Financial Decision Making – The Role of Realization, Portfolio Composition and Biased Belief Formation

Inauguraldissertation

zur Erlangung des akademischen Grades eines

Doktors der Wirtschaftswissenschaften

der Universität Mannheim

vorgelegt von

Jan Müller-Dethard, M.Sc.

im FSS 2020

Dekan: Joachim Lutz

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

Korreferent: Prof. Dr. Erik Theissen

Tag der mündlichen Prüfung: 15. Oktober 2020

Acknowledgements

This thesis is the result of a three years and six months working period during my doctoral student time at the University of Mannheim between 2017 and 2020. I want to thank many people for their encouragement and support during this period.

I am deeply grateful to my supervisor, Prof. Dr. Dr. h.c. Martin Weber, for many helpful suggestions, enriching discussions, and his contagious enthusiasm for behavioral finance which definitely spilled over to me. Many thanks also go to Prof. Dr. Erik Theissen, my second examiner.

I am thankful to Prof. Dr. Christoph Merkle and Pascal Kieren for the great collaboration on our joint projects and Prof. Dr. Alex Imas for his invaluable support during the first project of my dissertation. I am also very thankful to Sabine Bernard for her support. I appreciate insightful remarks by Dr. David Becker, Dr. Maximilian Germann, Dr. Zwetelina Illiewa, Prof. Dr. Heiko Jacobs, Dr. Christine Laudenbach, Prof. Dr. Benjamin Loss, Dr. Philipp Marquardt, Prof. Dr. Oliver Spalt, Dr. Michael Ungeheuer, Dr. Maximilian Wimmer and the seminar participants at the University of Mannheim.

Finally, I wish to acknowledge the encouragement and great love of my family. My heartfelt thanks also due Anna for her carrying support during the final phase of my thesis. This work would not have been possible without all of your input.

Contents

List of Figures

List of Tables

Für Mama und Papa

Chapter 1

Introduction

1.1 Relevance and Foundations of Financial Decision Making

People make hundreds of decisions every day. All of these decisions involve instances of *uncertainty* and *risk*. Depending on the degree of these factors, some decisions are associated with greater complexity than others. Financial decisions certainly belong to the more complex ones, for both, financial professionals and even more for private investors. Not only is the decision complex because of the shear amount of information regarding financial products, but also because of the limited experience and financial literacy of many investors. Confronted with these difficulties, investors need to make long-lasting financial decisions which strongly impact their future living standard.

Taking all of this into account, it seems not surprising that there are so many people who are overwhelmed, disinterested, or have a flawed understanding of financial decision making. In particular, it puzzles researchers why in Germany – one of the most developed and richest economies – only very few people invest in the stock market. With a rate of barely 15.2% not even one out of six people invests directly or indirectly in stocks.[1] This low stock market participation rate is striking because many scientific studies demonstrate that from a pure rational perspective most people should invest at least a small part of their wealth in stocks (Haliassos and Bertaut, 1995). Thus, many people forgo the so-called market risk premium, i.e. the difference between the return of a broadly diversified market index and a risk-free investment. For example, over a long-term investment horizon of 50 years, the average yearly return of the German stock market index DAX 30 was around 7%.[2]

There are various reasons why stock market participation is so low and a straightforward answer does not exist. In a recent, large-scale survey people who are not investing in stocks were asked why they refrain from stock market participation. The top three answers were fear of high losses due to economic catastrophes (67% of respondents agreed), limited financial resources (66%), and lack of elementary financial knowledge (65%).[3] It is particularly interesting that the researchers of this survey determine errors in risk and probability estimation as well as the lack of knowledge about how to reduce risk as the potential reasons for the above-listed answers. The authors argue that if risks are systematically overstated, the already high risk aversion of many people becomes even more severe and consequently keeps them away from investments in the stock market.

Why should politicians and researchers be concerned about this finding? The low stock market participation is especially problematic for long-term financial decisions such as the decision about how to save for retirement. If

[1] See report 2019 of the "Deutsche Aktieninstitut" on the stock market participation rate in Germany. The stock market participation rate subsumes investments in stocks and mutual funds.

[2] See DAX-Rendite Dreieck of the "Deutsche Aktieninstitut", December, 31 2019, the average yearly return of the DAX from 1969-2019 was 7.3%.

[3] See "Zum Rätsel der Aktienmarktteilnahme in Deutschland" (2019), a study by researchers of the Frankfurt School of Finance & Management and the Goethe-University Frankfurt on behalf of the Deutsche Börse AG.

people forgo the market risk premium when they save for retirement, they will likely have a hard time to secure their living standard after they retire. On top of this, many pension schemes — not only in Germany – face significant demographic challenges. Less and less pension contributors have to pay for the large generation of pension receivers in the present and near future. Thus, people cannot rely only on statuary pension insurance, but should instead complement those with private pension schemes. This means, they have to take precautions themselves. In other words, they need to deal with financial decisions about how to save for retirement, and thereby cannot circumvent the confrontation with risk and uncertainty in the stock market.

Adding to this, one might argue that the necessity to deal with financial decisions in the stock market has increased due to the ongoing low interest rate environment. In the past, considering a stock market investment might not have been as necessary as it is today. Saving accounts yielded quite remarkable nominal risk-free interest rates which were 5% p.a. around the year 2000.[4] However, nowadays, within the low interest rate environment, classic savings accounts seem to be an unpromising vehicle to save appropriately for retirement. At the same time, financial markets offer more and more products which allow retail investors to participate in the stock market at relatively low costs. A good example are so-called index funds which usually are, and initially only were, intended to replicate a broad equity market index. On the one hand, one might argue that the increased variability of products makes financial decisions more complex because of the overload of choices. On the other hand, the increased variability can also be understood as a benefit since investors are likely to find a product which fits their individual preferences more closely.

[4] Average yearly interest rate of banks for deposits/savings bonds with maturity of four years was 5.37% p.a. between 1990 and 2002 taken from the Deutsche Bundesbank Time Series BBK01.SU0031: Deposit rates of banks/bank savings bonds with regular interest payments, maturity of 4 years/average interest rate.

The example of retirement investing intends to show that people have to deal more with financial decisions which concern direct or indirect stock market participation. This comes with several aspects about which people need to think before making an investment decision such as an appropriate assessment of risk and probabilities in financial markets. However, this is difficult and may overwhelm potential investors. Instead and as mentioned above, people may falsely evaluate risk and may over- or understate the likelihood of extreme events. Hence, it is and should be an important concern of politicians and researchers to help people with their financial decisions. However, to support people in their financial decision making process, it is in a first step necessary to examine how their actual behavior looks like relatively to what normative theory advices them to do. The objective of this thesis is to analyze actual investment and risk-taking behavior of individuals to identify and better understand potential drivers for the observed behavior, to compare whether and when they systematically deviate from normatively optimal decisions, and to potentially derive policy implications on how to help them.

The question of how individuals make investment decisions under risk and uncertainty, and in particular, what determines investors' risk-taking has been the object of theoretical and empirical research in economics for centuries (Machina, 1987). Economists have developed various models to understand and predict investment behavior and asset prices in financial markets. Among the neoclassical models is the fundamental risk-return model by Markowitz (1952, 1959). It became the foundation for ground-breaking work in finance such as the Capital Asset Pricing Model (CAPM) (Sharpe, 1964; Lintner, 1965) and the Arbitrage Pricing Theory (APT) (Ross, 1976; Roll and Ross, 1980).

Given the importance of Markowitz work in the literature and its implications for the topics investigated throughout this thesis, I will briefly describe the main intuition of a common risk-return model. The common risk-return model resembles a fundamental trade-off which investors face in various financial decision making problems. An investor's willingness to pay (WTP) for a risky asset X can be modeled as a function of the asset's expected return R(X) and its risk, usually measured by the variance of the asset's return distribution Var(X). In this model, researchers assume investors to minimize the level of risk for any given level of return. Investors trade-off expected return against risk. This trade-off can, in its simplest form, be described as follows:

$$WTP(X) = R(X) - bVar(X), \tag{1.1}$$

where the parameter b describes an individual's risk aversion. In this model, an investor's willingness to invest in a risky asset depends on three integral parameters: (i) the *expected return R(X)* of the investment option, (ii) the *risk*, i.e. *Var(X)*, of the investment option, and (iii) the *risk attitude (b)* of the investor. While there is still debate how these parameters are defined and measured (e.g. whether variance is the "true" and only measure of risk, how risk aversion is measured), researchers agree that risk-taking in the neoclassical world depends on three parameters: expected return, risk, and risk attitude (see Sarin and Weber, 1993, Weber et al., 2013).[5]

$$RiskTaking = f(ExpectedReturn, Risk, RiskAttitude) \tag{1.2}$$

[5] Further neoclassical parameters of risk-taking are background risk and intertemporal consumption preferences. For the case of a stock market investment, background risks are health risks, job insecurity, or for example the loss of real estate due to a natural catastrophe. Intertemporal consumption preferences describe how a person plans to consume, i.e. spend money, over time. These preferences can affect the amount of investment and the investment horizon. For reasons of simplification, we neglect these parameters.

Based on this model, financial decision-making can be described as a two-stage process: First, an investor forms beliefs about the expected return and the risk of an investment option. Second, using the estimates about expected return and risk (i.e. given his/her beliefs), the investor evaluates based on his/her risk preferences whether to invest in the risky asset or not. However, how exactly do investors form beliefs and evaluate risk? In particular, which assumptions do researchers make about how investors form beliefs about expected returns and risk? Which assumptions do they make about how investors evaluate risk?

The Traditional Framework

The "traditional" framework in financial economics builds on simple assumptions about individual psychology: (i) Individuals have *rational beliefs*. This means, they update their beliefs promptly and correctly, according to Bayes' Theorem, when new information arrives. (ii) *Given their beliefs*, individuals make investment decisions according to the *Expected Utility Theory* (Bernoulli, 1954; Von Neumann and Morgenstern, 1947), whereby the utility function is monotonically increasing, concave, and defined over final wealth (i.e. consumption outcomes). Many models in financial economics build on the assumptions of the traditional framework.

Applying the assumptions of the traditional framework to the fundamental risk-return model implies that the investor forms beliefs about the first two parameters (expected return and risk) of the risky investment option according to Bayes' Rule. In classic portfolio theory (Markowitz, 1952), this means that the investor estimates the expected return (first moment) and the volatility (second moment) of the return distribution of the risky asset. Further assuming that all investors form beliefs from the same set of information and by using the same rule, beliefs about one and the same risky asset should be independent and identical across all investors who evaluate this asset. In contrast to that, the third parameter (an individual's risk aversion) captures the investor's personal attitude towards risk and is as such investor-specific.

The Behavioral Framework

By the 1990s, a new framework, the "behavioral finance" framework, emerged. The behavioral finance framework deviates from the widely-accepted, rational approach in financial economics by allowing for psychologically more realistic assumptions about how individuals update their beliefs and how they evaluate risk (see Barberis, 2018). The driving force of the emergence of behavioral finance was and still is to better understand empirical facts about the trading behavior and portfolio choice of individuals (e.g. non-participation in financial markets, under-diversification, the disposition effect, etc.) and asset returns (e.g. the equity premium puzzle, momentum and long-term reversal of stock returns, excess volatility, etc.) which "traditional" approaches have problems to explain.

According to Barberis (2018), behavioral finance aims to improve the psychological realism of models in financial economics along three dimensions: (i) *Beliefs* can be less than fully rational. (ii) *Preferences* can be more realistic. (iii) People suffer from *cognitive limits*. Inspired by the work in psychology about "judgment and decision making", behavioral finance has developed various models of belief formation (e.g. extrapolative beliefs, overconfidence, etc.) and preferences (e.g. prospect theory, ambiguity aversion, and other preference specifications) which prove to be very useful in understanding investor behavior and asset prices.

Integrating the assumptions of the behavioral finance framework to the fundamental risk-return model has important implications. If investors update their beliefs about a risky investment in a not-fully rational way, what does this imply for their investment decision? In particular, if investors systematically deviate from Bayes' Rule, how and in which situations does this deviation occur? Furthermore, if investors evaluate risk differently from what the Expected Utility Theory implies, how do their risk preferences actually look like and which characteristics should models featuring alternative risk preference specifications incorporate?

Taken together, it is essential to learn more about the underlying drivers of financial decision making which are (i) how investors *form beliefs* about risky assets and (ii) how they *evaluate risk* (i.e. how their risk preferences look like). Answers to these fundamental questions will enable researchers to better understand investment behavior on the individual investor level as well as trading volume and asset prices in equilibrium on the aggregate market level.

Most Important Behavioral Concepts Used in the Thesis

Throughout this dissertation thesis, I will build on well-established frameworks and theories in behavioral finance. Therefore, I will briefly explain three of these frameworks which are core to the research questions I am examining in the following chapters.

The first framework is *mental accounting*. Mental accounting origins from the psychology literature and describes "the set of cognitive operations used by individuals and households to organize, evaluate, and keep track of financial activities" (Thaler, 1999). A fundamental part of mental accounting is the categorization and grouping of outcomes to certain mental accounts. The way in which people assign outcomes to distinct mental accounts affects how they evaluate outcomes and how they make investment decisions. Categorization is one of the clearest mechanisms of the human thought as it improves cognitive efficiency by facilitating the processing of complex information (Rosch and Lloyd, 1978; Henderson and Peterson, 1992). Given the complexity of financial decisions, financial decision making presents a promising field for the application of mental accounting theories. While mental accounting can reduce the cognitive effort associated with financial decisions, it can also cause systematic errors. For example, mental accounts violate the economic principle of fungibility of money (Shefrin and Thaler, 1988; Thaler, 1985, 1999).

Mental accounting does not follow the same clear-cut rules as traditional accounting does. However, work in psychology has established that in mental accounting funds or financial outcomes are grouped based on their sources and uses (Thaler, 1999), a concept which is related to choice bracketing (Read et al., 1999). Choices or outcomes can be defined broadly (assigned to one broadly defined mental account) or narrowly (assigned to many distinct narrowly defined mental accounts). How broadly or narrowly they are defined often depends on the similarity of their sources and uses (Heath and Soll, 1996).

The way how outcomes are assigned to mental accounts affects how these outcomes are evaluated. Outcomes within the same mental account are evaluated jointly, while outcomes assigned to different mental accounts are evaluated separately. This has implications for investing. Shefrin and Statman (1987) proposed a model of mental accounting to explain the selling behavior of investors in financial markets. An investor opens a mental account when making an investment and closes the mental account upon selling the respective asset. The purchase price is the reference price against which relative gains and losses are evaluated. In a recent model on realization utility, Barberis and Xiong (2012) follow this logic. However, they call it an investment episode instead of a mental account which an investor starts and ends or opens and closes, respectively.

The second framework is the *Prospect Theory* (Kahneman and Tversky, 1979). Prospect theory in its original form deviates from the Expected Utility Theory in three main features: (i) It replaces the utility function which is defined over absolute outcomes (wealth levels) by a value function which is defined over relative gains and losses, i.e. changes in value from a reference point. (ii) It maintains the property of "marginal decreasing sensitivity", but introduces reference-dependent evaluation. The value function has two distinct parts: It is concave in the gain domain, implying risk-averse behavior and convex in the loss domain, implying risk-seeking behavior. (iii) There is an asymmetry in the slope of the value function that evaluates gains and losses, with a steeper function for losses. This kink in the value function incorporates the empirically observed loss aversion (Rabin, 1998). Later, this framework has been extended by the feature of probability weighting leading to Cumulative Prospect Theory (Tversky and Kahneman, 1992). Probability weighting incorporates that individuals tend to overweight small probabilities and underweight large probabilities. This feature is implemented in cumulative prospect theory by transforming objective probabilities into subjective decision weights.

These seemingly more realistic assumptions about how individuals evaluate risk have proven to be very helpful in better understanding investor behavior as well as asset prices. Barberis and Xiong (2009, 2012) and Ingersoll and Jin (2013) have shown that prospect theory in combination with realization utility can explain patterns in investor trading behavior such as the disposition effect (Shefrin and Statman, 1985; Odean, 1998; Weber and Camerer, 1998).[6] Combining prospect theory with mental accounting, Barberis and Huang (2001) as well as Barberis et al. (2001) have pointed out that

[6] Realization utility depicts the idea that individuals receive a burst of utility or disutility at the point in time they realize a gain or loss. This burst of utility depends on the size of the realized gain or loss. The first formal model of realization utility was developed by Barberis and Xiong (2012). Ingersoll and Jin (2013) extended the model.

models incorporating prospect theory preferences can even account for various time-series and cross-sectional patterns in stock market returns such as momentum and long-term reversal as well as excess volatility and the value premium.

The final part of this brief review will be attributed to *beliefs* and which *biases* and *heuristics* investors engage in when forming them. This literature, however, is much less neat and organized compared to the literature on risk preferences in behavioral finance (Barberis, 2018). Multiple biases and heuristics in belief formation have been identified over the past, including the belief in the law of small numbers (Tversky and Kahneman, 1971), gambler's fallacy (Alberoni, 1962), conservatism bias (Phillips and Edwards, 1966), base-rate neglect (Tversky and Kahneman, 1973), representativeness heuristic (Kahneman and Tversky, 1972), and the confirmation bias (Nickerson, 1998). With respect to the research questions covered in this thesis, I will focus the review on two "types of biases" which are prior-biased inference and base-rate neglect. Prior-biased inference subsumes biases which lead to inference which is biased towards current beliefs (for example the confirmation bias). Base-rate neglect describes the fact that individuals on average under-use prior information. Interestingly, these two types of biases can point in opposite directions. While prior-biased inference implies that individuals overinfer and update too much if new information confirms their prior beliefs, base-rate neglect implies that they overinfer and update too much if new information disconfirms prior information as individuals tend to neglect prior information (Benjamin, 2019). There are several studies providing evidence for prior-biased inference and others providing evidence for base-rate neglect.

Overview of Chapters

This dissertation thesis contributes to research in finance which investigates individual investor financial decision making and its underlying drivers. Drawing from various streams of the finance, psychology, and economics literature, each chapter of this dissertation thesis focuses on a particular factor in either investors' risk preferences or investors' belief formation that ultimately influences their investment decisions. Chapter 2 and Chapter 3 examine *how investors evaluate risk* and thus contribute primarily to research on *risk preferences*. Chapter 4 and Chapter 5 are concerned about *how investors form expectations* in financial markets and hence contribute mostly to research on *belief formation*.

In the following paragraphs, I will give a very brief overview of the main research questions covered in each chapter of the dissertation thesis. Afterwards, Section 1.2 provides a more detailed summary which focuses on the main findings and contributions to the literature.

The first two chapters of this thesis examine how individuals frame and evaluate investment episodes *over time* (Chapter 2) and *across assets* (Chapter 3). Research in the field of mental accounting has advocated that individuals follow certain, but so far inconclusively explored rules when grouping and valuing financial outcomes. These rules define so-called investment episodes which in turn impact the way individuals evaluate risk (see Barberis and Xiong, 2012). A key question in this framework is, when investors start and when they end investment episodes. Using a sequential risk-taking design, Chapter 2 investigates experimentally whether and under which conditions the *framework of realization* ends an investment episode and consequently affects risk-taking. The study focuses on whether dynamic risk-taking over time is differently affected by realized gains and losses versus unrealized gains and losses.

In Chapter 3, the findings from Chapter 2 are complemented and the focus shifts from a single risky asset towards a portfolio of stocks. Instead of one single asset, Chapter 3 analyzes experimentally how individuals evaluate and frame gains and losses (and as such investment episodes) across many assets in a portfolio. More precisely, it examines whether individuals evaluate and frame gains and losses in a portfolio on the *overall portfolio level* (i.e. portfolio-level mental accounting) and/or on the *individual stock level* (individual stock-level mental accounting) to finally learn more about their portfolio investment decisions. To do so, a novel counting-based performance measure is defined which is the composition of the number of winner stocks relative to loser stocks in a portfolio. The experimental insights are applied to financial market data to show that portfolio composition matters not only for individual investment decisions in an experiment, but also for the demand of exchange-traded funds on leading equity market indices.

In Chapter 4 and Chapter 5, the main focus is on research questions related to belief formation, selected biases in belief updating, and the impact of biased beliefs on investment decisions. One general question in financial economics is why risk-taking varies strongly and systematically with market cycles: Investors take more risk during boom markets and less risk during recessions. One reason for this observed behavior could be that investors' attitude towards risk (i.e. their level of risk aversion) changes, albeit preferences are usually assumed to be a stable construct in economics. Another possible reason is that investors' expectations about returns and risk change. This question goes to the basis of financial decision making and is – even after years of scientific work – still causing heated debates among researchers on both, the question in terms of content, and the necessity of finding an answer per se. With regard to the effectiveness of potential policy implications, it is however essential to know more about the underlying driver(s) of the observed differences in risk-taking over time. Chapter 4 takes a step in this direction. It uses an experimental approach to show that the *way individuals form beliefs* across boom and bust markets differs and that the resulting *biased beliefs* can explain *differences in risk-taking* over time and in particular across macroeconomic cycles.

In a recent review on errors in probabilistic reasoning and judgment biases, Benjamin (2019) argues that "despite so much work by psychologists [. . .] and modern behavioral economics, to date belief biases have received less attention from behavioral economics than time, risk, and social preferences" (p. 71). Thus, prior to incorporating biases into applied economic models making them fit better to the observable data, it is relevant to learn more about which biases in belief formation are likely to occur under which conditions. In other words, according to Benjamin (2019), it should be a major objective of future research in behavioral economics and finance to study the interaction between biases to better understand when people will update too much or too little. Finally, Chapter 5 aims to add to this agenda. It investigates whether people follow a simple counting rule implied by Bayes' Theorem when incorporating sequential, binary information about the quality of a risky asset. Based on an empirical framework, Chapter 5 tests experimentally how individuals update their prior beliefs after same-directional and opposite-directional signals to identify *whether* and *in which situations* they *over-* or *underinfer*.

1.2 Contribution and Main Results

1.2.1 Closing A Mental Account: The Realization Effect for Gains and Losses

Chapter 2, coauthored with Christoph Merkle and Martin Weber, presents an experimental study of the realization effect. The realization effect was first documented by Imas (2016) and refers to the difference in risk-taking between unrealized (i.e. paper) and realized losses. After a realized loss, individuals become more risk-averse, while they become more risk-seeking after a paper loss. Imas (2016) explains its occurrence with cumulative prospect theory (Tversky and Kahneman, 1992) and choice bracketing (Read et al., 1999; Rabin and Weizsäcker, 2009), a concept which is directly related to mental accounting (Thaler, 1985, 1999). The framework of realization sheds light on an apparent inconsistency in the literature on dynamic risk-taking which is that some studies find risk-seeking behavior after prior losses (Coval and Shumway, 2005; Weber and Zuchel, 2005; Langer and Weber, 2008), while others provide evidence for risk-averse behavior after prior losses (Massa and Simonov, 2005; Shiv et al., 2005; Frino et al., 2008).

In this study, we contribute to the long-lasting debate on how prior outcomes affect subsequent risk-taking by examining two main research questions: (1) Does the realization effect exist for *gains* as well? (2) Under which conditions does a distinction between paper and realized outcomes lead to differential risk-taking behavior? In particular, does the realization effect depend on the *skewness of the underlying investment opportunity*?

We first derive theoretical predictions for risk-taking behavior after gains and investment opportunities with different skewness. The intuition of the framework is the following. We develop a model with loss-averse investors who open a mental account at the beginning of an investment episode and close it upon realization. Realization triggers the closure of a mental account and as such it affects whether prior outcomes are considered finite (when the account is closed) or temporary (when the account remains open). Paper gains act as a cushion against future losses which makes increased risk-taking attractive (Thaler and Johnson, 1990), while realized gains are internalized, considered as "own money" and are not available as a cushion anymore.

Therefore, risk-taking decreases after realized gains. Skewness alters risk-taking mainly via the magnitude of potential gains and losses relative to the mental account balance. Keeping the expected value constant, the less positively skewed the investment opportunity is, the less probable, but also the larger losses become, and the more probable, but the smaller gains are. This leads to an attenuated realization effect after both, gains and losses. Risk-seeking behavior becomes in both domains less attractive, because the downside risk threatens to exceed the paper gain cushion and the upside potential is too small to allow for breaking even after a paper loss.

In a series of experiments, we first replicate the realization effect for losses and then test our theoretical predictions for gains and investment opportunities with different skewness. The experimental design is based on a modified version of Gneezy and Potters (1997). Participants were endowed with EUR 8.00 which they could invest over the course of four rounds in a positively skewed (symmetric or negatively skewed) lottery. Each round, participants decide on the amount of money (between EUR 0.00 and EUR 2.00) they want to invest in the risky lottery, whereby the invested amount serves as a measurement tool for their level of risk-aversion. At the beginning of each experiment, participants are randomly allocated to either a *Paper* treatment or a *Realization* treatment. Participants in the Paper treatment were informed about their earnings after round 3 on the screen of the computer, and continued playing a final round. Participants in the Realization treatment were also informed about their earnings after round 3, but had to hand back money they lost, or received money they gained up until that round, before playing a final round, respectively. Consequently, outcomes remained unrealized, so to say "on the paper", in the Paper treatment, whereas they were realized, initiated by a physical transfer of money, in the Realization treatment. As such, the design allows us to test for differences in subsequent risk-taking following unrealized (Paper treatment) versus realized (Realization treatment) gains and losses conditional on the skewness of the investment opportunity.

The results can be summarized as follows. We replicate the realization effect for losses, albeit less pronounced than in Imas (2016). The realization effect, defined as the between-treatment difference of the within-treatment differences in risk-taking between round 3 and 4, is 16 cents and as such smaller than in the original experiment (38 cents). The smaller realization effect in our replication study is primarily caused by the less pronounced risk-seeking behavior we observe after paper losses. Consistent with our theoretical predictions, we find that the realization effect also exists for gains and that the effect is even larger for gains than for losses with a difference of 22 cents. This finding can also be confirmed in the original data by Imas (2016) which we analyze with respect to gains. Finally, we provide evidence – consistent with our theoretical predictions – that the realization effect reduces or even disappears for symmetric and negatively skewed investment opportunities. Participants in both treatments invest similarly after paper and realized outcomes.

Taken together, Chapter 2 proposes an extension of the theoretical framework of the realization effect by Imas (2016), tests it experimentally and confirms it. On the one hand, we find evidence that the **realization effect is more general, since it not only applies to losses, but also to gains**. Likewise, on the other hand, our findings suggest **boundary conditions for the realization effect** showing that **positive skewness** of the investment opportunity is **a necessary condition** to observe differential risk-taking after paper and realized gains and losses.

1.2.2 The Portfolio Composition Effect

Chapter 3, coauthored with Martin Weber, presents an experimental and empirical study of how investors evaluate portfolio investment decisions.

Portfolio theory (Markowitz, 1952) gives clear normative advice how wealth between two financial securities should be allocated: Portfolio evaluation should be reduced to two key parameters which are expected returns and variance of returns. Wealth should be allocated across financial securities such that the variance of returns is minimized for any given overall expected return. However, various studies in behavioral finance have shown that this is not the case and that actual investment behavior substantially diverges from basic portfolio theory (Barber and Odean, 2000; Benartzi and Thaler, 2001; Barber and Odean, 2013). Therefore, it is essential to learn more about how investors evaluate their portfolios to better understand how they make investment decisions.

To this end, we study one specific allocation decision of the portfolio choice problem, namely how investors' allocation decisions between given, pre-determined portfolios of stocks are affected by different levels of performance information. We ask the following two main research questions: (1) Do investors consider both, the *overall portfolio level* and the *individual stock level*, when evaluating a portfolio's performance? (2) How does this two-level informational setup affect their *portfolio investment decisions*?

So far, there is relatively little knowledge about how investors frame and evaluate gains and losses in a portfolio (i.e. how they frame and evaluate investment episodes for not just a single asset over time, but across various assets over time). The dominant, often implicit assumption of studies in this field is that investors consider stocks in isolation, so to say, detached from one another. In other words, gains and losses are framed narrowly on the individual stock level (Frydman et al., 2017). Consequently, it is not surprising that the role of the portfolio for the analysis of individual investor trading behavior has widely been ignored. However, this is questionable not at least because many retail investors hold either self-selected or pre-determined (e.g. index funds) portfolios of assets. Only a few recent papers challenge the narrow framing assumption and consider also the portfolio level when analyzing investors' trading behavior (see Hartzmark, 2015; An et al., 2019.

Gaining a better understanding of the role of portfolio-level and individual stock-level information for the evaluation of portfolios, is extremely important to derive potential interventions for the regulator helping individuals to make better investment decisions. For example, when it comes to the disclosure of past performance information of portfolio-like securities (e.g. index funds) in the Key Investor Information Document (KIID) it is of great importance to know whether there is a discrepancy between information investors actually care about and information they should care about.

To investigate how different levels of performance information affect portfolio investment decisions, we define a simple, counting-based measure which is determined from performance information of the portfolio's individual stocks. The portfolio composition measure is calculated as the number of winner stocks (positive return since purchase) relative to the number of loser stocks (negative return since purchase).

In a series of three experiments, we let participants allocate an endowment between two portfolios which differ in either (i) the portfolio composition or (ii) the overall realized (expected) portfolio return and variance, or (ii) in both dimensions. In the baseline experiment, we hold the overall realized returns across portfolios identical and only differ the composition of winner and loser stocks. More precisely, one portfolio consists of 70% winner/30% loser stocks, while the other portfolio consists of the reversed composition of 30% winner/70% loser stocks. Within our baseline design, we find a strong portfolio composition effect which is that participants allocate

26% (22%) more of their endowment to the portfolio consisting of 70% winner/30% loser stocks than to the alternative portfolio with the reversed portfolio composition. We also find that participants report more optimistic return expectations and lower risk evaluations for those portfolios which consist of more winner than loser stocks.

In a next step, we try to eliminate the effect by controlling participants' beliefs about expected returns and variance. Portfolios are now not only identical with respect to the overall realized return, but also with respect to the expected return and variance. To implement this feature in our design, we explain to participants the underlying data generating process of returns such that they can learn about expected returns and variance before making an investment decision. Yet, the effect persists even among those participants who state the same beliefs about expected returns for both portfolios.

In our third and last experiment, we put the effect to a severe test. We extent the learning phase, provide computational support for the calculation of expected returns and variance, and clearly display the resulting expected return and variance of each portfolio. Portfolios are designed such that there is a unique mean-variance efficient allocation. Even under these conditions, we still find a pronounced portfolio composition effect.

Motivated by how participants in an experimental task evaluate portfolios and make investment decisions, we apply our insights to real market data. We investigate whether historical fund flows of exchange-traded funds on leading equity market indices from the period 2016-2019 are influenced by the index composition of winner and loser stocks. We find that the proposed portfolio composition measure affects future fund flows of exchange-traded funds on leading equity market indices controlling for the index return. Several robustness analyses show that the effect is of rather short-term, daily nature, it does not depend on extreme portfolio compositions, and persist when controlling for an index return dispersion.

In summary, Chapter 3 shows that it seems to matter to investors **how an overall portfolio return has been achieved in the past and is expected to be achieved in the future** with respect to the portfolio's composition of winner and loser stocks. Strikingly, we find that the here documented portfolio composition effect is **not predicted by theories that assume mean-variance efficient portfolio selection** (Markowitz, 1952). The results from Chapter 3 have implications on **how investors evaluate and frame gains and losses in a portfolio** and as such contributes to theoretical work on risk preference specifications which combines Prospect Theory with different levels of mental accounting as proposed by Barberis and Huang (2001).

1.2.3 Why So Negative? Belief Formation and Risk-Taking in Boom and Bust Markets

Chapter 4, coauthored with Pascal Kieren and Martin Weber, presents an experimental study of belief formation in boom and bust markets and its role on financial risk-taking. Various studies in financial economics find that investors' risk-taking varies over time and in particular across market cycles (Malmendier and Nagel, 2011; Weber et al., 2013; Guiso et al., 2018). Investors take more risk during boom markets and less risk during recessions. While there is broad empirical consensus on *how* investment behavior differs across macroeconomic cycles, there is more of a controversy *why* this is the case. In particular, the literature is at odds whether differential risk-taking over time is caused by *changes in investors' risk aversion* and/or by *changes in investors' beliefs*.

One strand of literature argues in favor of time-varying, instable risk preferences. In these models, the utility function of the representative agent is usually modified such that it accounts for the countercyclical equity risk premium which effectively generates countercyclical risk aversion (Campbell and Cochrane, 1999; Barberis et al., 2001). The notion “countercyclical risk aversion” describes that individuals become more risk averse during bust markets, and consequently demand a higher risk premium, whereas they become less risk averse during boom markets, demanding a lower risk premium. Experimental evidence for countercyclical risk aversion is found by Cohn et al. (2015), while survey evidence is presented by Guiso et al. (2018).

A key assumption of these models is that the representative agent forms rational expectations according to Bayes’ Theorem which means that following the countercyclical nature of risk premiums the agent should have more pessimistic return expectations during boom markets (i.e. the relatively high asset prices during booms markets will lead on average to lower future returns) and more optimistic return expectations during bust markets (i.e. the relatively low asset prices during bust markets will lead on average to higher future returns).

However, Greenwood and Shleifer (2014) have shown that reported expectations of investors are inconsistent with “rational” expectations. They are highly correlated with past returns and as such exactly opposite to what rational expectation models assume. Greenwood and Shleifer (2014) find survey evidence that investors’ subjective return expectations are more optimistic during boom markets and more pessimistic during bust markets. In essence, they seem to be pro-cyclical rather than countercyclical. This finding is in line with recent evidence by Amromin and Sharpe (2014) as well as Giglio et al. (2019) who use survey data to show that stock return expectations are pro-cyclical. Weber et al. (2013) also challenge the notion of countercyclical risk aversion by observing that changes in return expectations rather than changes in risk attitude explain changes in risk-taking of a sample of online-broker customers over the financial crisis of 2008. This is in line with König-Kersting and Trautmann (2018) who cannot replicate the findings on countercyclical risk aversion by Cohn et al. (2015) within a student subject pool.

Taken together, there is an ongoing debate on what drives changes in risk-taking over time. While there is much more work on risk preferences in the economics literature so far (see also Benjamin, 2019), research on beliefs has been continuously catching up over the recent years. This is mainly due to the easier accessibility of survey data and the decreasing, initial skepticism against survey data being too noisy.

Chapter 4 contributes to this debate by showing that distorted belief formation rules can explain differences in risk-taking across recessions and boom markets. In particular, we aim to answer the following main research questions: (1) How do different learning environments affect the formation of return expectations? (2) How do systematic differences resulting from different learning environments affect risk-taking? (3) Do different learning environments affect not only investors' beliefs, but also their risk preferences?

To answer these questions, we run two experiments. The general idea of all experiments is to combine an abstract belief formation (forecasting) task in an adverse (bust) or favorable (boom) learning environment with an incentive-compatible investment task in a financial environment. In the forecasting task, participants learn to form beliefs about the quality of a risky asset in an environment which resembles either key characteristics of a boom market (Experiment 1: only positive returns, Experiment 2: positive expected return) or a bust market (Experiment 1: only negative returns, Experiment 2: negative expected return). Importantly, while the outcomes of the lotteries from which participants learn are framed differently, the underlying probability distributions are exactly the same in both learning environments. In the subsequent, independent investment task, we develop a *between-subject measure of belief- and preference-based risk-taking*. In particular, we randomly assign participants to either an ambiguous lottery with unknown probabilities or a risky lottery with known probabilities. The risky lottery is used as a measurement tool for possible changes in risk aversion across learning environments, whereas the ambiguous lottery intentionally provides participants with room to form beliefs about the underlying probability distribution which may then translate to the investment in the ambiguous lottery. The design allows us to isolate the effect of differences in beliefs from differences in risk aversion on financial risk-taking conditional on the boom or bust learning environment.

We predict a pessimism bias in participants' beliefs in the forecasting task which is that they report significantly more pessimistic beliefs in the bust as compared to the boom treatment and significantly more pessimistic beliefs compared to Bayes, extending work by Kuhnen (2015). With respect to the investment task, we predict lower investments in the ambiguous lottery in the bust treatment than in the boom treatment and no significant differences across treatments for the risky lottery.

The results can be summarized as follows. We confirm our hypotheses and find that the induced pessimism from adverse learning environments translates to lower investments in the ambiguous lottery. However, we do not find any differences in investment across treatments for the risky lottery which implies that risk preferences remain unaltered by the environments in which participants formed beliefs. Further analyses provide evidence for beliefs and the biased way in which they are formed being the underlying mechanism for our main finding.

To conclude, Chapter 4 provides and tests an **alternative channel to countercyclical risk aversion** that can also lead to the empirically observed changes in risk-taking over time. We show that **biased belief formation rules** caused by different **learning environments** can induce **overly pessimistic expectations which translate to lower risk-taking**. This finding is consistent with survey evidence reporting **pro-cyclical expectations**. However, we do not find that adverse learning environments affect participants' risk aversion. Our findings and the reported mechanism have important policy implications. If investors are overly pessimistic in recessions, they may expect lower returns and reduce their equity share. This, in consequence, may amplify the intensity and length of recessions.

1.2.4 Can Agents Add and Subtract When Forming Beliefs?

Chapter 5, coauthored with Pascal Kieren and Martin Weber, presents an experimental study on how individuals incorporate confirming and disconfirming information signals when sequentially updating their beliefs about the quality of a risky asset. Thereby, we test in the standard updating paradigm whether individuals follow a simple, but fundamental counting rule, implied by Bayes' Theorem, when forming beliefs. This common updating task is characterized by subjects receiving binary information signals about a risky asset that can be in one of two states of the world (Grether, 1980). The rule we are testing within this framework states that two opposite-directional signals should cancel out such that prior beliefs remain constant.

Probabilistic beliefs are essential in various economic problems such as for example investments in the stock market or purchasing insurance. Traditional models in economics assume that individuals update their beliefs promptly and correctly, according to Bayes' Theorem, when new information arrives. However, a large body of studies in psychology and economics has shown that individuals' beliefs often deviate from what Bayes' Rule implies. In other words, there is consensus in the literature that individuals are not perfect Bayesian. Instead, they do under- or overinfer from new information. However, there is less of a consensus on the question *when* one may expect to observe one versus the other. The literature is in need of a clarification of *when* people update too much or too little and seeks for a parsimonious model that can explain when one versus the other is more likely to occur.

Benjamin (2019) proposes in a recent overview on errors in probabilistic reasoning and judgment biases that by and a large, people update too little in the above-sketched updating paradigm. He also proposes that there are exceptions: When signals go in the same direction or when priors are extreme and signals go in the opposite direction of the priors, people overinfer and update too much. Given the large and often apparently inconsistent evidence in the literature and the unifying suggestion by Benjamin (2019), it is imperative to develop a framework and to systematically test and identify *when* individuals are more likely to over- and when they are more likely to underreact to new information. We take a step in this direction by investigating *whether* and *in which situations* individuals follow the simple counting rule, implied by Bayes' Theorem and if not, *in which situations* they deviate and instead over- or underreact.

We first develop a simple empirical framework and then test the hypotheses derived from this framework experimentally. A key feature of our framework is that we aim to investigate how individuals react to (i) a single disconfirming signal (i.e. opposite-directional signal) conditional on the number of previously observed confirming signals (i.e. same-directional signals) and how they react to (ii) a confirming signal which directly follows the disconfirming signal (i.e. a reversion of the disconfirming signal). We define a confirming signal as a signal which confirms the underlying state of the world and a disconfirming signal as a signal which does not confirm the underlying state of the world. Additionally, we define three phases of how Bayesian beliefs can evolve over a sequence of outcomes. Phase 1 ("confirming signals") is characterized by a sequence of at least two same-directional signals, Phase 2 ("disruptive signal") resembles the period in which the disconfirming signal occurs, and Phase 3 ("correction") defines the situation when a previously observed disconfirming signal gets directly reverted. The counting rule makes clear predictions how participants should update their beliefs in Phase 2 and Phase 3: an agent should reduce his prior probability estimate after a disconfirming signal by the same magnitude than he increased it after the previous confirming signal.

We test this prediction using the standard, incentivized updating paradigm by Grether (1980). Participants learn over six periods about the quality of a risky asset from binary signals (good or bad) which are drawn either from a "good distribution" or a "bad distribution". We exogenously manipulate the period in which the disconfirming signal occurs. This provides us with twelve stratified price paths (six for the good and six for the bad distribution).

Our main findings are as follows. Participants violate the simple counting rule and strongly overreact whenever a sequence of confirming signals is interrupted by a single disconfirming signal. The documented overreaction is relatively independent of the number of previously observed confirming signals, occurs already after a sequence of only two confirming signals, and thus does not critically depend on participants having extreme priors. Interestingly however, participants adhere to the counting rule and fully correct their prior overreaction when the disconfirming signal gets directly reverted. In addition to this, participants generally underinfer in situations in which they cannot or do not violate the counting rule. This is the case when there are only signals of same direction or signals of alternating sign.

Our findings have implications for various fields of research, in particular on belief formation in financial market, trading behavior and asset prices. It contributes to the early literature on over- and underreactions (Bondt and Thaler, 1985; Barberis et al., 1998; Daniel et al., 1998; Hong and Stein,

1999) as well as to the recent literature on extrapolative beliefs (Barberis and Shleifer, 2003; Greenwood and Shleifer, 2014; Barberis et al., 2015, 2018), by showing that participants seem to already over-extrapolate from and as such overreact to a single opposite-directional signal which interrupts a sequence of previous same-directional signals. Thus, our findings suggest that individuals even over-extrapolate from a single opposite-directional signal (e.g. bad earnings news) if it occurs after a relatively long prior history of same-directional signals (e.g. many good earnings news) and as such in situations in which they state and should be quite sure about the underlying state of the world. This in turn might add to a better understanding of the empirically observed high trading volume in stock markets in general and during bubbles in particular (Hong and Stein, 2007) as well as the excessive volatility of stock prices (Shiller, 1981).

In summary, Chapter 5 contributes to one important objective in research on probabilistic belief formation which is – according to Benjamin (2019) – to **identify when individuals update too much and when they update too little**. Within the common paradigm of Grether (1980), our results coherently suggest that individuals **update too much whenever they violate the simple counting rule**, implied by Bayes' Theorem. Across all of our experiments, this is the case in situations when **a sequence of same-directional signals is interrupted by a single opposite-directional signal**.

Chapter 2

Closing A Mental Account: The Realization Effect for Gains and Losses *

2.1 Introduction

Many risky endeavors, be it a night at the casino or an investment in a stock, involve instances in which individuals must decide whether to continue, to abandon, or to double down on a previous decision. They often view such episodes in isolation, even though normative theory suggests integrating them into a broader perspective of total wealth. They instead engage in mental accounting (Thaler, 1985, 1999), which refers to a cognitive process to categorize outcomes by their source or purpose. Prior outcomes within a mental account, perceived as a gain or a loss, obtain special relevance for this account and affect subsequent risk-taking (Thaler and Johnson, 1990).

* Authors: Christoph Merkle, Jan Müller-Dethard, and Martin Weber. Christoph Merkle is at the Aarhus University BSS. Jan Müller-Dethard and Martin Weber are at the University of Mannheim. For valuable comments, we thank Nick Barberis, Peter Bossaerts, Michaela Pagel, Yuval Rottenstreich, participants of the Experimental Finance 2018, the Research in Behavioral Finance Conference 2018, the Tiber Symposium on Psychology and Economics 2018, the Annual Meeting of the Society for Judgment and Decision Making 2018, and seminar participants at the University of Basel, the Max Planck Insitute in Bonn, and the University of Mannheim. We especially thank Alex Imas for his comments. We gratefully acknowledge financial support by the Stiegler Foundation and the German Research Foundation (DFG grant WE 993/15-1). The paper is *forthcoming in Experimental Economics*.

The direction of this influence has been subject to a long-standing debate. After losses, many studies find that individuals become more risk-seeking (Coval and Shumway, 2005; Weber and Zuchel, 2005; Langer and Weber, 2008; Andrade and Iyer, 2009), while others report they become more risk-averse (Massa and Simonov, 2005; Shiv et al., 2005; Frino et al., 2008). Similarly, after gains, investors will either exhibit more risk-seeking behavior (Thaler and Johnson, 1990; Weber and Zuchel, 2005; Suhonen and Saastamoinen, 2018) or more risk-averse behavior (Kahneman and Tversky, 1979; Clark, 2002; Coval and Shumway, 2005).

Existing theory can account for these different reactions by a variety of models or arguments. On the one hand, risk-seeking behavior after a prior loss and risk-averse behavior after a prior gain are often explained by prospect theory (Kahneman and Tversky, 1979). After a loss, the relevant part of the prospect theory value function to evaluate further outcomes is convex, which implies risk-seeking behavior. In contrast, a prior gain will situate a person in the gain domain for which the value function is concave, which implies risk-averse behavior.

On the other hand, more risk-seeking behavior after gains and more risk-averse behavior after losses can be motivated by the house money effect (Thaler and Johnson, 1990) and the hedonic editing hypothesis (Thaler, 1985). The house money effect describes a situation in which prior gains can be used to wager in subsequent gambles. People find it easier to part with money not coming from their own pocket. In addition, hedonic editing allows them to offset future losses against earlier gains. For losses, it is argued that they become more painful when they follow on the heels of prior losses (Barberis et al., 2001).

A unifying framework to resolve the conflicting evidence has been recently proposed by Imas (2016). It builds on the distinction between realized and unrealized outcomes, whereby a realization is defined "as an event in which money or another medium of value is transferred between accounts" (Imas, 2016, p. 2091). He argues that individuals behave differently depending on whether a loss is realized or whether it is still unrealized (a paper loss). Experimentally, he replicates prior findings that participants become more risk-averse after a realized loss, while they become more risk-seeking after a paper loss. He labels the difference in risk-taking between paper and realized losses the "realization effect" and explains its occurrence with cumulative prospect theory (Tversky and Kahneman, 1992) and choice bracketing (Read et al., 1999; Rabin and Weizsäcker, 2009), an idea directly related to mental accounting.

The proposed framework sheds light on why both, risk-averse as well as risk-seeking behavior, can be observed after the same prior outcome. However, drawing general conclusions from realization for subsequent risk-taking requires some caution. First, Imas's (2016) theoretical and experimental elaboration focuses exclusively on losses, and second, it tests the realization effect for an investment opportunity with a positively skewed distribution of outcomes. We argue that the literature is still in need of empirical and theoretical clarification about how prior outcomes – losses as well as gains – affect subsequent risk-taking, and in particular, under which conditions a distinction between paper and realized outcomes leads to differential risk-taking behavior. In this study, we contribute to this goal by examining two major research questions: (1) Does the realization effect exist for gains as well? (2) Does the realization effect depend on the skewness of the underlying investment opportunity?

To this end, we derive theoretical predictions for risk-taking behavior after gains and investment opportunities with positive skewness, no skewness, and negative skewness. We model loss-averse investors who open a mental account at the beginning of an investment episode and close it upon realization. Paper gains and losses alter the balance of the mental account and can thereby affect risk-taking. Paper gains act as a cushion against future losses and thus invite higher risk-taking, which is absent after gains are realized. We thus predict a realization effect for gains. Skewness comes into play mainly via the size of potential gains and losses relative to the account balance. With non-positive skewness, losses become less probable but larger. They threaten to exceed the paper gain cushion, attenuating the realization effect after gains. Likewise, after paper losses, more probable but smaller gains take away the potential to break even, which is a major motivation for higher risk-taking after losses. We thus predict a smaller or absent realization effect for non-positively skewed lotteries.

We conduct three well-powered experiments to test these predictions. In the first experiment, we replicate the main experiment by Imas (2016) using an identical design, which examines a series of positively skewed investment opportunities. The importance of replication for scientific progress in economics has been highlighted recently (Maniadis et al., 2014; Camerer et al., 2016; Christensen and Miguel, 2018). At the same time, the experiment allows us to address the first research question about a realization effect for gains. Not only is risk-taking after gains arguably as important as after losses, but it shares a similar conflict in previous empirical results and theory. If there is evidence for a realization effect in the gain domain as we predict, the proposed framework would have broader implications than those already suggested for the loss domain.

To answer the second research question, we analyze in two further experiments boundary conditions for the realization effect. In particular, we depart from positively skewed lotteries used so far and examine how symmetric or negatively skewed lotteries affect risk-taking behavior after paper and realized outcomes. Not only does positive skewness encourage risk-seeking behavior as it is often associated with gambling (e.g., lotteries or casinos), but the underlying distributions of most financial investment opportunities (e.g., stocks or funds investments) are less or not at all positively skewed. In order to establish the validity of the realization effect for these settings, it is essential to confirm whether the effect is indeed reduced as theory predicts.

The first experiment, which replicates study one by Imas (2016), involves a sequence of four positively skewed lotteries, each of which represents the throw of a die. One lucky number (out of six) wins seven times the stake invested in the lottery, while the stake is lost for all other outcomes. Up to EUR 2.00 can be invested in each lottery. After the third lottery, previous earnings are either paid out to participants or remain unrealized, which defines the two treatments in the experiment (realization treatment and paper treatment). The relevant comparison then is what participants do in the fourth and final lottery depending on realization. We use a larger sample size (N=203) than the original study to ensure sufficient statistical power and to be able to examine outcome histories that occur less frequently.

We first confirm that participants take less risk after a realized loss compared to a paper loss. However, the difference of 16 cents in average invested amounts between treatments is smaller than in the original experiment (38 cents), and the realization effect is not statistically significant. While we confirm a decrease in risk-taking in the realization treatment, we cannot corroborate an increase in risk-taking in the paper treatment. Standard replication measures show that the replication is at least partially successful.

Exploiting observations in which participants have obtained a gain at the time of realization, we find a similar investment pattern as for losses. Participants take significantly less risk after a realized gain than after a paper gain. The realization effect is larger for gains than for losses with a difference of 22 cents in average investment between treatments. In the paper treatment, participants seem to gamble with the house's money, while in the realization treatment, they have closed the mental account and regard gains from the lottery as their personal money. Given the consistent direction of the realization effect for gains and losses, we test for the realization effect unconditional of a particular outcome history. The results show a positive and strongly significant realization effect ($p < .01$) in the full sample.

