Christoph Merkle

Investor Beliefs and Actions

Inauguraldissertation
zur Erlangung des akademischen Grades
eines Doktors der Wirtschaftswissenschaften
der Universität Mannheim

vorgelegt im Herbst-/Wintersemester 2011/2012

Dekan der Fakultät für Betriebswirtschaftslehre: Dr. Jürgen M. Schneider

Referent: Prof. Dr. Dr. h.c. Martin Weber

Korreferent: Prof. Dr. Peter Albrecht

Tag der mündlichen Prüfung: 28.11.2011

Für Christa

Contents

List of Figures

List of Tables

Acknowledgements

This thesis was prepared during my time at the University of Mannheim between 2007 and 2011. I have to thank many for their encouragement and help during this period.

I am greatly indebted to my supervisor, Prof. Dr. Dr. h.c. Martin Weber, for most helpful suggestions, numerous stimulating discussions, and invaluable motivating support. Many thanks are also due to Prof. Dr. Peter Albrecht, my second examiner, and to Prof. Dr. Jürgen Schröder as a member of the dissertation committee.

I appreciate insightful remarks by Dr. Sina Borgsen, Christian Dick, Prof. Juan Dubra, Christian Ehm, Dr. Daniel Foos, Prof. Dr. Markus Glaser, Prof. Ernan Haruvy, Justus Heuer, Prof. Robin Hogarth, Heiko Jacobs, Yigitcan Karabulut, Christine Kaufmann, Prof. Michael Liersch, Dr. Sebastian Mller, Prof. Dr. Alexandra Niessen, Prof. Don Moore, Prof. Dr. Lars Norden, Dr. Alen Nosić, Prof. Chris Parsons, Prof. Dr. Thomas Post, Phillip Schreiber, Ulrich Seubert, Daniel Smith, Dr. Sascha Steffen, Kateřina Strážnická PhD, Michael Ziegelmeyer, and members of the Graduate School of Economic and Social Sciences (GESS). Special thanks go to Daniel Egan, who co-authored one study in this dissertation (chapter 3).

I am very grateful to Barclays Stockbrokers for providing access to their online investor client base and for joint design and execution of the survey. These data were the basis for chapter 2, 3, and 5 of this thesis. I would like to thank in particular the behavioral finance team at Barclays Wealth, Peter Brooks PhD, Greg Davies PhD, and Daniel Egan. Further thanks go to Dr. Alen Nosić, who introduced me to the data set, and to Robin Cindori, who was of great help in preparing the portfolio holdings and trading data for analysis.

Furthermore, I am thankful to Dr. Carsten Schmidt and the team of the economic laboratory of the SFB 504, and to many anonymous but highly motivated students, who participated in the experiments of chapter 4. For this chapter Daniel Smith took great efforts in proof-reading the manuscript. I thank Prof. William Bottom as the associate editor of OBHDP and three anonymous reviewers.

Research reported in this thesis was supported by the European Institute of Savings (OEE) and Deutsche Forschungsgemeinschaft (DFG, grant We993).

Mannheim, December 2011

Chapter 1

General Introduction

1.1 Motivation

Recent years saw the individual investor to become increasingly target of academic research in finance. In his 2006 presidential address to the American Finance Association, John Y. Campbell calls for further efforts in the field of household finance, which studies how households reach their financial objectives using assets and debt (Campbell, 2006). This interest in household finance can be justified already by the size of the private sector owning 45.5 trillion US-$ in financial assets in the US and 4.4 trillion Euro in Germany.[1] Apart from the sheer amount of assets in the hands of individual investors, the growing personal responsibility for old-age provisions is a second factor in rendering household finance increasingly important. This responsibility manifests itself in a shift from defined benefit pension plans to defined contribution plans in the US, and the introduction of privately funded savings schemes (e.g. Riester-Rente) in Germany. Similar trends can be observed for other countries as well.

[1]Sources: Federal Reserve System (http://www.federalreserve.gov/releases/z1/Current/z1r-4.pdf) and Statistisches Bundesamt Deutschland (http://www.destatis.de/jetspeed/portal/cms/Sites/destatis/Internet/DE/Content/Publikationen/Fachveroeffentlichungen/VolkswirtschaftlicheGesamtrechnungen/Vermoegensbilanzen). Both statistics report private households including non-profit organizations.

Under these institutional conditions financial well-being more and more depends on individual investment decisions.

At the same time, we now know that retail investors might not be well equipped to deal with the complex decisions in today's financial markets. On the one hand financial literacy in the population is often found to be poor (Lusardi and Mitchell, 2007), and on the other hand there is plenty of evidence for specific investment mistakes, for instance the disposition effect (Shefrin and Statman, 1985; Weber and Camerer, 1998; Odean, 1998b) or the equity home bias (French and Poterba, 1991; Coval and Moskowitz, 1999). A common research strategy in behavioral household finance is to identify these kinds of biases in investing behavior and to describe their psychological underpinnings. In some cases more general psychological dispositions are established that may hamper investors in a variety of ways, such as overconfidence (Odean, 1998; Glaser and Weber, 2007) or loss aversion (Kahneman and Tversky, 1979; Benartzi and Thaler, 1995). While this undoubtedly has generated many important insights into investing behavior (and even some of the material to be presented here is in this flavor), it may be worthwhile to take a step back and to look at the individual investing problem in a more integrated way.

Comparably little is known about the processes, which lead to investment decisions, and about how beliefs, expectations, and experiences are incorporated into these decisions. From a theoretic point of view Markowitz (1952) describes the process of portfolio selection as follows (p. 77): "The process of selecting a portfolio may be divided into two stages. The first stage starts with observation and experience and ends with beliefs about the future performances of available securities. The second stage starts with the relevant beliefs about future performance and ends with the choice of portfolio." Portfolio choice then gets extensive treatment as a maximization problem, which however does not capture the formation of beliefs and its actual (as opposed to normative) translation into investing behavior. It took a relatively long time before portfolio selection was first analyzed empirically by Lewellen, Lease, and Schlarbaum in the seventies, and still today our understanding of the individual investment process is limited.

When the importance and urgency of research addressing the process of household investing decisions is contrasted to the scarcity of research output on this topic, one may ask why this remarkable discrepancy exists. The reason is probably twofold, on the one hand in economics and finance, there is a traditional leaning towards theoretic (normative) modeling with mainly decision outcomes as objects of these models. Consequently, beliefs and preferences are treated either as given, or as if they can be completely inferred from outcomes (cp. Samuelson, 1938). But perhaps more importantly, it is very demanding to gather data on both input variables of financial decisions (beliefs, preferences, attitudes, dispositions) and output variables (transactions, portfolio choices, financial risk taking). For instance, the meanwhile prominent dataset of Odean (1998) about transactions of discount brokerage clients lacks the former, while investor surveys like the Michigan survey of consumer finance typically lack the latter.

In principle, a combination of the two approaches should reap more powerful and instructive results to understand the individual investment process. We were able to obtain a dataset that includes both, a detailed panel survey about beliefs, expectations, and preferences of online brokerage investors, and the trading activity and portfolio holdings of the very same investors. The survey was constructed in cooperation with the Behavioral Finance team of Barclays Wealth, and was conducted over a time horizon of two years between September 2008 and September 2010. In this capacity the dataset is special, as among the few studies that combine a survey with real trading data, there are even fewer that use a repeated survey over time, which allows studying the dynamic interaction of investor beliefs and trading behavior. Before I outline the different viewpoints under which the dataset is analyzed in this thesis, I will briefly review other literature which examines a combination of survey information and actual investing of private households.

1.2 Survey studies on investing behavior

In an extensive recent review Barber and Odean (2011) cover research on the behavior of individual investors, which is mostly based on trading records of private households. A main focus is on performance of individual investors, the disposition effect, attention and diversification of investors. As explained above I will discuss a subset of this literature, which also employs survey methodology, with the aim to highlight which research questions are addressed and can be addressed using this combined approach.

1.2.1 The beginnings and their lessons for today

Pioneering work in this area has been carried out by Lewellen, Lease and Schlarbaum, who explore data from a large US retail brokerage house, which covers seven years of individual trading data (1964–1970) in 2500 geographically representative accounts. Additionally they gather almost 1000 responses to a mail questionnaire, which contains questions about demographics, investment strategy, information sources, asset holdings, market attitudes and perceptions, and broker relationship. In a first step Lease, Lewellen, and Schlarbaum (1974) describe responses to the questionnaire and demographics of the investor population, in which men, well-educated and older investors are overrepresented (a typical pattern until today). The investors characterize their investment strategy as fundamental, diversified, and long-term oriented, and they mostly use public information and little time for investing. Self-reported asset portfolios contain 40% in direct equity holdings, other major portions are in real estate and personal property.

Investors report (on a five-point scale) to enjoy investing (4.1), they feel substantially better informed than the average investor (3.3) and able to beat institutionals (3.6). At the same time they believe that higher return goes along with higher risk (3.6) and that security prices are not predictable in the short run (3.9). Cohn, Lewellen, Lease, and Schlarbaum (1975) examine the more specific question of how risk aversion and portfolio composition are related. The analysis of risk taking behavior is based on self-reported portfolio compositions and finds a significantly positive influence of age, income and wealth for the proportion of risky assets in investors' portfolios. The results strongly favor a decreasing relative risk aversion among investors. Interestingly a question for risk attitudes which was part of the questionnaire is not used in the analysis.[2] Probably, the authors did not sufficiently trust responses to this qualitative question, which is a sign of the pre-behavioral finance nature of the research approach.

In Lewellen, Lease, and Schlarbaum (1977) the questionnaire is mated with the transaction records for a more in-depth analysis of investment strategies. The authors describe the investment process in a sequential model, starting with basic goals and objectives. Investors then engage in the collection and analysis of information leading to a choice, which is finally evaluated and provides feedback for another round in this sequence. While the resemblance to Markowitz' account of portfolio selection is still apparent, investment behavior is viewed by the authors as a "direct and systematic function of personal circumstances" (p.304). The central dilemma for any study on individual investment behavior is also formulated:

> Because we are dealing heavily with elements of internal attitude development and decision formulation, we should anticipate a substantial level of single-observation "noise" due to those particularities and aberrations of personal makeup which originate inevitably in circumstance dimensions unreachable by variables we can construct as measurement inputs. Hence, we should properly be skeptical of the possibilities for explaining well the fit of each individual investor in the sample into a neat behavior pattern (p.301).

[2] Risk attitude was measured as agreement to the statement "I like to take substantial financial risk to realize significant financial gains from investments" on a five-point Likert scale.

Although this is a problem generally encountered in studies on individual investing behavior (and definitely in later chapters of this thesis), there has been some evolution in the question which dimensions are "unreachable". We now use a variety of psychological constructs to characterize personal makeup that may have looked obscure in the time of the presented research.

Lewellen et al. (1977) then examine, whether there exists a link between stated investing objectives and actual behavior. A general consistency is confirmed, for instance the desire for short-term capital gains is reflected in a less diversified portfolio and a higher trading frequency. The authors find an association between risk attitude and aspects of investment behavior: risk tolerant investors are more likely to trade options, to trade on margin, and to engage in short-sales, but less likely to hold income securities. Furthermore there are interdependencies between risk attitude and investment objectives, risk tolerant investors often aim for short-term capital gains. Return expectations are also evaluated, though not for a specified time period (as e.g. in our survey), but "the average annual pretax percentage rate of return attainable from my portfolio". This average expected return turns out to be positively related to risk tolerance; investors thus seem to be aware that to earn high returns it is necessary to bear some risk. Moreover, return expectations are positively correlated with trading frequency and riskiness of portfolio holdings.

In three further articles Lewellen, Lease, and Schlarbaum report on the performance of the investors in their sample (Schlarbaum, Lewellen, and Lease, 1978; Schlarbaum, Lewellen, and Lease, 1978b; Lewellen, Lease, and Schlarbaum, 1979), which they find on par with institutional investors even after transaction costs. However, the validity of this result is questionable, as it is based on round-trip trades only, and later evidence mostly disproves the finding (cp. Barber and Odean, 2011). More interesting in the present context is the attempt to correlate demographics, strategies and opinions from the survey to performance in investing. Systematic relationships are not detected, which the authors, besides individual account noise, at least partly attribute to the “inevitable imprecision in questionnaire response data” (Lewellen et al., 1979, p.55). It is unclear to which extent this verdict has discouraged later research in this direction, but it is at least notable that for almost twenty years investor surveys were not in vogue.

The presented research program was ahead of its time in another respect, which has been alluded to before. Since behavioral finance had not yet emerged, and in fact it was the heyday of rational portfolio selection and asset pricing (consider only the eminent Merton (1973), published briefly before Lease et al. (1974)), some of the findings could not take their full effect. Exemplarily two quotes shall be reproduced here:

> It could be that what we observe is a psychological rather than an economic phenomenon: a tendency for investors to sell those securities which rise in price and hold the ones which fall, in hopes the latter ultimately recover. (Schlarbaum et al., 1978, p.322f.)

> A high degree of confidence in an ability to forecast does seem an essential emotional prerequisite to both a willingness to trade often and an optimistic view of the likely outcome. In effect, portions of the group may be deluded. (Lewellen et al., 1977, p.322f.)

This is the disposition effect and overconfidence, formulated years before the concepts would gain currency in finance. It is instructive though, how a prevalent mindset can preclude recognizing the importance of these findings. Instead Lewellen et al. (1977) hurry to note that “none of the patterns violates any tenets of rational behavior; by and large, they fit traditional hypotheses (p.320).” While the pre-behavioral finance view of the authors is apparent in the explanations for investing behavior they draw on, their treatment of subjective measures, and their focus on the aggregate rather than the individual, they still provide a work rich in aspects and with many lose ends to build on. Lewellen et al. (1979) themselves call for improved survey instruments to better explain cross-sectional patterns in investing behavior. These instruments are available now and proved fruitful for the research discussed in the next section.

1.2.2 Modern research

One reason for the revival of survey methodology in finance in recent years is the "missing link"-problem that characterizes many studies solely based on transaction data. While this data allows identifying patterns in investing behavior such as excessive trading or the disposition effect, it is hardly able to explain these phenomena. Often psychological findings are invoked to provide reasons for the observed behavior, but there remains a missing link as the potential psychological cause is only assumed and not empirically supported. One strategy to circumvent this problem is to construct a model, which derives the observed investing behavior directly from psychological findings, for instance as demonstrated in Odean (1998). However, this may shift the problem only a bit forward, as it is still unclear whether people show the hypothesized behavior in the financial domain under investigation. A further concern is that a behaviorally motivated model necessarily introduces additional degrees of freedom, which render these models somewhat arbitrary.[3]

A second strategy for filling the gap between psychological causes and investing behavior is to rely on observables in the data. These can be demographic variables such as gender (cp. Barber and Odean, 2001) or even trading behavior itself (cp. Goetzmann and Kumar, 2008). In the provided examples both proxies are chosen to represent overconfidence, in the case of gender based on the finding that men are more overconfident than women, and in the case of trading activity based on the finding that heavy traders are more overconfident. But gender can proxy for many different qualities, most notably risk-aversion (Jianakoplos and Bernasek, 1998; Croson and Gneezy, 2009). And to build a new result on the existence of a previous one (although meanwhile established), might be even more problematic.

[3] A good example for this arbitrariness is the disposition effect, where several alternative explanations have been proposed and partly modeled (for the recent debate cp. Barberis and Xiong, 2009; Kaustia, 2010; Dorn and Strobl, 2009; Hens and Vlcek, 2011).

In recognition of these obvious disadvantages, some researchers have turned to surveys to directly explore the link between psychological states and dispositions and investing behavior. Dorn and Huberman (2005) collect data from "one of Germany's three largest online brokers" with 5.5 years of individual trading data (1995–2000). Questionnaires are send out to a random sample of 2,420 clients and are also online available to 19,000 clients. The paper survey and the online survey yield a total of 1345 responses. Besides demographical variables including gender, age, education, occupation, income, and wealth, the questionnaire elicits unobservable psychological and subjective attributes; in particular measures for investment experience and knowledge, overconfidence, and risk aversion.

Dorn and Huberman (2005) examine the failure to hold diversified portfolios and the tendency of excessive portfolio churning. They find that portfolio volatility and portfolio concentration (as proxies for diversification) decline with age, wealth, experience, perceived knowledge, and (most importantly) risk aversion. While younger, male and less wealthy investors trade more, risk-aversion has a highly negative impact on turnover. Sophisticated investors (actual and self-assessed knowledge) trade more, while overconfidence proxies have no explanatory power. The contribution of the survey in this study is obvious, it allows measuring risk aversion and identifying it as an important and robust determinant of both, diversification and turnover. Any strategy based on observables would have failed to do so, as for example gender (as a possible proxy) is significant in the turnover regression, but not in the diversification regression.

In a second paper on the same dataset, Dorn and Sengmueller (2009) investigate the importance of entertainment as a motivation for trading—that is whether there exist non-pecuniary benefits from researching and executing trades. The entertainment motive is measured in the questionnaire by agreement to four statements ("I enjoy investing."; "I enjoy risky propositions."; "Games are only fun, when money is involved."; "In gambling, the fascination increases with the size of the bet."). As a main result it is found that those who agree to the statements, i.e. trade for entertainment, exhibit significantly higher turnover. Again survey methodology allows to link a psychological construct to trading behavior, and additionally provides the opportunity to test for competing explanations such as investor sophistication, overconfidence, risk aversion, and sensation seeking.

Glaser and Weber (2007) obtain similar data from a German online broker with more than 4 years of individual trading records (1997–2001) covering 3,079 investors and a total of 563,104 transactions. An accompanying online survey is answered by 215 investors. The focus in this questionnaire is on overconfidence measures (miscalibration, better-than-average effect) and on estimates of own performance. This allows for the first time to empirically substantiate the claim in finance models that overconfidence (usually miscalibration) is related to trading volume. As a first result, the authors find that miscalibration measures and better-than-average measures are correlated among themselves, but not mutually. This means that different facets of overconfidence are independent of each other and may also be of differential importance for trading. Glaser and Weber (2007) examine trading volume and turnover in dependence of elicited overconfidence measures. Contrary to theory, miscalibration remains insignificant in these regressions. Instead the better-than-average effect has explanatory power, investors who belief to be better than others, trade more. The study of Glaser and Weber (2007) can serve as a prime example for the power of combined survey and investing data. It does not only enable a direct test of the suggested theory, which in this case yields no evidence in favor of the theory, but also provides a nuanced explanation that rests on another type of overconfidence.

Survey methodology transcends the possibilities of pure trading data in yet another way; it allows to compare actual outcomes to what people believe about these outcomes. With the same dataset, Glaser and Weber (2007b) explore why inexperienced investors do not learn from their (often poor) results in financial markets by analyzing their estimates of past performance. Obviously, knowing what happened in the past is a necessary condition to learn. But interestingly, they find no correlation between return estimates and realized returns. Instead there is a mostly positive difference between estimation and realization, which suggests that overconfidence manifests itself in overestimating past performance. Experienced investors tend to give better estimates, and there is no correlation between relative self-assessments and actual percentile ranking, but the self-assessment is correlated with perceived absolute performance.

The presented studies mark typical examples of the combined data approach, but there are sometimes other ways to obtain similarly powerful datasets. Grinblatt and Keloharju (2009) possess trading data from the "central register of shareholdings" for Finnish stocks (FCSD), which stores recordings of trading of virtually all Finnish investors, both retail and institutional investors from 1995 to 2002. In this case of a very large investor population, conducting a survey would decimate the number of observations substantially, and the anonymized format of the data presents another obstacle. But the authors find a solution by extracting survey-like information from existing sources. They use a mandatory psychological test administered by the Finnish Armed Forces and information on driving related judgments collected by the Finnish Vehicle Administration.

The type of questions in the psychological test at least partly resembles typical questions in investor surveys; and a proxy for overconfidence can be constructed using a measure for self-confidence controlled for actual talent. The number of speeding convictions is employed as a proxy for sensation seeking, assuming that the desire for sensations in driving translates to an analogous desire in investing. As Grinblatt and Keloharju (2009) report, indeed a higher number of speeding convictions increases the probability to trade at all, the number of trades, and turnover. Sensation seeking remains strongly significant even when controlled for overconfidence. Overconfidence itself proves more relevant for the choice to trade, less for trading frequency and turnover.

Some caution is appropriate here, as test responses and behavior fall apart by several years and overconfidence might not be stable but dependent on personal experiences (see later chapters). The approach of Grinblatt and Keloharju thus seems better suited to investigate more stable personality characteristics such as intelligence. The military questionnaire includes an IQ-test, which Grinblatt, Keloharju, and Linnainmaa (2011) use to analyze the link between intelligence and trading behavior. In this respect the quasi-survey bears an advantage compared to deliberately constructed investor questionnaires, as in the latter an IQ-test will usually not be included for institutional, ethical, or merely practical reasons. Grinblatt et al. (2011) find a connection of intellectual ability with stock-picking skill, and with the mitigation of trading costs. Intelligent investors earn superior returns even after control for other demographic characteristics or commonly used risk factors. Intelligent investors also tend to be better diversified, more likely to hold mutual funds and experience less risk (Grinblatt, Keloharju, and Linnainmaa, 2011b).

Juxtaposed to the Finish transaction records enriched by survey-like data, there is the case of genuine investor surveys that lack actual transaction data. To some extent these can be replaced by self-reported investing behavior within the questionnaire. For example Vissing-Jorgensen (2003) and Graham, Harvey, and Huang (2009) use the UBS/Gallup investor survey, which includes questions about stock market expectations and portfolio shares for broad investment categories, and also asks for trading frequency and home bias ("In general, how often do you trade in financial markets?"; "What percent of your portfolio is currently in assets of foreign countries or foreign currencies?"). While this approach certainly has its limitations, it can overcome other problems often encountered with brokerage data. For instance, accounts from one particular broker will mostly cover only part of an investor's portfolio, whereas self-reported data may provide more comprehensive information. Additionally UBS/Gallup conducts random telephone interviews not subject to potential selection issues with brokerage data. Vissing-Jorgensen (2003) finds that investors with higher expectations for stock returns tend to hold higher equity positions, a result that also holds for more specific expectations for an industry (IT stocks). Graham et al. (2009) establish a positive effect of perceived competence on trading frequency and on foreign investment.

1.2.3 Investor panels

So far the discussed studies were static in the sense that the analysis was performed for one point in time or for an average over a time period. The limiting factor is the survey part, as the transaction data in all cases span several years of trading. The next evolutionary step is thus to conduct a repeated survey that enables to match dynamic trading data with dynamic survey responses about beliefs, expectations and preferences. In addition to the work presented in this thesis, there are few contemporaneous studies, which follow this approach and are discussed in this subsection.

Amromin and Sharpe (2009) obtain data from the Michigan Survey of Consumer Attitudes, which conducts monthly interviews with each time over 500 participants. In the time period 2000–2005, 22 of the survey rounds contain questions about knowledge, beliefs, and expectations regarding the stock market. The authors follow the approach of Vissing-Jorgensen (2003) in relying on self-reported portfolio choices within the survey. As the composition of respondents changes in each survey round, the data represent rather a repeated survey than a true panel survey. Nevertheless, if one regards participants as comparable, one can infer trends and reactions over time. In fact, return expectations for the stock market are found to be subject to extrapolation and to depend on past returns, macroeconomic conditions, and individual investor characteristics. Most influences are reverse for risk, which results in negative risk-return correlation and negative Sharpe-ratios.

Based on self-reported portfolio shares, Amromin and Sharpe (2009) find higher equity exposure with those investors who expect higher portfolio Sharpe ratios. A regression of the portfolio share invested in stocks on expected risk premium and expected variance for equity yields a positive effect of expected risk premium and a negative effect of expected variance. In a similar vein, Weber, E.U.Weber, and Nosić (2010) identify a positive impact of return expectations for the stock market on the proportion of risky investing in a hypothetical investment task, whereas risk expectations exert a negative influence. Moreover, changes in return and risk expectations predict changes in risky investment. However, it remains unclear whether the expressed preference in the hypothetical choice task translates into actual investing behavior.

Besides the material presented in this thesis, a contemporaneous study by Hoffmann, Post, and Pennings (2011) is the only one to relate survey responses to actual trading data within a panel format. Their data include trading records of 1,510 clients of a discount broker in the Netherlands and a monthly survey of these investors over one year (2008–2009). The survey primarily elicits return expectations and risk perceptions for the Dutch stock market, and risk tolerance of investors. Investor perceptions strongly react to past market developments, with return expectations and risk tolerance rising and risk perceptions falling in response to past market returns. High levels and positive changes of return expectations increase the likelihood for investors to engage in trading and also have a positive impact on turnover. While higher risk tolerance renders it more likely that an investor trades at all, there is no effect on turnover. Instead turnover positively reacts to levels and changes in risk perception. Comparibly little influence of expectations is found for risk taking (volatility of own portfolios), only levels of risk perception are positively associated with risk taking. Levels and positive changes in risk tolerance induce higher risk taking among investors.

However, the main focus in Hoffmann et al. (2011) is on investor performance. Their results suggest that higher levels and upward revisions of return expectations are related to higher returns and higher risk-adjusted performance (modified Sharpe ratios). In contrast, higher levels of risk tolerance lead to worse performance. A closer inspection over the time of the financial crisis reveals that investors with superior performance within the crisis were already better before the crisis, but are not able to maintain this outperformance after the crisis. The authors interpret this as an indication for investors developing overconfidence, yet they do not present any evidence in investors' expectations supporting this claim.

1.3 Outline of the thesis and main results

Most part of this thesis is based on the above mentioned panel survey of individual investors, which we conducted in collaboration with the Behavioral Finance team of Barclays Wealth. The dataset combines trading records of 617 individual investors with 9 rounds of questionnaire data, administered in three-month intervals between September 2008 and September 2010. The extensive dataset will be described in more detail in the individual chapters, as different aspects of trading behavior and survey responses are important for different research questions. These research questions can be formulated as follows:

1. What determines financial risk taking behavior of investors? Are return expectations, risk expectations and risk preferences relevant? (Chapter 2)

2. What is the role of second-order beliefs in investing? Are second-order beliefs accurate, and if not, which biases can be identified? (Chapter 3)

3. Is there robust evidence for true overconfidence in form of a better-than-average effect? How can overconfidence be measured properly? (Chapter 4)

4. How are different forms of overconfidence related to trading activity, diversification, and risk taking of investors? Does investor overconfidence depend on past investment success? (Chapter 5)

In chapter 2 (joint work with Martin Weber) we examine the basic economic assumption that people act on their beliefs, in particular the premise of portfolio theory that investors form expectations about return and risk of securities and trade accordingly. We test this theory using the described panel survey of self-directed online investors, which asks for return and risk expectation of these investors in three-month intervals between 2008 and 2010. We combine the survey data with investors' actual trading data and portfolio holdings. We find that investor beliefs have little predictive power for immediate trading behavior. However, portfolio risk levels and changes are systematically related to return and risk expectations. In line with financial theory, risk taking increases with return expectations and decreases with risk expectations.

In chapter 3 (joint work with Daniel Egan from Barclays Wealth and Martin Weber) we show that investors use their beliefs about the stock market expectations of others in their investment decisions. These second-order beliefs play a role beyond own risk and return expectations. However, second-order beliefs are inaccurate and exhibit several well-known psychological biases. We document these again in our panel survey of active private investors, who are asked for their return expectations and their beliefs about the return expectations of others. First-order and second-order beliefs differ greatly and investors have only a vague idea what other market participants are thinking. Among the biases we observe is investors' belief that their own opinion is relatively more common among the population. They further assert that others who hold divergent expectations are biased. We interpret these findings as evidence for a false consensus effect and a bias blind spot. The influence of second-order beliefs on investment decisions is mediated by the identified biases.

The third of the presented research questions stands out a bit and so does chapter 4 (joint work with Martin Weber), as it is not directly related to the dataset and only indirectly to investing behavior in general. Instead a theoretical argument about overconfidence in form of the better-than-average effect is made and tested in an experiment. The better-than-average effect describes the tendency of people to perceive their skills and virtues as being above average. We derive a new experimental paradigm to distinguish between two possible explanations for the effect, namely rational information processing and overconfidence. Experiment participants evaluate their relative position within the population by stating their complete belief distribution. This approach sidesteps recent methodology concerns associated with previous research. We find that people hold beliefs about their abilities in different domains and tasks which are inconsistent with rational information processing. Both on an aggregated and an individual level, they show considerable overplacement.[4] We conclude that overconfidence is not only apparent overconfidence but rather the consequence of a psychological bias.

In chapter 5 the concept of overconfidence is applied to investing behavior, for which it has been among the most popular psychological explanations for quite some time. Overconfidence in its nuanced facets has been linked to portfolio turnover, diversification and risk taking, with mostly negative consequences for investors. Several overconfidence measures are derived from stock market and portfolio expectations of the investors in the panel. I find that in general overconfidence is present in our

[4] Overplacement and the better-than-average effect are used synonymously, the same holds for overprecision and miscalibration. Definitions are explained in detail in chapters 4 and 5; we in general follow the terminology introduced by Moore and Healy (2008), but to connect with the existing literature occasionally refer to the older terms.

sample. When the survey data are matched with investors' actual transactions and portfolio holdings, an influence is revealed of overplacement on trading activity, of overprecision and overestimation on degree of diversification, and of overprecision and overplacement on risk taking. Overconfidence hereby leads to increased trading activity, higher risk taking, and less diversification. I explore the evolution of overconfidence over time and indentify a role of past success and hindsight on subsequent degree of overconfidence.

Chapter 2

Do Investors Put Their Money Where Their Mouth Is? Stock Market Expectations and Investing Behavior

2.1 Introduction

People acting on their beliefs is a primitive to economic theory that has seldom been challenged. Portfolio theory assumes that investors form expectations about return and risk of securities and trade accordingly (Markowitz, 1952). We test this theory by empirically investigating the relationship between investors' beliefs and their trading behavior. To this end we collect return and risk expectations in a repeated panel survey of self-directed private investors at a large UK online brokerage provider. In three-month intervals these investors are queried for numerical and qualitative expectations and their risk tolerance. We then match expectations of investors to their actual transactions in their online brokerage accounts. We observe volume, timing, and direction of all trades within the survey period. and are able to calculate portfolio holdings of participants.

We develop different measures of financial risk taking based on trading behavior and portfolio holdings of investors. In a first step we consider the direction of trading and calculate two ratios of buys versus sells. We find that the absolute levels of expectations for market return and risk do not predict buying and selling behavior. An explanation could be that previous expectations are already reflected in investors' portfolios and there is no need for investors to engage in further transactions. We therefore also test whether changes in expectations explain buying and selling behavior corresponding to trades reflecting *changes* in portfolios. Indeed improving return expectations have a positive impact on buy-sell ratios. Thus, quite intuitively, positive return expectations foster buying activity but there is no effect of changes in risk expectations or risk attitude on buy-sell ratios.

Instead we find a pronounced impact of a variable that indicates whether investors regard current market prices as "cheap". Investors viewing prices as cheap tend to be on the buying side of the market. This suggests that the predominant focus of classical finance on returns should be complemented—although theoretically equivalent—by a view on prices. Price levels seem to be easier to remember and to process, and hereby provide orientation for investors; this corresponds to the earlier finding that investor expectations may be very different for prices compared to returns (Glaser, Langer, Reynders, and Weber, 2007). We add that investors also act on their price expectations in addition to their return expectations.

Selling decisions in turn might be closer connected to expectations for the own portfolio rather than to market expectations. We therefore test for the impact of portfolio expectations on sales. In particular there is a group of investors, who hold negative return expectations for their own portfolio. These investors should have a high incentive to change their portfolio composition. Contrary to this hypothesis neither the general propensity to sell is influenced by portfolio expectations nor do investors with negative portfolio expectations react any different from other participants. The only significant effect even suggests that contrary to theory higher return expectations render selling activity more likely. As selling decisions have been shown to be based on realized rather than expected returns (Odean, 1998b), we include investors' portfolio returns as explanatory variable. We find a pronounced positive effect of realized returns on selling propensity; this is in line with the disposition effect (Shefrin and Statman, 1985) which suggests that investors are more likely to sell securities after previous gains. We conjecture that investors selling decision is primarily based on backward looking realized returns rather than forward looking expected returns.

While immediate trading behavior and direction of trade is a means to alter one's risk position, we also directly investigate portfolio risk. We calculate portfolio volatility and beta for investors in our panel as standard risk measures. This is complemented by additional measures such as relative volatility and average component volatility (Dorn and Huberman, 2005). We consider both, levels of portfolio risk at the point in time of survey rounds and changes in portfolio risk between survey rounds. Levels of risk taking of investors can be well explained by their beliefs, preferences and demographics. All portfolio risk measures are positively related to return expectations and risk tolerance, and negatively related to risk expectations, age, and wealth of investors. These results are consistent with financial theory and previous literature.

An advantage of our dataset is that it allows studying the dynamics of this relationship between expectations and risk taking, i.e. whether investors react to changes in expectations by changing their portfolio composition and thus alter risk exposure. For the volatility measures this is the case, as we find a positive change in volatility when return expectations improve and a negative change if investors expect increasing stock market risk. The relationship is weakest for short-term volatility and portfolio beta, indicating that investors manage their portfolios rather based on long-term volatility as a proxy for risk taking.

2.2 Theory and literature

Private investors have been shown to trade frequently (Odean, 1999; Barber and Odean, 2000), to hold underdiversified portfolios (Goetzmann and Kumar, 2008), and to follow various investment strategies (Grinblatt and Keloharju, 2000; Lewellen et al., 1977). Often this behavior is costly for them in terms of lower portfolio returns or higher portfolio risk; in many cases they would be better off by a passive investment in the market index. While part of their motivation to trade actively might be entertainment (Dorn and Sengmueller, 2009) or sensation seeking (Grinblatt and Keloharju, 2009), it is safe to assume that the desire for return and avoidance of (unnecessary) risk is the main objective of private investors.

To reach this objective, most economic models of behavior under uncertainty suggest to evaluate probabilities of future courses of events and then to choose an optimal strategy given these beliefs. This decision is typically characterized by a trade-off between value and risk of different options (Sarin and Weber, 1993). In context of financial markets, portfolio theory advises investors to form expectations about return and risk of securities and to trade accordingly (Markowitz, 1952). Our aim is to test this theory using a panel survey and trading data of self-directed online investors. For this purpose we formulate a very general functional relationship of the form:

$$Financial\ risk\ taking = f(return\ expectations, risk\ expectations, risk\ attitude, x) \quad (2.1)$$

In this function "financial risk taking" will be represented by several measures related to risk taking in actual trades and portfolios of investors. For expectations, we consider numerical and qualitative expectations collected in the survey. It has been demonstrated that qualitative expectations often are a better predictor of investment behavior than numerical expectations (E.U.Weber, Siebenmorgen, and Weber, 2005; Weber et al., 2010). As a parameter for risk preferences, we elicit risk attitude of investors; there might be interdependencies between risk expectations (risk perception) and risk attitudes (E.U.Weber and Milliman, 1997). The variable "x" stands for any other influences on financial risk taking that we might control for.

While this function is very flexible, it does not yield straightforward predictions beyond the fact that expectations and preferences matter in some way for risk taking. Additional assumptions are necessary, which in portfolio theory usually include that return and risk expectations correspond to the first two moments of asset returns, and that risk attitude is represented by the curvature of a utility function. Under standard preferences a simple form to represent this relationship is a linear mean-variance function, which depends on expected returns μ, expected variance σ^2, and a risk aversion coefficient γ.[1]. Financial risk taking (in terms of the proportion of wealth allocated to the risky investment) is then related to expectations and preferences in the following way:

$$\frac{\partial f}{\partial \mu} > 0; \quad \frac{\partial f}{\partial \sigma^2} < 0; \quad \frac{\partial f}{\partial \gamma} < 0 \tag{2.2}$$

Risk taking increases with return expectations and decreases with risk expectations and risk aversion. While this theoretical result may not be valid for other definitions of financial risk taking or qualitative assessments of expectations or preferences, we nevertheless take it as a general prediction for investment behavior in our panel.

[1]For details about which utility functions give rise to linear mean-variance optimization and its performance relative to utility maximization consider e.g. Baron (1977), Kroll, Levy, and Markowitz (1984).

The literature closest related to our study is the work by Dorn and Huberman (2005), Amromin and Sharpe (2009), Weber et al. (2010) and Hoffmann et al. (2011). Dorn and Huberman (2005) also combine survey data with real trading choices of investors and relate psychological attributes to various portfolio characteristics. However, the set of variables they use differs from those we are interested in. They primarily examine the role of overconfidence and risk-aversion for diversification, portfolio turnover, return, volatility, and other relevant measures. In particular they do not consider investor expectations and restrict their analysis to a static survey. We expand this approach by conducting a panel survey which allows us to track changes in opinions and trading behavior over time.

Similar to us Amromin and Sharpe (2009) use panel data, in their case coming from the Michigan Survey of Consumer Attitudes. They concentrate on the interrelation of return expectations and risk expectations, but also provide some evidence of the influence of these variables on portfolio composition. Consistent with financial theory higher return expectations and lower risk expectations increase the share of equity in portfolios of investors. However, Amromin and Sharpe (2009) use self-reported portfolio shares of survey participants and do not have access to their transactions or actual portfolios.

Hoffmann et al. (2011) study an investor survey in the Netherlands which is matched to brokerage account data. Their data spans a time period from April 2008 to March 2009 and survey rounds are administered monthly. By eliciting expectations and portfolio characteristics, Hoffmann et al. (2011) establish a link between the beliefs of investors and their trading behavior concentrating on investor performance. They find that high return expectations, low risk expectations, and low risk tolerance contribute to high returns and Sharpe ratios. We show how these variables influence financial risk taking.

In a previous analysis of our dataset, Weber et al. (2010) report a relationship between expectations and investing decisions. They analyze a survey question which asks participants to split a hypothetical amount of £100,000 between an investment in the UK stock market and a riskless asset. With this investment task they are able to show a strong influence of changes in expectations and risk attitude on changes in the proportion of risky investment; this influence is in the expected direction, increases in expected returns or risk tolerance lead to an increase in risky investment, while higher risk expectations render investors more cautious. We extend this research by relating return and risk expectations to the actual trades and portfolios of investors. By analyzing various aspects of investing behavior, we present a more complete portrayal of the underlying relationships. We also exploit the full time series of the survey which was not available to the earlier study by Weber et al. (2010).

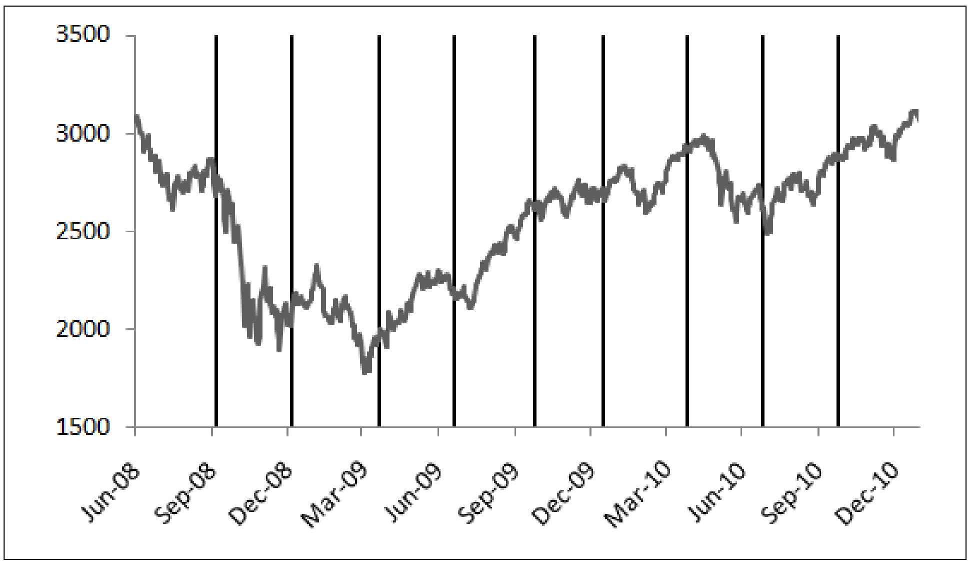
2.3 Data

We obtain survey responses and transaction data for a sample of clients at Barclays Stockbrokers, a UK direct brokerage provider. Barclays is one of the largest brokers in the UK and attracts a wide variety of customers (for demographic characteristics of its clients see chapter 3). The accounts are self-directed in the sense that customers can inform themselves on special webpages provided by the bank, but receive no direct investment advice. Most transactions are processed online.

