -0.16 |
| Kurtosis | | -0.77 | -0.87 | -1.16 | -1.07 | 0.14 | 0.05 | -0.77 | -1.36 | -1.30 | -0.85 | -1.12 | -1.05 |

**Table 4.A.6: Descriptive statistics of overconfidence scores during Stage 2 of the experiment ($OC_{area}$ short)**

This table reports summary statistics of overconfidence scores of all participants that have initially decided to short the stock over 12 rounds of Stage 2 of the experiment following the $OC_{area}$ method.

| Round | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Mean | | 0.42 | 0.37 | 0.45 | 0.43 | 0.63 | 0.58 | 0.43 | 0.37 | 0.60 | 0.65 | 0.55 | 0.44 |
| 95% CI for Mean | Lower B. | 0.32 | 0.26 | 0.33 | 0.32 | 0.52 | 0.46 | 0.33 | 0.26 | 0.50 | 0.55 | 0.44 | 0.33 |
| | Upper B. | 0.51 | 0.47 | 0.56 | 0.54 | 0.75 | 0.69 | 0.53 | 0.48 | 0.71 | 0.75 | 0.66 | 0.56 |
| 5% Trimmed Mean | | 0.42 | 0.36 | 0.45 | 0.43 | 0.66 | 0.60 | 0.43 | 0.37 | 0.63 | 0.68 | 0.56 | 0.45 |
| Median | | 0.54 | 0.37 | 0.47 | 0.45 | 0.80 | 0.74 | 0.45 | 0.37 | 0.74 | 0.75 | 0.61 | 0.52 |
| Variance | | 0.08 | 0.09 | 0.11 | 0.10 | 0.11 | 0.11 | 0.09 | 0.11 | 0.10 | 0.09 | 0.10 | 0.11 |
| Std. Deviation | | 0.29 | 0.29 | 0.33 | 0.31 | 0.34 | 0.33 | 0.30 | 0.33 | 0.31 | 0.29 | 0.32 | 0.34 |
| Minimum | | -0.10 | -0.09 | -0.10 | -0.09 | -0.10 | -0.10 | -0.10 | -0.10 | -0.10 | -0.10 | -0.10 | -0.10 |
| Maximum | | 0.83 | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 |
| Range | | 0.93 | 0.99 | 1.00 | 0.99 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Interquartile Range | | 0.43 | 0.58 | 0.67 | 0.56 | 0.46 | 0.49 | 0.51 | 0.52 | 0.44 | 0.33 | 0.61 | 0.61 |
| Skewness | | -0.41 | -0.11 | -0.36 | -0.19 | -1.07 | -0.74 | -0.14 | 0.16 | -1.11 | -1.33 | -0.57 | -0.26 |
| Kurtosis | | -1.17 | -1.22 | -1.20 | -1.10 | -0.23 | -0.81 | -1.00 | -1.22 | 0.20 | 0.80 | -0.97 | -1.29 |