In addition to our own experimental data, we analyze data from the original study by Imas (2016) with respect to gains.[1] Although limited in the number of observations, the realization effect for gains is strong and consistent with our results. Thus, we find evidence for a realization effect for gains in two independent samples. Moreover, pooling the data from both studies, we find a positive and strongly significant realization effect ($p < 0.001$) for gains and losses. To test for the theoretical relation between the realization effect after gains and the house money effect, we examine the invested amounts after a paper gain. In almost all cases, participants do not invest more than what they have gained in the lotteries. This implies that they gamble with the house's money, but do not touch their initial experimental endowment.

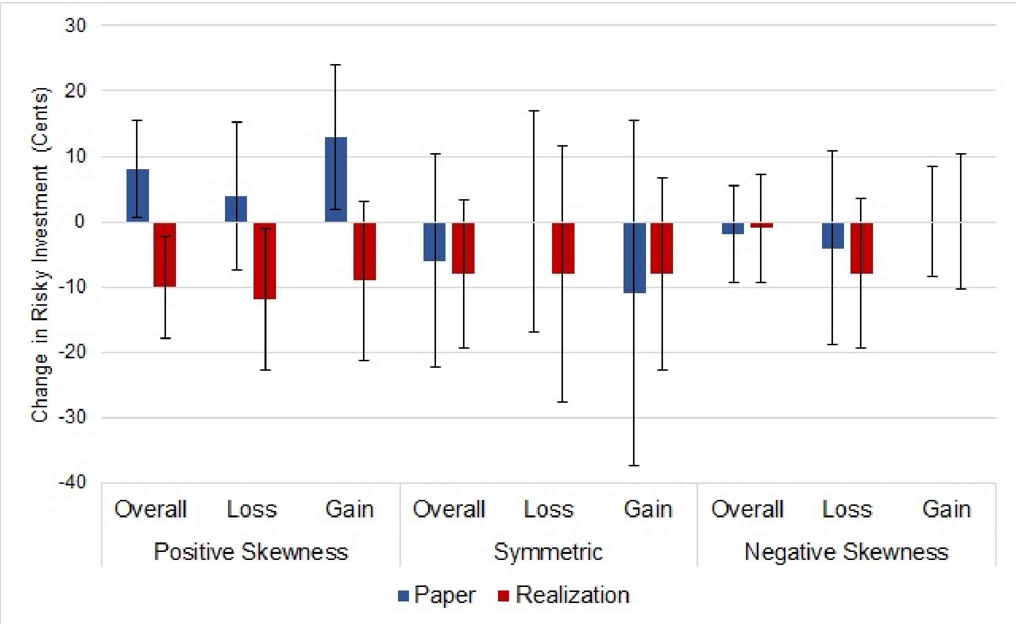
In experiments two and three, we examine how other distributions of outcomes affect risk-taking behavior after paper and realized gains and losses. We keep the basic experimental setup but change the probability of gains. Instead of a positively skewed lottery, participants invest in a symmetric or negatively skewed lottery, respectively. By construction this also increases the heterogeneity of outcome histories prior to realization. We find neither in the symmetric lottery nor in the negatively skewed lottery a statistically significant realization effect for gains or losses (total sample size N=304). In contrast to the positively skewed environment in the first study, participants tend to invest similarly after a paper outcome and a realized outcome. This finding is in line with theoretical work by Barberis (2012) and Imas (2016) in

[1] The data is publicly available via the AER website. Imas (2016) restricts his analysis to participants, who have lost in all lotteries up to round three (when realization takes place).

which individuals form contingent plans over a sequence of lotteries.

The results across all experiments suggest boundary conditions for the realization effect. Figure 2.1 depicts the magnitude of the realization effect we find, conditional on the outcome history as well as the skewness of the investment opportunity.

Figure 2.1: The Realization Effect Across All Experiments



Note: The figure displays average changes in risk-taking after paper and realized outcomes unconditional of the prior outcome history, and split by loss and gain for positively skewed, symmetric, and negatively skewed lotteries. Reported are 90%-confidence intervals.

Increased risk taking after paper gains and losses requires positive skewness, while decreased risk taking after realized gains and losses does not. The absence of the realization effect for non-positively skewed lotteries is thus primarily driven by an absence of increased risk taking after paper outcomes. This includes the absence of loss chasing, which seems to be limited to positive skewness environments.

The remainder of the paper is organized as follows. In Section 2.2, we derive theoretical predictions for the experiments, in particular for risk-taking behavior after gains and lotteries with different skewness, and review the prior literature. Section 2.3 presents the experimental design and the main results. A final section concludes.

2.2 Theory and Literature

To understand the behavior of participants in the experiments, we build on the model by Barberis et al. (2001). In addition to standard consumption-based utility, they consider utility derived directly from the fluctuations of financial wealth. In particular, agents react to gains and losses from their risky assets, which makes the model suitable for the analysis of behavior after gains and losses. Prior theory used to motivate the realization effect does not generate clear predictions for risk-taking behavior after gains. We introduce two departures from the main model in Barberis et al. (2001), which are the distinction between paper outcomes and realized outcomes, and a different value function after losses. The first is a natural extension to accommodate the treatment of paper and realized outcomes, the second takes into account the empirically observed behavior in the loss domain.

2.2.1 Basic Framework

The full utility specification in Barberis et al. (2001) includes utility from consumption $u(C_t)$ and utility derived from fluctuation of financial wealth $v(X_t, B_t, Z_t)$. We concentrate on the latter as it represents the important part of evaluating risk-taking behavior after gains and losses. X_t is the gain or loss a participant experiences in lottery t.[2] B_t is the bet size a participant selects for lottery t. And Z_t is a mental account, which reflects whether a participant perceives himself up or down in the game. Mental accounting describes the cognitive processes people use to organize and evaluate their financial activities (Thaler, 1985, 1999). A key implication is that people do not consider money across different mental accounts as perfect substitutes, but rather categorize money based on its origin or purpose and assign

[2] The original model defines X_{t+1} as the outcome over the time period from t to $t+1$. As we deal with discrete events, we use t to refer to successive lotteries.

it to separate accounts. Outcomes within a mental account are evaluated jointly, whereas outcomes in different mental accounts are evaluated separately (Thaler, 1999).

The three variables X_t, B_t, and Z_t, jointly determine the utility derived from fluctuations of financial wealth. A difference to the more general model arises from the fact that only part of a participant's endowment is invested in the risky lottery. Still, B_t can be interpreted as a participant's risky asset holdings. The outcome of lottery t is $X_t = R_t B_t - B_t$ with gross return R_t. We abstract from a risk-free rate, as no return is paid on money not invested in the lottery. If a participant loses in the lottery, then $X_t = -B_t$. If a participant wins, then $X_t = (x-1)B_t$ with $x > 1$ as the multiple that is applied to a winning bet. The lottery will thus either generate a loss or a gain. Besides these potential outcomes, participants take their prior gains and losses into account. Z_t is the mental account, which reflects prior outcomes:

$$Z_t = \sum_{\tau=1}^{t} X_{\tau-1} \tag{2.1}$$

While Barberis et al. (2001) leave open what exactly this mental account (or "historical benchmark") is, in our context, we will assume that it is the sum of prior gains and losses. A participant can thus be in the gain domain ($Z_t > 0$), in the loss domain ($Z_t < 0$), or at break-even ($Z_t = 0$). In particular, $Z_1 = 0$ as no lottery has yet been played. In this situation, utility from changes in financial wealth is described by:

$$v(X_t, B_t, 0) = \begin{cases} X_t & \text{for } X_t \geq 0 \\ \lambda X_t & \text{for } X_t < 0 \end{cases} \tag{2.2}$$

The parameter $\lambda > 1$ captures loss aversion. We further assume that realizing a gain or a loss resets the benchmark to zero as the mental account is closed. The intuition is that when a stock is sold, the proceeds are mentally transferred from the account investment to consumption. Paper losses may consequently not be regarded as final and possess the potential to rebound (Shefrin and Statman, 1985). The idea that realization affects decision making has been tested in an experimental asset market (Weber and Camerer, 1998). When stocks are automatically sold after each period, the disposition effect is significantly reduced. The automatic selling procedure closes existing mental accounts, and stocks are no longer charged by prior experiences of gains or losses.[3] This means that after realizing lottery outcomes, a participant is effectively in the same decision situation as before entering the first lottery:

H1. After a gain or a loss is realized, risk-taking behavior will be similar as in a decision without prior history.

Barberis and Xiong (2009) study the implications of realized and paper outcomes as well. In two alternative models, they define prospect theory preferences either over total gains and losses or realized gains and losses. They discover that the model based on realized outcomes predicts the disposition effect more reliably.

2.2.2 Behavior After Gains

One main idea of the model is that prior gains serve as a cushion against losses that are felt less severely as long as they do not exceed prior gains. This is consistent with the "house money effect," predicting that people take more risk in the presence of a prior gain (Thaler and Johnson, 1990). When offered a risky lottery, individuals evaluate prior paper gains (house money) and the risky prospect jointly within the same mental account. Since the house money is integrated with future outcomes, losses can be offset and are perceived as less painful than usual.[4] Formally, losses up to the level of prior gains are not subject to loss aversion:

[3] Barberis et al. (2001) consider this plausible although they exclude this possibility for their analysis: "However, larger deviations – a complete exit from the stock market, for example – might plausibly affect the way [Z_t] evolves. In supposing that they do not, we make a strong assumption, but one that is very helpful in keeping our analysis tractable (p.13)." We assume that realizing all gains or losses is perceived similarly to an exit from the market.

[4] The idea is consistent with Arkes et al. (1994) who argue that windfall gains are spent more readily than other types of assets and Peng et al. (2013) who argue that the psychological value of losing parts of a prior gain is relatively low.

$$v(X_t, B_t, Z_t) = \begin{cases} X_t & \text{for } X_t \geq -Z_t \\ \lambda(X_t + Z_t) - Z_t & \text{for } X_t < -Z_t \end{cases} \tag{2.3}$$

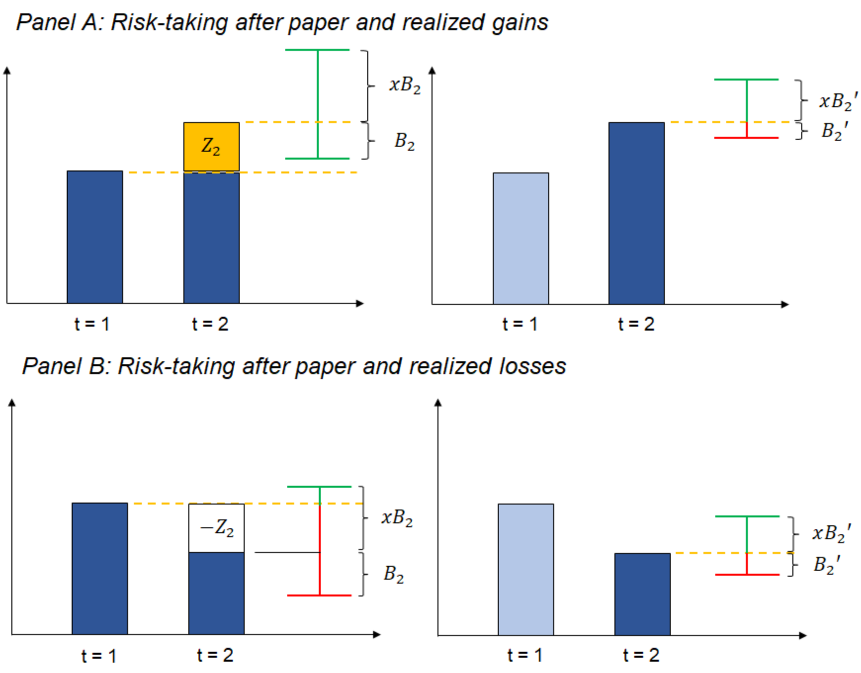
This means that losses up to Z_t are evaluated at the gentler rate of 1 instead of λ. Accordingly, a paper gain reduces loss aversion when compared to a realized gain. This is particularly true for small bet sizes $B_t < Z_t$, which do not jeopardize the whole gain cushion. Realization closes the respective mental account for prior gains and triggers the internalization of house money. Prior gains are no longer available to offset potential losses. Without integration, individuals evaluate a risky lottery separately from the previous gain and do not use the gentler rate of 1 instead of λ anymore. This reasoning is also graphically illustrated in Panel A of Figure 2.2. We hypothesize:

H2. After a paper gain people are more prone to take risks than after a realized gain.

H2a. They avoid bet sizes that run the risk to lose more than the sum of prior gains.

Hypothesis 2 may shed light on seemingly contradictory results in the empirical literature: Less risk taking after a prior gain versus more risk taking after a prior gain. While the house money effect predicts a higher propensity to gamble after a prior gain than before (or after a loss), the disposition effect describes the opposite behavior. Investors show a tendency to sell winning stocks too early and to keep losing stocks too long (Shefrin and Statman, 1985; Odean, 1998; Weber and Camerer, 1998). Intuitively, the trading behavior behind the disposition effect is in line with prospect theory (Kahneman and Tversky, 1979). A winning stock moves an investor into the gain domain of the prospect theory value function. As the value function is assumed to be concave for gains, it implies risk-averse behavior and a higher likelihood of selling the stock.

Figure 2.2: Risk-Taking After Paper and Realized Outcomes



Note: The figure illustrates risk-taking after gains in Panel A and after losses in Panel B depending on realization. For illustrative purposes, only two rounds of a lottery are displayed, and outcomes are either on paper (left diagrams) or realized after the first round (right diagrams). Each diagram plots the round of the lottery on the x-axis and the earnings on the y-axis. Endowments are the same in t=0, which then adjust depending on the outcome of the first lottery in t=1. In round two, the chosen investment B_2 determines the potential earnings indicated by the horizontal bars. Color coding shows whether outcomes are evaluated as gains (green) or losses (red). Whether an outcome is evaluated as a gain or loss depends on the mental account and its reference point. For example, in the left diagram of Panel A, the paper gain from the first lottery enters a newly opened mental account shown in yellow. Outcomes in round two are evaluated against this previous gain which offsets potential losses. The right diagram of Panel A shows the same situation when instead the gain is realized. The respective mental account is closed, the previous gain is internalized, and the reference point shifts to the new wealth level. In round two, there is no cushion against a potential loss which is indicated in red.

Further tests are similarly inconclusive for risk taking after gains. Weber and Zuchel (2005) show in lottery experiments that participants become more risk-seeking after a gain, while Franken et al. (2006) find in a gambling task that previous gains lead to less risk-taking. Clark (2002) does not find evidence in either direction following gains in a public goods experiment. However, bettors on the horse track take more risk after a previous gain (Suhonen and Saastamoinen, 2018), as do novice investors in the stock market (Hsu and Chow, 2013). Recently, Lippi et al. (2018) support this finding by showing that clients of an Italian bank engage in more risk-seeking behavior after unrealized gains. However, Coval and Shumway (2005) analyze the trading behavior of futures traders and find that traders with gains in the morning take less risky positions in the afternoon. In a similar setting, Frino et al. (2008) report the opposite result.

2.2.3 Behavior After Losses

When a mental account is in the red, i.e., a participant has experienced an overall loss, then the outcomes of a lottery are evaluated in the following way:

$$v(X_t, B_t, Z_t) = \begin{cases} -\lambda Z_t + (X_t + Z_t) & \text{for } X_t \geq -Z_t \\ \lambda X_t & \text{for } X_t < -Z_t \end{cases} \tag{2.4}$$

The expression represents the mirror image of the situation after gains and again reflects the idea of an open mental account in which a loss is not final. Gains that make up for prior losses are particularly attractive and are valued at a rate of λ. Barberis et al. (2001) assume that losses on the heels of prior losses are more painful than usual and let loss aversion rise in Z_t. However, the results by Imas (2016) for paper losses question this idea, as people take more risk after a series of losses. The traditional view inspired by prospect theory also favors higher risk-taking after losses (Kahneman and Tversky, 1979). While the channel in prospect theory is higher risk tolerance, in the piecewise linear (risk-neutral) utility function used here, it could manifest in a decreasing loss aversion parameter (consistent with a learning effect documented by Merkle (2020)). We thus depart from the assumption of higher loss aversion after a prior loss and instead propose a constant loss aversion parameter. The extent of loss chasing will depend on how people's preferences react to prior losses.

When offered a risky lottery, individuals evaluate prior paper losses and the risky lottery jointly within the same mental account. They thus evaluate further losses at the same rate as gains reducing these losses. By contrast, realization closes the respective mental account, internalizes the prior losses, and resets the reference point to $Z_t = 0$ (see also Panel B of Figure 2.2). Note that equations 2.3 and 2.4 simplify to equation 2.2 in this case. We thus expect participants to take more risk when confronted with a paper loss (mental account still open) than with a realized loss (mental account closed):

H3. After a paper loss people are more prone to take risks than after a realized loss.

H3a. They favor bet sizes that give them the opportunity to break even.

For risk-taking after losses similarly inconclusive empirical evidence as for gains has been found. There is strong empirical support for an increase in risk-taking after experiencing a loss, which has been demonstrated in the lab (Gneezy and Potters, 1997; Weber and Zuchel, 2005; Langer and Weber, 2008; Andrade and Iyer, 2009) as well as in the field (Coval and Shumway, 2005; Meier et al., 2020). Such loss chasing has been identified as a source for gambling problems (Zhang and Clark, 2020), and might be driven by impulsive action (Verbruggen et al., 2017). On the other hand, several studies report a decrease in risk-taking after losses (Massa and Simonov, 2005; Shiv et al., 2005; Frino et al., 2008). Imas (2016) points out how the different results can be reconciled by distinguishing paper losses and realized losses (in line with H3). The presented findings almost exclusively rely on positively skewed gambles, for other skewness environments, there is hardly any evidence (see also Nielsen, 2019).

Hypothesis 3a does not follow directly from the introduced theory, as gains are treated equally up to the point where they exceed prior losses ($X_t > -Z_t$). However, already Thaler and Johnson (1990) report such a break-even effect. Moreover, there is evidence that finally realizing an outcome is associated with an immediate burst of utility (Barberis and Xiong, 2012; Frydman et al., 2014). Such realization utility implies that agents also care about the level of Z_t, in particular when they anticipate that the respective mental account will be closed. In the experiment, the final lottery represents the last opportunity to influence cumulative outcomes Z_T which are automatically realized at the end of the experiment. Lotteries that allow changing the sign of Z_T should be especially attractive. A sufficiently large multiplier x, as found in positively skewed lotteries, usually allows to break even. Depending on accumulated losses, it might not even be necessary to increase risk.

2.2.4 The Realization Effect and Skewness

In our model, a positively skewed lottery is prone to the realization effect as it offers a high potential gain and limited loss. In the gain domain, the cushion provided by Z_t will be able to absorb most of a possible loss and induce risk-taking unless the mental account is closed. In the loss domain, the lottery almost always offers the chance to break even, as the multiplier x applied on the bet B_t is sufficiently high. This also induces risk-taking, which is why a strong realization effect can be expected for positively skewed lotteries independent of the prior outcome.

In contrast, symmetric and negatively skewed lotteries are characterized by a lower but more probable gain and a higher but less probable loss. A reasonable assumption is that probabilities and payoffs of the lotteries are altered simultaneously so that their expected payoff remains (about) constant.[5] It is then more likely that previous gains cannot completely cushion a potential loss, which might deter people from risk-taking. Figure 2.2 illustrates this by the size of the mental account balance Z_2 in period two relative to the bet size B_2 in period two. The smaller account balance Z_2 after an initial gain only allows for smaller bets if people do not want to risk their endowment. We predict no reaction to skewness for risk-taking behavior after realized gains, as it is independent of prior history (see H1). Consequently, the realization effect should be reduced.

In the loss domain, symmetric or negatively skewed lotteries offer less potential to break even. Initial losses ($-Z_2$) are larger relative to potential

[5] Changing skewness without adjusting payoffs would just make the lottery more and more attractive. This would increase risk-taking across the board and represents a less interesting case to study.

gains xB_2. However, it is still possible to recoup prior losses at least partly, making the lottery somewhat more attractive than after losses are realized and mental accounts are closed.

H4. The realization effect is reduced or absent for symmetric and negatively skewed lotteries.

Previous empirical studies have shown in various domains that skewness influences risk-taking and that positively skewed lotteries tempt individuals to engage in more risk-taking. For example, individual investors have a preference for lottery-type stocks, characterized by low prices, high volatility, and large positive skewness (Kumar, 2009). Further evidence for positive skewness-loving investment behavior comes from horse race betting and state lotteries (Golec and Tamarkin, 1998; Garrett and Sobel, 1999). This is in line with Grossman and Eckel (2015), who find increased risk-taking in an experimental study with positively skewed lotteries. While most of the literature on dynamic risk-taking concentrates on positively skewed lotteries, there are many situations in every-day decision making in which outcome distributions are less or not at all positively skewed. For example, investors in the stock market or corporate managers usually face less lottery-like investment opportunities. Given this gap in the literature on risk-taking for non-positively skewed lotteries, the second objective of this study is to investigate whether the realization effect can be generalized to symmetric and negatively skewed lotteries.

Our model is broadly consistent with the theory provided by Imas (2016). The common prediction is that risk-taking after a paper loss is higher than a) before a paper loss and b) after a realized loss. However, we explicitly model a mental account (represented by Z_t), while Imas (2016) invokes a mere shift in the reference point. This difference becomes apparent when deriving predictions for the gain domain. An agent with a paper gain might take *less* risk in his model compared to an agent with a realized loss or no history.[6] As this defies, for example, the presence of a house money effect, we find this approach not appealing for understanding behavior after gains.

In the main model by Imas (2016), the proof for the general existence of a realization effect after losses relies on features of a positively skewed lottery. The effect is not necessarily absent for symmetric or negatively skewed lotteries, but in these cases depends on preferences (e.g., the degree of loss aversion). Similar to our model, a reduced aggregate realization effect can be expected in a population with heterogeneous preferences. Both models rely on myopic decision makers, who take only the next round of a lottery into account. An alternative is allowing for people to make contingent plans on their investments after gains and losses (e.g., Barberis, 2012). Contingent plans may alter the existing skewness of asset returns, for example, make them more positively skewed by planning to cut losses. In Online Appendix A.1, we discuss such models in more detail.

2.3 Experimental Design and Results

The design of the experiments is based on Imas (2016), who studies a version of the investment lottery by Gneezy and Potters (1997). Participants receive a total endowment which can be invested over several rounds in the same lottery. In each round, participants can invest a maximum amount E in the lottery, which is a constant fraction of the total endowment. They thus decide on their lottery investment (B_t) and how much they want to invest risk-free ($E - B_t$). For simplicity, the risk-free investment provides no interest. With probability p, the lottery returns the invested amount times a multiple x, with probability $1 - p$ the investment is lost. A participant can thus either make a gain of $(x - 1)B_t$ or a loss of $-B_t$. The expected payoff in each round is:

$$p \cdot (xB_t + E - B_t) + (1 - p) \cdot (E - B_t) = E + (px - 1)B_t. \tag{2.5}$$

[6] This depends on the chosen parameters. As Imas (2016) considers only risk-taking behavior after losses, he does not explicitly derive these predictions. His model is neither intended nor tested to work in the gain domain.

Lotteries are structured in such a way that $px > 1$, which means that the lottery has a positive expected payoff, and the expected payoff increases in the bet size B_t. Otherwise, the lottery would be unattractive to risk-averse participants. After the investment decision is made, the outcome of the lottery is determined and revealed to participants. In the following round, the same lottery is played again. Importantly, investment possibilities in later rounds do not depend on prior payoffs as the maximum investment E is a constant fraction of the total endowment.

The total number of lottery rounds in all experiments is four. In the realization treatment, participants invest over three rounds, and outcomes are realized at the end of the third round. After this, an additional lottery takes place. In the paper treatment, all four rounds are played consecutively, and there is no special significance of the turn between the third and final round. However, to keep information between treatments constant, participants in both treatments are informed about their earnings on the screen at the end of the third round. The main analysis thus relies on the risk-taking behavior in the final round, as the first three rounds are identical between treatments.

2.3.1 Experiment 1

Design and Participants

In the first experiment, we replicate the original design by Imas (2016). In each round, participants decided how much to invest in a positively skewed lottery. The lottery succeeded with a probability of 1/6 and paid seven times the invested amount, or it failed with a probability of 5/6 and the invested amount was lost. Considering this experimental design, the conditions under which the realization effect occurs turn out to be arguably restrictive. Imas (2016) focuses his attention on sequences of prior losses, excluding all histories involving a gain.[7] In addition, the nature of the lotteries is such that participants bet on the throw of a six-sided die and win (seven-fold) if their predetermined "lucky number" comes up. This results in a positively

[7] In expectation, only $(5/6)^3 = 58\%$ of observations enter the analysis.

skewed lottery. In the first experiment, we extend the analysis to the gain domain, while in experiments two and three, we introduce different types of skewness.

Participants were randomly assigned to either a realization treatment or a paper treatment as described above. After entering the laboratory, each participant received an envelope which contained the endowment of EUR 8.00. The instructions asked participants to count the money (see Online Appendix A.2 for the experimental instructions). The lotteries were framed in terms of the throw of a six-sided die and always proceeded in the same way. First, each participant was randomly assigned a success number between 1 and 6, which was displayed on the computer screen. Then participants decided how much to invest in the lottery up to a maximum of EUR 2.00. As soon as all participants had entered the amount, the experimenter rolled a large die in front of the room. All participants received the opportunity to check whether the die was fair. If the success number matched the rolled number, the participant won the lottery and obtained seven times the invested amount (plus the amount invested risk-free). If the success number did not match the rolled number, the participant lost the invested amount and kept the amount not invested. For the next round, a new success number was assigned. As in the original experiment, all results of the die roll were written on a board in front of the room.

In the realization treatment, outcomes were realized at the end of the third round. Participants who lost money by that time took the lost amount out of the envelope and handed it back to the experimenter. Participants who won received additional money from the experimenter. After this, participants made one last investment decision in a final round and were paid accordingly. In the paper treatment, outcomes were not realized at the end of the third round. Outcomes were merely communicated on the screen as in the realization treatment, but no physical transfer of money took place.[8] At the end of round four, all outcomes were realized for both groups. As in the

[8] Screenshots of a representative lottery round in the experiment and of the earnings update after round three for both treatments are provided in the Online Appendix A.2.

original experiment, the time between rounds was normalized across treatments. Consistent with hypotheses H2 and H3, we predict that participants in the paper treatment (after gains and losses) will invest more in the final lottery than participants in the realization treatment.

Experiment one was programmed in z-Tree (Fischbacher, 2007) and conducted in the Mannheim Experimental Laboratory (mLab). We selected a sample size of $N > 200$ participants to obtain statistical power of at least 90% to detect an effect of the size of the original realization effect at the 5% significance level (Camerer et al., 2016). We recruited 203 people via ORSEE (Greiner, 2015) from a university-wide subject pool to participate in a study on decision making. Participants were on average 23 years old, and the number of female ($n = 108$) and male ($n = 95$) participants was relatively similar (see Table 2.1).

Table 2.1: Summary Statistics of Experiment Participants

	Experiment 1 mLab	Experiment 2 mLab	Experiment 3 mLab & AWI Lab
Number of participants	203	95	209
Gender (male=1)	0.47	0.43	0.43
Age	22.7	22.1	23.3
Semesters studied	6.00	5.38	6.61
Risk aversion (0−10)	5.22	4.28	4.10
Loss aversion	2.13	1.86	1.82
Time preference (0−10)	7.68	6.92	6.31
Financial literacy (0−8)	5.16	4.45	3.73
Illusion of control (1−5)	2.15	2.36	2.10
Cognitive reflection (0−7)	5.16	3.64	3.60

Note: The table presents means of demographic variables, preferences, and cognitive variables for participants in experiments 1-3. Gender is an indicator variable (male=1), age is measured in years, and a semester corresponds to half a year of study (at least undergraduate level). Risk aversion, loss aversion, time preference, financial literacy, illusion of control, and cognitive reflection are measured as described in Online Appendix A.3

Replication Results

We first examine the replication of the realization effect for losses. The analysis centers on the change of investment between rounds three and four, as realization takes place before round four. To test for the realization effect, we are mainly interested in three comparisons: The difference in the change of investment between the paper and realization treatment (between-treatment comparison) and the change of investment for each treatment separately (within-treatment comparisons). Panel A of Table 2.2 shows the amounts invested in the lottery for participants who have a total loss by the end of round three, which means that they lost in each of the first three rounds.[9] Investments do not differ significantly across treatments over the first three rounds. In the final round, participants in the paper treatment invest slightly more, while participants in the realization treatment invest less. This pattern is consistent with a realization effect as stated in hypothesis H3, which predicts a positive difference in differences (DiD $= 0.16$, $t(113) = 1.58$, $p = 0.12$).

However, compared to results of study one by Imas (2016) (DiD $= 0.38$, $t(51) = 3.19$, $p < 0.01$), our data show a less pronounced effect with respect to economic and statistical significance. The found effect size is 42% of the original effect size, which is smaller than the mean replicated effect size of 66% reported by Camerer et al. (2016) in a large-scale study on the replicability of laboratory experiments in economics. To further assess replicability, we apply confidence intervals and a meta-analysis they propose as standard measures. The original effect size is outside, but close to the upper bound of the 95% confidence interval of the replicated effect size $[-0.04, 0.34]$.

Interestingly, when focusing on the investment behavior within treatment, the realization effect we find is primarily driven by a decrease in risk-taking in the realization treatment (-0.12, $t(57) = 1.64$, $p = 0.11$), while the effect in the original data is primarily driven by an increase in risk-taking in the paper treatment. We can confirm that participants tend to take less risk after a realized loss, but we cannot replicate that participants increase risk-taking after a paper loss (0.04, $t(56) = 0.57$, $p = 0.57$). The magnitude of the decrease in risk-taking after a realized loss in the original study (-0.15)

[9] We follow Imas (2016) who restricts the sample to those participants who experienced three consecutive losses. Most participants who won once ended up in the gain domain due to the positive skewness of the lottery.

Table 2.2: Risk-Taking in the Positively Skewed Lottery

Panel A: Risk-taking after losses						
	Invested amount				Change	
Treatment	Round 1	Round 2	Round 3	Round 4	R4−R3	N
Paper	0.98	0.91	0.78	0.82	0.04	57
					(0.57)	
Realization	0.90	0.73	0.80	0.68	−0.12	58
					(1.64)	
Difference	0.08	0.18	−0.02	0.14	0.16	
	(0.72)	(1.56)	(0.13)	(1.05)	(1.58)	

Panel B: Risk-taking after gains						
	Invested amount				Change	
Treatment	Round 1	Round 2	Round 3	Round 4	R4−R3	N
Paper	0.94	0.73	0.71	0.84	0.13	35
					(1.75)	
Realization	0.71	0.77	0.73	0.64	−0.09	36
					(1.27)	
Difference	0.23	−0.04	−0.02	0.20	0.22	
	(1.78)	(0.22)	(0.13)	(1.32)	(2.16)	

Note: The table shows the average invested amounts in the lottery for all rounds of experiment 1 (in Euro). Panel A is restricted to participants who lost in the first three rounds of the experiment, Panel B shows averages for all participants with at least one gain in the first three rounds. Both panels show results by treatment (paper and realization) and differences between treatments. Change is the difference between the investment in the final round and round three. N provides the number of participants for each treatment-outcome combination. T-values of a two-sided t-test are shown in parentheses.

is well inside the 95% confidence interval in the replication $[-0.25, 0.02]$. However, the increase in risk-taking after a paper loss in the original study (0.23) is not compatible at the 95% confidence level with the replication $[-0.09, 0.17]$.

In other words, we do not find loss chasing in the paper treatment, which ultimately explains the overall less pronounced realization effect for losses as compared to Imas (2016). One reason for the non-robust results after paper losses might be that the positively skewed lottery offers participants the chance to break even without necessarily having to increase risk-taking. Whether or not some participants still increase their risk-taking will depend on their prospect theory preference parameters.[10]

When comparing the invested amount in round four to the invested amount in round one, we find that participants are more risk-averse after a realized loss than without any prior outcome (-0.22, $t(57) = 2.49$, $p = 0.02$). This is inconsistent with hypothesis H1, but in line with the idea of Barberis et al. (2001), who argue that individuals become more sensitive to future losses after a previous loss. In general, the changes in risk-taking between rounds three and four are not particularly large when compared to the changes observed for earlier rounds (see Online Appendix A.4). We find some significant results for earlier rounds across all three experiments, but we cannot identify a systematic pattern behind these changes. Significance occurs mostly between round one and round two, which suggests that participants try out the lottery first before making considerable adjustments to their bet size. Importantly, by round three, risk-taking behavior is very similar between treatments.

As a further test for replication, we pool our data with the original data by Imas (2016). Thus, we are able to obtain a meta-analytic estimate of the effect (Camerer et al., 2016). In the pooled data we obtain a strongly significant realization effect after losses (DiD $= 0.24$, $t(165) = 3.10$, $p < 0.01$). We conclude that the evidence on the outcome of the replication is mixed. We

[10] The result that loss chasing is parameter-dependent, but risk-taking after a realization is not is also present in the framework by Barberis (2012) and Imas (2016).

find a weaker but directionally consistent realization effect after losses.

Results For Gains

Next, we examine participants with a gain at the end of the third round. Given the considerable upside potential of the lottery, most participants who succeeded in at least one lottery faced positive net earnings at the end of the third round. The overall sample of 203 participants splits into 115 participants with a loss by the end of round 3 analyzed above, 71 participants with a gain by the end of round three, and 17 participants who have zero net earnings by the end of round three (due to not investing in the lottery at all). Of the 71 participants with a gain, 65 won the lottery once, and 6 won twice.[11]

Panel B of Table 2.2 shows the invested amounts for these participants. In most cases, changes in investment in rounds one to three do not differ significantly across treatments.[12] Consistent with Hypothesis H2, the change in risk-taking between rounds three and four is significantly different between the paper and the realization treatment (DiD = 0.22, $t(69) = 2.16$, $p = 0.03$). This realization effect for gains is somewhat larger than the replicated effect for losses. Within treatment, participants in the paper treatment take significantly more risk (0.13, $t(34) = 1.75$, $p = 0.09$), while participants in the realization treatment take less risk (-0.09, $t(35) = 1.27$, $p = 0.21$), yet statistically insignificant. However, in line with hypothesis H1, individuals invest similarly after a realized gain compared to the case of no prior outcome. The difference between the invested amount in round one and round four after a realized gain is insignificant (0.06, $t(35) = 0.61$, $p = 0.54$).

To back-up this finding, we turn again to the original data by Imas (2016), which has not been analyzed with regard to risk-taking after gains. As before, we only use observations of participants with a gain at the end of round three. Despite the relatively small sample size (N=24), we nevertheless find evidence for a realization effect after gains in his data. As shown in Table 2.3,

[11] Table A.4 in Online Appendix A.4 provides more details about participants' average earnings after round three conditional on the outcomes in each round.

[12] We also do not find significant changes in investment before and after the round in which a participant wins across treatments; see Online Appendix A.4, Table A.5.

participants take more risk in the paper treatment than the realization treatment considering changes between rounds three and four. Consistent with the results from our experiment, the realization effect is positive and statistically significant (DiD $= 0.55$, $t(22) = 2.29$, $p = 0.03$). Within treatment, participants take more risk after a paper gain (0.47, $t(8) = 1.99$, $p = 0.08$) and tend to take less risk after a realized gain (-0.08, $t(14) = 0.67$, $p = 0.51$).

Table 2.3: Risk-Taking After Gains in Imas (2016) Study 1

	Invested amount				Change	
Treatment	Round 1	Round 2	Round 3	Round 4	R4−R3	N
Paper	0.81	0.78	0.75	1.22	0.47 (1.99)	9
Realization	0.83	0.68	0.83	0.75	−0.08 (0.67)	15
Difference	−0.02 (0.17)	0.10 (0.54)	−0.08 (0.38)	0.47 (2.00)	0.55 (2.29)	

Note: The table shows the average invested amounts in the lottery for all rounds of the experiment (in US-Dollar) for all participants with at least one gain in the first three rounds. Data are obtained from the AER website. Displayed are results by treatment (paper and realization) and differences between treatments. Change is the difference between the investment in the final round and round three. N provides the number of participants for each treatment. T-values of a two-sided t-test are shown in parentheses.

When we pool the data from both studies, we find a strong realization effect for gains (DiD $= 0.29$, $t(93) = 2.96$, $p < 0.01$). We thus find experimental evidence for a realization effect for gains in two independent samples. The studies were conducted with student populations from different universities, in different countries, and at different points in time. While the p-value in both samples is similar ($p = 0.03$), the combined evidence provides far stronger support to hypothesis H2.

Irrespective of whether the prior outcome is a gain or loss, risk-taking is thus higher when outcomes remain unrealized. This finding allows us to analyze the existence and strength of the effect independent of the sign of the prior outcome. Therefore, we run OLS regressions for the entire sample with the change in invested amount between rounds three and four as the dependent variable. We include a treatment indicator taking a value of one for the realization treatment. Table 2.4 shows in column (1) the results of the baseline regression. We observe a strong realization effect, with those in the realization treatment taking significantly less risk. Unsurprisingly, the economic magnitude is in between those estimated for gains and losses separately. The positive constant provides evidence for an increase in risk-taking in the paper treatment. Controlling for gains and losses after round three by a gain indicator (gain=1) does not affect the main result (Column 2). Interacting the treatment and gain variables allows us to test whether the realization effect is stronger after previous gains or losses. The negative but insignificant coefficient of the interaction term hints at a stronger realization effect after gains.

Table 2.4: The Realization Effect for Gains and Losses

	Data from Experiment 1 Change in invested amount			Data from Imas (2016) Change in invested amount		
	(1)	(2)	(3)	(4)	(5)	(6)
Realization	-0.182^{***}	-0.183^{***}	-0.165^{**}	-0.420^{***}	-0.437^{***}	-0.390^{***}
	(0.066)	(0.066)	(0.082)	(0.107)	(0.108)	(0.128)
Gain		0.046	0.071		0.134	0.231
		(0.069)	(0.097)		(0.118)	(0.183)
Gain x Realization			-0.051			-0.166
			(0.138)			(0.240)
Constant	0.083^{*}	0.068	0.059	0.295^{***}	0.264^{***}	0.242^{***}
	(0.046)	(0.051)	(0.056)	(0.077)	(0.082)	(0.088)
Observations	203	203	203	81	81	81
R^2	0.037	0.039	0.039	0.163	0.177	0.182

Note: The table shows the results of OLS regressions with the change in the invested amount between rounds three and four as the dependent variable based on data from experiment 1 and data from Imas (2016) study 1. Realization is an indicator variable taking a value of one for the realization treatment. Gain is an indicator variable taking a value of one for participants with a prior gain. Gain x Realization is the interaction between the two variables. Robust standard errors are shown in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

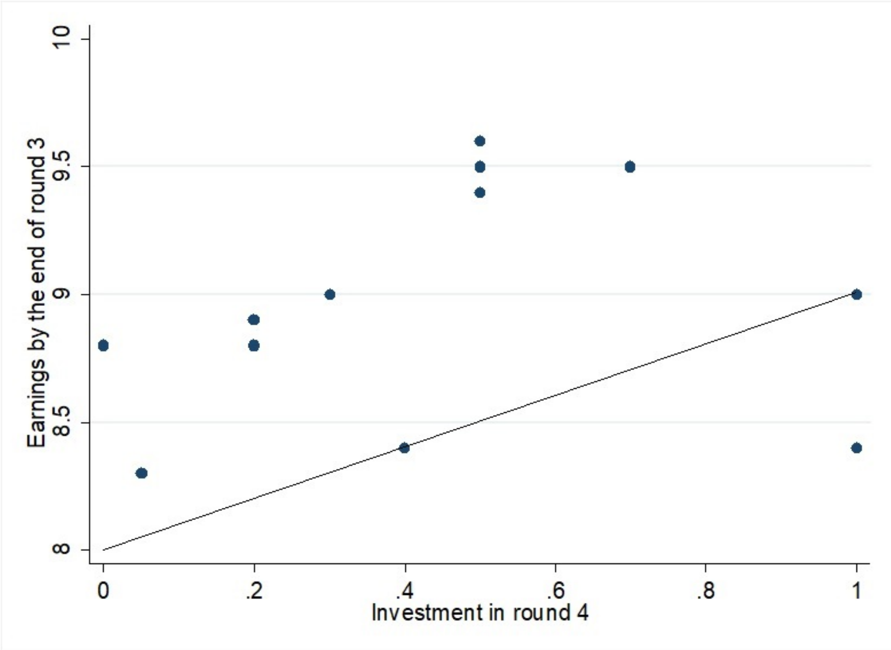
We run the same regressions on the data from Imas's (2016) study 1. Columns (4) to (6) in Table 2.4 display the results. A strong realization effect also exists in his data independent of prior gains and losses. The effect in his data is even more pronounced in economic magnitude than in our data.

The combined effect independent of the prior outcome in the pooled data is (DiD $= 0.25$, $t(283) = 4.38$, $p < 0.001$).

A relevant assumption about the realization effect is that people are less loss averse for money they keep in the mental account for house money (paper gains) than for their own money that they keep in a different mental account (realized gains). This assumption has testable implications for the amount people are willing to bet (hypothesis H2a). We predict that participants avoid bet sizes that run the risk losing more than the sum of prior gains. Since participants can invest up to EUR 2.00 in each round and lose at a maximum their invested amount, the subsample of interest are participants who have earnings between EUR 8.00 and EUR 10.00 after round three (i.e., gains between EUR 0 and EUR 2). If mental accounting is important, participants are expected not to invest more than their current paper gains (house money) in round four. Figure 2.3 plots the earnings after round three against the invested amount in round four. The maximum invested amount of participants in this subsample was EUR 1.00. All dots above the line represent participants who invest less than their house money in round four, which restricts their potential losses to less than their previous gains. Dots below the line represent participants who risk to lose more than their prior gains. Consistent with hypothesis H2a, 11 out of 12 participants invest less or exactly as much money as they previously gained.

Essential for the realization effect is that the used realization mechanism is effective in closing a mental account. We tested an alternative realization mechanism in two versions of an online experiment, one of which is an identical replication of the online study in Imas (2016). As a physical transfer of money is not feasible online, participants in the realization treatment initiate a transfer of money between accounts by typing the command "closed." We successfully replicate the realization effect using this alternative realization mechanism in the original design by Imas (2016) but discover that the effect is rather fragile when modestly changing the design. We find that the framing of how the last round is related to the preceding three matters for whether risk-taking increases or decreases in the realization treatment of the online experiment.[13] We conclude that in an online environment, proper realization is more difficult to achieve, and mental accounts may remain open using the described procedure. Complete results are reported in Online Appendix A.5.

Figure 2.3: Testing the Mental Accounting Assumption



Note: The figure plots the earnings by the end of round three against the investment in round four for each participant who has earnings between EUR 8.00 and EUR 10.00 by the end of round three. Participants with earnings below EUR 8.00 are excluded as they made a loss and participants with earnings above EUR 10.00 are excluded as they cannot lose more than what they previously gained (given than the investment per round cannot be more than EUR 2.00 which also presents the highest possible loss per round). All dots above the diagonal line represent participants who invest less than what they previously gained, and all dots below the diagonal line represent participants who invest more than what they previously gained.

2.3.2 Experiment 2 and 3

Design and Participants

Experiment two and three address the question of whether the realization effect depends on the skewness of the underlying investment opportunity. We

[13] Effects of different exchange media (cash, tokens, e-coins) are examined in a similar experimental paradigm by Stivers et al. (2020). They find that reduced moneyness alters risk-taking behavior as well.

take the same experimental design as in experiment one except for the investment opportunity, which we change to either a symmetric (experiment two) or a negatively skewed lottery (experiment three). In line with hypothesis H4, we predict a reduced or absent realization effect in these settings.

Participants were again endowed with EUR 8.00 at the beginning of the experiment and could invest up to EUR 2.00 in each of four subsequent lottery rounds. In experiment two (symmetric lottery), participants could invest in a lottery that succeeded with a probability of 1/2 and paid 2.33 times the invested amount. With a probability of 1/2, the lottery failed and the invested amount was lost. Instead of one success number for the role of the die, participants received three success numbers. In experiment three (negative skewness), participants could invest in a lottery which succeeded with a probability of 5/6 and paid 1.4 times the invested amount or failed with a probability of 1/6. Instead of a success number, they received one failure number.

The multiplier for the gain case was adjusted to keep the expected payoff of each lottery equal to the expected payoff of the lottery in experiment one. While the objective of experiment three was to create a mirror image of the original positively skewed lottery, a complete reversal of gains and losses was infeasible as losses cannot exceed the endowment (by laboratory rules). Instead of a seven-fold loss, we thus have to restrict the loss to the invested amount. Still, participants are expected to experience many small gains and occasionally (relatively) large losses.

As before, participants were randomly assigned to either a realization treatment, in which outcomes were realized by the end of the third round or a paper treatment. The procedure in the two treatments was the same as in experiment one. Both experiments were conducted in the Mannheim Experimental Laboratory (mLab) and the AWI Experimental Laboratory at the University of Heidelberg.[14] We recruited 304 participants in total, 95 of them were assigned to experiment two and 209 to experiment three. A smaller

[14] The additional lab was added to obtain a larger subject pool. Participants who had already participated in experiment one were excluded.

sample size was required in experiment two as a symmetric lottery generates sufficient observations for gains and losses more easily. The demographics of participants in experiments two and three are similar to those in experiment one (see Table 2.1).

Results of Experiment 2 (Symmetric Lottery)

We first analyze the investment behavior of participants who accumulate a loss by the end of round three. Panel A of Table 2.5 presents the invested amounts for those participants. Investments do not differ significantly across treatments in the first three rounds. Comparing the changes in investment between rounds three and four across treatments, the realization effect points in the expected direction ($DiD = 0.08$, $t(35) = 0.54$, $p = 0.59$), but is small and statistically insignificant. When analyzing the invested amounts within each treatment, we find that participants who have a paper loss by the end of round three do not increase their investment (0.00), and participants who have a realized loss tend to slightly decrease their investment (-0.08, $t(14) = 0.73$, $p = 0.48$). Participants thus seem not to invest differently after a paper or a realized loss. In particular, we do not observe more risk-taking after paper losses.

Panel B of Table 2.5 shows the invested amounts of participants with an accumulated gain by the end of round three. Similar to losses, the realization effect cannot be observed in the symmetric lottery setting ($DiD = -0.03$, $t(55) = 0.18$, $p = 0.86$) for gains. The change in investment between rounds three and four in the paper treatment (-0.11, $t(26) = 0.69$, $p = 0.50$) and the realization treatment (-0.08, $t(29) = 0.89$, $p = 0.38$) points in the same direction. After a paper as well as a realized gain, participants tend to invest similarly. Consistent with hypothesis H4, we find no evidence for a realization effect after gains or losses when the investment opportunity is symmetric.

Looking at investments on participant level, we find that 53% of the participants do not change their invested amount between rounds three and four (fairly independent of treatment). Any overall effect would thus have to rely on a subset of participants to make strong changes in their investments. We also find that the absence of the realization effect does not depend on the round(s) in which participants win in the lottery.

Table 2.5: Risk-Taking in the Symmetric Lottery

Panel A: Risk-taking after losses						
	Invested amount				Change	
Treatment	Round 1	Round 2	Round 3	Round 4	R4−R3	N
Paper	1.40	1.37	1.45	1.45	0.00	22
					(0.00)	
Realization	1.57	1.50	1.53	1.45	−0.08	15
					(0.73)	
Difference	−0.17	−0.13	−0.08	0.00	0.08	
	(0.94)	(0.67)	(0.39)	(0.04)	(0.54)	
Panel B: Risk-taking after gains						
	Invested amount				Change	
Treatment	Round 1	Round 2	Round 3	Round 4	R4−R3	N
Paper	1.35	1.15	1.19	1.08	−0.11	27
					(1.75)	
Realization	1.43	1.56	1.38	1.30	−0.08	30
					(1.27)	
Difference	0.08	−0.41	−0.19	−0.22	−0.03	
	(0.55)	(2.97)	(1.17)	(1.16)	(0.18)	

Note: The table shows the average invested amounts in the lottery for all rounds of experiment 2 (in Euro). Panel A is restricted to participants who have a loss by the end of round three, Panel B shows averages for all participants who have a gain by the end of round three. Both panels show results by treatment (paper and realization) and differences between treatments. Change is the difference between the investment in the final round and round three. N provides the number of participants for each treatment-outcome combination. T-values of a two-sided t-test are shown in parentheses.

Finally, we test whether participants in the paper treatment do not increase their investment after a loss because their losses are too high to break even in the final lottery. In contrast, the positively skewed lottery always allowed to break even. We split the sample of participants with accumulated losses into those who have earnings by the end of round three that are smaller than EUR 5.34 and those who have earnings between EUR 5.34 and EUR 8.00 (the highest possible gain in the final lottery is $2.33 * 2 - 2 =$ EUR 2.66). Despite the resulting small sample size, we find that participants with paper losses tend to invest differently depending on whether break-even is possible or not. Those who cannot break even tend to decrease the invested amount in round four by on average EUR 0.38, whereas participants who can break even tend to increase the invested amount by EUR 0.11. That people favor bet sizes that allow them to break even is consistent with hypothesis H3a. However, given the small sample size of participants with a paper loss (N=22), the effect remains insignificant and has to be interpreted with caution.

Results of Experiment 3 (Negatively Skewed Lottery)

We again start by examining the investment behavior of participants who accumulated a loss by the end of round three. Most of these participants lost only once but remained in the loss domain. Panel A of Table 2.6 shows the investments in all rounds for these participants by treatment. Levels and changes in investment between rounds do not differ significantly across treatments. Considering the difference of the changes in investment from round three to round four across treatments, the realization effect points in the expected direction ($DiD = 0.05$, $t(68) = 0.36$, $p = 0.72$), but is small and statistically insignificant. Participants in both treatments react similarly to a loss by slightly reducing their investments (-0.04, $t(31) = 0.45$, $p = 0.66$ and -0.09, $t(37) = 1.17$, $p = 0.25$).