2.3.1 Survey data

In collaboration with Barclays Wealth, we conduct a repeated survey taking place every three months, beginning in September 2008 and ending in September 2010. Figure 2.1 shows the development of the UK stock market represented by the FTSE all share index and the timing of survey rounds. Our panel consists of nine rounds covering a time period of highly volatile market environment. We thus expect participants to express changing beliefs about market prospects; in the standard model this would in turn lead to changes in their portfolios.

Figure 2.1: FTSE all-share index and survey rounds



Notes: Development of the FTSE all-share index (covers 98% of UK market capitalization) between June 2008 and December 2010. Vertical lines represent the timing of the nine survey rounds.

In the initial survey a stratified sample of the banks client base was invited via e-mail to participate in the online questionnaire (for details on the sampling procedure see Weber et al., 2010). In total 617 clients of the bank participated in the survey, 394 of which participated multiple times. 189 participants have completed at least five rounds, and 52 have participated in all nine rounds. We have a minimum of 130 observations for each of the nine rounds.

We elicit beliefs about return and risk expectations in two ways, by a numerical question asking for return expectations in percentage terms and a more subjective evaluation of risk and return on a bipolar scale. The wording of the numerical question is as follows:

We would like you to make three estimates of the return of the UK stock market (FTSE all-share) by the end of the next three month.

- *Your best estimate should be your best guess.*
- *Your high estimate should very rarely be lower than the actual outcome of the FTSE all-share (about once in 20 occasions)*
- *Your low estimate should very rarely be higher than the actual outcome of the FTSE all-share (about once in 20 occasions)*

Please enter your response as a percentage change.

The question asks participants to predict the three-month return of the UK stock market. We use this time horizon to avoid overlapping observations as the distance between survey rounds is three month as well. In a question design similar to Glaser and Weber (2005), participants have to submit a best estimate as well as a high and a low estimate, which together yield a 90%-confidence interval. We take the best estimate to represent an investor's return expectation about the UK stock market. The high and low estimates allow calculating implicit expected volatility of investors which we use as numerical risk estimate (applying the method of Keefer and Bodily, 1983).

It has been shown that people often have difficulties with numeric estimates (Windschitl and Wells, 1996; Dave, Eckel, Johnson, and Rojas, 2010). This is why already the volatility estimate is indirectly backed out from confidence intervals. Furthermore numeric estimates may not cover all aspects of expected risks and benefits which are partly emotional (Loewenstein, Hsee, E.U.Weber and Welch, 2001). We therefore also include qualitative questions which ask people to evaluate return and risk on a seven-point scale.

- *How would you rate the returns you expect from an investment in the UK stock market (FTSE all-share) over the next 3 months?*

- *Over the next 3-months, how risky do you think the UK stock market (FTSE all-share) is?*

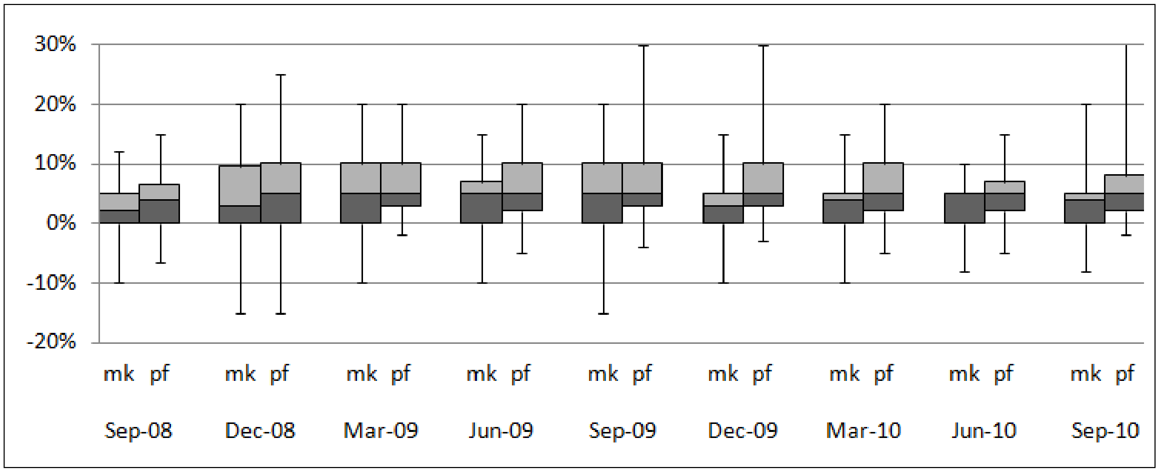
In the first question answer alternatives range from "extremely bad" to "extremely good", in the second question from "not risky at all" to "extremely risky". We ask equivalent questions for investors' own portfolios held with Barclays. In total we thus collect eight belief items per investor per round. Risk attitude of investors is measured using the psychometrical approach described in Brooks, Davies, and Egan (2008). Besides these core variables the survey contains further queries about psychological dispositions and investment objectives. We will refer to these in the result section where appropriate.

2.3.2 Survey responses

Numeric return expectations are relatively low before the peak of the financial crisis, then rise during the crisis and fall again, when the UK stock market recovers. Figure 2.2 shows the pattern in detail. In general investors tend to be more optimistic about their own portfolios: the displayed percentiles for the distribution of expectations are always at least as high and often higher for portfolio expectations. In contrast to market expectations, portfolio expectations remain high in later rounds.

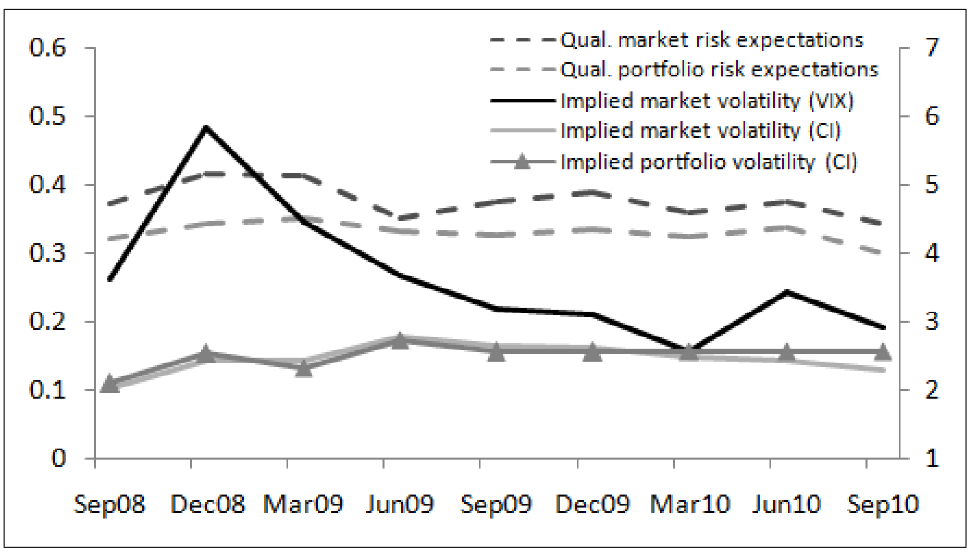
Investors in our panel (numerically) underestimate stock market risk. The implied volatilities calculated from the confidence intervals of investors' return expectations are much lower than volatility expectations of sophisticated market participants (represented by implied option volatilities, see figure 2.3). While confidence intervals are too narrow in the initial survey round, investors seem to learn from observed outcomes that extreme realizations are possible and enlarge their confidence intervals. Expected volatility thus increases, but is still below implied option volatility. Furthermore after the initial adjustment the confidence intervals remain insensitive to subsequent market developments.

Figure 2.2: Numerical return expectations of investors



Notes: Boxplot of numerical return expectations (percentiles:5th,25th,50th,75th,95th) for market expectations (mk) and portfolio expectations (pf).

Figure 2.3: Risk expectations of investors



Notes: Qualitative risk expectations for market and own portfolio (scale 1-7, right axis), and numerical risk expectations as implied by confidence intervals (volatilities, left axis). For comparison implied option volatility (FTSE 100 VIX, left axis)

Qualitative risk expectations elicited on a seven-point scale seem to reflect more closely implied market risk expectations. While it is not possible to compare the absolute magnitudes, we find a correlation of 0.78 ($p < 0.02$) between average qualitative risk expectations and implied option volatilities. Quite intuitively risk expectations rise with the peak of the financial crisis and then fall afterwards. However, there are two further rises in panelists' risk expectations: one, which is not reflected by a rise in option market expectations (Sep to Dec 2009), and another, which falls together with the European debt crisis (June 2010). In general expectations for own portfolio risk follow this trend but are on average slightly lower and more smooth than market expectations.

For investigating trading behavior over time, changes in expectations are particularly important. Table 2.1 shows average changes for all expectation variables. We observe a significant increase in average return and risk expectations between round

Table 2.1: Changes in expectations of investors

		market				own portfolio			
Round	Δ risk tolerance	Δ num. return	Δ qual. return	Δ num. risk	Δ qual. risk	Δ num. return	Δ qual. return	Δ num. risk	Δ qual. risk
2 (Dec08)	0.23**	0.020**	0.12	0.023***	0.43***	0.026**	-0.09	0.023***	0.28***
3 (Mar09)	-0.10	0.014*	0.20**	-0.001	0.03	0.030***	0.20***	-0.007*	0.13*
4 (Jun09)	0.07	-0.010*	0.30***	0.014**	-0.77***	-0.003	0.33***	0.018***	-0.24***
5 (Sep09)	0.15	-0.008	0.01	-0.008	0.38***	0.019**	0.10	-0.008	0.02
6 (Dec09)	-0.14	-0.016	-0.03	-0.011*	0.07	0.014	-0.17**	-0.008	0.06
7 (Mar10)	0.03	-0.004	0.05	0.004	-0.22**	-0.047**	0.11	0.017	-0.12
8 (Jun10)	0.21*	0.008	-0.27**	-0.001	0.17*	-0.010	-0.09	-0.010*	0.13
9 (Dec10)	0.22*	0.009	0.45***	-0.015*	-0.29***	0.045***	0.35***	0.011	-0.29**

Notes: The table states changes in risk tolerance and changes in numerical and qualitative expectations of investors (compared to the previous survey round). Changes are significantly different from zero at *10%-level, **5%-level, or ***1%-level (one-sided t-test).

one and three followed by a very mixed pattern from round three to four (further increase of qualitative return and numerical risk expectations, but sharp drop of qualitative risk expectations). Changes in expectations are less pronounced for the time after the immediate crisis. An exception is the very last survey round for which we observe strongly increasing return expectations and decreasing risk expectations. Similar to Weber et al. (2010), we find that the correlations between changes of numeric and qualitative expectations are often low (return) or insignificant (risk). Stronger correlations exist between market and portfolio expectations. Average risk tolerance remains fairly stable over the whole survey period.

2.3.3 Trading data

Our data also include the trading records of all investors active in the panel survey. In the period between June 2008 and December 2010 we observe 49,372 trades with a total trading volume of £258,940,694.[2]. Of these trades 37,022 or 75% are in stocks (63% of trading volume). In some parts of the analysis we will concentrate on these equity transactions as they are closest related to the expectations we elicit among investors. The remaining trades include bonds, derivatives, mutual funds and ETFs.

The average trader in the panel trades 84.1 times within the 2.5 year period (about three times per month), with a total trading volume of £441,126. However, the distribution is strongly skewed, the median trader trades only 33 times (about once a month; volume £72,805). When we consider only trades in stocks, the values for number of trades and trading volume are reduced by about one third. 30 survey participants do not trade at all during the observation period.[3] An additional 20 investors do not trade in stocks. The mean (median) transaction value amounts to £5,245 (£2,406), and there are many trades worth less than a thousand pounds.

[2]We include three month prior to our first survey round and three month after our last survey round

[3]These investors were not considered in the calculation of means and medians above.

Table 2.2 shows trading frequencies per round of the panel. Rounds are matched to the three month period following each survey. As one can see, there is a falling trend in trading participation and trading frequency among investors. This trend is not caused by panel attrition as those who do not participate in later survey rounds still remain in the trader panel. Instead it is likely that the financial crisis has produced this pattern as the peak of trading coincides with the climax of the crisis. Investors seem to feel the need to react to the turbulent times on financial markets. There is an opposite trend in trading volume, which can be explained by the fact that we take means and medians over those who trade only. Consequently, when the occasional traders stay out of the market (as participation rates suggest), more affluent traders remain and we observe an increase in average volumes. Additionally low volume (high leverage) transactions such as derivatives trades mostly occur during the crisis. The percentage of buys which is always above 50% reflects that given their age most

Table 2.2: Trading frequency and volume over survey rounds

Round	% of panel who trade	Average (median) # of trades		Median trading volume		% of buys
		total	equity	total	equity	
Pre-survey (Jun08-Sep08)	72.4	14.1 (5)	9.3 (3)	8,792	6,740	63.1
Round 1 (Sep08-Dec08)	75.0	16.3 (6)	11.1 (4)	8,770	6,353	66.4
Round 2 (Dec08-Mar09)	71.3	13.8 (5)	10.0 (4)	7,796	4,336	61.4
Round 3 (Mar09-Jun09)	70.0	13.0 (5)	10.1 (3)	9,304	5,362	59.0
Round 4 (Jun09-Sep09)	65.3	12.4 (5)	9.9 (4)	9,229	6,336	55.6
Round 5 (Sep09-Dec09)	67.6	10.3 (4)	8.2 (3)	9,601	6,106	55.5
Round 6 (Dec09-Mar10)	64.2	8.1 (4)	6.2 (3)	9,528	6,934	56.6
Round 7 (Mar10-Jun10)	69.0	10.5 (5)	8.3 (4)	11,730	8,064	58.8
Round 8 (Jun10-Sep10)	66.9	9.8 (4)	7.9 (4)	10,001	7,127	59.8
Round 9 (Sep10-Dec10)	65.5	9.9 (4)	8.0 (3)	10,190	6,840	57.7

Notes: The table shows for all survey rounds the % of survey participants who trade, the average number of trades and median trading volume for those who trade, and percentage of buys (equity only).

inverstors are still saving rather than dissaving. Higher values for the early survey rounds suggest that investors try to profit from low market valuations in the crisis.

Combining trading data with a snapshot of investors' portfolios we are able to calculate portfolio statistics for our survey period. The median portfolio is worth £41,687 (average £314,663) and median portfolio turnover on a per round basis (three month) is 19% (mean 77%), which means that the median investor turns over his portfolio about twice in the survey period of 2.5 years, and some turn over their portfolio ten times or more.

2.3.4 Measures of risk-taking

As a next step we use the transaction records to develop several measures of trading behavior to serve as left-hand side variables in equation 2.1. In Weber et al. (2010), the investment task directly corresponds in space and time to the elicited expectations. In a survey question, participants indicate their desired allocation between the stock market and a riskless asset. By using real trading data we face a different situation: trades occur very rarely at the exact date of the survey but anytime between the survey rounds, and hardly the market index is traded but a variety of securities.

In most cases, extending one's equity position still corresponds to an increase in financial risk taking, while a reduction of one's equity position corresponds to a decrease in risk taking. We thus consider purchases and sales of stocks in the trading records of investors.[4] We form two ratios, based on the number and amount of buys and sales of equity among investors' transactions:

$$\textit{buy-sell ratio} = \frac{number\ of\ buys}{total\ trades} \tag{2.3}$$

[4]A more exact approach would be to look at the proportion of stocks in investors' overall portfolios. However, many investors use the Barclays stockbroker account only for transactions in risky securities. Moreover, we are aware that this account in many cases does not constitute their total investments. This is a common problem in studies on brokerage accounts as noted already by Schlarbaum et al. (1978).

$$\textit{buy-sell volume ratio} = \frac{\textit{buying volume}}{\textit{total trading volume}} \tag{2.4}$$

Similar ratios have been used by Ritter (1988),Grinblatt and Keloharju (2000), and Bhattacharya, Holden, and Jacobsen (2010). Equation 2.3 corresponds to the summary statistics already shown in the rightmost column of table 2.2. The ratio based on volume follows the same pattern, the correlation between both ratios is very high (0.94, $p < 0.01$).

We expect these ratios to be related to investors' stock market expectations: with high return expectations for the stock market, the propensity to buy should rise relative to the propensity to sell; the opposite effect is predicted for high risk expectations. More precisely, for changes in portfolios only changes in expectations should be relevant (cp. Weber et al., 2010). However, as this is a stark theoretical assumption, we analyze both levels and changes of expectations.

Odean (1999, p.1294) argues that a symmetric treatment of buying and selling behavior might not be warranted, since in most cases investors consider only stocks they currently hold in their portfolio for selling. We conclude that investors' expectations for their own portfolio should be more relevant for their selling decision than overall market expectations. To test this conjecture we construct an indicator variable (*sell dummy*) that takes the value of one if a sale occurs in the respective round and zero otherwise. In contrast to the buy-sell ratios this variable disregards any buying activity and allows to study selling decisions in isolation. In a reverse application of the predictions in equation 2.2 more negative portfolio expectations (lower return, higher risk) are expected to augment the propensity to sell.

A second strategy to assess financial risk taking of investors is to use well-known measures of portfolio risk such as volatility and beta (cp. Dorn and Huberman, 2005; McInish, 1982). The assumption is that investors either manage their portfolios according to these risk measures, or that their subjective concept of risk is at least related to these measures. Then for portfolio risk the theoretical predictions in equation 2.2 apply. We calculate volatility of portfolios over one year and over three months horizons. We calculate portfolio beta over a one year horizon using the FTSE all-share index as corresponding market index (this choice seems justified as survey participants hold most of their investments ($> 90\%$) in the UK stock market). And taking into account that within a volatile market environment a large part of the changes in portfolio volatility will be caused by changes in market volatility and may thus not reflect "voluntary" risk taking decisions by investors (although of course a fully rational investor should provide for this), we define relative volatility as another measure of portfolio risk:

$$\textit{relative pf vola} = \frac{\sigma_{pf,t}}{\sigma_{m,t}} \tag{2.5}$$

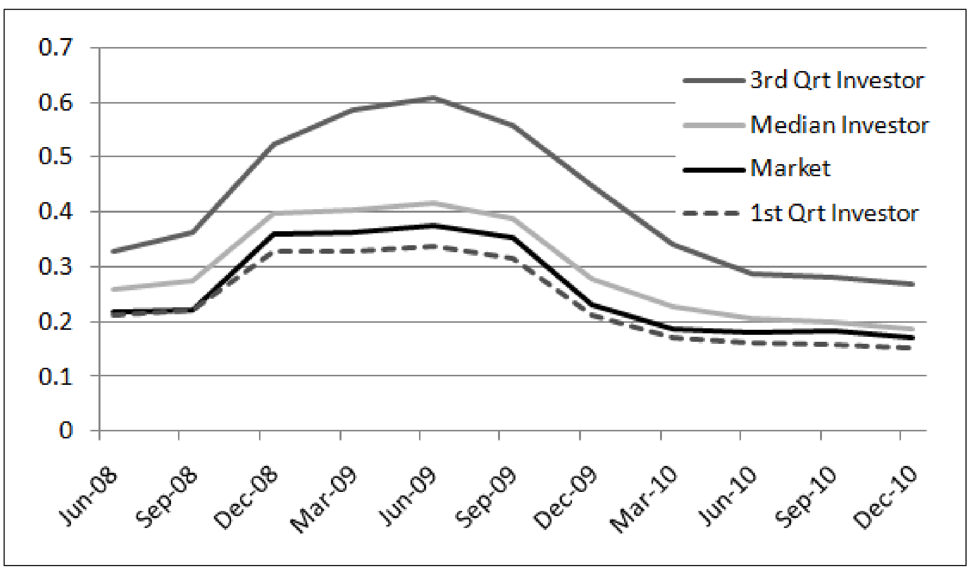
Relative volatility normalizes portfolio volatility by the current level of market risk and thus reflects the current relative riskiness of an investor's portfolio. As we aim to understand risk taking over time, in addition to levels of volatility and beta we also consider changes in these variables between survey rounds, calculated as percentage changes between consecutive rounds (e.g. $\textit{change pf vola} = \frac{\sigma_{pf,t}-\sigma_{pf,t-1}}{\sigma_{pf,t-1}}$). We prefer percentage changes to absolute changes as they are independent of the level of variables.

Dorn and Huberman (2010) argue that portfolio volatility is not the correct measure of risk if investors disregard correlations between securities. They propose a value-weighted average of the return volatilities of portfolio components; resulting average component volatility (ACV) reflects risk taking if investors mainly orient themselves at the volatility of individual securities rather than portfolio volatility. We calculate the ACV and changes in ACV for the investors in our panel.

2.3.5 Descriptive statistics of investor risk-taking

Figure 2.4 displays portfolio volatilities of the median investor, the first-quartile and third-quartile investor of our panel at the time of each survey round. The volatility of the FTSE all-share index serves for comparison. Median portfolio volatility in our panel rises from 0.26 in June 2008 to about 0.40 during the crisis, before falling to values around 0.18 for the last year of the survey. It remains constantly above market volatility, which indicates that a majority of investors hold portfolios that are riskier than the UK market portfolio. The difference between median and market volatility is strongly significant in all rounds ($p < 0.01$, Wilcoxon signed-rank test). The third quartile shows that many investors hold very volatile portfolios compared to the market index, while the first quartile is still close to that index.

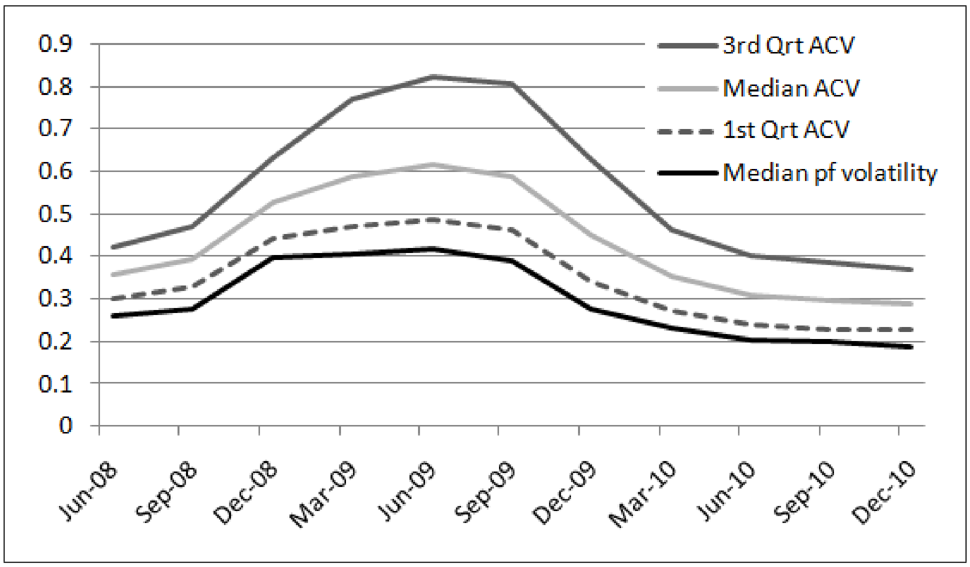
Figure 2.4: Portfolio volatility of investors and UK stock market volatility



Notes: Portfolio volatility is one-year standard deviation of daily portfolio returns at point in time of survey rounds. Displayed are the median investor, the first-quartile and third-quartile investor. UK stock market volatility uses the FTSE all-share index.

The average component volatility (ACV) of course even exceeds these portfolio volatilities as it does not account for diversification effects. Figure 2.5 demonstrates that the ACV follows the general trend of portfolio and market volatilities. However, the median level of ACV is about 40% higher than the median portfolio volatility. In other words, the diversification effect in investors' portfolios reduces risk by nearly one third.

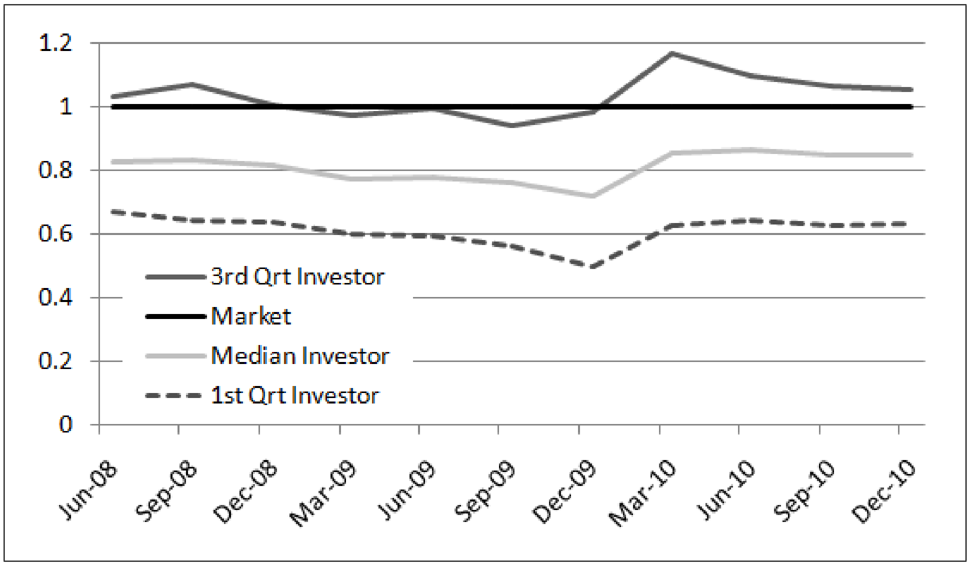
Figure 2.5: Average component volatility of investors



Notes: Average component volatility (ACV) is the value-weighted average of one-year standard deviation calculated from daily returns of portfolio components. Displayed are the ACV for the median investor, the first-quartile and third-quartile investor.

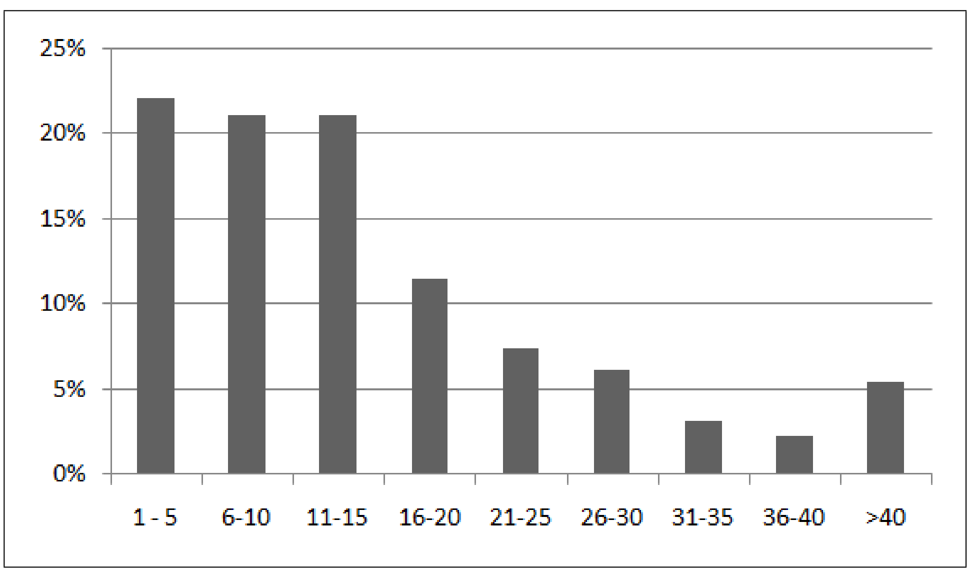
A logical question given investors' high portfolio volatilities is whether they take on high levels of systematic risk, which would be reflected in high betas. Figure 2.6 shows that this is not the case, the median beta is around 0.8, and most investors hold portfolios with a beta smaller than one. Their high volatility seems to be driven by idiosyncratic risk as a result of a low degree of diversification. This is confirmed by the number of portfolio positions investors hold (see figure 2.7). Most investors own between 1 and 15 securities, with the average (median) at 15.7 (12). This is more than in other studies on retail investors (Barber and Odean (2000): average 4.3, Goetzmann and Kumar (2008): average 4.7, Glaser and Weber (2009): average 6.8), but still less than theoretically necessary to obtain a "well-diversified" portfolio (> 30 stocks, Statman, 1987).

Figure 2.6: Portfolio beta of investors



Notes: Portfolio beta is calculated over a one-year window using the FTSE all-share index as the benchmark index.

Figure 2.7: Portfolio positions of investors



Notes: Distribution of the number of positions in the portfolios of survey participants.

As a further measure for the degree of diversification, we calculate a Herfindahl-Hirschmann-Index (HHI) by taking for each investor the sum of squared portfolio weights. Following the methodology of Dorn and Huberman (2005) we treat funds as if they consisted of 100 equally-weighted positions. The median HHI amounts to 0.14, which corresponds to a portfolio of seven equally weighted securities. This means that the above reported number of portfolio positions might overstate the degree of diversification among investors. The effectiveness of portfolio diversification becomes apparent by the fact that for investors with a HHI below the median, the ACV exceeds portfolio volatility by almost 60%.

Table 2.3: Change in portfolio risk

	change in...					
	volatility		relative volatility			
Round	1y	3m	1y	3m	beta	ACV
Pre-survey (Jun08-Sep08)	0.081	0.282	0.069	-0.035	-0.005	0.110
Round 1 (Sep08-Dec08)	0.444	1.026	-0.117	-0.194	-0.057	0.344
Round 2 (Dec08-Mar09)	0.071	-0.277	0.060	0.386	-0.046	0.154
Round 3 (Mar09-Jun09)	0.042	-0.171	0.007	0.015	0.012	0.074
Round 4 (Jun09-Sep09)	-0.058	-0.330	0.002	0.020	-0.034	-0.022
Round 5 (Sep09-Dec09)	-0.248	0.025	0.151	-0.091	-0.010	-0.215
Round 6 (Dec09-Mar10)	-0.196	-0.119	0.002	0.243	0.195	-0.208
Round 7 (Mar10-Jun10)	-0.096	0.490	-0.064	-0.137	-0.003	-0.127
Round 8 (Jun10-Sep10)	-0.014	-0.189	-0.032	-0.005	-0.022	-0.012
Round 9 (Sep10-Dec10)	-0.047	-0.166	0.017	0.120	-0.003	0.013

Notes: The table shows for all survey rounds the change in portfolio volatility, relative portfolio volatility (see eq. 2.5), portfolio beta, and average volatility of portfolio components (ACV).

To summarize, the changes of portfolio risk measures are shown in table 2.3: portfolio volatility and ACV of investors follow the general market trend, a pattern that has already become apparent in the previous figures. Naturally, three-month volatility reacts more quickly and more strongly to these changes. Relative volatilities suggest that investors in the immediate phase of the financial crisis try to reduce their risk exposure relative to the market, while they increase it again afterwards. Changes in beta confirm a reduction in systematic risk for the first phase of the crisis, while for later rounds the results remain inconclusive.

2.4 Results

2.4.1 Investor trading behavior

We first investigate whether market expectations drive the decision of investors to increase or decrease their stock market exposure, which is measured by buy-sell ratios (equations 2.3 and 2.4). We estimate a panel tobit model with random effects as the buy-sell ratios are limited on the interval between 0 and 1, and values on the boundaries occur frequently. We consider two specifications, one in which the absolute *levels* of expectations are relevant for investors, and another in which investors react more responsive to *changes* in expectations.

Column 1 and 5 of table 2.4 show the results of the buy-sell ratios regressed on expectation levels. Levels of expectations at the time of a survey seem to have little effect on subsequent buying and selling behavior. Among the few marginally significant effects is a negative coefficient for risk tolerance. An explanation might be that risk tolerant investors already hold high equity positions and tend to reduce their exposure during the financial crisis. However, this effect is not robust to the inclusion of additional explanatory variables.

Changes in expectations are defined over the same time horizon (between surveys), for which buy-sell ratios are calculated. Among the changes variables, changes in numeric return expectations exert a significant effect on buy-sell behavior (column 2

Table 2.4: Buying and selling behavior

	buy-sell ratio				buy-sell volume ratio			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
num. return	0.108		0.435*	0.239	0.065		0.406*	0.241
num. risk	0.095		-0.068	0.015	0.066		-0.200	-0.029
qual. return	-0.022*		0.015	0.028	-0.022		0.018	0.026
qual. risk	-0.011		0.028	0.024	-0.013		0.032	0.025
risk tolerance	-0.016*		-0.016	-0.013	-0.019*		-0.017	-0.011
Δ num. return		0.238**	0.465***	0.493**		0.302**	0.519***	0.558**
Δ num. risk		-0.092	-0.117	-0.063		-0.101	-0.186	-0.026
Δ qual. return		0.016	0.022	0.034		0.019	0.026	0.033
Δ qual. risk		0.011	0.026	0.023		0.016	0.033	0.025
Δ risk tolerance		0.019*	0.011	0.011		0.017	0.009	0.009
constant	1.050***	0.855***	0.712***	0.324	1.051***	0.830***	0.677***	0.370
n	1376	769	769	457	1376	769	769	457
demographics	no	no	no	yes	no	no	no	yes

Notes: The table shows results of a panel tobit regression with random effects and round dummies. Dependent variable is buy-sell ratio as defined in equation 2.3 for columns (1)-(4) and buy-sell volume ratio as defined in equation 2.4 for colums (5)-(8). Column (1) and (5) include levels of expectations and column (2) and (6) changes of expectations as explanatory variables. Column (3) and (7) show regressions on both, levels and changes; in column (4) and (8) it is additionally controlled for demographics. Demographic variables include age, gender, number of dependents, marital status, income, wealth, financial literacy and market view. Coefficients are significant at *10%-level, **5%-level, or ***1%-level.

and 6). If return expectations improve investors tend to move to the buying side of the market, which is consistent with the prediction in 2.2. For additional equity purchases thus not the absolute level of return expectations is relevant, but rather changes in these expectations; this result is robust to the inclusion of the levels variables (column 3 and 7) and of demographic variables: age, gender, number of dependents, marital status, income, wealth, financial literacy and market view (column 4 and 8).[5] Of these demographic variables (results not reported) only income, wealth, and market view have a significant impact. Income quite intuitively has a positive effect on buy-sell ratios as it is a proxy for additional liquidity investors might want to invest ($p < 0.01$). Wealth has a negative impact on buy-sell ratios ($p = 0.02$). Finally market view, which is an indicator whether investors consider the market as "cheap", has a positive effect ($p = 0.05$).

For robustness we exclude heavy traders (the top 10% in number of trades and trading volume), as these investors might be engaged in trading activity independent of their current beliefs or other situational factors. When investors, who trade less frequently, place an order, this order might be more closely related to personal return and risk expectations. However, there is almost no change in the results under this restriction. We also test a different way to measure increases or decreases in risky exposure by calculating the change in portfolio value that is not due to portfolio returns. This portfolio expansion or contraction corresponds to investors that predominantly reside on the buy or sell side of the market. Similarly to buy-sell ratios, the effect of expectations on changes in portfolio value is weak.

To disentangle the buying and selling decision, we now focus on sales of securities only. For these transactions the choice set for investors effectively narrows to the securities they hold in their portfolios (as we do not observe any short sales). Consequently portfolio expectations should drive these decisions.[6] Table 2.5 shows results of a panel probit model with the previously defined *sell dummy* as dependent variable, which takes a value of one if a selling transaction occurs between two survey

[5] The varying number of observations in the different regressions is due to the fact that for changes in expectations we need investors to participate in the survey for two consecutive rounds. Additionally certain demographics (in particular income and wealth) are not stated by all participants.

[6] We also test for market expectations as determinants of selling behavior, results are in general weaker than for portfolio expectations.

dates. Again we first consider levels of portfolio expectations (columns 1 and 2) and then changes of portfolio expectations (columns 3 and 4) as explanatory variables.

Table 2.5: The selling decision

	sell dummy					
	(1)	(2)	(3)	(4)	(5)	(6)
num. pf return	0.192	0.060			-0.685	-1.613*
num. pf risk	0.662*	0.565			0.605	0.157
qual. pf return	0.114***	0.129***			0.208***	0.275***
qual. pf risk	0.042	0.030			0.122*	0.205**
risk tolerance	0.070***	0.061**			0.038	0.012
Δ pf num. return			-0.272	-0.321	-0.064	-0.832
Δ pf num. risk			0.083	0.015	0.297	0.075
Δ pf qual. return			-0.075*	-0.084*	0.019	0.057
Δ pf qual. risk			-0.020	-0.021	0.040	0.081
Δ risk tolerance			-0.042	-0.038	-0.021	-0.057
pf realized return		0.651**		0.886**	0.938**	0.995**
pf realized vola		0.192		0.312	0.269	-0.007
constant	-1.046***	-0.923***	0.124	0.065	-1.408***	-3.269***
n	2016	1972	1126	1101	1101	788
demographics	no	no	no	no	no	yes

Notes: The table shows results of a panel probit regression with random effects and round dummies. Dependent variable is a sell dummy, which takes a value of one if an investor sells a security between to survey dates and zero otherwise. Demographic variables include age, gender, number of dependents, marital status, income, wealth and financial literacy. Coefficients are significant at *10%-level, **5%-level, or ***1%-level.

In the levels regressions high risk tolerance increases the propensity to sell, which results from a general increase in trading frequency among risk tolerant investors. Consequently, when we exclude heavy traders as defined above, risk tolerance is no longer significant for selling activity. High (qualitative) portfolio return expectations make it more likely to engage in a selling transaction. This result seems counterintuitive, as one rather would expect investors with poor return expectations to sell off portions of their portfolio. Indeed, some investors in the panel hold negative return expectations for their portfolio, but they are *less* inclined to change their portfolio composition. An explanation could be that regarding sales previous returns may be a more relevant consideration for investors than forward looking expectations, as the disposition effect suggests (Shefrin and Statman, 1985; Odean, 1998b; Weber and Camerer, 1998). In column (2) we thus include realized portfolio performance and volatility of investors. We find a very robust positive effect of realized returns, which is in line with investors selling after gains. Following this logic a positive effect of return expectations could be observed, if sales increase after part of the expectations has materialized. The regression on changes then explains that the point in time to sell is when expectations worsen (negative coefficients for changes in return expectations). However, this effect is only marginally significant and does not hold up to the inclusion of the level variables.

The explanatory power of change variables (columns 3-6) is generally poor; regression (3) even fails to attain joined significance. The results are robust to an alternative specification using the fraction of portfolio sold as dependent variable. For the fraction of portfolio sold, the level of perceived portfolio risk is more important: the higher the risk perception the greater the portion of the portfolio that is subsequently sold. Taken together with the previous buy-sell ratio regressions, our findings suggest that it is hard to predict immediate trading activity from investor expectations. Beliefs do not as directly and as stably determine the direction of trades as function 2.1 suggest, and the influence of other factors represented by “x” may be large. We have proposed some of these factors such as liquidity[7], realized returns or price levels. Additionally, expectations on market or portfolio level may be to coarse to explain trading decisions in individual securities.

[7]Income is again among the significant demographic variables in column 6 of table 2.5 with a negative effect on propensity to sell, confirming an influence of liquidity

2.4.2 Investor portfolio risk

We now turn to investor portfolio risk, which should be a more stable measure of investor risk taking. In our analysis, we interpret the volatility levels of investors' portfolios when the survey takes place as the *level* of risk an investor is taking at this point in time. Consequently, changes in volatility correspond to *changes* in risk taking.[8] Similarly, we use levels and changes of other portfolio risk measures (beta, relative volatility, average component volatility).

Table 2.6: Correlation of portfolio risk measures

PANEL A	Levels of portfolio risk				
	Vol 1y	Vol 3m	Rel. Vol	Beta	ACV
Volatility 1y	1.00				
Volatility 3m	0.76	1.00			
Rel. volatility	0.89	0.59	1.00		
Portfolio beta	0.42	0.28	0.43	1.00	
ACV	0.64	0.54	0.50	0.23	1.00
PANEL B	Changes of portfolio risk				
	Δ Vol 1y	Δ Vol 3m	Δ Rel. Vol	Δ Beta	Δ ACV
Δ Volatility 1y	1.00				
Δ Volatility 3m	0.60	1.00			
Δ Rel. volatility	0.39	0.11	1.00		
Δ Portfolio beta	0.13	0.05	0.40	1.00	
Δ ACV	0.60	0.32	0.06	-0.05	1.00

Notes: The table shows pairwise Pearson correlations of levels (Panel A) and changes (Panel B) of portfolio risk measures. All correlations are significant at 1%-level.