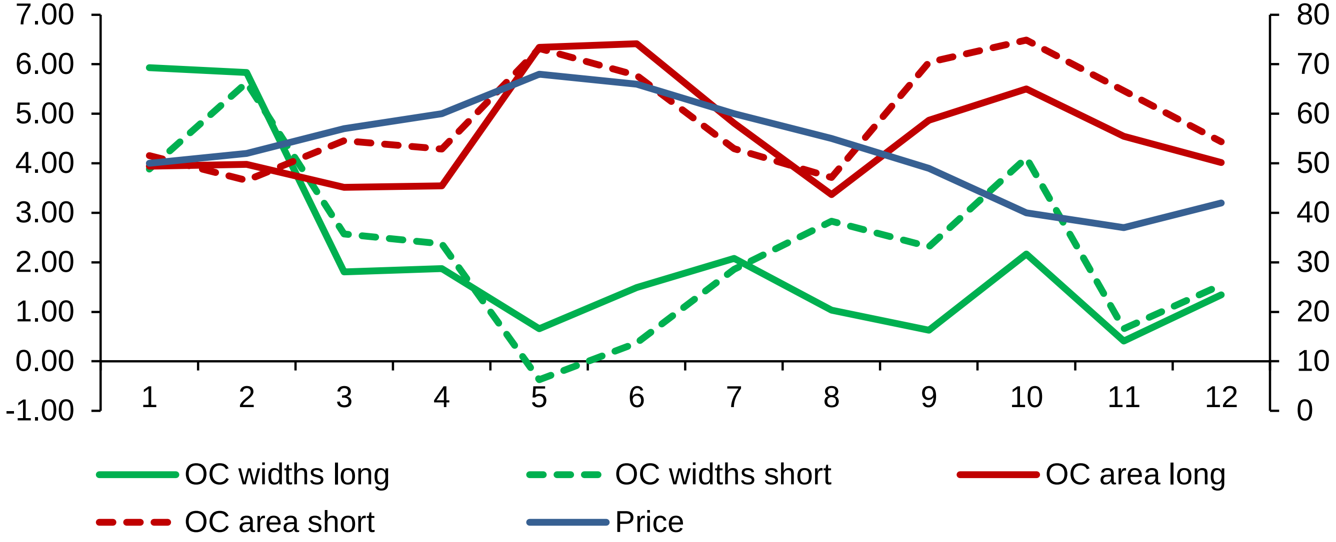



**Figure 4.A.1: Mean overconfidence scores over 12 stock price prediction rounds**

This chart illustrates mean overconfidence scores over 12 stock price prediction rounds among buyers/sellers of the stock, following two methodologies to compute overconfidence. $OC_{area}$, which is closer related to Glaser et al. (2013) and $OC_{width}$, which is an alternative method used in this paper.

# Chapter 5

# Conclusion

This research is devoted to making further contributions to the literature by extending our understanding of overconfidence. Overconfidence is considered one of the most robust findings in the psychology of human judgement (DeBondt and Thaler, 1994b), and, hence, is considered to offer a 'microfoundation' for a variety of models to explain phenomena in behavioural finance (Daniel and Hirshleifer, 2015).

In the second chapter, I develop a new measure of aggregate investor confidence, with the motivation being the lack of a measure that captures variations, in aggregate, in investor confidence from feedback signals with this new measure able to be applied in a wide range of domains.

The existing proxies of investor confidence typically capture individuals' confidence in their environment, such as their belief in the future development of the economy, their future purchase power or the likelihood that the stock market is properly valued, or in a bubble, or has a bullish or bearish outlook.

These perspectives are only difficult to align with theoretical models proposing behavioural explanations for stock market phenomena, such as excessive trading activity, or the size or price momentum effect. For instance, in the case of investor confidence and trading activity (e.g. Statman et al., 2006; Odean, 1998, 1999; Gervais and Odean, 2001; Cooper et al., 2004), investors are assumed to become more confident in their *own* abilities to trade stocks which is typically higher (lower) after UP (DOWN) markets, as most investors hold long positions. In other words, investors who have recently experienced portfolio gains are likely to credit their own abilities for the success and, consequently, are likely to trade more.

Chapter 2 develops a measure that is more closely related to the notion that feedback signals influence the level of investor (over)confidence, which is being conceptually and statistically distinct from other measures of investor confidence and sentiment used in the literature.

In a cross-disciplinary approach, I borrow a model of the formation of (over)confidence by Griffin and Tversky (1992). In the model, one's confidence depends on the strength and quality of evidence supporting or contradicting a belief, which Griffin and Tversky (1992) define as strength and weight of evidence. Strength is the magnitude of evidence, while weight characterises its reliability.

For instance, someone's belief about the skill of a football player depends on his observed skill during matches. If he scored many goals in the last few games, one may come to the conclusion that he is an excellent player. In this scenario, the magnitude of evidence is high. In other words, a high strength signal of evidence stimulates the observer's belief about the skill of the player.

However, his recent outstanding performance may either be due to his actual skill, or simply to luck. If the observer forms a belief about the player's level of skill which was, in fact, only due to luck, the observer becomes overconfident. In other words, he erroneously extrapolates observed results into the future.

This overconfidence, however, will not prevail with sufficient *weight* of evidence. Griffin and Tversky (1992) suggest sample size as a potential application of weight. If the observer now keeps watching matches, he will eventually adjust his belief about the player's skill to an appropriate level. The longer the player keeps performing well, the smaller the probability that his performance is due to luck. If, however, the player stops performing as well, the observer will soon adjust his belief.

Consequently, one's confidence about something increases when the strength of evidence is high. If, however, the weight of such evidence is low, the observer is likely to become overconfident as the strength of evidence is not justified with sufficient support (weight).