The investments for participants with gains by the end of round three are displayed in Panel B of Table 2.6. As for losses, we do not find a significant realization effect for gains in this setting. Participants do not invest differently after a paper and a realized gain (0.00 and 0.00). In fact, the investments on average do not change at all between round three and round four. Results change very little if we restrict the sample to those participants who experience three successes in a row (N=121). In line with hypothesis H4, we do not find evidence for a realization effect when participants invest in a negatively skewed lottery. This supports theoretical predictions that the realization effect depends on the positive skewness of the lottery.

Table 2.6: Risk-Taking in the Negatively Skewed Lottery

Panel A: Risk-taking after losses						
		Invested amount			Change	
Treatment	Round 1	Round 2	Round 3	Round 4	R4−R3	N
Paper	1.53	1.66	1.67	1.63	−0.04	32
					(0.45)	
Realization	1.70	1.79	1.78	1.69	−0.09	38
					(1.17)	
Difference	−0.17	−0.13	−0.11	−0.06	0.05	
	(1.24)	(1.11)	(0.91)	(0.49)	(0.36)	

Panel B: Risk-taking after gains						
		Invested amount			Change	
Treatment	Round 1	Round 2	Round 3	Round 4	R4−R3	N
Paper	1.45	1.60	1.60	1.60	0.00	70
					(0.00)	
Realization	1.58	1.71	1.67	1.67	0.00	64
					(0.00)	
Difference	−0.13	−0.09	−0.07	−0.07	0.00	
	(1.38)	(1.15)	(0.71)	(0.76)	(0.00)	

Note: The table shows the average invested amounts in the lottery for all rounds of experiment 3 (in Euro). Panel A is restricted to participants who have a loss by the end of round three, Panel B shows averages for all participants who have a gain by the end of round three. Both panels show results by treatment (paper and realization) and differences between treatments. Change is the difference between the investment in the final round and round three. N provides the number of participants for each treatment-outcome combination. T-values of a two-sided t-test are shown in parentheses.

2.4 Conclusion

In this paper, we examine whether and under which conditions a distinction between realized and unrealized prior outcomes leads to differential subsequent risk-taking. We formalize our thoughts in a model of mental accounts that people use to keep track of their paper gains and losses. A mental account is closed when an investment episode ends and outcomes are realized. For losses, recent experimental evidence finds that individuals take less risk after a realized loss and more risk after a paper loss, which is referred to as the realization effect. It is tempting to conclude from this result that realization per se has a strong effect on subsequent behavior. We first ask whether – as our theory predicts – the finding generalizes to the gain domain, i.e., whether a realization effect can also be observed after gains. Second, we identify positive skewness as a necessary condition to observe the realization effect. As such, our results show that conclusions about the universality of the realization effect have to be drawn with some caution.

The main objectives and findings from our study can be summarized as follows: We replicate the result by Imas (2016) for losses, extend the analysis to gains and test the boundary conditions of the effect with respect to the skewness of the investment opportunity. Using the same experimental setting and a larger sample size than the original study, we show that the realization effect also exists for gains. We thus show that the framework of realization is independent of the sign of prior outcomes as it holds not only for losses but also for gains. However, at the same time, the effect turns out to be sensitive to changes in the skewness of the underlying investment opportunity. We do not find differential risk-taking after paper and realized outcomes for non-positively skewed lotteries. This finding documents the importance of learning more about the conditions under which the effect arises and informs judgments about its external validity.

The results confirm theoretical predictions that a realization effect mostly occurs in positively skewed lotteries. The analysis of risk-taking in non-positively skewed lotteries, in particular, in negatively skewed lotteries has received less attention in the literature. One recent exception is contemporaneous work by Nielsen (2019), who examines risk-taking under negatively skewed outcome distributions for realized and unrealized losses. Using a different realization mechanism and a different investment task in which individuals can choose the skewness of their preferred option, she finds no realization effect for negatively skewed outcomes. Her finding is in line with our results and further supports the conclusion that the realization of outcomes does not always induce differences in risk-taking compared to settings in which outcomes remain unrealized.

Chapter 3

The Portfolio Composition Effect *

3.1 Introduction

How do investors evaluate their portfolio investment decisions? In models such as portfolio theory (Markowitz, 1952), portfolio evaluation reduces to two integral parameters: expected returns and variance of returns. The fundamental rule concerning choice of portfolio is that wealth between two financial securities (e.g. stocks or even entire portfolios) should be allocated such that the overall expected return is maximized for any given variance of returns. While this is an excellent normative advice given the assumptions of the model, various studies have shown that individuals' actual investment behavior substantially diverges from what basic portfolio theory implies (Barber and Odean, 2000; Benartzi and Thaler, 2001; Barber and Odean, 2013).

In this article, we study one specific allocation decision of the portfolio choice problem, namely, how individuals' investment decisions in given, predetermined portfolios are affected by different levels of performance information. In particular, we ask whether investors consider information from both, the *overall portfolio level* as well as *the individual stock level*, when evaluating a portfolio's performance and, if so, how this two-level informational

* Authors: Jan Müller-Dethard and Martin Weber. Jan Müller-Dethard and Martin Weber are at the University of Mannheim. For valuable comments, we thank Alex Imas, participants of the SPUDM 2019, Annual Meeting of the SJDM 2019, MPI Workshop 2019 in Bonn, and seminar participants at the University of Mannheim. We gratefully acknowledge financial support by the Stiegler Foundation, the Reinhard Selten Scholarship by the GfeW, and the German Research Foundation (DFG grant WE 993/15-1).

setup ultimately affects their portfolio investment decisions. We document a new stylized fact about how individuals evaluate and allocate funds across portfolios: a person's willingness to invest in a portfolio depends on the portfolio's composition of winner and loser stocks. We term the effect that individuals allocate larger funds to those portfolios which consist of more winner than loser stocks than to alternative portfolios with more loser than winner stocks – despite *identical overall realized* and *expected returns* as well as *variance* – the portfolio composition effect.

We explain its occurrence by combining two well-established frameworks from psychology which are category-based thinking and mental accounting (Rosch and Lloyd, 1978; Thaler, 1999; Shefrin and Statman, 1987). Investors assign stocks to individual mental accounts, whereby they are reluctant to integrate outcomes across different accounts (Frydman et al., 2017). However, once they evaluate a whole portfolio, and are presented with all information together, they deviate from this strong form of narrow framing and engage in a "semi-joint" evaluation of individual stock outcomes. This means, they assign stocks based on the most salient difference across them, into one of two categories which is either "winner" stocks or "loser" stocks. Given the complexity of full integration and individuals' reluctance to integrate outcomes across different mental accounts which are assigned to different or one and the same category, they simply engage in a counting heuristic to evaluate a portfolio investment decision. This is they count the number of mental accounts (i.e. stocks) which are assigned to either the "winner" or the "loser" category, compare these values to one another, and evaluate portfolios based on their composition of winner and loser stocks rather than their overall (expected) return and overall (expected) risk.

To test this framework and as such to investigate how different levels of performance information influence portfolio investment decisions, we define a simple, counting-based measure calculated from performance information on the individual stock level. This is the number of stocks with positive return since purchase (hereafter called "winner stocks") relative to the number of stocks with negative return since purchase (hereafter called "loser stocks").

Using data from a series of investment experiments with almost 1200 participants, we show that the proposed portfolio composition measure influences participants' willingness to invest in a pre-determined portfolio. Motivated by this finding, we turn to financial market data and show that portfolio composition matters not only for individual investment decisions in an experimental setting, but also for the demand of exchange-traded funds replicating leading equity market indices.

In chronological order, we first show that the documented effect exists in an arguably simple investment task in which realized portfolio returns are identical. Within our baseline experimental scenario, individuals invest on average 26% (22%) more of their endowment in a portfolio which consists of 70% winner/30% loser stocks than in an alternative portfolio with identical realized positive (negative) return, but the reversed composition of 30% winner/70% loser stocks. Participants are also more optimistic in their return expectations and report lower risk evaluations for those portfolios which consist of more winner than loser stocks.

Second, we try to get rid of the effect by controlling participants' beliefs about expected returns and variance. In particular, we test whether the effect still persists if portfolios are identical not only with respect to realized returns, but also with respect to expected returns and variance. We rerun the baseline experiment, but explain the underlying data generating process of returns to participants. In the spirit of a Bayesian updating task, participants can now learn about the expected returns of each portfolio before they make an investment decision. Still, we find a strong portfolio composition effect among those participants who state the same beliefs about expected returns.

Finally, we put the effect to a severe test. This means, we (i) extent the learning phase prior to the investment decision, (ii) provide computational support for the calculation of expected returns, and (iii) clearly display both, the resulting expected returns as well as the variance of each portfolio. Importantly, we design portfolios in a way that there is a unique mean-variance efficient allocation which suggests an equal split of wealth between portfolios. Even in this setting, participants invest more in the more favorably composed portfolio and by doing so choose a mean-variance suboptimal allocation. Compared to the baseline result, the effect gets even stronger with a 43% larger investment of the endowment in the 70% winner/30% loser portfolio relative to the alternative portfolio with identical realized and expected return as well as variance, but the reversed portfolio composition.

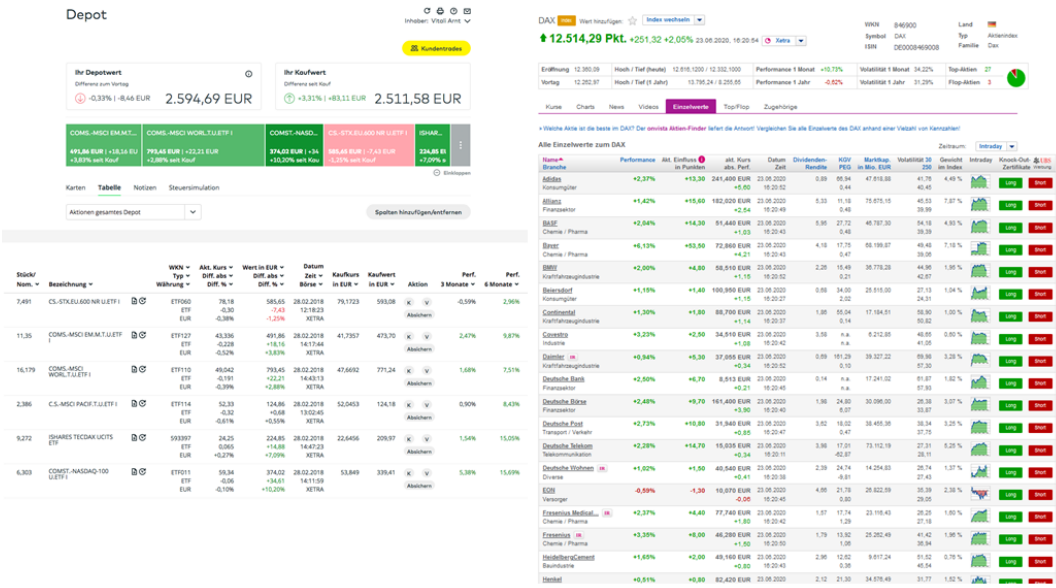
Taken together, we show experimentally that a portfolio's composition of winner and loser stocks affects an investor's willingness to invest in a portfolio. Specifically, this effect persists when controlling for investors' beliefs about expected returns and variance, and as such is not predicted by theories that assume mean-variance efficient portfolio selection (Markowitz, 1952). Consistent with our framework of category-based thinking and mental accounting, our findings suggest that individuals evaluate overall portfolio investment decisions not only on the portfolio level, but also on the individual stock level. It seems to matter to investors how an overall portfolio return has been achieved with respect to the performance of its individual components.

In a next step, we apply our findings on the evaluation of portfolios from a controlled experimental setting to real market data. In particular, we investigate whether historical fund flows of exchange-traded funds on leading European and North-American equity market indices from the period 2016-2019 are affected by the index composition of winner and loser stocks. Leading equity market indices represent ideal portfolio settings to test our hypothesis as they resemble relatively stable and transparent predetermined portfolios with respect to the members of the index over time. Moreover, market indices capture a lot of attention in the media and press of the respective country since they are often referred to as indicators of a country's economy.

We estimate that our portfolio composition measure, defined as the number of winner stocks on day *t* divided by the sum of the number of winner and loser stocks on day *t*, is positively correlated to fund flows on the subsequent day *t+1*. Across all leading equity market indices in our sample, we estimate that a portfolio composition of 100% winner stocks leads on average to $1,119,000 higher inflows on the subsequent day than a portfolio composition of 50% winner and 50% loser stocks. In relative terms, this inflow presents roughly 19% of the average daily fund inflow of an ETF in our sample. Interestingly, when splitting our sample by region, we do not find a similar effect for the two North-American market indices (Dow Jones Industrial Average and S&P500/TSX). Several robustness analyses show that the effect is of rather short-term, daily nature, does not crucially depend on the tails of the portfolio composition distribution, and persists when controlling for the mathematically related measure of return dispersion. In light of the determinants of index ETF fund flows (Elton et al., 2004; Clifford et al., 2014), our findings suggest that an index composition of winner and loser stocks affects future fund flows of the underlying ETF in addition to the index return. In essence, the composition of winner versus loser stocks in a portfolio matters not only for individual investment decisions in an experimental environment, but also for net flows of exchange-traded funds on leading equity market indices.

Throughout this paper, we argue that the evaluation of portfolio investment decisions is impacted by information on how the entire portfolio performs as well as by information on how each individual position in the portfolio performs. However, this reasoning implies that investors receive or at least have access to this information (on the portfolio level as well as on the individual stock level) when they evaluate their pre-determined (e.g. index funds) or self-selected portfolios of stocks. A look at how performance information is displayed by most online brokers and financial websites gives indication that this is indeed the case. The left part of Figure 3.1 shows exemplary which performance information investors usually receive by their online brokers when they log into their accounts. Performance information is provided on the overall portfolio level (e.g. the current portfolio value and the purchase value) as well as on the individual asset level (e.g. the return of each position in the portfolio). The information is similarly displayed if investors search online for the performance of pre-determined portfolios such as for example equity market indices. The right part of Figure 3.1 shows exemplary which performance information of the leading German equity market index DAX 30 an investor receives on the publicly available financial website onvista. Again, the overall portfolio (index) performance as well as the performance of each stock is clearly displayed.

Figure 3.1: How Portfolio Performance Information Is Displayed



Note: The left part of the figure shows a screenshot of how performance information of a portfolio is usually displayed to investors by online brokers (here: Comdirect). The right part of the figure shows how performance information of leading equity market indices is presented to investors on most financial websites (here: onvista).

In addition to this, the format of how performance information is displayed to investors suggests that they may easily gain an impression of a portfolio's composition of winner and loser stocks. Especially, the color coding of gains and losses facilitates the distinction between winner and loser stocks. Some financial websites report composition measures similar to ours. For various equity market indices, the financial website onvista depicts the number of "Top stocks" (i.e. winners) and "Flop stocks" (i.e. losers) of an index in a pie chart close to the overall index performance (see Figure 3.1).

This paper contributes to several strands of literature. First, we contribute to research on household and retail investor trading behavior in financial markets. Research in this field has advocated – for a long time – the simple assumption that investors consider stocks in a portfolio in isolation, so to say, detached from one another (Frydman et al., 2017). In particular, the role of the portfolio for investment decisions on the individual-stock level has widely been ignored. So far, most analyses of well-known trading patterns (e.g. the disposition effect) focus on individual assets rather than the portfolio. This is especially questionable not at least because many households and retail investors hold portfolios of assets (self-selected or predetermined). A paper which takes a step in this direction and analyzes trading behavior by taking the portfolio setting into account is Hartzmark (2015). He shows that a selected stock in a portfolio is traded differently depending on how the other stocks in the portfolio perform (e.g. the rank effect). In another study, An et al. (2019) find that the portfolio's overall return matters for individual stock trading (e.g. the portfolio-driven disposition effect). Even though, our paper focuses on the question whether the willingness to invest in a predetermined portfolio depends on the performance of its stocks, our findings may also have implications for how individual stocks are traded given the performance of the portfolio, in particular the composition of winner and loser stocks of the portfolio. As such, the proposed measure of portfolio composition might not only help to better understand households' and retail investors' buying and selling decisions of entire portfolios, but also of individual assets within portfolios.

Beyond individual-stock trading behavior in the context of self-selected portfolios, our findings have implications for portfolio-type assets such as ETFs and mutual funds. So far, a vast majority of studies has examined how retail investors evaluate and trade individual assets. The main focus, by and large, has been on individual stocks (Odean, 1998; Barber and Odean, 2000, 2001, 2008, 2013; Grinblatt and Keloharju, 2000; Feng and Seasholes, 2005).[1] More recently, studies started to investigate the selling behavior of investors in and across asset classes other than single stocks, such as equity mutual funds and index funds (Calvet et al., 2009; Boldin and Cici, 2010; Chang et al., 2016; Bhattacharya et al., 2017). We show in our analysis that portfolio composition affects net fund flows of leading equity market index ETFs. A next

[1] Besides stocks, trading behavior has been examined for executive stock options (Heath et al., 1999), real estate (Genesove and Mayer, 2001), and online betting (Hartzmark and Solomon, 2012).

step could be to investigate whether portfolio composition affects investors' index and mutual fund trading behavior on the individual investor account level.

Third, our findings contribute to experimental and theoretical work on how individuals evaluate risk. Already Slovic and Lichtenstein (1968) and more recently Anzoni and Zeisberger (2016) as well as Holzmeister et al. (2020) find that the probability to experience losses is a much stronger predictor of risk perception than return volatility. Our findings suggest that participants perceive portfolios with more loser stocks to be riskier than portfolios with more winner stocks, although the portfolios have the identical return volatility. Our paper is also related to the work by Barberis and Huang (2001) on risk preference specifications. They test whether narrowly framed (individual-stock level) or broadly framed (portfolio level) fluctuations in the value of an investment, assuming prospect theory preferences, can explain various patterns in the time series and cross section of historical stock returns. Our paper provides experimental evidence of some of the assumptions made in Barberis and Huang's (2001) model. Our findings indicate that portfolio value fluctuations alone cannot explain participants' investment decisions as well as their risk assessments. As proposed by Barberis and Huang (2001) and consistent with our experimental results, a combination of the narrowly framed and the broadly framed risk preference specification is most likely to fit best to how individuals evaluate risk in a portfolio setting.

The remainder of the paper is structured as follows. In Section 3.2, we provide a theoretical framework and experimental evidence of the portfolio composition effect. In Section 3.3, the insights from our experiments are applied to financial market data. In the final section, we discuss the implications of the effect and conclude.

3.2 Experimental Evidence

The evaluation of portfolio investment decisions is complex. Investors are faced with much information and should – if they take normative advice –

solve an optimization problem. The area in psychology referred to as judgment and decision-making has shown that individuals often tend to simplify the world to cope with its complexity. Thereby, one of the strongest tendencies of humans is to classify objects into categories based on some similarity among them (Rosch and Lloyd, 1978). Already in the 1950s, Allport et al. (1954) concludes that "categorical thinking is a natural and inevitable tendency of the human mind" (p. 171). A framework which builds on this finding is mental accounting (Thaler, 1985, 1999; Shefrin and Statman, 1987). It describes the rules individuals engage in when grouping and evaluating outcomes and choices.

A common assumption of mental accounting theories that are applied to portfolio investment decisions is that investors assign each stock to a distinct mental account (i.e. stock-by-stock accounting, see Hartzmark, 2015; Frydman et al., 2017), whereby each mental account defines a separate investment episode (Barberis and Xiong, 2012). Outcomes within one and the same mental account are evaluated jointly, whereas outcomes across different mental accounts are evaluated separately. In particular, this framework implies that individuals are reluctant to integrate gains and losses across different mental accounts, which – applied to a portfolio – suggests that they do not evaluate outcomes across different stocks jointly, but rather distinctly as individual, stock-specific gains and losses.

However, once individuals evaluate a whole portfolio of stocks, information is often presented together, which suggests a joint rather than a separate evaluation. In situations in which information is presented together, research in psychology has shown that individuals focus on differences between the alternatives, when comparing information (Hsee, 1996; List, 2002; Kahneman, 2003). The most salient difference of stocks in a portfolio is probably whether a stock trades at a gain or at a loss. In terms of categorical thinking, this suggests that mental accounts and hence stocks are assigned to one of two distinct categories, namely "winner" stocks or "loser" stocks. Given that the evaluation of outcomes across mental accounts – even across stocks which are all assigned to the same category – requires investors to integrate outcomes which they are reluctant to do and which is difficult, they may rather follow a simple counting heuristic when they evaluate portfolio investment decisions: They may count the number of distinct mental accounts (i.e. stocks) they have assigned to one and the same category rather than aggregating outcomes across different mental accounts within and/or across different categories. As a consequence, investors compare the number of "winner" stocks to the number of "loser" stocks rather than the overall (expected) portfolio return to the overall (expected) portfolio risk.

To deal with large amounts of information of complex decision problems, research in psychology has shown that individuals tend to use simplifying decision procedures such as counting heuristics or so called "tallying strategies" (i.e. equal weighting of cues) (Dawes, 1979; Rieskamp and Hoffrage, 1999). An application of this insight for finance has recently been proposed by Ungeheuer and Weber (2020). They find that individuals understand dependence between assets, but not in terms of correlation. Instead, individuals invest as if they apply a counting heuristic for the frequency of comovement of returns.

To test the predictions of our framework and as such the effect of a portfolio's composition of winner and loser stocks on the portfolio investment choice, we define a simple, counting-based measure of portfolio composition:

$$\frac{Number\ of\ winner\ stocks}{Number\ of\ winner\ stocks + Number\ of\ loser\ stocks} \tag{3.1}$$

A stock is counted as a winner stock, if the stock has a positive realized return since purchase and it is counted as a loser stock, if it has a negative realized return since purchase.[2] Stocks with zero return are not included in the measure. Based on the proposed framework, we make the following predictions:

[2] Later, in the fund flow analysis, we will define winner and loser stocks based on their daily returns.

H1: Holding overall *realized returns* constant across portfolios, participants invest more in the portfolio with the larger portfolio composition measure (i.e. more winner than loser stocks).

H2: Holding overall *realized* and *expected returns* as well as *variance* constant across portfolios, participants invest more in the portfolio with the larger portfolio composition measure (i.e. more winner than loser stocks).

To test the two hypotheses and to analyze whether portfolio composition influences investment decisions, we conduct three investment experiments with in total 1193 participants. In all experiments, participants are asked to allocate an endowment between two portfolios which differ in the proposed composition measure, but are identical with respect to overall realized returns (baseline treatments in experiment 1) and in addition with respect to overall expected return and variance (baseline treatments in experiment 2 and 3). We choose an allocation decision with only two portfolios and no risk-free security to keep the investment task as simple as possible. Both portfolios are similar in the sense that each portfolio consists of ten different, equally weighted stocks and that the underlying portfolio return distributions share similar first and second moments resulting from the used return generating process on the individual stock level, as described later on. In our baseline treatments, portfolios only differ in the composition of winner and loser stocks, but have identical overall realized portfolio returns (and identical overall expected portfolio returns and variance). We run further treatments in which we also differ the realized (and expected) overall portfolio returns across portfolios.

3.2.1 Treatments in the Experiments

The first treatment dimension is the portfolio composition of winner and loser stocks. We analyze two different portfolio compositions which are mirrored images of one another. The "winner" portfolio composition (W_s) consists of seven winner (i.e. positive realized return) and three loser (i.e. negative realized return) stocks. The "loser" portfolio composition (L_s) consists of three winner and seven loser stocks. Importantly, the magnitude of the returns is determined such that the cross-sectional return variance is constant across portfolios. The second treatment dimension of our experimental design is the overall portfolio return. A portfolio can either trade at a gain of +10$ (G_p) or at a loss of -10$ (L_p). We combine the two treatment dimensions to generate different *types of portfolios*. The following four *types of portfolios* result from all possible combinations of our treatment dimensions: G_pW_s, G_pL_s, L_pW_s, and L_pL_s, where the first character denotes the overall portfolio return (marked by the index p for portfolio-level information) and the second character the portfolio composition (marked by the index s for stock-level information).

Since we are interested in within-subject comparisons, i.e. participants' allocation decision of an endowment between two portfolios (i.e. two *types of portfolios*), we combine *two types of portfolio* to *one portfolio pair*. Treatments are then defined by portfolio pairs which in turn are defined by the differences in the respective treatment dimensions. We will first focus on the two portfolio pairs $G_pW_s - G_pL_s$ and $L_pW_s - L_pL_s$. These portfolio pairs define our baseline treatments. They allow us to directly isolate the effect of portfolio composition on investment decisions holding overall realized returns constant across portfolios. In section 1.4, we will discuss the experimental results of further treatments which result from the remaining possible combinations of the two treatment dimensions. Table 3.1 provides an overview of all treatments.

Table 3.1: Overview of Treatments in Experiment 1, 2, and 3

	Treatment dimensions				
Treat-ment	Portfolio return	Portfolio composition	Portfolio return displayed	Portfolio pair 1	Portfolio pair 2
1	same	different	yes	$G_pW_s - G_pL_s$	$L_pW_s - L_pL_s$
2	same	different	no	$G_pW_s - G_pL_s$	$G_pW_s - L_pL_s$
3	different	same	yes	$G_pW_s - L_pW_s$	$G_pW_s - L_pL_s$
4	different	same	no	$G_pW_s - L_pW_s$	$G_pW_s - L_pL_s$
5	different	different	yes	$G_pW_s - L_pL_s$	$G_pW_s - L_pW_s$
6	different	different	no	$G_pW_s - L_pL_s$	$G_pW_s - L_pW_s$

Note: Each experiment has six treatments (except for experiment 3 with only one treatment). Each treatment consists of two portfolio pairs. A portfolio pair consists of two portfolios. Portfolios differ in one or several of three treatments dimensions which are (1) overall portfolio return, (2) portfolio composition and (3) the display format of the portfolio return. Portfolio pairs are described by letter pairs (e.g. $G_pW_S - G_pL_S$). The first letter of each pair corresponds to the overall portfolio return (G_p: Portfolio trades at a gain, L_p: Portfolio trades at a loss) and the second letter corresponds to the portfolio composition (W_S: More winner than loser stocks, L_S: More loser than winner stocks). For example, portfolio pair 1 in treatment 1 is denoted as $G_pW_S - G_pL_S$. The label $G_pW_S - G_pL_S$ means that both portfolios of this pair trade at the same gain denoted by the first letter G_p, but differ in the portfolio composition denoted by the second letter W_S and L_S. All treatments are run in experiment 1 and 2. In experiment 3 only treatment 1 portfolio pair 1 is run.

Figure 3.2 demonstrates how the pairs of portfolios are presented to participants. Exemplary, the portfolio pair $G_pW_s - G_pL_s$ is shown. Both portfolios have the same realized positive return. Portfolio G_pW_s is mainly composed of winner stocks, while portfolio G_pL_s is mainly composed of loser stocks. The amount of information is deliberately reduced to a minimum to ensure a simple design which focuses on the main research question. At the same time, we ensure to provide the set of information investors usually obtain on the overview page of an exemplary online broker account. There are two levels of information. First, investors receive information on the individual stock level. They can see a list of their stock holdings and for each position the return in US dollar over the entire investment horizon. Second, they receive information on the overall portfolio level. They can observe the total return in US dollar of their portfolio which is the sum of the returns of the individual positions. The way we present return information by color coding gains and losses in green and red, respectively, is motivated by how investors usually observe returns in their online broker accounts and on financial websites (see Figure 3.1).

Figure 3.2: Portfolio Pair $G_pW_s - G_pL_s$ in Experiment 1 and 2

Portfolio X	
Stock A	4
Stock B	10
Stock C	-5
Stock D	-7
Stock E	2
Stock F	5
Stock G	2
Stock H	-9
Stock I	5
Stock J	3
Total	10

Portfolio Y	
Stock K	-2
Stock L	-4
Stock M	-2
Stock N	8
Stock O	-5
Stock P	5
Stock Q	-1
Stock R	-2
Stock S	14
Stock T	-1
Total	10

Note: The figure presents the portfolio pair $G_pW_s - G_pL_s$. On the left hand side, portfolio G_pW_s, labeled Portfolio X, and on the right hand side portfolio G_pL_s, labeled Portfolio Y, are demonstrated.

Besides the return information, participants are told about the number of shares held of each stock, the investment horizon and other relevant information in the introduction to the experiment. More details on the instructions can be found in Appendix B.1.

3.2.2 The Return Generating Process

We build on experiment one in experiment two and three by keeping the basic design the same, but we reduce the degrees of freedom participants have when forming beliefs about future returns. More precisely, while we do not tell participants in experiment one the stochastic process of stock returns, we introduce more structure in experiment two and three by telling them the return generating process of individual stocks. From this information and the observed return realizations, participants can infer the expected overall portfolio returns. This extension of the design allows us to keep both, the overall *realized* return as well as the overall *expected* return identical across portfolios. By doing so, we put the effect of portfolio composition on investment decisions to a severe test. In essence, we test whether the effect still exists if participants know that the portfolios will make exactly the same return on expectation.

The return generating process used in experiment two and three is a Bayesian updating task motivated by Grether (1980) and recently adopted by Kuhnen (2015). There are two types of stocks, "good" stocks that draw returns from a good distribution and "bad" stocks that draw returns from a bad distribution. Both distributions are binary and have symmetric stock-specific outcomes ($-X_i$ or X_i). In the good distribution, the probability that stock i increases in value by X_i is 70%, while the probability that it decreases in value by X_i is 30%. In the bad distribution, the probabilities are reversed, i.e. stock i increases in value by X_i with probability of 30%, while it decreases in value by X_i with probability of 70%. The expected return can easily be calculated and is $0.4X_i$ for a good stock and $-0.4X_i$ for a bad stock.

At the beginning of the experiment, participants do not know whether a stock draws from the good or bad distribution, i.e. it is equally likely that a stock draws from either of the two distributions. Over the course of the experiment, participants observe stock return realizations from which they can learn about a stock's type and thus its expected return. From this information and the fact that stocks are equally weighted, they can calculate the expected return of the portfolio from the expected returns of the individual stocks within the portfolio. The computer helps subjects to do the calculations. In particular, subjects are asked to assess a stock's quality and then, based on the assessment, the computer calculates the expected return of the stock. We want to emphasize that while subjects do not need to do the calculations on their own, we explain to them and also test their understanding of how the computer calculates expected returns by the answers they give to comprehension questions at the beginning of the experiment (see Appendix B.1).

Besides expected returns, we design portfolios such that the portfolio return volatility (i.e. the variance of portfolio returns in the time series) is also identical across portfolios. In other words, we ensure that the portfolios in our baseline treatments share identical expected risk-return characteristics measured by an identical expected Sharpe ratio. As a consequence of this design feature, we can also easily mathematically demonstrate how an expected utility maximizing agent with mean-variance preferences should invest given the data generating process and the chosen portfolio options in our experiments. Based on standard portfolio theory (Markowitz, 1952), an agent achieves the largest Sharpe ratio by investing equal amounts in each of the two portfolios in our baseline treatments.[3]

Figure 3.3: Portfolio Pair $G_pW_s - G_pL_s$ in Experiment 3

Portfolio X

Stock		Number of positive return days	Number of negative return days	Total change in value
Stock A	(+/-4)	21	9	48
Stock B	(+/-10)	22	8	140
Stock C	(+/-6)	7	23	-96
Stock D	(+/-7)	8	22	-98
Stock E	(+/-2)	22	8	28
Stock F	(+/-5)	20	10	50
Stock G	(+/-2)	24	6	36
Stock H	(+/-9)	10	20	-90
Stock I	(+/-6)	21	9	72
Stock J	(+/-3)	22	7	42
Total change in portfolio value				132

Portfolio Y

Stock		Number of positive return days	Number of negative return days	Total change in value
Stock K	(+/-2)	3	27	-48
Stock L	(+/-3)	12	18	-18
Stock M	(+/-2)	11	19	-16
Stock N	(+/-8)	24	6	144
Stock O	(+/-5)	8	22	-70
Stock P	(+/-6)	21	9	72
Stock Q	(+/-1)	11	19	-8
Stock R	(+/-2)	11	19	-16
Stock S	(+/-12)	19	11	96
Stock T	(+/-1)	13	17	-4
Total change in portfolio value				132

Note: This figure presets the portfolio pair $G_pW_s - G_pL_s$. On the left hand side, portfolio G_pW_s, labeled Portfolio X, and on the right hand side portfolio G_pL_s, labeled Portfolio Y, are presented. For each stock, the binary outcomes are displayed in parentheses, the number of positive return days, the number of negative return days and the total change in value are shown.

In experiment two and three, portfolio composition is now defined by the number of good (i.e. positive expected return) stocks relative to bad (i.e. negative expected return) stocks. In experiment two, we use the same composition ratios and the same number of return realizations as in experiment one to allow direct comparability between experiments (winner portfolio: seven good/three bad stocks, loser portfolio: three good/seven bad stocks). In experiment three, we also use the same composition ratios, but increase the number of return realizations from one to thirty. This extension of the design ensures that the uncertainty about a stock's type, and consequently the uncertainty about its expected return, is reduced close to zero. This means that participants can be sure about the expected portfolio returns they calculate after observing thirty realizations per stock. Since we provide participants

[3] Appendix Part B.3 provides more details.

in experiment three with more return realizations than in experiment one and two, we have to adjust the way information is presented to participants. Figure 3.3 shows how information is displayed to participants in experiment three. Participants can see for each stock the number of positive return realizations, the number of negative return realizations, and the resulting total change in value of each stock. Summing up these individual changes in value leads to the total change in portfolio value, which is clearly displayed below all portfolio holdings.

3.2.3 Experimental Procedure, Participants, and the Third Treatment Dimension

In all experiments, participants are told to imagine that they have invested $1000 in each of two equally weighted portfolios of stocks one month ago (at t = *-1*) and that they can now (at t = *0*) observe the performance of their investment over the last month. Afterwards, they make an investment decision (i.e. they allocate $1000 between the two portfolios) for another one-month investment period (till t = *1*) at the end of which all returns are realized. They are told that each portfolio consists of ten different stocks and that they invested equal amounts of money in each stock. In experiment one, each participant sees two pairs of portfolios one after the other in randomized order. In experiment two and three, the number of portfolio pairs is reduced to one pair per participant.

In all experiments, there are two periods framed as months: a learning period of one month and an investment period of one month. In experiment one and two, participants are presented one return realization per stock per month. In experiment three, we increase the number of observations from one to thirty per month (similar to daily returns assuming trading on the weekend). The learning period as well as the investment period consist of thirty daily return realizations per stock instead of one monthly return realization. Furthermore, portfolios are re-balanced and the end of the learning period to ensure equal weights of stocks when participants make the investment decision.

At the beginning of all experiments, we explain to participants how the performance of stocks has to be read. We clearly tell them that returns are shown as absolute changes in value of each stock over one month (one day). We repeat this information each time stock returns are displayed to ensure that participants know how to read the returns.

Besides the investment decision, we elicit additional variables. We ask participants to estimate the expected portfolio return and to assess the riskiness of the portfolio. In experiment one, we also ask participants about their satisfaction with the performance of the portfolio and about the confidence in their investment decision.

On each screen, we first show participants the portfolios. Then, we ask one of the above-listed questions including the investment task. Each question was displayed on a separate screen and participants could not return to change a previous answer. At the end of all experiments, participants' demographics, statistics skills, stock market experience and risk aversion were asked.[4]

Besides portfolio composition and overall realized/expected portfolio return, we also investigate whether providing portfolio-level performance information to subjects affects investment choice. In particular, we add a third treatment dimension that is whether overall portfolio returns are explicitly displayed or not. Taken together, this results in four baseline treatments ($G_pW_s - G_pL_s$ with portfolio returns displayed, $G_pW_s - G_pL_s$ without portfolio returns displayed, $L_pW_s - L_pL_s$ with portfolio returns displayed, and $L_pW_s - L_pL_s$ without portfolio returns displayed). We run all of these treatments in experiment one and two. In experiment three, we only run the treatment $G_pW_s - G_pL_s$ with portfolio returns displayed, but conduct two different conditions with respect to whether the expected portfolio variance is displayed in addition to the expected portfolio returns.

[4] Screenshots of the experiments can be seen in Appendix B.2

We conducted all experiments online with 1193 participants from Amazon Mechanical Turk (MTurk). MTurk facilitates access to a wide and diverse pool of participants. Furthermore, there are studies showing that the data obtained via the online platform is at least as reliable as those obtained via traditional methods (Buhrmester et al., 2011). The average time it took participants to finish experiment one (experiment two, experiment three) was 4 minutes and 20 seconds (10 minutes, 18 minutes). 61% (66%, 68%) of the participants were male and the mean age of all participants was 34.7 years (33.9 years, 32.6 years).

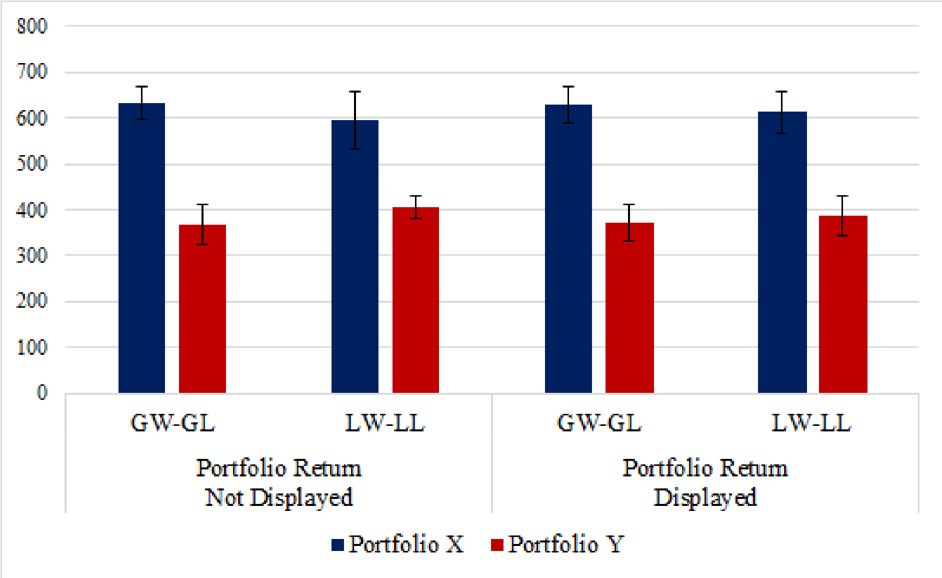
3.2.4 Results Experiment 1: The Portfolio Composition Effect

In experiment one we test the effect of varying portfolio compositions on portfolio investment decisions holding realized portfolio returns constant. In line with Hypothesis H_1, we expect that participants invest more in the portfolio with the more favorable portfolio composition.

Our first main result provides evidence that this is indeed the case. Figure 3.4 shows the average investments in each portfolio for the baseline treatment in which the overall portfolio returns are not explicitly displayed to participants (left part of the figure) and for the baseline treatment in which the overall portfolio returns are explicitly displayed to participants (right part of the figure). We first discuss the results of the treatment in which the overall portfolio return is not explicitly displayed. For those portfolios which have the same realized gain ($G_pW_s - G_pL_s$), participants invest on average \$265 out of \$1000 (t(77)=6.24, $p<0.001$) more in the portfolio with the larger number of winner relative to loser stocks. The difference in average investment is smaller for those portfolios which have the same realized loss ($L_pW_s - L_pL_s$). In the loss case, participants invest on average \$187 out of \$1000 (t(77)=4.22, $p<0.001$) more in the more favorably composed portfolio.

Is the effect simply caused by the fact that it is not obvious to participants that both portfolios have identical realized returns? To test his, we also run

Figure 3.4: Investment in Experiment 1 (Baseline Treatments)

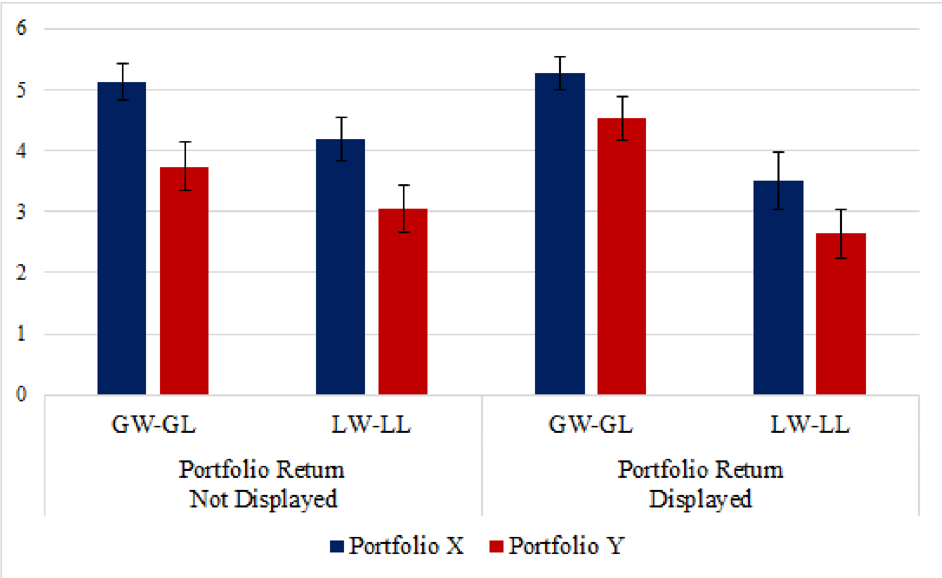


Note: The figure shows participants' mean investments in US dollar in each portfolio for the two portfolio pairs $G_pW_s - G_pL_s$ and $L_pW_s - L_pL_s$. The blue bars refer to Portfolio X which corresponds to the first two letters of each portfolio pair (e.g. G_pW_s for the first portfolio pair) and the red bars refer to Portfolio Y which corresponds to the second two letters of each portfolio pair (e.g. G_pL_s for the first portfolio pair). Displayed are 95% confidence intervals.

the baseline treatment with overall portfolio returns clearly displayed. We still find a strong portfolio composition effect. If both portfolios realized the same gain and this information is clearly displayed, participants invest on average \$258 out of \$1000 (t(78)=6.37, $p<0.001$) more in the portfolio which consists of more winner stocks. If both portfolios realized the same loss and, again, this same loss is clearly displayed, participants invest on average \$224 (t(78)=5.12, $p<0.001$) more in the portfolio which consists of more winner stocks. As such, we can confidentially rule out that the effect depends on whether the overall portfolio return is displayed or not. Even if the identical overall realized return is shown to participants, they are still more willing to invest in the portfolio with the larger number of winner than loser stocks. In addition, the effect size, measured by the magnitude of the differences in investment, is remarkably high in economic terms with 26% in the gain domain and 23% in the loss domain.

Besides the investment, we also elicit participants' satisfaction with the performance of the portfolios, their beliefs about expected portfolio returns and the risk of the portfolios. We find that all of these variables are consistent with participants' investment decisions. Figure 3.5 summarizes participants' average satisfaction levels (measured on a Likert scale from 1: low to 7: high). Irrespective of whether the portfolio return is displayed or not, we find that satisfaction levels are higher for those portfolios which consist of more winner than loser stocks, even though realized portfolio returns are identical.

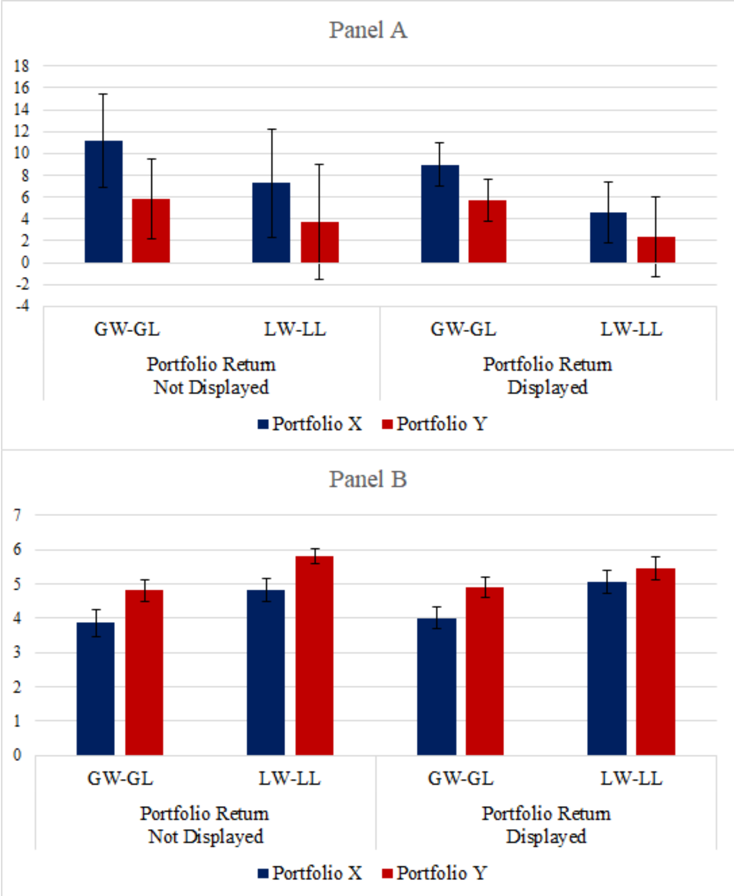
Figure 3.5: Satisfaction in Experiment 1 (Baseline Treatments)



Note: The figure shows participants' mean satisfaction levels for each portfolio elicited on a Likert scale from 1: low to 7: high for the two portfolio pairs $G_pW_s - G_pL_s$ and $L_pW_s - L_pL_s$. The blue bars refer to Portfolio X which corresponds to the first two letters of each portfolio pair (e.g. G_pW_s for the first portfolio pair) and the red bars refer to Portfolio Y which corresponds to the second two letters of each portfolio pair (e.g. G_pL_s for the first portfolio pair). Displayed are 95% confidence intervals

In addition, we find that participants tend to provide more optimistic return expectations as well as lower risk assessments for those portfolios which have more winner than loser stocks. Panel A of Figure 3.6 present participants' average return expectations and Panel B their risk assessment. Taken together, the composition of winner and loser stocks affects individuals' portfolio investment decision. In line with the investment decision, participants also report more optimistic return expectations and lower risk assessments for those portfolios which consist of more winner than loser stocks.

Figure 3.6: Return Expectations and Risk Assessment in Experiment 1 (Baseline Treatments)

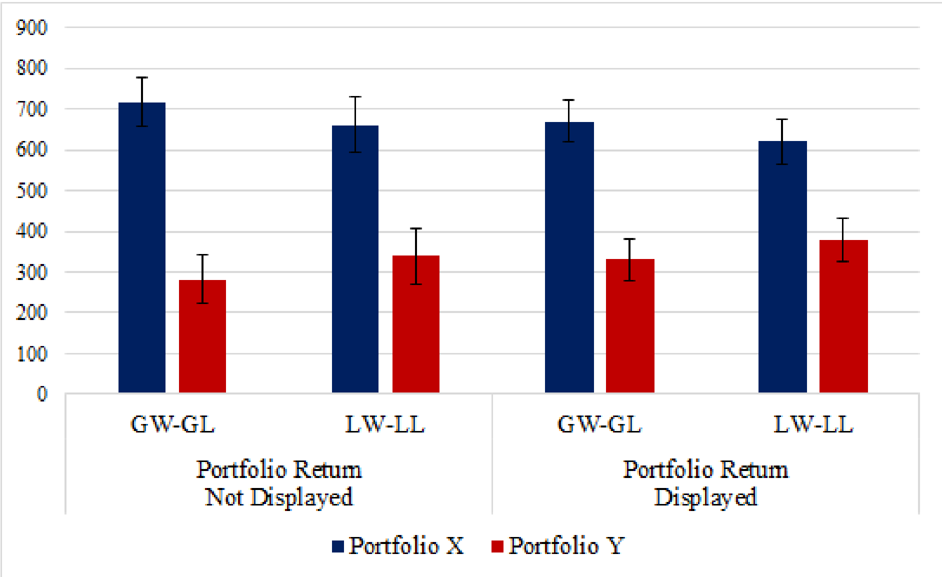


Note: Panel A shows participants' mean expected returns in US dollar and Panel B shows participants' mean risk perception for each portfolio elicited on a Likert scale from 1: low to 7: high for the two portfolio pairs $G_pW_S - G_pL_S$ and $L_pW_S - L_pL_S$. The blue bars refer to Portfolio X which corresponds to the first two letters of each portfolio pair (e.g. G_pW_s for the first portfolio pair) and the red bars refer to Portfolio Y which corresponds to the second two letters of each portfolio pair (e.g. G_pL_s for the first portfolio pair). Displayed are 95% confidence intervals.

3.2.5 Results Experiment 2: Learning About Expected Returns I

In a simple investment task, we have shown that a portfolio's composition of winner and loser stocks affects investors' willingness to invest in a portfolio. In experiment two, we aim to replicate the main finding from experiment one, but under the important modification that we keep not only the overall realized returns constant across portfolios, but also the overall expected returns (and variance). In line with Hypothesis H_2, we expect that participants invest more in the portfolio with a more favorable portfolio composition.

Figure 3.7: Investment in Experiment 2 (Baseline Treatments)

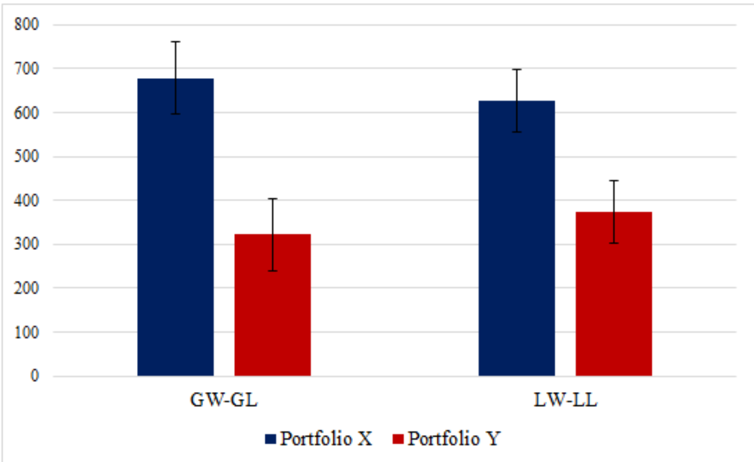


Note: The figure shows participants' mean investments in US dollar in each portfolio for the two portfolio pairs $G_pW_S - G_pL_S$ and $L_pW_S - L_pL_S$. The blue bars refer to Portfolio X which corresponds to the first two letters of each portfolio pair (e.g. G_pW_S for the first portfolio pair) and the red bars refer to Portfolio Y which corresponds to the second two letters of each portfolio pair (e.g. G_pL_S for the first portfolio pair). Displayed are 95% confidence intervals.