Panel A of table 2.6 shows correlations between the levels of these measures; all correlations are positive as they share a common concept of risk, but the variables

[8] This is a deliberate analogy to levels and changes in the hypothetical risk taking task analyzed by Weber et al. (2010). In this task investors had to divide £100,000 between the FTSE-all share and a riskless asset. If we assume a volatility of 0 for the riskless asset, the volatility of the chosen portfolio is monotonically increasing with the fraction invested in the FTSE.

also capture different aspects of risk as correlations are not perfect. In particular, portfolio beta shows the weakest relation to other risk measures with coefficients between 0.23 and 0.43. When considering changes (Panel B) the picture becomes even more mixed. All but one correlation are still positive, but especially for beta and three-month volatility (which is the only measure calculated over a shorter time horizon) coefficients are low. As portfolio risk measures seem to differ, we consider all of them in our regression analysis.

We run regressions with one-year and three-month portfolio volatilities as dependent variables; as volatilities are skewed we take the natural logarithm of these variables. We use market expectations as explanatory variables, because there exists a reverse causality problem (endogeneity) with portfolio expectations: current portfolio volatility will determine expectations for future portfolio returns and volatility. Table 2.7 shows the results of a pooled OLS regression (columns 1-4) and a panel GLS regression (columns 5-8). We find that the volatility level investors take on in their portfolios strongly depends on their expectations. In all regressions, a positive impact of numerical return expectations on volatilities and a negative impact of numerical risk expectations can be observed. Both effects are strongly significant and robust to the change of the model specification and inclusion of additional variables. Higher risk tolerance also contributes positively to the volatility investors take on. In contrast, qualitative expectations have no explanatory power for portfolio volatility. Among the demographic variables, we find significant effects for age and wealth.[9] Younger investors hold more volatile portfolios, while wealthier investors tend to own less risky portfolios. This result is consistent with the findings of Dorn and Huberman (2005).

We will now analyze the other portfolio risk measures analogously. Table 2.8 shows regression results for relative volatility, portfolio beta and average component volatility (ACV). In general the results support our findings for portfolio volatility. Numer-

[9]For an explanation of the demographic variables consider the notes to table 2.7.

Table 2.7: Portfolio volatility

	Pooled OLS				Panel GLS			
	ln(Vol 1y)		ln(Vol 3m)		ln(Vol 1y)		ln(Vol 3m)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
num. return	0.471***	0.475***	0.467***	0.459***	0.168***	0.169***	0.166***	0.115**
num. risk	-0.467***	-0.498***	-0.470***	-0.524***	-0.132***	-0.148**	-0.192***	-0.239***
qual. return	0.002	0.005	0.001	0.002	-0.001	-0.003	0.000	-0.003
qual. risk	0.014	-0.010	-0.025	-0.022	-0.001	0.003	-0.015	-0.009
risk tolerance	0.032***	0.022**	0.034***	0.024**	0.003	0.005	0.010**	0.012**
Age		-0.004*		-0.004**		-0.006***		-0.006***
Gender		0.034		0.034		0.070		0.114
Dependents		0.000		-0.001		0.008		0.009
Couple		-0.049		-0.047		-0.053		-0.045
Income		-0.001		0.003		-0.013		-0.011
Wealth		-0.043***		-0.039***		-0.034***		-0.034***
Fin. lit.		0.010		0.010		-0.047		-0.045
constant	-1.280***	-0.907***	-1.145***	-0.762***	-1.226***	-0.604***	-1.110***	-0.564
n	1935	1394	1922	1381	1935	1394	1922	1381
R^2	0.326	0.415	0.470	0.548	0.303	0.389	0.460	0.533

Notes: The table shows results of a pooled OLS regression (columns 1-4) and a panel GLS regression with random effects (columns 5-8). All regressions contain round dummies and standard errors are clustered by participant. Dependent variable is the natural logarithm of portfolio volatility calculated over a 1 year or 3 month horizon. Demographics: Age in years, gender dummy (male=1), number of dependents, couple dummy (married or co-habiting=1), income in categories, wealth in categories, financial literacy (number of correct answers using four of the basic literacy questions by van Rooij, Lusardi, and Alessie (2011)). Coefficients are significant at *10%-level, **5%-level, or ***1%-level.

ical return expectations are positively related to risk taking, the opposite holds for numerical risk expectations. When qualitative expectations pick up some significance, they do so in the same direction as numerical expectations. Risk tolerance has a positive impact on risk taking except for portfolio betas; betas are least well explained by our model which is confirmed by low R^2. Among the demographic variables, again age and wealth exert a negative influence on risk taking. For robustness we complete our demographic variables by linear imputation of missing values for wealth and income, which extends the number of observations for the regressions that include demographics. Non-reported results are similar to the shown specifications in tables 2.7 and 2.8.

Table 2.8: Other portfolio risk measures

	Pooled OLS			Panel GLS		
	rel.Vol	Beta	ACV	rel Vol	Beta	ACV
	(1)	(2)	(3)	(4)	(5)	(6)
num. return	0.475***	0.301***	0.269*	0.169***	0.055	0.162***
num. risk	-0.498***	-0.376**	-0.439**	-0.148**	-0.025	-0.159
qual. return	0.005	0.029**	0.013	-0.003	-0.005	-0.001
qual. risk	-0.010	-0.012	-0.023	0.003	-0.003	-0.014***
risk tolerance	0.022**	-0.005	0.025***	0.005	-0.001	0.013***
Age	-0.003*	-0.000	-0.001	-0.006***	-0.002	-0.004**
Gender	0.034	0.093	0.058	0.070	0.103	0.107
Dependents	0.000	0.007	0.019	0.008	0.002	0.019
Couple	-0.049	-0.000	-0.066	-0.053	-0.014	-0.063
Income	-0.001	-0.002	-0.015	-0.013	-0.002	-0.021
Wealth	-0.043***	-0.020**	-0.020**	-0.034***	-0.017*	-0.019*
Fin. lit.	0.010	0.001	-0.001	0.047	-0.073	-0.019
constant	0.610***	0.861***	0.705***	0.914***	1.278***	-0.604***
n	1394	1396	1323	1394	1396	1323
R^2	0.175	0.078	0.319	0.137	0.030	0.305

Notes: The table shows results of pooled OLS regressions (columns 1-3) and panel GLS regression with random effects (columns 4-6). All regressions contain round dummies and standard errors are clustered by participant. Dependent variables are the natural logarithm of relative volatility (see eq. 2.5), portfolio beta, and the natural logarithm of average component volatility; all calculated over a one-year horizon . Coefficients are significant at *10%-level, **5%-level, or ***1%-level.

Up to this point we dealt with state variables that give us some information which portfolio risk investors choose depending on their expectations, risk tolerance, and demographics. The panel structure of our data allows us to investigate at more detail the dynamics of these relationships. We now analyze changes of portfolio volatility for one-year and three-month horizons. Changes in portfolio volatility are related to contemporaneous changes in investor expectations. Table 2.9 shows the results of these regressions. With changes in one-year portfolio volatility (column 1 and 5) we observe the same patterns as in the levels regression. Positive changes in numerical return expectations are accompanied by increased risk taking, while higher numerical risk expectations result in decreased risk taking. Counterintuitive is the negative coefficient for changes in risk attitude, an increase in risk tolerance leads to less risk taking. In column 2 and 6 we replace the round dummies in the regression by changes in market volatility. This variable is constant in the cross-section and will

Table 2.9: Changes in portfolio volatility

	Pooled OLS				Panel GLS			
	change pf vol1y		change pf vol3m		change pf vol1y		change pf vol3m	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Δ num. return	0.066***	0.063**	0.099	0.126	0.061**	0.056**	0.112	0.139
Δ num. risk	-0.081**	-0.074*	-0.109	-0.084	-0.078*	-0.074*	-0.126	-0.099
Δ qual. return	0.001	0.002	0.002	0.001	0.003	0.004	0.005	0.003
Δ qual. risk	0.002	-0.001	-0.007	-0.008	0.001	-0.002	-0.009	-0.011
Δ risk tolerance	-0.006**	-0.007**	-0.009	-0.010	-0.005*	-0.006**	-0.008	-0.010
Δ market vol.		0.719***		0.678***		0.717***		0.676***
constant	-0.013	-0.007	-0.187***	-0.003	-0.024**	-0.008	-0.183***	0.004
round dummies	yes	no	yes	no	yes	no	yes	no
n	1040	1040	1033	1033	1040	1040	1033	1033
R^2	0.683	0.668	0.653	0.642	0.683	0.668	0.653	0.642

Notes: The table shows results of a pooled OLS regression (columns 1-4) and a panel GLS regression with random effects (columns 5-8). All standard errors are clustered by participant. Dependent variable is percentage change of portfolio volatility between survey rounds, calculated over a 1 year or 3 month horizon. Coefficients are significant at *10%-level, **5%-level, or ***1%-level.

thus capture the part of changes in portfolio volatility caused by a change in overall market volatility. The coefficients are around 0.7 which means that about 70% of changes in portfolio volatilities are driven by changes in market volatility; this is in line with average betas we found for participants.

In the regressions of changes in three-month volatilities on changes in expectations (see table 2.9), the coefficients for numerical expectations maintain their direction, but no longer reach statistical significance. This may be due to the diminished statistical power of the changes regressions, as we can only consider investors who participate in two subsequent survey rounds. However, another interpretation is that investors have more long-term objectives and do not manage their portfolios according to three-month volatilities. In our questionnaire, most investors state an investment horizon of three to five years.

Changes regressions of other portfolio risk measures are shown in table 2.10. For relative volatility and average component volatility similar patterns as for volatility emerge. In particular numerical return expectations positively influence risk taking. Interestingly changes in market volatility have a negative impact on relative volatility (columns 2 and 6). As already conjectured in the discussion of summary statistics, investors seem to counteract rising market volatility in an attempt to reduce their portfolio risk relative to the market. We observe no effects of changes in expectations on changes in portfolio beta (columns 3 and 7). As already found for levels, beta is the risk measure least related to expectations. It is likely that beta has little relevance to investors in managing the risk of their portfolios. Many private investors may not even know about this concept.

Table 2.10: Changes in other portfolio risk measures

	Pooled OLS				Panel GLS			
	change rel. vol.		Δ beta	Δ acv	change rel. vol.		Δ beta	Δ acv
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Δ num. return	0.060***	0.054**	0.001	0.077**	0.058***	0.052**	0.001	0.079**
Δ num. risk	-0.071**	-0.064*	-0.024	0.025	-0.072**	-0.068*	-0.019	0.021
Δ qual. return	-0.000	0.001	-0.006	-0.005	0.000	0.001	-0.005	-0.006
Δ qual. risk	0.000	-0.000	0.003	-0.004	0.000	-0.000	0.002	-0.003
Δ risk tolerance	-0.005*	-0.006**	-0.001	-0.004	-0.004	-0.005*	-0.000	-0.004
Δ market vol.		-0.206***				-0.212***		
constant	-0.031***	0.008*	-0.005	-0.001	-0.038***	0.009	-0.015	-0.003
round dummies	yes	no	yes	yes	yes	no	yes	yes
n	1040	1040	1011	1020	1040	1040	1011	1020
R^2	0.241	0.190	0.140	0.418	0.241	0.190	0.140	0.418

Notes: The table shows results of a pooled OLS regression (columns 1-4) and a panel GLS regression with random effects (columns 5-8). All standard errors are clustered by participant. Dependent variable is change of relative volatility (as defined in equation 2.5), change of portfolio beta, and change of average component volatility. All variables are calculated over a one-year horizon. Coefficients are significant at *10%-level, **5%-level, or ***1%-level.

2.4.3 Volatility of trades

We combine the two approaches of measuring financial risk taking and examine the volatility of securities investors are trading. For this purpose, all securities traded by survey participants (and for which a sufficient time series of returns is available) are sorted by return volatility throughout the survey period. We form ten volatility deciles and hereby establish a ranking of securities by their relative riskiness. We then calculate the value-weighted average of volatility decile each investor trades in. We also compute the volatility of purchases and the volatility differential between purchases and sales. The latter we interpret similar to buy-sell ratios as an indicator

Table 2.11: Volatility of securities traded

Round	trade volatility	buy volatility	buy-sell vol. diff.
Pre-survey (Jun08-Sep08)	6.20	6.42	0.33^{***}
Round 1 (Sep08-Dec08)	6.23	6.37	0.29^{**}
Round 2 (Dec08-Mar09)	6.02	6.05	-0.24^{**}
Round 3 (Mar09-Jun09)	5.99	5.79	-0.39^{***}
Round 4 (Jun09-Sep09)	6.00	5.88	-0.32^{***}
Round 5 (Sep09-Dec09)	5.96	6.07	-0.02
Round 6 (Dec09-Mar10)	6.03	6.04	-0.06
Round 7 (Mar10-Jun10)	5.75	5.60	-0.36^{**}
Round 8 (Jun10-Sep10)	5.86	5.84	-0.29^{**}
Round 9 (Sep10-Dec10)	5.82	5.76	-0.23^{*}

Notes: The table shows for all survey rounds the average volatility decile of trades and purchases, and the average volatitity differential between purchases and sales. This difference is significant by a Wilcoxon signed-rank test at *10%-level, **5%-level, or ***1%-level.

of financial risk taking; if the difference is positive an investor shifts money to more volatile securities.

Table 2.11 shows population averages of volatility of trades, of volatility of purchases, and the average buy-sell volatility differential. We observe that investors trade securities that are slightly more volatile than the total sample of securities (which of course has an average decile rank of 5.5). This is due to the fact that mutual funds and ETFs are less frequently traded than more volatile securities such as stocks and options. Volatility of trades and purchases is highest in the first two rounds of the survey; these are also the only rounds where the buy-sell volatility differential is positive which confirms the earlier finding that private investors in our sample seem to view the crisis as an opportunity to buy risky securities. This behavior then turns around, in particular for a period of high stock market gains in mid-2009 (cp. also figure 2.1). Investors move back into safer securities, a behavior that repeats itself for the final survey rounds, for which the average volatility of trades is lowest on average.

When we regress the three measures defined above on the levels of investors' expectations and risk tolerance (table 2.12), we find a pronounced impact of (numerical) return expectations: high return expectations induce trades in riskier securities, especially for purchases. The effect is strongest for the volatility differential between purchases and sales which means that investors shift capital towards riskier securities in presence of high return expectations. We also find that less risk-averse investors trade and buy securities with higher volatility, in line with risk habitat theory which

Table 2.12: Volatility of trades explained by expectations

	Panel GLS					
	trade volatility		buy volatility		buy-sell vol. diff.	
	(1)	(2)	(3)	(4)	(5)	(6)
num. return	0.591	1.097**	0.968*	1.690***	1.282***	2.262***
num. risk	-0.159	0.341	-0.579	0.266	0.370	0.918
qual. return	0.002	0.014	0.007	0.002	-0.002	-0.037
qual. risk	-0.052	0.011	-0.045	-0.001	-0.047	-0.070
risk tolerance	0.063**	0.059*	0.099***	0.077**	0.035	0.010
Age		-0.017**		-0.017**		0.007
Gender		0.585		0.328		0.027
Dependents		0.026		0.066		0.002
Couple		-0.171		-0.246		-0.211
Income		-0.030		-0.036		0.050
Wealth		-0.120***		-0.131***		-0.042
Fin. lit.		-0.166		-0.138		0.097
constant	6.217***	7.515***	6.229***	7.833***	0.341	-0.108
n	1478	1045	1354	958	894	626
R^2	0.038	0.076	0.060	0.103	0.038	0.037

Notes: The table shows results of panel GLS regression with random effects, all regressions contain round dummies and standard errors are clustered by participant. Dependent variables are the volatility of trades, the volatility of purchases and the difference between volatility of purchases and sales. Coefficients are significant at *10%-level, **5%-level, or ***1%-level.

states that investors select securities of which volatilities are commensurate with their risk aversion (Dorn and Huberman, 2010). There is no such effect for buy-sell volatility differentials, but this is not surprising as both risk-averse and risk tolerant investors will sometimes augment and sometimes reduce risk (though on different levels). Again older and wealthier investors trade less volatile securities. We do not report results for a regression on changes in expectations in this case, as we find no significant results.

2.5 Discussion

A main problem any research in beliefs and expectations encounters is whether the responses in a survey are valid representations of the internal beliefs of participants. The challenge is twofold, questions need to be stated in a way that participants are able to answer them in a sensible way, and participants need to be motivated to do so. For the latter we rely on the intrinsic motivation of participants as they completed the survey voluntarily, and many found it interesting enough to take part multiple times. As in most large-scale surveys, monetary incentives were not feasible but we are in this case not aware of any obvious reason to conceal or distort beliefs in their absence.[10] Additionally we build on the finding of Weber et al. (2010)—who use the same survey—that the elicited expectations are effective and consistent predictors of decisions, which should attenuate concerns about their validity.

The other concern that participants might not be able to express their beliefs in the question format provided to them is taken into account by the use of both, numerical and qualitative elicitation of expectations. While the numerical estimates are more demanding, in particular with respect to confidence intervals, they have the advantage of being comparable across participants. On the other hand qualitative

[10] For a discussion about when monetary incentives are useful see Camerer and Hogarth (1999). Other surveys that do not incentivize participants include the Michigan Survey of Consumers, the German Socioeconomic Panel and most surveys on investing behavior.

estimates may capture aspects of value and risk not comprised in the first two moments of a distribution. Interestingly, we find with one exception that only numerical expectations are relevant for actual financial risk taking decisions which is in contrast to the results of Weber et al. (2010) who establish a strong influence of qualitative expectations on allocations in the hypothetical investment task. We test whether the explanatory power of numerical expectations changes over to qualitative expectations if we drop numerical expectations from the regressions. In general, this is not the case and the impact of qualitative expectations remains weak.

An explanation thus has to consider the decision process in the hypothetical investment task compared to actual investing. First of all, we find our measures of financial risk taking only weakly correlated with the proportion of risky investment in the survey task which already hints at the two being different. In particular the changes of risk taking in the task and investors' portfolios are unrelated. We conjecture that the qualitative expectations are affective evaluations of the market situation, while the numerical estimates draw on more cognitive resources (cp. Kempf, Merkle, and Niessen, 2009). We would then expect these evaluations to be predictive for decisions that are made in the same "mode" of thinking.[11] If the actual investment decisions of investors are preceded by a more deliberate thought process than the allocations in the hypothetical task, this would at least partly explain the greater predictive power of numerical expectations for these decisions.

We also consider the time structure of expectations and trading, and throughout the chapter we opt for an approach that tries to explain changes in investing behavior by contemporaneous changes in expectations. Another possibility would be that investors need some time to react on changes in expectations, for example because of inertia. Indeed, when we use lagged variables many of the described relationships between expectations and investment behavior can still be observed. However, the

[11] Support for this dual-process theories of information processing and decision making can be found e.g. in Kahneman (2003).

effects are in general equal or weaker than for contemporaneous expectations. We thus conclude that investors tend to implement their beliefs in a timely manner.

As a complement to our research, the investor survey of Hoffmann et al. (2011) has an overlap of seven month with our data. Elicited expectations and portfolio characteristics show some similarities: For instance return expectations of Dutch investors also rise from September to December 2008 and further to March 2009, and trading and buying activity increases in response to the crisis. Similar to us Hoffmann et al. (2011) find that median portfolio volatility is higher than market volatility and closely tracks the market index. However, there are some differences as well, e.g. risk perceptions fall gradually after a peak in September 2008, while in our data they rise and then stay on a high level until March 2009. This might be due to the different wording of the question, which in Hoffmann et al. (2011) refers to current risk perception while our approach is more forward looking. Nevertheless taken together the findings suggest that there exist some more general properties in expectations of private investors that are not limited to a particular dataset.

In a regression of buy-sell ratios on beliefs and preferences, Hoffmann et al. (2011) find very different results. They use qualitative measures of expectations—which in our case remained insignificant—and demonstrate a negative effect of return expectations (levels and changes) and a positive effect of risk perception (levels and changes). This is inconsistent with financial theory and we are unable to confirm this result. However, it contributes to our impression that immediate trading behavior is hard to predict from elicited beliefs. For portfolio volatility both datasets share the intuitive positive result for risk tolerance and the insignificant result for qualitative return expectations. However, Hoffmann et al. (2011) identify a positive effect of risk perception on portfolio volatility. While this might again be a result of the different measurement, our findings for numerical risk and return expectations strongly point in an opposite direction.

2.6 Conclusion

We investigate the functional relationship between beliefs and preferences of investors and their trading behavior. While we are still far from suggesting a definite functional form for equation 2.1, our findings are a first step to improve the understanding of this complicated but fundamental relationship. We provide evidence that expectations are relevant for risk taking of investors, and that they are used in a predominantly rational and intuitive way.

Higher return expectations lead to increased risk taking in terms of volatility among investors, while higher risk expectations have the opposite effect. Even more, changes in portfolio risk are predicted by contemporaneous changes in return and risk expectations. We find evidence that investors counteract changes in market volatility by reducing their portfolio volatility relative to the market. In general the best fit of our model is achieved for long-term portfolio volatility. Changes in short-term portfolio volatility and changes in portfolio beta are less well or not at all predicted by changes in expectations. This relates directly to the question how private investors manage their portfolio risk and which risk measure is closest to their subjective experience of risk. As long-term volatility measures react strongest to investor expectations, we take this as tentative evidence that they are a good proxy for experienced risk.

Expectations have less predictive power for immediate trading activity of investors. We do not find robust patterns for direction of trade and selling behavior, with the exception of return expectations, which have a positive effect on buying activity. However, we also observe some odd or confusing results for immediate trading behavior. We conclude that trading is often too noisy, influenced by liquidity and other exogenous trading motives. We are able to identify some influences such as realized returns or price levels.

Taken together our results suggest that financial theory may overstate the role of return and risk expectations for trading behavior. Private investors take their expectations into account to determine whether to buy or sell and whether to increase or decrease risk taking. But at the same time they rely on various other factors which to describe and identify is beyond the scope of this chapter. This finding creates leeway for alternative explanations of trading behavior provided by the young field of household finance (Campbell, 2006).

Chapter 3

The Beliefs of Others — Naive Realism and Investment Decisions

3.1 Introduction

In financial markets where prices are determined by the interaction of market participants, the expectations of investors are an important factor. We are interested in investors' beliefs about the expectations of others, commonly referred to as *second-order* beliefs. Investors may find second-order beliefs very useful to gain a more complete understanding of financial market movements. Compared to (for example) weather forecasters, they need to estimate a combination of fundamental values and others' beliefs regarding these values. While the weather remains unimpressed by the convictions of forecasters, investors do influence market prices by their opinions and actions.

This view of financial markets is closely related to the economic idea of the market as a beauty contest (Keynes, 1936) or voting mechanism (Graham and Dodd, 1934).

In these models price movements come about by changes in investor beliefs and expectations that are not necessarily induced by changes in fundamentals. Investors instead try to gauge what other market participants are thinking by forming higher-order beliefs. The modern finance expression for these aggregated beliefs in the investor population is *investor sentiment* (Barberis, Shleifer, and Vishny, 1998); and second-order beliefs can thus be interpreted as personal estimate of current investor sentiment. They can lead to an adjustment of own beliefs (as per Bayesian updating) and to expectations becoming more homogeneous in markets. However, they can differ greatly from actual investor sentiment if the underlying judgments of others' beliefs are systematically biased and thus inaccurate.

The objective of this chapter is to analyze the role of second-order beliefs in investment decisions. The main question in this context is whether investors find second-order beliefs informative beyond their own return and risk expectations. We furthermore assess the formation and accuracy of investors' second-order beliefs, their relationship to own (first-order) expectations, and their sensitivity to different market environments. We test for investors' susceptibility to commonly observed biases in second-order beliefs and investigate the repercussions of these biases on their investing.

To these ends we collect first and second-order return expectations in a longitudinal panel survey of private investors. The participants represent an especially interesting group as they are affluent self-directed online investors who frequently trade in stocks and other securities. The survey covers September 2008 to December 2009, a time period of widely varied returns in stock markets, which allows us to track investor expectations and second-order beliefs through different market environments. The survey is conducted every three months, and also elicits financial market views and psychological variables.

We reveal that second-order beliefs are indeed meaningful to investors. We analyze responses to an investment question which asks participants to divide a fixed amount between the stock-market and a riskless asset. The relationship between second-order beliefs and the amount of risky investment is positive, meaning that participants tend to invest more when they perceive others to be more optimistic. This suggests that they view financial markets as operating like a beauty contest or being positively influenced by investor sentiment.

Accurately estimating aggregate beliefs of others is likely to be difficult however, and this may render their use problematic. We first compare investors' first-order and second-order beliefs and analyze divergences between the two. For all survey rounds the majority of investors in our panel remain optimistic about the quarterly return prospects of the stock market. However, they believe their peers are far less sanguine, and thus hold inaccurate second-order beliefs. A test of individual accuracy of second-order beliefs reveals that investors' estimates of others' beliefs are mostly worse than a random guess.

To explain biases in estimating others' expectations we draw on the psychological paradigm of *naive realism* (Ichheiser, 1943; Ross and Ward, 1996), which holds that people take their own view on the world for objective reality. It provides a common framework for several well-known patterns in second-order beliefs, among them the false consensus effect (Ross, Greene, and House, 1977), which describes the tendency of people to overestimate the commonness of their own beliefs or behaviors, and the bias blind spot (Pronin, Lin, and Ross, 2002), which refers to people's unawareness of own judgmental biases.

We detect both biases in second-order beliefs of investors in the panel. In general participants perceive their own position as relatively more common in the investor population. Compared to investors with contrary first-order beliefs, they estimate a greater fraction of people to share their own expectations indicating a false consensus effect. As individuals' second-order beliefs may be driven by believing others are simplistically effected by negative or positive market environments, we check whether second-order beliefs are influenced by sentiment. We indeed find that the false consensus effect is especially pronounced for participants holding expectations in line with current market sentiment.

The bias blind spot is evident in investors' belief that their own expectations are based on more rational considerations than those of their peers. Participants also think that a surprisingly large fraction of others hold very unlikely expectations. These effects are stronger for investors holding beliefs contrary to current market sentiment. This subgroup of investors believe themselves to be in the minority and it would be a bold decision to stand against a well-informed crowd. Assuming others are biased allows them to maintain their opinion contrary to (perceived) sentiment.

We finally test for an interaction of the naive realism biases with participants' use of second-order beliefs in investing. We find that those investors who assume others to be biased incorporate second-order beliefs much less in their decision. Those who assume a consensus rely more on the estimated beliefs of others as they confirm their own views. Market views and financial literacy also play a role in the way investors incorporate second-order beliefs in their decision process.

In summary this chapter contributes to the extant literature in several ways. We empirically investigate the neglected field of second-order beliefs in financial economics and introduce a unique data set of longitudinally surveyed real investors. We demonstrate the presence of known psychological biases in an environment where natural incentives should motivate accurate predictions, especially for those who believe second-order beliefs are useful. We also develop new bias measures that allow us to describe the working of false consensus effect and bias blind spot more precisely.

Finally we show the relevance of second-order beliefs and these effects for trading behavior.

Section 3.2 reviews economic and psychological literature on second-order beliefs and derives hypothesis for our analysis. Section 3.3 introduces the data set and the questions asked in the survey. Section 3.4 presents and discusses the results, a final section concludes.

3.2 Second-order beliefs in finance and psychology

3.2.1 The beauty contest, investor sentiment, and belief heterogeneity

The most well-known account of second-order beliefs in economics is Keynes' (1936) metaphor of the stock-market as a beauty contest. It implies that security prices react to market sentiment rather than to fundamentals, and that stock market predictions should be derived from the perceived beliefs of others. Keynes likewise argued that personal return expectations in the stock market will not materialize unless they are shared by a significant proportion of investors. In the words of Graham and Dodd (1934) the market works as a "voting machine" rather than a "weighing machine".

Lately a model-theoretic discussion has evolved around the role of higher-order beliefs in asset pricing. Allen, Morris, and Shin (2006) strongly suggest prices to reflect average opinion and find higher-order beliefs useful in forecasting prices. Without assuming any irrationality, their model can explain financial market phenomena such as bubbles and underreaction. While others come to slightly different conclusions (cp. Banerjee, Kaniel, and Kremer, 2009; Makarov and Rytchkov, 2006), the importance of higher-order beliefs is generally acknowledged in these theoretic models. Consequently investors need to consider the opinions and expectations of other market participants as well as fundamentals of companies.

The literature on investor sentiment builds on this idea. Barberis et al. (1998) define investor sentiment broadly as "how investors form beliefs" (p. 308) and model overreaction and underreaction as a consequence of sentiment. Baker and Wurgler (2006) empirically find that investor sentiment influences the cross-section of stock returns. High sentiment stocks earn relatively low subsequent returns, while the result reverses for low sentiment stocks. This suggests that investor sentiment should be used as a contrarian indicator. However, due to noise trader risk it does not offer direct arbitrage opportunities. Chan and Fong (2004) investigate where and how the effects of such sentimental price pressures are likely to be observed and indeed they find that the effect is strongest for small and less liquid stocks (where arbitrage may be limited).

The implications of second-order beliefs for own investing behavior are thus not straightforward. While a beauty contest model would suggest aligning your investment with the crowd by selecting securities that one believes are or will be commonly regarded as attractive, objective investor sentiment research finds that high sentiment stocks subsequently perform poorly. Investor sentiment could thus also be interpreted as a contrarian indicator.

The most direct and intuitive way to measure sentiment is to survey investors. The return expectations we collect may therefore be interpreted as indicating current sentiment among the panelists. Fisher and Statman (2000) show that sentiment from such surveys is negatively related to future stock returns. This supports the view of sentiment as a contrarian indicator and stands in contrast to the beauty contest model, where the winning strategy is to pick the stocks that are commonly regarded most attractive.

In either case, individuals may believe knowing investor sentiment is helpful in predicting future returns. Experimental evidence confirms that expectations of investors can be informative beyond fundamentals and historical price data (Haruvy, Lahav, and Noussair, 2007). In order to use current market sentiment, investors must form second-order beliefs. We thus can see the practical relevance of second-order belief formation in their close connection to investor sentiment and the beauty contest view of financial market functioning.

However, we do not rely on investor sentiment to actually affect observed prices. It is sufficient that investors *perceive* market functioning as a beauty contest or as influenced by investor. Ongoing interest in analysts' recommendations and opinions of market pundits reflect this view. Brown and Cliff (2004, p.2) conclude from similar evidence that "market watchers and participants seem to believe in sentiment". Survey studies show that investors consider financial markets to be influenced by the opinion and expectation of other market participants. For example, Fisher and Statman (2002) find that investors believe markets may continue to rise due to sentiment driving prices, even though they recognize it is (fundamentally) already overvalued. Some participants in our panel confirm these results when asked for their view of how stock markets work (see section 4).

Social interaction between stock market participants is another facet of second-order beliefs in investing behavior. Hong, Kubik, and Stein (2004), and Brown, Ivković, Smith, and Weisbenner (2008) show that individuals' propensity to invest in stocks is influenced by the stock market participation of their peers. They identify sociability and word-of-mouth effects as drivers of these results. Kaustia and Knüpfer (2010) find a direct influence of neighbors' stock returns on investing behavior of individuals. Sources of these interaction maybe direct communication, verbal accounts, or beliefs about others. In financial markets (and elsewhere) people care for what other people do and what other people think. Whenever these actions or opinions remain unobservable, second-order beliefs play a vital role in social interaction.

Our first hypothesis thus is that investors believe second-order beliefs are relevant for their investment decisions. The direction of this influence can either be confirmatory, i.e. more positive second-order beliefs induce higher investment levels, which would reflect a belief in a beauty contest or voting mechanism. Or it could be negative, with more positive second-order beliefs resulting in lower investment levels if investors follow a contrarian investment strategy. The strength and direction of the influence should be mediated by investors' views on market functioning and by biases in estimating the expectations of others. The former determines whether second-order beliefs play a role within the convictions of investors. The latter affects the perception the individual investor has of others's expectations and will be discussed in detail in the next section.

H1: Investors use second-order beliefs in making their investment decision.

H1a: The strength of H1 depends on how investors view market functioning and on the judgmental biases they make.

To estimate second-order beliefs accurately and to use them in decision making correctly, it is important to understand what causes differences in beliefs between individuals. Traditional finance literature generally links beliefs to information (e.g. Fama, 1970). According to Black (1986, p.531) "differences in beliefs must derive ultimately from differences in information." In this view, the beliefs of others simply reflect different information sets and meta-thinking is reduced to an attempt to infer others' information. Behavioral finance adds other sources of interpersonal differences to the picture. Perception, attention, memory, cognitive biases and limitations, and emotions have all been invoked to explain heterogeneity in beliefs.[1]

Within the behavioral paradigm stock market expectations depend on participants' recall of previous events, the way they perceive new information, and on the attention they pay to the task at hand. Expectations further depend on participants' ability to process the various input factors sensibly, their proneness to judgmental biases, and the time and resources they spend on the task. Given these many and mostly unobservable factors which contribute to belief heterogeneity in financial markets, we consider it as unlikely that investors are able to correctly estimate beliefs of others.

H2: Second-order beliefs describe actual beliefs of other investors inaccurately.

3.2.2 Naive realism and judgmental biases

Financial economics' insights into second-order beliefs are useful to understand their role in stock markets, but lack predictions of how these beliefs are generated. Social psychology has argued that people perceive the world through a lense of naive realism — they *believe* that they experience and observe entities, events, and people in an objective and unbiased way (Ichheiser, 1943). Three basic tenets of naive realism were derived by Ross and Ward (1996): First is one's own felt objectivity and unmediated, factual interpretation of available evidence. As a consequence people secondly assert that others will share their beliefs and opinions, if they only analyze the situation in a reasonable manner. The third tenet of naive realism describes the reaction to disagreements in opinion. Given that the objectivity of one's own position is not contested, people often conclude that those who disagree with them either lack

[1] Among the many existing studies consider exemplarily for attention: Barber and Odean (2008), memory: Biais and Weber (2009), perception: E.U.Weber et al. (2005), cognitive biases and limitations: De Bondt and Thaler (1985), and emotions: Kempf et al. (2009).

information, are unable or unwilling to process the facts at hand rationally, or are motivationally biased.

This framework leads to two specific biases we will examine in detail. As a result of the second tenet people will perceive their own judgments to be more common and appropriate than alternative responses. This false consensus effect manifests itself regarding personal traits, preferences, characteristics, and expectations of others (Ross et al., 1977). For example, adolescents who smoke provide higher estimates for the prevalence of smoking in the population than their non-smoking peers (Sherman, Presson, Chassin, Corty, and Olshavsky, 1983). In a meta-analysis of 115 items Mullen et al. (1985) show the pervasiveness and robustness of the effect.[2]

Causes of the false consensus effect broadly fall into four categories (Marks and Miller (1987)). The availability heuristic attributes it to the ease with which instances of (dis)similarity can be recalled. Given that people associate with other people of similar status, profession, and preferences, selective exposure leads to a biased assessment of the overall population. A second ingredient is the salience of one's own reasoning. Introspection emphasizes features supporting one's own position. Third, people tend to attribute their behavior and beliefs to situational rather than dispositional causes. Individuals assume that others will behave similarly in the same situation, neglecting differences in personality, tastes, and processing. Finally motivation plays a role as an existing consensus validates the correctness and appropriateness of an opinion. This way it bolsters self-esteem and perceived social support.

Despite the seemingly obvious applicability of this long established psychological bias to finance, we are not aware of many similar studies. Academic financial economists were shown to exhibit a false consensus effect when surveyed about the equity risk premium (Welch, 2000). Experimentally Hsee and E.U.Weber (1997), Faro

[2]It has been argued that rational Bayesian updating using first-order beliefs leads to similar results (Dawes, 1989; Dawes and Mulford, 1996). However, the strength of the false consensus effect mostly goes beyond the (limited) informational value of own judgment (Krueger and Clement, 1994).

and Rottenstreich (2006), and Borgsen and Weber (2008) investigate financial lotteries, where subjects have to predict risky choices of other participants. An integral component of such predictions are one's estimates of the risk tolerance of others. They find that a false consensus about risk tolerance and lottery choices exists, but can be moderated by different experimental designs. We anticipate a false consensus effect among the investor population when estimating the beliefs of other investors.

H3: Survey participants exhibit a false consensus effect in predicting future returns.

To gain a deeper understanding of the effect we will enhance the classic false consensus paradigm to analyze who is particularly prone to the effect. Usually the bias is defined in binary subgroups of the population (endorsers and non-endorsers of a statement, entrants and non-entrants of a financial lottery), meaning that *independent* of the position a person holds he or she will overstate the *relative* commonness of this position. However, in stock markets the prevalent market environment may give investors reason to believe in a specific consensus. Particularly those who submit return predictions in line with market sentiment will perceive this to be more of a consensus than those who hold contrarian views. We define current market sentiment by several standard variables including market trend, news, consumer confidence, and implied volatility.

H3a: The false consensus effect is mainly present among investors with expectations consistent with current market sentiment.

To better understand those investors who hold contrarian views, we now turn to the third tenet of naive realism. It seems unlikely that investors can believe that a majority of well-informed, rational market participants hold beliefs widely divergent from their own. This would undermine their own position for two reasons: first, there is the informational value of others' expectations. Second, if one believes that market sentiment influences asset prices (in the sense of a voting mechanism), one should also adjust one's own expectations towards others.

The naive realism model presents a solution for this dilemma — it posits that survey participants will regard their own stock market expectations as unbiased, and the opinions of those who disagree with them as biased. This asymmetry in the perception of bias has been coined the "bias blind spot" (Pronin et al., 2002). Blindness here refers to the inability to recognize potential bias in one's own judgments, and thus asserting bias in others (Kruger and Gilovich, 1999). In its self-serving capacity the bias blind spot has been linked to the better-than-average effect, the tendency of people to view themselves as better than average in various domains (Alicke and Govorun, 2005). But non-egotistical causes for the bias blind spot have been proposed as well. Observing discrepancies in opinions requires an explanation for how the differences have been generated. Regarding others as biased is a simple strategy to resolve this dissonance. A reliance on introspection partakes in and adds to this explanation. The search for traces of bias in oneself often proves fruitless as non-cognitive processes are involved in biased judgments (Pronin and Kugler, 2007). This introspection illusion keeps the self-image of an unmediated view on reality intact.

Menkhoff and Nikiforow (2009) extend the bias blind spot approach to finance. In an experimental study with fund managers, they show that participants see themselves as less prone to fall for financial biases such as home bias, herding, or the disposition effect. This result holds almost equally for endorsers or non-endorsers of behavioral finance.

H4: Investors are subject to a bias blind spot in generating expectations.

As noted before, the mental need to resolve dissenting opinions is more pronounced for those investors who regard their position as a minority position. We thus expect a stronger bias blind spot for these investors.

H4a: The bias blind spot is stronger for survey participants expressing a view contrary to current market sentiment.

Our hypotheses allow to study the judgmental biases through different market environments and to make predictions for their occurrence among different groups of investors. The dataset we use to test our hypotheses will be described in detail in the next section.

3.3 Dataset and survey questions

In collaboration with the Behavioural Finance team at Barclays Wealth, we conduct a panel survey of online self-directed investors at Barclays Stockbrokers, a UK direct brokerage provider. The first survey started in September 2008, shortly before what is in retrospect widely regarded as the climax of the financial crisis (the events around Lehman Brothers and AIG in the US, Northern Rock, HBOS, and other banks in the UK). Subsequent rounds occurred in three month intervals in December 2008 and throughout the year 2009. Within figure 2.1, which shows a chart of FTSE all-share index and the timing of the rounds, the first six rounds are relevant for this study. The survey covers the sharp stock market decline in late 2008 as well as the recovery in spring and summer 2009. The great variation in terms of realized returns resembles a remarkable natural experiment for studying reactions to market uncertainty, and is a strength of our dataset. It augments our testing of the role of investor sentiment and market phases, as we expect considerable changes for these variables throughout the survey period.

In the initial survey a stratified sample of the banks client base was invited via e-mail to participate in the online questionnaire (for details on the sampling procedure see Weber et al., 2010). In total 617 clients of the bank participated in at least one round, 200 have participated in at least 4 rounds, and 67 have participated in all 6 rounds. We have a minimum of 198 observations for each of the six rounds.

Investor surveys like ours have been pioneered by Lease et al. (1974) and were more recently used e.g. by Dorn and Huberman (2005), Glaser and Weber (2005), and Amromin and Sharpe (2009). Survey methodology has gained in importance and acceptance in finance, and has delivered notable evidence also in other subfields such as corporate finance (Graham and Harvey, 2001; Lins, Servaes, and Tufano, 2009).