Using this model as a starting point, chapter 2 develops a simple measure of aggregate investor confidence that uses stock market data as feedback signals. The rationale is as follows. If the majority of investors hold long positions, and the stock market has recently performed well, investors will, on average, interpret this arriving information as feedback on prior investment decisions. For example, if recent portfolio performance was particularly strong, investors receive a high strength signal to boost their confidence. As argued in the previous example, high performance can either be due to skill or luck.

If investors tend to credit their own abilities in the case of success and blame externalities for failure (self-attribution bias), the weight should be low if the spread in returns is high.

In other words, aggregate investor confidence is high when recent stock market returns were high and returns are subject to self-attribution bias.

Chapter 2 demonstrates that the newly proposed measure aligns with major economic events, and is conceptually and statistically distinct from measures of investor confidence existing in prior literature.

The empirical findings are in line with expectations of higher trading activity subsequent to time periods of high aggregate investor confidence, as well as an increase in the trading activity of stocks associated with higher risk. The new measure of aggregate investor confidence is a better predictor of trading activity than past returns, as used in prior studies (Statman et al., 2006).

Chapter 3 extends the empirical applications of the new measure to test a range of hypotheses. The profitability of size and momentum strategies increases when aggregate investor confidence is high which complements the overconfidence hypothesis by Daniel et al. (1998) and an early attempt to explain the size premium (Roll, 1981). Additional encompassing tests reveal that the new measure is a better predictor of momentum returns than market states (Cooper et al., 2004).

Chapter 4 takes a qualitative perspective. One shortcoming of an empirical measure of confidence is that overconfidence cannot be distinguished from high levels of confidence, as that requires the assessment of an *observed* level of confidence in comparison with a level of confidence that an individual *should* have. Obviously, assessment of a market-wide level of overconfidence is impossible, especially in retrospective. Borrowing a recent methodology proposed by Glaser et al. (2013), we explore the anatomy of overconfidence in extreme stock market situations in an experiment with experienced and inexperienced participants.

Findings from the experiment suggest that individual overconfidence increases when recent gains are high which complements rationale of chapters 2 and 3, and prior literature (e.g. Cooper et al., 2004; Odean, 1999, 1998; Daniel and Hirshleifer, 2015; Glaser and Weber, 2007). However, if individuals are confronted with evidence that strongly contradicts their prior decisions, they tend to abruptly lose their overconfidence, that is, 'confidence crashes'. We also find evidence that after such crashes, overconfidence can re-emerge when evidence arrives that strongly supports prior decision making.

These additional experimental findings further complement suggestions by Daniel and Hirshleifer (2015) and a hypothesis by Hong and Stein (2007) and Scheinkman and Xiong (2003). They suggest that overconfidence is associated with investors interpreting arriving information about stocks differently, and that this disagreement is particularly pronounced for strong information signals. In line with the hypothesis, we find that time-series variance in stock price prediction tasks increases dramatically when strong feedback signals arrive. Furthermore, we find that the effect is particularly strong when arriving information contradicts prior beliefs.

The interpretation of this finding aligns with heterogeneous investor personalities. When information arrives that reinforces prior decision making, most individuals will conclude that they are indeed brilliant investors. However, when opposing information arrives, those investors suffering from self-attribution bias may neglect such information and fail to update their valuation. When more self-aware investors indeed revise their valuations, higher disagreement as manifested in high time-series variance will be observed.

In addition, we found consistent evidence with Langnickel and Zeisberger (2016) and Biais et al. (2005) who criticise the methodology of using confidence intervals that is commonly used to assess overconfidence. Participants tend to ignore stated confidence levels when producing their responses. In other words, they tend to fail to adjust the widths of confidence intervals with their changing confidence levels. We find similar evidence for stock price prediction tasks. Self-assessed confidence levels were consistently lower than the confidence levels stated in assessment tasks. As a result, overconfidence assessment methodologies tend to systematically produce inflated overconfidence scores, with this presenting a gap for future research. A new method free from this bias would be a great contribution to the literature.

This research provides an avenue for future studies in various ways. Firstly, the INVCON measure may be a suitable tool to explore if investor overconfidence can be linked to liquidity (Pastor and Stambaugh, 2001).

Secondly, the new measure could potentially be used to test the hypothesis of Burnside, Han, Hirshleifer, and Wang (2011) who propose a relationship between investor overconfidence and the forward premium puzzle.