We replicate the findings from experiment one. Figure 3.7 depicts participants' average investments in each portfolio. Again, the left part of Figure 3.7 presents average investments when total portfolio returns are not displayed and the right part when total portfolio returns are displayed. Irrespective of whether the portfolio returns are displayed or not, we find that participants invest significantly more in the portfolio which consists of more winner than loser stocks. For the treatment in which both portfolios have the same realized and expected positive return, participants invest on average \$339 (\$436 if portfolio returns are not displayed) more in the portfolio with the larger number of winner to loser stocks (t(50)=6.62, $p<0.001$; t(49)=7.25, $p<0.001$). For the treatment in which both portfolios have the same realized and expected negative return, participants invest on average \$240 (\$322 if portfolio returns are not displayed) more in the portfolio with the larger number of winner to loser stocks (t(54)=4.46, $p<0.001$; t(40)=4.74, $p<0.001$). In the entire sample, unconditional of subjects' self-reported beliefs about expected returns, we find as in experiment one a portfolio composition effect.

To test whether the effect still persists if subject's self-reported beliefs about expected returns are identical, we rerun the analysis on the subsample of subjects who report – as Bayes' rule implies – the same expected returns for both portfolios. Even though, the sample size decreases quite significantly with this restriction, we find for those participants who report exactly the same beliefs about expected portfolio returns, a portfolio composition effect. Figure 3.8 reports the average investment for this subsample.

Figure 3.8: Investment in Experiment 2 Conditional on Return Expectations (Baseline Treatments)



Note: The figure shows participants' mean investments in US dollar in each portfolio of those participants who state the same expected returns for the two portfolios of a pair. The portfolio pairs are $G_pW_S - G_pL_S$ and $L_pW_S - L_pL_S$. The blue bars refer to Portfolio X which corresponds to the first two letters of each portfolio pair (e.g. G_pW_S for the first portfolio pair) and the red bars refer to Portfolio Y which corresponds to the second two letters of each portfolio pair (e.g. G_pL_S for the first portfolio pair). Displayed are 95% confidence intervals.

For the treatment in which portfolios have the same positive realized return and participants report the same beliefs about expected portfolio returns, we find that participants invest \$356 more in the portfolio with more winner than loser stocks (t(35)=4.38, $p<0.001$). For the treatment in which portfolios have the same negative realized return and participants report the same beliefs about expected portfolio returns, we find that participants invest $254 more in the portfolio with more winner than loser stocks than in the alternative portfolio with more loser than winner stocks (t(34)=3.60, p=0.001).

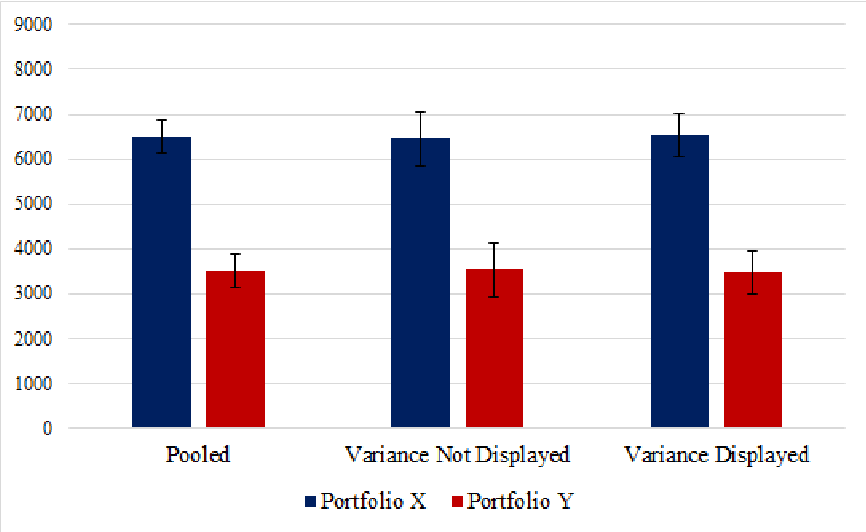
Besides the investment decisions and the return expectations, we also elicit participants' perception of risk. The results are reported in Appendix B Figure B.13. Participants rate those portfolios which consist of more loser than winner stocks to be riskier than those portfolios which consist of more winner than loser stocks. The results are in line with the investment decisions and replicate findings from experiment one.

3.2.6 Results Experiment 3: Learning About Expected Returns II

In experiment three, we further test the robustness of our findings from our two previous experiments. We build on experiment two (same realized and expected portfolio returns), but extend the learning phase such that participants can learn from a larger number of return realizations before they make their investment decision. In addition to that, we explicitly display to one group of participants not only the calculated expected returns, but also the *expected portfolio return volatility*. This modification allows us to test whether the documented portfolio composition effect still exists if subjects' beliefs about expected portfolio returns as well as their beliefs about expected portfolio return volatility are identical across portfolios.

Figure 3.9 reports the average investment in each portfolio pooled and split by condition (portfolio volatility not displayed and displayed) unconditional of participants' beliefs. Figure 3.10 displays the results for those participants who report beliefs about expected returns that are in line with Bayes. In the entire sample, unconditional of participants' beliefs about expected returns, we find a strong portfolio composition effect. Participants invest on average $2994 (out of $10000) more in the portfolio which consists of more winner than loser stocks (t(101)=7.86, p<0.001). This finding is independent

Figure 3.9: Investment in Experiment 3

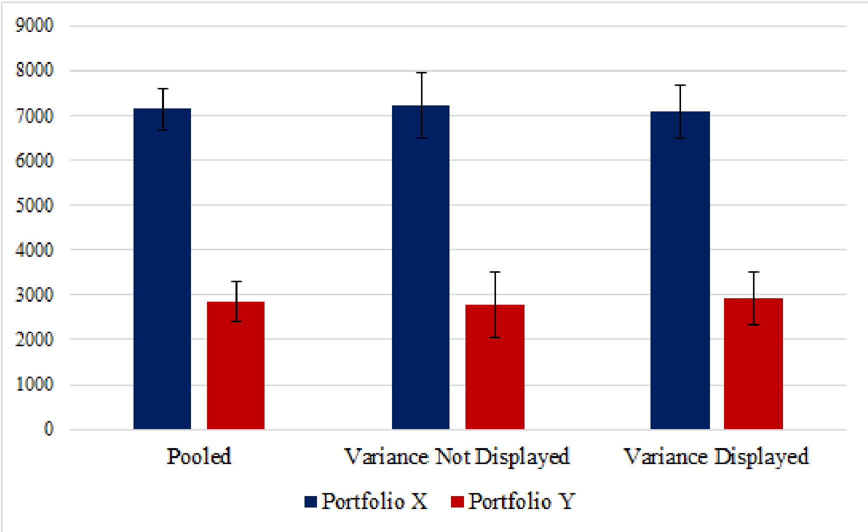


Note: The figure shows participants' mean investments in US dollar in each portfolio for the portfolio pair $G_pW_S - G_pL_S$. The blue bar refers to Portfolio X which corresponds to the first two letters of the portfolio pair (e.g. G_pW_S for the first portfolio pair) and the red bar refers to Portfolio Y which corresponds to the second two letters of the portfolio pair (e.g. G_pL_S for the first portfolio pair). Displayed are 95% confidence intervals.

of whether the portfolio variance is displayed or not. If we restrict the sample to those participants who report beliefs about expected returns which are in line with Bayes, the effect persists and gets even stronger. Participants in this subsample invest on average \$4295 (out of \$10000) more in the portfolio with more winner than loser stocks than in an alternative portfolio with more loser than winner stocks (t(58)=9.49, $p<0.001$). Again, and interestingly, this finding is unaffected by whether the identical portfolio variance is displayed to subjects or not. In other words, the portfolio composition effect persists even in situations in which we can confidentially rule out that differences in participants' beliefs about expected portfolio returns and expected portfolio return volatility can drive the observed differences in investments.

In addition to the investment choice and the beliefs about expected returns, we also ask participants about a risk assessment for the portfolios. Figure B.14 in Appendix B displays the average risk assessments unconditional of subjects' expected portfolio returns and Figure B.15 in Appendix B for those subjects who report identical beliefs about expected portfolio returns. Consistent with results from previous experiments, we find that

Figure 3.10: Investment in Experiment 3 Conditional on Expected Returns



Note: The figure shows participants' mean investments in US dollar in each portfolio of those participants who state the objective expected returns for the two portfolios of a pair. The portfolio pair is $G_pW_S - G_pL_S$. The blue bar refers to Portfolio X which corresponds to the first two letters of the portfolio pair (e.g. G_pW_S for the first portfolio pair) and the red bar refers to Portfolio Y which corresponds to the second two letters of the portfolio pair (e.g. G_pL_S for the first portfolio pair). Displayed are 95%confidence intervals.

participants evaluate the portfolio with more winner than loser stocks to be less risky by 1.77 scores than the portfolio with more loser than winner stocks (t(102)=9.55, p<0.001). If we restrict the sample to those subjects who report identical beliefs about expected returns and volatility, we interestingly still find that subjects evaluate the portfolio with more winner than loser stocks to be less risky than the portfolio with more loser than winner stocks. The difference in risk evaluation is even larger with 2.59 scores for this group of participants (t(58)=11.77, p<0.001).

3.2.7 Further Experimental Results

Besides our baseline treatments ($G_pW_s - G_pL_s$ and $L_pW_s - L_pL_s$), we run four additional treatments in experiment one and two (see Table 3.1). The additional treatments allow us to further test the robustness of the portfolio composition effect. In the first two additional treatments, we hold the portfolio composition identical across portfolios and differ the overall realized (expected) portfolio return ($G_pW_s - L_pW_s$ and $G_pL_s - L_pL_s$). In the second two additional treatments, we differ both, the portfolio composition and the overall realized portfolio return across portfolios ($G_pW_s - L_pL_s$ and $G_pL_s - L_pW_s$).

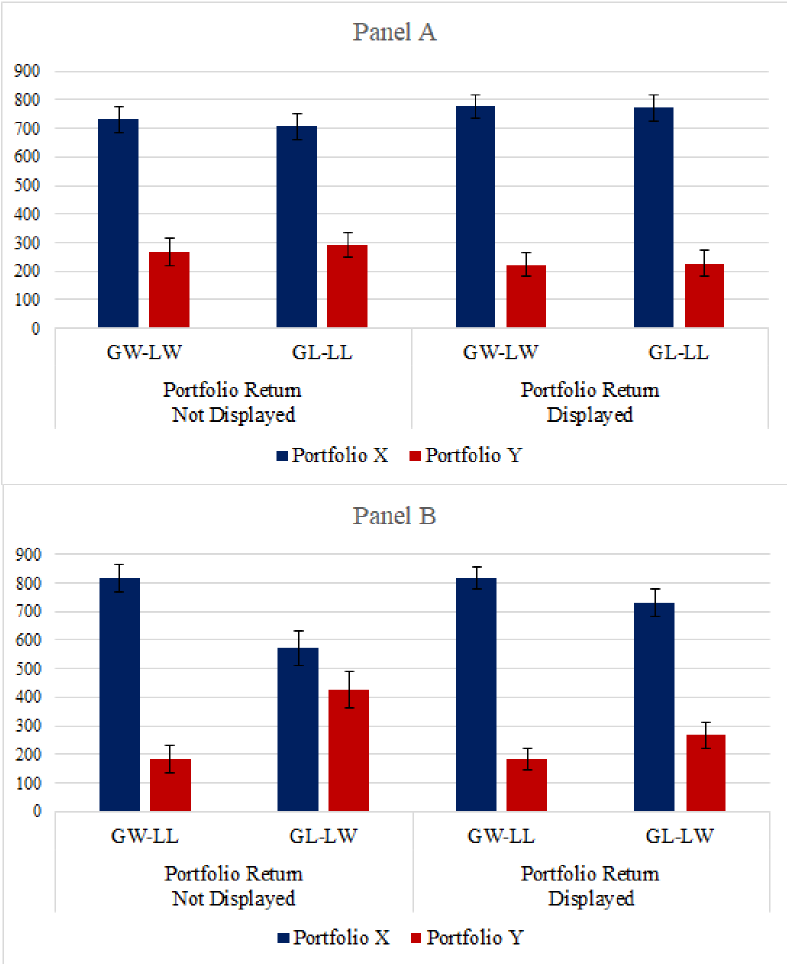
These additional treatments allow us to make several within- and between-treatment comparisons to better understand how portfolio investment decisions are affected by portfolio compositions and overall portfolio returns: (1) How strong is the effect of differences in portfolio composition on investment decisions as compared to the effect of differences in overall realized and expected portfolio returns? (2) How does "consistent" performance information (the portfolio with a positive realized return consists mainly of winner stocks and the portfolio with a negative realized return consists mainly of loser stocks) affect investment decisions as compared to "inconsistent" performance information (the portfolio with a positive realized return consists mainly of loser stocks, while the portfolio with a negative realized return consists mainly of winner stocks)?

Figure 3.11 displays the mean investment in each portfolio for the four additional treatments in experiment one.[5] First, we find for those treatments in experiment one in which we keep the portfolio composition constant and differ the overall realized portfolio return (i.e. $G_pW_s - L_pW_s$ and $G_pL_s - L_pL_s$) that participants invest on average \$550 more in the portfolio with a positive realized return than in the alternative portfolio with a negative realized return. This difference is unaffected by whether both portfolios consist mainly of winner stocks or whether both portfolios consist mainly of loser stocks ($\Delta_{G_pW_s - L_pW_s}$ versus $\Delta_{G_pL_s - L_pL_s}$).

Second, we provide further evidence of the portfolio composition effect by comparing differences between the portfolio pairs $G_pW_s - L_pL_s$ and $G_pL_s - L_pW_s$. Across both portfolio pairs, we keep the difference in overall realized (and expected) portfolio returns constant, but flip the portfolio composition (i.e. the portfolio with a positive realized and expected return consisting mainly of winners/losers is changed to the portfolio with a negative realized and expected return consisting mainly of losers/winners). As

[5] The results of experiment two are similar and reported in Appendix B Figure B.19.

Figure 3.11: Investment in Experiment 1 (Additional Treatments)



Note: The figure shows participants' mean investments in US dollar in each portfolio for the four portfolio pairs $G_pW_S - L_pW_S$ and $G_pL_S - L_pL_S$ (Panel A) and $G_pW_S - L_pL_S$ and $G_pL_S - L_pW_S$ (Panel B). The blue bars refer to Portfolio X which corresponds to the first two letters of each portfolio pair (e.g. G_pW_S for the first portfolio pair) and the red bars refer to Portfolio Y which corresponds to the second two letters of each portfolio pair (e.g. L_pW_S for the first portfolio pair). Displayed are 95%confidence intervals.

an alternative test of the portfolio composition effect, we compare differences in investment across two portfolio pairs instead of the investment between two types of portfolios within one portfolio pair (see baseline treatments). If portfolio composition does not matter for investment decisions, we expect to observe no significant difference between participants' mean investment decisions across the portfolio pairs ($\Delta_{G_pW_s - L_pL_s} = \Delta_{G_pL_s - L_pW_s}$). However, we find significant differences in investment decisions. In particular, participants in experiment one invest on average \$633 more in portfolio G_pW_s than in portfolio L_pL_s. This difference in investment reduces significantly by \$171 (t(83)=3.60, p<0.001) to \$462 for the portfolio pair $G_pL_s - L_pW_s$. The resulting difference of the differences in investment provides further evidence of a portfolio composition effect.

Table 3.2: Investment Behavior Across All Treatments

Dependent Variable	*Investment*			
	Experiment 1		**Experiment 2**	
	(1)	(2)	(3)	(4)
Gain	311.2***	260.6***	264.9***	220.0***
	(17.61)	(24.99)	(21.52)	30.61
Winner	116.5***	131.9***	147.3***	151.5***
	(16.17)	(23.60)	(20.49)	30.41
Gain x Winner	28.08	41.73	36.17	62.12
	(21.82)	(31.72)	(27.57)	39.43
Display		−28.43		−22.72
		(19.82)		(23.36)
Display x Gain		101.2***		88.89**
		(35.00)		(42.70)
Display x Winner		−30.75		−8.746
		(32.20)		(40.94)
Display x Gain x Winner		−27.30		−54.89
		(43.58)		(54.90)
Constant	279.1***	293.3***	284.8***	296.4***
	(9.919)	(14.50)	(11.67)	(16.93)
Observations	1,936	1,936	1,213	1,213
R^2	0.346	0.353	0.323	0.327

Note: The table shows the coefficients of OLS regressions of investment on a gain dummy variable (1 if portfolio trades at a gain), a winner dummy variable (1 if portfolio has more winner than loser assets), the interaction term of gain and winner, a display dummy variable (1 if total portfolio return is displayed) and multiple interaction terms of the display, gain and winner dummy variable. Regression (1) and (2) are run with data from experiment 1, regression (3) and (4) are run with data from experiment 2. We cluster standard errors on the individual investor level and on the portfolio pair level, standard errors are reported in parentheses, and *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

We also run multivariate ordinary least squared regressions to test for a portfolio composition effect across all treatments of experiment one and experiment two. The dependent variable $Investment_{ij}$ is the invested amount of subject i in portfolio j, $Gain_j$ is a dummy variable which is one if portfolio j made a gain, $Winner_j$ is a dummy variable which is one if portfolio j has more winner than loser stocks and $Display_j$ is a dummy variable which is one if the overall portfolio return is displayed. We use robust standard errors and cluster on the subject and the portfolio pair level. Table 3.2 reports the results for each experiment individually. In both experiments, we find a strong portfolio composition effect. Subjects invest on average $116.50 ($147.30) more in the portfolio which is mainly composed of winner stocks than in the portfolio which is mainly composed of loser stocks. The effect is slightly stronger if the total return is not displayed, although not statistically different.

Like in the baseline treatments, we find that participants' self-elicited level of satisfaction with the performance of the portfolios, their beliefs about expected returns and risk assessment are in line with the observed investment decisions. Appendix B.4 displays the results.

3.3 From the Experiment to Financial Market Data

In a series of experiments, we have identified that participants make portfolio investment decisions as if they evaluate portfolios based on a simple counting heuristic of their compositions of winner and loser stocks. In what follows, we take this finding outside the laboratory environment and test whether portfolio composition also plays a role in how portfolio-like securities are bought and sold in financial markets. More precisely, we investigate whether the demand for leading equity market index funds is influenced by the proposed composition measure.

Leading equity market indices of national economies represent ideal portfolio settings for our analysis. First, leading equity market indices are relatively stable and transparent predetermined portfolios with respect to the members of the index over time. There are clear rules when a stock leaves or enters a national equity market index and these changes of the members are communicated. Second, leading equity market indices capture a lot of attention in the daily media as well as press of a national economy since they are often referred to as indicators of a country's overall economic condition. Moreover, various publicly available financial websites as well as news channels on television report not only the overall performance of a national equity market index, but also the performance of its individual stocks (see websites such as finanzen.net and onvista.com or news channels such as n-tv and CNN). The information needed to calculate our portfolio composition measure are thus easily assessable and even prominently placed for retail as well as professional investors.

As a measure for investor demand, we use fund flows of exchange-traded funds replicating the respective equity market index. The exchange-traded fund industry has grown tremendously over the past decade and exchange-traded funds (ETFs) have become a popular financial security to invest in usually broad indices at relatively low costs. Moreover, ETFs have distinct advantages compared to usual index mutual funds. ETFs are traded on an intraday basis with a continuously observable price, while mutual funds can only be traded once a day at their NAV. However, this also comes with a key difference between the ETF's and mutual fund's investment mechanism. While for ordinary mutual funds, investors can directly buy shares at the end-of-trading-day NAV (i.e. they exchange cash for shares), for ETFs, the authorized participants (AP) and not the individual investors directly deal with the ETF (i.e. the AP buys a portfolio of the ETF's underlying stocks and exchanges it for shares of the ETF). Although, this mechanism effectively separates investors from ETFs, fund flows of ETFs can still be interpreted as net investor demand for an ETF given that the AP usually creates ETF shares if demand exceeds supply and redeems ETF shares otherwise (Clifford et al., 2014).

Building on Hypothesis H_1 and the experimental findings, we expect that portfolio composition affects net fund flows of exchange-traded funds replicating leading equity market indices. Like in our experiments, the positive relationship we expect between our portfolio composition measure and future net fund flows should even hold after controlling for the fund's return. In other words, we test whether future net fund flows are affected by the composition of winner and loser stocks of an equity market index in addition to the index return.

H3: A more favorable portfolio composition (i.e. more winner relative to loser stocks) leads to larger future net fund flows. This relation should hold even after controlling for a fund's overall return.

There is a large body of literature on the relation between fund flows and fund returns. Several studies find return chasing behavior of actively-managed mutual fund investors indicated by the positive relation between future net flows of mutual funds and their returns (Ippolito, 1992; Gruber, 1996; Warther, 1995; Sirri and Tufano, 1998; Edelen and Warner, 2001; Coval and Stafford, 2007; Ben-Rephael et al., 2011). Besides actively-managed mutual funds, return-chasing behavior has even been observed for index mutual funds (Elton et al., 2004). For ETFs, the return-flow relation has received much less attention in the literature so far and from those studies which exist, there is less clear-cut evidence of whether ETF flows are influenced by returns. Clifford et al. (2014) use monthly data to test drivers of ETF flows and find return-chasing behavior by investors, while Kalaycıoğlu (2004) does not find return-chasing behavior for ETFs at the daily level. Our paper contributes to the relatively unexplored literature of ETF investor return-chasing behavior.

3.3.1 Data

We test our hypothesis using fund flow data of leading equity market index ETFs for the period 2016-2019. Our sample consists of twelve leading equity market indices. Table 3.3 summarizes all market indices in our sample.

Our sample comprises ten European equity market indices as well as two North-American equity market indices. For each national economy in our sample, we chose the leading equity market index of the respective country (e.g. the CAC 40 for France, the IBEX 35 for Spain, the DAX 30 for Germany) and then search for ETFs replicating the index. Importantly, ETFs only enter the sample if their investment objective is to replicate the index as closely as possible. We exclude all index ETFs which use hedging strategies or claim in

Table 3.3: Summary of Equity Market Indices and ETFs in the Sample

Market Index	Country	Number of stocks	Number of ETFs
ATX	Austria	20	3
BEL 20	Belgium	20	1
CAC 40	France	40	5
DAX 30	Germany	30	10
Dow Jones	US	30	4
Euro STOXX 50	Eurozone	50	20
FTSE 100	Great Britain	100	8
FTSE MIB 40	Italy	40	4
IBEX 35	Spain	35	1
PSI 20	Portugal	18	1
S&P/TSX 60	Canada	60	2
SMI	Switzerland	20	2

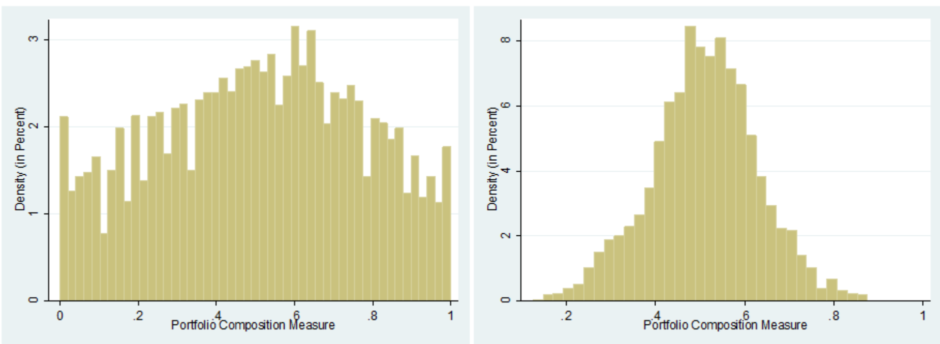
Note: The table lists the leading equity market indices of various European and North-American countries, the number of stocks of the index and the number of ETFs in our sample replicating the respective index.

their investment objective that they use other strategies to systematically deviate from the index (e.g. minimum variance, excluding financial industry). We verify the investment objective of all index ETFs in our sample by hand on the ETF provider's website. As seen in Table 3.3, the number of ETFs per index varies from ten for the DAX 30 to one for the IBEX 35 which depends on the availability of fund flow data from Morningstar. We use daily data for our analyses and test for a direct relation between our portfolio composition measure on a respective day and the index ETFs' net fund flow on the next day. In addition to our analyses with daily data, we also run our regression models with weekly data. For these analyses, we define a weekly composition measure which is the average of the daily portfolio compositions within a week.

We obtain fund-level data from Morningstar. For each ETF (identified by its SecId and FundId), we download the ETF's daily net asset value (NAV), index return, number of shares outstanding, total net assets (TNA) and the total expense ratio. We calculate net fund flows following Morningstar and common in the literature as difference between two consecutive day TNAs (calculated as number of shares outstanding times NAV) adjusted for the respective day's index return.[6] For the calculation of our portfolio composition measure, we download stock return data from Thomson Reuters Datastream. Each day, we define each stock as either a winner stock (positive daily return) or a loser stock (negative daily return). Stocks with zero daily return do not enter the composition measure on that day. Indices change their stocks from time to time. To account for these changes, we hand collect from Bloomberg the days on which an index in our sample experiences a change in its stocks and identify which stock leaves and which enters the index. Based on the stock return data and the changes of the members of an index, we calculate our portfolio composition measure as defined in section 3.2.

Before we turn to the main analysis, we provide summary statistics for our measure of portfolio composition. All summary statistics are calculated based on daily data as well as weekly data. Figure 3.12 displays the distribution of the portfolio composition measure for all market indices pooled. The distribution of our composition measure of winner and loser stocks is relatively normally distributed for all equity market indices pooled as well as individually.

Figure 3.12: Distribution of the Portfolio Composition Measure

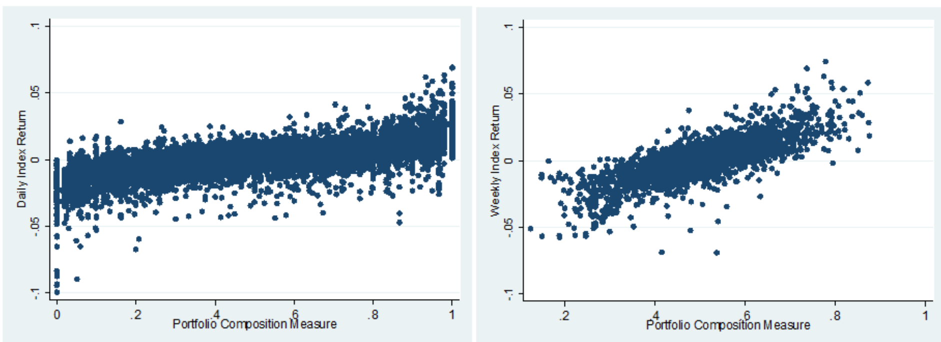


Note: The figure shows the distribution of our portfolio composition measure for the sample of twelve leading equity market indices. We show daily data in the left part of the figure and weekly data, i.e. the average portfolio composition over all trading days within a week, in the right part of the figure.

[6] Net cash flow on day t = (Shares on day t * NAV on day t) – (Shares on day $t-1$ * NAV on day $t-1$) * (1+ return on day t), see estimated net cash flow methodology by Morningstar.

Next, we take a look at how our measure of portfolio composition is related to index returns. Figure 3.13 illustrates the relation between our portfolio composition measure and the index return in a dot plot. As expected, there is a positive relation between the index return on a given day and its portfolio composition. A more favorable portfolio composition is related to a larger index return. However, and crucial for our study, there is a considerable variability in the portfolio composition for a given (fixed) index return. That means an index return can be achieved with different portfolio compositions. The left part of Figure 3.13 shows that a daily index return of 1.00% can be achieved by mainly winner stocks (i.e. a composition of more than 90% winner stocks) or by more loser than winner stocks (i.e. a composition of 80% loser stocks). Using weekly data as in the right part of Figure 3.13, there is still a considerable variability in the portfolio composition. For example, a 1.00% weekly index return can be achieved by more than 80% winner stocks or by up to 80% loser stocks. As a side remark, the portfolio compositions used in our experiments (e.g. 70% winners versus 30% winners for a portfolio return of 1.00%) are comparable to the empirically observed portfolio compositions in our sample.

Figure 3.13: Relation Between Portfolio Composition and Index Return



Note: The figure shows the relation between our portfolio composition measure and the index return for the sample of twelve leading equity market indices. We show daily data in the left part of the figure and weekly data, i.e. the average portfolio composition over all trading days within a week, in the right part of the figure.

3.3.2 Main Result

Our unique dataset of fund-level as well as stock-level data allows us to test our hypothesis. We run the following regression model (similar to Clifford et al., 2014 and Staer, 2017):

$$Flow_{i,j,t} = \beta_0 + \sum_{l=1}^{3} \beta_{R,l} FundReturn_{i,j,t-l} + \sum_{l=1}^{3} \beta_{C,l} CompositionMeasure_{i,j,t-l} + \epsilon_{i,j,t} \quad (3.2)$$

In the panel regression, the dependent variable $Flow_{i,j,t}$ represents the net fund flow of ETF i on index j on day t, $FundReturn_{i,j,t-l}$ represents the fund (index) return of ETF i on index j on day $t-l$, where l represents the number of lags, and $CompositionMeasure_{i,j,t-l}$ represents the value of the composition measure of ETF i on index j on day $t-l$, where l represents the number of lags. The panel model includes fund and day fixed effects. We cluster residuals by index and use robust standard errors. The results are summarized in Table 3.4.

For our sample of leading equity market indices, we find a positive relation between the portfolio composition and net fund flows. In particular, we find that today's net fund flows of an equity market index ETF are affected by yesterday's composition of winner and loser stocks of the index. Across all leading equity market indices in our sample, we estimate that a portfolio composition of 100% winner stocks leads on average to 1,119,000 US dollar higher inflows on the subsequent day than a portfolio composition of 50% winner and 50% loser stocks. In relative terms, this inflow presents roughly 19% of the average daily fund inflow of an ETF in our sample. The effect remains statistically significant and decreases only slightly in magnitude when controlling for the index return (column 2). We estimate that a portfolio composition of 100% winner stocks leads on average to 808,000 US dollar higher inflows on the subsequent day than a portfolio composition of 50% winner and 50% loser stocks. In relative terms, this inflow presents roughly 14% of the average daily fund inflow of an ETF in our sample. The results change only marginally if we include the portfolio composition and the index return of the day of the observed net fund flow to the regression model (columns

Table 3.4: Portfolio Composition and ETF Fund Flows – Daily Data

Dep. Variable	$NetFlow_t$			
	(1)	(2)	(3)	(4)
$Composition_t$			24240.1	374159.1
			(0.03)	(1.16)
$Composition_{t-1}$	2238859.2**	1616734.1**	2274420.5**	1648174.4**
	(2.66)	(2.23)	(2.72)	(2.29)
$Composition_{t-2}$	2750528.8	2294166.7	2799108.8	2329743.1
	(1.23)	(1.04)	(1.24)	(1.05)
$Composition_{t-3}$	3289634.8	2998568.9	3261656.1	2980088.7
	(1.46)	(1.34)	(1.45)	(1.34)
$FundReturn_t$				−20547840.6
				(−0.61)
$FundReturn_{t-1}$		29472179.8		29466988.4
		(1.21)		(1.24)
$FundReturn_{t-2}$		33695795.5		33953226.3
		(1.36)		(1.38)
$FundReturn_{t-3}$		29043137.2		28965826.4
		(1.26)		(1.25)
$Constant$	−825335.9	318431.2	−859524.6	78955.2
	(−0.67)	(0.18)	(−0.70)	(0.05)
Observations	68332	65327	68223	65207
R^2	0.018	0.019	0.018	0.019
Fund FE	YES	YES	YES	YES
Time FE	YES	YES	YES	YES

Note: The table summarizes results of panel regressions of the dependent variable Net Flow on day t on a Composition variable on day t and up to three days lagged and a Fund Return variable on day t and up to three days lagged. T-statistics are reported in parentheses, and *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

3 and 4). Moreover, we find a tendency of return-chasing behavior for ETF investors which in line with Clifford et al. (2014) and with several studies on mutual fund flow data (Ippolito, 1992; Gruber, 1996; Warther, 1995; Sirri and Tufano, 1998; Edelen and Warner, 2001). Compared to the effect of portfolio composition on future net fund flows, the effect of past index returns on future fund flows is economically considerably larger. Overall, our findings confirm Hypothesis H_3.

3.3.3 Robustness Analyses

We run several robustness analyses in this section. Can the effect be observed in weekly data, too? Does the effect exist for both, European and North-American equity market indices? What drives the effect? How is the effect related to comparable measures such as the return dispersion of a portfolio (i.e. the cross-sectional variance of stock returns within an index)?

First, we replicate the main finding using weekly instead of daily data. We calculate the weekly portfolio composition measure as the arithmetic mean of all daily portfolio compositions over a week. Table 3.5 reports the results. We find two main results: The portfolio composition of week t is positively related to the net fund flows of week t. In numbers, a weekly portfolio composition of 75% winner and 25% loser stocks leads on average to a 5,568,000 US dollar higher inflow in this week than a portfolio composition of 50% winner and 50% loser stocks. In relative terms, this inflow presents roughly 30% of the average weekly fund inflow of an ETF in our sample. The effect remains statistically significant and decreases only marginally in size when controlling for the index return. Interestingly, the previous week's portfolio composition has no significant effect on this week's net fund flows. This result hints that the effect is short-living. People may rather remember and act upon the observation that the majority of stocks of an index achieved a positive daily return yesterday and potentially also two days ago, but may have problems to remember and do not act anymore upon the same observation one week ago. Additionally, we find return-chasing behavior of ETF investors when using weekly data. Net fund flows tend to be larger in a given week if the index return of the previous week was larger. This finding is consistent with the above-cited literature on return-chasing behavior of ETF as well as mutual fund investors.

Table 3.5: Portfolio Composition and ETF Fund Flows – Weekly Data

Dep. Variable	$NetFlow_t$			
	(1)	(2)	(3)	(4)
$Composition_t$	24177218.8***	22592605.2***	24775876.3***	22273359.8***
	(3.60)	(4.35)	(3.55)	(4.00)
$Composition_{t-1}$			10205171.3	1359932.4
			(0.51)	(0.07)
$FundReturn_t$		29344623.9		43014072.6
		(0.65)		(0.89)
$FundReturn_{t-1}$				166208228.2**
				(3.08)
$Constant$	−10229770.8	−9517556.7	−9356166.9	−3738440.2
	(−1.09)	(−1.14)	(−1.29)	(−0.70)
Observations	17340	17340	17255	17255
R^2	0.017	0.017	0.017	0.018
Fund FE	YES	YES	YES	YES
Time FE	YES	YES	YES	YES

Note: The table summarizes results of panel regressions of the dependent variable Net Flow in week t on a Composition variable in week t and one week lagged and a Fund Return variable in week t and one week lagged. T-statistics are reported in parentheses, and *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

Second, we examine whether there are regional differences of the effect. We mean by regional differences whether the effect exists for European equity market indices as well as for North-American equity market indices. To test this, we split our sample into a European sample consisting of ten European national equity market indices and a North-American sample consisting of two North-American national equity market indices. We run the same panel regression model as run on the pooled sample, but now for each sample individually. The results are reported in Table 3.6 for the European market indices and in Table 3.7 for the North-American market indices.

Table 3.6: Portfolio Composition and ETF Fund Flows for European Market Indices

Dep. Variable	$NetFlow_t$			
	(1)	(2)	(3)	(4)
$Composition_t$			645485.5**	577509.6**
			(3.41)	(2.50)
$Composition_{t-1}$	1153508.8***	1068401.1***	1186908.3***	1109099.2***
	(5.47)	(5.71)	(5.61)	(5.85)
$Composition_{t-2}$	393176.9	318625.1	396590.9	324963.3
	(1.11)	(0.74)	(1.11)	(0.75)
$Composition_{t-3}$	975472.4**	847429.0*	988932.4**	861908.8*
	(2.86)	(2.11)	(2.86)	(2.14)
$FundReturn_t$				2183637.4
				(0.60)
$FundReturn_{t-1}$		1313132.7		1208931.6
		(0.32)		(0.29)
$FundReturn_{t-2}$		5648479.1		5756267.6
		(0.85)		(0.87)
$FundReturn_{t-3}$		5996211.0		5948771.6
		(0.95)		(0.94)
$Constant$	698347.8	576787.3	670390.8	576752.9
	(1.39)	(1.27)	(1.32)	(1.29)
Observations	60045	57344	59964	57253
R^2	0.024	0.025	0.024	0.025
Fund FE	YES	YES	YES	YES
Time FE	YES	YES	YES	YES

Note: The table summarizes results of panel regressions of the dependent variable Net Flow on day t on a Composition variable on day t and up to three days lagged and a Fund Return variable on day t and up to three days lagged. The sample is restricted to European market indices. T-statistics are reported in parentheses, and *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

Table 3.7: Portfolio Composition and ETF Fund Flows for North-American Market Indices

Dep. Variable	$NetFlow_t$			
	(1)	(2)	(3)	(4)
$Composition_t$			−12847118.9**	−6843806.4
			(−50.66)	(−0.76)
$Composition_{t-1}$	5414442.7	502328.0	4435847.1	1224740.8
	(1.55)	(0.05)	(1.44)	(0.17)
$Composition_{t-2}$	16640010.8	7850856.6	16932770.4	9107176.2
	(5.88)	(3.45)	(6.15)	(2.95)
$Composition_{t-3}$	12189596.4**	−245854.0	12882808.2**	706072.1
	(16.61)	(−1.24)	(16.69)	(2.56)
$FundReturn_t$				−338333893.4
				(−0.59)
$FundReturn_{t-1}$		455609698.1		348185116.3
		(1.13)		(1.22)
$FundReturn_{t-2}$		689106596.4***		613082885.0*
		(74.98)		(11.16)
$FundReturn_{t-3}$		782152623.3**		755689615.6**
		(17.06)		(13.92)
$Constant$	−9331621.2	40078961.5**	−8326689.1	33932165.3
	(−1.29)	(29.19)	(−1.38)	(4.42)
Observations	6402	6233	6374	6204
R^2	0.155	0.160	0.156	0.161
Fund FE	YES	YES	YES	YES
Time FE	YES	YES	YES	YES

Note: The table summarizes results of panel regressions of the dependent variable Net Flow on day t on a Composition variable on day t and up to three days lagged and a Fund Return variable on day t and up to three days lagged. The sample is restricted to North-American indices. T-statistics are reported in parentheses, and *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

While we find a statistically highly significant portfolio composition effect in our sample of European market indices, we do not find a significant effect of portfolio composition on future net fund flows in our sample of North-American market indices. As such, the effect in our pooled sample seems to be primarily driven by the European market indices. One potential reason for the non-existence of the effect in the North-American sample could be a power issue as the observations in this sample are only 10% of those in the European sample. On the contrary, we find a more pronounced return-chasing behavior in the North-American sample than in the European sample. Apart from potential power issues, this might be one reason why we do not observe a portfolio composition effect in the North-American sample.

Third, we analyze potential drivers of the portfolio composition effect. One potential driver of the effect could be macroeconomic news. We want to understand whether the portfolio composition effect is primarily driven by days on which macroeconomic news are priced in. We argue that macroeconomic news such as for example a political event (e.g. the passing of a trade agreement, the declaration of war, etc.), the announcement of a base rate change by the Federal Reserve Bank of America or the European Central Bank, or the spread of a disease are likely to affect all stocks of an index in a similar direction. As such, it is likely that macroeconomic news lead to extreme portfolio compositions. After unexpected bad macroeconomic news, it is likely that all stocks of an index trade at a daily loss, whereas after unexpected good macroeconomic news it is likely that all stocks of an index trade at a daily gain. The reason for these changes in portfolio composition are likely to be systematic (in the sense that all stocks are affected) rather than firm-specific. In what follows, we analyze whether our effect is primarily driven by these systematic changes in our portfolio composition measure or whether idiosyncratic changes caused by firm-specific information drive the documented portfolio composition effect. To test this, we include an "all-winner-dummy" for days on which all stocks of an index trade at a gain to our regression model and an "all-loser-dummy" for days on which all stocks of an index trade at a loss. We also add these dummies lagged by one, two, and three days. The results are reported in Table 3.8. We find that none of the all-winner/all-loser-dummies gains statistical significance. Even after controlling for days with extreme portfolio composition, the coefficient of the one-day lagged portfolio composition variable remains statistically significant and changes only slightly in economic magnitude compared to the result from Table 3.4.

Finally, we exclude that the portfolio composition measure introduced in our study proxies for return dispersion. More precisely, we analyze to what extent the cross-sectional standard deviation of stock returns of an index captures something similar to our measure of portfolio composition. There is literature on the impact of return dispersion on fund returns (Stivers and Sun, 2010; Liu et al., 2019). The composition measure we investigate is mathematically related to return dispersion. If the daily standard deviation of returns of the index members is large, it is also likely that the portfolio composition measure reflects the high return dispersion by neither being close to zero (i.e. all stocks exhibit a loss) nor being close to one (i.e. all stocks exhibit a gain). However, while return dispersion measures the absolute deviations of stock returns from the cross-sectional mean return of the index, the portfolio composition measure takes the direction into account. In other words, a small return dispersion can result from many winner stocks, many loser stocks, or even winner and loser stocks which all have a similar return. As such, our portfolio composition measure is likely to capture more than return dispersion. We examine whether the portfolio composition effect persists once we control for the return dispersion of an index. The results are shown in Table 3.9. We find that the portfolio composition effect persists even after controlling for the return dispersion of the index. While none of the included cross-sectional standard deviation variables gains statistical significance, the coefficient of the one-day lagged portfolio composition remains statistically significant, but decreases in magnitude.

Table 3.8: Portfolio Composition and ETF Fund Flows with Extreme Portfolio Composition Dummies

Dep. Variable	$NetFlow_t$			
	(1)	(2)	(3)	(4)
$Composition_t$			224485.5	535205.8
			(0.36)	(1.47)
$Composition_{t-1}$	2418668.1**	1791111.4*	2446369.2**	1815214.7**
	(2.47)	(2.18)	(2.51)	(2.25)
$Composition_{t-2}$	2530589.0	2129727.2	2581434.9	2166800.2
	(1.25)	(1.04)	(1.26)	(1.05)
$Composition_{t-3}$	3291765.6	2968248.8	3266929.0	2949603.8
	(1.50)	(1.35)	(1.48)	(1.35)
$AllWinner_t$	−223863.4	−72152.9	−462818.2	−257578.9
	(−0.22)	(−0.07)	(−0.52)	(−0.31)
$AllWinner_{t-1}$	−2029012.8	−2175879.6	−1994955.8	−2143162.4
	(−1.26)	(−1.27)	(−1.25)	(−1.26)
$AllWinner_{t-2}$	1012316.9	1170430.5	1003359.0	1164442.2
	(0.60)	(0.70)	(0.59)	(0.69)
$AllWinner_{t-3}$	−1151365.9	−1428280.0	−1157479.5	−1429012.9
	(−0.81)	(−0.92)	(−0.79)	(−0.90)
$AllLoser_t$	1944337.3	2203534.4	1982511.0	2162818.6
	(1.42)	(1.46)	(1.50)	(1.60)
$AllLoser_{t-1}$	−12920.9	106484.1	−229768.3	−125118.9
	(−0.07)	(0.56)	(−1.47)	(−0.71)
$AllLoser_{t-2}$	−2139776.7	−2249652.1	−2127471.9	−2230290.9
	(−1.17)	(−1.18)	(−1.15)	(−1.16)
$AllLoser_{t-3}$	−395123.0	−627694.0	−417019.3	−665206.2
	(−0.30)	(−0.49)	(−0.32)	(−0.52)
$FundReturn_t$				−18759053.5
				(−0.58)
$FundReturn_{t-1}$		30662222.7		30596262.5
		(1.21)		(1.24)
$FundReturn_{t-2}$		31123848.6		31465329.7
		(1.40)		(1.42)
$FundReturn_{t-3}$		30529448.1		30398365.5
		(1.27)		(1.28)
$Constant$	−972241.1	1105880.5	−1012963.3	915338.1
	(−1.59)	(0.57)	(−1.66)	(0.51)
Observations	68332	65327	68223	65207
R^2	0.018	0.019	0.018	0.019
Fund FE	YES	YES	YES	YES
Time FE	YES	YES	YES	YES

Note: The table summarizes results of panel regressions of the dependent variable Net Flow on day t on a Composition variable on day t and up to three days lagged, All Winner dummy which is one if all stocks are winners on day t and the dummy lagged up to three days, All Loser dummy which is one if all stocks are losers on day t and the dummy lagged up to three days and a Fund Return variable on day t and up to three days lagged. T-statistics are reported in parentheses, and *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

Table 3.9: Portfolio Composition and ETF Fund Flows with Cross-Sectional Standard Deviation of Stock Returns

Dependent Variable	$NetFlow_t$	
	(1)	(2)
$Composition_t$		391495.0
		(1.23)
$Composition_{t-1}$	1556442.7*	1567143.8**
	(2.16)	(2.26)
$Composition_{t-2}$	2231755.5	2278076.3
	(1.02)	(1.04)
$Composition_{t-3}$	3035496.0	3002865.0
	(1.37)	(1.35)
$StdDev_t$		49412435.1
		(0.94)
$StdDev_{t-1}$	95377775.6	88697012.6
	(1.54)	(1.58)
$StdDev_{t-2}$	−29853436.3	−35526246.2
	(−1.28)	(−1.35)
$StdDev_{t-3}$	−10173499.4	−14822118.0
	(−0.61)	(−0.95)
$FundReturn_t$		−20527500.7
		(−0.61)
$FundReturn_{t-1}$	29324218.8	29656748.1
	(1.21)	(1.24)
$FundReturn_{t-2}$	34374791.7	34974925.6
	(1.36)	(1.38)
$FundReturn_{t-3}$	29711655.8	30050861.7
	(1.25)	(1.25)
$Constant$	−315423.3	−1081120.3
	(−0.16)	(−0.60)
Observations	65315	65193
R^2	0.019	0.019
Fund FE	YES	YES
Time FE	YES	YES

Note: The table summarizes results of panel regressions of the dependent variable Net Flow on day t on a Composition variable on day t and up to three days lagged, the cross-sectional standard deviation on day t and up to three days lagged and a Fund Return variable on day t and up to three days lagged. T-statistics are reported in parentheses, and *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

3.4 Conclusion

We run three experiments to investigate how individuals evaluate portfolio investment decisions. Across all experiments, we find that participants are more willing to invest in a portfolio with a larger number of winner relative to loser stocks than in an alternative portfolio with a larger number of loser relative to winner stocks, although the portfolios have realized identical overall returns. The documented effect persists, if we keep the expected returns and volatility across portfolios identical. The observed investment decisions are consistent with our proposed theoretical framework of categorical thinking and mental accounting which implies that individuals use a counting heuristic to evaluate portfolio investment decisions.

We then use our well-identified experimental evidence on individuals' evaluation of portfolio investment decisions to test whether portfolio composition also matters in financial markets. In particular, we analyze the relation between net fund flows of national equity market index ETFs and our measure of portfolio composition for leading European and North-American equity market indices over the period 2016-2019. Consistent with our experimental evidence, we find that historical fund flows of leading equity market index ETFs are affected by the index previous-day composition of winner and loser stocks. Importantly, the effect remains stable and statistically significant even after controlling for the index return.

To better understand how individuals evaluate a portfolio and consequently how they make portfolio investment decisions, this paper proposes a simple measure: the composition of winner and loser stocks of a portfolio. This measure is arguably simple since it ultimately boils down to a counting heuristic. However, doesn't this measure capture an impression about the performance of a portfolio which people can easily and quickly gain? Winner stocks can easily be distinguished from loser stocks such that people can gain a good impression of how many stocks in their portfolio or of an index are winners relative to how many stocks are losers.

While the counting heuristic proposed in this paper can be grounded in well-established psychological frameworks, we do not claim that our measure of portfolio composition is the only performance measure that can be thought of when testing how performance information on the individual-stock level and performance information on the overall-portfolio level affect portfolio investment decisions. Future research may identify alternative or even complementary performance measures to deepen the understanding of how households and retail investors make portfolio investment decisions.

Chapter 4

Why So Negative? Belief Formation and Risk-Taking in Boom and Bust Markets *

4.1 Introduction

How do individuals form expectations about future stock returns? The answer to this question is crucial to understand differences in risk-taking over time and in particular across market cycles. A key assumption in models that generate time-variation in risk-taking is that investors have rational expectations, which are immediately updated according to Bayes' rule when new information arrives (Barberis et al., 2001; Campbell and Cochrane, 1999; Grossman and Shiller, 1981). Their authors assume implicitly that agents know the objective probability distribution in equilibrium and are as such fully aware of the counter-cyclical nature of the equity risk premium (Nagel and Xu, 2019). Yet, a number of recent surveys of investors' expectations show that this is not the case, and that investors – if anything – have rather pro-cyclical expectations: they are more optimistic in boom markets and less optimistic

* Authors: Pascal Kieren, Jan Müller-Dethard, and Martin Weber. All authors are at the University of Mannheim. For valuable comments, we thank Nick Barberis, Camelia Kuhnen, Oliver Spalt, Stefan Trautmann, Georg von Weizsäcker, as well as participants of the Masstricht Behavioral Experimental Economics Symposium 2019, the Experimental Finance 2019, the ESA 2019, the Annual Meeting of the SJDM 2019, the Annual Meeting of the AEA 2020, the Lynne & Andrew Redleaf Foundation Graduate Student Conference 2020 at Yale University, and seminar participants at the University of Mannheim. We gratefully acknowledge funding by the German Research Foundation (DFG grant WE993/15-1).

in recessions (Amromin and Sharpe, 2014; Giglio et al., 2019; Greenwood and Shleifer, 2014).