Table 3.1: Descriptive statistics

Panel A – Demographics	n	Mean	Median	Std.Dev.	Min	Max
Age (in years)	613	51.4	53	12.9	21	84
Gender (male=1)	617	0.93	1	0.25	0	1
Couple (married or cohabiting=1)	616	0.74	1	0.44	0	1
Investment experience (in years)	197	19.6	20	10.3	1	41
Wealth (categories see below)	502	4.80	5	2.39	1	9
Income (categories see below)	494	3.88	4	1.80	1	8
Panel B – Market views	**n**	**Mean**	**Agree**	**Neutral**	**Disagree**	
Statement 1 (alpha)	154	5.37	135	8	11	
Statement 2 (unpredictable)	154	3.48	37	24	93	
Statement 3 (mispricings)	154	5.53	134	11	9	
Statement 4 (index best)	154	3.03	25	36	93	
Statement 5 (minority investing)	193	4.15	61	89	43	
Statement 6 (majority beliefs)	192	3.16	48	31	113	

Notes: Number of observations varies due round when questions were asked or refusals.

Wealth categories: (1) 0–10,000£ (2) 10,000–50,000£ (3) 50,000–100,000£ (4) 100,000–150,000£ (5) 150,000–250,000£ (6) 250,000–400,000£ (7) 400,000–600,000£ (8) 600,000–1,000,000£ (9) >1,000,000£.

Income categories: (1) 0–20,000£ (2) 20,000–30,000£ (3) 30,000–50,000£ (4) 50,000–75,000£ (5) 75,000–100,000£ (6) 100,000–150,000£ (7) 150,000–200,000£ (8) >200,000£.

One £ is approximately 1.60 $, average gross yearly income in the UK is about 30,000£.

Statements are evaluated on a seven-point scale, from 1 (strongly disagree) to 7 (strongly agree). Column “Agree” comprises 5–7, “Disagree” 1–3 on that scale. Full statements can be found in the text.

Panel A of table 3.1 shows demographic statistics of survey participants. The older, more affluent, and male-dominated sample does not reflect the general British population (for an explicit comparison consider again Weber et al., 2010). But it does represent typical investor populations found in other studies (cp. the examples mentioned above). In particular most investors in our sample are experienced and do well in a financial literacy test.[3] Investors report that they spend on average six hours a week on trading or researching potential investments. About a third states they have significant training in finance, economics, mathematics or statistics. We are thus confident that participants are not only able to understand and answer the questions meaningfully, but have also developed independent expectations and opinions about market prospects.

The focus for this study will be on the following questions:

1. We would like you to make three estimates of the return of the UK stock market (FTSE all-share) by the end of the next three month.

- *Your best estimate should be your best guess.*
- *Your high estimate should very rarely be lower than the actual outcome of the FTSE all-share (about once in 20 occasions)*
- *Your low estimate should very rarely be higher than the actual outcome of the FTSE all-share (about once in 20 occasions)*

Please enter your response as a percentage change.

2. Think carefully about the best estimate question above, and how other people in this survey will respond. What percentage of respondents to this survey do you think will give a response falling into each of the categories below?

- *Fall 10% or more.*

[3] We use four questions by van Rooij et al. (2011) and obtain on average 87.3% correct responses (while van Rooij et al. (2011) report 65.5% for average households).

- *Fall 3% to 10%.*
- *Stay about the same.*
- *Rise by 3% to 10%.*
- *Rise by 10% or more.*

Question one asks participants to forecast the three-month return of the UK stock market. They have to submit a best estimate as well as a high and a low estimate, which together yield a 90%-confidence interval. The question design is similar to Glaser and Weber (2007) and allows to calculate implicit expected volatility and miscalibration of investors. We use the best estimate to represent an investor's first-order belief about stock market return and the confidence interval as the range of outcomes he perceives as likely. Question two elicits second-order beliefs for the best estimate with survey respondents serving as the reference population. Second-order beliefs are to be stated in five intervals ranging from large losses to large gains. Due to space and time considerations question 2 was omitted in round 5 of the survey (September 2009).

In our investment task question participants have to allocate a hypothetical endowment between the UK stock market and a riskless asset. In offering the FTSE all-share for investment and in using a time horizon of three months the task corresponds to the previously estimated expectations and serves to analyze investment decisions in dependence of first and second-order beliefs. Apart from the answers to question one and two we use risk tolerance and subjective risk perception as controls. Further descriptions of these variables and the investment task can be found in Brooks et al. (2008) and Weber et al. (2010), who use the same dataset.

To reveal if individuals believe that market functioning reflects the influence of sentiment and therefore the value of second-order beliefs, we ask for participants' agreement to several statements:

1. *It is possible for skilled investors to earn above-average risk-adjusted returns.*
2. *Price movements are completely unpredictable.*
3. *There are often temporary mis-pricings of stocks in the market.*
4. *The best way to invest is to just buy and hold a market index.*
5. *When making investing decisions, it is better to be in the minority than the majority.*
6. *When many people are saying the same thing about markets, it makes me believe they are correct.*

The first four statements relate to the notion of efficient markets and its implications for investment behavior. However, as panel B of table 3.1 shows, survey participants strongly believe in the presence of mis-pricings, the predictability of price movements, and the possibility to earn above-average risk-adjusted returns. Accordingly, they do not regard index investing as the best way to invest. We did not ask directly how investors assume mis-pricings to come about, and whether the ability to earn above-average returns and predict prices are seen as a consequence. But the very definition of mis-pricings is that they occur by trading activity driving prices away from "correct" values. When trading activity is based on (first-order) beliefs of investors, second-order beliefs or investor sentiment are a natural source for predicting these price movements. We interpret investors' answers to statements 1–4 as indicating this view of market functioning.

The last two statements explore whether people prefer to be in the minority or majority with regard to investment decisions and opinions (see table 3.1). If investors found this irrelevant, they would likely opt for the middle category. 46% (statement 5) and 16% (statement 6) do so, with these numbers probably being inflated by those who are just unsure and by a general central tendency bias. The remaining participants hold mixed opinions, consistent with the two alternative interpretations of second-order beliefs: beauty contest and investor sentiment. If anything, we observe a tendency towards acting and thinking contrary to the majority. The empirical analysis will show whether these opinions become effective in the investment task.

3.4 Results

3.4.1 A comparison of first-order and second-order beliefs

We first briefly analyze investors' first-order expectations, as they serve as a benchmark for later results. Participants submitted best estimates for return of the FTSE all share index over the next three month. Table 3.2 shows descriptive statistics for these estimates across rounds of the survey. The mean estimate is around 2% in September 2008, rises to 5.4% in March 2009, and stays on a relatively high level until it drops off to 3.4% in December 2009. On a yearly basis this return expectations appear high, which can be either a sign of over-optimism (Weinstein, 1980; Taylor

Table 3.2: Expectations of investors

	n	Mean	Median	Std.Dev.	5% Perc.	95% Perc.	width of CI
Round 1 (Sep08)	479	1.99	2.00	7.88	-10.00	12.00	16.7
Round 2 (Dec08)	380	3.35	3.00	14.57	-15.00	20.00	23.4
Round 3 (Mar09)	223	5.42	5.00	12.84	-10.00	20.00	23.3
Round 4 (Jun09)	188	4.25	5.00	15.87	-10.00	15.00	29.2
Round 5 (Sep09)	217	5.81	5.00	19.95	-15.00	20.00	26.8
Round 6 (Dec09)	195	3.43	3.00	16.81	-10.00	15.00	26.7

Notes: The table states summary statistics for return expectations (best estimates) of investors in %. Width of CI (confidence interval) reports the average difference between high and low estimate in percentage points, inconsisten observations with low > high estimate are dropped.

and Brown, 1988) or of some investors misinterpreting the question in terms of annual values[4].

The standard deviation and inter-quartile range of expectations increases dramatically within the financial crisis, and remains high throughout 2009. We can look beyond between-subject measures, as we also asked each individual for confidence intervals regarding return expectation. The average width of confidence intervals mirrors the cross-sectional standard-deviation over time, becoming larger for later rounds of the panel. Participants seem to learn from the crisis that extreme outcomes are not as unlikely as they previously thought.

For the following analysis of first-order and second-order beliefs, we define two groups of investors: optimists and pessimists. Optimists are participants who submit return expectations falling into the top two categories of the scale used for question 2, i.e. who expect the FTSE to rise at least by 3%. Similarly pessimists represent the bottom two categories, which correspond to a fall of the FTSE by at least 3%. Those in between we define as neutral. Figure 3.1 shows the proportions of optimistic and pessimistic investors (solid lines). Through all survey rounds investors are predominantly optimistic, their proportion ranges between 45% (September 2008) and 64% (March 2009). The fraction of pessimists is relatively stable and never exceeds one quarter of the population.

Individual first-order beliefs are persistent over time. The Pearson correlation between estimates in consecutive survey rounds is 0.32 ($p < 0.001$), and correlations remain positive for rounds further apart in time. A transition matrix using all five bins of question 2 reveals a probability to stay in the same belief category of 39% compared to 20% if allocation was random. Only rarely do very optimistic investors turn very pessimistic within three month time.

[4] A comparison of three month and one year expectations (the latter where elicited for two rounds of the panel) reveals that most investors do differentiate between the time horizons. Moreover, as a potential mis-interpretation would occur across both own and others' expectations, it remains inconsequential for most of our analysis.

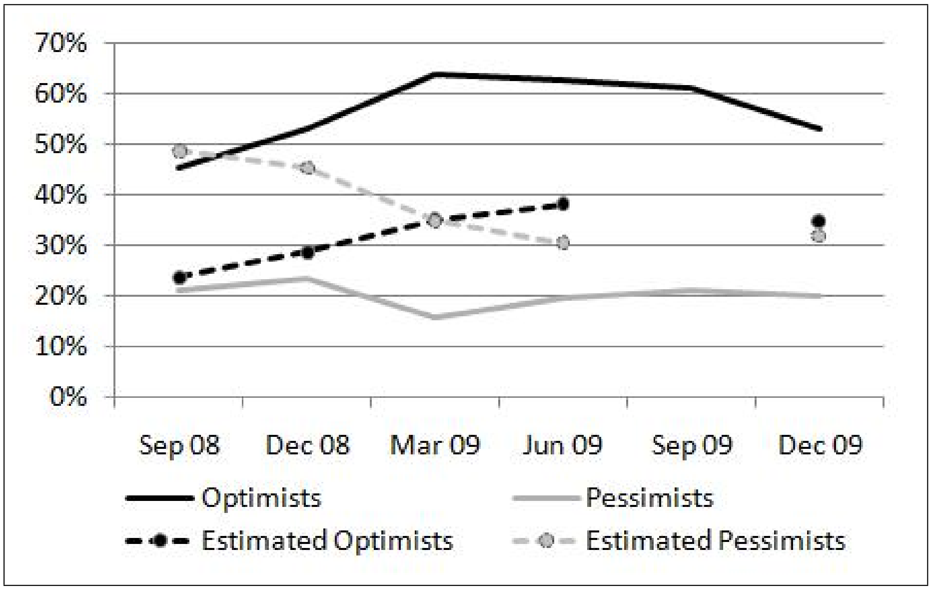


Figure 3.1: Actual and estimated investor expectations
Notes: The figure shows proportions of optimistic and pessimistic investors (solid lines) and the estimated proportions from second-order beliefs (dashed lines).

We begin our analysis of second-order beliefs with a simple illustration of how individuals responded to the question. Given the difficulty of the task, participants may use the question format to cue for reasonable responses or express ignorance (Schwarz, 1999). If investors do not have strong opinions about beliefs of others or attempt to communicate ignorance, we would likely see a lack of directionality and excessive use of the middle category (central tendency bias) in their estimates.

We find that only a small percentage of respondents expresses symmetry in second-order beliefs (10.7%). Table 3.3 shows that the middle category (-3% to $+3\%$) is slightly over-represented, but taking into account that average historical (three-month) return falls into this category one may have expected even greater proportions allocated to this category. A positive correlation between symmetry of responses and estimated proportion in middle category confirms that these two facets of responses may indeed be related to ignorance or lack of opinion. However, for the majority of

Table 3.3: Average second-order beliefs

	second-order beliefs					differences second - first order				
Round	1	2	3	4	6	1	2	3	4	6
<= -10%	23.5	23.0	13.2	12.5	13.4	14.5***	6.4***	4.7***	0.3	3.2***
-10% to -3%	25.4	22.5	21.6	17.9	18.5	13.3***	15.7***	14.4***	10.4***	8.8***
-3% to 3%	27.3	25.7	30.1	31.4	33.2	-6.1***	2.6***	9.9***	13.8***	6.5***
3% to 10%	16.9	17.8	23.2	26.2	23.9	-14.4***	-10.7***	-10.9***	-16.9***	-12.5***
>= 10%	6.9	11.0	11.8	12.0	10.9	-7.3***	-14.0***	-18.2***	-7.7***	-6.0***
Exp. return	-3.0	-2.1	-0.1	0.4	0.0	-5.0***	-5.5***	-5.5***	-3.8***	-3.4***

Notes: Estimated proportions of expectations of others in %, differences to first-order beliefs in %-points. Expected return for second-order beliefs is calculated according to the following conversion rule for categorical responses: -15%, -6.5%, 0%, 6.5% and 15% respectively for the five categories. Differences are significant at *10%-level, **5%-level, or ***1%-level.

participants we conclude that they believe they know something about the expectations of other investors.

Table 3.3 shows the average distribution of second-order beliefs for all rounds. In the first two rounds participants believe that a greater proportion of investors are pessimistic than optimistic (see also the dashed lines in figure 3.1). In round three the relation becomes balanced and turns around slightly for round 4 and 6.[5] It is clear that participants assume other investors are much more pessimistic than they actually are. Relative to actual expectations, second-order beliefs are too pessimistic for all rounds. Investors think that between 10 and 28%-points more pessimists are in the sample than there are, while they underestimate the fraction of optimists by 18 to 29%-points. All but one of the differences between first-order and second-order beliefs are highly significant.

[5]Round 5 is excluded from further analysis as it did not contain the second-order belief estimation.

If one calculates average expected return from categorical second-order beliefs according to the conversion rule stated in table 3.3, it becomes apparent that investors believe their peers to expect on average negative or zero returns. This is in contrast to previously reported first-order beliefs which were positive between 2% and 5.8%.

3.4.2 Investment decisions

While investors' second-order beliefs might be interesting in their own right, their relevance depends on their consequences for investing decisions. We now study market timing investment decisions in the investment task, which was part of the panel survey in each round. This task asks investors to divide a hypothetical £100,000 between the UK stock market (represented by FTSE all-share) and a riskless investment.

Weber et al. (2010) show that risk taking behavior in this investment task is driven by risk tolerance, risk perception and return expectations. We thus take these variables as given and include them as controls in our regression. Risk tolerance is measured by a psychometric risk-tolerance score (see Brooks et al., 2008). For risk perception we take subjective risk expectation expressed on a seven-point scale. Return expectations are three month best estimates for the FTSE all-share (first-order beliefs). As main variable of interest we consider second-order beliefs represented by expected return calculated from submitted belief distributions (using the conversion rule introduced before).

We employ a panel tobit regressions as the proportion of risky investment is censored on both sides. We use a model with individual random effects and add round dummies to account for round specific effects. Table 3.4 shows in column one the result of the baseline regression. Risk tolerance, return expectation, and risk perception have the expected influence on risk taking. The higher risk tolerance, the higher the proportion invested in the risky investment, and the higher risk expectation, the lower

Table 3.4: Investment behavior

Risk taking behavior	(1)	(2)	(3)	(4)	(5)	(6)
Risk tolerance	0.026***	0.025***	0.026***	0.026***	0.023***	0.026***
Return expectation	0.208***	0.207***	0.188***	0.207***	0.194***	0.168***
Risk expectation	-0.043***	-0.041***	-0.041***	-0.043***	-0.046***	-0.040***
Second-order beliefs (SOB)	0.660***	0.828***	0.320	0.935***	1.171 ***	0.719*
Bias Blind Spot (BBS)		0.001				0.002
Interaction BBS*SOB		-0.310				-0.269
Relative FCE			0.010			0.018
Interaction FCE*SOB			0.633*			0.834*
Financial Literacy (FL)				0.014		0.011
Interaction FL*SOB				-0.571		-0.790*
Market view (MV)					-0.052	...
Interaction MV*SOB					-0.583	...
Round 2	-0.036**	-0.034*	-0.035**	-0.034*	-0.053*	-0.030*
Round 3	-0.110***	-0.112***	-0.107***	-0.108***	-0.129***	-0.106***
Round 4	-0.070***	-0.077***	-0.067***	-0.067***	-1.110***	-0.069***
Round 6	-0.043*	-0.032	-0.040*	-0.039*	-0.067**	-0.024
Constant	0.695***	0.691***	0.679***	0.682***	0.785***	0.656***
n	1430	1336	1430	1427	600	1333

Notes: The table shows coefficients of a panel tobit regression with random effects. Dependent variable is percentage of 100,000£ invested in UK stock-market. Risk tolerance is survey based risk tolerance score, return and risk expectations are investors first-order beliefs. Second-order beliefs is the estimated return expectation of other investors calculated from second-order beliefs. Bias blind spot is the proportion outside confidence intervals as a dummy variable (median split). Relative false consensus error is the individual FCE as a dummy variable (split at 0). Financial literacy is one for investors answering all four financial literacy questions correct, zero otherwise. Market view is first principal component of opinion expressed for questions 1-4 (see table 3.1) as a dummy. BBS, FCE, MV, and financial literacy are interacted with SOB. Round dummies control for round effects (round 5 is excluded as SOB were not surveyed in that round). Coefficients are significant at *10%-level, **5%-level, or ***1%-level.

is risky investment. Return expectations positively influence the amount of risky investment. These factors are stable across all specifications we will discuss. The round dummies show that compared to round 1 in all subsequent rounds on average investors allocate a smaller amount to the risky investment, regardless of their expectations.

Beyond first-order beliefs second-order beliefs also impact the proportion invested in the stock market. The coefficient for the mean of second-order beliefs is positive and significant. This means participants consider perceived aggregated expectations of others as influential when making investment decisions. Interestingly the impact of second-order beliefs on investment behavior exceeds in magnitude the effect of first-order beliefs. The two are comparable as both are expressed in percent. With the necessary caution, given that the coefficients in a tobit model express the effect of the independent variables on the latent variable, an increase of 10%-points in second-order beliefs result in about 6.6%-points more risk taking in the investment task. The result supports hypothesis H1, the more optimistic investors perceive others to be, the higher is their own investment in the stock market.

The positive coefficient for second-order beliefs is in line with the notion of a beauty contest. Participants invest more, when they believe others are holding positive expectations and are thus likely to invest more as well. In contrast investor sentiment literature suggests contrarian behavior, which we do not find in the data on average. While in the statements about market views (see section 3) investors were split between favoring the majority or minority position, their investment decisions clearly speak for acting with the majority.

As apparently investors rely on their second-order beliefs in investing, the question how investors generate these beliefs, how accurate they are, and which psychological biases are involved comes to the fore. Once we have discussed these issues in detail we will return to table 3.4 and the evaluation of H1a.

3.4.3 Accuracy of investors' estimates

Already the large discrepancies between first-order and second-order beliefs presented in section 4.2 raise concerns about the accuracy of second-order beliefs. Especially for the first rounds of the panel investors perceived others to be much more pessimistic than they actually were. We now turn from the summary statistics to the individual accuracy of participants.

We employ for each investor the sum of absolute errors between estimated second-order and actual belief distribution to look at individual level accuracy. The error measure $\delta = \sum_{i=1}^{5} |\hat{p}_i - p_i|$ is calculated over the five categories, a δ of zero conveys the investor estimated the distribution perfectly, a δ of two corresponds to the maximal possible error. Table 3.5 displays the average error of participants, as well as the error produced by a simple guess of a uniform or normal distribution. Although there is great heterogeneity in accuracy only about a quarter of investors submit estimates that are more precise than a benchmark of a naive random guess. We also test whether financial literacy helps in predicting the beliefs of others. The error for financially

Table 3.5: Accuracy in estimation

Absolute error	Sep 2008	Dec 2008	Mar 2009	Jun 2009	Dec 2009
Investor panel (mean)	0.72	0.70	0.70	0.66	0.57
Investor panel (10 perc.)	0.31	0.36	0.35	0.30	0.28
Investor panel (90 perc.)	1.18	1.13	1.08	1.01	0.87
Uniform distribution	0.49	0.33	0.49	0.46	0.46
Normal distribution	0.19	0.38	0.42	0.47	0.24
Better than uniform distr.	28%	6%	29%	26%	29%

Notes: Sum of absolute error $\delta = \sum_{i=1}^{5} |\hat{p}_i - p_i|$ in estimates of investors (mean and 10th and 90th percentile). Uniform distribution assumes equal proportions per category, normal distribution uses mean and stdev. of historical returns. Better than uniform distr. is the fraction of investors more accurate than a uniform distribution.

literate participants is on average 0.06 lower compared to the remaining participants ($p < 0.001$), but they are still far less accurate than the random benchmark.

We conclude that second-order beliefs of investors are a poor representation of first-order beliefs and thus confirm hypothesis H2, which stated that investors are inaccurate in estimating financial market expectations of others. Moreover their bias is systematically negative, as they hold too pessimistic beliefs about others' expectations. We will now continue by exploring the effects responsible for this bias.

3.4.4 False consensus effect

In most studies false consensus is demonstrated for binary choices between alternative judgments or behaviors. To apply the classic false consensus paradigm we therefore use the above defined groups of optimists and pessimist (leaving out neutral expectations). Table 3.6 shows how optimists and pessimist evaluate the beliefs of other investors. Each group thinks that their own expectations are shared by a relatively greater proportion of the population. Thus there is a positive difference between judgments of optimists and pessimists when the fraction of optimists is concerned and vice versa.[6]

This result is confirmed by positive correlations between own expectations and the mean of second-order beliefs. Independent of the way of calculation, using either numerical or categorical expectations, correlations are between 0.12 and 0.47 for the individual rounds ($p < 0.01$). The more positive participants' own view, the more positive they think the common evaluation of financial market prospects is. Krueger and Clement (1994) suggest another measure for a "truly" false consensus effect (TFCE), which is the correlation between the estimation error (estimated – actual beliefs) and the own position. The values for TFCE are reported in the last column of table 3.6. A positive correlation suggests the presence of a false consensus effect, which

[6] Extending the analysis to all five categories of expectation, in general endorsers of a certain opinion submit significantly higher estimates for the prevalence of this opinion in the population.

Table 3.6: False consensus effect

		optimists	pessimists	difference	p-value	TFCE
Sep 2008	proportion optimistic	29.4%	16.7%	12.6	< 0.001	0.33***
n = 319	proportion pessimistic	43.4%	60.0%	-16.7	< 0.001	0.33***
Dec 2008	proportion optimistic	35.8%	20.3%	15.5	< 0.001	0.38***
n = 283	proportion pessimistic	39.8%	60.2%	-20.4	< 0.001	0.42***
Mar 2009	proportion optimistic	39.9%	29.6%	10.3	< 0.01	0.24***
n = 176	proportion pessimistic	31.3%	43.3%	-11.9	< 0.001	0.29***
Jun 2009	proportion optimistic	43.7%	28.0%	15.7	< 0.001	0.39***
n = 152	proportion pessimistic	25.9%	46.1%	-20.2	< 0.001	0.47***
Dec 2009	proportion optimistic	38.3%	29.8%	8.5	< 0.01	0.22***
n = 142	proportion pessimistic	28.3%	42.8%	-14.5	< 0.001	0.40***

Notes: The table shows proportions of optimistic and pessimistic investors as estimated by optimists and pessimists, differences between the two groups and p-values of two-sample t-tests. TFCE (true false consensus effect, Krueger and Clement (1994)) is the correlation between estimation error and own position. Number of observations is after exclusion of participants with neutral expectation. Correlations are significantly different from 0 at *10%-level, **5%-level, or ***1%-level.

is the case for all ten prediction items. Collected evidence thus supports hypothesis H3 that a false consensus effect is present among investors.

Taking round 1 as an example figure 3.2 shows average second-order beliefs for strong pessimists (return expectation $< -10\%$) and strong optimists. Strong pessimists estimate about 40% of the population to be as pessimistic as they are, and optimism is an unlikely opinion from their point of view. Although less extreme, strong optimists share the pessimistic second-order beliefs and perceive themselves to be in the minority. While we find a relative false consensus effect for both groups only pessimists seem to believe in a real (absolute) consensus for their expectations. Moreover, this consensus is not backed by actual first-order beliefs which were mostly

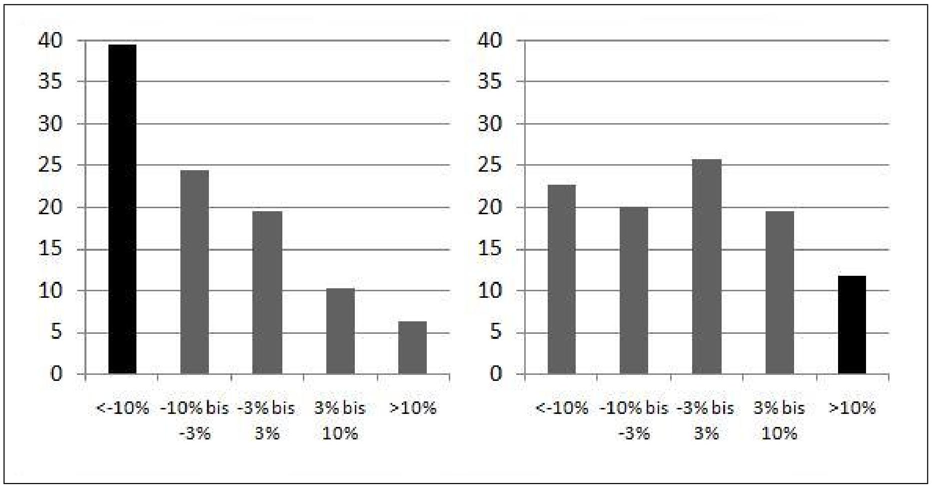


Figure 3.2: Second-order beliefs of pessimists vs. optimist

Notes: The left figure shows average second-order beliefs of strong pessimists, the right figure average second-order beliefs of strong optimists, both for survey round 1. The estimates for the own belief category are highlighted.

optimistic. We thus observe a consensus which is false in two ways, in absolute and relative terms.

To analyze this interesting feature of second-order beliefs, we have to depart from classic false consensus literature which usually defines the effect only as a relative bias (Ross et al., 1977). The argument of the classic paradigm is that it does not make sense to compare absolute second-order estimates for the own position, as these are influenced by actual consensus[7]. Given the low accuracy in estimating actual consensus we observe, we can treat participants as ignorant with respect to the estimation task. We define the absolute consensus error as the difference between second-order beliefs and actual beliefs in the population for one's own belief category, i.e. the fraction of participants an investor believes to share his opinion minus the fraction who actually does. This corresponds to the left panel of table 3.3 conditional on own first-order beliefs. On average this absolute false consensus amounts to 1.42 percentage points. It is significantly greater 0 ($p < 0.01$) and thus further supports H3. Its rather low

[7]For example in the case of smoking prevalence (Sherman et al., 1983), both smokers and non-smokers know that smoking is a minority activity. It follows naturally that non-smokers submit higher estimates for non-smoking than smokers do for smoking.

magnitude is due to the considerable number of participants we already referred to, who underestimate absolute consensus for their position.

3.4.5 Bias blind spot

We now turn to how individuals understand and explain beliefs divergent from their own. Some investors hold return expectations that are 12 or 13%-points apart from what they believe other investors are expecting. To justify this difference it is not sufficient to assume different information. If beliefs of others revealed information, a rational investor would be required to incorporate this information and the gap in beliefs would narrow or close. When investors disregard the beliefs of others however, it is likely that they suffer from a bias blind spot. They perceive other investors as biased and themselves as unbiased, which renders any adjustment of own opinion unnecessary.

Our approach to detect the bias blind spot follows Pronin, Gilovich, and Ross (2004). In round 6 of our survey respondents are asked to what degree their own expectations, and the expectations of those who disagree with them, have been influenced by a number of factors. Of the factors some represent normative or objective considerations (investing expertise, evaluation of economic conditions) and others nonnormative or biasing considerations (emotions, own recent performance). Typically own beliefs are attributed to objective thoughts, while beliefs of those who disagree are attributed to bias.

Table 3.7 shows that participants perceive economic conditions as significantly more important for their own expectations than for those of others. The opposite pattern can be observed for emotions and own recent performance. Investors believe they personally rely more on valid cues, while others are influenced more by biasing factors. In this capacity the bias blind spot is related to the better-than-average effect,

Table 3.7: Belief attribution

Influencing factor	Own beliefs	Others' beliefs	Difference	p-value
Economic conditions	5.23	4.69	0.54	< 0.001
Investing expertise	4.55	4.69	-0.14	0.24
News and media	4.72	5.57	-0.85	< 0.001
Recent performance	4.13	5.14	-1.01	< 0.001
Emotions	3.88	5.18	-1.30	< 0.001

Notes: Estimated importance of factors for own beliefs and beliefs of others who disagree on a seven-point scale (1-7), differences, and p-values of one-sided t-tests.

the tendency of people to view themselves more favorable than others. Investing expertise is hence expected to yield a positive difference as well. However, we find that the result in this case is confounded by own perceived financial expertise. Only investors who rate themselves highly in investing skill attribute a stronger influence of expertise to their own judgments.

From the results of table 3.7 we construct a bias blind spot measure aggregating the differences across factors (aligned in sign). We do not include investing expertise due to its confoundedness, and news and media for containing both factual information and normatively irrelevant aspects. Since this measure is available only for one round of the survey we use another more indirect approach to confirm these results. Confidence intervals submitted by investors define a range of outcomes they perceive as likely. Reversely beliefs of others outside these confidence intervals are seen as improbable and (presumably) biased. For each investor we calculate the proportion of own second-order beliefs that falls outside own confidence intervals and interpret the result as a sign for a bias blind spot.

The first row of table 3.8 reports the results for both bias blind spot measures. On average the asymmetry in belief attribution to normative and biased factors amounts to 2.8 response categories. There is great heterogeneity in this bias blind spot assessment, but only about a quarter of the participants is unbiased. Investors further assume on average that 22.4% of their peers hold return expectations that fall outside own confidence intervals and are thus unlikely. This fraction is large given that an investor with more accurate confidence intervals (e.g. from implied volatility) and accurate second-order beliefs would exclude a mere 2.5% of others. Both measures show a general presence of a bias blind spot among investors and thus confirm hypothesis 4.

Table 3.8: Biases and market sentiment

	False Consenus Effect		Bias Blind Spot	
Measure	Absolute FCE	Relative FCE	Belief attribution	Outside CI
All participants	1.4	6.7	2.82	22.4
In line with sentiment	7.2	8.2	2.44	18.9
Contrary to sentiment	−5.5	6.6	4.08	27.4
Difference	12.7***	1.6*	−1.64***	−8.5***

Notes: Absolute FCE is second-order beliefs compared to actual beliefs in the population, *Relative FCE* is second-order beliefs compared to average second-order beliefs, both for the own belief category. *Belief attribution* aggregates differences in self-other belief attribution (see table 3.7). *Outside CI* displays proportions of second-order beliefs that fall outside own confidence interval (CI) in %. Subgroups are investors who hold beliefs in line with and contrary to current investor sentiment (neutral beliefs not considered). Differences are significant at *10%-level, **5%-level, or ***1%-level.

3.4.6 Market sentiment and biases

We now analyze subgroups of the investor population. It was pointed out that the bias blind spot helps to resolve cognitive dissonance when facing a majority of others who hold different opinions. In contrast the feeling of being in the minority might moderate or offset the false consensus effect. Both biases thus seem to depend on the prevalent market sentiment and to affect groups of investors differently.

To test this proposition we construct a sentiment indicator based on four sentiment measures that have been proposed in the literature. Brown and Cliff (2004) find that recent stock market returns are an important determinant of sentiment. We take the past three month return of the FTSE all share as it is the object of elicited expectations and the time span allows for non-overlapping observations. Another directly market-based sentiment measure is implied volatility, which has been linked to investor fear (Whaley, 2000). We consider implied volatility of the FTSE 100 represented by the FTSE VIX index. Media reports have a strong influence on market sentiment as well (Tetlock, 2007), in addition their impact on false consensus has been documented (Christen and Gunther, 2003). We use a Google news search to identify positive and negative stock market news in the month of the survey rounds (for details see table 3.9). Finally consumer confidence has been established as a sentiment measure (Fisher and Statman, 2003; Lemmon and Portniaguina, 2006). We get UK consumer confidence data from Nationwide.

Table 3.9 reports the four sentiment measures for the rounds of our survey. The average interitem correlation is high (0.71), Cronbach's alpha is 0.91. This confirms that

Table 3.9: Market sentiment

	Sep 2008	Dec 2008	Mar 2009	Jun 2009	Sep 2009	Dec 2009
Recent return	−12%	−25%	−8%	14%	18%	3%
News (pos.-neg.)	−3100	1700	400	3900	3800	3700
Implied volatility	29.5	44.7	36.4	27.4	21.5	21.1
Consumer confidence	53	49	47	64	77	71
Sentiment indicator	−1.47	−2.24	−1.45	1.26	2.37	1.55

Notes: Recent return is previous three month return of UK stock market. News is the net number of news items retrieved via a Google News search using the term stock market combined with positive keywords (rally, rise, boom, grow, gain, positive, hope) and negative keywords (crash, fall, drop, fear, worries, negative, bad). Implied volatility is based on the FTSE 100 volatility index of NYSE Euronext. Consumer confidence is the consumer confidence index of Nationwide Building Society. Sentiment indicator is the first principal component of these four factors.

the items share a common underlying construct. We perform a principal component analysis and take the first principal component as a sentiment indicator. To combine several measures to an indicator has proven useful in investor sentiment research (cp. Baker and Wurgler, 2006). The values of the sentiment indicator are shown in the bottom row of table 3.9. It captures 79% of the variance of the individual measures. The general picture that emerges is that sentiment is negative during the first half of our survey and positive in the second half.

As second-order beliefs can be interpreted as participants' estimates of investor sentiment, we check whether they are consistent with our sentiment indicator. Indeed the correlation between mean second-order beliefs and market sentiment is 0.72, much higher than between mean first-order beliefs and sentiment (0.44). Investors seem to believe that others are forming their expectations by simplistically taking over the prevalent market sentiment.

We can now divide the investor population into those holding expectations in line with investor sentiment and those who hold expectations contrary to sentiment. For the first three rounds of the survey pessimists are in line with market sentiment, for the last three rounds optimists are in this position. Each group should be particularly prone to the false consensus effect at different points in time. Table 3.8 shows that the absolute and relative false consensus error is larger for investors with expectations in line with market sentiment. They overestimate the absolute consensus for their position by 7.2 percentage points, and relative to other participants by 8.2 percentage points. Investors with expectations contrary to prevalent market sentiment still show a relative false consensus effect, but on absolute terms they even underestimate the consensus for their position. Differences between the two groups are significant and in the predicted direction.

Results for the two bias blind spot measures are reversed. Investors with expectations contrary to market sentiment attribute beliefs of others less to normative factors and more to biasing factors than their own beliefs. They also exclude a greater proportion of other investors from their own confidence intervals. Differences are again strongly significant. This is likely due to the fact that "contrarians" need to justify their own minority beliefs by assuming bias in others. The opposed directionality of false consensus and bias blind spot in dependence of market sentiment simultaneously supports H3a and H4a. The distinct pattern allows to predict the occurrence of judgmental biases in different market phases.

3.4.7 Consequences of psychological biases for investing

The important role of second-order beliefs for investing has been established before (see section 4.2). It should be clear that inaccuracy and biasedness of second-order beliefs provide a threat to sound decision making of investors. But also the false consensus effect and bias blind spot itself should alter the way second-order beliefs are used in investing, as expressed by H1a.

In a next step we thus interact second-order beliefs with the judgmental biases described in the previous sections. Of the two bias blind spot measures we take the bias blind spot from confidence intervals as it is available for all rounds. For the false consensus effect we take relative false consensus, i.e. the degree the prevalence of one's own position is overestimated compared to the average estimate of all participants. For both measures we create dummy variables and interact these dummies with expected return from second-order beliefs. The natural prediction is that investors, who perceive others as biased, will rely on second-order beliefs to a lesser extent than those who do not. In contrast a felt consensus is expected to increase the impact of second-order beliefs as it lends support to own expectations and reduces the ambiguity in the decision.

We perform a similar interaction for market views of investors by using the first principal component of their answers to statements 1-4 (see section 3). A higher value here signifies a view more in line with efficient markets, which would contradict a role of second-order beliefs in investing. Finally we consider financial literacy, as financially literate participants may trust more in their own expectations than in the expectations of others.

Colums (2) to (6) of table 3.4 present the results of the described regressions. Indeed the signs of the coefficients are as expected and the magnitude of the effects is large compared to the baseline regression. Investors who view the market as more efficient, perceive others as biased or have higher financial literacy rely less on second-order beliefs, while a perceived consensus strengthens the influence of second-order beliefs. The effects are stable to the inclusion of all interactions in the full model of column (6)[8]. However, we are reluctant to emphasize this result too much as the significance of the interactions is at best weak. Overall the directionality of the results provides some evidence that the judgmental processes involved when coming up with second-order beliefs impact the way second-order beliefs are used in investment decisions. We take this as tentative support for H1a.

3.5 Conclusion

Investors exhibit systematic errors in forming second-order beliefs, yet appear to use these biased estimates in making investment decisions. They regard themselves as objective in judging stock market prospects, and believe others either to agree with

[8]Market view is not considered in this regression as it overly reduces the number of observations

their views or to hold somewhat biased expectations. Still they only have a very vague idea of the beliefs of other investors. The accuracy in estimating second-order beliefs is on average worse than a random guess. Investors mostly hold relatively optimistic expectations, while believing others are strongly influenced by current sentiment. We consider two biases to explain this asymmetry, a false consensus effect and a bias blind spot. We show a strong false consensus effect for the participants in our study. Its significance exceeds typical findings in psychology (see Mullen et al., 1985), which might be due to the nature of our financial judgment task. While people often have an idea of actual consensus (e.g. for smoking prevalence) and only slightly over- or underestimate it in direction of their own behavior, it is more challenging to guess return expectations of others in a volatile market environment. It is thus likely that investors rely on the same evidence both for own judgments and second-order beliefs.

We find pronounced differences in false consensus conditional on own expectations. We develop absolute consensus as an additional measure, which compares estimates to actual expectations of investors. Only investors with views in concert with current market sentiment assume a majority of other investors to share their views. We argue that these differences are produced by the market environment in different rounds of our survey. A sentiment indicator based on return trend, implied volatility, news, and consumer confidence, makes a case for a negative stock market outlook in the first half of the survey and a positive one thereafter. Investors in line with this sentiment indicator easily gain the impression that most people agree with them.

Conversely investors standing against current sentiment underestimate the commonness of their responses. They see themselves as minority and remote from mean expectation. To justify their expectation against a perceived majority a mental excuse is needed. Consequently contrarian investors believe that the majority accept trends, news and events at face value and fall for the more obvious prediction of stock markets to follow current sentiment. This implicit bias blind spot becomes explicit when asked for plausible ranges of stock market outcomes. Investors exclude a substantial fraction of their beliefs about others' expectations from what they themselves perceive as likely. We interpret this as investors thinking of these people as biased, as otherwise they would hold more reasonable expectations.

A general tendency to assume that others are using less credible information sources is confirmed by the weights investors give to certain factors for their own judgment and the judgment of disagreeing others. Investors consider normative considerations such as economic conditions or investing expertise (when controlled for own expertise) as more important for their expectations than others' expectations. We find the opposite pattern for biasing factors such as emotions or own past performance. The bias blind spot seems to be anchored both in the perception of holding expectations against the current market sentiment, and a more general propensity to act contrary to mainstream opinion.

Finally we show that second-order beliefs influence investing decisions. The more optimistic about stock returns participants assume other investors to be, the more money they allocate themselves to stocks. If investors perceive a consensus, i.e. first-order and second-order beliefs are aligned, then the impact of second order beliefs is even stronger. Quite logically investors who see other market participants as biased rely less on second-order beliefs. The effect further depends on investors' financial literacy and view of market functioning.