Thirdly, further investigation of the behaviour of overconfident investors in experimental settings could be a fruitful approach to shed light on investor behaviour and risk perception. Do investors suffer from dynamic overconfidence, as suggested in chapter 4? If so, do they indeed change their risk perception and tilt their preference in favour of riskier stocks? Furthermore, a methodology of overconfidence assessment is needed that is robust against measurement bias due to participants' insensitivity to confidence levels.

The future of the field of behavioural finance looks promising. While many recent studies have shed much light on a multitude of puzzles that are only difficult to explain with an entirely rational perspective, much work has to be done to further deepen our understanding of financial markets.

# Bibliography

Agrawal, A. and K. Tandon (1994). Anomalies or illusions? Evidence from stock markets in eighteen countries. *Journal of International Money and Finance 13*(1), 83–106.

Alicke, M. D. and O. Govorun (2005). The better-than-average effect. *The Self in Social Judgment 1*, 85–106.

Alicke, M. D., M. L. Klotz, D. L. Breitenbecher, T. J. Yurak, and D. S. Vredenburg (1995). Personal contact, individuation, and the better-than-average effect. *Journal of Personality and Social Psychology 68*(5), 804.

Asness, C. S., T. J. Moskowitz, and L. H. Pedersen (2013). Value and momentum everywhere. *The Journal of Finance 68*(3), 929–985.

Baker, M. and J. Wurgler (2007). Investor sentiment in the stock market. *Journal of Economic Perspectives 21*, 129–157.

Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics 9*(1), 3–18.

Barber, B. M. and T. Odean (2000). Trading is hazardous to your wealth: The common stock investment performance of individual investors. *The Journal of Finance*, 773–806.

Barber, B. M. and T. Odean (2001). Boys will be boys: Gender, overconfidence, and common stock investment. *The Quarterly Journal of Economics 116*(1), 261–292.

Barberis, N., A. Shleifer, and R. Vishny (1998). A model of investor sentiment. *Journal of Financial Economics 49*, 307–343.

Bem, D. J. (1965). An experimental analysis of self-persuasion. *Journal of Experimental Social Psychology 1*(3), 199–218.

Biais, B., D. Hilton, K. Mazurier, and S. Pouget (2005). Judgemental overconfidence, self-monitoring, and trading performance in an experimental financial market. *The Review of Economic Studies 72*(2), 287–312.

Burnside, C., B. Han, D. Hirshleifer, and T. Y. Wang (2011). Investor overconfidence and the forward premium puzzle. *The Review of Economic Studies 78*(2), 523–558.

Camerer, C. and D. Lovallo (1999). Overconfidence and excess entry: An experimental approach. *American Economic Review*, 306–318.

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance 52*(1), 57–82.

Cesarini, D., Ö. Sandewall, and M. Johannesson (2006). Confidence interval estimation tasks and the economics of overconfidence. *Journal of Economic Behavior & Organization 61*(3), 453–470.

Chan, K., A. Hameed, and W. Tong (2000). Profitability of momentum strategies in the international equity markets. *Journal of Financial and Quantitative Analysis 35*(02), 153–172.

Chordia, T. and L. Shivakumar (2002). Momentum, business cycle, and time-varying expected returns. *The Journal of Finance 57*(2), 985–1019.

Clayson, D. E. (2005). Performance overconfidence: metacognitive effects or misplaced student expectations? *Journal of Marketing Education 27*(2), 122–129.

Conrad, J. and G. Kaul (1993). Long-term market overreaction or biases in computed returns? *The Journal of Finance 48*(1), 39–63.

Cooper, M. J., R. C. Gutierrez, and A. Hameed (2004). Market states and momentum. *The Journal of Finance 59*(3), 1345–1365.

Daniel, K. and D. Hirshleifer (2015). Overconfident investors, predictable returns, and excessive trading. *The Journal of Economic Perspectives 29*(4), 61–87.

Daniel, K., D. Hirshleifer, and A. Subrahmanyam (1998). Investor psychology and security market under- and overreactions. *The Journal of Finance 53*(6), 1839–1885.

Daniel, K., D. Hirshleifer, and A. Subrahmanyam (2001). Overconfidence, arbitrage, and equilibrium asset pricing. *The Journal of Finance*, 921–965.

Daniel, K. and T. J. Moskowitz (2013). Momentum crashes. *Columbia Business School Research Paper* (11-03).