In the light of this inconsistency, it is imperative to obtain a deeper understanding of how investors incorporate new information when they form expectations, and whether this could ultimately explain differences in risk-taking across macroeconomic cycles. Prior research has shown that investors put too much probability weight on new information, if the information looks representative of previously observed data (Kahneman and Tversky, 1972). Gennaioli and Shleifer (2010) as well as Gennaioli et al. (2012, 2015) show that such a representativeness can generate and amplify boom/bust financial crises based entirely on investors' beliefs. Besides the representativeness of the outcome history, Kuhnen (2015) shows that agents learn differently from outcomes in the negative domain than from the same outcome history in the positive domain. Both findings together and individually can lead to systematic distortions in how investors learn from outcomes and how they incorporate beliefs in their decision-process.

In this study, we investigate whether distorted belief formation rules (i.e. systematic violations of Bayes' rule) can explain differences in risk-taking across recessions and boom markets. To examine this relation, we conduct an experimental study with two different learning environments that closely resemble key characteristics of financial market cycles. The first learning environment characterizes a market setting in which subjects exclusively learn either in the positive (i.e. boom) or in the negative (i.e. recession) domain. The second learning environment characterizes a potentially more realistic market setting in which subjects learn from mixed-outcome distributions with either positive expected value (i.e. boom) or negative expected value (i.e. bust). We test 1) how different learning environments affect the formation of return expectations, 2) how systematic differences in beliefs resulting from different learning environments translate to risk-taking, and 3) whether different learning environments not only affect subjects' beliefs but also their risk preferences.

While recent survey data on expectations are helpful to establish a link between subjective beliefs and investment decisions, they do not allow inference about how investors depart from rational expectations without imposing strong assumptions. In an experiment however, we can establish a setting in which we have direct control over objective (rational) expectations and can compare them to participants' subjective beliefs. This allows us to document systematic errors in the belief formation process, which we can then relate to the subjects' investment choice.

In our experiment, we combine an abstract Bayesian updating task (similar to Grether, 1980; and more recently adopted by Glaser et al., 2013, or Kuhnen, 2015) with an unrelated incentive-compatible investment task in a financial environment. In the Bayesian updating task, subjects have to incorporate a sequence of information signals into their beliefs to estimate the likelihood that an asset pays dividends drawn from one of two distributions. Depending on the learning environment, the information subjects receive is either exclusively positive (boom treatment) or negative (bust treatment) in Experiment 1, or both positive and negative but drawn from distributions with either positive (boom treatment) or negative expected value (bust treatment) in Experiment 2. The underlying probability distribution, however, from which the information is drawn, is completely identical in both learning environments. In other words, a Bayesian agent should make identical forecasts, irrespective of whether he learns in the positive or negative environment.

After subjects completed the forecasting task, they make an unrelated investment decision in either a risky or an ambiguous lottery, which serves as a *between-subject measure of belief- and preference-based risk-taking*. In the ambiguous lottery, we purposefully give participants room to form subjective beliefs about the underlying true probability distribution. In the risky lottery, we have perfect control over subjects' return and risk expectations since both probabilities and outcomes are known. As such, investments in the ambiguous lottery are affected by both subjects' risk preferences and their beliefs about the underlying probability distribution, while investments in the risky lottery serve as a measurement tool for risk aversion. The between-subject comparison finally allows us to isolate the effect of belief-induced risk-taking caused by outcome-dependent learning environments.

Our findings can be summarized as follows. First, we find that subjects who learn to form beliefs in adverse market environments take significantly less risk in an unrelated ambiguous investment task than subjects who learn to form beliefs in favorable market environments. Once there is room to form subjective beliefs, subjects in the bust treatment invest on average 20% less in the ambiguous lottery compared to subjects in the boom treatment. In line with their lower willingness to take risks, subjects who have learned to form beliefs in adverse market environments are also substantially more pessimistic about the success probability of the ambiguous lottery (by about 19 percentage points). In the risky lottery, when expectations are fixed, we can directly test whether adverse learning environments also affect the subjects' risk aversion. However, we do not find any significant difference between treatments on subjects' investment in an unrelated risky investment option. This indicates that subjects' risk preferences (i.e. their risk aversion) remained stable and were unaltered by the environment in which they learned to form beliefs. Effectively, this finding suggests that when individuals form expectations in adverse learning environments (as is frequently the case in recessions), they become substantially more pessimistic about future prospects. However, this pessimism only translates to lower risk-taking when there is uncertainty in the investment process.

Second, we investigate how adverse learning environments induce pessimism in subjects' return expectations. We find that subjects who forecast the probability distribution of an asset in an adverse learning environment (bust treatment) are significantly more pessimistic in their average probability estimate than those subjects who forecast the identical probability distribution in a favorable learning environment (boom treatment). This indicates that the frame of the learning environment crucially affects subjects' belief formation, although the actual learning task is identical. In other words, in our setting a Bayesian forecaster would make identical probability forecasts irrespective of the underlying learning environment. The resulting asymmetry in belief formation resembles a pessimism bias as subjects' beliefs in the bust treatment show larger deviations from Bayesian beliefs compared to subjects' beliefs in the boom treatment. This finding is independent of whether subjects learn exclusively from negative outcome lotteries (Experiment 1) or from mixed-outcome lotteries with negative expected value (Experiment 2), and extends previous work by Kuhnen (2015).

Third, we seek to better understand the link of how forecasting in different learning environments affects risk-taking and for whom the effect is most pronounced. We find that those subjects who show above-median forecasting ability in the learning task of the experiment critically drive the results. In particular, these subjects show a stronger link between the pessimism induced by the initial adverse learning environment and the subsequent (lower) risk-taking. However, and importantly, even these subjects still exhibit a pronounced pessimism bias in their probability assessment, which subsequently translates to more pessimistic beliefs about the success probability of the ambiguous asset. To rationalize why the risk-taking of the seemingly better performing agents is more affected by the learning environment, we test whether they share particular socio-demographic characteristics or whether they are more involved in the experimental task. We find that above-median forecasters spend significantly more time on reading the instructions and make significantly less basic, directional wrong updating errors than below-median forecasters. As such, our analyses rather support the latter argument, which suggests that the effect reported here might be even stronger in the real economy, where stakes and involvement are presumably higher.

Finally, we provide evidence that the pessimism induced by adverse learning environments within our experimental setup even affects subjects' return expectations in the real economy. When asked to provide a return forecast of the Dow Jones Industrial Average, subjects in the bust treatment are significantly more pessimistic about the future performance of the index than their peers in the boom treatment. In addition to the more pessimistic expectations, we find that subjects who learn in adverse financial conditions provide negative return estimates, while those learning in rather favorable financial conditions provide positive return estimates. Given that we are able to systematically manipulate return expectations for real world market indices even in a short-living learning environment as in our experiment, we believe that the effect reported here is even more generalizable in the real economy.

Our findings contribute to several strands of literature. Most importantly, our results provide a direct and causal link of how systematic distortions in investors' expectations can affect their willingness to take financial risks. The most prominent rational expectations models that generate high volatility of asset prices and the countercyclical equity risk premium introduce modifications into the representative agent's utility function, which effectively generates countercyclical risk aversion (Campbell and Cochrane, 1999; Barberis et al., 2001). This implies that during bust markets investors become more risk averse and consequently demand a higher risk premium, and they become less risk averse during boom markets, thus demanding a lower risk premium. Recently, Cohn et al. (2015) present experimental evidence supporting this notion, while Guiso et al. (2018) present survey evidence in line with this argument.[1]

However, in our experimental design, we can confidently rule out that a change in preferences can explain our findings. Instead, we show that expectations and how they are formed can generate similar feedback loops as implied by countercyclical risk aversion without having to assume unstable risk preferences. If bust markets systematically induce pessimistic expectations about future returns for a substantial subset of investors, this may reduce the aggregate share invested in risky assets of an economy, which in turn generates downward pressure on prices due to excess supply. In line with our results, Amromin and Sharpe (2014) find that households' lower willingness

[1] There are also recent papers who challenge the notion of countercyclical risk aversion as tested in Cohn et al. (2015) such as Alempaki et al. (2019) and König-Kersting and Trautmann (2018).

to take risks during recessions is rather driven by their more pessimistic subjective expectations than by countercyclical risk aversion. Similarly, Weber et al. (2013) show that changes in risk-taking of UK online-broker customers over the financial crisis of 2008 were mainly explained by changes in return expectations and to a lesser degree by changes in risk attitudes.

Furthermore, our study also relates to the findings reported in recent surveys of investor return expectations (Amromin and Sharpe, 2014; Giglio et al., 2019; Greenwood and Shleifer, 2014). A common finding is that survey expectations of stock returns are pro-cyclical (i.e. investors are more optimistic during boom markets and more pessimistic during recessions), and as such inconsistent with rational expectation models. A first attempt to reconcile this puzzling finding was made by Adam et al. (2020), who test whether alternative expectation hypotheses proposed in the asset pricing literature are in line with the survey evidence. However, they reject all of them. In our study, we also find that investors' expectations are pro-cyclical, as they are more optimistic when learning in favorable environments then when learning in adverse environments. As such, the belief formation mechanism tested in our study may provide an interesting starting point for alternative theories of belief updating featuring pro-cyclical expectations.

Finally, our finding also relates to the literature on investors' experience (Graham and Narasimhan, 2004; Malmendier and Nagel, 2011, 2015; Malmendier and Tate, 2005; Malmendier et al., 2011). The literature posits that events experienced over the course of an investor's life have persistent and long-lasting effects. In the spirit of this literature, learning rules, if more frequently applied throughout investors' lives, may exert a greater influence on the way they form beliefs and ultimately on their willingness to take risks. For example, investors who experienced the Great Depression in their early career were more frequently exposed to negative stock returns, which might have affected the way they form beliefs about future economic events. As a result, these investors are more pessimistic in their assessment of future stock returns and less willing to take financial risks compared to those who experienced the post-war boom until the 1960s in their early life.

The mechanism reported here and its effect on risk-taking may have important policy implications. For example, if investors exhibit overly pessimistic expectations in recessions, they may expect lower returns and reduce their equity share. As a consequence, the pro-cyclical nature of beliefs resulting from partly distorted belief formation rules reported in our study may amplify the intensity and the length of market phases.

The remainder of the paper is organized as follows. In Section 4.2, we outline the experimental design, and briefly discuss the most important design aspects. In Section 4.3, we state our hypotheses, while in Section 4.4 we describe summary statistics of our sample and randomization checks. In Section 4.5, we present our findings, and in Section 4.6 we conclude.

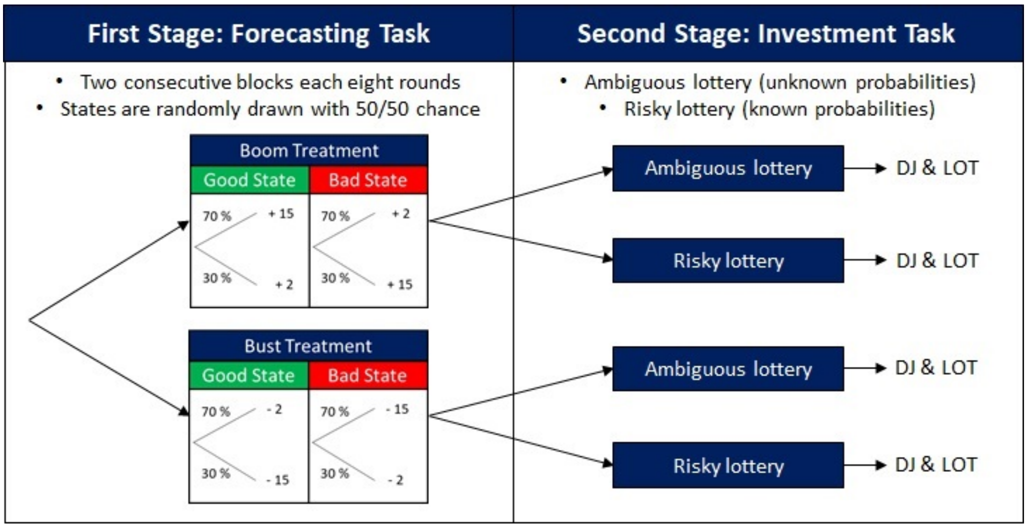
4.2 Experimental Design

Seven-hundred fifty-four individuals (458 males, 296 females, mean age 34 years, 10.3 years standard deviation) were recruited from Amazon Mechanical Turk (MTurk) to participate in two online experiments. MTurk advanced to a widely used and accepted recruiting platform for economic experiments. Not only does it offer a larger and more diverse subject pool as compared to lab studies (which frequently rely on students), but it also provides a response quality similar to that of other subject pools (Buhrmester et al., 2011; Goodman et al., 2013).

Both experiments consist of two independent parts, a forecasting task (Bayesian updating) and an investment task (see Figure 4.1). The experiments differ with respect to the forecasting task, but are identical with respect to the investment task. In the forecasting task, we create a learning environment which resembles key characteristics of boom and bust markets.

In Experiment 1, we focus on the domain (positive vs. negative returns) in which subjects primarily learn across different market cycles. As such, we let subjects learn from either exclusively positive outcome-lotteries (boom-scenario) or negative outcome-lotteries (bust-scenario). However, even in recessions agents occasionally observe positive returns, but the magnitude is

Figure 4.1: Structure and Flow of the Experiments



Note: this figure documents the structure and the flow of our two experiments. Subjects do a forecasting task which is followed by an investment task. At the beginning, subjects are randomly assigned to either a boom treatment or a bust treatment (here the lotteries of Experiment 1 are illustrated). In the first stage of the experiment, they make 16 forecasts in total split in two blocks of eight rounds. In the second stage of the experiment, they are assigned either to invest in an ambiguous lottery (unknown probabilities) or a risky lottery (known probabilities) and are asked to make a 6-month return forecast for the Dow Jones Industrial Average (DJ) as well as to answer a 10-item life orientation test (LOT).

on average smaller than the magnitude of observed negative returns. During the last two financial crises, the frequency of observing a negative monthly return of the MSCI AC World index was 66.67 % for the DotCom Crisis and 68.42 % for the 2008 Financial Crisis, while the average realized monthly return was -1.17 % and -2.11 %, respectively, as displayed in Figure 4.1.[2]

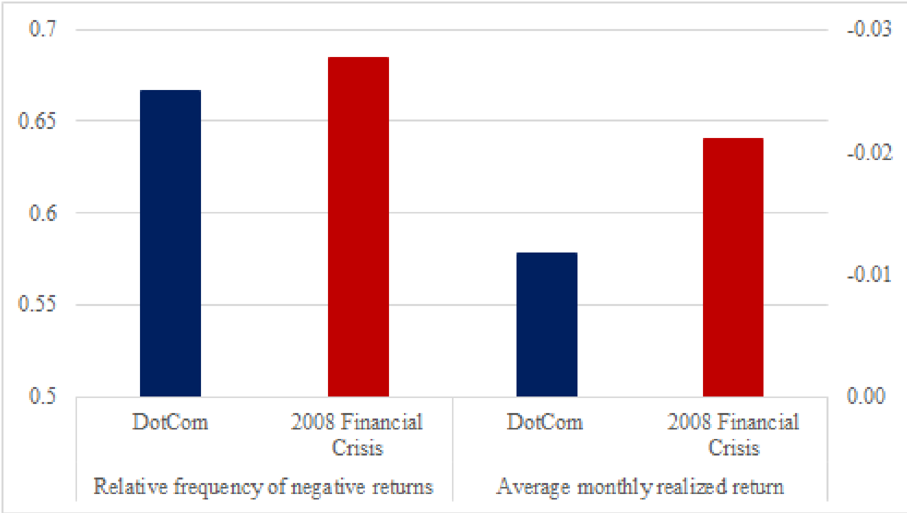
To account for this fact, we conduct another experiment with an even more realistic learning environment. In Experiment 2 subjects learn from mixed outcome-lotteries, which either have a positive expected value (boom-scenario) or a negative expected value (bust-scenario).

4.2.1 Detailed Description of the Experiment

In the forecasting task of both experiments, subjects receive information about a risky asset, whose payoffs are either drawn from a “good distribution” or from a “bad distribution”. Both distributions are binary with

[2] Business cycles are defined using the NBER Business Cycle Expansion and Contractions Classification.

Figure 4.2: Characteristics of Boom and Bust Market Phases



Note: This figure documents both the relative frequency of observing a negative monthly return of the MSCI All Country World Index as well as the average monthly return for the last two financial recessions. Recessions are defined according to the NBER US Business Cycle Contraction classification. The left y-axis refers to the relative frequency of negative returns. The right y-axis (reversed scale) refers to the average monthly realized returns.

identical high and low outcomes. In the good distribution, the higher payoff occurs with a 70 % probability while the lower payoff occurs with a 30 % probability. In the bad distribution, the probabilities are reversed, i.e. the lower payoff occurs with a 70 % probability while the higher payoff occurs with a 30 % probability. The actual payoffs depend on both the experiment and the treatment to which subjects are assigned. In both experiments, subjects are randomly assigned to either a "boom" treatment or a "bust" treatment. In the first experiment, the payoffs of the risky asset are either exclusively positive or negative, which resembles domain-specific learning. The payoffs in the boom treatment are either +15, or +2, whereas they are −2, or −15 in the bust treatment. In the second experiment, the payoffs of the risky asset are drawn from mixed-outcome lotteries, with either a positive or a negative expected value. The payoffs in the boom treatment are either +15, or −2, whereas they are +2, or −15 in the bust treatment. While the payoffs across treatments are mirrored, the underlying probability distributions of the risky asset from which outcomes are drawn are identical.

In both experiments, subjects make forecasting decisions in two consecutive blocks each consisting of eight rounds. At the beginning of each block, the computer randomly determines the distribution of the risky asset (which can be good or bad). In each of the eight rounds, subjects observe a payoff of the risky asset. Afterwards, we ask them to provide a probability estimate that the risky asset draws from the good distribution and how confident they are about their estimate. As such, subjects will make a total of 16 probability estimates (8 estimates per block). To keep the focus on the forecasting task and to not test their memory performance, we display the prior outcomes in a price-line-chart next to the questions. At the beginning of the experiment and before they could continue, subjects had to correctly answer three questions the answers to which indicated their understanding of the experiment (see Appendix C).

In the second part of each experiment, the investment task, subjects were randomly assigned to invest in either an *ambiguous* or a *risky* lottery with an endowment of 100 Cents (Gneezy and Potters, 1997). In both lotteries, the underlying distribution to win is 50 %. However, to introduce uncertainty and to provide subjects the freedom to form beliefs, the success probability remains unknown to them in the ambiguous lottery. In both lotteries, subjects can earn 2.5 times the invested amount if the lottery succeeds, whereas they lose the invested amount if the lottery fails. Subjects can keep the amount not invested in the lottery without earning any interest. In addition to the lottery investment, subjects in the ambiguous treatment are asked to provide an estimate of the success probability of the ambiguous lottery. Subjects in the risky treatment are not asked about a probability estimate as the objective success probability is known and clearly communicated.

The experiments concluded with a brief survey about subjects' socio-economic background, a 10-item inventory of the standard Life Orientation Test (Scheier et al., 1994), self-assessed statistic skills, stock trading experience and whether a participant was invested during the last financial crisis. In addition, subjects were asked to provide a 6-month return forecast of the Dow Jones Industrial Average index on a twelve-point balanced Likert scale.

Both parts of the experiment were incentivized. In the first part, participants were paid based on the accuracy of the probability estimate provided. Specifically, they received 10 cents for each probability estimate within 10 % $(+/-5\%)$ of the objective Bayesian value. In the second part of the experiment, subjects received the amount not invested in the lottery plus the net earnings from their lottery investment. Both studies took approximately 9 minutes to complete and participants earned $1.93 on average.

4.2.2 Discussion of Important Aspects

Overall, our design allows us to test whether asymmetric belief formation in boom and bust markets can account for time variation in risk taking. As it is imperative for our design to ensure that risk preferences remain constant and are unaffected by the forecasting task, a few aspects warrant a brief discussion. First, feedback regarding the accuracy of subjects' probability estimates was only provided at the very end of the experiment. This was done to not only avoid wealth effects, but also to ensure that subjects do not hedge the lottery investment against their earnings from the forecasting task, which would inevitably affect their risk-taking. Second, we abstract from using predisposed words like "boom", "bust", or similar financial jargon. This circumvents evoking negative or positive emotions (such as fear), experience effects, and other confounding factors, which would distort a clear identification of belief-induced risk-taking. Third, by exploiting the between-subject variation in the lottery tasks, we can directly investigate whether the forecasting task in different domains unintentionally affects risk preferences. More precisely, we can exclude that learning from adverse market conditions affects risk preferences.[3]

[3] Although we can directly control for the effect of positive and negative numbers on risk preferences in our design, Kuhnen (2015) concludes as well that risk preferences remain unaffected.

4.3 Hypotheses

We have two main hypotheses, one regarding the forecasting task and one regarding the investment task. First, we test whether forecasting in adverse learning environments systematically induces pessimism in subjects' belief formation. In the first experiment, we investigate the effect of domain-specific learning environments on subjects belief formation as originally tested by Kuhnen (2015). In the second experiment, we examine whether this effect is restricted to domain-specific learning or whether it generalizes to mixed-outcome learning environments as frequently observed in both boom markets and in recessions.

H1: Pessimism Bias

Subjects in the bust treatment are significantly more pessimistic in their average probability forecast both relative to the objective Bayesian forecast and relative to the subjects in the boom treatment.

Next, we investigate the main treatment effect of our study. In particular, we aim to examine whether asymmetric belief formation in boom and bust markets could explain differences in risk-taking. To do so, we introduce a between-subject measure of belief- and preference-based risk-taking. In the risky treatment, we have perfect control over subjects' return and risk expectations since both probabilities and outcomes are known and clearly communicated. As such, the risky treatment serves as a measurement tool for risk aversion. In the ambiguous treatment however, we intentionally give participants room to form subjective beliefs as there is uncertainty about the true probability. If the induced pessimism leads to more pessimistic expectations, we should observe a stronger treatment effect in the ambiguity treatment as the absence of perfect certainty about the success probability of the ambiguous lottery leaves more room for expectations (Klibanoff et al., 2005).

H2a: Belief-Induced Risk-Taking

Subjects in the bust treatment invest significantly less in the ambiguous lottery than subjects in the boom treatment.

H2b: Preference-Based Risk-Taking

Investments in the risky lottery should not significantly differ across treatments.

4.4 Summary Statistics and Randomization Checks

Table 4.1 presents summary statistics, Panel A for Experiment 1 and Panel B for Experiment 2. Overall 754 subjects participated in our studies, with an average age of 35.15 years in Experiment 1 (33.53 years in Experiment 2). Forty-five percent (thirty-four percent) were female. Subjects reported average statistical skills of 4.19 out of 7 (4.47) and are medium experienced in stock trading, with a self-reported average score of 3.64 out of 7 (3.94). Roughly thirty-nine percent (forty-four) were invested during the 2008 Financial Crisis.

Additionally, we tested whether our randomization successfully resulted in a balanced sample. Table 4.1 also reports the mean and standard deviation of each variable split by treatment. Differences were tested using rank-sum tests, or χ^2-tests for binary variables. As we find no significant difference between our treatments for any variable, our randomization was successful. As such, we cannot reject the null hypothesis that the socio-economic background of the subjects is balanced between our boom and bust treatment.

4.5 Results

We present answers to the following questions: 1) Do agents learn to form beliefs differently across market cycles?; 2) if belief formation is systematically different across market cycles, do the resulting beliefs translate to systematic differences in risk-taking?; 3) what is the mechanism behind the effect?; 4) who is most affected?; and 5) what are the boundaries?

Table 4.1: Summary Statistics on Subjects

Panel A: Experiment 1 **Variable**	**Full Sample (N=350)**	**Boom (N=174)**	**Bust (N=176)**	**Differ-ence**	**p-value**
Age	35.15	34.76	35.54	0.78	0.76
	(11.52)	(11.18)	(11.86)		
Female	0.45	0.47	0.43	0.04	0.44
	(0.50)	(0.50)	(0.50)		
Statistical Skills	4.19	4.22	4.16	0.06	0.91
	(1.62)	(1.51)	(1.72)		
Experience Stock Trading	3.64	3.73	3.56	0.17	0.42
	(1.88)	(1.84)	(1.92)		
Invested Financial Crisis	0.39	0.39	0.39	0	1
	(0.49)	(0.49)	(0.49)		

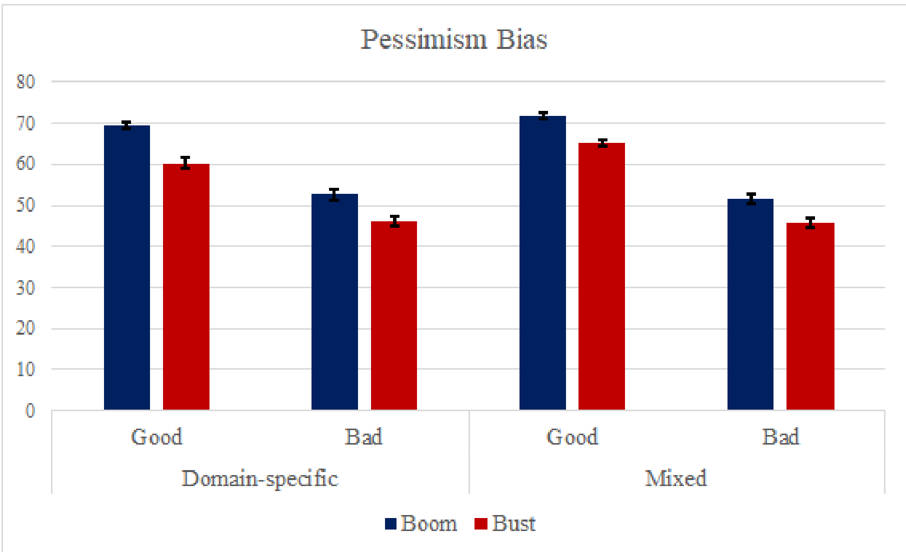
Panel B: Experiment 2 **Variable**	**Full Sample (N=403)**	**Boom (N=207)**	**Bust (N=196)**	**Differ-ence**	**p-value**
Age	33.53	32.73	34.37	1.63	0.07
	(9.03)	(8.46)	(9.55)		
Female	0.34	0.33	0.35	0.02	0.69
	(0.48)	(0.47)	(0.48)		
Statistical Skills	4.47	4.40	4.55	0.15	0.42
	(1.67)	(1.69)	(1.65)		
Experience Stock Trading	3.94	3.89	3.98	0.09	0.52
	(1.99)	(1.95)	(2.03)		
Invested Financial Crisis	0.44	0.41	0.47	0.06	0.24
	(0.50)	(0.49)	(0.50)		

Note: This table shows summary statistics for our experimental data. Reported are the mean and the standard deviation (in parentheses) for the whole sample (Column 1) and split across treatments (Column 2 and 3). Column 4 presents randomization checks. Differences in mean were tested using rank-sum tests, or χ^2-tests for binary variables. The p-value is reported in Column 5. *Female* is an indicator variable that equals 1 if a participant is female. *Statistical skills* denotes participants' self-assessed statistical skills on a 7-point Likert scale. *Experience in stock trading* is the self-reported experience participants have in stock trading, assessed by a 7-point Likert scale. *Invested financial crisis* is an indicator that equals 1 if participants were invested in the stock market during the last financial crisis.

4.5.1 Distorted Belief Formation

First, we examine whether belief formation in bust markets differs from belief formation in boom markets and to what extend the effect depends on the underlying characteristic of the learning environment. While participants learn exclusively from either only positive or negative outcome lotteries (i.e. domain-specific learning) in Experiment 1, they learn from mixed outcome lotteries with either positive or negative expected value (i.e. mixed-outcome dependent learning) in Experiment 2. Figure 4.3 displays the average probability estimate over eight rounds for good and bad distributions, separated by treatment and experiment.

Figure 4.3: Pessimism Bias



Note: This figure documents the pessimism bias. It depicts participants' average probability forecasts split by the underlying distribution they had to forecast (good or bad), the treatment they were in (boom or bust), and the experiment in which they participated (domain-specific forecasting or mixed-outcome forecasting). Displayed are 95 % confidence intervals.

In the domain-specific learning environment (Experiment 1), we find that subjects who forecast the distribution of an asset from negative numbers only (i.e. bust treatment) are significantly more pessimistic in their average probability estimate than those who forecast the identical distribution from positive numbers (i.e. boom treatment). This finding is independent of the type of distribution subjects witnessed (good or bad) and in line with previous work by Kuhnen (2015).

Interestingly, and perhaps more importantly for market cycles, this finding is not limited to domain-specific learning environments. Instead, those subjects who forecast distributions from mixed-outcome lotteries with negative expected value (bust treatment) are also more pessimistic in their average probability assessment than those who learn from mixed-outcome lotteries with positive expected value (boom treatment). In contrast, a Bayesian forecaster would provide completely identical probability estimates irrespective of the learning environment given the identical underlying distribution from which outcomes are drawn. To control for the objective posterior probability, we also run regressions of subjects' probability estimates on a bust-indicator and the objective Bayesian probability that the stock is in the good state. Results for both experiments pooled and individually are reported in Table 4.2.

Across both experiments, we find that beliefs expressed by subjects in the bust treatment are on average 6.43% lower (i.e. more pessimistic) than in the boom treatment ($p < 0.001$), confirming Hypothesis **H1**. This means that - holding the objective posterior constant - subjects update their priors differently when learning in adverse market environments compared to favorable environments. Remarkably, the magnitude of this pessimism bias does not significantly differ across experiments. In other words, the reported pessimism bias does not critically depend on whether subjects observe exclusively negative outcomes or mixed outcomes drawn from a distribution with negative expected value. In essence, our results imply that the way subjects form beliefs is different in bust markets than in boom markets.

Table 4.2: Pessimism Bias

Dependent Variable	*Probability Estimate (Subjective Posterior)*		
	Pooled Data	Domain-specific	Mixed
Bust	-6.425***	-6.218***	-6.742***
	(-6.16)	(-3.86)	(-4.88)
Objective Posterior	0.378***	0.370***	0.384***
	(23.94)	(17.21)	(17.09)
Constant	46.31***	45.96***	47.01***
	(10.82)	(7.02)	(8.24)
Observations	12048	5600	6448
R^2	0.262	0.244	0.279

Note: This table reports the results of three OLS regressions on how subjective posterior beliefs about the distribution of the lottery depend on the treatment. The dependent variable in the regression model, *Probability Estimate*, is the subjective posterior belief that the asset is paying from the good distribution. Independent variables include the *Bust* dummy, an indicator variable that equals 1 if participants were in the bust treatment and zero otherwise, as well as *Objective Posterior*, which is the correct Bayesian probability that the stock is good, given the information seen by the participant up to trial t in the learning block. Controls include age, gender, statistical skills, self-reported experience in stock trading, whether subjects were invested in the stock market during the last financial crisis, and the order of outcomes they experienced in the forecasting task. Reported are coefficients and t-statistics (in parentheses) using robust standard errors. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

4.5.2 Belief Formation and Risk-Taking

So far, we have shown that belief formation is systematically distorted by whether subjects learn during boom periods or during bust periods. Next, we investigate whether the induced pessimism resulting from biased belief formation in bust markets translates to lower risk-taking, without altering risk preferences. Table 4.3 summarizes subjects' average investment in the ambiguous and risky lottery, split by treatment.

The results reported in Table 4.3 provide a simple first test for our main hypothesis. In particular, while subjects in the bust treatment invest on average 36 out of 100 Cents into the ambiguous lottery, subjects in the boom treatment invest roughly 45 Cents into the ambiguous lottery ($p < 0.01$, two-sided t-test). As such, we find a significant treatment effect of learning to form beliefs in adverse market conditions on subjects' willingness to take risks. That is, subjects in the bust treatment invest on average 20 % less in the ambiguous lottery than subjects in the boom treatment. However, we find no such effect for investments in the risky lottery. While subjects in the boom treatment invest on average 39 Cents in the risky lottery, subjects in the bust treatment invest roughly 43 Cents, with no significant difference between the two ($p = 0.32$, two-sided t-test). Effectively, this result indicates that the pessimism induced by adverse market environments only translates to significantly lower risk-taking when there is room to form subjective expectations (i.e. the decision involves ambiguity). However, when expectations are fixed, risk-taking is not affected, which implies that asymmetric learning in different market environments does not alter individuals' inherent risk preferences.

To jointly test our main hypotheses while controlling for demographics and other potentially confounding factors, we specify the following regression model:

$$Investment_i = \beta_0 + \beta_1 Bust_i + \beta_2 Ambiguous_i + \beta_3 Bust_i x Ambiguous_i + \sum_{j=1}^{n} \beta_j X_{ij} + \epsilon_i \tag{4.1}$$

Table 4.3: Risk-Taking Across Macroeconomic Cycles I

	Treatment			
	Bust	Boom	Difference	p-value
Investment Ambiguous	36.31	44.82	-8.51***	< 0.01
Investment Risky	42.57	39.38	3.19	0.32

Note: This table summarizes the average investments (0 - 100) of participants in the ambiguous lottery and the risky lottery split by the treatment variable. Differences in investment between the treatments with the respective p-values from two-sided t-tests are also reported. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

where the dependent variable $Investment_i$ is the amount individual i invested in the risky/ambiguous asset. $Bust_i$ is a dummy that denotes if a subject learned to form beliefs in the bust treatment, while $Ambiguous_i$ is a dummy that denotes that the investment decision was made under ambiguity (i.e. unknown probabilities in the investment task). The interaction $Bust_i$ *x* $Ambiguous_i$ allows us to examine our main hypothesis, i.e. that subjects who learned to form beliefs in adverse environments invest significantly less in the ambiguous lottery where they have room to form subjective expectations. Finally, $\boldsymbol{X_{ij}}$ is a set of control variables including gender, age, statistic skills, stock trading experience, a life orientation test, the order of good and bad distributions in the forecasting task, and an indicator whether subjects were invested in the last financial crisis. We estimate our regression model using OLS with robust standard errors. However, results remain stable if we use a Tobit model instead.

In Table 4.4, we report our main finding for each experiment pooled and separately. In the pooled data, the negative interaction term indicates that individuals in the bust treatment invest significantly less in the ambiguous lottery compared to those in the boom treatment ($p = 0.011$), providing further evidence in favor of Hypothesis **H2a**. In the risky lottery, when expectations are fixed, we can directly test the effect of our forecasting task on subjects' risk aversion. However, we do not find any significant difference between treatments on subjects' investment in the risky lottery ($p = 0.47$), confirming Hypothesis **H2b**. This means that we cannot reject the null hypothesis that risk aversion for subjects who learned to form beliefs in adverse market

Table 4.4: Risk-Taking Across Macroeconomic Cycles II

Dependent Variable		*Investment*	
	Pooled Data	Domain-specific	Mixed
Bust	2.271	3.948	-0.948
	(0.72)	(0.86)	(-0.21)
Ambiguous	5.149*	5.540	4.473
	(1.71)	(1.26)	(1.04)
Bust x Ambiguous	-11.23**	-13.57**	-8.229
	(-2.54)	(-2.21)	(-1.25)
Constant	15.82*	20.32*	10.69
	(1.70)	(1.67)	(0.74)
Observations	753	350	403
R^2	0.060	0.080	0.069

Note: This table examines subjects' risk-taking across treatments. We report the results of OLS regressions for the whole sample, and for each experiment individually (Experiment 1: Domain-specific; Experiment 2: Mixed). The dependent variable is *Investment*, which denotes participants' invested amount (0 - 100) in the lottery they were assigned to. *Bust* is an indicator variable that equals 1 if participants were in the bust treatment. *Ambiguous* is an indicator variable that equals 1 if participants were asked to invest in the ambiguous lottery, and 0 if they invested in the risky lottery. Controls include age, gender, statistical skills, self-reported experience in stock trading, whether subjects were invested in the stock market during the last financial crisis, and the order of outcomes they experienced in the forecasting task. Reported are coefficients and t-statistics (in parentheses) using robust standard errors. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

environments is similar compared to subjects who learned to form beliefs in favorable market environments.

When looking at the results of each experiment separately, we find a strong and similar-sized effect for the domain-specific learning environment and a weaker - albeit statistically insignificant - effect for the mixed-outcome learning environment. Moreover, and consistent with the pooled data, we find no effect on subjects' risk preferences in neither the domain-specific nor the mixed-outcome learning environment. To better understand whether the effect in the pooled sample is primarily driven by domain-specific outcomes, or whether other factors are at play, we will run further regressions in Section 4.5.4.

4.5.3 Mechanism

In this section, we test whether expectations are indeed the driving mechanism behind our main effect. We designed the ambiguous treatment in such a way that we can assess participants' subjective beliefs about the success probability of the lottery and directly relate them to their investment decision. If expectations are the main driver of differences in risk-taking, we should observe that subjects who learned to form beliefs in either the negative domain-specific or in the negative expected value mixed-outcome learning environment are more pessimistic about the success probability of the ambiguous lottery. In addition, we would expect a positive correlation between the subjective probability estimate of the success chance of the ambiguous lottery and the amount invested in the ambiguous lottery. In order to directly test the implied mechanism, we estimate the following two OLS regression models for our pooled sample and for each experiment separately:

$$Probability_i = \beta_0 + \beta_1 Bust_i + \sum_{j=1}^{n} \beta_j X_{ij} + \epsilon_i \tag{4.2}$$

$$InvestmentAmbiguous_i = \beta_0 + \beta_1 Probability_i + \sum_{j=1}^{n} \beta_j X_{ij} + \epsilon_i \quad (4.3)$$

where $Probability_i$ is the subjective success probability of the ambiguous lottery of subject i, and $InvestmentAmbiguous_i$ is the investment of subject i in the ambiguous lottery. Findings for the first model are reported in Table 4.5 and for the second model in Table 4.6.

Table 4.5: Relation Between Treatment Variable and Probability Estimates

Dependent Variable	*Success Probability Estimate of Ambiguous Asset*		
	Pooled Data	Domain-specific	Mixed
Bust	-18.86***	-11.83***	-25.59***
	(-8.59)	(-3.74)	(-8.57)
Constant	55.83***	68.72***	41.10***
	(6.15)	(5.25)	(3.59)
Observations	377	177	200
R^2	0.241	0.176	0.349

Note: This table examines the underlying mechanism of how our treatment variable affects subjects' beliefs about the success probability of the ambiguous lottery. We report the results of OLS regressions for the whole sample, and for each experiment individually (Experiment 1: Domain-specific; Experiment 2: Mixed). The dependent variable is *Success Probability*, which denotes participants' beliefs about the success probability of the ambiguous lottery. *Bust* is an indicator variable that equals 1 if participants were in the bust treatment. Controls include age, gender, statistical skills, self-reported experience in stock trading and whether subjects were invested in the stock market during the last financial crisis. Reported are coefficients and t-statistics (in parentheses) using robust standard errors. *, **, *** indicates statistical significance at the 10%, 5%, and 1% levels, respectively.

In the pooled data, we find a strong and highly significant effect of our treatment indicator on the subjective success probability of the ambiguous lottery. In particular, those subjects who learned to form expectations in the bust treatment are about 19 percentage points ($p < 0.001$) more pessimistic about the success probability than subjects who learned to form beliefs in the boom treatment (average success probability estimate for boom treatment: 68 %; for bust treatment: 49 %). The finding remains stable and statistically highly significant for each learning environment separately, even though the effect seems to be stronger in the mixed-outcome learning environment. As such, the induced pessimism resulting from distorted belief formation translates to other – independent – investment environments.

Table 4.6: Relation Between Beliefs About Success Probability and Investment

Dep. Variable	*Investment in Ambiguous Asset*					
	Pooled Data	Pooled Data	Domain-specific	Domain-specific	Mixed	Mixed
Success Probability	0.412***	0.409***	0.365***	0.341***	0.470***	0.521***
	(6.45)	(5.70)	(3.88)	(3.42)	(5.47)	(4.83)
Bust		-0.372		-3.846		4.571
		(-0.11)		(-0.93)		(0.82)
Constant	-3.304	-2.985	-5.350	-2.458	2.166	0.00936
	(-0.26)	(-0.23)	(-0.36)	(-0.16)	(0.10)	(0.00)
Observations	377	377	177	177	200	200
R^2	0.146	0.146	0.162	0.166	0.157	0.160

Note: This table examines whether subjects in our experiment act upon their beliefs about the success probability of the ambiguous asset. We report the results of OLS regressions for the whole sample, and for each experiment individually (Experiment 1: Domain-specific; Experiment 2: Mixed). The dependent variable is *Investment Ambiguous*, which captures subjects' invested amount in the ambiguous lottery. *Success Probability* denotes participants' beliefs about the success probability of the ambiguous lottery. *Bust* is an indicator variable that equals 1 if participants were in the bust treatment. Controls include age, gender, statistical skills, self-reported experience in stock trading and whether subjects were invested in the stock market during the last financial crisis. Reported are coefficients and t-statistics (in parentheses) using robust standard errors. *, **, *** indicates statistical significance at the 10%, 5%, and 1% levels, respectively.

In Table 4.6, we test whether differences in subjective expectations regarding the success probability of the ambiguous lottery also translate to changes in risk-taking. In essence, we test whether subjects adhere to a basic economic principle: keeping everything else constant, do subjects increase their investment in an ambiguous asset when their beliefs about the outcome distribution are more optimistic? Our results across all specifications confirm that subjects act upon their beliefs. In other words, the more optimistic they are about the success probability of the ambiguous asset, the more they invest

($p < 0.01$). In addition, in Columns (2), (4), and (6), we include the Bust indicator as an additional control variable to exclude the possibility that our manipulation affects factors unrelated to expectations. Even after including the Bust indicator, the effect of subjective probability estimates on investments remains of similar magnitude and statistical significance. Moreover, we find no additional effect of our manipulation on the investment decision. Effectively, this means while our manipulation does induce pessimism, it does not affect factors unrelated to expectations.

Taken together, our findings suggest that: 1) Learning to form beliefs in adverse market environments induces pessimism caused by systematic errors in the belief updating process. 2) This pessimism translates to lower risk-taking even in independent investment environments when there is room to form beliefs. 3) Pessimism causes agents to assign lower probabilities to more favorable outcomes. 4) Learning in adverse market environments and the resulting errors in the belief updating process do not affect risk preferences.

4.5.4 Who is Most Affected?

In this section, we seek to establish a more profound understanding of how the subjects' forecasting abilities in the first part of the experiments affect their subsequent risk-taking. To investigate this relation, we define the squared deviation of subjects' probability estimate in each round from the objective posterior probability as a measure of forecasting quality. Next, we conduct median splits with respect to this measure to distinguish above-median forecasters from below-median forecasters. To assess the validity of our measure, we compare the number of correct forecasts (defined in the payment scheme by being in the range of 10 % of the objective forecast) between below- and above-median forecasters. Across both experiments, those subjects who are classified as "above-median" have on average three more correct forecasts than those classified as "below-median" ($p < 0.001$, t-test). Moreover, both measures are highly correlated (Pearson correlation of 0.57, $p < 0.001$).

To better understand to what extent the resulting pessimism through learning from adverse market outcomes is a necessary condition for belief-induced changes in risk-taking, we repeat the previous analyses and split by the forecasting ability of our participants. Table 4.7 reports our main finding.

Table 4.7: Risk-Taking Across Macroeconomic Cycles Split by Forecasting Quality

Dependent Variable	*Investment*					
	Pooled Data		**Domain-specific**		**Mixed**	
	Above Median	Below Median	Above Median	Below Median	Above Median	Below Median
Bust	6.126	-1.109	6.424	0.652	3.437	-2.713
	(1.38)	(-0.25)	(0.86)	(0.11)	(0.59)	(-0.41)
Ambiguous	10.94***	-1.448	11.48*	-1.582	10.56*	-2.073
	(2.65)	(-0.33)	(1.92)	(-0.24)	(1.75)	(-0.34)
Bust x Ambiguous	-21.49***	-1.454	-22.15**	-4.501	-19.14**	1.881
	(-3.54)	(-0.23)	(-2.44)	(-0.52)	(-2.25)	(0.19)
Constant	1.238	22.65	1.822	37.77**	5.365	4.365
	(0.10)	(1.58)	(0.11)	(2.09)	(0.29)	(0.20)
Observations	377	376	169	181	208	195
R^2	0.095	0.072	0.139	0.070	0.119	0.114

Note: This table examines subjects' risk-taking across treatments split by above and below median forecasting ability as defined in the text. We report the results of OLS regressions for the whole sample, and for each experiment individually (Experiment 1: Domain-specific; Experiment 2: Mixed). The dependent variable is *Investment*, which denotes participants' invested amount (0 - 100) in the lottery they were assigned to. *Bust* is an indicator variable that equals 1 if participants were in the bust treatment. *Ambiguous* is an indicator variable that equals 1 if participants were asked to invest in the ambiguous lottery, and 0 if they invested in the risky lottery. Controls include age, gender, statistical skills, self-reported experience in stock trading, whether subjects were invested in the stock market during the last financial crisis, and the order of outcomes they experienced in the forecasting task. Reported are coefficients and t-statistics (in parentheses) using robust standard errors. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Interestingly, we find that the previously reported effect is both stronger in absolute terms and in terms of statistical significance but only for participants with above-median forecasting ability. In other words, the risk-taking of those agents who achieve more correct forecasts is stronger affect by the learning environment than the risk-taking of agents who achieve less correct forecasts. While this effect is roughly twice as big as for the full sample, it is also independent of the learning environment and even slightly stronger for the mixed-outcome learning environment.

In a next step, we investigate whether the learning environment affects the estimated success probability of the ambiguous asset differently depending on the forecasting ability. The results are reported in Table 4.8.

Table 4.8: Relation Between Treatment and Probability Estimates Split by Forecasting Quality

Dep. Variable	*Success Probability Estimate of Ambiguous Asset*					
	Pooled Data		**Domain-specific**		**Mixed**	
	Above Median	Below Median	Above Median	Below Median	Above Median	Below Median
Bust	-25.58***	-13.38***	-13.55***	-11.40**	-35.34***	-15.48***
	(-8.20)	(-4.50)	(-3.09)	(-2.51)	(-8.57)	(-3.75)
Constant	57.97***	53.54***	84.00***	50.75**	33.84**	54.92***
	(4.19)	(4.33)	(4.75)	(2.57)	(2.14)	(3.40)
Observations	187	190	85	92	102	98
R^2	0.333	0.194	0.228	0.185	0.516	0.244

Note: This table examines the underlying mechanism of how our treatment variable affects subjects' beliefs about the success probability of the ambiguous lottery split by above and below median forecasting ability as defined in the text. We report the results of OLS regressions for the whole sample, and for each experiment individually (Experiment 1: Domain-specific; Experiment 2: Mixed). The dependent variable is *Success Probability*, which denotes participants' beliefs about the success probability of the ambiguous lottery. *Bust* is an indicator variable that equals 1 if participants were in the bust treatment. Controls include age, gender, statistical skills, self-reported experience in stock trading and whether subjects were invested in the stock market during the last financial crisis. Reported are coefficients and t-statistics (in parentheses) using robust standard errors. *, **, *** indicates statistical significance at the 10%, 5%, and 1% levels, respectively.

Across all specifications, we consistently find that subjects in the bust treatment are significantly more pessimistic in their assessment of the success probability of the ambiguous asset. For the mixed-outcome learning environment, we find that above-median forecasters are even more pessimistic in their probability assessment than below-median forecasters, which is consistent with our previous findings. Across both experiments, above-median forecasters rate the success probability on average 25 percentage points lower if they are in the bust treatment than their peers in the boom treatment. This effect shrinks substantially to only 15 percentage points for below-median forecasters. Similar to previous analyses, we also find that independently of their forecasting ability subjects act upon their beliefs by investing more in the ambiguous asset if they rate the success probability to be higher (see Table C.1 in the Appendix C).

But how is it possible that the risk-taking of the seemingly better performing agents (i.e. the better forecasters) is more affected by the learning environment? One possible explanation could be that our proxy of forecasting ability is related to other factors such as socio-demographic background. Alternatively, our proxy might capture participants' involvement in the experimental task. Effectively, this would suggest that the documented effect is more generalizable outside of the experimental environment but limited by the difficulty of the Bayesian updating task. To test the first explanation, we investigate whether agents with above-median forecasting ability share specific socio-demographic characteristics. The results are reported in Table 4.9.

Overall, neither gender nor age can explain differences in participants' forecasting abilities. In addition, and importantly, we find no treatment differences between both groups. As such, the share of above-median forecasters is rather evenly distributed among our boom and bust treatment. Somehow surprisingly, we find differences in subjects' self-reported statistical skills. However, the sign of the coefficient is rather unexpected as the group of above-median forecasters self-reports on average lower statistical skills, which might hint at overconfidence. Similar findings can be observed for subjects' self-reported experience in stock trading. Taken together, our results – while not conclusive – provide no basis to support the first explanation.

To test whether subjects with above-median forecasting ability are more involved in the experiment, we investigate the time it took to finish the experiment and the strength of the pessimism bias. Interestingly, we find that above-median forecasters spent on average 112 seconds to read the instructions of the forecasting task, while below-median forecasters only

Table 4.9: Socio-Demographic Determinants of Forecasting Ability

Variable	Full sample (N=753)	Above median (N=377)	Below median (N=376)	Difference	p-value
Age	34.72	34.32	34.23	0.09	0.90
	(10.28)	(9.79)	(10.77)		
Female	0.39	0.39	0.40	0.01	0.77
	(0.49)	(0.49)	(0.49)		
Statistical Skills	4.35	4.14	4.56	0.42	< 0.01
	(1.65)	(1.56)	(1.71)		
Experience Stock Trading	3.80	3.29	4.31	1.02	< 0.01
	(1.94)	(1.92)	(1.83)		
Invested Financial Crisis	0.41	0.33	0.50	0.17	< 0.01
	(0.49)	(0.47)	(0.50)		
Bust	0.50	0.47	0.53	0.06	0.11
	(0.50)	(0.50)	(0.50)		

Note: This table shows demographics for our sample split by above- and below-median forecasting ability. Reported are the mean and the standard deviation (in parentheses) for the whole sample (Column 1) and split by median (Column 2 and 3). Column 4 presents randomization checks. Differences in mean were tested using rank-sum tests, or χ^2-tests for binary variables. The p-value is reported in Column 5. *Female* is an indicator variable that equals 1 if a participant is female. *Statistical skills* denotes participants' self-assed statistical skills on a 7-point Likert scale. *Experience in stock trading* is the self-reported experience participants have in stock trading, assessed by a 7-point Likert scale. *Invested financial crisis* is an indicator that equals 1 if participants were invested in the stock market during the last financial crisis. *Bust* is an indicator variable that equals 1 if participants were in the bust treatment.

spent roughly 86 seconds ($p < 0.05$). Additionally, the overall time to finish the experiment is roughly 580 seconds for above-median forecasters, and about 553 seconds for below-median forecasters ($p < 0.10$). The difference is largely driven by the additional time above-median forecasters spent to read the instructions more carefully. Besides investigating the time subjects take to read the instructions, we also look at the number of basic errors subjects make during the forecasting task. We define a basic error as a situation in which a participant updates his prior belief in the wrong direction (i.e. reporting a lower posterior probability after observing a high outcome signal or reporting a higher posterior probability after observing a low outcome signal). While above-median forecasters make basic errors in roughly 11 % of their forecasts, below-median forecasters make such errors in roughly 30 % of their forecast ($p < 0.001$, two-sided t-test). In other words, below-median forecasters make a basic error in approximately every third forecast, even though a comprehension question following the instructions exactly tested this relation (see Appendix C). Taken together, the lower time below median-forecasters take to read the instructions paired with the large frequency of basic errors they make, hint at a significantly lower involvement in the experimental task.