There are several implications of our findings. First, if one interprets joint investor expectations as a form of market sentiment, then actual and perceived sentiment can be two very different things. Investors are largely unaware what others are thinking, and if they base strategies on their second-order beliefs, such as market-timing, they will most likely fail. Second, given that most investors submit too narrow confidence intervals (miscalibration), it would help them to consult other opinions and to widen own confidence intervals accordingly. Confidence intervals that account for the full range of second-order beliefs are usually large enough and thus less susceptible to miscalibration. Third, financial intermediaries and advisors should be aware of these financial judgment biases. This aids them in identifying own biases and biases on the side of their clients. It further prevents them from projecting their own expectations on their clients and to be more careful with predictions in general.

Chapter 4

True Overconfidence: The Inability of Rational Information Processing to Account for Apparent Overconfidence[1]

4.1 Introduction

Overconfidence is not just an artifact of psychological experiments but seems present in many real life situations where considerable stakes are involved. Overconfident decision making has been observed in financial markets (Odean, 1998), corporations (Malmendier and Tate, 2005), with business entries (Cooper, Woo, and Dunkelberg, 1988) or even marriages (Mahar, 2003). Indeed, overconfidence is perhaps the behavioral bias most readily embraced by academic researchers in economics and fi-

[1]This chapter was published 2011 in *Organizational Behavior and Human Decision Processes*, Vol. 116(2), pp. 262-271.

nance. In particular the better-than-average effect, which is the tendency of people to rate their skills and virtues favorably relative to a comparison group, yields direct predictions for economic decision making.

In a recent paper, Benoît and Dubra (2009) challenge the notion of overconfidence as it was previously analyzed in psychology and economics. The subject of their criticism is the conventionally used research methodology to demonstrate the better-than-average effect. In a signaling framework, Benoît and Dubra show that rational information processing can lead to the very results formerly interpreted as evidence for overconfidence. This does not rule out true overconfidence as an explanation for these findings, but instead also allows for straightforward Bayesian updating as an alternative explanation.[2]

Despite this setback for the overconfidence literature, it is not sufficient to take a methodological viewpoint on the matter; we have to ask ourselves about the psychological reality of this bias and its relation to other self-serving biases. The assertion that people are overconfident is an appealing explanation for behavior, both on the financial markets and elsewhere. In contrast, rational updating is demanding in terms of people's information processing capacity and the underlying signal structure necessary to produce apparent overconfidence. It therefore seems worthwhile to design a research strategy that would be able to demonstrate the presence of true overconfidence by improving previous research methodology in such a way that it becomes capable of withstanding the critisism of Benoît and Dubra (2009).

We identify the aggregation of beliefs as the feature most damaging to the interpretational value of the traditional experimental setting. The simplest setup asks people to judge whether they believe themselves to be above average in a certain domain,

[2]In accordance with Benoît and Dubra, we use the term "true overconfidence" for truly biased self-evaluations. In contrast "apparent overconfidence" stands for data that seems to reflect overconfidence, but where it is not possible to prove the presence of a better-than-average effect. The term "apparent overconfidence" thus includes cases of true overconfidence and other possible causes such as rational information processing.

as for example in the famous account on driving ability by Svenson (1981). More advanced designs ask participants to specify the percentile of a distribution they believe themselves to belong to (e.g. Dunning, Meyerowitz, and Holzberg, 1989). Both approaches have the common feature that as they only retrieve a single estimate, a lot of information gets lost, thus leaving room for alternative explanations. Many sets of beliefs can produce the same result when aggregated in this manner; Bayesian posteriors and true overconfidence are just two of them.

We design two experiments to elicit more detailed beliefs of participants concerning a number of domains that have previously been associated with overconfidence, overoptimism, or underconfidence. Self-evaluations are given along a quantile scale that describes the ability distribution relative to a peer group. Along this scale, participants provide estimates representing their subjective probability of themselves falling into each skill quantile. The extended assessment allows us to directly test whether the findings are in line with rational information processing.

Our central result is that considerable overconfidence is present in the belief distributions of experiment participants. We test various conditions for population averages of these probability distributions and find them incompatible with rational information processing. Bayesian updating can be rejected as an explanation for apparent overconfidence at conventional significance levels. Most people find it highly probable that they rank among the higher quantiles of the ability distribution and not at all likely that they are below average. On an individual level, they often fall short of their expectations, and especially the unskilled exhibit pronounced overconfidence. We conclude that true overconfidence is the main driver of our results.

4.2 Types of overconfidence

While this is not the place to review an abundant overconfidence research in psychology and economics (consider e.g. Glaser and Weber, 2010), it is nevertheless useful to divide the field into three subareas which can be summarized following Moore and Healy (2008):

1. Judgments of one's absolute performance or ability (overestimation)
2. Confidence in the precision of one's estimates (miscalibration or overprecision)
3. Appraisal of one's relative skills and virtues (better-than-average effect or overplacement)

Overestimation is diagnosed if people's absolute evaluation of their own performance (e.g. correct answers in a knowledge test) exceeds their actual performance (Lichtenstein, Fischhoff, and Phillips, 1982; Moore and Healy, 2008). Miscalibration or overprecision denotes the observation that people choose overly narrow confidence intervals when asked for a range that is supposed to contain a true value with a certain probability (Alpert and Raiffa, 1982; Russo and Schoemaker, 1992). Overplacement often occurs when people try to evaluate their competence in a certain domain relative to others. Typically, most people rate themselves above average, which is why this effect is also called better-than-average effect (Alicke and Govorun, 2005). The relationship between these different forms of overconfidence is discussed for instance, in Glaser, Langer, and Weber (2009), Healy and Moore (2007), and Larrick, Burson, and Soll (2007).

Apart from the aforementioned, overoptimism (Weinstein, 1980) and illusion of control (Langer, 1975) are associated with overconfidence in a broad interpretation of the term. We will concentrate on the better-than-average effect (overplacement) and occasionally on overoptimism, as the elicitation techniques for these biases are similar.

Criticism which has been raised against all types of overconfidence is usually directed either at research methodology and experimental design or the underlying concept itself; the list of authors in psychology who have questioned the reality of overconfidence or the research design includes Gigerenzer (1991), Gigerenzer, Hoffrage, and Kleinbölting (1991), Juslin (1994), Erev, Wallsten, and Budescu (1994), Dawes and Mulford (1996), and Klayman, Soll, González-Vallejo, and Barlas (1999). In economics—where the rationality assumption was long prevalent—the emphasis was a different one: in recent years, various approaches were pursued to reconcile overconfidence with rational behavior (Bénabou and Tirole, 2002; Brocas and Carrillo, 2002; Compte and Postlewaite, 2004; Healy and Moore, 2007; Köszegi, 2006; Santos-Pinto and Sobel, 2005; Van Den Steen, 2004; Zábojník, 2004). These models differ mainly in their assumptions, their relevance for different forms of overconfidence and the degree of rationality they are based on. In many ways, this literature has contributed to improving and clarifying methodology, but the debate whether overconfidence exists at all is far from being settled.

4.3 Criticism by Benoît and Dubra

The reasoning of Benoît and Dubra (2009) to some extent combines the two mentioned strands of criticism. They identify a problematic feature in the conventional procedure to demonstrate the better-than-average effect, namely relative imprecise inquiries for an appraisal of relative skills and virtues. Based on a parsimonious signaling model, they then employ rational Bayesian argumentation to illustrate that this kind of research cannot show overconfidence in the form of the better-than-average effect. We will now examine their reasoning in detail.

Probably the most prominent account of the better-than-average effect is given in Svenson (1981), who finds that a great majority of subjects rated themselves to be safer drivers than the median driver (77% of his Swedish and 87% of his US sample). He explains his findings by a general tendency of people to view themselves more favorably than they view others, possibly accompanied by cognitive effects such as low availability of negative memories. Similar results could be reproduced for other domains, for example for people evaluating their personal virtues relative to others (Alicke, Klotz, Breitenbecher, Yurak, and Vredenburg, 1995).

These overplacement studies have a common research methodology, which often simply consists of asking participants whether they view themselves as better as or worse than the median or average of a comparison group with respect to some skill or virtue. Researchers occasionally require more precise estimates, i.e. other quantiles (often percentiles or deciles) are used instead of the median. Overconfidence is usually diagnosed if significantly more than half of the participants place themselves above the median, or more generally if more than x% place themselves above the (100-x)-percentile.

Some concerns regarding this design were raised earlier; for instance, people may interpret the skill in question differently or they may lack information about its distribution within the population. Additionally, the sample of participants might not be representative of the population, and the meaning of "average" can be understood in various ways. These problems can nevertheless be addressed by a more careful experimental design including precise and unambiguous formulation of questions and a fairly large and representative choice of subjects. Combined with the assumption that participants use best estimates of their own and others' abilities and skills, the general result remains valid—it seems intuitive that no more than a certain fraction of the population can rate themselves above a respective percentile.

However, Benoît and Dubra (2009) show that exactly this is possible even when people update their beliefs in a perfectly rational manner. In order to illustrate this, we shall briefly reproduce their example capitalizing on Svenson's study of driving ability here. In a uniformly distributed population of low, medium and highly skilled drivers, people are assumed to evaluate their driving skills depending on whether they have previously had an accident or not. Probabilities for causing an accident are given as $p_L = 0.8$, $p_M = 0.4$ and $p_H = 0$ for the different groups. If drivers do not know their initial skill level but interpret the occurrence of an accident as a signal, they will update their beliefs according to Bayes' law. Given the prior probabilities, all people who did not experience an accident will arrive at a posterior probability of $\frac{5}{9}$ that they are of high skill; it seems reasonable for this group to rate themselves above average. As 60% of all drivers have had no accident, this implies that these 60% are expected to regard themselves as highly skilled. Beyond this concrete example, Benoît and Dubra (2009) show that—within the traditional experimental design—almost any distribution of respondents on the percentile scale can be explained by rational information processing.

There are some immediate concerns with the Benoît-Druba-criticism. One concern refers to the way people deal with signals in classic overconfidence domains. A very early study of Preston and Harris (1965) suggests that even drivers hospitalized after an accident exhibit the same overplacement patterns when asked for their driving performance. The authors find no evidence that participants adjust their evaluation according to the received signal. Benoît and Dubra (2009) discuss this study and some more recent evidence on how people interpret adverse signals. Although the results are quite mixed, a general self-serving bias in the perception of signals is well documented (Bradley, 1978; Zuckerman, 1979). People tend to ascribe bad outcomes to external forces rather than to their own performance or ability. For this reason, it is at the least unclear whether good and bad signals are perceived symmetrically within a Bayesian model.

The framework of Benoît and Dubra also imposes some requirements on signal structure. It is obvious that if the number of signals becomes large (or alternatively the quality of signals very good), overconfidence can no longer be explained rationally. With perfect signals and people allocating themselves reasonably to the percentiles, the distribution will correspond to underlying probabilities. If one extends the aforementioned driving setup by an additional period and maintains the same probabilities as before, this becomes clear: after observing two signals, only 47% would reasonably consider themselves as highly skilled and 27% each as medium or low skill drivers. After ten periods, the Bayesian result is practically indistinguishable from the real probabilities. If, for instance, ten signal observations are interpreted as ten years of driving experience, there is no room for overconfidence afterwards. While it is possible to construct different examples that still show rational overconfidence after many periods, this comes at the cost of a highly asymmetric signal structure with the rare occurrence of (very) negative signals. Indeed, the asymmetric signal structure is one key ingredient to the emergence of rationally explicable overconfidence in Benoît and Dubra (2009), but this is a good portrayal of reality only for some domains (e.g. driving).

However, with respect to signal frequency and quality, feedback is far from perfect in many situations—even in financial markets where new information arrives almost continuously. It would therefore be premature to dismiss the Benoît-Dubra-criticism solely on these grounds. We instead design an experiment to distinguish between two possible sources of apparent overconfidence, namely rational information processing and true overconfidence.

4.4 Derivation of an alternative experimental design

The distinction of how people arrive at overconfident judgments is crucial, as Bayesian updaters are not biased in a special direction; whether they appear over- or underconfident is simply a result of prior probabilities and signal distribution. Any claim that human beings are persistently overconfident must be based on a non-rational formation of beliefs. Overconfidence is consequently mostly modeled as over- or underreaction relative (and thus distinct) to rational Bayesian updating (cp. e.g. Odean, 1998).

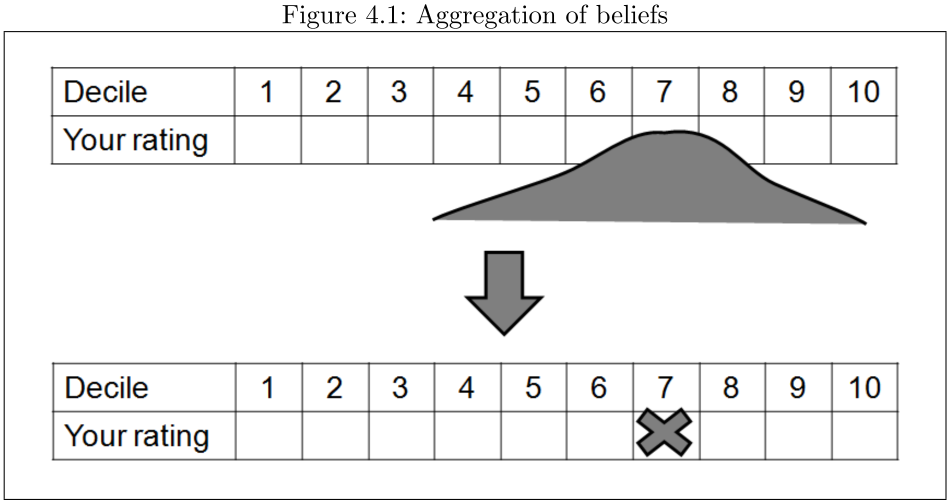
Classic experiments, however, are unable to distinguish between apparent and true overconfidence. To overcome this problem, Benoît and Dubra (2009) propose using a stronger requirement to test for overplacement. Based on their proof that maximally 2*x% can rate themselves rationally among the top x% of the population, they suggest using this hurdle for future experiments. This rule is unsuitable for the often used median condition and represents a very strong requirement to find overconfidence for other percentiles. For example (following this logic), more than 60% of the subjects must place themselves among the top 30% of a population before one can deduce a better-than-average effect. Although many studies observe overconfidence among their participants, it is rarely this pronounced. Even for the large levels of overconfidence observed by Svenson (1981) this rule allows to identify overconfidence only for some intervals of his US subject sample.

Furthermore, this rule applies only in the case that people indeed use the median of their own beliefs to arrive at their rating. Another perfectly prudent way of answering such a question is to take the average of one's beliefs; in that case almost any possible distribution of self-evaluations can be rationalized (for a proof, see again Benoît and Dubra, 2009).

The difficulty in showing true overconfidence within the traditional framework lies in the aggregation of information resulting from subjects placing themselves in one specific half, decile, quartile or other category. In the example mentioned, drivers that experienced no accident had a posterior probability of $\frac{5}{9}$ of being of high skill, $\frac{1}{3}$ of being of medium skill and $\frac{1}{9}$ of being of low skill. This distributional information gets lost if one observes only a point estimate. Figure 4.1 shows how a rich belief distribution containing information about probabilities for all deciles is represented by a single rating. First, subjects enjoy some discretion concerning how they determine this rating given their beliefs, i.e. how to summarize their beliefs in a single parameter. Second, and more importantly, the resultant ratings yield much less information to distinguish between true overconfidence and alternative explanations. This is why one can often only speak of "apparent overconfidence" in such situations.

The experimental setting proposed here is to ask subjects directly for the probabilities with which they would place themselves in the different quantiles (e.g. deciles). This avoids the complication of different aggregation methods and preserves the ad-

Figure 4.1: Aggregation of beliefs



Notes: The top panel of the figure shows a belief distribution over ten deciles a person may possess about a skill or ability. The typical assessment of the better-than-average effect asks people to aggregate this belief distribution in a one point-estimate.

ditional information coming from people's distributional beliefs. The setup imposes clear restrictions on what is possible under rational Bayesian updating. Posterior probabilities calculated by Bayes' law, weighted by their occurrence, must add up to the relative frequencies within the population. In a quantile framework, these real probabilities are simply defined by the chosen partition of the scale: for instance, for any decile, there are 10% of the population who belong to that decile. To determine whether the conditional beliefs for a state A are consistent with Bayesian updating, one has to check whether

$$\sum_i P(A|S_i) \times P(S_i) = P(A), \tag{4.1}$$

where the signals S_i form a disjoint partition of the universe.

To make this restriction clearer, we will again refer to the driving skill example: in the example, 60% of the population had no accident. These people share the beliefs mentioned above: $P(H|\textit{no accident}) = \frac{5}{9}$, $P(M|\textit{no accident}) = \frac{1}{3}$, and $P(L|\textit{no accident}) = \frac{1}{9}$. Among the 40% who experienced an accident, posterior probabilities are 0 for being highly skilled, $P(M|\textit{accident}) = \frac{1}{3}$ and $P(L|\textit{accident}) = \frac{2}{3}$ for being of medium and low skill. To translate the example into equation (4.1), signal S_1 corresponds to "no accident" and signal S_2 to "accident". Together, these two signals describe all possible scenarios. If we plug in the different skill levels for A, we arrive at the following equations:

$$\begin{aligned}
P(H|\textit{no accident}) \times P(\textit{no accident}) + P(H|\textit{accident}) \times P(\textit{accident}) &\overset{!}{=} P(H)\\
P(M|\textit{no accident}) \times P(\textit{no accident}) + P(M|\textit{accident}) \times P(\textit{accident}) &\overset{!}{=} P(M)\\
P(L|\textit{no accident}) \times P(\textit{no accident}) + P(L|\textit{accident}) \times P(\textit{accident}) &\overset{!}{=} P(L)
\end{aligned}$$

We know from the given distribution of driving skill within the population that $P(H) = P(M) = P(L) = \frac{1}{3}$. This provides us with three conditions that have to hold when beliefs are updated rationally. As the posterior probabilities stated above were calculated by Bayes' law, the conditions are of course met.

Note that in the example, the signal and the probability of the signal for each group were known; this is not necessarily the case. In an experimental setting, $P(S_i)$ are unobservable and signals may be much more complicated than the binary "accident" versus "no accident". We treat the S_i as elements of a set of possible signals S. One may think of these signals as idiosyncratic life-time experiences in a certain domain. We do not impose any further restrictions on these signals except for the standard assumption that signal realizations for experiment participants are randomly drawn from S.

We now define K ability quantiles Q_k to provide a common understanding of skill levels. The probability $P(Q_k)$ of falling into each quantile is evaluated conditional on the observed signal S_i and subjects in an experiment will thus report $P(Q_k|S_i)$. Inserting this into equation 4.1 we obtain:

$$\sum_i P(Q_k|S_i) * P(S_i) \stackrel{!}{=} P(Q_k) = \frac{1}{K} \tag{4.2}$$

The right-hand side of equation 4.2 is defined by the choice of the scale's partition. Conditional probabilities $P(Q_k|S_i)$ may differ from $1/K$ but—weighted by the probability of the signals in the population—they must equal $P(Q_k)$. As $P(S_i)$ corresponds to the fraction of subjects observing signal S_i, the average reported probability for a quantile must again equal 1/K. We arrive at K equations of this form as the condition needs to be satisfied for each quantile.

In an experimental setting, concrete signals and signal probabilities are usually unknown. However, under random sampling for the number of participants $n(S_i)$ observing each signal S_i it holds that $E[n(S_i)] = P(S_i) * n$; we can thus replace

$P(S_i)$ by $E[n(S_i)/n]$. Moving the expectation operator outside the sum, we arrive at equation 4.3:

$$E\left[\sum_i P(Q_k|S_i) * \frac{n(S_i)}{n}\right] \stackrel{!}{=} \frac{1}{K} \tag{4.3}$$

Our final simplification is to assume that $n(S_i)$ equals one, which corresponds to the notion of idiosyncratic signals—we nevertheless allow for several subjects to observe the same signal or for elements of S not to be observed.

In the experiments we will mostly use a decile setup. Equation 4.3 then generates ten conditions of the form:

$$E\left[\frac{1}{n}\sum_{j=1}^{n} P(Q_k|S_j)\right] = 0.1 \qquad k = 1, ..., 10 \tag{4.4}$$

The left-hand side represents the average reported probability for each ability decile, which in expectation equals 0.1 for a population of perfect Bayesian updaters. It enables us to compare the realized average belief distribution in the experiments to the uniform benchmark distribution.

To test for true overconfidence, we will additionally rely on two conditions: first, the average reported probability mass for the upper half of the ability quantiles should not exceed 50%. This follows directly from equation 4.1 if A represents the state of "being above average". Likewise, in the decile setup of equation 4.4, the probability of the union of the top five deciles must in expectation equal 0.5 across participants. In contrast, true overconfidence would predict that

$$\frac{1}{n}\sum_{k=6}^{10}\sum_{j=1}^{n} P(Q_k|S_j) > 0.5 \tag{4.5}$$

Of course, similar relations also exist for the top 30%, top 20% and other fractions of the scale. Since the better-than-average effect takes its name from the notion of being above average, we will mainly concentrate on 4.5 but report other results occasionally.

As a second indicator of overconfidence, we consider the mean of the individual belief distribution means. This mean of means should correspond to the middle point of the ability scale in a population of rational updaters, which follows from the definition of the mean for individual belief distributions $\sum_{k=1}^{K} P(Q_k|S_i) * k$ and the aggregation using the signal weights:

$$\sum_i P(S_i) * \sum_{k=1}^{K} P(Q_k|S_i) * k = \sum_{k=1}^{K} \sum_i P(Q_k|S_i) * P(S_i) * k = \sum_{k=1}^{K} P(Q_k) * k = \frac{K+1}{2} \quad (4.6)$$

The second equality uses equation 4.2 and shows that the mean of the belief distribution means must equal $(K+1)/2$. For a decile scale, the mean of individual belief distribution means should thus be at 5.5. True overconfidence predicts a mean of means > 5.5.

4.5 Experiment one

4.5.1 Method

Experiment one was conducted in 2008 at the University of Mannheim. 68 business students completed the paper-based questionnaire; 69% of the participants were male, the median age was 24. We excluded four participants who left blank substantial portions of the questionnaire. Subjects answered questions about their skills and abilities in several domains. We selected various domains of skills and abilities to reflect different levels of overconfidence. Subjects were asked for their performance as a student, their abilities in choosing investments, their ability to get along with other students, their programming skills, their sense of humor and their risk of suffering a heart attack before the age of 40.

The ability to get along with other people is a domain in which people are prone to high overconfidence (Moore and Cain, 2007). The same applies to sense of humor (Kruger and Dunning, 1999). In contrast, the performance as a student is regularly objectively reported by a relatively efficient feedback mechanism (grades), thus less overconfidence is expected here. For investment abilities, we anticipate considerable overconfidence in line with the behavioral finance literature (e.g. Odean, 1998). Computer skills were previously related to underconfidence (Kruger, 1999), the risk of a heart attack is one of many incidents where overoptimism has been observed (Weinstein, 1980).

Appendix A contains part of the questionnaire used in the experiment, which explains to subjects the meaning of the quantile scale and how to fill out the input fields. They were asked to state their probabilities for the quantiles of the scale according to their belief distribution. We alternately used a quartile and a decile scale; the decile scale has the advantage of being more precise at the expense of being more demanding to complete.

It is crucial for our analysis of belief distributions that people understood the scale and answered the question for their probabilities of falling into each quantile reasonably. They were informed by the instructions that the probabilities had to add up to one (see Appendix A). For 96% of the entries, people obeyed this rule. In the remaining cases, the sum of probabilities was almost always close to one, suggesting mistakes in calculation and not in comprehension; we nevertheless exclude these cases.

Additionally, we asked subjects for a point estimate along the quantile scale following the traditional approach of demonstrating overconfidence; they thus made both judgments as illustrated in figure 4.1. This enables us to compare the two evaluations and—most likely—infer how subjects tend to aggregate their beliefs. We varied the order of point and probability judgments and for two domains we used a between-group design in which one group was asked only for probabilities while the other stated only a point estimate; this allows us to test for order effects and interdependencies between the two types of evaluations.

4.5.2 Results and discussion

4.5.2.1 Classic overconfidence

Subjects classifying themselves by point estimates into ability quantiles for the different domains corresponds to the traditional way of testing for overplacement. We can conclude whether there is apparent overconfidence or not and—more importantly—later compare whether true overconfidence shows up in the same domains. Table 1 shows the results.

Table 4.1: Point estimates of own skills and virtues on quantile scales (Experiment 1)

Domain	Expectation	Scale	n	Mean	Median	% above avg.
Study performance	overplacement (> 5.5)	deciles	64	6.39^{***}	6^{***}	78.1^{***}
Financial markets	overplacement (> 2.5)	quartiles	61	2.56	3	54.1
Sense of humor	overplacement (> 5.5)	deciles	25	7.80^{***}	8^{***}	96.0^{***}
Programming skills	underplacement (< 2.5)	quartiles	25	2.32	2	32.0^{*}
Getting along with others	overplacement (> 5.5)	deciles	64	7.08^{***}	7^{***}	84.4^{***}
Risk of heart attack	overoptimism (< 2.5)	quartiles	63	1.83^{***}	2^{***}	25.4^{***}

Notes: The table shows the tested skill domains, the experimental expectation (including its numerical meaning), and the partition of the scale used for each domain in experiment one. It contains number of observations, mean, median and percentage of subjects that placed themselves above average. For the decile scale the midpoint is 5.5, for the quartile scale 2.5. We use a two-tailed t-test (mean), Wilcoxon signed-rank test (median), and binominal probability test (percentage above average). $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

In line with earlier research, participants appear to be exceedingly overconfident when judging their sense of humor or their ability to get along with others. Most people see themselves above average (96% and 84%, respectively) and many place themselves in the 7th or 8th decile of the ability distribution. However, one needs to keep in mind that even this extreme case is not sufficient for proving the existence of true overconfidence; for instance, with respect to sense of humor, exactly 60% place themselves among the top 30% of the distribution, which does not violate the condition set up by Benoît and Dubra (2009). As discussed before, their requirement is too strong for the ratios typically found in overconfidence experiments.

For study performance (a domain where better feedback is available), overconfidence is less pronounced, but still mean, median and percentage of participants viewing themselves to be above average are significantly greater than the corresponding neutral values. We find only slight underconfidence for programming skills and a neutral result for investing abilities. A young student population accustomed to computers may feel more competent in programming while at the same time having little financial market experience. Kruger (1999) shows that the self-assessed ability in a domain is a strong driver of overplacement. In particular Glaser et al. (2009) find higher levels of overconfidence for finance professionals than for students. The result for risk of a heart attack is again as anticipated: most participants are overoptimistic and assess their personal risk as lower compared to their peers; in fact, 75% of participants believe that their risk is below average.

We emphasize that—for the purpose of this chapter—it is less important whether the results for each domain match precisely the expectations derived from the literature as we are primarily interested in the relationship between apparent and true overconfidence.

4.5.2.2 Probability assessments

The key data of our study consist of the probability assessments supplied by experiment participants. While participants exhibit many shapes of belief distributions—some skewed and others symmetric, some flat and others very steep—the distributions have in common that they are unimodal, i.e. exhibiting a single probability peak in one quantile or several adjacent quantiles sharing the same probability. We almost never observe a first peak followed by a drop in probability followed by another peak; this makes sense intuitively as one mostly feels either skillful or not for any given domain. The tightness of the distribution hints at how sure subjects are about their self-assessment: very often, subjects indicate a zero probability for several deciles, i.e. they are sure that they could conceivably fall only into a certain range of the scale.

When we ask for the whole distribution of beliefs, it is no longer possible for a rational population to be predominantly above average in the sense that neither the mean of the individual distribution means, nor a major part of the probability mass of the distributions can be significantly above average. In a decile framework, the mean of individual distribution means must be at 5.5, in a quartile setup it must lie at 2.5; the aggregated probability mass must be split equally between the lower and upper half of the quantiles. In fact, in a population of true Baysians, equation 4.2 has to hold in expectation for every quantile. This is far more restrictive than the direct assessment analyzed before, where people only indicated to which quantile of the distribution they believed themselves to belong to.

The results for the probability assessment in table 2 appear similar to those of table 1. We again find strong overconfidence for sense of humor and the ability to get along with others, and somewhat weaker overconfidence for study performance, all significant at 1%-level (t-test). The rightmost column of table 2 refers directly to equation 4.5. Values significantly above 50% indicate true overconfidence (over-

Table 4.2: Population averages of probability assessments for skills and virtues (Experiment 1)

Domain	Scale	n	Mean of distr. means	Total probability mass above average
Study performance	deciles	64	6.20***	69.0%***
Financial markets	quartiles	63	2.51	50.0%
Sense of humor	deciles	39	7.08***	80.6%***
Programming skills	quartiles	38	2.15***	36.2%***
Getting along with others	deciles	64	7.08***	80.6%***
Risk of heart attack	quartiles	64	2.02***	31.0%***

Notes: The table shows the tested skill domains and scale used for each domain in experiment one. It contains number of observations, mean of individual distribution means, and the total probability mass above average in %. For the decile scale the midpoint is 5.5, for the quartile scale 2.5. Two-sided t-test: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

placement). Overplacement is found in all expected domains with the exception of "financial market investment". Analogously, underplacement can be diagnosed for programming skills and risk of a heart attack (overoptimism). We did not find any order effects, neither for the order of domains nor for the order of probability estimate and point estimate. The between-subject domains ("sense of humor" and "programming skills") reveal that the degree of overplacement is similar between point estimates and probability estimates even if compared across groups.

To test whether the elicited probabilities coincide approximately with the rational benchmark (and thus might have been derived by Bayesian updating), we use a chi-square goodness-of-fit test and a Kolmogorov-Smirnov test for equality of distributions. Table 3 shows the p-values of these tests for the skills and abilities used in experiment one.

For four domains, both tests reject rational information processing at 1%-level, for the remaining two domains at least the chi-square test is significant.[3] The in

[3] The differences between the two tests arise from the fact that the chi-square test penalizes any deviation from the distribution, while the Kolmogorov-Smirnow test is sensitive to deviations in the cumulative distribution function.

Table 4.3: Tests for compatibility of average belief distributions with the prediction of rational information processing (Experiment 1)

	p-values	
Domain	χ^2-test	KS-test
Study performance	0.000	0.002
Financial markets	0.005	0.351
Sense of humor	0.000	0.001
Programming skills	0.041	0.211
Getting along with others	0.000	0.000
Risk of heart attack	0.004	0.009

Notes: The table reports two tests whether the average probabilities submitted by experiment participants correspond to the theoretical prediction of Bayesian updating. It shows the p-values of a chi-square test with nine degrees of freedom (deciles) and 3 degrees of freedom (quartiles) and the p-values of a Kolmogorow-Smirnov test for all domains of experiment one.

general pronounced asymmetric shape of the average belief distribution cannot be reconciled with the normative result derived in section 4. While in the theoretical driving skill example optimistic assessments of those who received a positive signal were counterbalanced by the beliefs of those with a negative signal, this does not seem to happen in the experiment. We will analyze individual belief distributions in more detail in experiment two.

4.5.2.3 Limitations of experiment one

It has been argued that ambiguity is a problem in questions concerning skills and virtues (Dunning et al., 1989; Van Den Steen, 2004). People might interpret the skill in question differently, consequently allowing everyone to rightfully reach the conclusion that they are above average with respect to their subjective definition of the skill in question. There also were no monetary incentives in experiment one, primarily because a convincing incentive scheme was not available. However, it has been demonstrated that behavioral biases may disappear with proper incentivization, although evidence in psychological and economic experiments is mixed (Camerer and Hogarth, 1999; Hertwig and Ortman, 2001). To account for these possibilities, we design a second experiment in which subjects have to evaluate their performance in incentivized laboratory tasks.

4.6 Experiment two

4.6.1 Method

Experiment two took place at the University of Mannheim in 2010. 50 students of various faculties (50% business and economics) were recruited via an online recruitment system for economic experiments (ORSEE; Greiner, 2004). 48% of the participants were male, the median age was 24. Experiment two was computer-based and programmed in z-tree (Fischbacher, 2007). In this experiment we elicited probability assessments for four tasks conducted in the laboratory. We chose tests for intelligence, memory, creativity, and general knowledge as tasks for the experiment (see Appendix B); these domains should represent meaningful and desirable qualities for our subjects.

Trait desirability often goes along with self-serving ability assessments (Alicke, 1985); in particular, Burks, Carpenter, Goette, and Rustichini (2009) find overplacement in an IQ test and Moore and Healy (2008) demonstrate a similar effect in knowledge tests. It has been further shown that in social comparisons, easy tasks typically produce more pronounced overplacement than do difficult tasks (Larrick et al., 2007; Moore and Healy, 2008), as people seem to focus on their own result and do not fully account for the task being easy or difficult for most of the other participants as well.[4] We thus expect true overconfidence of subjects for the test domains, probably moderated by increasing task difficulty.

As subjects had given appropriate responses in the more precise decile framework in experiment one, we used solely this design in experiment two. The wording of the experimental instructions remained the same (see Appendix A). It was automatically checked whether probabilities summed to one and subjects were prompted to correct their entries if not. Two participants repeatedly failed to correct their answers and were excluded from the analysis.

Participants completed tasks prior to their evaluation of probabilities, implying that at the time they had to make their judgments, there was little ambiguity about which performance they should evaluate. We used a quadratic scoring rule to incentivize subjects (Selten, 1998), as a quadratic scoring rule makes it optimal for (risk-neutral) subjects to submit their true belief distributions. If anything, risk-averse subjects would bias their response to a more uniform distribution which would counteract our results. In overconfidence research, Moore and Healy (2008) apply the quadratic scoring rule in a similar experimental design. Participants were told that tied scores would be resolved by chance.

4.6.2 Results and discussion

4.6.2.1 Probability assessments

Participants find it most likely that they rank between the sixth and ninth decile for the tasks in experiment two. Table 4 shows the average probabilities that participants stated for the ability scale used. Relatively few subjects believe their performance to be in the very top decile compared to their peers. While the seventh and eighth

[4] This finding is a reversed form of the classic hard-easy effect (Lichtenstein et al., 1982; Juslin, Winman, and Olsson, 2000).

Table 4.4: Average probabilities assigned to deciles (Experimtent 2)

	worst 10%				decile scale					best 10%
Domain	1	2	3	4	5	6	7	8	9	10
Intelligence	0.014	0.020	0.045	0.073	0.120	0.140	0.189	0.185	0.129	0.085
	(0.061)	(0.061)	(0.105)	(0.094)	(0.127)	(0.126)	(0.152)	(0.147)	(0.146)	(0.165)
Memory	0.016	0.023	0.014	0.025	0.076	0.115	0.167	0.200	0.189	0.176
	(0.065)	(0.074)	(0.049)	(0.055)	(0.131)	(0.135)	(0.149)	(0.132)	(0.169)	(0.245)
Creativity	0.056	0.084	0.088	0.115	0.125	0.147	0.155	0.132	0.058	0.040
	(0.130)	(0.135)	(0.119)	(0.125)	(0.113)	(0.128)	(0.153)	(0.135)	(0.101)	(0.082)
Knowledge	0.033	0.071	0.078	0.062	0.073	0.107	0.152	0.188	0.153	0.085
	(0.083)	(0.131)	(0.135)	(0.093)	(0.096)	(0.129)	(0.159)	(0.167)	(0.153)	(0.162)

Notes: The table shows the average probabilities assigned to the deciles of the ability scale for the four tasks of experiment two. Standard deviatons are in parentheses.

decile are the most popular choice (with average probabilities assigned to these deciles mostly exceeding 0.15), participants submit very low probabilities for the bottom deciles, sometimes as low as between 0.01 and 0.03 for the domains of intelligence and memory. Kruger and Dunning (1999) show that for many tasks, very few people believe that they perform very badly compared to their peers. We add that people do not even find it *probable* that they could be bad.

Table 5 generalizes these results to the statistics of the distribution that we are especially interested in. In three out of four domains, we find significant overplacement of participants measured by both the mean of the distribution means and the average probability mass above the middle point of the scale. The extent of overplacement is comparable to the untested abilities of experiment one. Ambiguity may have contributed to overconfidence in domains such as "sense of humor" (cp. Dunning et al., 1989), but true overconfidence is present also in controlled tasks with little interpretational flexibility.

Table 4.5: Population averages of probability assessments in experimental tasks (Experiment 2)

Domain	Scale	n	Mean of distr. means	Total probability mass above average	Proportion of correct responses
Intelligence	deciles	48	6.74***	72.8%***	69.5%
Memory	deciles	48	7.50***	84.7%***	85.9%
Creativity	deciles	48	5.52	53.2%	32.8%
Knowledge	deciles	48	6.45***	68.4%***	67.8%

Notes: The table shows the experimental tasks and the scale used for each domain of experiment two. It contains number of observations, the mean of individual distribution means, the total probability mass above average in %, and the proportion of correct responses for each task. For the decile scale the midpoint is 5.5. Two-sided t-test: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Extending the analysis to other thresholds than the average or middle point of the quantile scale reveals patterns which were already suggested by the descriptive statistics. For the top 40%, we still find overplacement similar to the presented results for being above average. For high quantiles, however, overconfidence becomes markedly weaker or even disappears altogether. Only in one domain of experiment two (memory test), overplacement is still significant for the top 20%; the better-than-average effect thus appears to be a "slightly-better-than-average effect".

The pattern of overplacement found is in line with the reversed hard-easy effect for relative judgments (Larrick et al., 2007; Moore and Healy, 2008). The right column of table 5 shows the proportion of correct answers given in the tasks. With only 32.8% correct responses, the creativity test clearly qualifies as hard, and we find no significant overplacement; on the other hand, overplacement is most pronounced in the easiest task (memory). We explain this finding by the egocentric nature of relative judgments (Kruger, 1999; Moore and Cain, 2007): subjects react more strongly to variations in their own performance than to possible variations in the performance of other participants.

We again test whether the submitted probabilities coincide approximately with the rational benchmark, and hence might have been derived by Bayesian updating. Table 6 shows the p-values of the chi-square test and of the Kolmogorov-Smirnov test for the tasks used in experiment two.

Table 4.6: Test for compatibility of average belief distributions with the prediction of rational information processing (Experiment 2)

	p-values	
Domain	χ^2-test	KS-test
Intelligence	0.000	0.002
Memory	0.000	0.000
Creativity	0.013	0.343
Knowledge	0.007	0.033

Notes: The table reports two test whether the average probabilities submitted by experiment participants correspond to the theoretical prediction of Bayesian updating. It shows the p-values of a chi-square test with nine degrees of freedom (deciles) and 3 degrees of freedom (quartiles) and the p-values of a Kolmogorow-Smirnov test for all domains of experiment two.

For three out of four domains, rational information processing can clearly be rejected. The asymmetry of belief distributions already visible in table 4 is again not compatible with the uniform distribution postulated by Bayesian updating. The statistical tests suggest that deviations are too strong to be a result of imperfect sampling of experiment participants alone, as this type of randomness should be small in magnitude and not systematic in a manner observed in the presented results.

4.6.2.2 Overconfidence on an individual level

Besides population averages, the controlled tasks allow us to test for overconfidence on an individual level. In the probability framework, participants provide a range of deciles they believe to be possible for themselves with different probabilities. If the actual result is below or above this range, this is far stronger evidence for over- or underplacement than if it were to fall short of—or exceed—a point estimate. Table 7 shows the fraction of participants who ranked below their worst expectation or above their best expectation. For the domains of intelligence, memory, and general knowledge, about a third of the subjects end up in a decile below all of the deciles they had assigned a probability greater than zero, i.e. they perform worse than they had even considered possible. The other extreme—reaching a decile above one's best expectation—happens far less often (between 2% and 10%). Except for the domain of creativity, the difference between the two proportions is strongly significant (z-test). This impression of asymmetry is backed by the fraction of subjects reaching a decile below their mean expectation. It is rather common for subjects to fall short of their mean expectation, especially for the intelligence and memory test.