Deaves, R., E. Lueders, and G. Y. Luo (2008). An experimental test of the impact of overconfidence and gender on trading activity. *Review of Finance*, 555–575.

DeBondt, W. F. M. and R. H. Thaler (1985). Does the stock market overreact? *The Journal of Finance 40*(3), 793–805.

DeBondt, W. F. M. and R. H. Thaler (1987). Further evidence on investor overreaction and stock market seasonality. *The Journal of Finance 42*(3), 557–581.

DeBondt, W. F. M. and R. H. Thaler (1994a). Financial decision-making in markets and firms: A behavioral perspective. Technical report, National Bureau of Economic Research.

DeBondt, W. F. M. and R. H. Thaler (1994b). Financial decision-making in markets and firms: A behavioral perspective. Technical report, National Bureau of Economic Research.

Edwards, W. (1968). Conservatism in human information processing. *Formal Representation of Human Judgment*, 17–52.

Einhorn, H. J. and R. M. Hogarth (1978). Confidence in judgment: Persistence of the illusion of validity. *Psychological Review 85*(5), 395.

Erb, C. B. and C. R. Harvey (2006). The strategic and tactical value of commodity futures. *Financial Analysts Journal*, 69–97.

Fama, E. F. (1965). The behavior of stock-market prices. *The Journal of Business 38*(1), 34–105.

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

Fama, E. F. (1995). Random walks in stock market prices. *Financial Analysts Journal*, 75–80.

Fama, E. F. and K. R. French (1992). The cross-section of expected stock returns. *the Journal of Finance 47*(2), 427–465.

Fama, E. F. and K. R. French (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics 33*, 3–56.

Fama, E. F. and K. R. French (1996a). The CAPM is wanted, dead or alive. *The Journal of Finance 51*(5), 1947–1958.

Fama, E. F. and K. R. French (1996b). Multifactor explanations of asset anomalies. *The Journal of Finance 51*(1), 55–84.

Fama, E. F. and K. R. French (2004). The capital asset pricing model: theory and evidence. *The Journal of Economic Perspectives 18*(3), 25–46.

Fama, E. F. and K. R. French (2012). Size, value and momentum in international stock returns. *Journal of Financial Economics 105*(3), 457–472.

Fama, E. F. and J. D. MacBeth (1973). Risk, return, and equilibrium: Empirical tests. *The Journal of Political Economy*, 607–636.

French, K. R. (1980). Stock returns and the weekend effect. *Journal of Financial Economics 8*, 55–69.

Gervais, S. and T. Odean (2001). Learning to be overconfident. *Review of Financial Studies 14*(1), 1–27.

Glaser, M., T. Langer, and M. Weber (2010). Overconfidence of professionals and lay men: individual differences within and between tasks?

Glaser, M., T. Langer, and M. Weber (2013). True overconfidence in interval estimates: Evidence based on a new measure of miscalibration. *Journal of Behavioral Decision Making 26*(5), 405–417.

Glaser, M. and M. Weber (2007). Overconfidence and trading volume. *The Geneva Risk and Insurance Review 32*(1), 1–36.

Glaser, M. and M. Weber (2009). Which past returns affect trading volume? *Journal of Financial Markets 12*(1), 1–31.

Gompers, P. A. and A. Metrick (1998). Institutional investors and equity prices. Technical report, National Bureau of Economic Research.

Gregersen, N. P. (1996). Young drivers' overestimation of their own skill—an experiment on the relation between training strategy and skill. *Accident Analysis & Prevention 28*(2), 243–250.

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

Griffin, J. M., F. Nardari, and R. M. Stulz (2007). Do investors trade more when stocks have performed well? Evidence from 46 countries. *Review of Financial Studies 20*(3), 905–951.

Grinblatt, M. and M. Keloharju (2009). Sensation seeking, overconfidence, and trading activity. *The Journal of Finance 64*(2), 549–578.

Hilton, D., I. Regner, L. Cabantous, L. Charalambides, and S. Vautier (2011). Do positive illusions predict overconfidence in judgment? A test using interval production and probability evaluation measures of miscalibration. *Journal of Behavioral Decision Making 24*(2), 117–139.

Hong, H. and J. C. Stein (1999). A unified theory of underreaction, momentum trading, and overreaction in asset markets. *The Journal of Finance 54*(6), 2143–2184.