We also investigate the strength of the pessimism bias in both groups. The results are reported in Table 4.10. As expected the bias is less pronounced for subjects with above-median forecasting ability (who also have more correct forecasts). However, and more importantly, the pessimism bias still persists and is statistically highly significant. Across all experiments, we consistently find that above-median forecasters exhibit a 34 % less pronounced pessimism bias. Nevertheless, these findings show that even the above-median forecasts suffer from a pessimism bias which subsequently translates to lower risk-taking. One indication of this might be that the above-median forecasters are more involved in the overall experiment and in particular the forecasting task given the additional time they need to finish the experiment. The higher involvement is also reflected in the high explanatory power for this particular subgroup as seen by the relatively high R^2 of roughly 0.70 compared to the rather low R^2 of around 0.10 for the subgroup of below-median forecasters. Given the strength of the pessimism bias even in the group of more sophisticated forecasters paired with the higher involvement of the aforementioned group in our experiment, we believe that the effect of different learning environments on risk-taking might be even more pronounced in the real economy.

Table 4.10: Pessimism Bias Split by Forecasting Quality

Dep. Variable	*Probability Estimate (Subjective Posterior)*					
	Pooled Data		**Domain-specific**		**Mixed**	
	Above Median	Below Median	Above Median	Below Median	Above Median	Below Median
Bust	-4.529***	-6.813***	-4.261***	-7.247***	-4.997***	-5.661***
	(-6.13)	(-4.54)	(-3.72)	(-3.11)	(-5.18)	(-2.86)
Objective Posterior	0.671***	0.133***	0.641***	0.165***	0.693***	0.107***
	(48.14)	(7.46)	(34.13)	(6.33)	(35.46)	(4.37)
Constant	20.92***	58.92***	14.88***	66.86***	27.49***	50.78***
	(6.75)	(9.62)	(3.22)	(6.82)	(6.38)	(6.60)
Observations	6032	6016	2704	2896	3328	3120
R^2	0.69	0.10	0.68	0.08	0.70	0.12

Note: This table reports the results of OLS regressions on how subjective posterior beliefs about the distribution of the lottery depend on the treatment split by above and below median forecasting ability as defined in the text. We report the results of OLS regressions for the whole sample, and for each experiment individually (Experiment 1: Domain-specific; Experiment 2: Mixed). The dependent variable *Probability Estimate* is the subjective posterior belief that the asset is paying from the good distribution. Independent variables include the *Bust* dummy, an indicator variable that equals 1 if participants were in the bust treatment and zero otherwise, as well as *Objective Posterior*, which is the correct Bayesian probability that the stock is good, given the information seen by the participant up to trial *t* in the learning block. Controls include age, gender, statistical skills, self-reported experience in stock trading, whether subjects were invested in the stock market during the last financial crisis, and the order of outcomes they experienced in the forecasting task. Reported are coefficients and t-statistics (in parentheses) using robust standard errors. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

4.5.5 Boundaries and External Validity

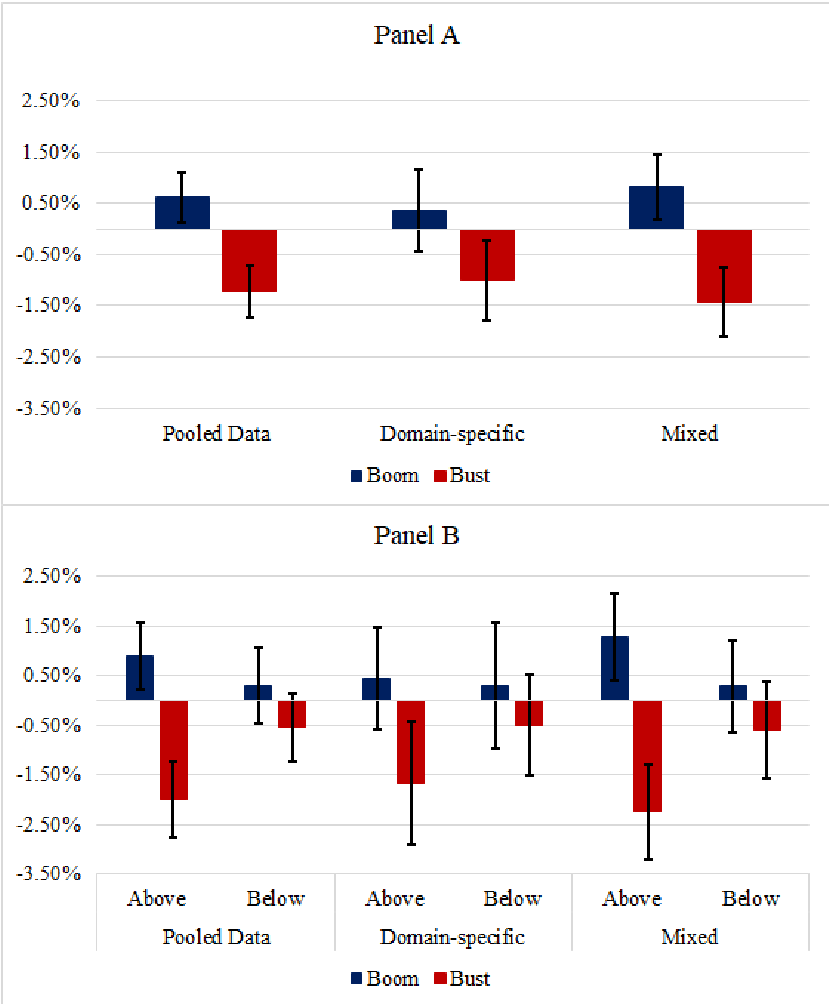
To test both the external validity and the boundaries of the induced pessimism resulting from asymmetric learning in boom and bust markets, we analyze subjects' responses to two additional set of questions, which deal with expectations outside the experimental setting. The first question tests to which extent the induced pessimism translates to expectations in the real economy. We gave subjects the at the time current level of the Dow Jones Industrial Average, and asked them to provide a 6-month return forecast on a balanced 12-point Likert scale (see Appendix C). The second set of questions tests to which degree the induced pessimism from the underlying learning environment permeates to different contexts. As a measure of dispositional optimism/pessimism across different life situations, we included a 10-item general Life Orientation Test borrowed from Scheier et al. (1994), which is frequently used in psychological research (see Appendix C). Results for the Dow Jones return estimate are reported in Figure 4.4 (Panel A for entire sample and Panel B split by forecast quality).

For the Dow Jones return estimates, we consistently find across all learning environments that subjects in the bust treatment are significantly more pessimistic in their return expectations. More strikingly, subjects in the bust treatment provide not only lower return estimates but also negative return estimates, while those in the boom treatment provide positive return estimates on average. Moreover, the effect seems to be stronger in absolute magnitude for the negative return estimates, consistent with a pessimism bias. When split by forecast quality, we observe that the effect is again mainly driven by subjects with above-median forecasting ability. As such, even while above-median forecasters show a less pronounced pessimism bias overall (see previous section), their pessimism still translates to lower return expectations in the real economy and thus outside the experimental setting.

For the below-median forecasters however, we do not find significant differences even though they also suffer from a pessimism bias. This fact paired with a potentially lower involvement may explain why we cannot observe differences in risk-taking in the ambiguous lottery between treatments for this subgroup. It remains to stress, that even in such a simple and short-learning environment as in our experiment, we are able to systematically manipulate return expectations for real world market indices.

Finally, we investigate the boundaries of how the pessimism induced by adverse learning environments affects subjects overall psychological well-being. Results are reported in Figure 4.5 (Panel A for entire sample and Panel B split by forecast quality).

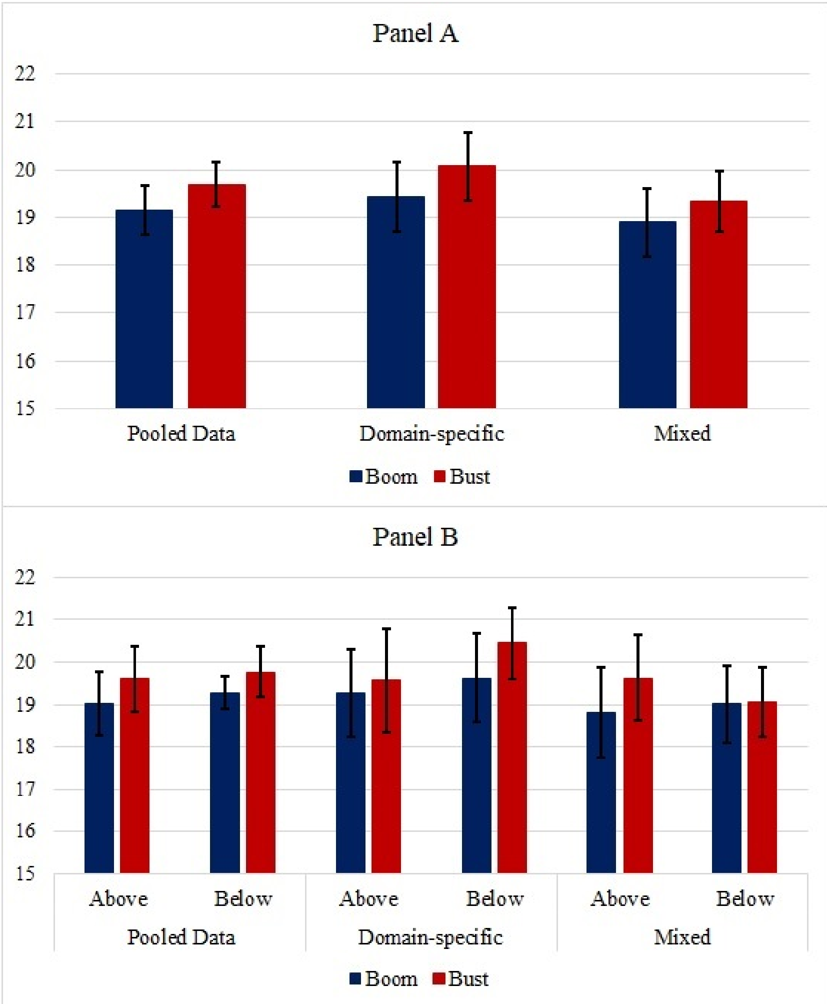
Figure 4.4: Dow Jones Estimates



Note: This figure displays subjects' self-reported return expectations of the Dow Jones Industrial average. Dow Jones return expectations were assessed on a 12-point Likert scale Results are displayed separately for subjects across treatments (boom / bust) and across experiments. Panel A displays return expectations for the entire sample and split by experiment. Panel B displays return expectations split by above- and below-forecasting ability and by experiment. Displayed are 95% confidence intervals.

Across all experiments and splits we do not find any significant difference in dispositional optimism/pessimism depending on whether subjects were in the boom or bust treatment. Taken together, our results suggest that the environment in which subjects learn strongly affects their return expectations for even unrelated financial investments, but does not affect subjects' inherent psychological traits such as neuroticism, anxiety, self-mastery, or self-esteem as assessed by the Life Orientation Test.

Figure 4.5: Life Orientation Test



Note: This figure displays subjects' answers to a general life orientation test. The life orientation test Scheier et al. (1994) is a 10-item inventory where subjects rate statements on a 7-point Likert scale. Displayed is the cumulated score separated by treatment (boom / bust) and by experiment. Panel A displays the cumulated score for the entire sample and split by experiment. Panel B displays the cumulated score split by above- and below-forecasting ability and by experiment. Displayed are 95 % confidence intervals.

4.6 Conclusion

In this paper, we present experimental evidence on an alternative channel to countercyclical risk aversion for time-varying risk-taking. While rational expectations models introduce modifications in the representative agent's utility, we test whether systematic deviations from rational expectations can cause the same observed investment pattern without assuming time-varying degrees of risk aversion.

We place subjects in a learning environment which resembles key characteristics of boom and bust markets and measure their risk-taking under risk (i.e. known probabilities) or under uncertainty (i.e. unknown probabilities) in an independent investment task. Subjects who learned to form beliefs from adverse outcomes (resembling a bust market) take significantly less risk in investments under uncertainty. However, we do not find any significant difference in their level of risk aversion.

Overall, the mechanism described in our experiment implies that agents may form pro-cyclical return expectations, i.e. they are more optimistic in boom markets and more pessimistic in recessions. These results are consistent with recent survey evidence on investors' return expectations. While traditional models (i.e. rational expectations models) assume that agents are fully aware of the implied counter-cyclical nature of the equity premium (Nagel and Xu, 2019), these surveys find that – if anything – investors form rather pro-cyclical expectations.

Additionally, the investigated systematic deviation from rational expectations can produce similar self-reinforcing processes as countercyclical risk aversion. The countercyclical nature of risk preferences implies that investors are more risk averse during recessions, which leads investors to reduce their equity share. This process then generates additional downward momentum for prices. Yet, similar dynamics can also be generated assuming time-varying changes in expectations. If bust markets systematically induce pessimistic expectations about future returns for a substantial subset of investors, this may reduce the aggregate share invested in risky assets of an economy, which in turn generates downward pressure on prices due to excess supply.

Chapter 5

Can Agents Add and Subtract When Forming Beliefs? *

5.1 Introduction

Probabilistic beliefs are essential to decision-making under risk in various economic domains, including investments in financial markets, purchasing insurance, attaining education, or when searching for employment. Standard models assume that individuals update their prior beliefs according to Bayes' Theorem. Besides the prescription of how individuals should form posterior probabilities, Bayes' Theorem has an implicit, fundamental rule of how subjects should incorporate information *signals of opposite direction*. In the usual case of updating about two states of the world from independent binomial signals, two unequal signals should cancel out. Thus, taken together they should not affect prior beliefs. Importantly, this relation is independent of whether individuals' prior beliefs are consistent with Bayes.

To illustrate this idea, imagine you think about visiting a restaurant which recently opened in your city. Before making a reservation, you call two of your friends who know the restaurant. Suppose, both of them recommend the new restaurant, making you rather optimistic about its quality. Yet, since the restaurant is quite expensive, you decide to call two more friends. Assume, the first one did not like the restaurant, whereas the second did like

* Authors: Pascal Kieren, Jan Müller-Dethard, and Martin Weber. All authors are at the University of Mannheim. We gratefully acknowledge funding by the German Research Foundation (DFG grant WE993/15-1).

it. Would you still be just as optimistic as you were after the first two calls? In other words, are *two* recommendations just as good as *three* recommendations and *one* critique, as prescribed by Bayes' Theorem?

In this article, we ask whether individuals follow this simple, counting-based rule when updating their beliefs. To test this, we create an environment in which subjects repeatedly observe binary signals to learn about an underlying state of the world. While such a binary decision-making problem appears to presents a specific, commonly used and simplified setting in experimental research, it applies to many every-day decision problems (e.g. are we in a good or bad stock market regime, should I take an umbrella for the walk or not, or as in our example above, is the restaurant good or bad?).

Throughout this paper, we refer to signals that are in line with the true underlying state of the world as *confirming* signals and otherwise as *disconfirming* signals. We exogenously manipulate the number of subsequent confirming signals that gets interrupted by a single disconfirming signal. This setup allows us to test (i) how subjects update their priors after a *disconfirming* signal conditional on the number of previously observed confirming signals; and (ii) the extent to which they revise their priors after the disconfirming signal is followed by another confirming signal (i.e. *corrected*). In both cases, Bayes' Rule makes a simple, yet important prediction: An agent should reduce (increase) his prior after a disconfirming (confirming) signal by the same magnitude than he increased (reduced) it after the previous confirming (disconfirming) signal.

To implement this framework, we conduct three bookbag-and-poker-chip experiments in the spirit of Grether (1980) with 1800 participants. All experiments follow the same basic design. Over the course of six periods, we provide subjects with information signals about a risky asset which can either draw from a "good distribution" or from a "bad distribution". Both distributions are binary with a high outcome of +5 and a low outcome of -5. In the good distribution, the higher payoff occurs with 70% probability while the lower payoff occurs with 30% probability. In the bad distribution, the probabilities are reversed, i.e. the lower payoff occurs with 70% probability while the higher payoff occurs with 30% probability. To create situations which are consistent with our framework, we use a stratified sample of price paths. More precisely, we examine six price paths for the good distribution and six price paths for the bad distribution. In each of the six periods of a price path, subjects subsequently observe payoffs of the risky asset. After each payoff, we ask them to provide a probability estimate that the risky asset draws from the good distribution and how confident they are about their estimate.

In Experiment 2 and 3 we run variations of our baseline experiment to test the robustness and underlying drivers of our findings. In Experiment 2, we change the informational content of the positive signal (i.e. the diagnosticity). In Experiment 3, we reduce the uncertainty about the underlying distribution by providing subjects with the full outcome history in advance. For comparability, the price paths we use in both variations remain identical to the baseline experiment.

To detect whether subjects follow a simple, counting-based heuristic when updating their beliefs after a disconfirming signal, we compare the change in probability estimate after a disconfirming signal to the change in probability estimate after a confirming signal which is directly observed prior to the disconfirming signal. The same logic applies to the case when the disconfirming signal is reverted (i.e. corrected).

Our findings can be summarized as follows. First, we consistently find that subjects strongly overreact whenever a sequence of confirming signals is interrupted by a single disconfirming signal. Across all experiments, subjects update their prior beliefs on average by 3.54 % immediately before observing the disconfirming signal, whereas they update their prior beliefs on average by 15.38 % after the subsequent disconfirming signal. In relative terms, subjects update their priors by 334 % too much after a disconfirming signal, thereby acting as if one single disconfirming signal would carry the weight of up to three confirming signals.

Second, we find that this overreaction is almost entirely corrected once subjects observe another confirming signal following the disconfirming signal. More precisely, after observing a confirming signal directly following the disconfirming signal, they update their prior beliefs again by 13.65 %, compared to their initial overreaction of 15.38 %. In other words, subjects almost completely correct their initial overreaction if the disconfirming signal gets reverted.

Third, we find that both the overreaction and the subsequent correction do not critically depend on subjects having extreme priors. Even with a diagnosticity of only 60 %, two subsequent confirming signals are sufficient to observe a pronounced overreaction after a disconfirming signal. In such a setting not only the experimentally observed subjective priors, but also the objective Bayesian probabilities are low with on average 72 % and 69 %, respectively.

Fourth, the observed overreaction after a disconfirming signal becomes stronger the more confirming signals individuals previously encountered. Even though – in absolute terms – the observed overreaction should become smaller as subjective priors converge to one, we find that a single disconfirming signal can completely revert up to five confirming signals the later it occurs. This implies that – in contrast to the Bayesian prediction – signals are not invariant to the order in which they occur. In other words, observing one single disconfirming signal followed by five confirming signals is different compared to observing five confirming signals that are followed by a single disconfirming signal. Whereas subjects mostly correct their strong overreaction if they can, the violation of the counting heuristic is most severe when subjects have no opportunity to collect further information.

Motivated by previous work showing that agents react most strongly to unexpected events, we finally investigate whether the observed overreaction still exists if subjects (i) have little uncertainty about the underlying distribution and (ii) know in which period the disconfirming signal will occur. However, even under these circumstances subjects still strongly overreact after a disconfirming signal.

Overall, our findings suggest that when observing a disconfirming signal after a sequence of confirming signals subjects fail to follow the simple counting heuristic implied by Bayes' Theorem. Instead of reverting one previous signal, they revert up to five signals. In other words, they strongly overreact. Interestingly however, this is not the case, if a disconfirming signal is immediately reverted. Then, subjects appear to follow the counting heuristic and fully correct their prior overreaction. Referring to our introductory restaurant example, a single critique would cancel out both prior recommendations, while another recommendation following the critique would be considered as two recommendations.

Our paper contributes to several strands of literature. First, we contribute to the various studies that document biases and heuristics in probabilistic reasoning (for an overview see Camerer, 1987, 1995; Benjamin, 2019). A common finding, by and large is that people update too little, with three exceptions as noted by Benjamin (2019): (i) People overinfer from signals if the diagnosticity is low, (ii) people may overinfer when signals go in the same direction of the priors (i.e. prior-biased updating), and (iii) people may overinfer when priors are extreme and signals go in the opposite direction of the priors (due to base-rate neglect). Especially, (ii) and (iii) push in opposite directions which makes it important to understand when one or the other dominates. Our study suggests that whenever subjects violate the simple counting heuristic implied by Bayes' Theorem, individuals generally overreact to signals of opposite direction of their priors. A violation occurs whenever a sequence of signals that go in the same direction is interrupted by a signal of opposite direction. Importantly, we find that this overreaction is independent of subjects having extreme priors and requires only a sequence of two signals that go in the same direction. Conversely, we find that subjects generally underinfer in situations in which they cannot or do not violate the counting heuristic. This is either because there are (i) only signals of same direction, or (ii) positive and negative signals alternate.

Second, our study also contributes to the recent literature on tipping points. In psychology, a tipping point describes "the point at which people begin to perceive noise as signal" (O'Brien and Klein, 2017, p. 161). In other words, a tipping point defines the first point when people infer that a pattern is no longer an anomaly and thus believe that one state of the world is more likely to be the true state (O'Brien, 2019). So far, research has uncovered two robust findings: tipping points are asymmetric across valence (i.e. people reach conclusions faster for negative events than for positive events) and asymmetric across time (i.e. people predict slower tipping points than they actually express). Our findings suggest that tipping points regarding probabilistic beliefs about an underlying state of the world (i.e. one of two possible probability distributions) are symmetric across domains. One possible reason for this difference is both, the signal structure and the underlying stochastic process. Whereas our study employs objective and randomly distributed signals with a predefined underlying stochastic process, previous studies employ more realistic (and thus more subjective) signals with no clear underlying stochastic process. This distinction is in line with the discussion on the use of neutral versus more realistic quantities in the experimental literature on information processing (see Eil and Rao, 2011). Interestingly, our findings also suggest that individuals are quick to revise their priors once they observe a disconfirming signal, which might be important for the formation of tipping points and the persistence of subsequent beliefs.

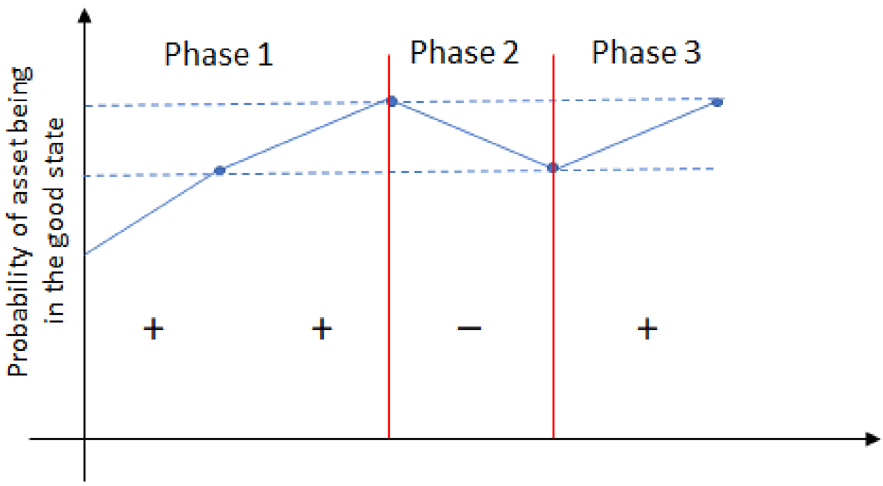
Finally, we also contribute to the literature on over- and underreactions to unexpected news in financial markets (Bondt and Thaler, 1985; Barberis et al., 1998; Daniel et al., 1998; Hong and Stein, 1999). Our results suggest that the violation of a simple counting heuristic implied by Bayes' Rule presents a potential mechanism underlying over- and underreactions. In situations in which agents observe a sequence of signals that go in the same direction (e.g. consensus favorable earnings forecasts) agents initially underreact. If such a sequence is interrupted by a single signal that goes in the opposite direction (e.g. an unfavorable earnings surprise), they strongly overreact and partly neglect previous signals. Interestingly, our findings also suggest that the strength of the overreaction only partly depends on the underlying signal being unexpected. In other words, the violation of a simple counting heuristic in probabilistic belief updating does not crucially depend on the fact that agents are surprised.

Our paper is structured as follows. In Section 5.2, we present an empirical framework, briefly review the existing literature and state our hypotheses. In Section 5.3, we describe the experimental design and summary statistics. Finally, in Section 5.4 we discuss our results and conclude in Section 5.5.

5.2 Empirical Framework and Hypotheses

In this section, we describe the framework which serves as a basis for our hypotheses and the later empirical analyses and then relate the existing literature to our established framework. Suppose there is an agent who wants to learn about the quality of a risky asset. The risky asset can either be in a good or bad state. Over a number of periods, the agent may receive good (+) or bad (−) signals from which he can learn about the quality of the risky asset. This framework of how the agent's beliefs about the asset being in the good state should evolve can best be illustrated using the following graph.

Figure 5.1: Empirical Framework



Note: The figure illustrates the empirical framework of this study. We examine subjects' belief updating behavior over three phases: Phase 1 describes a sequence of signals that go in the same direction (i.e. confirming an underlying distribution). Phase 2 describes a situation in which a sequence of previously observed same-directional signals is interrupted by a single signal of opposite direction (i.e. disconfirming signal). Finally, Phase 3 defines the situation when a disconfirming signal is immediately reverted (i.e. correction). The blue dots present the objective probabilities (i.e. the beliefs according to Bayes' Theorem) that the asset pays from the good distribution given the sequence of signals.

Figure 5.1 illustrates three phases of how Bayesian beliefs evolve over a sequence of four outcomes. The first phase ("confirming signals") resembles a sequence of same-directed signals. A signal which (i) confirms the underlying distribution and (ii) follows another same-directed signal will be referred to as a *confirming signal*. Thus, if a signal is to be referred as a confirming signal, an agent must have observed at least two signals. The second phase ("disruptive signal") defines the situation when a sequence of confirming signals (phase 1) is disrupted by a signal of opposite direction than the previously observed signal. A signal which disrupts a sequence of same-directed signals will be referred to as a *disconfirming signal*. The third phase ("correction") resembles the case when a previously observed disconfirming signal is reverted. A signal which follows on a disconfirming signal and has the opposite direction than the previously-observed disconfirming signal is referred to as a *correction*.

In our framework with binary information signals, an agent should update his prior beliefs according to the following formula:

$$P_t^{Bayes} = P(G|\delta_t)^{Bayes} = \frac{\theta^{\delta_t}}{\theta^{\delta_t} + (1-\theta)^{\delta_t}}, \quad \delta_t = g_t - b_t \tag{5.1}$$

where P_t^{Bayes} is the posterior probability that the risky asset pays from the good distribution (G) and θ refers to the diagnosticity of the good signal. The number of good signals observed until period t is referred to as g_t, while the number of bad signals observed until period t is referred to as b_t.

Applying the formula to our described framework from Figure 5.1 provides several implications on how agents should update their beliefs. Overall, note that the Bayesian agent in our setting is indifferent regarding the order of the signals, since only the difference delta t is relevant. This feature of the described framework has implications which are especially relevant for the second and the third phase in Figure 5.1. For the second phase this implies that an agent should reduce the probability estimate after a disconfirming signal by the same magnitude than he increased it after the previous confirming signal. In other words, a Bayesian agent would report the same probability estimate than he did two signals ago. As such he simply cancels the previously observed confirming signal. Referring to the framework in Figure 5.1, the Bayesian agent would state the same probability estimate as he did after observing the first positive signal. For the third phase, a similar logic applies. In particular, after observing a correction (i.e. the reversion of the disconfirming signal) agents should also only cancel the previously observed disconfirming signal and should again, end up with the same probability estimate as they did two signals ago. In both scenarios (disruption and correction), a Bayesian agent would follow a counting heuristic which means that one positive and one negative signal simply cancel out.

In contrast, agents in the first phase cannot rely on a simple counting heuristic in determining the precise probability estimate. That means after observing two same-directional signals, the counting heuristic does not provide any insight by how much they need to adjust the prior estimate. In other words, to state the correct magnitude of the change in probability estimate, the agent needs to know Bayes' Rule.

Based on the established framework, we formulate the following hypotheses:[1]

Hypothesis H1: Disruption (Phase 2)

After observing a disconfirming signal, an agent should reduce his prior probability estimate by the same magnitude than he increased it after the previous confirming signal.

Hypothesis H2: Correction (Phase 3)

After a previous disconfirming signal got reverted, an agent should cancel the previously observed disconfirming signal and end up with the same probability estimate as he did two signals ago.

[1] The hypotheses are formulated for the good distribution. In the bad distribution, subjects should adjust their priors in the opposite direction.

It is important to stress that our framework and the later experimental design do not crucially depend on agents being Bayesian. Instead, it is sufficient for agents to know that two directionally inconsistent signals cancel each other out. In other words, for the basic updating rule we are testing, it is not essential that agents state the correct *absolute* Bayes estimate. We are rather interested in the *changes* in probability estimates after subjects incorporate new signals into their prior beliefs.

As discussed, Bayes Theorem provides clear and testable predictions on how individuals should revise their beliefs after a sequence of confirming signals is interrupted by a single disconfirming signal as well as after its subsequent reversal (i.e. correction). While this is perfect normative advice, the literature on probabilistic reasoning has identified various situations in which individuals systematically deviate from Bayes and either over- or underinfer. Using bookbag-and-poker-chip experiments, some studies find underinference when a new signal confirms the prior hypothesis and no or only very little revision of beliefs when a new signal disconfirms the prior hypothesis, consistent with prior-biased inference (Pitz et al., 1967; Geller and Pitz, 1968; Pitz, 1969). In contrast to this, DuCharme and Peterson (1968) observe in experiments with normally distributed signals overinference in response to a disconfirming signal. However, Eil and Rao (2011) as well as Möbius et al. (2014) find no evidence for prior-biased inference at all. Recently, Charness and Dave (2017) establish a conceptual framework which combines both under- and overinference and test it experimentally. They find prior-biased inference. In particular, they observe overinference after a confirming signal in updating problems with equal prior probabilities of the states and high diagnosticity of 70%. However, and opposing to Charness and Dave (2017), Pitz et al. (1967), find for the identical level of diagnosticity underinference after a confirming signal. In brief, while there are several studies showing that individuals deviate from Bayes, the evidence *in which way* and *when* they deviate is mixed and apparently inconsistent.

5.3 Experimental Design

One-thousand-eight-hundred-and-seven individuals (1159 males, 648 females, mean age 34 years, 10 years standard deviation) were recruited from Amazon Mechanical Turk (MTurk) to participate in three online experiments. MTurk advanced to a widely used and accepted recruiting platform for economic experiments. Not only does it offer a larger and more diverse subject pool as compared to lab studies (which frequently rely on students), but it also provides a response quality similar to that of other subject pools (Buhrmester et al., 2011; Goodman et al., 2013).

An environment to study the role of disconfirming information signals requires (i) a sequential set-up with room for subjective belief formation, (ii) control over Bayesian beliefs, (iii) variation in the number of confirming signals prior to a disconfirming signal, and (iv) an incentive-compatible belief elicitation. Our design accommodates all of these features.

5.3.1 Baseline Design

To study the role of disconfirming information signals, we provide subjects with information about a risky asset. In all of our experiments, the risky asset has an initial value of 50 which either increases or decreases over the course of six periods depending on the asset's payoffs. The payoffs are either drawn from a "good distribution" or from a "bad distribution". Both distributions are binary with a high outcome of +5 and a low outcome of -5. In the good distribution, the higher payoff occurs with 70 % probability while the lower payoff occurs with 30 % probability. In the bad distribution, the probabilities are reversed, i.e. the lower payoff occurs with 70 % probability while the higher payoff occurs with 30 % probability.

Since we only focus on a single disconfirming signal within six periods, we differentiate between six possible price paths per distribution. These price paths resemble our treatments. The first treatment dimension depicts the underlying distribution and therefore the domain (good or bad), while the second treatment dimension depicts the period in which the disconfirming signal occurs (from period one to period six). Table 5.1 provides an overview of all twelve treatments.

Table 5.1: Overview of Treatments

			Good Distribution			
Treatment	1	2	3	4	5	6
G-1	−	+	+	+	+	+
G-2	+	−	+	+	+	+
G-3	+	+	−	+	+	+
G-4	+	+	+	−	+	+
G-5	+	+	+	+	−	+
G-6	+	+	+	+	+	−
			Bad Distribution			
Treatment	1	2	3	4	5	6
B-1	+	−	−	−	−	−
B-2	−	+	−	−	−	−
B-3	−	−	+	−	−	−
B-4	−	−	−	+	−	−
B-5	−	−	−	−	+	−
B-6	−	−	−	−	−	+

Note: This table provides an overview of all treatments in our experiments. Overall, there are twelve treatments, six in the good distribution and six in the bad distribution, defined by the period in which the disruptive signal occurs. The "−" sign represents a negative (bad) signal and the "+" sign a positive (good) signal.

For example, in treatment G-3, the risky asset pays from the good distribution and the disconfirming signal appears in period three after two confirming signals (i.e. the sequence would be: positive, positive, negative, positive, ... signal). A key feature of our design is that we shift the single disconfirming signal between a sequence of six signals. That allows us to test how subjects update their beliefs after observing a single disruptive, disconfirming signal conditional on the number of previously observed confirming signals. Additionally, the design makes it possible to investigate how subjects update their beliefs after the disconfirming signal is reverted.

Across all experiments, subjects make forecasting decisions in six consecutive periods. At the beginning of the experiment, the computer randomly determines the distribution of the risky asset (which can be good or bad) and the period in which the disconfirming signal will occur (which can be from one to six). In each of the six rounds, subjects observe a payoff of the risky asset. After each round, we ask them to provide a probability estimate that the risky asset draws from the good distribution and how confident they are about their estimate. To keep the focus on the forecasting task and to not test their memory performance, we display the prior outcomes in a price-line-chart next to the questions. To ensure that subjects have a sufficient understanding of the forecasting task, they had to correctly answer four comprehension questions before they could continue (see Appendix D.1).

The experiment concluded with a brief survey about subjects' socio-economic background, self-assessed statistic skills, as well as a measure of risk preferences and financial literacy adopted from Kuhnen (2015). Subjects' belief elicitation was incentivized. Participants were paid a participation fee and a variable fee based on the accuracy of the probability estimates provided. Specifically, they received 25 cents for each probability estimate within 10 % ($+/-5\%$) of the objective Bayesian value. Across all studies, it took participants approximately 7 minutes to complete the experiment and participants earned $1.50 on average.

5.3.2 Experimental Variations

We conducted two variations of our baseline experiment, referred to as *Reduced Diagnosticity* and *Reduced Uncertainty*. The two additional experiments are designed to identify whether the belief updating after a disconfirming signal depends on (i) the diagnosticity of the signal (i.e. its informational content), (ii) subjects' uncertainty about the distribution (i.e. whether the asset turns out to be good or bad), and (iii) whether subjects do not anticipate the disconfirming signal (i.e. are surprised about the disruption of the sequence of confirming signals).

Experiment Reduced Diagnosticity: In the experiment *Reduced Diagnosticity* we change the informational content that subjects can infer from a positive signal. This means, we change the probability of the higher outcome in the good distribution from 70 % to 60 % and of the lower outcome from 30

% to 40 %, respectively. In the bad distribution, we change the probability of the lower outcome from 70 % to 60 % and of the higher outcome from 30 % to 40 %, respectively. On the one hand, we expected to observe - as Bayes' Theorem implies - lower (higher) absolute levels of probability estimates in the good (bad) distribution given the reduced diagnosticity of signals. On the other hand, we expect to observe no impact of diagnosticity on the fundamental counting rule we are testing. Within our empirical framework, the increase (decrease) in posterior probability after a confirming signal in the good (bad) distribution should remain exactly as much as the decrease (increase) in posterior probability after a subsequent disconfirming signal, irrespective of how informative the signal is.

Experiment Reduced Uncertainty: In the experiment *Reduced Uncertainty* we combine aspects (ii) and (iii) from above. To do so, we change the previously framed forward-looking updating task to a backward-looking updating task. In detail, subjects in the baseline experiment are asked to make a forecasting decision without knowing the future outcome history. In the *Reduced Uncertainty* experiment, we show subjects the full outcome history beforehand. Then, we ask them to provide probability estimates period by period as in the baseline experiment for exactly the same outcome history they have seen in advance. Importantly, subjects were still incentivized to provide probability forecasts which only incorporate the information subjects had in each period. In other words, the objective Bayesian probabilities are identical to the baseline experiment. By showing subjects the entire outcome history beforehand, we already eliminate most of the uncertainty regarding the underlying distribution and any of the potential surprise related to the period in which the disruptive signal occurs. Additionally, before the first period, we directly ask subjects two questions: (i) we ask them to count the number of positive and the number of negative payoffs in the outcome history and (ii) we ask them to state the period in which the disconfirming signal occurs.

5.3.3 Demographics

Table 5.2 presents summary statistics for all our three experiments. Overall 1807 subjects participated in our studies, with an average age of 33.79 years in Experiment 1 (33.59 years in Experiment 2, and 35.01 years in Experiment 3). Thirty-five percent (forty-one percent, thirty-two percent) were female. Subjects reported average statistical skills of 4.46 out of 7 (4.42, 4.42) and their level of risk aversion, measured by how much of an endowment of 10,000 they are willing to invest risky in a broad equity index, is as follows. Subjects invest on average 4,470 (4,420, 5,000) in the risky asset. Across all experiments subjects report medium financial literacy. In particular, they make 1.73 (1.70, 1.70) out of three possible basic errors.

Table 5.2: Summary Statistics on Subjects

Variable	**Experiment 1** Baseline (N=601)	**Experiment 2** Reduced Diagnosticity (N=602)	**Experiment 3** Reduced Uncertainty (N=604)
Age	33.79	33.59	35.01
	(9.89)	(9.17)	(9.83)
Female	0.35	0.41	0.32
	(0.48)	(0.49)	(0.47)
Statistical Skills (1-7)	4.46	4.42	4.42
	(1.64)	(1.64)	(1.68)
Risk Preferences	44.7%	44.2%	50.0%
	(2.94)	(2.89)	(2.98)
Financial Literacy (1-3)	1.73	1.70	1.70
	(0.93)	(0.91)	(0.93)

Note: This table shows summary statistics for our experimental data. Reported are the mean and the standard deviation (in parentheses) for each experiment individually. *Female* is an indicator variable that equals 1 if a participant is female. *Statistical skills* denotes participants' self-assessed statistical skills on a 7-point Likert scale. *Risk preferences* are elicited by asking subjects to split an endowment between a risky and a risk-free asset (reported is the fraction invested risky). *Financial literacy* was assessed by asking subjects to identify the correct formula for calculating the expected value of the portfolio they selected. Through multiple choice answers, participants could make three basic errors (reported is the number of basic errors).

5.4 Results

5.4.1 Main Results

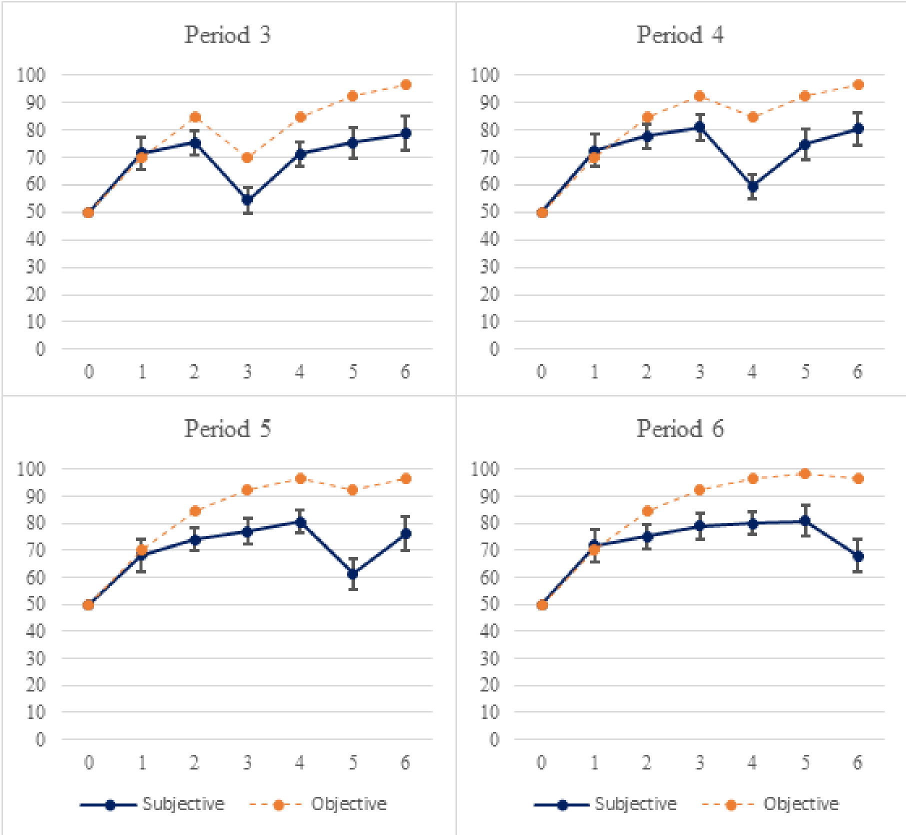
In this section, we first present results of our baseline experiment of how individuals update their beliefs after disconfirming signals as well as of how they revise their probability estimates after a correction. Then, we test the robustness of our findings with respect to the diagnosticity of the information signals and finally examine how the reduction of uncertainty with respect to the underlying distribution affects subjects' updating behavior.

Baseline Results

Figure 5.2 and Figure 5.3 present subjects' average updating tendency over all periods for each treatment G-3 to G-6 of our baseline experiment. Figure 5.2 shows the results of those treatments in which the underlying distribution is *good* and Figure 5.3 shows the results of those treatments in which the underlying distribution is *bad*.

To be consistent with our framework in Section 5.2, we focus our analysis on the treatments in which subjects observe at least two subsequent same-directional signals before a disconfirming signal occurs. This is the case for our treatments G-3, G-4, G-5, and G-6 (B-3, B-4, B-5, and B-6). We will analyze the results of treatments G-1 and G-2 (B-1 and B-2) in a separate section at the end of this chapter. From Figure 5.2, we observe that subjects in the good distribution increase their prior beliefs by 6.44 % on average after a confirming signal, whereas they decrease their prior beliefs by 18.63 % on average after observing a disconfirming signal. In the bad distribution, the findings look similar as seen in Figure 5.3. Subjects decrease their prior beliefs by 5.38 % on average after a confirming signal, while they increase their prior beliefs by 16.94 % on average after observing a disconfirming signal. In relative terms, this means that subjects in the good distribution update their prior beliefs after a disconfirming signal with a magnitude that is approximately three times as large as if they update after a confirming signal. This ratio is more or less

Figure 5.2: Subjects' Average Updating Behavior in the Good Distribution – Experiment 1

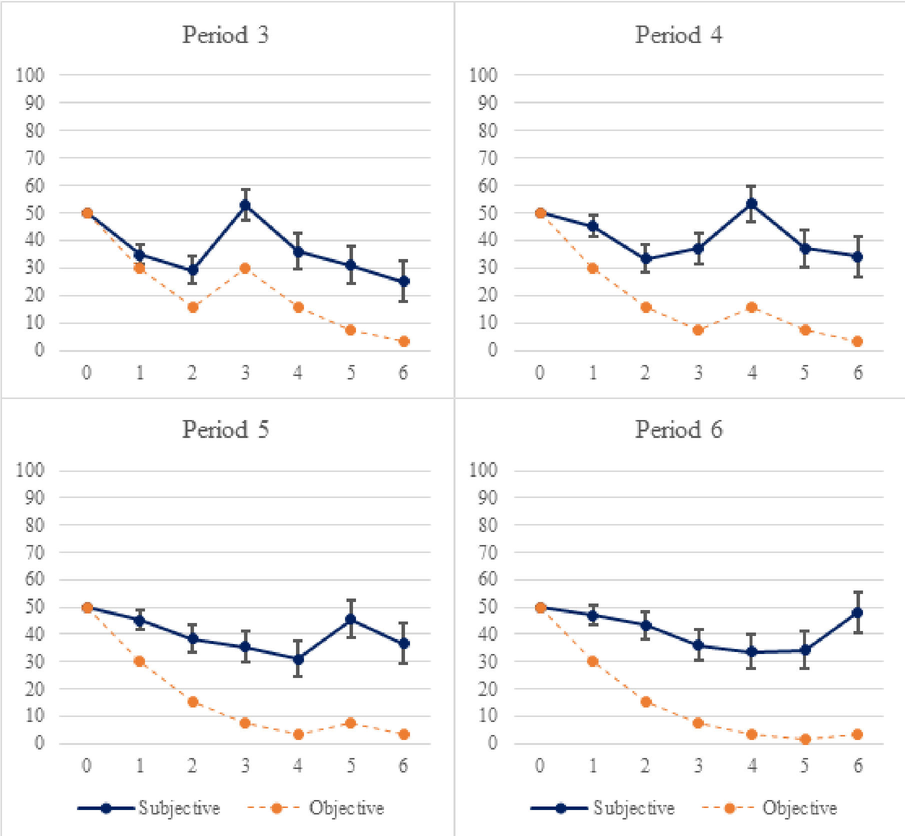


Note: The figure displays subjects' average probability estimates over six consecutive periods in the good distribution for each treatment G-3 to G-6 individually. The dashed line shows the objective Bayesian posterior probabilities and the solid line shows subjects' average probability estimates. Displayed are 95% confidence intervals.

independent of the distribution, albeit a little bit stronger in the bad distribution. Participants update their beliefs after a disconfirming signal as if they failed to incorporate three previously observed confirming signals. In other words, subjects strongly overreact after a disconfirming signal. In particular, it appears that a disconfirming signal destroys up to five prior confirming signals.

Next, we investigate how individuals update their prior beliefs after a disconfirming signal gets reverted. In particular, we examine whether and to what extent subjects correct the observed overreaction after a disconfirming signal. We find that subjects in the good distribution increase their probability estimate on average by 17.11 %. Similarly, in the bad distribution, subjects decrease their probability estimates on average by 14.16 %. In essence,

Figure 5.3: Subjects' Average Updating Behavior in the Bad Distribution – Experiment 1



Note: The figure displays subjects' average probability estimates over six consecutive periods in the bad distribution for each treatment G-3 to G-6 individually. The dashed line shows the objective Bayesian posterior probabilities and the solid line shows subjects' average probability estimates. Displayed are 95% confidence intervals.

the previously observed overreaction after a disconfirming signal is almost entirely corrected. This finding holds independent of the distribution.

From these descriptive statistics alone, it becomes already evident that subjects fail to follow a simple counting heuristic when they incorporate inconsistent signals in their beliefs. In other words, they do not adhere to the simple updating rule in which they count the difference between positive and negative signals. Instead, they strongly overreact after a disconfirming signal. Interestingly however, this is not the case, if an inconsistent (i.e. disconfirming) signal is reverted. Then, subjects appear to follow the counting heuristic implied by Bayes' Rule and fully correct their prior overreaction.

Besides the descriptive analysis, we also run regressions, in which we can control for the objective posterior probability. To investigate how individuals update their prior beliefs both in response to disconfirming signals and subsequent confirming signals (i.e. the correction of the disconfirming signal), we estimate the following model[2]:

$$\Delta p_{i,t} = \beta_1 \Delta ObjectivePrior_{i,t} + \beta_2 Disconfirm_{i,t} + \beta_3 Correction_{i,t} + \varepsilon_{i,t}, \quad (5.2)$$

where $\Delta p_{i,t}$ is the difference in subjects' probability estimates between two subsequent periods and $\Delta ObjectivePrior_{i,t}$ is the difference in the objective Bayesian probability between two subsequent periods. Finally, $Disconfirm_{i,t}$ and $Correction_{i,t}$ are two indicator variables which equal one if subject i observes a disconfirming signal or a correction in period t, respectively. In the above specification we can test both for Bayesian behavior and in which way individuals depart from it. If subjects were perfect Bayesian, we would expect that $\widehat{\beta_1} = 1$, and $\widehat{\beta_2} = \widehat{\beta_3} = 0$. In other words, subjects always update their prior beliefs according to Bayes' Rule, while neither a disconfirming signal (which disrupts a sequence of confirming signals) nor a subsequent correction would explain any additional variation. Conversely, $\widehat{\beta_1} < (>) 1, \widehat{\beta_2} < (>) 0$, and $\widehat{\beta_3} < (>) 0$ would signal underinference (overinference) to subsequent confirming signals, to disconfirming signals, and to corrections, respectively. The results are reported in Table 5.3.

The findings support our previously drawn conclusions. Even after controlling for the objective posterior, we find an economically strong and statistically highly significant overreaction after a disconfirming signal. Additionally, we find that the initial overreaction is almost entirely corrected if the disconfirming signal is reverted. While in the bad distribution, both effects are of similar magnitude and thus cancel out, we find a slightly asymmetric effect in the good distribution. Whereas the correction is of similar strength as in the bad distribution, the overreaction is stronger. As such the overreaction in the good distribution is not entirely corrected.