We have previously in parts explained the failure of average belief distributions to match the rational benchmark by the pronounced overplacement of unskilled participants (cp. Kruger and Dunning, 1999; Ehrlinger, Johnson, Banner, Dunning and Kruger, 2008): those who receive a negative signal should adjust their probabilities accordingly, and (as in the driving skill example) should hence submit high probabil-

Table 4.7: Individual overplacement in experimental tasks (Experiment 2)

	Percentage of subjects ranked...		
Domain	...below worst expectation	...below mean expectation	...above best expectation
Intelligence	27.1%	70.8%	10.4%**
Memory	37.5%	70.8%	2.1%***
Creativity	10.4%	45.8%	8.3%
Knowledge	31.2%	56.3%	4.2%***

Notes: The table shows the proportion of subjects for which their actual decile rank is below their worst expectations, below their mean expectations, and above their best expectations in experiment two. Worst (best) expectations are defined as the lowest (highest) decile for which subjects submit a probability > 0. Asterisks stand for significant differences between the proportion below worst expectation and above best expectation using a two-sample z-test of proportion. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

ities for low quantiles. We thus now examine the belief distributions of two specific groups, namely the skilled and the unskilled, where we define the groups as those who finish in the top three and bottom three deciles in each task, respectively. We assume that participants hold neutral priors before they enter the tasks.[5] A good or bad performance in the task should then inflate their subjective probabilities of falling into low or high quantiles, at least if subjects interpret their task performance correctly and update their beliefs in a rational manner.

However, table 8 shows that unskilled participants recognize their negative signal only partially: with the exception of the memory task, they understand that they are less likely to reach the top 30% but only slightly and occasionally increase their probability for the bottom 30%. For two tasks, the intelligence and memory test, they state even smaller probabilities than the neutral prior probability of 0.3 for the bottom 30%. Skilled participants seem to react more strongly to their positive signal: while they assign high probabilities to the top 30%, they regard the bottom 30% as almost impossible. This asymmetry in signal processing is one major cause for the disparity between the belief distributions and the rational benchmark. The right column of table 8 displays the average probability subjects assign to the correct decile, i.e. the decile they actually fall into according to their task performance. With the exception of creativity, skilled participants here submit higher probabilities and are thus better able to recognize their true skill level; they consequently earn more under the quadratic scoring rule regime. This finding is consistent with the idea that poor performers also lack metacognitive skills (Ehrlinger et al., 2008). While part of the result may be due to a regression effect (cp. Burson, Larrick, and Klayman, 2006), it cannot explain the asymmetry we observe between the judgments of unskilled and skilled participants.

[5] This is a conservative assumption; if anything, experience should already have shifted the skilled and unskilled towards more realistic priors.

Table 4.8: Probability assessment of skilled and unskilled participants (Experiment 2)

Domain	Skill level	Estimated probability to be in top 30%	Estimated probability to be in bottom 30%	Estimated probability for actual decile
Intelligence	skilled	0.54**	0.04***	0.20**
	unskilled	0.20	0.19	0.06*
Memory	skilled	0.83***	0.00***	0.33***
	unskilled	0.39	0.07***	0.04**
Creativity	skilled	0.42	0.05***	0.15
	unskilled	0.12**	0.46*	0.18*
Knowledge	skilled	0.65***	0.03***	0.22**
	unskilled	0.17	0.42*	0.08

Notes: The table shows the average estimated probabilities of skilled and unskilled participants for different ranges of the decile scale in experiment two. We test for deviations from neutral prior probability using a Wilcoxon signed-rank test. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

4.7 General discussion

We propose a new methodology to measure overconfidence. Experiment participants evaluate their relative position within the population for different skills and tasks by stating their complete corresponding belief distribution; they provide probability estimates for each decile or quartile instead of a single point estimate. This approach avoids many problems that were shown to be detrimental to previous research in overconfidence. Belief distributions yield clear restrictions as to what is possible for a population of rational Bayesian updaters.

There is considerable overplacement in the belief distributions of experiment participants. Probability estimates closely resemble results based on traditionally used point estimates. Population averages for different characteristics of belief distributions are inconsistent with Bayesian updating. Participants on average state high probabilities for quantiles above average while they regard it as unlikely that they should fall into the bottom quantiles. Because of this pattern, the aggregated belief distribution fails to match the rational benchmark. Individual level results confirm these observations, with people often underperforming even their worst expectations. Overplacement is particularly pronounced for unskilled participants who apparently do not fully account for the negative signals they receive.

4.7.1 Causes of true overconfidence

We believe that motivational and non-motivational factors account for the existence of overconfidence (and in particular overplacement). It has been argued that positive illusions contribute to mental health and well-being (Taylor and Brown, 1988). They foster self-esteem and enhance the motivation to act (Bénabou and Tirole, 2002). Among the non-motivational factors, selective recruitment of information, focalism, and egocentrism have been put forward (cp. Alicke and Govorun, 2005). As discussed before, ambiguity, desirability, and controllability of the judgment item moderate the degree of overplacement. We favor these explanations as they examine the psychological roots of the phenomenon and seem more plausible than a logically rigorous but less realistic model.

Healy and Moore (2007) provide such a rational explanation for the occurrence of the reversed hard-easy effect, which we observe in experiment two. In their model, people hold incorrect prior beliefs but then update these beliefs rationally. If subjects perform better or worse than their expectation, they will attribute this partly to chance and partly to their ability; we cannot fully exclude this possibility. However, we use tests that relate to abilities and virtues such as memory or creativity for which subjects should have more accurate prior beliefs (compared to the trivia quizzes used in Healy and Moore, 2007). To additionally reduce surprise potential in these tasks, we mostly use questions of a type that people may have seen before, e.g. typical IQ-test questions. In the post-experiment questionnaire, subjects rate the tasks according to their perceived reliability to test for the ability in question. Predominantly high ratings support the impression that the test content was in line with the expectation of participants.

4.7.2 Remaining caveats

A caveat to the proposed methodology is that participants may have difficulties with meaningfully completing the probability evaluation. People possess underlying beliefs about their skills but may not be able to express them in a probability distribution. We tried to address this concern by a careful analysis of what subjects actually do in the experiment: individual responses seem reasonable (as submitted belief distributions are unimodal without jumps or breaks), but this is of course only indicative evidence. Additionally, the incentive scheme in experiment two should motivate subjects to represent their true beliefs as closely as possible.

We further do not measure priors directly in experiment two and also cannot observe the signals participants receive from having completed the tasks. It is thus hard to determine precisely at what point during the information processing procedure the biases occur. However, in combination with experiment one (which elicits unconditional beliefs in several domains) our impression is that both priors and interpretation of signals are biased.

4.7.3 Implications

Overconfidence is among the behavioral biases most readily adopted by academic researchers in economics and finance. In the literature, it is related to excessive trading volume (Barber and Odean, 2000; Glaser and Weber, 2007; Odean, 1998), to the emergence of stock market bubbles (Scheinkman and Xiong, 2003; Shiller, 2002), to corporate investment decisions (Gervais, Heaton, and Odean, 2003; Malmendier and Tate, 2005), and to the predictability of market returns (Daniel, Hirshleifer, and Subrahmanyam, 1998). Most of these articles take overconfidence as a given result from psychology and not as a subject of further scrutiny. For instance, Odean (1998) states that "a substantial literature in cognitive psychology establishes that people are usually overconfident and, specifically, that they are overconfident about the precision of their knowledge (p. 1888)." Some caution seems to be appropriate here: whereas excessive trading for instance is an observed reality, its link to overconfidence is established only on argumentative grounds; it relies on the existence of overconfidence as a robust feature of human behavior.

Consequently, if the existence of overconfidence is challenged in psychology, this will directly affect the mentioned research in economics and finance. Alternative explanations appear less compelling in many situations, thus without overconfidence these results lose much of their appeal. On that account, our findings contribute to behavioral explanations built on overconfidence remaining intact. They still might inspire some research to directly relate behavioral phenomena to economic reality.

4.7.4 Conclusion

The evidence collected suggests that the theoretically valid criticism of Benoît and Dubra (2009) has only little practical consequences for overconfidence research. In general, apparent overconfidence represents underlying true overconfidence which is reflected in belief distributions. It is not necessary to discard the literature on the better-than-average effect or to redo the entire research with a methodology that is robust against this objection. For future research, scientists may want to adopt a design like ours to avoid potential concerns.

Appendix to chapter 4

Appendix A: Questionnaire

This is an example for the questionnaire in experiment one. Phrasing differed slightly when a point estimate was requested first. Experiment two used similar wording, except that "decile" was replaced by the more intuitive "rank" as participants were invited in groups of ten (in this case deciles and ranks are of course equivalent).

Question 3: Get along with other students

We would like to know how you think about your ability to get along with other students. We ask you to compare your ability with that of the other participants in this experiment. You should by means of a scale evaluate your estimated position within this group. The partition of the scale represents the ten deciles (i.e. 10%) of all students in ascending order. That is, the first decile contains the 10% of students who get along worst with other students, the second decile the following 10% up to the tenth decile, where you find the 10% of students getting along best with other students. You should state for each decile state the probability with which you think you are among this rank group. A value of 0.3 in one of the ten boxes thus means that you assume to belong to this decile of students with a probability of 30%. A value of 0 states that you definitely not belong to this ability group, while a value of 1 indicates that you are absolutely sure to be among this 10%-category.

Please enter the probabilities in the third row of the table and notice that the values must add up to 1.

worst 10%									best 10%
1	2	3	4	5	6	7	8	9	10

Appendix B: Sample questions of the experimental tasks

The intelligence test and the memory test in experiment two were taken from the Italian psychology platform www.nienteansia.it, translated to German and adjusted where necessary. Questions in the general knowledge domain come from Studenten Pisa, a knowledge test administered by German news magazine “Der Spiegel”. The test for creativity is a self-designed remote associates test (Mednick, 1968). We reproduce here three sample questions for each task.

Intelligence test (21 questions)

How does this series of numbers continue? 1 - 4 - 10 - 22 - 46 - 94 - ...

A 188, B 190, C 200, D 47

Which of the following words does not fit the rest?

A Mouse, B Whale, C Snake, D Cat, E Seal

Please complete the sentence: ”Car is to chassis as body is to...”

A Skin, B Blood, C Brain, D Skeleton

Solutions: B, C, D

Memory test (18 questions)

(Participants first read an excerpt from Oscar Wilde’s short story “The Remarkable Rocket”.)

How long had the King’s son waited for his bride?

A One month, B One year, C Two years

What nationality was the bride?

A Russian, B Finnish, C None of these

What means of transportation did the bride use?

A Coach, B Sledge, C Ship

Solutions: B, A, B

Creativity test (12 questions)

(Participants are asked to think of a word that relates to the other three words. We do not present original examples as the task is very language specific.)

Bass—Complex—Sleep

Chamber—Staff—Box

Desert—Ice—Spell

Solutions: deep, music, dry

General knowledge test (24 questions)

In which century did the Thirty Years' War take place?

A 16th century, B 17th century, C 18th century, D 19th century

In which city is the novel "Buddenbrooks" situated?

A Lübeck, B Danzig, C Husum, D Kiel

Which sensory cells in the human eye are responsible for color vision?

A Cones, B Rods, C Plugs, D Studs

Solutions: B, A, A

Chapter 5

Financial Overconfidence Over Time — Foresight, Hindsight, and Insight of Investors

5.1 Introduction

Overconfidence is now for more than a decade among the most popular psychological explanations for investing behavior of private households. It has been linked to portfolio turnover (Odean, 1998; Glaser and Weber, 2007), diversification (Goetzmann and Kumar, 2008) and risk taking (Dorn and Huberman, 2005; Nosić and Weber, 2010) of investors. The implications of overconfidence in this context are mostly viewed negatively, leading to excessive trading, underdiversification, and increased risk taking. However, little is known about the development of financial overconfidence over time and its dynamic interaction with trading behavior. We fill this gap by providing for the first time a comprehensive study of financial overconfidence in its different facets, its consequences for various aspects of real investing behavior and its development over time.

In a panel survey of private investors at a large UK bank, we ask participants for their return expectation and risk perceptions regarding the UK stock market and their own portfolios. From these expectations we construct three measures of overconfidence related to the three types of overconfidence commonly identified in the literature: overestimation, overplacement, and overprecision (cp. Moore and Healy, 2008). The survey is administered every three months between September 2008 and September 2010 resulting in a total of nine survey rounds, which cover one of the most interesting times in recent stock market history. Participants are affluent, self-directed investors, who own on online brokerage account at the bank. Their transactions are recorded, which allows us to combine the survey responses with the actual trades and portfolio holdings of participants. The trading and portfolio data include information about trading frequency, turnover, diversification, and risk taking of investors. This enables us to analyze the most prevalent phenomena that have been linked to overconfidence in a dynamic setting.

We extent existing literature in several ways, as before the link between overconfidence and real trading behavior has often been only postulated (Odean, 1998) or verified by proxies (Barber and Odean, 2001; Goetzmann and Kumar, 2008), or the analysis has been limited to one particular form of overconfidence (Graham et al., 2009). In rare cases two types of overconfidence are considered Glaser and Weber (2007), but then again the dependent variable is confined to trading volume. Besides our systematic and multi-dimensional treatment of overconfidence and investing behavior, we aim for a better understanding of dynamic development of overconfidence as suggested by Gervais and Odean (2001).

We first document the presence of overconfidence in its various facets in our panel. Participants overestimate their actual portfolio returns on average by a large degree and also expect portfolio Sharpe ratios to be higher than ex post realized values. They further believe that their portfolios will outperform the market, while at the same time they perceive own portfolios as less risky on average. This finding of overplacement is supported by survey responses in which participants describe themselves to be better informed, more experienced and more skillful in investing. Overprecision is also widespread in the investor population. Depending on whether confidence intervals are compared to historical volatilities or option implicit volatilities, elicited confidence intervals are too narrow by a factor two to four. Miscalibration for own portfolios tends to be worse than for the market in general. With all overconfidence measures we observe great cross-sectional heterogeneity, which is the prerequisite to explain differential behavior in trading and risk-taking. However, for most measures a majority of participants exhibit overconfidence.

As our main result we establish a relationship between overplacement and trading activity, between overestimation, overprecision, and degree of diversification, and between overplacement, overprecision, and risk-taking. We do not find overprecision to be relevant for trading activity, which is in contrast to the theoretic literature on overconfidence in finance (Odean, 1998; Daniel et al., 1998), but in line with previous empirical results (Glaser and Weber, 2007; Graham et al., 2009). All findings are established in dynamic panel regressions, suggesting a dynamic relationship between overconfidence and trading behavior which goes beyond the view of overconfidence as a stable personality trait. The findings suggest a nuanced role of financial overconfidence for all examined aspects of investing behavior, and they round off and clarify the previous view on this complex relationship.

To learn more about the dynamics of overconfidence we examine persistence and possible causes of overconfidence. We find rank-correlations of overconfidence measures over time mostly between 0.1 and 0.4, suggesting some cross-sectional stability in the degree of overconfidence, but also considerable variability over time (cp. also Glaser et al., 2009). It has been proposed that this variability might be driven by past success and failure, with success leading to an increase in overconfidence through a self-attribution bias (Gervais and Odean, 2001; Daniel et al., 1998). Hereby success can be actual investment success or estimation success with regard to the elicited expectations. It is even possible that perceived success not backed by reality affects overconfidence (cp. Barberis and Thaler, 2003).

We find a strong influence of past investment success on overestimation for the subsequent estimation period, but no effect on overplacement. This means that after good portfolio returns investors overestimate their returns in the future, but after having outperformed the market, they do not necessarily expect to outperform the market again. Analyzing the role of potentially erroneous beliefs about past performance, we observe a positive correlation between perceived and actual portfolio returns, showing that investors have some idea about their realized performance, but still the estimation error is large. Although investors do not consistently overrate their past performance, perceived past success nevertheless contributes to increased levels of overconfidence in foresight. We show that investors, who hold inflated views about past portfolio returns are subsequently subject to higher levels of overplacement and overestimation.

The remainder of the chapter is organized as follows: in section 5.2 we review the overconfidence literature in finance and develop hypotheses for the empirical analysis. Section 5.3 presents the data set, which consists of investors' survey responses and matched trading and portfolio data. In section 5.4 we define measures of overconfidence and report descriptive statistics. Section 5.5 contains our main results with regard to investing behavior, section 5.6 reports findings about the dynamics of overconfidence, a final section concludes.

5.2 Literature and hypotheses

Overconfidence is a well-documented bias in the psychology of judgment and has readily found its way into finance literature. While the notion of overconfident investors seems to have some immediate appeal in describing the behavior of financial market participants, recent evidence suggests that the underlying mechanisms are more complex. First of all the term overconfidence encompasses at least three distinct phenomena to which we will refer to as "types of overconfidence". In analyzing these different types, we adopt the terminology of Moore and Healy (2008), who distinguish between overestimation, overplacement, and overprecision.[1]

People can be overconfident with regard to their absolute ability or performance in a domain; they overestimate their personal outcome, for example the grade they will achieve in an exam or the time they will need to run a Marathon (Grieco and Hogarth, 2009). Overestimation is often demonstrated in performance judgments after experimental tasks, and it has been shown that levels of overestimation increase with difficulty and personal importance of the tasks (Moore and Healy, 2008; Frank, 1935). Investing ranks high on both dimensions, therefore we expect considerable overestimation in judgments of financial performance. The counterpart to overestimation in relative comparisons is overplacement, also known as the better-than-average effect (for a review cp. Alicke and Govorun, 2005). It describes the tendency of people to view themselves above average in many domains, for example almost 90% of a sample of US drivers claim to be above average with regard to driving safety Svenson (1981). Overplacement is present in judgments of skills and abilities Kruger and Dunning (1999) as well as personality traits (Alicke et al., 1995). In contrast to overestimation, levels of overplacement are highest for easy tasks (Kruger, 1999; Moore and Cain, 2007).

[1] Often also overoptimism (Weinstein, 1980) and illusion of control (Langer, 1975) are associated with overconfidence in a broad interpretation of the term.

Finally, overprecision occurs in questions for ranges in which an unknown value will fall with a certain probability. Usually people submit far to narrow intervals (Alpert and Raiffa, 1982; Klayman and Soll, 2004), regardless whether general knowledge questions (e.g. "length of the Nile", cp. Russo and Schoemaker, 1992) or financial values (e.g. "value of the Dow in one year", cp. Glaser et al., 2009) are target of the estimation. Often less than 50% of the true values fall within 90% confidence intervals, indicating severe miscalibration. While it has been argued that the three presented types of overconfidence share a common psychological basis (Larrick et al., 2007), correlations are often found to be low or even negative (Glaser et al., 2009; Moore and Healy, 2008). We will thus document the different types of overconfidence separately and state as our first hypotheses that investors are overconfident.

H1: Private investors are overconfident with regard to their financial market expectations. Overconfidence materializes in the form of overestimation, overplacement, and overprecision.

H1a: Different types of overconfidence are unrelated.

Overconfidence has been introduced to finance to explain long-standing anomalies in investor behavior such as underdiversification (Blume and Friend, 1975) and excessive trading (Odean, 1999; Barber and Odean, 2000). The underlying reasoning is that overconfident investors believe in their investing skills and their private information, and therefore they engage in frequent trades, concentrate on few favorable assets, and take on additional risk. In financial models, overprecision has been linked to increased trading volume and overreaction in stock markets (Odean, 1998; Daniel et al., 1998). This theoretic link is used by Odean (1999) and Barber and Odean (2000) to motivate their findings of high individual trading activity, but the relationship is established solely on argumentative grounds. Barber and Odean (2001) introduce gender as a proxy for overconfidence and find that men trade more frequently than women. However, gender might as well proxy for risk-aversion or other properties.

More recent studies directly relate empirical measures of overconfidence from surveys to trading behavior of investors. Glaser and Weber (2007) find a positive effect of overplacement on individual trading volume, but contrary to theory they do not observe an influence of overprecision. Graham et al. (2009) confirm the role of overplacement using a measure based on return expectation. They further demonstrate a strong effect of investors' perceived competence which might be related to overconfidence. Dorn and Huberman (2005) also report a positive impact of overplacement on turnover, while they find no effect of measures for biased self-attribution and illusion of control. Finally, Grinblatt and Keloharju (2009) show a positive influence of overconfidence (self-confidence corrected for competence) on trading activity.

Neither of the three overconfidence measures in Dorn and Huberman (2005) is related to portfolio diversification, and results for risk-taking are mixed, depending on whether fraction of risky investments or portfolio volatility is analyzed. Goetzmann and Kumar (2008) employ a different strategy to identify an effect of overconfidence on diversification; they use turnover as a proxy for overconfidence and find a negative effect of turnover on diversification. This result relies on a robust relationship between overconfidence and turnover and leaves open which type of overconfidence drives the effect. Nosić and Weber (2010) experimentally test for a relationship between overconfidence and risk-taking and show an increase in risk-taking with higher overprecision. Financial literature in general supports the view that there is a fairly established effect of overconfidence (namely overplacement) on trading volume, while its role for diversification and risk-taking is less well understood. Our contribution is in systematically analyzing the effect of all types of overconfidence on trading, diversification and risk-taking.

H2: Overconfident investors trade more frequently, hold underdiversified portfolios and take more financial risk.

The panel structure of our data allows us to explore the evolution of overconfidence among investors over time. The models of Gervais and Odean (2001) and Daniel et al. (1998) suggest that overconfidence will change dynamically with success and failure. Outcomes that confirm a persons' beliefs and actions tend to elevate confidence too much, while contradicting outcomes weaken confidence too little. This biased self-attribution leads to greater overconfidence after successes and reduced overconfidence after failure. Or, as Gervais and Odean (2001) put it, investors "learn to be overconfident". Deaves, Lüders, and Schröder (2010) provide some evidence for financial professionals that indeed success in estimating returns of a stock market index leads to higher overprecision in the subsequent period. The alternative hypothesis that feedback helps people to become less overconfident or better calibrated has found little support even in controlled experiments under ideal feedback conditions (cp. Pulford and Coleman, 1997, and the literature therein).

H3: Overconfidence depends on previous outcomes. Success (estimation success or investment success) leads to increased overconfidence.

While financial market outcomes can be measured objectively, it is possible that subjectively perceived success is more important for the genesis of overconfidence. In particular, when actual and perceived values fall apart and investors view the past rosier than justified. Hindsight effects have been assumed to contribute to overconfidence (Hoch and Loewenstein, 1989; Hawkins and Hastie, 1990). In their survey on behavioral finance Barberis and Thaler (2003, p.1064) write: "Overconfidence may in part stem from two other biases, self-attribution bias and hindsight bias. [...] If people think they predicted the past better than they actually did, they may also believe that they can predict the future better than they actually can." Winman, Juslin, and Björkman (1998) show a systematic relation between hindsight bias and overconfidence in foresight, which they describe as the confidence-hindsight mirror effect. In contrast, Biais and Weber (2009) do not find a relationship between hindsight and overconfidence in a financial context. We adopt a broader notion of hindsight here, which does not only include the classical hindsight bias (overestimation of the probability of having predicted the correct outcome), but also other hindsight effects such as overestimation of past investment success. We expect hindsight to contribute to overconfidence.

H4: There exists a positive and systematic relationship between hindsight and overconfidence.

5.3 Data

To test for overconfidence and its consequences for investing behavior, we obtain survey responses and transaction data for a sample of clients at Barclays Stockbrokers, a UK direct brokerage provider. Barclays is one of the largest banks in the UK and attracts a wide variety of customers (for demographic characteristics of its clients see chapter 3). The accounts are self-directed in the sense that customers receive no direct investment advice, but they can inform themselves on special web pages provided by the bank; most transactions are processed online.

In cooperation with Barclays Wealth, we conduct a repeated survey, which takes place every three months, starts in September 2008 and ends in September 2010. For the initial survey a sample of the banks client base is invited via e-mail to participate in the online questionnaire (for details on the stratified sampling procedure see Weber et al., 2010). In total 617 clients of the bank participate in the survey, 394 of which participate multiple times. 189 participants complete at least five rounds, and 52 answer the questionnaire in all rounds. We have a minimum of 130 observations for each of the nine rounds. In particular, the survey includes questions for expectations of investors, which are our main source for overconfidence assessments. Additionally we elicit several psychological constructs related to overconfidence.

Our data also include the trading records of all investors, who at least once respond to the panel survey. For the time period between June 2008 and December 2010 we observe 49,372 trades with a total trading volume of £258,940,694.[2]. Of these trades about 75% are in stocks, the remaining trades include bonds, derivatives, mutual funds and ETFs. Panel A of table 5.1 shows descriptive statistics of investors trading activity on a per round basis (3 month). On average 70% of investors trade within each survey period, among those who trade the mean trading frequency is 11.9 (median 5) and mean trading volume is £62,943 (median £9,499). Trading activity is highly skewed with some very frequent and high volume traders. We calculate turnover by dividing trading volume through the sum of the portfolio value at the beginning and the end of each survey period.[3] Average three-month turnover is 39% of portfolio value (median 9.8%), if investors who do not trade are taken into account this values drops to 28% (3.1%). We will later use the variables of Panel A as our measures for trading activity.

Combining trading data with a snapshot of investors' portfolios we are able to calculate portfolio statistics for our survey period. The median portfolio is worth £41,687 (mean £314,663). Panel B of table 5.1 shows our measures for degree of diversification. A simple indicator is the number of portfolio positions investors hold. Most investors own between 1 and 15 assets, with the average (median) at 15.7 (12), which exceeds typical values in other studies on individual investors (Barber and Odean (2000): average 4.3, Goetzmann and Kumar (2008): average 4.7, Glaser and Weber (2009): average 6.8), but is still less than theoretically necessary to obtain a "well-diversified" portfolio (> 30 stocks, Statman, 1987). We calculate a Herfindahl-

[2] We include three month prior to our first survey round and three month after our last survey round

[3] The convention to use twice the portfolio value (or half of trading volume) has been introduced by Odean (1999). He finds average monthly turnover of 6.5%, while Dorn and Huberman (2005) report 18%. Scaled up to three month our value lies between these results.

Table 5.1: Trading and portfolio statistics

Panel A – Trading activity	n	Mean	Median	Std.Dev.	5th Perc.	95th Perc.
Trade dummy	6020	0.704	1	0.456	0	1
Number of trades	4241	11.92	5	26.31	1	46
Trading volume	4241	62,943	9,499	333,499	201	237,777
Turnover	3903	0.386	0.098	0.684	0.003	2.690
Panel B – Diversification	**n**	**Mean**	**Median**	**Std.Dev.**	**5th Perc.**	**95th Perc.**
Portfolio positions	6406	15.72	12	226.80	2	42
HHI	6384	0.242	0.144	0.253	0.021	0.936
Normalized variance	6264	0.545	0.481	0.260	0.195	1
Panel C – Risk taking	**n**	**Mean**	**Median**	**Std.Dev.**	**5th Perc.**	**95th Perc.**
Portfolio volatility	6472	0.370	0.301	0.306	0.137	0.824
Relative volatility	6472	1.410	1.148	0.851	0.672	2.885
Portfolio beta	6473	0.840	0.810	0.420	0.222	1.541
ACV	6304	0.472	0.418	0.292	0.077	1.016
Risk taking task	2114	0.535	0.500	0.278	0.000	1.000

Notes: The table shows descriptive statistics for measures of trading activity, degree of diversification, and risk taking. The trade dummy equals one if an investor trades within a survey round (=time between to surveys), number of trades is the number of sales and purchases, and trading volume is the value of all purchases and sales. Turnover is trading volume divided by the sum of portfolio value at the beginning and end of survey round (we exclude portfolios <£5,000 and winsorize turnover from above at the 5%-level). Portfolio positions is number of different securities hold in portfolio, HHI is sum of squared portfolio weights, and normalized variance the ratio of portfolio variance and average variance of portfolio components. Portfolio volatility is the volatility of portfolio returns over a one year horizon, relative volatiliy is portfolio volatility devided by market volatility (winsorized at 1%-level from above), portfolio beta is the covariance of portfolio and market returns divided by the variance of market returns, and ACV is the average volatility of portfolio components. Risk taking task is the fraction of money invested in the stock market as opposed to a risk-free asset in a hypothetical investment task.

Hirschmann-Index (HHI) of portfolio diversification by taking for each investor the sum of squared portfolio weights. We follow the methodology of Dorn and Huberman (2005) in treating mutual funds as if they consisted of 100 equally-weighted positions. The median HHI amounts to 0.14, which corresponds to a portfolio of seven equally weighted securities. The above reported number of portfolio positions thus overstates the degree of diversification among investors. As a further measure of portfolio diversification, we consider normalized variance as defined by Goetzmann and Kumar (2008). Normalized variance is the ratio of portfolio return variance and average return variance of assets in the portfolio. The mean value of 0.55 suggests that portfolio diversification reduces variance by about one half; again there is great cross-sectional heterogeneity.

Panel C of table 5.1 displays measures of portfolio risk we use to analyze financial risk taking of investors. We calculate portfolio return volatility over a one-year horizon. The mean of 0.37 is well above average UK stock market volatility during the same period of time (0.26). This is reflected also in relative volatility, which we define as portfolio volatility divided by market volatility. But higher risk taking is not explained by higher exposure to market risk, as the average portfolio beta is just 0.84. Instead, the above documented lack of diversification is responsible for high risk levels of investors. Dorn and Huberman (2010) argue that portfolio based measures are hard to evaluate for investors, as they orient themselves at individual securities and disregard correlations between securities; they instead propose a value-weighted average of the return volatilities of portfolio components. This average component volatility (ACV) with a mean of 0.47 exceeds portfolio volatility, as already the related measure of normalized variance suggested. Finally, there is also a survey-based measure of risk taking which is the fraction of a hypothetical £100,000 participants would allocate to the stock market as opposed to a risk-free investment. The average is about 54%, but the range is between investing nothing and investing everything (Cp. Weber et al., 2010, for an in-depth analysis of this investment task.).

5.4 Financial overconfidence

5.4.1 Overprecision

We begin our analysis of investor overconfidence with overprecision (or miscalibration), as it is constituent for financial models of overconfidence. In Odean (1998) traders overestimate the precision of a private signal they receive, which then renders their posteriors more precise than those of a well-calibrated trader. They consequently underestimate the volatility of assets. Empirically, confidence intervals are frequently used to elicit a measure for precision (e.g. Glaser and Weber, 2007; Deaves et al., 2010; Glaser et al., 2009). Our survey questions are designed similarly to those studies:

We would like you to make three estimates of the return of the UK stock market (FTSE all-share) by the end of the next three month.

- *Your best estimate should be your best guess.*
- *Your high estimate should very rarely be lower than the actual outcome of the FTSE all-share (about once in 20 occasions)*
- *Your low estimate should very rarely be higher than the actual outcome of the FTSE all-share (about once in 20 occasions)*

Please enter your response as a percentage change.

Participants are asked to submit a best estimate for the return of the UK stock market and, in an additional question, also for their own portfolios. Both estimates are made for a time horizon of three month to avoid overlapping observations. The best estimate is in each case followed by questions for a high and a low estimate, which together yield a 90%-confidence interval. Applying the method of Keefer and Bodily (1983), it is possible to back out implicit expected volatility from these confidence intervals. We prefer this indirect way of elicitation to direct volatility estimates, as it is simpler and more intuitive (cp. Dave et al., 2010).

Table 5.2 shows the width of predicted confidence intervals for stock market returns and portfolio returns. On average participants submitted slightly larger confidence intervals for own portfolios, but the median confidence interval is in both cases 20 percentage points wide. This translates into a volatility of 0.12 on a yearly basis. Given that people hold very different portfolios, it is no surprise that the standard deviation of confidence intervals is larger for portfolio returns. To assess whether these confidence intervals are well calibrated, we use historical volatility of the stock market and of investors' portfolios as a benchmark. For the market additionally option implied volatility is available. Our miscalibration measures are then defined as the difference between benchmark volatility and estimated volatility from confidence intervals.[4] The

[4]While Glaser et al. (2009) simply use one divided by estimated volatility as a measure for overprecision, but in a panel study a benchmark volatility is crucial, as the same estimated volatility can be appropriate or overprecise depending on the current market enviroment.

Table 5.2: Overprecision

	n	mean	median	Std.Dev.	% overprecise
Width confidence interval market	1957	23.3	20.0	18.6	
Width confidence interval portfolio	2002	24.0	20.0	22.9	
Miscalibration market (hist. vola)	1997	0.142	0.119	0.191	82.7%
Miscalibration market (impl. vola)	1997	0.150	0.151	0.155	89.0%
Miscalibration portfolio	1981	0.228	0.157	0.363	84.9%
Hit rate market	1957	0.45	0.00	0.49	
Hit rate portfolio	2002	0.49	0.00	0.50	

Notes: The table shows the width of confidence intervals (difference between high and low return estimate) in percentage points for UK stock market return and investors' own portfolio return. We exclude investors who violate the condition that high estimate=>low estimate. Miscalibration is historical (or implied) volatility devided by estimated volatility (from confidence intervals). Historical volatility is calculated over a one-year horizon, implied volatility is option implied volatility as represented by the FTSE VIX volatility index. Miscalibration measures are winsorized at a value of 10. Hitrate is a dummy that equals 1 if a confidence interval covered the realized value.

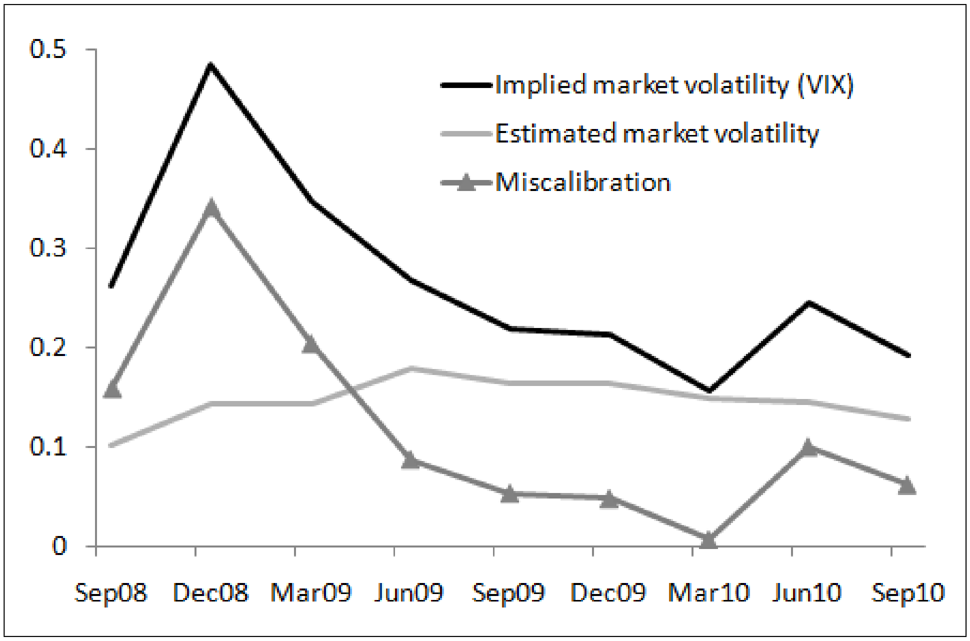
descriptive statistics in table 5.2 show that investors exhibit strong overprecision. Their volatility estimates fall short benchmark volatilities by on average 0.14 to 0.23 (median 0.11 to 0.16), meaning that their confidence intervals are too narrow by a factor 2 to 2.5. Miscalibration significantly exceeds the neutral value of zero ($p < 0.001$). Depending on the benchmark between 83% and 89% of all observations are overprecise (> 1) for the stock market and 85% for own portfolios.

The high levels of overprecision are due to the highly volatile markets during our survey period. Our benchmarks reach volatilities between 0.4 and 0.5, which implies that at certain times confidence intervals of up to 80 percentage points would have been appropriate. Figure 5.1 shows the development of miscalibration over time, and indeed it spikes with the high volatilities in the immediate financial crisis, as investors do not adjust confidence intervals sufficiently. However, there is overprecision for all rounds of the survey. While with different benchmarks the level of miscalibration slightly changes, the rank-order of participants is preserved.

Another way to assess calibration are hit rates, which should be close to 90% for 90% confidence intervals. We calculate hit rates in the cross section of investors (see table 5.2): Only about half of the time do confidence intervals cover the later realized value. This shows that miscalibration is not only present compared to some abstract benchmark, but also leads to investors frequently missing the real outcome by their estimates.[5] It has been argued that experts give more precise (and thus more informative) estimates, while hit rates remain constant (McKenzie, Liersch, and Yaniv, 2008). If this was the case, our miscalibration measure would punish these more knowledgeable investors. However, the correlation between miscalibration and hit rate is significantly negative (-0.37), suggesting that narrower confidence intervals are not more informative. In general, our results support H1 with regard to overprecision.

[5] In a strict sense the 90%-criterion applies only on an intra-participant basis, for which we have too few observations. We nevertheless calculate the personal hit rate for those who participated at least five times. The average individual hit rate is 0.48 and thus close to the overall hit rate.

Figure 5.1: Miscalibration over time



Notes: Average volatility estimates of survey participants and implied option volatility for the FTSE (right axis). Average level of miscalibration with implied volatility as a benchmark (left axis).

5.4.2 Overplacement

Measures for overplacement (or the better-than-average effect) involve a comparison to other investors. We ask survey participants for their agreement to the following statements:

1. I believe my investing skill is above average.

2.Compared to an average person, I am informed about current financial conditions.

3.Compared to an average person, I am informed about investing in general.

4. I have more experience with investing than the average person.

Responses are given on a seven-point Likert scale ranging from "strongly disagree" to "strongly agree". The question format of the first question is similar to a better-than-average question in Glaser and Weber (2007), while questions two and three are related to a better-than-average measure in Dorn and Huberman (2005) and Dorn and Sengmueller (2009). We add investing experience as it covers another facet of investing ability and knowledge. We admit that the comparison to the average person instead to other investors might inflate levels of overplacement, but this should be inconsequential for the cross-sectional distribution of overplacement, which is our main interest for the later analysis. Panel A of table 5.3 shows how investors rate themselves with regard to the statements. Participants believe that their investing skill is slightly above average with a mean of 4.4. Higher levels of overplacement are observed for being informed and being experienced (5.4 and 5.6). All three values are significantly above the neutral value of 4 ($p < 0.001$). While a moderate fraction of 45% thinks to possess above average skills, a large majority agrees to be better informed and more experienced than others. Correlations between the three evaluations are relatively high (0.5-0.7), Cronbach's alpha amounts to 0.75. We therefore combine the statements in a single overplacement measure by taking the average of the three ratings.

Due to time considerations, these statements were included only in the entry questionnaire of the survey and were not repeated in each survey round. If investors hold stable views of their skills and abilities (as argued for example by Larrick et al., 2007), we can still use them in our panel analysis. However, to generate time-varying measures of overplacement we will apply an alternative strategy. Graham et al. (2009) propose the difference between forecasts of own portfolio return and forecasts of the stock market return as a proxy for the better-than-average effect. The interpretation is straightforward: Market returns define the average an investor can expect to earn and any expected outperformance means to be better than this average. The above stated questions for best estimates of portfolio and stock market return allow us to calculate this expected outperformance. One may object that this measure does not take portfolio risk into account, and investors who take more financial risk are

Table 5.3: Overplacement and overestimation

PANEL A	n	mean	median	Std.Dev.	% > 4
Bta skill	617	4.40	4.00	1.28	45.2%
Bta information	617	5.55	5.50	1.00	88.3%
Bta experience	617	5.38	5.00	1.22	79.9%
Bta combined	617	5.11	5.00	0.98	85.1%
PANEL B	n	mean	median	Std.Dev.	% > 0
Expected outperformance	2106	0.029	0.000	0.086	47.8% (33%=0)
Sharpe ratio difference	1883	0.308	0.000	1.287	48.8% (27%=0)
Overestimation	2081	0.043	0.017	0.193	54.7%

Notes: Panel A shows responses to better-than-average statements (see text) on a 7-point Likert scale (1="strongly disagree" to 7="strongly agree"). Bta combined aggregates the three individual measures. Panel B: Expected outperformance is the difference between estimated portfolio return and estimated market return for the next three month. Sharpe ratio difference is the difference between expected portfolio Sharpe ratio and expected market Sharpe ratio (winsorized at 1%-level). Overestimation is the difference between expected portfolio return and realized portfolio return. Expected outperformance and overestimation are winsorized at 1%-level.