Hong, H. and J. C. Stein (2007). Disagreement and the stock market. *The Journal of Economic Perspectives 21*(2), 109–128.

Jegadeesh, N. (1990). Evidence of predictable behavior of security returns. *The Journal of Finance 45*(3), 881–898.

Jegadeesh, N. and S. Titman (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance 48*(1), 65–91.

Jegadeesh, N. and S. Titman (2001). Profitability of momentum strategies: An evaluation of alternative explanations. *The Journal of Finance 56*(2), 699–720.

Johnson, T. C. (2002). Rational momentum effects. *The Journal of Finance 57*(2), 585–608.

Kahneman, D. and A. Tversky (1974). Subjective probability: A judgment of representativeness. In *The Concept of Probability in Psychological Experiments*, pp. 25–48. Springer.

Kahneman, D. and A. Tversky (1979). Prospect theory: An analysis of decision under risk. *Econometrica 47*(2), 263–292.

Keim, D. B. (1983). Size-related anomalies and stock return seasonality: Further empirical evidence. *Journal of Financial Economics 12*(1), 13–32.

Kelly, P. A. and P. Haidet (2007). Physician overestimation of patient literacy: a potential source of health care disparities. *Patient Education and Counseling 66*(1), 119–122.

Keysar, B. and A. S. Henly (2002). Speakers' overestimation of their effectiveness. *Psychological Science 13*(3), 207–212.

Kindleberger, C. P. (2000). *Manias, Panics, and Crashes: A History of Financial Crises*. John Wiley and Sons.

Kirchler, E. and B. Maciejovsky (2002). Simultaneous over- and underconfidence: Evidence from experimental asset markets. *Journal of Risk and Uncertainty 25*(1), 65–85.

Koellinger, P., M. Minniti, and C. Schade (2007). "I think I can, I think I can": Overconfidence and entrepreneurial behavior. *Journal of Economic Psychology 28*(4), 502–527.

Koriat, A., S. Lichtenstein, and B. Fischhoff (1980). Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory 6*(2), 107.

Langer, E. J. and J. Roth (1975). Heads I win, tails it's chance: The illusion of control as a function of the sequence outcomes in a purely chance task. *Journal of Personality and Social Psychology 32*, 951–955.

Langnickel, F. and S. Zeisberger (2016). Do we measure overconfidence? A closer look at the interval production task. *Journal of Economic Behavior & Organization 128*, 121–133.

Larrick, R. P., K. A. Burson, and J. B. Soll (2007). Social comparison and confidence: When thinking you're better than average predicts overconfidence (and when it does not). *Organizational Behavior and Human Decision Processes 102*(1), 76–94.

Lehmann, B. (1990). Fads, Martingales, and Market Efficiency. *The Quarterly Journal of Economics 105*(1), 1–28.

Lemmon, M. and E. Portniaguina (2006). Consumer confidence and asset prices: Some empirical evidence. *Review of Financial Studies 19*(4), 1499–1529.

Lichtenstein, S. and B. Fischhoff (1977). Do those who know more also know more about how much they know? *Organizational Behavior and Human Performance 20*(2), 159–183.

Liew, J. and M. Vassalou (2000). Can book-to-market, size and momentum be risk factors that predict economic growth? *Journal of Financial Economics 57*(2), 221–245.

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics 47*(1), 13–37.

Lo, A. W. and A. C. MacKinlay (1990). When are contrarian profits due to stock market overreaction? *Review of Financial Studies 3*(2), 175–205.

Loewenstein, G. F., E. U. Weber, C. K. Hsee, and N. Welch (2001). Risk as feelings. *Psychological Bulletin 127*(2), 267–286.

Maio, P. and P. Santa-Clara (2012). Multifactor models and their consistency with the ICAPM. *Journal of Financial Economics*.

Malkiel, B. G. (1999). *A random walk down Wall Street: including a life-cycle guide to personal investing*. WW Norton & Company.

Malmendier, U. and G. Tate (2005). CEO overconfidence and corporate investment. *The Journal of Finance 60*(6), 2661–2700.

Malmendier, U. and G. Tate (2008). Who makes acquisitions? CEO overconfidence and the market's reaction. *Journal of Financial Economics 89*(1), 20–43.

McKenzie, C. R., M. J. Liersch, and I. Yaniv (2008). Overconfidence in interval estimates: What does expertise buy you? *Organizational Behavior and Human Decision Processes 107*(2), 179–191.