[2] Since we investigate changes in subjective probability estimates, we estimate the model without constant to be consistent with the theoretical benchmark. However, results are qualitatively similar if we estimate the model on levels or with constant. For the ease of interpretation, we report the specification without constant.

Table 5.3: Updating Behavior After Disconfirming Signal and Correction – Experiment 1

Dependent Variable	*Change in Posterior Probability Estimate*			
	Good Distribution		**Bad Distribution**	
Change in Bayes	0.770***	0.377***	0.718***	0.384***
	(14.64)	(8.02)	(13.03)	(7.51)
Disconfirm		−15.94***		12.38***
		(−9.15)		(7.37)
Correction		11.57***		−11.05***
		(7.36)		(−6.76)
Observations	1782	1782	1824	1824
R^2	0.138	0.218	0.097	0.142

Note: This table reports the results of four OLS regressions on how subjects update their posterior beliefs after a disconfirming signal and a correction in the baseline experiment. We report the results of OLS regressions for each distribution individually (good and bad distribution). The dependent variable in the regression model, *Change in Posterior Probability Estimate*, is the change in subjective posterior beliefs that the asset is paying from the good distribution between period *t* and period *t-1*. Independent variables include the *Disconfirm* dummy, an indicator variable that equals 1 if participants observe a disconfirming signal and zero otherwise, the Correction dummy, an indicator variable that equals 1 if a disconfirming signal is subsequently reverted, as well as *Change in Bayes*, which is the change in the correct Bayesian probability that the stock is good between period *t* and period *t-1*. Reported are coefficients and t-statistics (in parentheses) using robust standard errors. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Next, we examine how our model in which we explicitly control for a disconfirming signal and a subsequent correction performs compared to the standard Bayes model. When comparing the explanatory power of the two models, we find that the standard Bayesian model explains roughly 14 % (10 %) in the good (bad) distribution, while our model explains roughly 22 % (14 %). Irrespective of the distribution, our model explains roughly 50 % more of the variation of subjects' probability estimates than the standard Bayesian model.

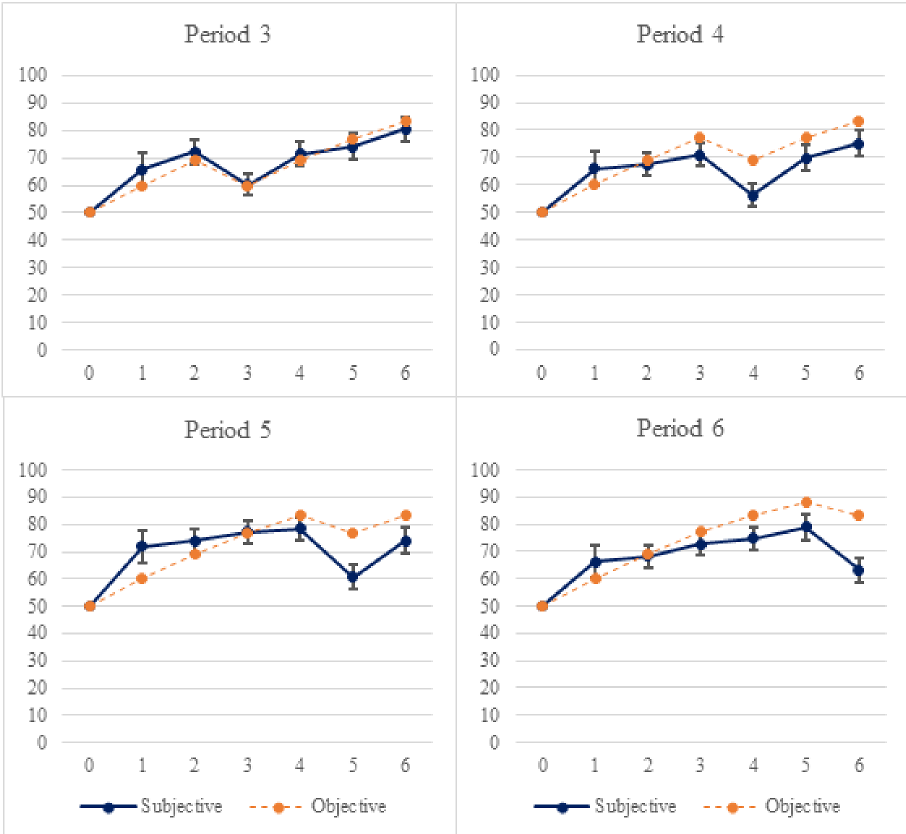
Moreover, Table 5.3 implies that subjects generally underinfer which is consistent with several studies on Bayesian updating (see Benjamin, 2019). Interestingly, our results suggest that the observed underinference is mostly driven by subsequent confirming signals. When differentiating between the good and the bad distribution, we find that the observed underinference is stronger when subjects update their beliefs from a sequence of confirming bad signals than when updating their beliefs from a sequence of confirming good signals. This finding is consistent with the recently identified good news-bad news effect reported by Eil and Rao (2011) as well as Möbius et al. (2014). However, for our main finding, it remains to stress that we do not find such an asymmetric effect across domains.

Reducing the Diagnosticity of Information Signals

In this section, we report results of our second experiment in which we vary the informational content of the signals. Like in our baseline experiment, Figure 5.4 and Figure 5.5 present subjects' general updating behavior in the *good* and the *bad* distribution, respectively, over all periods for each treatment G-3 to G-6.

Overall, the findings look very similar to our baseline experiment. In particular, we find that subjects in the good distribution increase their prior beliefs by 7.15 % on average after a confirming signal, whereas they decrease their prior beliefs by 14.81 % on average after observing a disconfirming signal. In the bad distribution, the findings look similar. Subjects decrease their

Figure 5.4: Subjects' Average Updating Behavior in the Good Distribution – Experiment 2

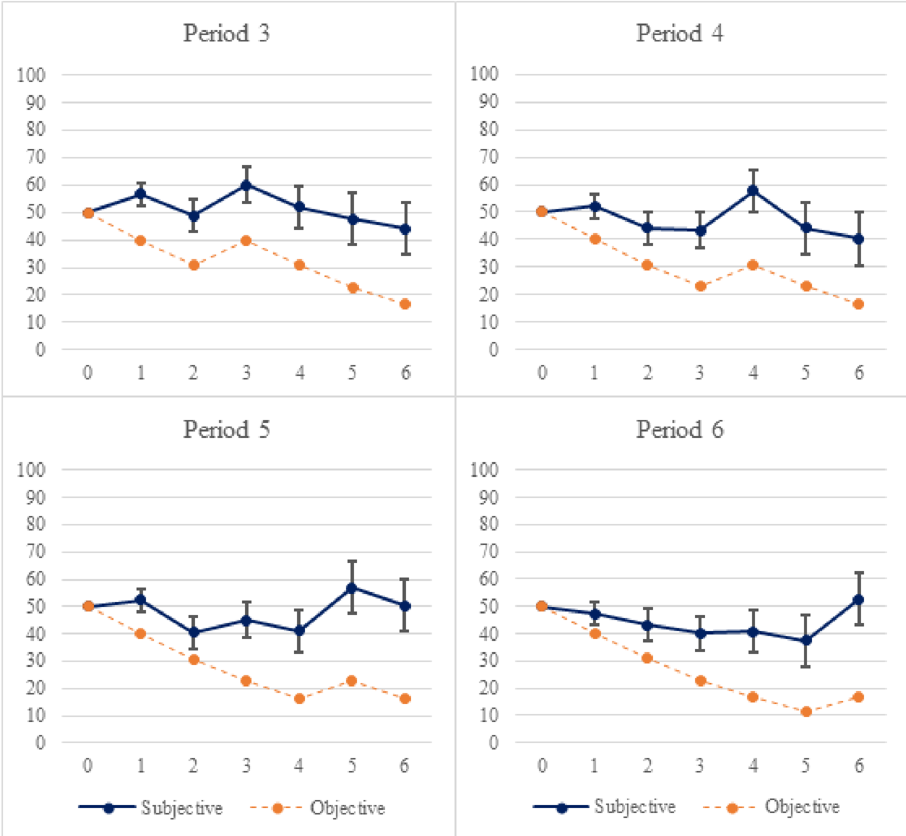


Note: The figure displays subjects' average probability estimates over six consecutive periods in the good distribution for each treatment G-3 to G-6 individually. The dashed line shows the objective Bayesian posterior probabilities and the solid line shows subjects' average probability estimates. Displayed are 95% confidence intervals.

prior beliefs by 3.65 % on average after a confirming signal, while they increase their prior beliefs by 7.15 % on average after observing a disconfirming signal. Like in our baseline experiment, subjects update their beliefs after a disconfirming signal as if they failed to incorporate up to three previously observed confirming signals. Despite the lower diagnosticity in the second experiment, the observed overreaction after a disconfirming signal persists.

This finding even holds after controlling for the objective Bayesian probability as to be seen in Table 5.4. The observed overreaction after a disconfirming signal remains economically large and statistically significant. In comparison to the results from our baseline experiment, the magnitude with which subjects update their prior after a disconfirming signal is smaller. However, this is to be expected since the updating magnitude strongly correlates with

Figure 5.5: Subjects' Average Updating Behavior in the Bad Distribution – Experiment 2



Note: The figure displays subjects' average probability estimates over six consecutive periods in the bad distribution for each treatment G-3 to G-6 individually. The dashed line shows the objective Bayesian posterior probabilities and the solid line shows subjects' average probability estimates. Displayed are 95% confidence intervals.

the diagnosticity. Consistent with our previous findings, we find that subjects correct their priors after a disconfirming signal is reverted. Interestingly, we find that in contrast to the baseline experiment, subjects seem to not sufficiently correct their previous overreaction which can especially be seen in the bad distribution. Overall, even in a setting with lower diagnosticity subjects still do not follow the simple counting heuristic when observing a disconfirming signal. Instead, they show a strong overreaction which they partly correct subsequently.

Table 5.4: Updating Behavior After Disconfirming Signal and Correction – Experiment 2

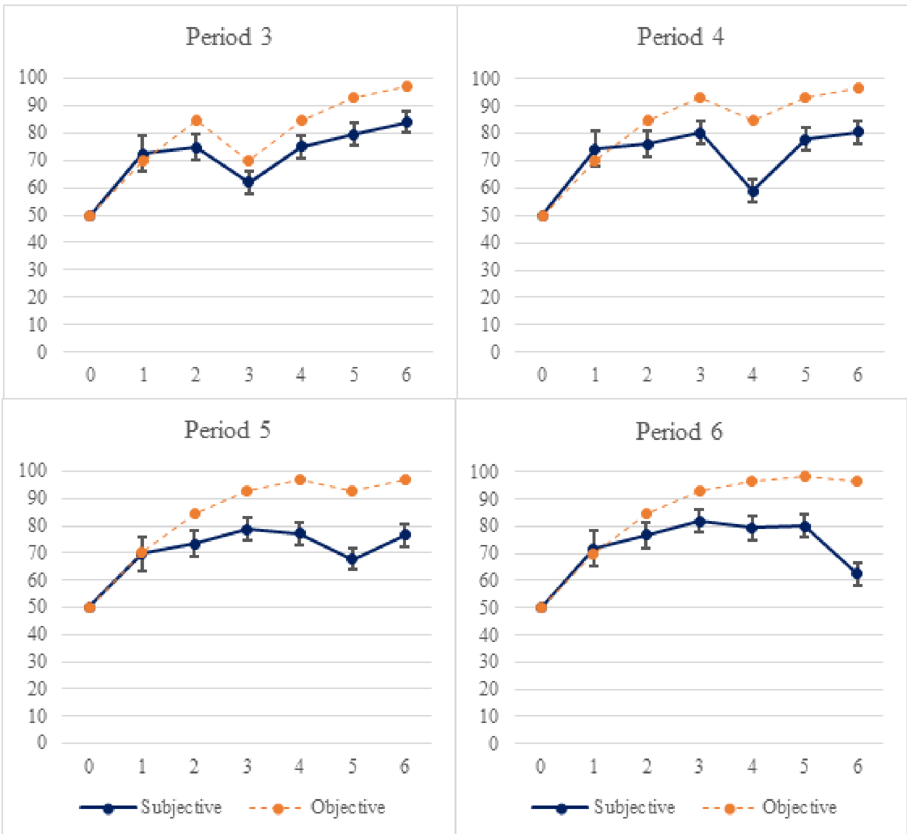
Dependent Variable	*Change in Posterior Probability Estimate*			
	Good Distribution		**Bad Distribution**	
Change in Bayes	0.860***	0.430***	0.877***	0.524***
	(16.47)	(9.08)	(13.98)	(8.84)
Disconfirm		−11.53***		10.06***
		(−8.82)		(6.74)
Correction		9.355***		−6.649***
		(6.57)		(−4.18)
Observations	1872	1872	1740	1740
R^2	0.112	0.169	0.087	0.116

Note: This table reports the results of four OLS regressions on how subjects update their posterior beliefs after a disconfirming signal and a correction in Experiment 2 with lower diagnosticity than in the baseline experiment. We report the results of OLS regressions for each distribution individually (good and bad distribution). The dependent variable in the regression model, *Change in Posterior Probability Estimate*, is the change in subjective posterior beliefs that the asset is paying from the good distribution between period *t* and period *t-1*. Independent variables include the *Disconfirm dummy*, an indicator variable that equals 1 if participants observe a disconfirming signal and zero otherwise, the Correction dummy, an indicator variable that equals 1 if a disconfirming signal is subsequently reverted, as well as *Change in Bayes*, which is the change in the correct Bayesian probability that the stock is good between period *t* and period *t-1*. Reported are coefficients and t-statistics (in parentheses) using robust standard errors. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Reducing the Uncertainty About the Underlying Distribution

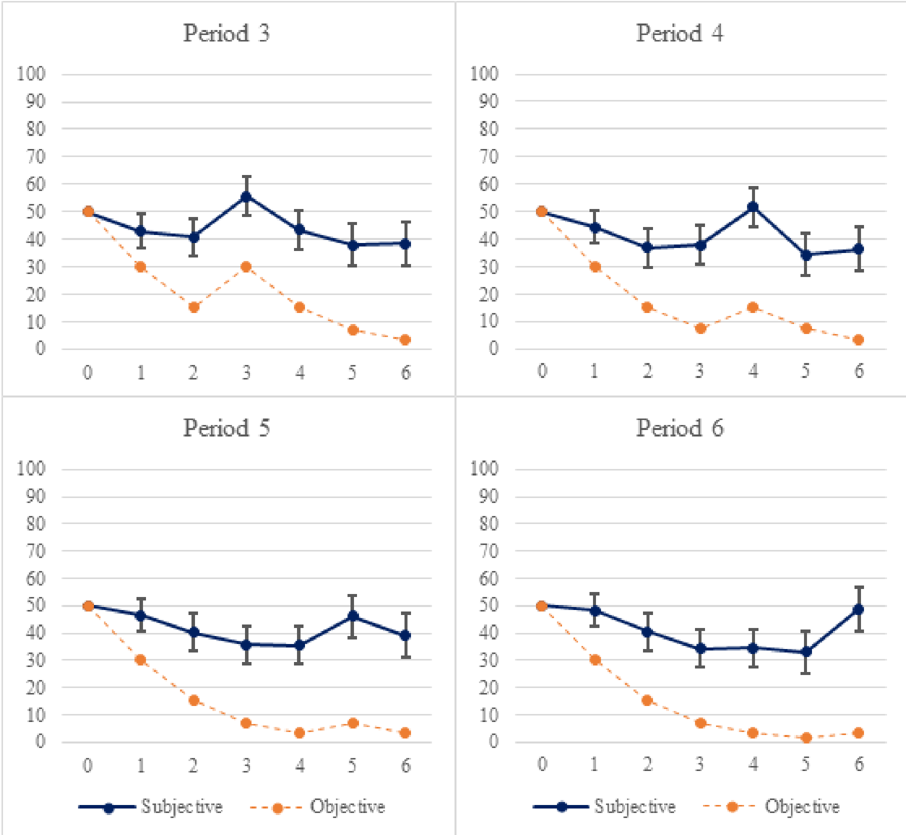
In the following, we discuss the results of our third experiment in which we reduce subjects' uncertainty about the underlying distribution. This variation of the design allows us to exclude the possibility that subjects falsely infer trends or price reversal. Additionally, we control for the possibility that subjects do not anticipate (i.e. are surprised by) the disconfirming signal as they observe the full outcome history in advance. The results on individuals' updating behavior are reported in Figure 5.6 and Figure 5.7. Again, Figure 5.6 shows the results of those treatments in which the underlying distribution is *good* and Figure 5.7 shows the results of those treatments in which the underlying distribution is *bad*.

Figure 5.6: Subjects' Average Updating Behavior in the Good Distribution – Experiment 3



Note: The figure displays subjects' average probability estimates over six consecutive periods in the good distribution for each treatment G-3 to G-6 individually. The dashed line shows the objective Bayesian posterior probabilities and the solid line shows subjects' average probability estimates. Displayed are 95% confidence intervals.

Figure 5.7: Subjects' Average Updating Behavior in the Bad Distribution – Experiment 3



Note: The figure displays subjects' average probability estimates over six consecutive periods in the bad distribution for each treatment G-3 to G-6 individually. The dashed line shows the objective Bayesian posterior probabilities and the solid line shows subjects' average probability estimates. Displayed are 95% confidence intervals.

We find that both, overreaction after a disconfirming signal and subsequent correction even persist in a setting in which the uncertainty about the underlying distribution is dramatically reduced. In particular, the Bayesian probability of the asset being in the good distribution is 96.74 %. As such after subjects observe the full outcome history there should be barely any uncertainty left about the distribution. Besides almost no uncertainty about the underlying distribution, there is also no uncertainty about the period in which the disconfirming signal will occur. First, the graphical representation of the full outcome history in the form of a price-line chart is known to subjects and makes the period in which the disconfirming signal occurs easily identifiable. Second, we also explicitly ask participants to state the period in which the disconfirming signal occurs prior to the forecasting task. As such our design should eliminate any potential surprise subjects may experience when observing a disconfirming signal. In the light of the still persistent overreaction, we can confidentially rule out that surprise effects or uncertainty about the underlying distribution drive the results. Moreover, we can also exclude that subjects overreact after a disconfirming signal because they potentially anticipate a new trend, given that they know that a disconfirming signal will subsequently be reverted.

We run the same regression as previously to control for the objective Bayesian posterior probability, while also investigating potential differences to the baseline experiment. The results are reported in Table 5.5.

Table 5.5: Updating Behavior After Disconfirming Signal and Correction – Experiment 3

Dependent Variable	*Change in Posterior Probability Estimate*			
	Good Distribution		**Bad Distribution**	
Change in Bayes	0.603***	0.294***	0.666***	0.362***
	(12.89)	(6.87)	(11.68)	(7.03)
Disconfirm		−11.77***		9.559***
		(−6.41)		(6.10)
Correction		9.978***		−11.03***
		(7.53)		(−6.77)
Observations	1884	1884	1740	1740
R^2	0.088	0.135	0.086	0.122

Note: This table reports the results of four OLS regressions on how subjects update their posterior beliefs after a disconfirming signal and a correction in Experiment 3. We report the results of OLS regressions for each distribution individually (good and bad distribution). The dependent variable in the regression model, *Change in Posterior Probability Estimate*, is the change in subjective posterior beliefs that the asset is paying from the good distribution between period *t* and period *t-1*. Independent variables include the *Disconfirm* dummy, an indicator variable that equals 1 if participants observe a disconfirming signal and zero otherwise, the Correction dummy, an indicator variable that equals 1 if a disconfirming signal is subsequently reverted, as well as *Change in Bayes*, which is the change in the correct Bayesian probability that the stock is good between period *t* and period *t-1*. Reported are coefficients and t-statistics (in parentheses) using robust standard errors. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

A direct comparability is given as Bayes' probabilities are identical across treatments in the baseline and the reduced uncertainty experiment. First, we can confirm all prior findings. Subjects strongly overreact after a disconfirming signal and subsequently correct the overreaction. Second, when comparing the effect sizes between the two experiments, we find that the overreaction as well as the subsequent correction are slightly more pronounced in the baseline treatment. Even though the reduced uncertainty experiment was designed to significantly decrease the overreaction resulting from disconfirming signals, the effect is still economically strong and statistically significant.

Additional Treatments G-1 and G-2 (B-1 and B-2)

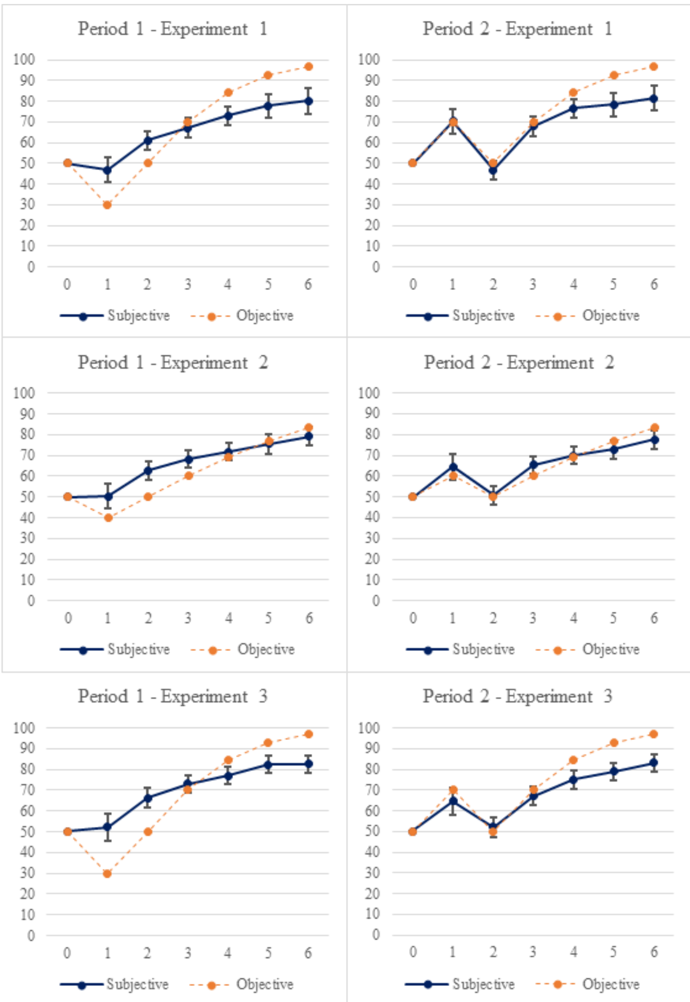
Finally, we analyze the results of treatments G-1 and G-2 (B-1 and B-2) for which – per definition – our empirical framework does not apply. In these treatments, the single opposite-directional signal occurs either directly in the first period or in the second period. As such these treatments describe price paths for which the pre-requisite for Phase 1 of our framework (i.e. at least two confirming signals prior to the disconfirming signal) is not fulfilled. Nevertheless, they allow us to analyze how subjects update their beliefs (i) in situations without prior outcome history (G-1 and B-1) and (ii) in situations with exclusively alternating signals (G-2 and B-2).

Figure 5.8 reports the results for the *good* distribution split by experiment. Figure 5.9 reports the results for the *bad* distribution split by experiment. Across all experiments, we find that subjects do not significantly update their beliefs downwards if the first signal is bad.[3] In contrast to that, subjects significantly update their beliefs upwards if the first signal is good. Their first probability estimate is almost identical to the objective Bayesian probability and this finding holds for both, the two experiments with high diagnosticity (70 %) and the experiment with low diagnosticity (60 %). In period 2,

[3] We follow the terminology used in the empirical framework section and also refer to a bad signal in the first period drawn from an asset with a good distribution as a disconfirming signal, even though subjects cannot know at this point in time that the signal disconfirms the true underlying distribution. The same logic applies to a good signal in the first period drawn from the good distribution which we refer to as a confirming signal.

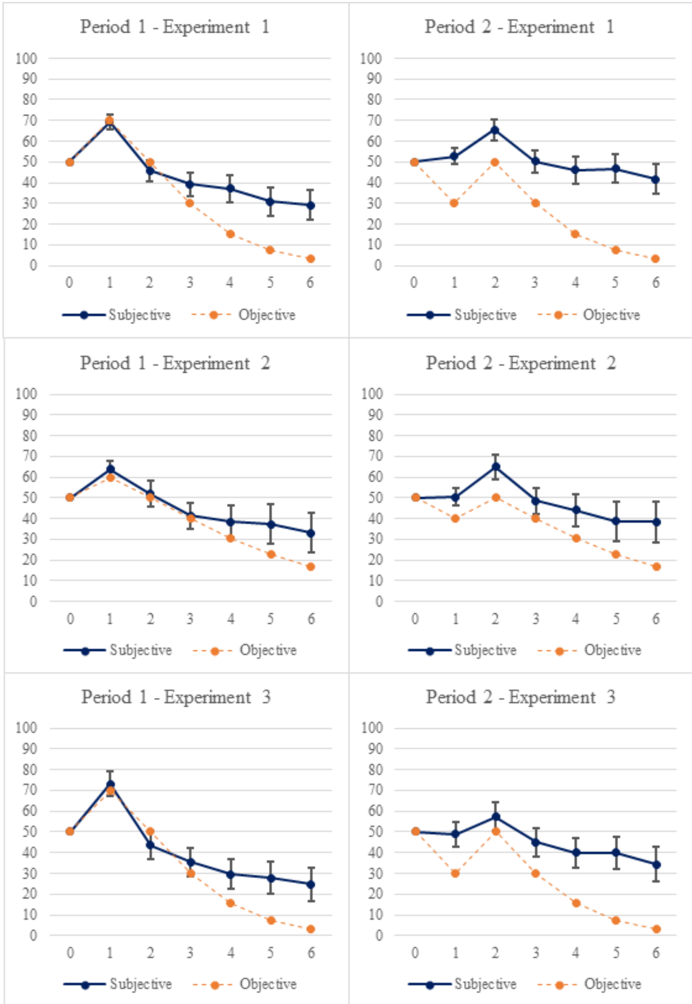
when the bad signal of period 1 is reverted, subjects state probability estimates significantly above the objective probability of 50 %, while when the good signal of period 1 is reverted, subjects are almost perfect Bayesian. In other words, subjects in the B-1 treatment almost perfectly adhere to the investigated counting rule implied by Bayes' Theorem, while subjects in the G-1 treatment clearly violate this rule. In particular, they seem to violate this rule because they ignored or were averse to adjust their beliefs downwards following the first bad signal.

Figure 5.8: Subjects' Average Updating Behavior in the Good Distribution – Treatments G-1 and G-2



Note: The figure displays subjects' average probability estimates over six consecutive periods in the good distribution for treatments G-1 and G-2 of experiment 1, 2, and 3. The dashed line shows the objective Bayesian posterior probabilities and the solid line shows subjects' average probability estimates. Displayed are 95% confidence intervals.

Figure 5.9: Subjects' Average Updating Behavior in the Bad Distribution – Treatments G-1 and G-2



Note: The figure displays subjects' average probability estimates over six consecutive periods in the bad distribution for treatments G-1 and G-2 of experiment 1, 2, and 3. The dashed line shows the objective Bayesian posterior probabilities and the solid line shows subjects' average probability estimates. Displayed are 95% confidence intervals.

This pattern is mirrored when looking at the treatments G-2 and B-2. In these treatments, the signals alternate up until period 3. Subjects, who observe first a good, second a bad, and then again a good signal, are almost perfect Bayesian. Across all experiments, they follow the counting rule and increase their probability estimate after the good signal in period 3 as much as they decreased it after the bad signal in period 2 which in turn they previously increased exactly as much as after the good signal in period 1. In contrast to that, subjects who first observe a bad, second a good, and then again a bad signal do only partly follow the counting rule. Like subjects in the G-1 treatment, they do not significantly adjust the probability estimate downwards if the first signal is bad, but correctly – as implied by the counting rule – decrease their probability estimate in period 3 by the amount by which they previously increased it in period 2. This robust pattern can be found across all experiments.

Taken together, we can complement our findings from treatments G-3 to G-6 (B-3 to B-6) as follows: We find that subjects adhere to the counting rule implied by Bayes' Theorem in situations with no prior sequence of same-directional signals and in situations with exclusively alternating signals. Interestingly however, subjects seem to have problems following this rule right at the beginning of the updating task, when the first signal is bad. In these cases, they act as if they ignore the bad signal and consequently update too much after the subsequent good signal.

5.4.2 Signal Ordering

One aspect of the counting heuristic we have not discussed so far is that Equation 5.1 of the established framework also implies that a Bayesian is indifferent regarding the order in which outcomes occur. In other words, observing a disconfirming signal followed by five subsequent confirming signals should lead to the same posterior probability as first observing five subsequent confirming signals followed by a disconfirming signal. Since our experimental design explicitly varies the round in which the single disconfirming signal occurs, we can directly test this relation. To do so, we estimate the following model:

$$P_{i,6} = \beta_0 + \beta_1 D_{i \mid R=2} + \beta_2 D_{i \mid R=3} + \beta_3 D_{i \mid R=4} + \beta_4 D_{i \mid R=5} + \beta_5 D_{i \mid R=6} + \varepsilon_{i,t}, \quad (5.3)$$

where $P_{i,6}$ is the subjective posterior in round 6, and $D_{i \mid R=t}$ are indicator variables denoting the round in which participants encountered the disconfirming signal (with round 1 being the baseline category). Note that the Bayesian posterior in our setting is the same for each treatment and only depends on the underlying distribution (good or bad) and the underlying diagnosticity. To accommodate this feature, we estimate the model separately for each distribution and split by diagnosticity of the signal. Results are reported in Table 5.6.

Table 5.6: Outcome Ordering

Dependent Variable	*Posterior Probability Estimate in Period 6*			
	Experiment 1 & 3		**Experiment 2**	
	Good Distribution	Bad Distribution	Good Distribution	Bad Distribution
Disconfirm Round 2	0.912	11.45**	−1.755	5.194
	(0.39)	(2.52)	(−0.53)	(0.77)
Disconfirm Round 3	−0.374	5.306	1.224	11.09*
	(−0.16)	(1.36)	(0.41)	(1.74)
Disconfirm Round 4	−1.070	8.198**	−4.059	7.177
	(−0.46)	(2.17)	(−1.21)	(1.21)
Disconfirm Round 5	−5.043**	10.63***	−5.145	17.34***
	(−2.00)	(2.68)	(−1.61)	(2.76)
Disconfirm Round 6	−16.09***	21.24***	−16.09***	19.67***
	(−5.15)	(5.10)	(−4.33)	(3.05)
Constant	80.45***	26.16***	78.25***	32.10***
	(45.94)	(9.72)	(34.38)	(6.69)
Observations	611	594	312	290
R^2	0.094	0.046	0.101	0.049

Note: This table reports the results of OLS regressions on how subjects updating behavior after a disconfirming signal and correction depends on their prior beliefs. We report the results of OLS regressions for each experiment (Experiment 1 and 3 pooled) and distribution (good and bad distribution) individually. The dependent variable in the regression model, *Posterior Probability Estimate in Period 6*, is the absolute subjective posterior belief that the asset is paying from the good distribution in period 6. Independent variables include *Condition t* dummies which are indicator variables for each period *t*. Reported are coefficients and t-statistics (in parentheses) using robust standard errors. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

We find that the round in which the disconfirming signal occurs plays an important role in how individuals form their posterior beliefs. In particular, the later the disconfirming signal occurs, the stronger the overreaction which ultimately leads to a lower final posterior after round 6. This result holds independent of the underlying distribution and is of similar magnitude across different diagnosticities. One potential driver of this further inconsistency is that individuals generally overreact after disconfirming signals, which is mostly corrected after subsequently observing another confirming signal. However, if subjects observe the disconfirming signal in the final period (where the objective prior in the good distribution is as high as 96.74 %!) subjects can no longer correct their strong overreaction, causing them to be substantially more pessimistic (or optimistic if the underlying distribution is the bad one) about the underlying distribution than they should be. This relation can be especially seen by the considerably higher coefficients of the disconfirming dummy for round 6.

Overall, this result highlights once more the fact that individuals consistently violate the counting heuristic after they encounter disconfirming signals. However, whereas they mostly correct their strong overreaction if they can, the violation is most severe when subjects have no opportunity to collect further information.

5.4.3 Robustness Checks

In this section we will replicate our main analyses on different subsamples to validate its robustness against extreme outliers or individuals who are inattentive and as such more likely to suffer from a bias in probabilistic reasoning. Besides validating the robustness of our main finding, such an analysis might also provide valuable insights into which subgroup is most likely to violate the counting heuristic.

In particular, we conduct splits regarding (i) extreme outliers; (ii) "speeders"; and (iii) below median forecasters. Extreme outliers are individuals whose subjective priors largely deviate from the Bayesian benchmark. Following the classification of Enke and Graeber (2019), we define extreme outliers as individuals who report a subjective posterior $p_s < 25\%$ ($> 75\%$) when the Bayesian posterior is $p_B > 75\%$ ($< 25\%$). Speeders are defined as subjects who are in the bottom decile of the response time distribution. Finally, we also investigate whether the here documented effect is only driven

Table 5.7: Forecasting Ability and Extreme Outliers

Panel A: Extreme Outliers				
Dependent Variable	*Change in Posterior Probability Estimate*			
	Good Distribution		**Bad Distribution**	
	No Outlier	Outlier	No Outlier	Outlier
Change in Bayes	0.397***	−0.0964	0.583***	−0.128*
	(15.00)	(−0.51)	(18.76)	(−1.72)
Disconfirm	−11.41***	−35.85****	10.45***	12.19***
	(−14.21)	(−4.45)	(12.32)	(5.22)
Correction	8.757***	36.26***	−9.312***	−10.33***
	(11.80)	(5.28)	(−9.87)	(−4.41)
Observations	5238	300	3882	1422
R^2	0.181	0.222	0.242	0.031

Panel B: Speeders versus Non-Speeders				
Dependent Variable	*Change in Posterior Probability Estimate*			
	Good Distribution		**Bad Distribution**	
	Non-Speeders	Speeders	Non-Speeders	Speeders
Change in Bayes	0.370***	0.149	0.415***	0.299***
	(12.86)	(1.63)	(12.20)	(3.43)
Disconfirm	−13.75***	−7.236**	11.25***	7.325***
	(−13.89)	(−2.36)	(11.46)	(3.15)
Correction	10.81***	5-825	−10.07***	−5.991***
	(12.57)	(2.08)	(−10.37)	(−1.90)
Observations	5028	510	4734	570
R^2	0.190	0.040	0.143	0.039

Panel C: Forecasting Ability				
Dependent Variable	*Change in Posterior Probability Estimate*			
	Good Distribution		**Bad Distribution**	
	Above Median	Below Median	Above Median	Below Median
Change in Bayes	0.625***	0.0137	0.823***	0.111**
	(23.44)	(0.30)	(26.97)	(2.48)
Disconfirm	−6.218***	−21.97***	5.924***	13.95***
	(−8.90)	(−11.48)	(7.29)	(10.35)
Correction	5.420***	16.59***	−5.780***	−12.13***
	(8.36)	(9.71)	(−6.62)	(−8.42)
Observations	3270	2268	2154	3150
R^2	0.267	0.154	0.388	0.079

Note: This table reports the results of OLS regressions on how subjects update their posterior beliefs after a disconfirming signal and a correction across all experiments split by extreme outliers (Panel A), the time it takes subjects to finish the experiment (Panel B), and subjects' forecasting ability (Panel C). We report the results of OLS regressions for each subsample of individuals (with above-median versus below-median updating ability, no outlier versus outlier, and speeders versus non-speeders) and for each distribution (good and bad distribution) individually. Speeders are defined as the fastest 10% of the subjects. Non-speeders are defined as the remaining 90% of the subjects. Reported are coefficients and t-statistics (in parentheses) using robust standard errors. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

by individuals who lack the statistical skills to correctly perform the forecasting task, or whether even individuals who are closer to Bayesian behavior exhibit a pronounced bias. To examine this relation, we define the squared deviation of subjects' probability estimate in each period from the objective posterior probability as a measure of forecasting quality and conduct median splits. The results are reported in Table 5.7. Panel A reproduces the analysis split by extreme outliers, Panel B splits the sample by speeders, and Panel C reports results split by forecasting ability.

Overall, results are very similar, with two sets of results warrant a brief discussion. First, throughout each subsample, we consistently find an economically strong and statistically significant overreaction following a disconfirming signal, which is mostly corrected after observing a subsequent confirming signal. While the overreaction is even more pronounced for outliers and individuals with below-median forecasting ability, it is mostly unaffected by individuals' response time. This suggest that systematic violations of the counting heuristic appear to be a general phenomenon even though they correlate with participants' statistical skills. Yet, given that response time does not play a major role, attention does not appear to be a major driver. Second, when splitting the sample by extreme outliers, it becomes apparent that outliers are mostly clustered in the bad distribution. This confirms our previous finding, that a greater fraction of individuals struggles to forecast the bad distribution, even though both tasks should be – at least from a Bayesian perspective – equivalent.

5.5 Conclusion

The goal of this study is to test whether subjects follow a simple counting heuristic in belief updating as implied by Bayes' Rule: two informationally equivalent signals of opposite direction should always cancel out. However, our study suggests that this is not the case. Whenever a sequence of signals that go in the same direction is interrupted by a signal of opposite direction, subjects violate the simple counting heuristic and strongly overreact to the signal of opposite direction. In contrast to that, subjects correctly follow the counting heuristic whenever opposite-directional signals alternate.

Our results show a clear and robust pattern of over- and underreaction following violations of a simple counting heuristic. This pattern does not depend on the diagnosticity of the signals, on individuals' limited memory capacity, on signals not being anticipated, and the uncertainty of the underlying state. While, we identify *when* people violate simple counting rules, it remains an open question *why* they do so.

Our findings have relevant implications for various fields of research, among others investors' belief formation and trading behavior in financial markets as well as asset prices. In particular, the observation that agents' expectations are overly influenced by a single opposite-directional signal after a sequence of already just two same-directional signals may have valuable implications for how investors form expectations in financial markets and consequently act upon them. By and large, one of the most important and widely-applied ideas in behavioral financial economics is that people put too much weight on recent past returns, i.e. they over-extrapolate (Hong and Stein, 1999; Barberis and Shleifer, 2003; Greenwood and Shleifer, 2014; Barberis et al., 2015, 2018). This finding has important applications for excess stock market volatility, bubbles, and cross-sectional phenomena of stock returns such as for example momentum and long-term reversal. In models of extrapolative returns a crucial input parameter is the relative weight investors put on recent versus distant past returns. So far, the exact characteristics of this input parameter are still incomprehensively understood. For example, Cassella and Gulen (2018) recently show that the weight parameter varies over time, but cannot explain why this is the case. Our findings may add to a better understanding of the characteristics of this parameter in extrapolative belief formation, as we find that (i) individuals already strongly over-extrapolate from a single opposite-directional signal which interrupts a sequence of previous same-directional signals and (ii) that the observed over-extrapolation is relatively independent of the number of previously observed same-directional signals. In other words, individuals even over-extrapolate from a single opposite-directional signal if it occurs after a relatively long history of same-directional signals which in turn means that they even over-extrapolate in situations in which they are and should be quite sure about the underlying state of the world.

Appendix A

Closing A Mental Account: The Realization Effect for Gains and Losses

A.1 The Realization Effect and Skewness

In what follows, we use a non-myopic framework to explain more precisely how the realization effect and the skewness of the investment opportunity are related. To do so, we build on previous work by Barberis (2012) and Imas (2016) [in an alternative to his main model]. Barberis (2012) shows that cumulative prospect theory can explain sequential risk-taking behavior and demonstrates how the skewness of a lottery affects people's propensity to take risk. Therefore, his model is relevant for analyzing the relationship between realization and skewness. To explain differential risk-taking behavior after paper and realized losses in a non-myopic case, Imas (2016) uses the findings from Barberis (2012) framework. In the following, we will, therefore, also refer to Imas (2016).

A key ingredient of cumulative prospect theory is probability weighting (Tversky and Kahneman, 1992). People tend to overweight small probabilities while underweighting large probabilities. As a result of probability weighting, Barberis (2012) shows that (1) people are willing to invest in a symmetric lottery with negative expected payoff, and (2) prospect theory predicts an inconsistency in subsequent investment behavior. To understand these predictions, we briefly review his model. In the model, cumulative prospect theory generates inconsistency in sequential risk-taking environments, which is captured by the difference between people's ex-ante plans and actual behavior. People initially optimize over a set of potential gambling plans. For a wide range of parameters, people prefer the "loss-exit" plan, where they plan to continue gambling if they win and stop gambling if they start accumulating losses. However, after they begin gambling, people actually deviate from this "loss-exit" plan: They continue gambling when they lose and stop gambling when they have a significantly large gain.

The reasoning is as follows: The "loss-exit" plan makes accepting risk initially attractive. A key characteristic of this plan is that its perceived distribution of outcomes over all rounds is positively skewed. Since small probabilities of winning in this plan are overweighted, gambling becomes highly attractive. In other words, following this plan limits the downside (they stop gambling after losing) while it retains the potential upside (they continue gambling after winning), making the overall lottery distribution much more positively skewed than the one of a single lottery. However, over the course of rounds, the probabilities of the prospective outcome distribution change, becoming less positively skewed. The difference in skewness over final outcomes before and after the individual starts gambling is what generates inconsistent behavior. Barberis (2012) shows that for a wide range of preference parameter values, the described probability weighting effect outweighs the loss aversion effect, and thus people are willing to begin gambling in the first place with the "loss-exit" plan in mind. The trade-off between probability weighting and loss aversion will be important to our line of argument.

When do people deviate from the plan? People only show inconsistent behavior over the course of lotteries if the plan makes it attractive enough to accept risk in the first place. This means that the overall lottery distribution needs to be sufficiently positively skewed such that small probabilities within this plan are initially overweighted but over the course of rounds become less overweighted. More precisely, the probability weighting effect needs to dominate the loss aversion effect. Otherwise, the person would stick to his plan and not deviate from it. Now, what does realization do to gambling behavior? Imas (2016) has theoretically shown for losses that realization brings people closer to their initial plan if they suffer from inconsistent behavior. The argument is the following: Realization closes the respective mental account and internalizes the paper outcome. For a loss, this means that there is no option to break-even anymore and, therefore, people stop chasing losses. In addition to this, investing in the lottery after a realization becomes less attractive as the overall distribution becomes less positively skewed the more rounds have already been played. How realization influences risk taking after gains requires a little more explanation: For a gain, realization removes the possibility to offset future losses by previous gains. Therefore, losses are more painful after realized gains than after paper gains, which decreases people's willingness to take risk after a realized gain compared to a paper gain. In addition to this, the progress in rounds decreases the attractiveness of investing in the lottery because probability weighting changes over time. Once gains which were initially unlikely and, therefore, overweighted occurred, they are not perceived as unlikely anymore and consequently less overweighted. The lower attractiveness of the lottery combined with the larger sensitivity to future losses makes investing in the lottery after a realized gain less attractive than after a paper gain. Realization decreases people's propensity to gamble after a gain - they deviate from the ex-ante plan.

Realization of a loss brings people closer to their initial plan (Imas, 2016), whereas realization of a gain does not. How can this prediction be used to explain the relationship between the realization effect and the skewness of lotteries? First, we consider a symmetric lottery and assume that it is played over four rounds (as in experiment 2). We assume that people form an optimal "loss-exit" plan as described above. However, this plan has an overall lottery distribution, which is not very positively skewed compared to, for example, the one that would emerge from a positively skewed lottery. This has a key implication: People are less likely to deviate from the optimal plan for losses and more likely to deviate from it for gains. They are predicted to act inconsistently after gains while they act consistently after losses. This implication follows directly from Barberis (2012), who finds that at least 26 rounds are necessary to observe inconsistency with symmetric lotteries for the usually assumed preference parameter values of Tversky and Kahneman (1992). This can easily be seen in our setting of either a symmetric lottery with p=1/2 versus a positively skewed lottery with p=1/6 played over four rounds. The most favorable outcome (4 successes) occurs with a probability of 1/16 for the symmetric lottery compared to 1/1296 for the positively skewed lottery. Therefore, the more positively skewed the lottery is, the fewer rounds are needed to observe inconsistency.

This implication is essential to understand why the realization effect is less likely to be found for symmetric lotteries. A necessary condition to find differential behavior between paper and realized outcomes is that people deviate from their optimal plan after a loss and stick to it after a gain. As explained, this occurs if probability weighting is very pronounced due to the degree of skewness of the lottery distribution. The skewness can be affected in two ways: Either it can be increased by providing a positively skewed lottery from the beginning or by extending the number of rounds people can invest in the lottery. Since the number of rounds is fixed over all our experiments, it is the skewness of the symmetric lottery that makes deviations from the ex-ante plan after a loss unlikely and after a gain likely. After a paper loss, people are actually willing to gamble but do not do so with a symmetric lottery over a few rounds because the probability weighting effect does not outweigh the loss aversion effect.

There is a second argument adding to this reasoning: when abstracting from the role of ex-ante plans, a symmetric lottery provides little opportunity to recover from a paper loss because the potential upside is relatively small compared to the downside. Consistent with the effect of probability weighting, the decreased chance to break-even makes deviations from the ex-ante plan after a paper loss less likely. As explained by Imas (2016), realization of a loss brings people closer to the initial plan and circumvents inconsistent behavior. However, if there is no inconsistent behavior, realization should have little effect: people will adhere to their ex-ante plan after a loss and behave similarly after a paper and realized loss.

Similar reasoning works for gains. After a gain, the loss aversion effect outweighs the probability weighting effect. Although, previous paper gains cushion future losses, which decreases loss aversion, the progress in rounds in a symmetric lottery is accompanied by a strong reduction in probability overweighting. Ultimately, the unattractiveness of the lottery dominates, and the person is less likely to continue gambling after a paper gain. This means that the person shows inconsistency in his investment decisions. There is another point adding to the unattractiveness of a symmetric lottery: The relatively large downside can potentially wipe out people's previous gains as well as parts of their own money if they gamble again. Consistent with our previous line of argument, the risk to be wiped off by large losses makes deviations from the ex-ante plan after a paper gain more likely. As explained above, realization presents another way to prevent people from gambling after a gain. However, if there is inconsistent behavior after paper gains as well as after realized gains, realization has little effect in preventing people who else gamble from gambling: people will not adhere to their ex-ante plan after a gain and behave similarly after a paper and realized gain.

The same reasoning as above applies to negatively skewed lotteries. As the probability weighting works in the opposite direction (losses are overweighted and gains underweighted) and the potential upside relative to the downside becomes even less attractive compared to the symmetric lottery, realization will have little effect because people are already predicted to stick to their ex-ante plan after paper losses and deviate from it after paper gains.

A.2 Experiment Instructions

This appendix contains the instructions of experiment 1. Moreover, it presents exemplary screenshots of experiment 1.

Instructions in the Paper Treatment of Experiment 1

Welcome to our experimental study on decision making. The experiment will take about 30 minutes. All the money you earn is yours to keep. You receive 8.00 Euro in an envelope. This is your money which you can use to participate in the experiment. Please check that the envelope contains 8.00 Euro. The experiment consists of 4 successive rounds of investment decisions. You will have 8.00 Euro in total to invest. Each round you must decide how much of 2.00 Euro you would like to invest in a lottery: With a probability of 1/6 (16%) the lottery will "succeed" and you will make 7 times the amount you invested. With a probability of 5/6 (84%) the lottery will "fail" and you will lose the amount you invested. The procedure in each round is the same.
First, you are assigned one success number between 1 and 6. It is displayed on the computer screen. Second, you enter the amount you would like to invest in the lottery. The amount can be up to 2.00 Euro. When everyone is ready, the experimenter will roll a six-sided die in front of the class. If the rolled number is your success number, you will win the round and you will earn 7 times the amount invested. If the rolled number is not your success number, you will lose the invested amount. The outcome of the lottery is reported each round. Afterwards, you get a new success number and make the same decision in the next round.
At the end of the four rounds, your game payment will be the 8.00 Euro you started with plus your net earnings from the investments. Note that net earnings can be positive or negative.

Instructions in the Realization Treatment of Experiment 1

Welcome to our experimental study on decision making. The experiment will take about 30 minutes. All the money you earn is yours to keep. You receive 8.00 Euro in an envelope. This is your money which you can use to participate in the experiment. Please check that the envelope contains 8.00 Euro. The experiment consists of 3 successive rounds of investment decisions. You will have 6.00 Euro in total to invest. Each round you must decide how much of 2.00 Euro you would like to invest in a lottery: With a probability of 1/6 (16%) the lottery will "succeed" and you will make 7 times the amount you invested. With a probability of 5/6 (84%) the lottery will "fail" and you will lose the amount you invested. The procedure in each round is the same.

First, you are assigned one success number between 1 and 6. It is displayed on the computer screen. Second, you enter the amount you would like to invest in the lottery. The amount can be up to 2.00 Euro. When everyone is ready, the experimenter will roll a six-sided die in front of the class. If the rolled number is your success number, you will win the round and you will earn 7 times the amount invested. If the rolled number is not your success number, you will lose the invested amount. The outcome of the lottery is reported each round. Afterwards, you get a new success number and make the same decision in the next round.

At the end of the three rounds, your game payment will be the 8.00 Euro you started with plus your net earnings from the investments. Note that net earnings can be positive or negative. After the three rounds, you begin with the next part of the experiment. In the next part, you make one more decision.