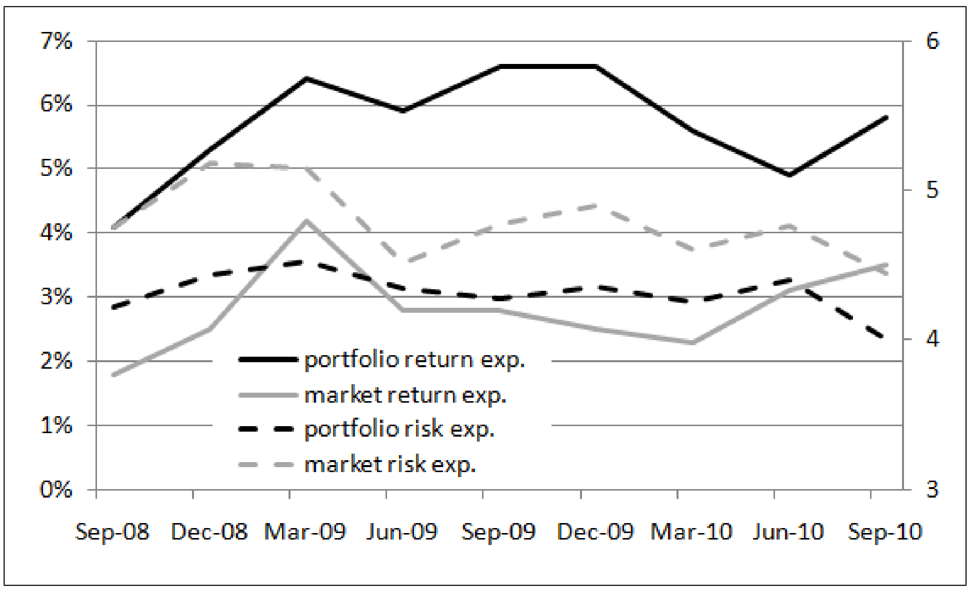
correct in expecting higher returns. To accommodate for this possibility we consider expected volatility of investors, both for their portfolio and the stock market. We construct an expected Sharpe ratio for portfolios and the stock market.[6] The difference between expected portfolio Sharpe ratio and expected market Sharpe ratio proxies for overplacement.

Panel B of table 5.3 displays descriptive statistic for these variables. On average investors expect to outperform the market by 2.9 percentage points ($p < 0.001$); given the time horizon of three month this amount appears large. About 48% of the time investors expect that their portfolios returns will exceed market returns, in 81% of all cases portfolio returns are expected to at least equal market returns. On a risk-return basis (Sharpe ratio) 49% of investors expect to beat the market, which is a large proportion given that in turn only 24% believe to be beaten by the

[6]We use the 3m-LIBOR as riskfree rate and set negative Sharpe ratios to zero, as these are not well defined.

market. The overplacement in portfolio expectations is perhaps best illustrated in figure 5.2: consistently over all survey rounds average portfolio return expectations are above average market return expectations (solid lines) and vice versa for risk expectations (dashed lines). Investors thus believe to achieve higher returns while taking less risk. This outperformance expectation in concert with the findings for skill-related statements confirms the presence of overplacement in our investor sample.

Figure 5.2: Return and risk expectations of investors



Notes: Average portfolio return expectations and market return expectations in % (left axis). Average risk perceptions for own portfolio and stock market on a seven-point scale (right axis).

5.4.3 Overestimation

While in finance there has been the casual observation that investors or managers are too optimistic about their return on investment (Dimson, Marsh, and Staunton, 2004; Heaton, 2002), a formal empirical treatment of overconfidence in the form of overestimation is still missing. The relationship between overestimation and optimism is obvious: those who are most optimistic are most likely to fall short of their expectations, in particular if performance involves a large chance element. Overestimation is defined as the difference between expected performance and actual performance (cp. Moore and Healy, 2008), in our case the difference between expected portfolio return and realized portfolio return. Even though it is possible to define this variable also for market returns, overestimation involves the notion of personal achievement, which is only present in portfolio returns.

On average investors in the panel overestimate their portfolio return by 4.3 percentage points (median 1.7, see table 5.3). While these figures are large and significant ($p < 0.001$), the proportion of overconfident investors is with 55% not overwhelmingly high. There is great heterogeneity with strong overestimation and strong underestimation both present in the data. Given the partly random nature of realized returns, one might suspect that effects in the time series drive this heterogeneity. For instance in the first round of the survey investors predict returns for Sep–Dec 2008, which turn out to be catastrophic due to the progress of the financial crisis. Indeed in this round overestimation reaches highs with an average of 17.9%-points and a proportion of 88% overconfident investors. However, within rounds the standard deviation is only slightly smaller than for the whole panel, indicating that also differences between individuals are considerable. We are thus confident that our measure of overestimation might contribute to explain investing behavior. We confirm H1 with respect to the presence of overestimation in the investor population.

As we have now collected evidence about all three types of overconfidence, we examine the relationship between these measures. Deaves, Lüders, and Luo (2009) report positive but low and not significant correlations between different types of overconfidence, while Glaser and Weber (2007) find even a negative correlation between overprecision and overplacement (also not significant). Table 5.4 shows a correlation matrix for the measures defined in this section. Miscalibration variables are highly correlated among each other (0.71) but show slightly negative correlations to other

Table 5.4: Correlations between overconfidence measures

	Misc. mk.	Misc. pf.	Bta	Exp. Outp.	Sharpe	Overest.
Miscalibration market	1					
Miscalibration portfolio	0.71***	1				
Bta combined	−0.14***	−0.16***	1			
Expected outperformance	−0.04*	−0.02	0.12***	1		
Sharpe ratio difference	−0.06***	0.09***	0.07	0.62***	1	
Overestimation	−0.09***	0.00	0.01	0.16***	0.15***	1

Notes: The table shows pairwise Spearman rank correlation coefficients between measures for overprecision (as defined in table 5.2), overplacement and overestimation (as defined in table 5.3). Miscalibration market using implied volatilies is not displayed as it is almost identical to miscalibration using historical volatility ($\rho = 0.98$). Correlations are significant at *10%-level, **5%-level, or ***1%-level

types of overconfidence. The agreement to better-than-average statements is positive, but lowly correlated to the expectation to outperform or beat the market. It seems that the two describe different aspects of overplacement and are only weakly related. As anticipated, expected outperformance and the expectation to beat the market on a risk-return basis (Sharpe ratio) are highly correlated (0.62). Overestimation is positively correlated to expected outperformance (0.16), as both measures share portfolio return expectations as one ingredient. But altogether we can confirm H1a in the sense that different types of overconfidence are at most very weakly related.

5.5 Investing behavior and overconfidence

We now bring together the two parts of our analysis: the investing behavior as portrayed in section 3 and the overconfidence measures as defined in section 4. Our goal is to systematically study their relationship, whereas previous literature has mostly examined only parts of this picture. Either the attention was limited to one aspect of investing behavior, or an incomplete typology of overconfidence was considered. Our main contribution is to fill this gap in a dynamic panel setting.

5.5.1 Trading activity

As excessive trading was most prominently related to overconfidence, we begin with trading activity measures as our dependent variables. We use a trade dummy, which equals one if an investor trades in the respective period, the number and volume of trades, and turnover (all variables as defined in table 5.1). We take the natural logarithm of number of trades, volume of trades, and turnover as these variables are strongly skewed. As explanatory variables, for overprecision we mostly employ miscalibration with respect to market returns (benchmark historical volatility). Due to the high correlation with the other miscalibration measures we abstain from including several of these variables simultaneously into the regression. However, our results are robust to the use of other specifications. For overplacement the previous analysis revealed that the combined better-than-average proxy and expected outperformance are only weakly related. We thus consider both in our regressions, the former as a time-invariant variable, as it was elicited only once. Occasionally we also report results for the expectation to beat the market on a risk-return basis (Sharpe ratio). Finally we include overestimation as defined in table 5.3.

Besides our overconfidence measures, we account for a set of further controls, in particular self-reported risk tolerance, which was identified by Dorn and Huberman (2005, p.437) as the “single most important determinant of both portfolio diversification and turnover”. We measure risk tolerance as agreement to the statement “It is likely I would invest a significant sum in a high risk investment” (on a seven-point scale). Weber et al. (2010) show that responses in this format are consistent to a more complete assessment of risk attitudes. Demographic variables include gender, age, and family status. We further control for income, wealth and financial literacy.

It was noted that on average about 70% of all investors trade within a three-month period. We thus first examine in a probit model the decision to engage actively in trading at all. Column (1) of table 5.5 shows the results of this regression. Participants who believe to be more skillful investors and to be more informed about financial markets have a higher propensity to become active in trading. In addition, men are more likely to trade, which contributes to earlier results of gender and trading activity (Barber and Odean, 2001; Deaves et al., 2009). The gender effect, while positive in almost all regressions, is at most marginally significant. This can be due to the fact that we explicitly control for factors gender might be a proxy for (such as risk tolerance or overconfidence), but could also be a consequence of the extreme composition of participants which is 93% male.

Column (2) to (5) report results for number of trades and trading volume. Both variables are far from normally distributed even after logarithmic transformation. Therefore we consider a Poisson regression as an alternative to GLS, yet results are similar. We find a strong impact of risk attitude; more risk tolerant investors trade more frequently and trade higher volumes. Among the overconfidence measures, expected outperformance has a positive influence. Investors who expect to beat the market return engage in more trades and also trade higher volumes. Not surprisingly wealthy investors place larger trades, which follows from no effect on number of trades but a significant effect on trading volume. Interestingly, investors with a higher number of dependents also trade more.

Finally, turnover represents the relation between trading volume and portfolio value. Column (6) of table 5.5 shows that investors who expect to outperform the market churn their portfolio more often. Male and more wealthy investors also exhibit

Table 5.5: Trading activity

	trade dummy	ln(# of trades)		ln(trading vol.)		ln(turnover)		
	Probit	GLS	Poisson	GLS	Poisson		GLS	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Risk tolerance	0.016	0.046***	0.028**	0.088***	0.015***	0.009	0.009	0.009
Miscalibration market	-0.450	-0.298	-0.084	-0.859**	-0.019	-0.081	-0.073	-0.070
Bta combined	0.181**	0.048	0.023	0.094	0.008	-0.005	0.001	-0.004
Exp. outperformance	0.083	0.535*	0.521**	1.184**	0.316***	0.357***		0.295***
Sharpe difference							0.012*	
Overestimation	0.208	-0.166	-0.055	-0.202	0.005	0.054	0.096*	0.016
Portfolio return exp.								0.128
Gender	0.525*	0.151	0.031	0.140	-0.012	0.070*	0.056	0.071*
Age	0.004	-0.001	0.000	0.000	0.000	-0.001	-0.001	-0.001
Couple	0.008	-0.007	-0.000	-0.258	-0.040	-0.004	-0.001	-0.001
Dependents	0.048	0.076*	0.035*	0.156**	0.013	0.024	0.025	0.024
Income	-0.024	-0.010	-0.011	0.047	0.004	0.005	0.004	0.005
Wealth	-0.012	-0.001	0.003	0.202***	0.026***	0.020**	0.019**	0.020**
Fin. lit.	0.025	-0.096	-0.037	0.168	0.017	0.015	0.017	0.018
constant	-0.669	1.875***	0.569**	6.632***	1.916***	0.043	0.025	0.022
n	1928	1480	1480	1480	1480	1363	1300	1363
R^2	—	0.036	—	0.125	—	0.093	0.079	0.095

Notes: The table shows panel regressions of trading activity variables on overconfidence measures and controls (including round dummies). Column (1) displayes a panel probit model, column (3) and (5) a panel poisson regression, and the remaining colums a general least squares model with random effects (standard errors are clustered by participant). Columns (2)-(8) are estimated for those who trade only. Dependent variables are a trade dummy (=1 if investor trades), logarithm of the number of trades, logarithm of trading volume and logarithm of turnover. Risk tolerance is measured on a seven-point scale, overconfidence measures are definded in section 4. Demographics: Gender dummy (male=1), age in years, couple dummy (married or co-habiting=1), number of dependents, income, wealth, financial literacy (number of correct answers using four of the basic literacy questions by van Rooij et al. (2011)). Income and wealth are measured in categories (for a full description see chapter 3), missing values are imputed. Coefficients are significant at *10%-level, **5%-level, or ***1%-level.

higher turnover.[7] In column (7) we replace expected outperformance by the related measure of beating the market on a risk-return basis. The effect is still positive and significant, but the Sharpe ratio difference proves to be less powerful in predicting trading activity. This result is confirmed when we introduce the variable in one of the other reported regressions. We conclude that the degree of overplacement is best reflected in the continuous measure provided by expected outperformance. One might argue that it is portfolio expectations driving our findings, as it is one main ingredient of expected outperformance. However, when we include portfolio return expectations as an additional control (column (8)), expected outperformance remains significant, while return expectations have no direct influence on turnover. Thus the difference between portfolio return expectations and market return expectations is crucial, not some general optimism for own portfolio returns.

We find in all regressions that a measure of overplacement (better-than-average measure or expected outperformance) exerts a positive effect on trading activity, while neither overprecision nor overestimation seems to be relevant. Quite intuitively investors who feel more skillful are more likely to trade, but for the degree of trading the more timely measure of expected outperfromance is relevant. By now there emerges a consensus that overplacement is the facet of overconfidence which is most predictive for trading activity (Glaser and Weber, 2007; Graham et al., 2009; Dorn and Huberman, 2005; Deaves et al., 2009, with the latter finding also a role for miscalibration). We confirm hypothesis 2 that overconfidence has a positive effect on trading activity.

For robustness we consider several alternative specifications. Given that we disentangle the decision to trade at all and trading volume, a Heckman selection approach might help to understand this two-step process. However, the inverse Mills ratio of

[7]The latter result is in contrast to previous literature (e.g. Dorn and Huberman, 2005); however, the effect turns around if we use portfolio value as a wealth proxy (similar to Glaser and Weber, 2007). Our interpretation thus is that especially investors with large outside wealth churn their portfolios often.

the participation regression remains insignificant in the second stage suggesting that selection is not severe. We also test a fixed effects model and find coefficients for risk attitude and overconfidence much reduced and mostly not significant. This is not surprising as these variables are at least in part stable individual characteristics. Still the diminished coefficients provide some evidence that part of the effect comes from the time series. We also estimate a static model with the average of turnover through our survey period as dependent variable. Again risk attitude and overplacement are the most important determinants of trading activity.

5.5.2 Diversification

With number of portfolio positions, a Herfindahl-Hirschman-Index of portfolio weights, and normalized variance, we introduced three indicators of portfolio diversification. For the regression analysis we take the natural logarithm of positions and HHI as these variables are skewed. As the measures represent different aspects of diversification we try to distill a common component by performing a principal component analysis (PCA). The first principal component of this PCA captures about two thirds of the variance and is highly correlated to the other diversification measures. We thus consider it as an additional diversification proxy. To facilitate the interpretation of our results, it is worth noting that diversification decreases with higher values of HHI and normalized variance, while it increases with number of portfolio positions and the PCA-measure.

The set of independent variables remains the same as in our analysis of trading activity. One exception is that we complement miscalibration with respect to market return by miscalibration of portfolio returns, as diversification is a matter of how the own portfolio is perceived. Columns (1),(3),(5) and (7) of table 5.6 show our baseline regression for the various dependent variables. Risk tolerant investors tend to hold

Table 5.6: Diversification

	ln(positions)		ln(HHI)		norm. variance		diversification (PCA)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Risk tolerance	0.016***	0.016***	-0.007	-0.010	-0.006*	-0.006*	0.023**	0.026**	0.028**
Miscalibration portfolio	-0.105**		0.201***		0.134***		-0.391***		-0.217***
Miscalibration market		0.018		-0.284**		0.078*		0.043	
Bta combined	-0.007	-0.014	-0.015	-0.015	0.006	0.005	-0.005	0.001	
Exp. outperformance	-0.020	-0.031	0.223	0.236	-0.068	-0.046	0.037	-0.029	0.191
Overestimation	-0.091	-0.058	0.240***	0.227***	0.067**	0.051*	-0.271**	-0.216*	-0.174**
Gender	0.138	0.125	-0.179	-0.153	0.002	0.009	0.165	0.132	
Age	0.014***	0.013***	-0.013***	-0.012***	-0.002**	-0.002**	0.020***	0.020***	
Couple	-0.085	-0.093	0.178*	0.162	0.039	0.027	-0.213	-0.184	
Dependents	-0.021	-0.016	0.025	0.023	-0.004	-0.005	-0.028	-0.025	
Income	0.002	0.000	-0.065*	-0.067*	0.003	0.004	0.033	0.037	
Wealth	0.080***	0.086***	-0.030	-0.035	-0.005	-0.007	0.078**	0.085***	
Fin. lit.	0.022	0.036	-0.030	-0.056	0.015	0.011	0.002	0.027	
constant	1.131***	1.069***	-0.747**	-0.542	0.573***	0.620***	-1.530***	-1.806***	0.004
n	1941	1907	1939	1904	1934	1883	1934	1883	1934
R^2	0.127	0.109	0.081	0.048	0.122#	0.086#	0.146	0.099	0.068

Notes: The table shows panel regressions of diversification variables on overconfidence measures and controls (including round dummies). Regressions are GLS, column (1)-(4), (7), (8) with random effects and clustered standard errors by participant, column (9) with fixed effects. Columns (5) and (6) are tobit regressions as normalized volatility is bounded at 1 ($^{\#}R^2$ from linear model). Dependent variables are logarithm of the number of portfolio positions, logarithm of a Herfindahl-Hirschmann-index of portfolio weights, normalized variance, and the first principal component of a PCA involving the three other variables. Risk tolerance is measured on a seven-point scale, overconfidence measures are definded in section 4. Demographics: Gender dummy (male=1), age in years, couple dummy (married or co-habiting=1), number of dependents, income, wealth, financial literacy. Coefficients are significant at *10%-level, **5%-level, or ***1%-level.

more positions, and are also better diversified with respect to the other measures (although not significantly so for HHI). This result stands in contrast to the finding of Dorn and Huberman (2005) that risk averse investors hold better diversified portfolios. An explanation could be that in a panel setting rounds of high trading activity often coincide with a more diverse inventory of assets, a connection that is much weaker for a larger time span. Indeed, when we run a static regression on average portfolio positions, the effect turns around.

We observe a strong effect of overprecision in portfolio return expectations on all diversification variables. Overprecision reduces the number of portfolio positions, increases HHI and normalized variance, and decreases the PCA-proxy of diversification. The intuition behind this effect is that investors who hold overly precise beliefs about future portfolio returns are unaware of the benefits of diversification, which mainly consist of reducing non-systematic portfolio risk and hereby narrowing the distribution of potential outcomes. However, it is the very nature of overprecision that investors underestimate the width of this distribution and are thus likely to underdiversify. Deaves et al. (2009) argue that calibration measures must be highly specific to explain investing behavior. Indeed, when we replace miscalibration of portfolio returns by miscalibration of market returns, the results differ considerably (see columns (2),(4),(6) and (8)). First of all the, explanatory power of the regressions goes down as becomes evident in R^2, but also the coefficients for miscalibration turn partly insignificant and sometimes change direction. It is obvious that underestimating stock market volatility might not have as direct consequences for portfolio diversification as underestimating portfolio volatility.

As a second overconfidence variable, overestimation contributes to portfolio underdiversification. The sign of coefficients is consistent for all diversification measures, implying that investors tend to hold fewer positions, less diverse portfolio weights (HHI), and higher normalized variance; the effect is significant for HHI, normalized variance, and diversification from PCA. With this finding the picture becomes more complete; investors who believe to be able to identify few very profitable stocks will also neglect diversification. Hypothesis 2 thus holds for diversification, with overprecision and overestimation to increase underdiversification.

Among the demographic variables we identify two effects: Older investors tend to hold better diversified portfolios, as do wealthier investors. For wealth the relation is particularly strong for number of portfolio positions, which is not surprising as wealthy investors hold larger portfolios. The weaker result for other diversification measures leaves open, whether the mechanical accumulation of assets or a genuine understanding of diversification drives this influence. With age it remains unclear, whether older investors become indeed wiser or if they just through the course of their lives have collected a larger number of portfolio holdings. Similar age and wealth effects on diversification are observed by Goetzmann and Kumar (2008) and Dorn and Huberman (2005).

Column (9) of table 5.6 displays a fixed effects regression of diversification on risk attitude and overconfidence (all time-invariant variables are dropped). It confirms the finding that miscalibration and overestimation decrease diversification, even under the restriction that the individual fixed effect will pick up stable personality characteristics. While the effect size is reduced, fluctuations of overconfidence over time seem to be important for the degree of diversification. For robustness we again perform static regressions with the average of diversification over the survey period as dependent variable. As mentioned above, the somewhat counterintuitive effect of risk attitude disappears in these regressions. Our conclusions with regard to overprecision and overestimation and their impact on diversification however remain valid in a static framework. A caveat to our findings is that investors with higher portfolio volatilities have a harder time to be well calibrated, and of course those who are less diversified tend to own more volatile portfolios.

5.5.3 Risk taking

Compared to other aspects of investing behavior, the relationship between overconfidence and risk taking has found less attention in finance literature. Using the same dataset, Weber et al. (2010) and chapter 2 identified determinants of risk taking behavior. They find a strong influence of return and risk expectations on portfolio risk and hypothetical risk taking. As our overconfidence variables are mainly based on expectations, we need to carefully control for these effects. Dependent variables are portfolio volatility, relative volatility, average component volatility (ACV), portfolio beta, and fraction of risky investment in the hypothetical risk taking task. We take the natural logarithm of volatility and ACV, as they are skewed.

The analysis of portfolio risk is marred by a potential reverse causality problem, as investors who take higher risk will as a consequence expect higher portfolio returns and higher portfolio risk. Therefore using portfolio expectations to explain risk taking can easily lead to wrong conclusions. We instead include market expectations into the regressions, which should be independent of own portfolio holdings (cp. chaper 2). A similar strategy can be applied to overprecision, instead of portfolio return miscalibration we rely on market return miscalibration as explanatory variable. Expected outperformance as an overplacement measure is based on the difference between portfolio and market expectations. As argued however, investors who take higher portfolio risk reasonably expect higher portfolio returns. But they have no reason to anticipate higher Sharpe ratios as this variable takes portfolio risk into account.

Table 5.7 displays regression results of risk taking behavior on overconfidence and controls; columns (1)-(4) report our findings for actual investing behavior. Risk taking increases for higher return expectations, but there is little effect of risk attitude and risk expectations (cp. chaper 2). Among the overconfidence measures we identify some influence of miscalibration and overplacement. In both cases overconfidence

Table 5.7: Risk taking

	ln(volatility)	rel. vola	ln(ACV)	beta	risk task	
	(1)	(2)	(3)	(4)	(5)	(6)
Risk tolerance	0.000	-0.001	0.016***	-0.003	0.021***	0.021***
Market return exp.	0.207***	0.369***	0.162***	0.101**	0.315***	0.332***
Market risk exp.	-0.000	-0.003	-0.009*	-0.003	-0.034***	-0.034***
Miscalibration market	0.123***	0.148*	0.072	0.016	0.161***	0.150***
Bta combined	-0.006	0.020	-0.026	0.013	0.036***	0.035***
Sharpe ratio difference	0.011**	0.028***	0.004	0.009*	0.007	
Exp. outperformance						0.088
Overestimation	0.028	0.064	0.010	-0.055	0.039	0.038
Gender	0.106*	0.164*	0.075	0.108*	0.104**	0.109**
Age	-0.006***	-0.009***	-0.003*	-0.002	-0.000	0.000
Couple	-0.026	-0.059	-0.049	-0.032	-0.022	-0.023
Dependents	0.006	0.007	0.017	0.011	0.019**	0.016*
Income	-0.020*	-0.027	-0.022*	-0.007	-0.002	-0.002
Wealth	-0.031***	-0.050***	-0.032***	-0.018**	0.003	0.003
Fin. lit.	-0.056**	-0.125**	-0.053*	-0.053	0.006	0.001
constant	-0.604***	2.432***	-0.333*	1.117***	0.297***	0.306***
n	1815	1817	1793	1817	1835	1930
R^2	0.388	0.129	0.243	0.042	—	—

Notes: The table shows panel regressions of risk taking variables on overconfidence measures and controls (including round dummies). Columns (1)-(4) report GLS-models with random effects (standard errors are clustered by participant), columns (5)-(6) a panel tobit regression. Dependent variables are logarithm of portfolio volatility (1y-horizon), relative volatility, logarithm of average component volatility(ACV), portfolio beta, and fraction of risky investment in the risk taking task. Risk tolerance and market risk expectation are measured on a seven-point scale, market return expectations in %, overconfidence measures are as definded in section 4. Demographics: Gender dummy (male=1), age in years, couple dummy (married or co-habiting=1), number of dependents, income, wealth, financial literacy. Coefficients are significant at *10%-level, **5%-level, or ***1%-level.

contributes to increased risk taking, participants who underestimate the range of possible market returns and those who expect to beat the market on a risk adjusted basis invest more riskily. While the strength of this result appears not to be overwhelming, one has to take into account that we are forced to use weaker overconfidence proxies. The coefficients for demographic variables show that older and wealthier investors take less risk, while male and financially less sophisticated participants invest more riskily.

The hypothetical risk taking task described in section 2 provides an opportunity to analyze determinants of risk taking in a quasi-experimental setting. Advantages are that results in this task are far less noisy than actual investing, and that aforementioned identification problems do not apply. Columns (5) and (6) show that the impact of expectations and risk attitude are much more distinct for the hypothetical decision, which we attribute partly to the absence of noise but also to the fact that the fraction of risky investment is chosen anew in each survey round, while actual portfolios only change gradually over time. Therefore participants adjust their hypothetical decisions much more to current situational circumstances. We observe a positive influence of miscalibration on risk taking: those who underestimate the variety of probable market returns invest more into the stock market. Additionally those, who believe to be more skillful or better informed, invest more. In contrast to real investing, we find no effect for the overplacement measures based on expected outperformance and Sharpe ratio. This is intuitive, as these variables are related to properties of investors own portfolios, but the risk taking task is not. While the positive impact of gender can be observed for the hypothetical task as well, demographics such as age or wealth are less relevant; a reason could be that in the task the amount to invest is the same for all participants, and they start off from a level situation.

As an overall finding from actual and hypothetical risk taking, we note that miscalibration is positively related to risk taking. In an experimental setting, Nosić and

Weber (2010) come to the same conclusion. But unlike overtrading and underdiversification, higher risk taking per se is not an investing bias. Given the miscalibration in the assessment of possible outcomes however, it seems that in particular those who invest more riskily are not fully aware of the potential consequences. They then would invest too much given their risk preferences, which is why we conclude that overconfidence might lead to excessive risk taking. Besides miscalibration, some form of overplacement contributes to risk taking, we thus confirm hypothesis 2. The results are robust to several alternative specifications, for example replacing the tobit model for the risk taking task by a linear model. Again a static model relying on averages of risk taking and controls confirms our findings.

5.6 Dynamics of overconfidence

5.6.1 Success and failure

Given the impact of different types of overconfidence on various aspects of investing behavior, it is important to understand the dynamic development of overconfidence over time. Our hypothesis 3 suggests a role of past outcomes, in particular success, on future overconfidence. We will define success (and failure) differently for overprecision as opposed to overplacement and overestimation. Overprecision is linked to estimation success as people have to predict confidence intervals for future market and portfolio returns and may adjust intervals in response to outcomes conflicting or justifying their prediction. In contrast, overplacement and overestimation are based on evaluations of relative or absolute performance and thus may be bolstered by past investment success.

Estimation success can be easily defined by confidence intervals that cover the realized value. We compare the group of investors with successful estimates to those who missed the true value by their confidence interval. From a Bayesian point of view it is unclear whether there should be an adjustment at all, as for any confidence interval with a probability of less than one, both outcomes will occur with a certain frequency. In our questions for 90%-confidence intervals, a miss definitely is a stronger signal than a hit. Table 5.8 displays how participants react to realizations that fall inside or outside their confidence intervals. A clear pattern emerges: those who missed the realized value in the previous round enlarge their confidence intervals by about 5%-points, while those who covered the true value narrow their confidence intervals by 3%-points. These changes are large given that the median confidence interval is about 20%-points wide. The proportion of people who increase their confidence interval is also much larger for those who missed the realized value; all differences are significant at 1%-level. These changes in confidence intervals have direct consequences for miscalibration, investors who previously estimated the range of returns successfully become more overprecise in the subsequent round.

Table 5.8: Estimation success and overprecision

	realization within confidence interval	realization outside confidence interval	difference
Δ CI market return	-0.033	0.047	-0.079***
Proportion of increase	0.341	0.598	-0.256***
Δ CI portfolio return	-0.032	0.056	-0.088***
Proportion of increase	0.347	0.589	-0.242***
Δ Miscalibration market	0.022	-0.017	0.039***
Δ Miscalibration portfolio	0.032	-0.001	0.033**

Notes: The table displays changes of confidence intervals and changes of miscalibration for two groups of investors: those who covered and those who missed the realized value by their confidence interval estimates in the previous round. Miscalibration (benchmark historical volatility) is demeaned by round to eliminate round effects. Proportion of increase shows for both groups the fraction of investors who increased their confidence interval subsequently. Differences between the group are significant at *10%-level, **5%-level, or ***1%-level by a standard t-test and a Wilcoxon rank-sum test.

We test this conjecture in a multivariate framework proposed by Deaves et al. (2010). They regress changes in confidence intervals on a dummy variable that indicates whether an investor was previously correct or not and control for changes in market return and market volatility. We expect the coefficient for the dummy variable to be negative, as the univariate statistics suggested that investors decrease confidence intervals after correct estimates. Column (1) of table 5.9 with the dummy as the sole independent variable yields the same result. Interestingly the coefficients (constant and dummy coefficient) are close to the values observed by Deaves et al. (2010).[8] Suggested control variables (column 2) do not change the negative relationship between correct estimates and subsequent confidence intervals, we further observe that after positive past returns investors tend to decrease intervals. It seems that the results so far unambiguously support hypothesis 3. However, given that investors might not fully recall their previous estimates, might not be aware of the realized value, and thus may only have a very vague idea of estimation success or failure, the findings are surprisingly clear-cut. A caveat to this interpretation is that width of confidence intervals and estimation success are not independent. Investors with confidence intervals that cover the realized value have on average submitted wider intervals. A regression toward the mean would predict narrower intervals subsequently. To control for this effect we add lagged width of confidence intervals to the regression (column 3). Lagged width has a strong negative impact on change of confidence intervals, but more importantly the influence of the estimation success dummy turns around. This raises some doubts about the above findings and the results of Deaves et al. (2010), as a regression toward the mean effect provides an alternative explanation.

Compared to estimation success in a survey task, actual investment success should be much more salient to participants. We again divide participants into two groups, those who experienced portfolio returns that exceeded market returns in the past three

[8] We choose a pooled regression here to most closely match the approach of Deaves et al. (2010). However, a panel fixed effects regression produces very similar results.

Table 5.9: Determinants of changes in confidence intervals

	Change confidence interval		
	(1)	(2)	(3)
Success dummy (correct=1)	-0.079***	-0.069***	0.033*
Past market return		-0.150**	-0.008
Δ Market volatility		-0.019	0.029
Lagged width of CI			-0.633***
constant	0.047***	0.039***	0.151***
n	1041	1041	1041
R^2	0.037	0.044	0.319

Notes: The table shows a pooled regression of changes in estimated confidence intervals for market returns on a success dummy and controls. The success dummy equals one, if the previous confidence interval covered the realized value. Past market return is the return of the FTSE all-share index between the previous and current survey round. Δ market volatility is the change in implied market volatility of the FTSE for the same time period. Lagged width of confidence interval is the estimated width of confidence interval in the previous period. Standard errors are clustered by participant. Coefficients are significant at *10%-level, **5%-level, or ***1%-level.

month (survey period) and those with portfolio returns lower than market returns. According to the learning-to-be-overconfident hypothesis those who outperformed the market are expected to become more overconfident (Gervais and Odean, 2001). Table 5.10 shows levels and changes of portfolio expectations and overconfidence measures for the two groups of investors. Successful investors expect higher portfolio returns in the future (7% vs. 6%), and they increase their return expectations compared to the previous round more strongly. This gives rise to more pronounced overplacement: levels and changes of expected outperformance and Sharpe ratio difference are always higher for previously successful investors. However, differences are in general not large and only marginally significant, for instance successful investors expect to outperform the market by 3.1% and unsuccessful investors by 2.7%. While we observe higher levels of overestimation among the previously unsuccessful, overestimation does indeed decrease more strongly for this group.

Table 5.10: Investment success and overconfidence

	portfolio return > market return	portfolio return < market return	difference
Portfolio return expectations	0.070	0.060	0.010$^{***/**}$
Δ Portfolio return exp.	0.013	0.004	0.009$^{**/-}$
Expected outperformance	0.031	0.027	0.004$^{*/-}$
Δ Expected outperformance	0.005	0.000	0.005
Sharpe ratio difference	0.347	0.256	0.091$^{-/*}$
Δ Sarpe ratio difference	0.086	0.037	0.049
Overestimation	-0.009	0.017	-0.027$^{***/***}$
Δ Overestimation	-0.012	-0.063	0.051$^{-/***}$

Notes: The table displays expectations, overconfidence measures and changes in these variables for two groups of investors: those who outperformed and those who underperformed the market in the previous round. Portfolio return expectations are three-month return estimates for own portfolios, overconfidence variables as defined in table 5.3. Differences between the group are significant at *10%-level, **5%-level, or ***1%-level by a Wilcoxon rank-sum test / standard t-test.

We draw on the model introduced before and regress changes in overplacement and overestimation on investment success and controls. For investment success we use either a dummy variable that indicates whether an investor outperformed the market in the past three months, or a continuous version, which is the difference between past portfolio returns and market returns. Table 5.11 reports the results of these regressions. Columns (1)-(4) show that we do not identify an effect of investment success on subsequent overplacement, the explanatory variables even fail to attain joint significance by an F-test. A different picture emerges for changes in overestimation (columns 5–6), investment success encourages participants to overestimate future portfolio returns more strongly; past market returns contribute to this tendency. As past levels of overestimation are not independent of past investment success (as with lower portfolio returns it is more likely to have overestimated returns), a similar concern applies as to changes in confidence intervals. However, when we introduced lagged levels of overestimation in the regression (column 7), the influence of investment success remains robust. A regression toward the mean effect might thus partake in the explanation, but not invalidate the role of previous investment success. Altogether we can only partly confirm hypothesis 3, investment success leads to increased overestimation but not to increased overplacement, and estimation success leads to higher overprecision, which may however be a result of a regression toward the mean effect.

Table 5.11: Determinants of changes in overplacement and overestimation

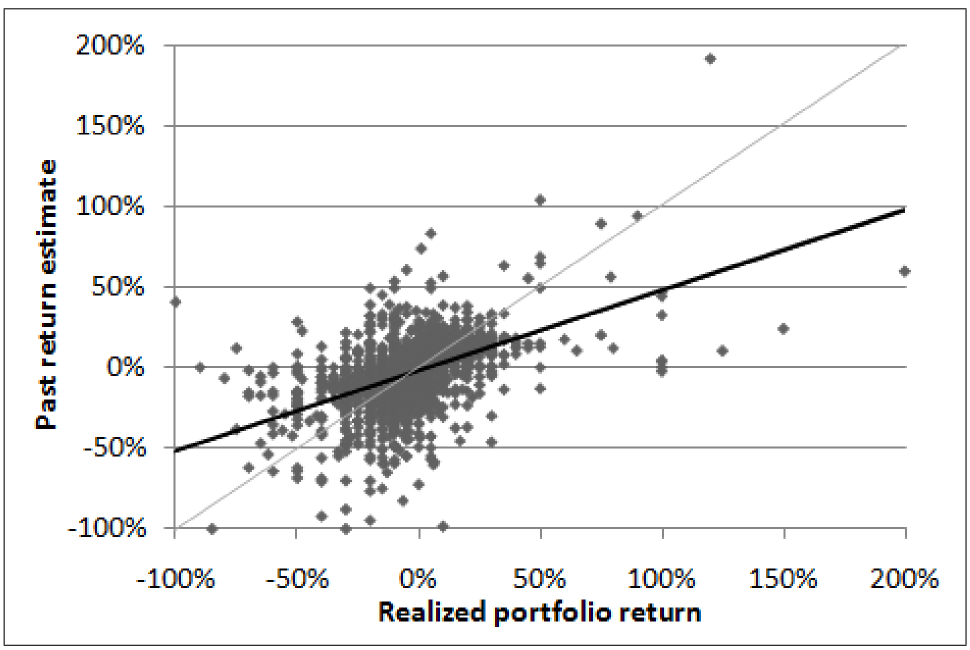
	Δ exp. outperf.		Δ Sharpe diff.		Δ overestimation		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Success dummy	0.007		0.052		0.161***		
Past outperformance		0.026		-0.054		1.024***	0.473***
Past market return	0.017	0.021	0.145	-0.008	0.804***	1.181***	0.667***
Δ Market volatility	0.001	0.001	0.098	0.063	-0.052*	0.034	0.040**
Lagged overestimation							-0.563***
constant	-0.001	0.002	0.038	0.071*	-0.115***	-0.056***	-0.022***
n	1192	1192	959	959	1194	1194	1194
R^2	0.001	0.001	0.000	0.000	0.301	0.503	0.571

Notes: The table shows pooled regressions of changes in overplacement and overestimation on investment success variables and controls. The success dummy equals one, if the previous portfolio returns exceeded market returns. Past outperformance is the difference between past portfolio return and past market return in %-points. Past market return is the return of the FTSE all-share index between the previous and current survey round. Δ market volatility is the change in implied market volatility of the FTSE for the same time period. Lagged overestimation is the level of individual overestimation in the previous period. Standard errors are clustered by participant. Coefficients are significant at *10%-level, **5%-level, or ***1%-level.

5.6.2 Hindsight

One reason for the mixed relevance of previous outcomes for subsequent overconfidence might be that investors are not fully aware of the values of these outcomes. The survey does not provide them with information on past returns, and investors beliefs in retrospect might differ from the returns we calculate. Therefore a survey questions asks investors explicitly for past market and portfolio returns. We can thus compare investors' hindsight on returns and the actual values, figure 5.3 plots post-estimates against realized portfolio returns. The relationship is positive, we find a Pearson-correlation of 0.50 for portfolio returns and of even 0.68 for market returns. This is contrary to Glaser and Weber (2007b), who report no correlation between hindsight portfolio return estimates and realized values. A possible explanation is that during our survey period attention to portfolio returns is amplified due to the financial crisis. Moreover, Glaser and Weber (2007b) use a time period of three years, which renders estimation harder.

Figure 5.3: Estimates and actual values for past portfolio return



Notes: The figure plots realized portfolio returns against (post-)estimates of portfolio return, including linear fitted values and 45-degree line.

But figure 5.3 also shows large deviations from the 45-degree line, suggesting severe estimation errors. We calculate the difference between estimates and realized values, which indicates whether investors systematically over- or underestimate their own past performance, and the absolute value of this difference, which signifies the magnitude of estimation error. Panel A of table 5.12 shows that investors overestimate past market returns by on average 2.4%-points ($p < 0.01$), but show no systematic bias for own past portfolio performance. Median estimation errors are large, 5.3%-points for market returns and even 8.5%-points for portfolio returns. However, one has to take into account that in the volatile market environment at the time of our survey pronounced changes in returns occur within days, which exacerbates errors if investors recall a time window only slightly departing from the three months used for calculation of realized returns.[9]

For a relative assessment of past performance we pose the following question to investors in each round of the survey:

Compared to the other investors, how well do you think your portfolio held with us performed in the past three months?

Responses are given on a seven-point scale ranging from "much worse" to "much better". Panel B of table 5.12 displays results for this evaluation of relative performance, which are on average slightly below the middle point of the scale suggesting that investors believe, their returns were somewhat worse than the returns of other participants. When we now sort investors' realized returns into seven quantiles (corresponding to the seven-point scale) and then subtract this realization from the relative performance evaluation, we obtain a measure for overplacement in hindsight, which indicates whether investors think that they performed better than they really did (in relative terms). On average there is slight underplacement, which directly follows from the low relative performance estimates. An explanation for this underplacement might be poor portfolio returns during the financial crisis. Of course other investors

[9]A source of estimation error participants are not responsible for are potential errors in calculation of realized returns. We use price data from Thomson Reuters Datastream which covers more than 90% of portfolio holdings and fill in last observed transaction price for securities without available price data. This should guarantee fairly accurate portfolio return calculations.