Menkhoff, L., U. Schmidt, and T. Brozynski (2006). The impact of experience on risk taking, overconfidence, and herding of fund managers: Complementary survey evidence. *European Economic Review 50*(7), 1753–1766.

Miffre, J. and G. Rallis (2007). Momentum strategies in commodity futures markets. *Journal of Banking & Finance 31*(6), 1863–1886.

Miller, D. T. and M. Ross (1975). Self-serving bias in the attribution of causality: Fact or fiction? *Psychological Bulletin 82*(2), 213–225.

Moore, D. A. and P. J. Healy (2008). The trouble with overconfidence. *Psychological Review 115*(2), 502.

Moskowitz, T. J., Y. H. Ooi, and L. H. Pedersen (2012). Time series momentum. *Journal of Financial Economics 104*(2), 228–250.

Odean, T. (1998). Volume, volatility, price, and profit when all traders are above average. *The Journal of Finance 53*(6), 1887–1934.

Odean, T. (1999). Do investors trade too much? *The American Economic Review 89*(5), 1279–1298.

Okunev, J. and D. White (2003). Do momentum-based strategies still work in foreign currency markets? *Journal of Financial and Quantitative Analysis 38*(2), 425–448.

Pastor, L. and R. F. Stambaugh (2001). Liquidity risk and expected stock returns. Technical report, National Bureau of Economic Research.

Plumert, J. M. and D. C. Schwebel (1997). Social and temperamental influences on children's overestimation of their physical abilities: Links to accidental injuries. *Journal of Experimental Child Psychology 67*(3), 317–337.

Ren, Y. and R. Croson (2013). Overconfidence in newsvendor orders: An experimental study. *Management Science 59*(11), 2502–2517.

Roll, R. (1981). A possible explanation of the small firm effect. *The Journal of Finance 36*(4), 879–888.

Rouwenhorst, K. G. (1998). International momentum strategies. *The Journal of Finance 53*(1), 267–284.

Sagi, J. S. and M. S. Seasholes (2007). Firm-specific attributes and the cross-section of momentum. *Journal of Financial Economics 84*(2), 389–434.

Samuelson, P. A. (1965). Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review 6*(2), 41–49.

Scheinkman, J. A. and W. Xiong (2003). Overconfidence and speculative bubbles. *Journal of Political Economy 111*(6), 1183–1220.

Schwert, G. W. (2003). Anomalies and market efficiency. *Handbook of the Economics of Finance 1*, 939–974.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance 19*(3), 425–442.

Shefrin, H. and M. Statman (1985). The disposition to sell winners too early and ride losers too long: Theory and evidence. *The Journal of Finance 40*(3), 777–790.

Shiller, R. J. (2000a). *Irrational exuberance.* Princeton, New Jersey: Princeton University Press.

Shiller, R. J. (2000b). Measuring bubble expectations and investor confidence. *The Journal of Psychology and Financial Markets 1*(1), 49–60.

Soll, J. B. and J. Klayman (2004). Overconfidence in interval estimates. *Journal of Experimental Psychology: Learning, Memory, and Cognition 30*(2), 299.

Statman, M., S. Thorley, and K. Vorkink (2006). Investor overconfidence and trading volume. *Review of Financial Studies 19*(4), 1531–1565.

Stivers, C. and L. Sun (2010). Cross-sectional return dispersion and time variation in value and momentum premiums. *Journal of Financial and Quantitative Analysis 45*, 987–1014.

Sunstein, C. R. and R. Zeckhauser (2011). Overreaction to fearsome risks. *Environmental and Resource Economics 48*(3), 435–449.

Svenson, O. (1981). Are we all less risky and more skillful than our fellow drivers? *Acta Psychologica 47*(2), 143–148.

Teigen, K. H. and M. Jørgensen (2005). When 90% confidence intervals are 50% certain: On the credibility of credible intervals. *Applied Cognitive Psychology 19*(4), 455–475.

Tversky, A. and D. Kahneman (1974). Judgment under uncertainty: Heuristics and biases. *Science 185*(4157), 1124–1131.

Wagenaar, W. A. and G. B. Keren (1986). Does the expert know? The reliability of predictions and confidence ratings of experts. In *Intelligent decision support in process environments*, pp. 87–103. Springer.