Table 5.12: Estimation of past returns

Panel A	n	Mean	Median	Std.Dev.	5th Perc.	95th Perc.
Overestimation mk (hindsight)	2108	0.024	0.012	0.124	-0.131	0.224
Overestimation pf (hindsight)	2088	0.005	0.001	0.197	-0.298	0.289
Absolute error market	2108	0.077	0.053	0.074	0.004	0.223
Absolute error portfolio	2088	0.128	0.085	0.684	0.008	0.413
Panel B	**n**	**Mean**	**Median**	**Std.Dev.**	**5th Perc.**	**95th Perc.**
Relative performance	2101	3.723	4.000	1.250	2.000	6.000
Overplacement (hindsight)	2074	-0.238	0.000	2.116	-4.000	3.000
Outperformance (hindsight)	2104	-0.009	0.000	0.133	-0.25	0.19

Notes: Panel A displays overestimation in hindsight, the difference between estimates of past return and the realized values for the UK stock market (mk) and own portfolios (pf), both winsorized at 1%-level. Absolute error is the absolute value of this difference. Panel B shows relative past performance judgments of investors, the difference of this judgments and realized perfomance (quantile-sorted), and the numeric difference between estimates for past portfolio return and past market return.

struggle with bad outcomes as well, but it has been shown that egocentrism is present in relative evaluations (Kruger, 1999). This means that individuals rely too heavily on their own performance neglecting the performance of others. To test this conjecture, we regress evaluations of past relative performance on the estimates for past portfolio return and past market return (results not reported). Indeed, relative performance is positively related to own performance and negatively to market performance, but the effect of own performance is about three times as large as the effect of market returns. Thus poor portfolio returns in conjunction with egocentrism may explain underplacement. We can construct a related overplacement measure also from the numeric estimates of past performance discussed above. The difference between estimates for past portfolio return and past market return represents investors' perceived outperformance. On average there is a belief of having underperformed the market by about 1%-point ($p < 0.01$). If survey participants are consistent, the degree of this perceived past outperformance should be closely related to judgments of past relative performance; we find a positive correlation of 0.40.

We have now gathered several measures of investors' perceptions of past performance as opposed to actual performance data used in the previous section. Hypothesis 4 claims that these perceptions in hindsight will induce overconfidence in foresight. We thus split investors into those who overestimate their past portfolio returns and those who do not, and compare the levels and changes in overconfidence variables for the two groups (see table 5.13). We find that investors who overestimate their past performance not only exhibit higher levels of overconfidence, but also increase these levels more strongly (except for overestimation). We consider an alternative partition based on outperformance in hindsight; again those who believe in having outperformed the market in the past are increasingly overconfident in the future. While in this case the differences mostly do not attain significance, the always posi-

Table 5.13: Perceived past performance and overconfidence

	Overestimation pf (hindsight)			Outperformance (hindsight)		
	> 0	< 0	difference	> 0	< 0	difference
Portfolio return exp.	0.065	0.065	$0.000^{*/-}$	0.073	0.066	0.007
Δ Portfolio return exp.	0.020	0.002	$0.018^{-/**}$	0.011	0.006	0.005
Expected outperformance	0.030	0.027	0.003	0.035	0.030	0.006
Δ Expected outperf.	0.011	-0.002	$0.013^{**/**}$	0.004	0.000	0.004
Sharpe ratio difference	0.326	0.279	0.046	0.373	0.305	$0.068^{*/-}$
Δ Sharpe ratio diff.	0.256	-0.053	$0.309^{**/***}$	0.098	0.000	0.098
Overestimation	0.066	0.018	0.048	0.056	0.035	0.021
Δ Overestimation	-0.069	-0.006	$-0.064^{***/***}$	-0.010	-0.064	$0.053^{**/***}$

Notes: The table displays expectations, overconfidence measures and changes in these variables for two partitions of investors: those who overestimate vs. those who underestimate past portfolio returns, and those who believe in having outperformed the market vs. those who believe in having underperformed the market. Portfolio return expectations are three-month return estimates for own portfolios, overconfidence variables as defined in table 5.3. Differences between the group are significant at *10%-level, **5%-level, or ***1%-level by a Wilcoxon rank-sum test / standard t-test.

tive sign still suggests some robustness of the effect. Differences get even larger if we consider only those investors who believe in having outperformed the market but in reality did not. These investors suffer under a form of hindsight bias that makes them particularly prone to become overconfident. Interestingly, the qualitative evaluations of past relative performance and overplacement do not produce the same pattern of persistently larger levels and changes in overconfidence.

In a multivariate setting we introduce overestimation and outperformance in hindsight as additional explanatory variables for changes in overconfidence. Table 5.14 shows results, which build on the regression used to test for the influence of actual investment success before. We find a strong impact of perceived past performance on changes in overplacement, and also a marginally significant influence on changes in overestimation. Investors, who believe in having outperformed the market or who

Table 5.14: Actual and perceived success as determinants of changes in overconfidence

	Δ exp. outperf.		Δ Sharpe diff.		Δ overestimation	
	(1)	(2)	(3)	(4)	(5)	(6)
Past outperformance	-0.041	0.036	-0.953	0.371	0.957***	1.022***
Past market return	-0.008	-0.023	-0.376	-0.584	1.164***	1.159***
Δ Market volatility	0.004	-0.010	0.072	-0.160	0.023	0.014
Outperformance (hindsight)	0.088***		1.828***		0.100*	
Overestimation (hindsight)		0.091***		1.525***		0.067
constant	0.006*	0.004	0.130*	0.089	-0.054***	-0.057***
n	1188	1189	956	957	1189	1190
R^2	0.001	0.003	0.001	0.007	0.488	0.489

Notes: The table shows fixed effects panel regressions of changes in overplacement and overestimation on actual and perceived investment success and controls. Past outperformance is the difference between past portfolio return and past market return in %-points. Past market return is the return of the FTSE all-share index between the previous and current survey round. Δ market volatility is the change in implied market volatility of the FTSE for the same time period. Outperformance (hindsight) is the difference between estimated past portfolio return and estimated past market return. Overestimation (hindsight) is the difference between estimated past portfolio return and realizedportfolio return. Coefficients are significant at *10%-level, **5%-level, or ***1%-level.

overestimate their past investment success, will become increasingly overconfident. We still control for actual success by past outperformance, which is again only relevant for changes in overestimation (results for the success dummy variable are similar). In many situations it might be more important for expectations of investors how successful they think to be than how suscessful they really are. We confirm the relevance of hindsight effects for overconfidence, which supports hypothesis 4. Results are mostly robust to alternative specifications such as a pooled or random effects regression or the addition of lagged overconfidence values.

5.6.3 Dynamics of overconfidence and investing

The presented results bear repercussions for investing behavior, as we before identified a distinct impact of overconfidence on trading activity, diversification and risk taking of investors. If overconfidence is fueled by past actual or perceived investment success, this will induce higher turnover, more concentrated portfolios, and increased risk taking. To directly track this dynamic relationship over time we estimate regressions of changes in investing behavior on changes in overconfidence measures (in an approach similar to chapter 2). Unreported results of these regressions mostly show no significant effects of changes in overconfidence on changes in investing behavior. However, such effects are notoriously hard to identify, as firstly the changes regressions require participation of investors in two consecutive survey rounds which cuts down the number of observations by one half and thus substantially lowers statistical power. Secondly, changes in overconfidence may be subtle as they are derived from the interaction of several expectation variables. Thirdly, and perhaps most importantly, changes regressions are particular susceptible to noise, both in the survey responses and in timing and execution of financial market transactions. Some of the examined dependent variables even change without deliberate activity of investors (e.g. HHI or portfolio volatility). Levels regressions are more robust against these effects, as the variable magnitude is comparibly large and the rank-order of participants is in general preserved. This is why we draw upon the findings of the previous section and assume that this is the channel through which dynamics in overconfidence translates into investing behavior.

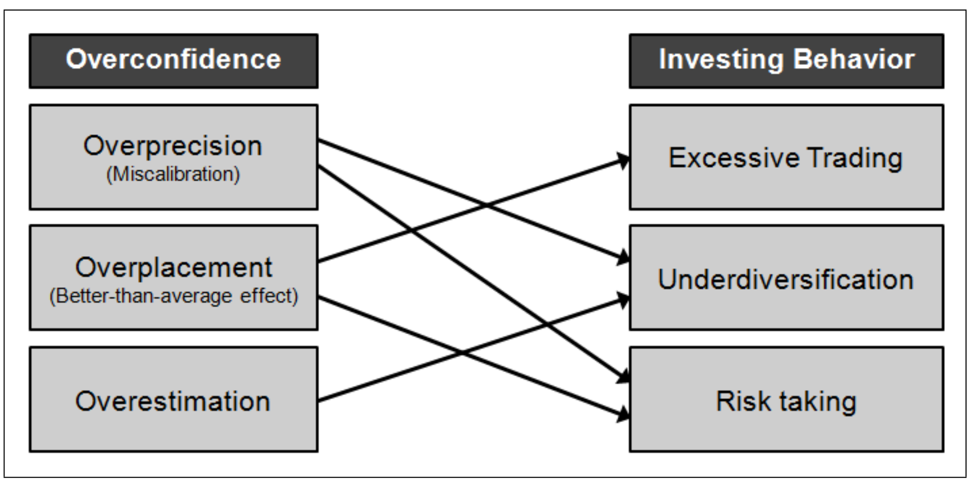
5.7 Conclusion

Investors are overconfident in various forms, for example they expect their portfolios to yield higher returns than the market and to be less risky at the same time. They also think to be better informed about financial markets and more skillful in investing than others. Investors are overprecise in their predictions of future returns, when asked for 90% confidence intervals submitted ranges contain the true value less than 50% of the time. And they do not only expect to outperform the market, but frequently overestimate portfolio returns compared to actual outcomes. We document these patterns in investor beliefs and ask the fundamentally important question: Is overconfidence in beliefs relevant for investing behavior?

The response is yes, and while this question has been partly answered in other studies, our approach systematically relates different aspects of investing behavior to all types of overconfidence; figure 5.4 summarizes our results. Trading activity is spurred by overplacement, the belief to be able to outperform the market. This is intuitive as it appears sensibly to trade only if one expects higher returns than from a buy and hold strategy in the market index. Portfolio diversification in contrast is influence by overprecision and overestimation, investors who are not aware of the range of possible outcomes feel less need to diversify, and those who expect overly high returns from their portfolio as well forgo diversification opportunities. Finally, portfolio risk taking depends on overprecision and overplacement. Investors, who believe to be better than others, take more risk, and those, who underestimate the variance of returns, also take more risk.

The dynamics of overconfidence is driven by actual and perceived success of investors. Estimation success is associated with higher miscalibration for subsequent confidence intervals, and investment success produces higher overestimation of future portfolio returns. However, the effect of realized investment performance is attenuated by the fact that investors are not fully aware of past outcomes. We identify estimation errors for past portfolio and market returns, and that perceived performance can differ a lot from actual performance. As overconfidence is a psychological phenomenon, perceived performance might be at least as important as realized values. Indeed, overestimation and outperformance of investors in hindsight are important determinants of overplacement and overestimation in foresight.

Figure 5.4: Overconfidence and investing behavior



Notes: Types of overconfidence and their identified pathways to several aspects of investing behavior.

Bibliography

Allen, F., Morris, S., & Shin, H. (2006). Beauty Contests and Iterated Expectations in Asset Markets *The Review of Financial Studies*, 19, 719–752.

Alicke, M. D. (1985). Global Self-Evaluation as Determined by the Desirability and Controllability of Trait Adjectives *Journal of Personality and Social Psychology*, 49, 1621–1630.

Alicke, M., & Govorun, O. (2005). The Better-Than-Average Effect. In M. Alicke, D. Dunning, & J. Krueger (Eds.), *The Self in Social Judgment* (pp. 83–106). Philadelphia, PA: Psychology Press.

Alicke, M. D., Klotz, M. L., Breitenbecher, D. L., Yurak, T. J., & Vredenburg, D. S. (1995). Personal contact, individuation, and the better-than-average effect *Journal of Personality and Social Psychology*, 68, 804–825.

Alpert, M., & Raiffa, H. (1982). A progress report on the training of probability assessors. In A. Tversky & D. Kahneman (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 294–305). New York, NY: Cambridge University Press.

Amromin, G., & Sharpe, S. A. (2009). Expectations of risk and return among household investors: Are their Sharpe ratios countercyclical? *Working paper.*

Baker, M., & Wurgler, J. (2006). Investor Sentiment and the Cross-Section of Stock Returns *Journal of Finance*, 61, 1645–1680.

Banerjee, S., Kaniel, R., & Kremer, I. (2009). Price Drift as an Outcome of Differences in Higher-Order Beliefs *The Review of Financial Studies*, 22, 3707–3734.

Barber, B. M. & Odean, T. (2000). Trading Is Hazardous to Your Wealth: The Common Stock Investment Performance of Individual Investors *The Journal of Finance*, 55(2), 773–806.

Barber, B. M. & Odean, T. (2001). Boys Will Be Boys: Gender, Overconfidence, And Common Stock Investment *Quarterly Journal of Economics*, 116, 261–292.

Barber, B., & Odean, T. (2008). All That Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors *The Review of Financial Studies*, 61, 785–818.

Barber, B. M., & Odean T. (2011). The Behavior of Individual Investors. In G. Constantinides, M. Harris, & R. Stulz (Eds.), *Handbook of the Economics of Finance* (in press). Amsterdam: Elsevier.

Barberis, N., Shleifer, A., & Vishny, R. (1998). A Model of Investor Sentiment *Journal of Financial Economics*, 49, 307–343.

Barberis, N. & Thaler, R. (2003). A Survey of Behavioral Finance. In G. M. Constantinides, M. Harris & R. Stulz (Eds.), *Handbook of the Economics of Finance* (pp.1051–1121). Amsterdam, Elsevier.

Barberis, N., & Xiong W. (2009). What Drives the Disposition Effect? An Analysis of a Long-Standing Preference-Based Explanation *The Journal of Finance*, 64(3), 751–784.

Baron, D. P. (1977). On the Utility Theoretic Foundations of Mean-Variance Analysis *The Journal of Finance*, 32(5), 1683–1697.

Bénabou, R., & Tirole, J. (2002). Self-confidence and personal motivation *The Quaterly Journal of Economics*, 117, 871–915.

Benartzi, S., & Thaler R. H. (1995). Myopic Loss Aversion and the Equity Premium Puzzle *The Quarterly Journal of Economics*, 110(1), 73–92.

Benoît, J., & Dubra, J. (2009). Overconfidence? *Working paper*, available at SSRN: http://ssrn.com/abstract=1088746.

Bhattacharya, U., Holden, C. W. & Jacobsen, S. (2010). Penny Wise, Dollar Foolish: The Left-Digit Effect in Security Trading *Working paper*.

Biais, B., & Weber, M. (2009). Hindsight Bias, Risk Perception and Investment Performance *Management Science*, 55, 1018–1029.

Black, F. (1986). Noise *The Journal of Finance*, 41, 529–543.

Blume, M. E. & Friend I. (1975). The Asset Structure of Individual Portfolios and Some Implications for Utility Functions *The Journal of Finance*, 30(2), 585–603.

Borgsen, S., & Weber, M. (2008). False Consensus and the Role of Ambiguity in Predictions of Others' Risk Preferences *Working Paper*.

Bradley, G. W. (1978). Self-serving biases in the attribution process: A reexamination of the fact or fiction question *Journal of Personality and Social Psychology*, 36, 56–71.

Brocas, I., & Carrillo, J. D. (2002). Are we all better drivers than average? Self-perception and biased behavior *CEPR discussion paper No. 3603*.

Brooks, P., Davies, G. B., & Egan, D. P. (2008). Linking Psychometrically Measured Risk Tolerance with Choice Behavior *Working paper*.

Brown, G. & Cliff, M. (2004). Investor Sentiment and the Near-Term Stock Market *Journal of Empirical Finance*, 11, 1–27.

Brown, J., Ivković, Z., Smith, P., & Weisbenner, S. (2008). Neighbors Matter: Causal Community Effects and Stock Market Participation *The Journal of Finance*, 63, 1509–1531.

Burks, S. V., Carpenter, J. P., Goette, L., & Rustichini, A. (2009). Is Overconfidence a Judgment Bias? Theory and Evidence *Working paper*.

Burson, K. A., Larrick, R. P., & Klayman, J. (2006). Skilled or unskilled, but still unaware of it: How perceptions of difficulty drive miscalibration in relative comparisons *Journal of Personality and Social Psychology*, 90, 60–77.

Camerer, C. F., & Hogarth, R. M. (1999). The Effects of Financial Incentives in Experiments: A Review and Capital-Labor-Production Framework *Journal of Risk and Uncertainty*, 19, 7–42.

Campbell, J. Y. (2006). Household Finance *The Journal of Finance*, 61(4), 1553–1604.

Chan, S. & Fong, W. (2004). Individual Investors' Sentiment and Temporary Stock Price Pressure *Journal of Business Finance and Accounting*, 31, 823–836.

Christen, C. & Gunther, A. (2003). The Influence of Mass Media and Other Culprits on the Projection of Personal Opinion *Communication Research*, 30, 414–431.

Cohn, R. A., Lewellen, W. G., Lease, R. C., & Schlarbaum, G. G. (1975). Individual Investor Risk Aversion and Investment Portfolio Composition *The Journal of Finance*, 30(2), 605–620.

Compte, O., & Postlewaite, A. (2004). Confidence-enhanced performance *The American Economic Review*, 94, 1536–1557.

Cooper, A. C., Woo, C. Y., & Dunkelberg, W. C. (1988). Enterpreneurs' perceived chances for success *Journal of Business Venturing*, 3, 97–108.

Coval, J. D., & Moskowitz, T. J. (1999). Home Bias at Home: Local Equity Preference in Domestic Portfolios *The Journal of Finance*, 54(6), 2045–2073.

Croson, R., & Gneezy, U. (2009). Gender Differences in Preferences *Journal of Economic Literature*, 46(2), 448–474.

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

Dave, C., Eckel, C. C., Johnson, C. A., & Rojas, C. (2010). Eliciting Risk Preferences: When is Simple Better? *Journal of Risk and Uncertainty*, 41(3), 219–243.

Dawes, R. (1989). Statistical Criteria for Establishing a Truly False Consensus Effect *Journal of Experimental Social Psychology*, 25, 1–17.

Dawes, R., & Mulford, M. (1996). The False Consensus Effect and Overconfidence: Flaws in Judgment or Flaws in How We Study Judgment? *Organizational Behavior and Human Decision Processes*, 65, 201–211.

Deaves, R., Lüders, E., & Luo, G. Y. (2009). An Experimental Test of the Impact of Overconfidence and Gender on Trading Activity *The Review of Finance*, 13(3), 575–595.

Deaves, R., Lüders, E., & Schröder, M. (2010). The dynamics of overconfidence: Evidence from stock market forecasters *Journal of Economic Behavior & Organization*, 75, 402–412.

De Bondt, W., & Thaler, R. (1985). Does the Stock Market Overreact? *The Journal of Finance*, 40, 793–805.

Dimson, E., Marsh, P., & Staunton, M. (2004). Irrational Optimism *Financial Analyst Journal*, 60(1), 15–25.

Dorn, D., & Huberman, G. (2005). Talk and Action: What Individual Investors Say and What They Do *Review of Finance*, 9(4), 437–481.

Dorn, D., & Huberman, G. (2010). Preferred risk habitat of individual investors *Journal of Financial Economics*, 97(1), 155–173.

Dorn, D., & Sengmueller, P. (2009). Trading as Entertainment? *Management Science*, 55(4), 591–603.

Dorn, D., & Strobl, G. (2009). Rational Disposition Effects: Theory and Evidence *Working Paper*.

Dunning, D., Meyerowitz, J. A., & Holzberg, A. D. (1989). Ambiguity and self-evaluation: The role of idiosyncratic trait definitions in self-serving assessments of ability *Journal of Personality and Social Psychology*, 57, 1082–1090.

Ehrlinger, J., Johnson, K., Banner, M., Dunning, D., and Kruger, J. (2008). Why the unskilled are unaware: Further explorations of (absent) self-insight among the incompetent *Organizational Behavior and Human Decision Processes*, 105, 98–121.

Erev, I., Wallsten, T. S., & Budescu, D. V., (1994). Simultaneous over- and underconfidence: The role of error in judgment processes *Psychological Review*, 101, 519–527.

Fama, E. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work *The Journal of Finance*, 25, 383–417.

Faro, D., & Rottenstreich, Y. (2006). Affect, Empathy, and Regressive Mispredictions of Others' Preferences Under Risk *Management Science*, 52, 529–541.

Fischbacher, U. (2007). Z-tree - Zurich toolbox for readymade economic experiments *Experimental Economics*, 10, 171–178.

Fisher, K., & Statman, M. (2000). Investor Sentiment and Stock Returns *Financial Analysts Journal*, 56, 16–23.

Fisher, K., & Statman, M. (2002). Bubble Expectations *The Journal of Wealth Management*, 5, 17–22.

Fisher, K., & Statman, M. (2003). Consumer Confidence and Stock Returns *Journal of Portfolio Management*, 30, 115–127.

Frank, J. D. (1935). Some Psychological Determinants of the Level of Aspiration *The American Journal of Psychology*, 47(2), 285–293.

French, K. R., & Poterba, J. M. (1991). Investor Diversification and International Equity Markets *The American Economic Review*, 81(2), 222–226.

Gervais, S., Heaton, J. B., & Odean, T. (2003). Overconfidence, investment policy, and executive stock options *Rodney L. White Center for Financial Research Working Paper No. 15-02*, available at SSRN: http://ssrn.com/abstract=361200.

Gervais, S., & Odean, T. (2001). Learning to Be Overconfident *The Review of Finacial Studies*, 14(1), 1–27.

Gigerenzer, G. (1991). How to make cognitive illusions disappear: Beyond "Heuristics and Biases" *European Review of Social Psychology*, 2, 83–115.

Gigerenzer, G., Hoffrage, U., & Kleinbölting, H. (1991). Probabilistic mental models: A Brunswikian theory of confidence *Psychological Review*, 98, 506–528.

Glaser, M., Langer, T., Reynders, J., & Weber, M. (2007). Framing Effects in Stock Market Forecasts: The Difference Between Asking for Prices and Asking for Returns *Review of Finance*, 11(2), 243–279.

Glaser, M., Langer, T., & Weber, M. (2009). Overconfidence of Professionals and Laymen: Individual Differences Within and Between Tasks?, *Working paper*.

Glaser, M., & Weber, M. (2005). September 11 and Stock Return Expectations of Individual Investors *Review of Finance*, 9(2), 243–279.

Glaser, M., & Weber, M. (2007). Overconfidence and Trading Volume *The GENEVA Risk and Insurance Review*, 32(1), 1–36.

Glaser, M., & Weber, M. (2007b). Why Inexperienced Investors Do Not Learn: They Do Not Know Their Past Portfolio Performance *Financial Research Letters*, 4(4), 203–216.

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

Glaser, M., & Weber, M. (2010). Overconfidence. In H. K. Baker & J. Nofsinger (Eds.), *Behavioral Finance: Investors, Corporations, and Markets*, (pp. 241–258). Hoboken, NJ: John Wiley & Sons.

Goetzmann, W. N., & Kumar, A. (2008). Equity Portfolio Diversification *Review of Finance*, 12(3), 433–463.

Graham, B., & Dodd, D. (1934). *Security Analysis*, New York: McGraw Hill.

Graham, J., & Harvey, C. (2001). The theory and practice of corporate finance: evidence from the field *Journal of Financial Economics*, 60, 187–243.

Graham, J. R., Harvey, C. R., & Huang, H. (2009). Investor Competence, Trading Frequency, and Home Bias *Management Science*, 55(7), 1094–1106.

Greiner, B. (2004). An Online Recruitment System for Economic Experiments. In K. Kremer & V. Macho (Eds.), *Forschung und wissenschaftliches Rechnen 2003* (pp. 79–93), Göttingen: Ges. fr Wiss. Datenverarbeitung.

Grieco, D., & Hogarth, R. M. (2009). Overconfidence in absolute and relative performance: The regression hypothesis and Bayesian updating *Journal of Economic Psychology*, 30, 756–771.

Grinblatt, M., & Keloharju, M. (2000). The investment behavior and performance of various investor types: a study of Finland's unique data set *Journal of Financial Economics*, 55, 43–67.

Grinblatt, M., & Keloharju, M. (2009). Sensation Seeking, Overconfidence, and Trading Activity *The Journal of Finance*, 64(2), 549–578.

Grinblatt, M., Keloharju, M., & Linnainmaa, J. T. (2011). IQ, Trading Behavior, and Performance *Working Paper*.

Grinblatt, M., Keloharju, M., & Linnainmaa, J. T. (2011b). IQ and Stock Market Participation *The Journal of Finance*, in press.

Haruvy, E., Lahav, Y., & Noussair, C. (2007). Traders' Expectations in Asset Markets: Experimental Evidence *The American Economic Review*, 97, 1901–1920.

Hawkins, S. A., & Hastie, R. (1990). Hindsight: Biased judgments of past events after the outcomes are known *Psychological Bulletin*, 107(3), 311–327.

Healy, P. J., & Moore, D. A. (2007). Bayesian overconfidence *Working paper*, available at SSRN: http://ssrn.com/abstract=1001820.

Heaton, J. B. (2002). Managerial Optimism and Corporate Finance *Financial Management*, 31(2), 33–45.

Hens, T., & Vlcek, M. (2011). Does Prospect Theory Explain the Disposition Effect? *Journal of Behavioral Finance*, in press.

Hertwig, R., & Ortman, A. (2001). Experimental practices in economics: A methodological challenge for psychologists? *Behavioral and Brain Sciences*, 24, 383–403.

Hoch, S. J., & Loewenstein, G. F. (1989). Outcome Feedback: Hindsight and Information *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15(4), 605–619.

Hoffmann, A. O., Post, T., & Pennings, J. M. (2011). Individual Investors and the Financial Crisis: How Perceptions Change, Drive Behavior, and Impact Performance *Working Paper*.

Hong, H., Kubik, J., & Stein, J. (2004). Social Interaction and Stock-Market Participation *The Journal of Finance*, 59, 137–163.

Hsee, C., & Weber, E. (1997). A Fundamental Prediction Error. Self-other Differences in Risk Preferences *Journal of Experimental Psychology*, 126, 45–53.

Ichheiser, G. (1943). Misinterpretations of Personality in Everyday Life and the Psychologist's Frame of Reference *Journal of Personality*, 12, 145–160.

Jianakoplos, N. A., & Bernasek, A. (1998). Are Women More Risk Averse? *Economic Inquiry*, 36(4), 620–630.

Juslin, P. (1994). The overconfidence phenomenon as a consequence of informal experimenter-guided selection of almanac items *Organizational Behavior and Human Decision Processes*, 57, 226–246.

Juslin, P., Winman, A., & Olsson, H. (2000). Naive empiricism and dogmatism in confidence research: A critical examination of the hard-easy effect *Psychological Review*, 107, 384–396.

Kahneman, D. (2003). Maps of bounded rationality: A perspective on intuitive judgement and choice *The American Economic Review*, 93(5), 1449–1475.

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision Under Risk *Econometrica*, 47, 263–291.

Kaustia, M. (2010). Prospect Theory and the Disposition Effect *Journal of Financial and Quantitative Analysis*, 45(3), 791–812.

Kaustia, M., & Knüpfer, S. (2010). Peer performance and stock market entry *AFA 2010 Atlanta Meetings Paper*.

Keefer, D. L. & Bodily, S. E. (1983). Three-Point Approximations for Continuous Random Variables *Management Science*, 29(5), 595–609.

Kempf, A., Merkle, C., & Niessen, A. (2009). Low Risk and High Return - How Emotional Attitudes Shape Stock Expectations *Working Paper*.

Keynes, J. (1936). *The General Theory of Employment, Interest and Money*, London: Macmillen.

Klayman, J., & Soll, J. B. (2004). Overconfidence in Interval Estimates *Journal of Experimental Psychology*, 30(2), 299–314.

Klayman, J., Soll, J. B., González-Vallejo, C., & Barlas, S.(1999). Overconfidence: It depends on how, what, and whom you ask *Organizational Behavior and Human Decision Processes*, 79, 216–247.

Köszegi, B. (2006). Ego Utility, Overconfidence, and Task Choice *Journal of the European Economic Association*, 4, 673–707.

Kroll, Y., Levy, H., & Markowitz, H. M. (1984). Mean-Variance Versus Direct Utility Maximization *The Journal of Finance*, 39(1), 47–61.

Krueger, J., & Clement, R. (1994). The Truly False Consensus Effect: An Ineradicable and Egocentric Bias in Social Perception *Journal of Personality and Social Psychology*, 67, 596–610.

Kruger, J. (1999). Lake Wobegon be gone! The "Below average effect" and the egocentric nature of comparative ability judgements," *Journal of Personality and Social Psychology*, 77, 221–232.

Kruger, J., & Dunning, D. (1999). Unskilled and Unaware of It: How Difficulties in Recognizing One's Own Incompetence Lead to Inflated Self-Assessments *Journal of Personality and Social Psychology*, 77(6), 1121–1134.

Kruger, J., & Gilovich, T. (1999). "Naive Cynism" in Everyday Theories of Responsibility Assessment: On Biased Assumptions of Bias *Journal of Personality and Social Psychology*, 76, 743–753.

Langer, E. J. (1975). The illusion of control *Journal of Personality and Social Psychology*, 32, 311–328.

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

Lease, R. C., Lewellen, W. G., & Schlarbaum, G. G. (1974). The Individual Investor: Attributes and Attitudes *The Journal of Finance*, 29(2), 413–433.

Lemmon, M., & Portniaguina, E. (2006). Consumer Confidence and Asset Prices: Some Empirical Evidence *The Review of Financial Studies*, 19, 1499–1529.

Lewellen, W. G., Lease, R. C., & Schlarbaum, G. G. (1977). Patterns of Investment Strategy and Behavior Among Individual Investors *The Journal of Business*, 50(3), 296–333.

Lewellen, W. G., Lease, R. C., & Schlarbaum, G. G. (1979). Investment Performance and Investor Behavior *Journal of Financial and Quantitative Analysis*, 14(1), 29–57.

Lichtenstein, S., Fischhoff, B., & Phillips, L. D. (1982). Calibration of probabilities: The state of the art to 1980. In A. Tversky & D. Kahneman (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 306–351). New York, NY: Cambridge University Press.

Lins, K., Servaes, H., & Tufano, P. (2009). What Drives Corporate Liquidity? An International Survey of Cash Holdings and Lines of Credit *Working paper*.

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

Lusardi, A., & Mitchell, O. S. (2007). Financial Literacy and Retirement Preparedness: Evidence and Implications for Financial Education *Business Economics*, 42(1), 35–44.

Mahar, H. (2003). Why are there so few prenuptial agreements? *Harvard Law School John M. Olin Center for Law, Economics and Business, Discussion Paper No. 436*.

Makarov, I., & Rytchkov, O. (2006). Forecasting the Forecasts of Others: Implications for Asset Pricing *Working paper*.

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

Markowitz, H. (1952). Portfolio Selection *The Journal of Finance*, 7(1), 77–91.

Marks, G., & Miller, N. (1987). Ten Years of Research on the False Consensus Effect: An Empirical and Theoretical Review *Psychological Bulletin*, 102, 72–90.

McInish, T. (1982). Individual investors and risk-taking *The Journal of Economic Psychology*, 2(2), 125–136.

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

McPherson Frantz, C. (2006). I AM Being Fair: The Bias Blind Spot as a Stumbling Block to Seeing Both Sides *Basic and Applied Social Psychology*, 28, 157–167.

Mednick, S. A. (1968). The Remote Associates Test *The Journal of Creative Behavior*, 3(2), 213–214.

Menkhoff, L., & Nikiforow, M. (2009). Professionals' Endorsement of Behavioral Finance: Does it Impact Their Perception of Markets and Themselves? *Journal of Economic Behavior and Organization*, 71, 318–329.

Merton, R. C. (1973). An Intertemporal Capital Asset Pricing Model *Econometrica*, 41(5), 867–887.

Moore, D. A., & Cain, D. M. (2007). Overconfidence and underconfidence: When and why people underestimate (and overestimate) the competition *Organizational Behavior and Human Decision Processes*, 103, 197–213.

Moore, D. A., & Healy, P. J. (2008). The trouble with overconfidence *Psychological Review*, 115, 502–517.

Mullen, B., Atkins, J., Champion, S., Edwards, C., Hardy, D., Story, J., & Vanderklok, M. (1985). The False Consensus Effect: A Meta-analysis of 115 Hypothesis Tests *Journal of Experimental Social Psychology*, 21, 262–283.

Nosić, A., & Weber, M. (2010). How Riskily Do I Invest? The Role of Risk Attitudes, Risk Perceptions, and Overconfidence *Decision Analysis*, 7(3), 282–301.

Odean, T. (1998). Volume, Volatility, Price, and Profit When All Traders Are Above Average *The Journal of Finance*, 53(6), 1887–1934.

Odean, T. (1998b). Are Investors Reluctant to Realize Their Losses? *The Journal of Finance*, 53(5), 1775–1798.

Odean, T. (1999). Do Investors Trade Too Much? *The American Economic Review*, 89(5), 1279–1298.

Pronin, E., Gilovich, T., & Ross, L. (2004). Objectivity in the Eye of the Beholder: Divergent Perceptions in Bias in Self Versus Others *Psychological Review*, 111, 781–799.

Preston, C. E., & Harris, S. (1965). Psychology of drivers in traffic accidents *Journal of Applied Psychology*, 49, 284–288.

Pronin, E., & Kugler, M. (2007). Valuing Thoughts, Ignoring Behavior: The Introspection Illusion as a Source of the Bias Blind Spot *Journal of Experimental Social Psychology*, 43, 565–578.

Pronin, E., Lin, D., & Ross, L. (2002). The Bias Blind Spot: Perceptions of Bias in Self Versus Others *Personality and Social Psychology Bulletin*, 28, 269–281.

Pulford, B. D., & Colman, A. M. (1997). Overconfidence: Feedback and item difficulty effects *Personality and Individual Differences*, 23(1), 125–133.

Ritter, J. R. (1988). The Buying and Selling Behavior of Individual Investors at the Turn of the Year *The Journal of Finance*, 43(3), 701–717.

Ross, L., Greene, D., & House, P. (1977). The "False Consensus Effect": An Egocentric Bias in Social Perception and Attribution Processes *Journal of Experimental Social Psychology*, 13, 279–301.

Ross, L., & Ward, A. (1996). Naive Realism in Everyday Life: Implications for Social Conflict and Misunderstanding. In E. Reed, E. Turie, & T. Brown (Eds.), *Values and Knowledge* (pp. 103–135). Mahwah, NJ: Lawrence Earlbaum.

Russo, J. E., & Schoemaker, P. J. H. (1992). Managing overconfidence *Sloan Management Review*, 33, 7–17.

Samuelson, P. A. (1998). A Note on the Pure Theory of Consumer's Behaviour *Economica*, 5(17), 61–71.

Santos-Pinto, L., & Sobel, J. (2005). A model of positive self-image in subjective assessments *The American Economic Review*, 95, 1386–1402.

Sarin, R. K., & Weber, M. (2009). Risk-Value Models *European Journal of Operational Research*, 70(2), 135–149.

Scheinkman, J. A., & Xiong, W. (2003). Overconfidence and Speculative Bubbles *Journal of Political Economy*, 111, 1183–1219.

Schlarbaum, G. G., Lewellen, W. G., & Lease, R. W. (1978). The Common-Stock-Portfolio Performance Record of Individual Investors: 1964-70 *The Journal of Finance*, 33(2), 429–441.

Schlarbaum, G. G., Lewellen, W. G., & Lease, R. W. (1978b). Realized Returns on Common Stock Investments: The Experience of Individual Investors *The Journal of Business*, 51(2), 299–325.

Schwarz, N. (1999). Self-reports: How the questions shape the answers *American Psychologist*, 54, 93–105.

Selten, R. (1998). Axiomatic Characterization of the Quadratic Scoring Rule *Experimental Economics*, 1, 43–62.

Shefrin, H., & Statman, M. (1985). The Disposition to Sell Winners Too Early and Ride Losers Too Long: Theory and Evidence *The Journal of Finance*, 40(3), 777–790.

Sherman, S., Presson, C., Chassin, L., Corty, E., & Olshavsky, R. (1983). The False Consensus Effect in Estimates of Smoking Prevalence *Personality and Social Psychology Bulletin*, 9, 197–207.

Shiller, R. J. (2002). Bubbles, Human Judgment and Expert Opinion *The Financial Analysts Journal*, 58, 18–26.

Statman, M. (1987). How Many Stocks Make a Diversified Portfolio? *Journal of Financial and Quantitative Analysis*, 22(3), 353–363.

Svenson, O. (1981). Are we all less risky and more skillful than our fellow drivers? *Acta Psychologica*, 94, 143–148.

Taylor, S., & Brown, J. (1988). Illusion and Well-Being: A Social Psychological Perspective on Mental Health *Psychological Bulletin*, 103, 193–210.

Tetlock, P. (2007). Giving Content to Investor Sentiment: The Role of Media in the Stock Market *The Journal of Finance*, 62, 1139–1168.

Van den Steen, E. (2004). Rational overoptimism (and other biases) *The American Economic Review*, 94, 1141–1151.

van Rooij, M., Lusardi, A., & Alessie, R. (2011). Financial Literacy and Stock Market Participation *Journal of Financial Economics*, 101(2), 449–472.

Vissing-Jorgensen, A. (2003). Perspectives on Behavioral Finance: Does 'Irrationality' Disappear with Wealth? Evidence from Expectations and Actions *NBER Working paper*.

Weber, E. U., & Milliman, R. A. (1997). Perceived Risk Attitudes: Relating Risk Perception to Risky Choice? *Management Science*, 43(2), 123–144.

Weber, E. U., Siebenmorgen, N., & Weber, M. (2005). Communicating Asset Risk: How Name Recognition and the Format of Historic Volatility Information Affect Risk Perception and Investment Decisions *Risk Analysis*, 15, 597–609.

Weber, M., & Camerer, C. F. (1998). The disposition effect in securities trading: an experimental analysis *Journal of Economic Behavior and Organization*, 33(2), 167–184.

Weber, M., Weber, E. U., & Nosić, A. (2010). Who takes risks when and why: Determinants of changes in investor risk taking *Working paper*.

Weinstein, N. (1980). Unrealistic Optimism about Future Life Events *Journal of Personality and Social Psychology*, 39, 806–820.

Welch, I. (2000). Views of Financial Economists on th Equity Premium and on Professional Controversies *Journal of Business*, 73, 501–537.

Whaley, R. (2000). The Investor Fear Gauge *The Journal of Portfolio Management*, 26, 12–17.

Windschitl, P. D., & Wells, G. L. (1996). Measuring Psychological Uncertainty: Verbal Versus Numeric Methods *Journal of Experimental Psychology*, 2(4), 343–364.

Winman, A., Juslin, P., & Björkman, M. (1998). The Confidence-Hindsight Mirror Effect in Judgment: An Accuracy-Assessment Model for the Knew-It-All-Along Phenomenon *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24(2), 415–431.

Zábojník, J. (2004). A model of rational bias in self-assessments *Economic Theory*, 23, 259–282.

Zuckerman, M. (1979). Attribution of success and failure revisited, or: The motivational bias is alive and well in attribution theory *Journal of Personality*, 47, 245–287.

Lebenslauf

Dipl.-Volkswirt Christoph Merkle
Lehrstuhl für ABWL und Finanzwirtschaft, insb. Bankbetriebslehre
Universität Mannheim

Akademischer Werdegang

2002-2007	Universität zu Köln, Studium der Volkswirtschaftslehre
2005	University of Economics Prague, Studium in Ökonomie und Finanzwirtschaft
2007-2011	Universität Mannheim, PhD-Student in Finance (GESS/CDSB)
Seit 2008	Wissenschaftlicher Mitarbeiter am Lehrstuhl für ABWL und Finanzwirtschaft

Forschung

Arbeitspapiere

- The Beliefs of Others - Nave Realism and Investment Decisions (2011, mit D.Egan und M. Weber)
- Low Risk and High Return - How Emotions Shape Expectations on the Stock Market (2011, mit A. Kempf und A. Niessen-Ruenzi)
- Financial Overconfidence Over Time - Foresight, Hindsight, and Insight of Investors (2011)
- Do Investors Put Their Money Where Their Mouth Is? Stock Market Expectations and Trading Behavior (2011, mit M.Weber)

Veröffentlichung

True Overconfidence: The Inability of Rational Information Processing to Account for Apparent Overconfidence, *Organizational Behavior and Human Decision Processes*, 116(2), S. 262-271, mit M. Weber.