Figure A.1: Experiment 1 Screen 1 Round 1

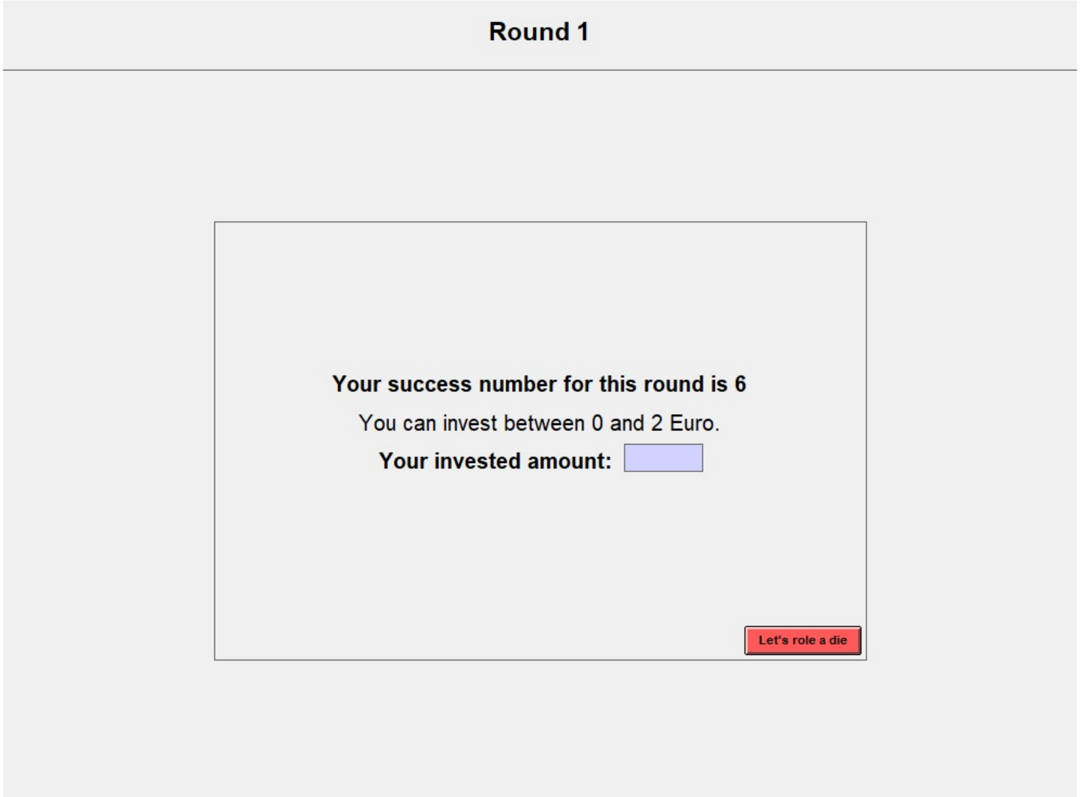


Figure A.2: Experiment 1 Screen 2 Round 1

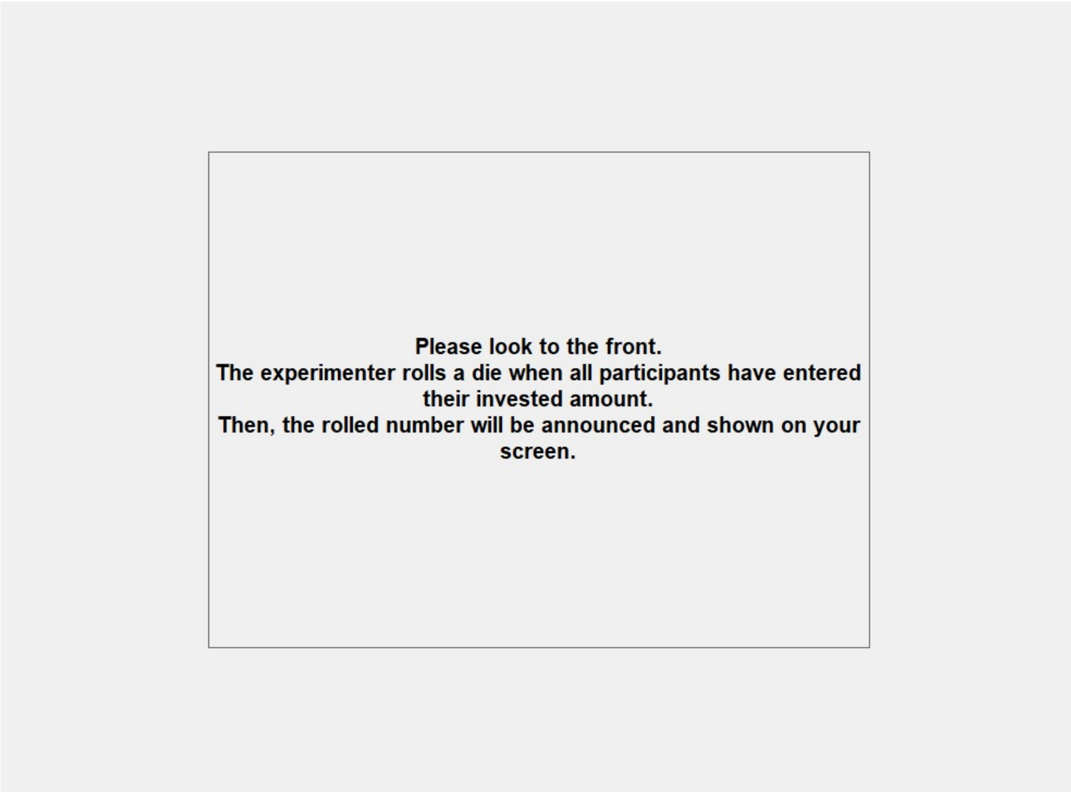


Figure A.3: Experiment 1 Screen 3 Round 1

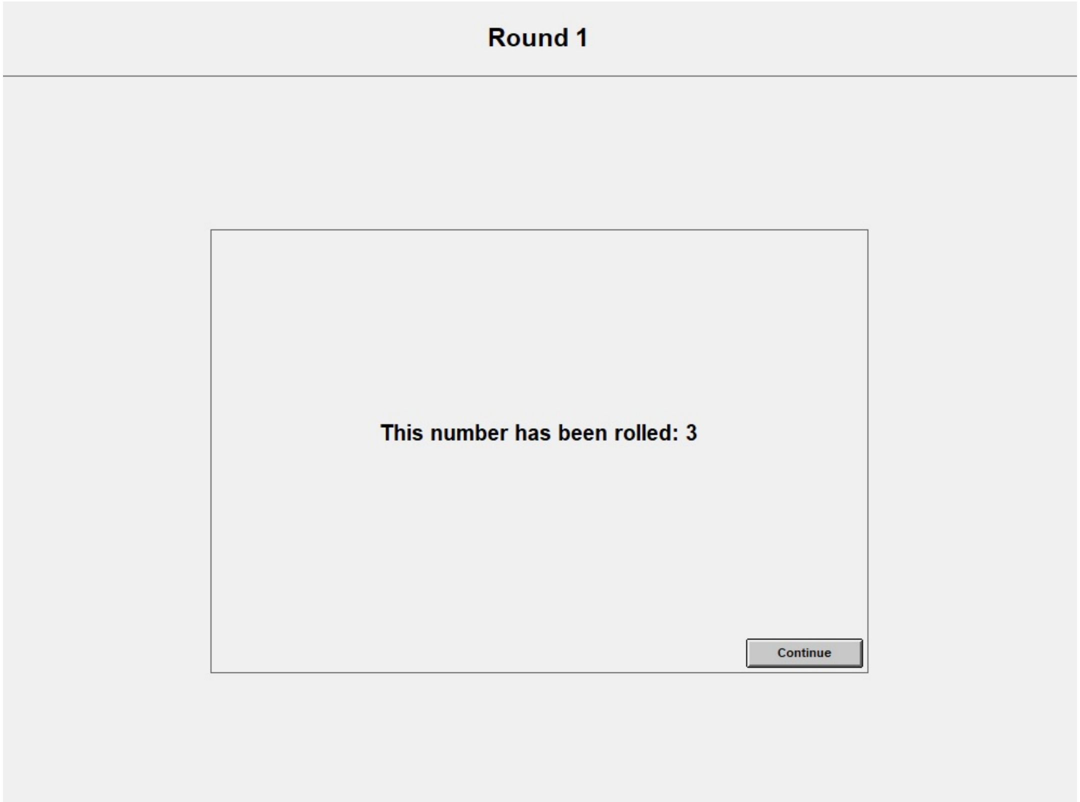


Figure A.4: Experiment 1 Screen 4 Round 1

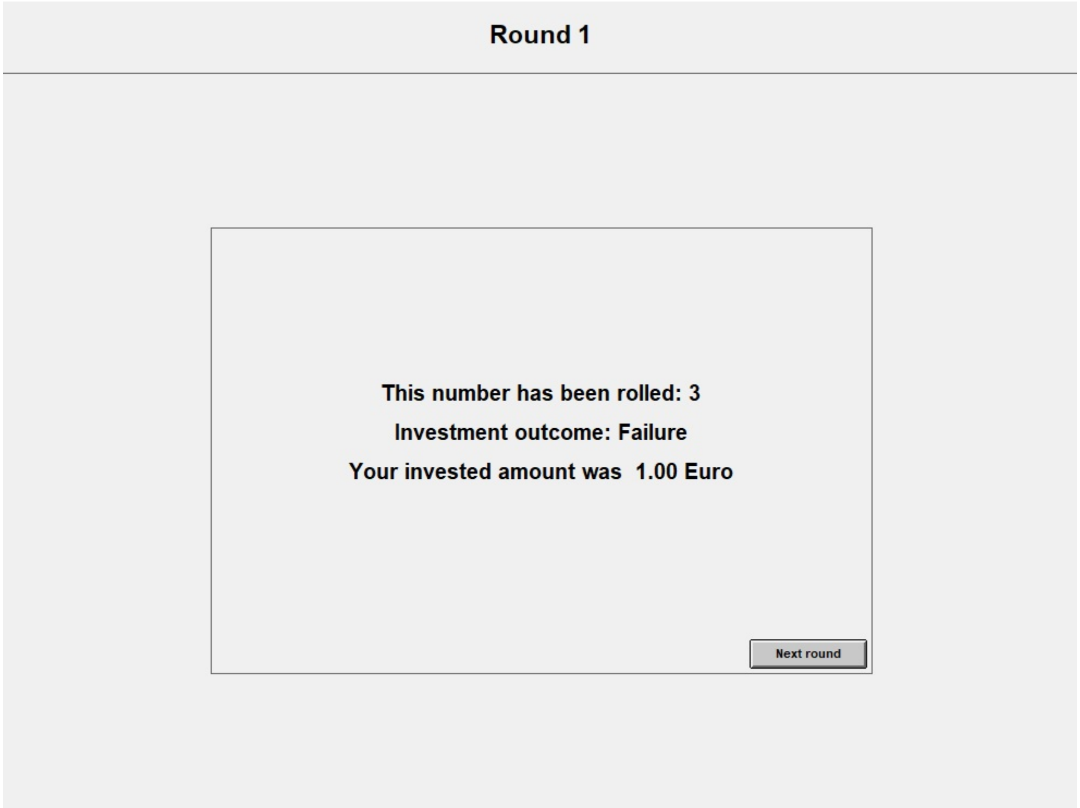


Figure A.5: Experiment 1 Earnings Update After Round 3 - Paper Treatment

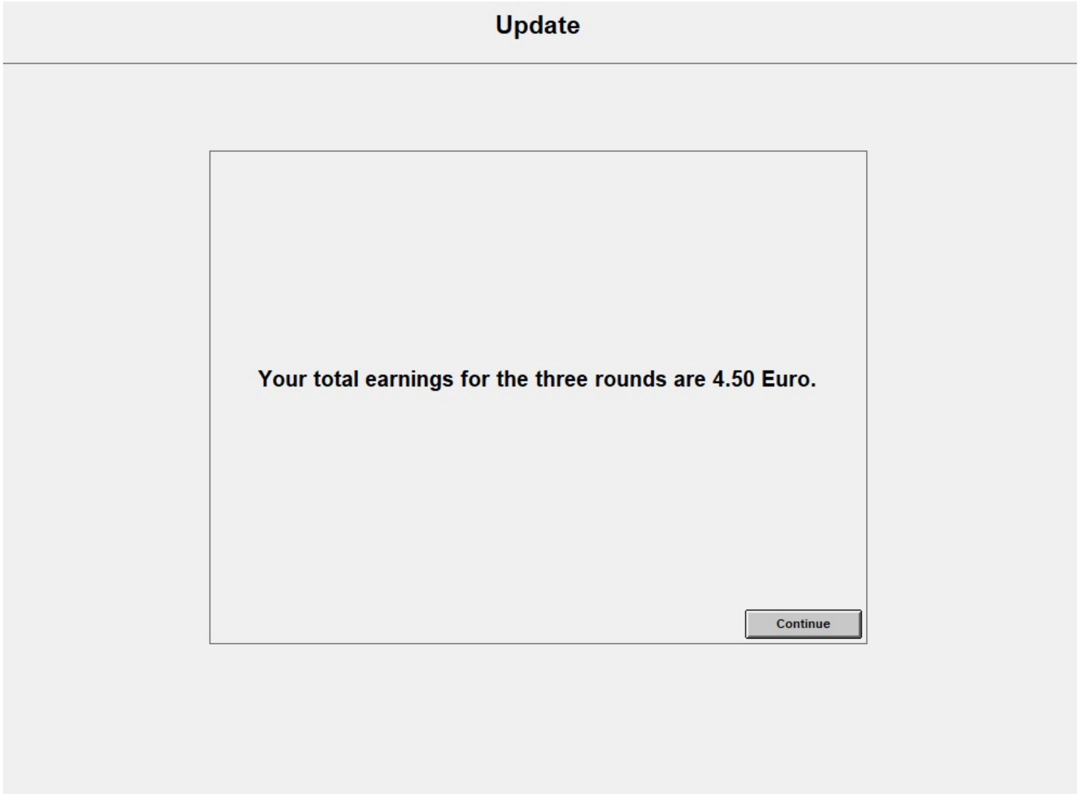
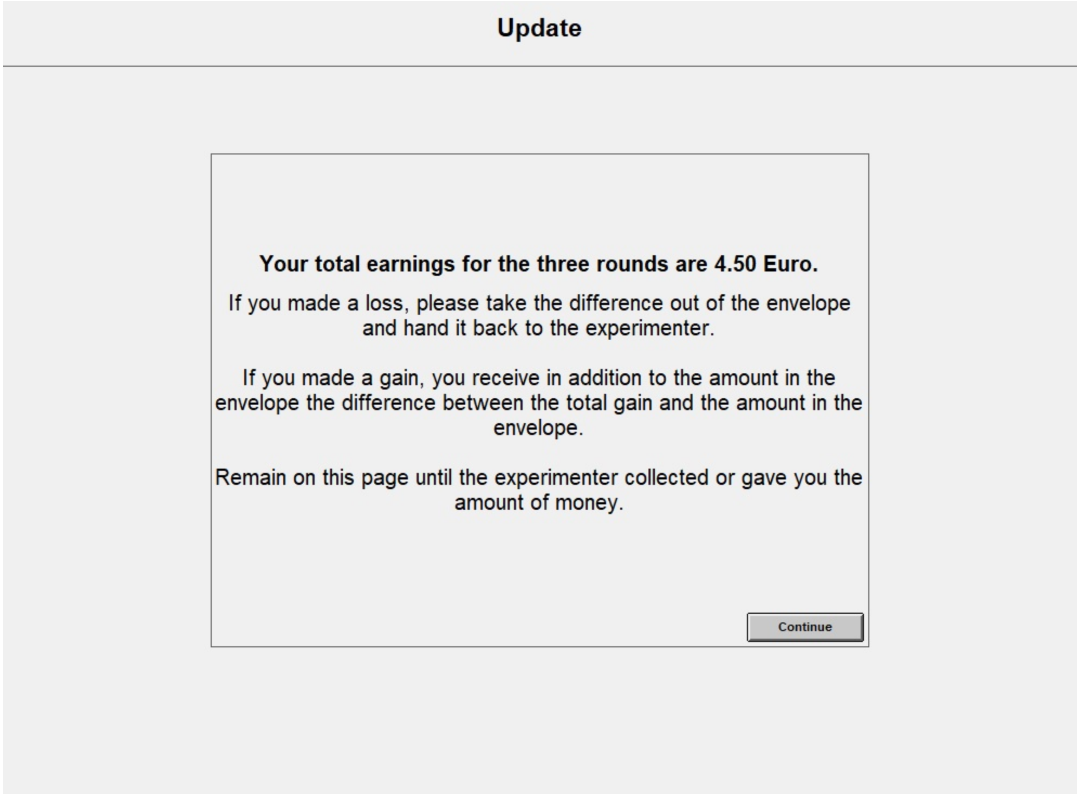


Figure A.6: Experiment 1 Earnings Update After Round 3 - Realization Treatment



A.3 Measures Used in the Experiments

This appendix details how different concepts were measured in the experiments including risk aversion, loss aversion, time preferences, illusion of control, financial literacy, and cognitive reflection.

Risk Aversion (Dohmen et al., 2011; Falk et al., 2016)

How do you see yourself: Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks? Please use a scale from 0 to 10 for your assessment. 0 means not at all willing to take risks and 10 means very willing to take risks.

Loss Aversion (Gächter et al., 2007)

Loss aversion is measured by the number of accepted gambles from Table A.1.

Table A.1: Choices to Accept or Reject a Coin Toss for Different Outcomes

	Accept	Reject
If the coin shows heads, you lose 20 Euro. If the coin shows tails, you win 60 Euro.		
If the coin shows heads, you lose 30 Euro. If the coin shows tails, you win 60 Euro.		
If the coin shows heads, you lose 40 Euro. If the coin shows tails, you win 60 Euro.		
If the coin shows heads, you lose 50 Euro. If the coin shows tails, you win 60 Euro.		
If the coin shows heads, you lose 60 Euro. If the coin shows tails, you win 60 Euro.		
If the coin shows heads, you lose 70 Euro. If the coin shows tails, you win 60 Euro.		

Time Preferences (Falk et al., 2016)

In comparison to others, are you a person who is generally willing to give up something today in order to benefit from that in the future or are you not willing to do so? Please use a scale from 0 to 10, where 0 means you are "completely unwilling to give up something today" and 10 means "you are very willing to give up something today." You can also use the values in-between to indicate where you are on the scale.

Illusion of Control (Wood and Clapham, 2005)

Read each of the following statements carefully. Rate to what extent you agree or disagree with each statement. 1 means that you strongly disagree and 5 means that you strongly agree.

1. There are secrets to successful casino gambling that can be learned.
2. It is a good advice to stay with the same pair of dice on a winning streak.
3. One should pay attention to lottery numbers that often win.
4. If a coin is tossed and comes up heads ten times in a row, the next toss is more likely to be tails.
5. The longer I have been losing, the more likely I am to win.

Financial Literacy (Van Rooij et al., 2011)

1. Which of the following statements describes the main function of the stock market?
   (a) The stock market helps to predict stock earnings.
   (b) The stock market results in an increase in the prices of stocks.
   (c) The stock market brings people who want to buy stocks together with people who want to sell stocks.
   (d) None of the above.
   (e) Do not know.
2. Which of the following statements is correct? If somebody buys the stock of firm B in the stock market:
   (a) He owns a part of firm B.
   (b) He hast lent money to firm B.
   (c) He is liable for firm B's debt.
   (d) None of the above.

(e) Do not know.

3. Which of the following statements is correct?

   (a) Once one invests in a mutual fund, one cannot withdraw the money in the first year.
   (b) Mutual funds can invest in several assets, for example invest in both stocks and bonds.
   (c) Mutual funds pay a guaranteed rate of return which depends on their past performance.
   (d) None of the above.
   (e) Do not know.

4. Which of the following statements is correct? If somebody buys a bond of firm B:

   (a) He owns a part of firm B.
   (b) He has lent money to firm B.
   (c) He is liable for firm B's debt.
   (d) None of the above.
   (e) Do not know.

5. Considering a long time period (for example 10 or 20 years), which asset normally gives the highest return?

   (a) Savings accounts
   (b) Bonds
   (c) Stocks
   (d) Do not know.

6. Normally, which assets displays the highest fluctuations over time?

   (a) Savings accounts
   (b) Bonds

(c) Stocks

(d) Do not know

7. When an investor spreads his money over different assets, does the risk of losing money:

   (a) Increase

   (b) Decrease

   (c) Stay the same

   (d) Do not know.

8. If the interest rate falls, what should happen to bond prices?

   (a) Rise

   (b) Fall

   (c) Stay the same

   (d) Do not know.

Cognitive Reflection Test (Toplak et al., 2014)

1. If John can drink one barrel of water in 6 days and Mary can drink one barrel of water in 12 days, how long would it take them to drink one barrel of water together?

2. Jerry receives both the 15th highest and the 15th lowest mark in the class. How many students are in the class?

3. A man buys a pig for 60 Euro, sells it for 70 Euro, buys it back for 80 Euro and sells it finally for 90 Euro. How much has he made?

4. Simon decided to invest 8,000 Euro in the stock market one day early in 2017. Six months after he invested, on July 17, the stocks he had purchased were down 50%. Fortunately for Simon, from July 17 to October 17, the stocks he had purchased went up 75%. At this point, Simon

(a) has broken even in the stock market

(b) is ahead of where he began

(c) has lost money

5. A bat and a ball cost 1.10 Euro in total. The bat costs 1.00 Euro more than the ball. How much does the ball cost?

6. If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets?

7. In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake?

A.4 Additional Results from the Experiment

This appendix presents additional results from the experiments. In Tables A.1, A.2, and A.3, we examine dynamic risk-taking more generally and report the changes in risk taking over all rounds prior to the final round for the positively skewed, symmetric, and negatively skewed lottery, respectively. In Table A.4, we provide an overview of the outcomes after round 3 for the participants in the positively skewed lottery. Table A.5 shows how risk taking after gains in the positively skewed lottery depends on the round of success. The regression analysis in Table A.6 presents whether, and to what extent the realization effect depends on individual characteristics such as, for example, gender, risk aversion and time preferences.

Table A.1: Dynamic Risk-Taking in the Positively Skewed Lottery (Round 1 to 3)

Panel A: Risk-taking after losses						
	Invested amount			Change		
Treatment	Round 1	Round 2	Round 3	R2−R1	R3−R2	N
Paper	0.98	0.91	0.78	−0.07	−0.13	57
				(1.11)	(1.77)	
Realization	0.90	0.73	0.80	−0.17	0.07	58
				(3.07)	(1.15)	
Panel B: Risk-taking after gains						
	Invested amount			Change		
Treatment	Round 1	Round 2	Round 3	R2−R1	R3−R2	N
Paper	0.94	0.73	0.71	−0.21	−0.02	35
				(2.86)	(0.28)	
Realization	0.71	0.77	0.73	0.06	−0.04	36
				(0.65)	(0.42)	

Note: The table shows the average invested amounts in the lottery and the respective changes in the average invested amounts between rounds for the first three rounds of experiment 1 (in Euro). Panel A is restricted to participants who lost in the first three rounds of the experiment, Panel B shows averages for all participants with at least one gain in the first three rounds. Both panels show results by treatment (paper and realization). Change is the difference between the investment of two subsequent rounds. N provides the number of participants for each treatment-outcome combination. T-values of a two-sided t-test are shown in parentheses.

Table A.2: Dynamic Risk-Taking in the Symmetric Lottery (Round 1 to 3)

Panel A: Risk-taking after losses						
	Invested amount			Change		
Treatment	Round 1	Round 2	Round 3	R2−R1	R3−R2	N
Paper	1.40	1.37	1.45	−0.03	0.08	22
				(0.41)	(1.45)	
Realization	1.57	1.50	1.53	−0.07	0.03	15
				(0.49)	(0.29)	
Panel B: Risk-taking after gains						
	Invested amount			Change		
Treatment	Round 1	Round 2	Round 3	R2−R1	R3−R2	N
Paper	1.35	1.15	1.19	−0.20	0.04	27
				(1.80)	(0.52)	
Realization	1.43	1.56	1.38	0.13	−0.18	30
				(1.81)	(1.57)	

Note: The table shows the average invested amounts in the lottery and the respective changes in the average invested amounts between rounds for the first three rounds of experiment 2 (in Euro). Panel A is restricted to participants who lost in the first three rounds of the experiment, Panel B shows averages for all participants with at least one gain in the first three rounds. Both panels show results by treatment (paper and realization). Change is the difference between the investment of two subsequent rounds. N provides the number of participants for each treatment-outcome combination. T-values of a two-sided t-test are shown in parentheses.

Table A.3: Dynamic Risk-Taking in the Negatively Skewed Lottery (Round 1 to 3)

Panel A: Risk-taking after losses						
	Invested amount			Change		
Treatment	Round 1	Round 2	Round 3	R2−R1	R3−R2	N
Paper	1.53	1.66	1.67	0.13	0.01	32
				(2.12)	(0.04)	
Realization	1.70	1.79	1.78	0.09	−0.01	38
				(1.64)	(0.26)	
Panel B: Risk-taking after gains						
	Invested amount			Change		
Treatment	Round 1	Round 2	Round 3	R2−R1	R3−R2	N
Paper	1.45	1.60	1.60	0.15	0.00	70
				(2.52)	(0.00)	
Realization	1.58	1.71	1.67	0.13	−0.04	64
				(1.86)	(0.67)	

Note: The table shows the average invested amounts in the lottery and the respective changes in the average invested amounts between rounds for the first three rounds of experiment 3 (in Euro). Panel A is restricted to participants who lost in the first three rounds of the experiment, Panel B shows averages for all participants with at least one gain in the first three rounds. Both panels show results by treatment (paper and realization). Change is the difference between the investment of two subsequent rounds. N provides the number of participants for each treatment-outcome combination. T-values of a two-sided t-test are shown in parentheses.

Table A.4: Overview of Outcomes After Round Three in the Positively Skewed Lottery

Outcome by end of round 3	N Total	Average Earnings	Std. Deviation of Earnings	N Paper	N Realization
Loss (Earnings $<$ EUR 8.00)	115	5.45	1.62	57	58
Gain (Earnings $>$ EUR 8.00)	71	12.04	1.92	35	36
Success only in round 1	21	11.83	2.24	11	10
Success only in round 2	17	12.50	3.46	12	5
Success only in round 3	27	11.39	2.86	10	17
Success in two rounds	6	14.38	2.95	2	4
No Gain/Loss (Earnings = EUR 8.00)	17	8.00	0	13	4

Note: The table shows how many participants have negative, positive or zero net earnings after round three, as well as the average earnings and the standard deviation of the earnings after round three. For those participants who have a gain after round three, we report in which round they won the lottery (those with a loss never won the lottery). We provide the number of observations for each outcome by treatment.

Table A.5: Risk-Taking After Gains in the Positively Skewed Lottery

Panel A: Success in round 1						
	Invested amount				Change	
Treatment	Round 1	Round 2	Round 3	Round 4	R4−R3	N
Paper	0.96	0.72	0.74	1.00	0.26	13
					(2.00)	
Realization	0.78	0.68	0.53	0.48	−0.05	14
					(0.73)	
Difference	0.19	0.04	0.21	0.52	0.31	
	(0.95)	(0.20)	(0.91)	(2.32)	(2.17)	
Panel B: Success in round 2						
	Invested amount				Change	
Treatment	Round 1	Round 2	Round 3	Round 4	R4−R3	N
Paper	1.00	0.87	0.68	0.77	0.08	13
					(0.64)	
Realization	0.81	1.01	0.67	0.83	0.16	7
					(1.04)	
Difference	0.18	−0.14	0.01	−0.06	−0.07	
	(0.84)	(0.42)	(0.03)	(0.21)	(0.34)	
Panel C: Success in round 3						
	Invested amount				Change	
Treatment	Round 1	Round 2	Round 3	Round 4	R4−R3	N
Paper	0.86	0.50	0.64	0.86	0.22	11
					(1.30)	
Realization	0.61	0.65	0.82	0.62	−0.20	19
					(1.91)	
Difference	0.25	−0.15	−0.18	0.24	0.42	
	(1.08)	(0.60)	(0.78)	(0.92)	(2.24)	

Note: The table shows the average invested amounts in the lottery for all rounds of experiment 1 (in Euro) conditional on the round of success. Panel A is restricted to participants who win in round one and have a gain by the end of round three, Panel B is displays averages for participants who win in round two and have a gain by the end of round three, and Panel C shows averages for all participants who win in round three and have a gain by the end of round three. All three panels show results by treatment (paper and realization) and differences between treatments. Change is the difference between the investment in the final round and round three. N provides the number of participants for each treatment-outcome combination. T-values of a two-sided t-test are shown in parentheses.

Table A.6: The Realization Effect for Gains and Losses with Controls

	Change in invested amount							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Realization	-0.175***	-0.182***	-0.193***	-0.173***	-0.179***	-0.174***	-0.176***	-0.213***
	(-0.0657)	(-0.0656)	(-0.066)	(-0.0662)	(-0.0661)	(-0.0663)	(-0.0664)	(-0.0667)
Male	0.115*							0.159**
	(-0.0659)							(-0.0784)
RA		-0.0275**						-0.0127
		(-0.0136)						(-0.0161)
LA			0.0037					0.0244
			(-0.0292)					(-0.0328)
TP				0.00288				0.00594
				(-0.0156)				(-0.0168)
IOC					0.0501			0.0417
					(-0.0445)			(-0.0487)
FL						-0.0013		-0.0162
						(-0.0165)		(-0.0198)
CRT							-0.00585	-0.0201
							(-0.0181)	(-0.0209)
Constant	0.0324	0.232***	0.0984	0.0623	-0.0208	0.0915	0.116	0.112
	(-0.0547)	(-0.0859)	(-0.0776)	(-0.13)	(-0.104)	(-0.0986)	(-0.108)	(-0.247)
Observations	200	200	189	200	200	200	200	189
R^2	0.048	0.054	0.044	0.034	0.04	0.034	0.034	0.081

Note: The table shows the results of OLS regressions with the change in the invested amount between rounds three and four in experiment 1 as the dependent variable. Realization is an indicator variable taking a value of one for the realization treatment. Control variables are a male indicator, risk aversion (RA; reversed scale of the original question, 0=not risk averse, 10=very risk averse), loss aversion (LA), time preferences (TP; 0=very impatient, 10=very patient), illusion of control (IOC; 1=low, 5=high), financial literacy (FL; 0=low, 8=high), and cognitive reflection (CR; 0=low, 7=high). Robust standard errors are shown in parentheses. ***, **, and * indicate significance on the 1%, 5%, and 10% levels, respectively.

A.5 Online Experiment

This appendix presents the results of an experiment, which was conducted online. The main difference to experiments 1-3 is that we used another realization mechanism and a different framing of the final investment round in the online experiment.

In his study, Imas (2016) tests the boundary conditions of the realization effect with respect to the realization mechanism.While in the original experiment money is physically transferred to participants by the experimenter, a robustness experiment considers an electronic transfer of money between different accounts. The process of sending money is initiated by participants. They have to type "closed" to transfer and realize outcomes from the first three lottery rounds. Based on the results, this realization mechanism is sufficient to produce a similarly strong realization effect. Imas (2016) concludes that a physical transfer of money is not necessary to show the realization effect.

We challenge this conclusion for two reasons. First, it is instrumental for the realization effect that the investor recognizes the difference between paper and realized outcomes and the point in time of realization. Not only is this distinction presumably less salient for a virtual transfer of money, but it remains open whether the two separate electronic accounts also constitute separate mental accounts. Secondly, literature in psychology shows that people perceive a physical transfer of money differently than an electronic transfer of money. For example, paying with a credit card is perceived as less painful than paying with cash explained by the transparency of the payment outflow (Raghubir and Srivastava, 2008). Likewise, realization utility might be felt less intensely when relying on this realization mechanism. We, therefore, replicate the two different mechanisms in our experiment.

Furthermore, we examine how framing affects the realization effect. We initially did not intend to investigate this question as part of the online experiment. However, after we ran the online experiment, we noticed that we deviated from the original design in how we labeled the final round in the realization treatment. This change in framing turns out to make a difference. We will, therefore, also report the results of an additional treatment in which we replicate the framing of the final round exactly as in Imas (2016).

A.5.1 Design and Participants

We discuss that the realization mechanism might be a major determinant for differential risk taking after realized and unrealized outcomes. The less strong realization effect identified in the prior experiments casts some doubt on whether a weaker realization mechanism is sufficient to generate the effect. We thus replicate study 2 by Imas (2016) in experiment 4, which uses an online experiment without physical transfer of money.

The structure of the online experiment was similar to the laboratory experiments with the main differences that the stakes were smaller, and the realization mechanism was modified. Participants were paid a fixed amount of \$0.30, plus their earnings at the end of the experiment. They received an endowment of \$1.00 at the beginning of the experiment to be used in four subsequent investment decisions. In each round, participants were randomly assigned a success number and decided how much of \$0.25 to invest in the same lottery as in experiment 1. Afterward, they rolled a virtual six-sided die, and the rolled number (randomly generated) was presented on the computer screen. If the success number matched the rolled number, the invested amount was multiplied by seven; if not, the invested amount was lost. Participants learned the outcome and continued with the next round with a new success number.

Participants were randomly assigned to either the paper or the realization treatment. In the realization treatment, earnings were reported at the end of the third round, and participants were asked to type "closed" in a dedicated window to realize their position. It was explained that any money they lost up to this round would be withdrawn from their account and transferred to the experimenter and any money they won would be credited to their account. Afterward, the same lottery was offered for an additional investment decision. Participants in the paper treatment viewed their earnings after round three. However, they did not initialize any transfer of money before continuing with the final investment decision. The design matches the original design by Imas (2016) as close as possible using the same stakes and, most importantly, the same realization mechanism.

As mentioned in the previous section, we unintentionally deviated from the original design with respect to the framing of the final round in the realization treatment. Instead of "Round 1" we labeled the final round "Additional Round".[1] We initially did not expect this change in framing to affect the outcome of the experiment since we assume a proper realization mechanism to be robust against relatively minor framing effects. Nevertheless, we run a second realization treatment (Realization Round 1) on Amazon mechanical Turk in which we revert the framing in the final round from "Additional Round" to "Round 1".

Experiment 4, including all three treatments, was programmed in SoSciSurvey, a platform to create academic survey studies and conducted online using the labor market of Amazon mechanical Turk, which allowed us to get access to a more representative sample of the population.[2] We recruited 471 individuals for the experiment, again guided by a power analysis. Participants were on average 36 years old, 42% of participants were male, 39% stated that they attended a statistics class, and the level of cognitive reflection was lower than in the student sample (1.84 correct answers out of 4).

A.5.2 Results

We first examine the investment decisions of participants who lost in each round prior to round four. Panel A of Table A.1 shows investments in all

[1] We noticed this deviation from the original design after we ran the online experiment.

[2] `https://www.soscisurvey.de/index.php?id=index&lang=en`.

rounds by treatment for this group. Investments are not statistically different in round one to three across treatments. Focusing on the changes in investment between rounds three and four across treatments, we cannot find evidence for a realization effect. The difference of the changes in investment between treatments is insignificant and points in the wrong direction ($DiD = -1.48$, $t(197) = 1.37$, $p = 0.17$). To understand this result better, we consider the investment by treatment group. In line with a realization effect, participants in the paper treatment take more risk after a paper loss, but only marginally and not statistically significant (0.44, $t(103) = 0.64$, $p = 0.52$). Unlike in experiment 1, participants in the realization treatment take significantly more risk after losses (1.92, $t(94) = 2.29$, $p = 0.02$). This result is contradictory to the realization effect, which predicts the opposite.

As an explanation for these findings, it appears natural to question the modified realization mechanism. In this experiment, an electronic transfer of money between accounts instead of a physical transfer of money is supposed to induce a mental realization of earnings. We argue that by merely typing "closed," participants do not perceive the difference between realized and unrealized outcomes and do not derive disutility from the realization of a loss (Barberis and Xiong, 2012). Since the realization is unrecognized, the mental account remains open, and the opportunity to break even by taking more risk persists. The investment pattern of participants in the realization treatment is consistent with this reasoning. The mechanism might rather emphasize the existence of a loss and stimulate more risk taking.

Further evidence for the assumption that participants did not perceive the electronic transfer of money between accounts as a realization might come from the gain domain. Panel B of Table A.1 summarizes the investments for participants who have a gain by the end of round three. Indeed, we neither find a realization effect for gains. Participants do not take less risk after a realized gain compared to a paper gain, but rather the opposite ($DiD = -1.67$, $t(152) = 1.38$, $p = 0.17$). As for losses, participants tend to take more risk after a realized gain than before a realized gain (1.50, $t(64) = 1.81$, $p = 0.07$).

Again this is inconsistent with a realization effect but consistent with the assumption that participants in the realization treatment did not part with the money they gained.

In the remaining part, we show the results of the "Realization (Round 1 framing)" treatment. To remind the reader, we only change the label of the final round in the realization treatment from "Additional Round" to "Round 1". The change in framing is implemented to exactly align this version with the original design of the online study by Imas (2016). Consistent with the realization effect, participants now tend to take less risk after a gain and a loss in the realization treatment when using the proposed change in framing (see Table A.1). In the slightly altered version, we thus successfully replicate the online study by Imas (2016). One possible reason is that the altered framing of the final round strengthens the triggered process of closing the respective mental account. The "Round 1" frame makes the beginning of a new investment episode clearer to participants, while the "Additional Round" frame might rather give participants the impression that they still have the chance to recover from the previous loss as they are offered an additional lottery. This would imply a continuation of the investment episode rather than an end.

We conclude that the realization effect is sensitive to the realization mechanism if money is not physically transferred. Presumably, participants in the realization treatment of the online experiment do not part with the money in the same way as those participants in the experiments with physical transfer do. Therefore, the realization effect becomes vulnerable to circumstantial effects such as framing. Future research could follow up on potential other effects that may be interrelated with the realization effect.

Table A.1: Risk-Taking in the Online Experiment

Panel A: Risk taking after losses						
	Invested amount				Change	
Treatment	Round 1	Round 2	Round 3	Round 4	R4−R3	N
Paper	12.03	9.30	10.07	10.51	0.44	104
					(0.64)	
Realization	13.06	10.78	11.32	13.24	1.92	95
					(2.29)	
Realization (Round 1 framing)	12.54	9.89	14.97	12.51	−2.46	37
					(2.51)	
Difference	−1.03	−1.48	−1.25	−2.73	−1.48	
	(0.84)	(1.20)	(1.00)	(2.00)	(1.37)	
Difference (Round 1 framing)	−0.51	−0.59	−4.91	−2.00	2.90	
	(0.32)	(0.37)	(3.04)	(1.14)	(2.23)	
Panel B: Risk taking after gains						
	Invested amount				Change	
Treatment	Round 1	Round 2	Round 3	Round 4	R4−R3	N
Paper	17.64	14.79	13.41	13.24	−0.17	70
					(0.19)	
Realization	15.87	13.38	13.39	14.89	1.50	84
					(1.81)	
Realization (Round 1 framing)	11.74	10.94	11.61	10.45	−1.16	31
					(0.81)	
Difference	1.77	1.40	0.02	−1.65	−1.67	
	(1.28)	(0.93)	(0.01)	(1.05)	(1.38)	
Difference (Round 1 framing)	5.90	3.85	1.80	2.79	0.99	
	(3.17)	(1.96)	(0.91)	(1.35)	(0.60)	

Note: The table shows the average invested amounts in the lottery for all rounds of experiment 4 (in Cent). Panel A is restricted to participants who have a loss by the end of round three, Panel B shows averages for all participants who have a gain by the end of round three. Both panels show results by treatment (paper, realization and realization round 1 framing) and differences between treatments. In this experiment the realization is a non-physical transfer of money. Participants in the realization treatment had to type the word "closed" in a respective window after round 3 to realize their earnings. In the realization round 1 framing treatment, the final round was named "Round 1" instead of "Additional Round". Change is the difference between the investment in the final round and round three. N provides the number of participants for each treatment-outcome combination. T-values of a two-sided t-test are shown in parentheses.

Appendix B

The Portfolio Composition Effect

B.1 Experiment Instructions

In this section we present the experiment instructions. First, the instructions of Experiment 1 are shown and then the instructions of Experiment 3. The instructions of Experiment 2 are similar to the instructions of Experiment 3.

Experiment 1

Dear participant,
You participate in an experiment on decision making which is part of a research study at the University of Mannheim.
In the following you will be presented with the performance of two portfolios of stocks. Each portfolio consists of ten different stocks. Please imagine that you bought the respective stocks one month ago. You invested equal amounts of money in each stock. Now you observe the performance of the stocks in each of your portfolios.
Please take your time and ask yourself how you would feel when observing the performance. There are two pairs of portfolios. It is possible that the second pair of portfolios is shown to you before the first pair of portfolios.
Overall, this study will take 3-5 minutes. You will be compensated $0.50 for the successful completion of this HIT on MTurk.

Experiment 3

Dear participant,

You participate in an experiment on decision making which is part of a research study at the University of Mannheim. Please read all instructions carefully. Your payment depends on your decisions. Overall this study will take approximately 10 minutes.

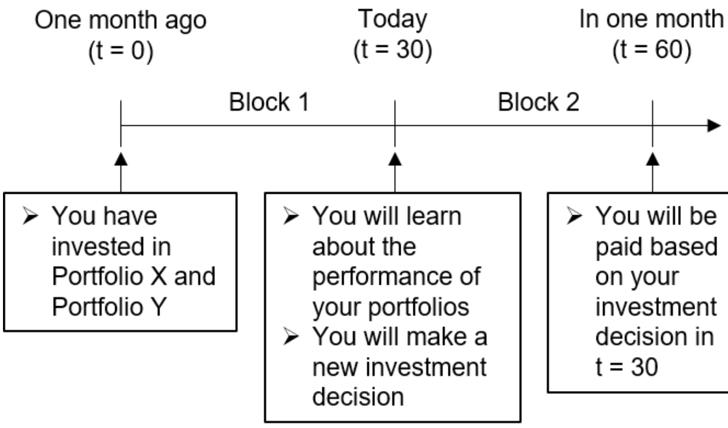
In the following you will be presented with the performance of two portfolios of stocks (**Portfolio X and Portfolio Y**). Each portfolio consists of ten different stocks. Please imagine that you have bought the respective stocks in period 0 (t = 0). To be precise, you have invested 10,000 ECU (experimental currency unit) in Portfolio X and 10,000 ECU in Portfolio Y in period 0. Within each portfolio, you have invested **equal amounts in each stock** (i.e. 1,000 ECU in each stock). More about the exchange rate between ECU and $ is described at the end of the instructions.

Today, you are in period 30 (see graph below) and you observe the performance of your portfolios. In particular, you will see how each stock in each of your portfolios has performed over 30 periods (block 1). Before you make any further decision, both portfolios will be rebalanced (the weight of each stock will be reset to 1/10). Then, at the beginning of block 2, you will be asked to make a **return forecast** for each portfolio and an **investment decision** for the next 30 periods. Importantly, while the weights of the stocks are reset between the blocks, the stocks themselves in your portfolios remain the same.

How do stock prices change over time?

Each period, the price of a stock can either increase by z or decrease by -z (z is supposed to be a variable that takes an absolute value). How likely it is that a stock price increases or decreases depends on its type. There are two types: A stock can be a **good stock** or a **bad stock**. If the stock is a good stock, the probability that the price increases is 70% and the probability that the price decreases is 30%. While, if the stock is a bad stock, the probability that the price increases is 30% and the probability that the price decreases is 70%.

Figure B.1: Timeline of the Experiment



In the beginning (t = 0), you do not know whether a stock is a good or a bad stock. As such, it is equally likely that a stock will be good or bad, i.e. the probability is exactly 50%. The table gives an overview of the types of stocks with the probability distributions.

Figure B.2: Probability Distribution of Stocks

	Good stock	**Bad stock**
Probability of **price increase by z**	0.70	0.30
Probability of **price decrease by -z**	0.30	0.70
Expected change in price	0.40 · z	-0.40 · z

Today, in period 30, you will observe 30 price changes for each stock. From this information, you can learn whether a stock is more likely to be a good or a bad stock. If you observe more increases in price than decreases, the stock is more likely to be a good stock, while if you observe more decrease in price than increases, the stock is more likely to be a bad stock.

Although, all stocks follow the same described rules, they differ in the magnitude of price change z. For each stock, z (and consequently $-z$) is randomly determined once and remains fixed over 60 periods. For example, the value of z may be 6 for one stock (e.g. Stock U $(+/-6)$), such that this stock can increase in price by 6 or decrease in price by -6. While for another stock (e.g.

Stock W (+/$-$10)), the value of z may be 10, such that this stock can increase in price by 10 or decrease in price by -10. Once again, how likely each outcome is, depends on the type of stock (see table). Consequently, **the expected price change of a stock** depends on its type (good or bad) and the magnitude of price change. The expected price change is calculated as $0.7z - 0.3z = 0.4z$ if you believe the stock is good or $0.3z - 0.7z = -0.4z$ if you believe the stock is bad.

Comfortably, the computer will do the calculations for you. Once you are asked to **make a return forecast**, the computer will support you by doing the calculations. However, one thing you need to do by yourself, is to decide whether the stock is more likely to be a "good" or a "bad" stock.

In addition to the portfolio return forecast, you will **make an investment decision** in period 30. You will be asked to allocate "fresh" money between Portfolio X and Portfolio Y for the investment horizon of 30 periods (between period 30 and period 60). This investment decision will be payoff-relevant.

Your payment:

You will be paid according to your performance which will be based on your **investment decision**. For the investment decision, you will be endowed with 10,000 ECU which can increase or decrease in value depending on your decision. This means that **you will earn the proportion of the change in portfolio value** between period 30 and period 60 (block 2) given the amount invested in each portfolio (e.g. assume, you invest x% in Portfolio Y which has a total increase in value of 30, you will earn x% of 30). Changes in price of 100 ECU correspond to $ 0.10 (e.g. a portfolio value increase of 150 units corresponds to a 15 cent gain).

Depending on your investment decision, you can gain money which will be added to your fixed payment of $ 1.00.

There is one last important information. We briefly want to make you familiar with the presentation format and then ask you some comprehension questions.

You can see an example of how the performance of the portfolios of stocks is presented to you below. On the left hand side, you can see the performance of Portfolio X and on the right hand side the performance of Portfolio Y. For each stock, we show ...

- the size of the positive and negative return (z and $-z$) in parentheses (e.g. Stock A ($+/-4$),
- the number of days with a positive return and the number of days with a negative return,
- and the total value change of the respective stock over 30 periods.

The total value change of each stock can easily be calculated by summing up the product of z times the number of positive return days and $-z$ times the number of negative return days.
On the following page, we will ask you some comprehension questions.

Figure B.3: Display Format of the Portfolios of Stocks

Portfolio X

Stock		Number of positive return days	Number of negative return days	Total change in value
Stock A	(+/-4)			
Stock B	(+/-10)			
Stock C	(+/-6)			
Stock D	(+/-7)			
Stock E	(+/-2)			
Stock F	(+/-5)			
Stock G	(+/-2)			
Stock H	(+/-9)			
Stock I	(+/-6)			
Stock J	(+/-3)			
Total change in portfolio value				

Portfolio Y

Stock		Number of positive return days	Number of negative return days	Total change in value
Stock K	(+/-2)			
Stock L	(+/-3)			
Stock M	(+/-2)			
Stock N	(+/-8)			
Stock O	(+/-5)			
Stock P	(+/-6)			
Stock Q	(+/-1)			
Stock R	(+/-2)			
Stock S	(+/-12)			
Stock T	(+/-1)			
Total change in portfolio value				

Comprehension Questions

Below we report the comprehension questions that participants had to answer correctly after reading the instructions to proceed to the Bayesian Updating task. Correct responses are displayed in italic.

1. Imagine you observe the following performance of Stock A (+/−4) in period 30: Number of positive return days = 18, number of negative return days = 12. Please evaluate whether Stock A is more likely to be a good or a bad type.
   (a) *Good type*
   (b) Bad type

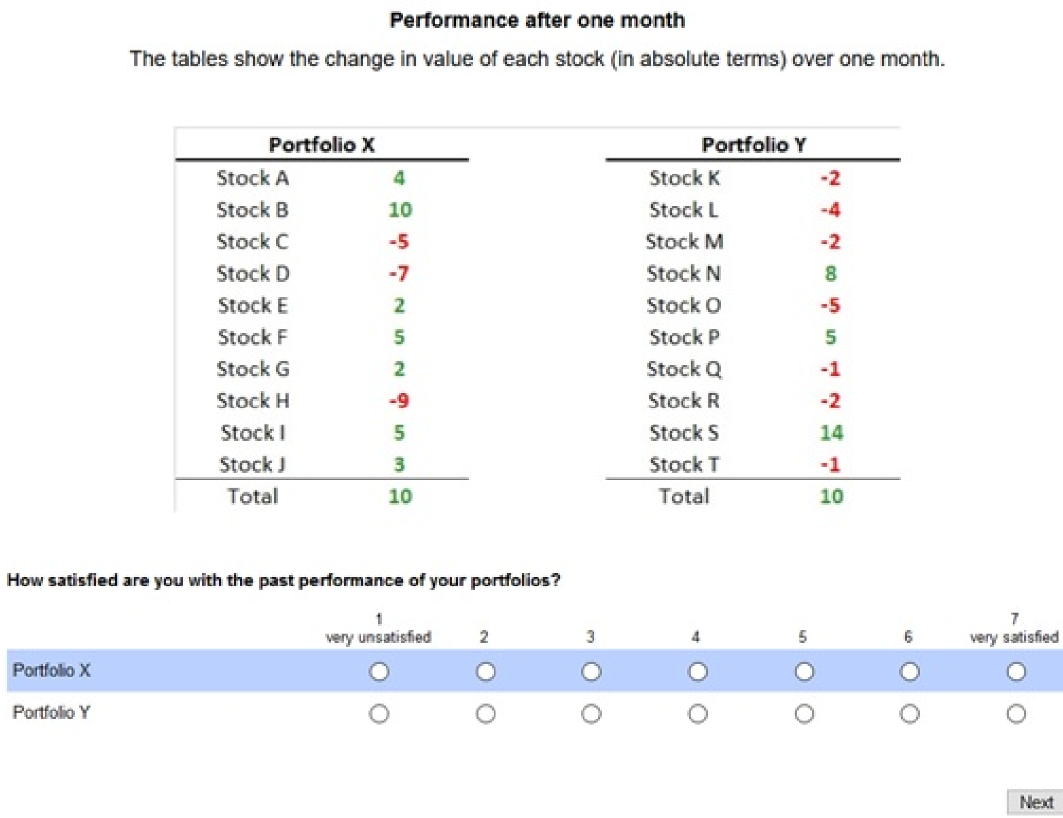
2. What is the expected return of Stock A (+/−4) for the next period given the following performance in period 30: Number of positive return days = 18, number of negative return days = 12? The computer will do the calculation in the investment task. However, we kindly ask you to do it on your own in this question such that you understand what the computer will do.
   (a) −2.0
   (b) −1.6
   (c) −1.2
   (d) −0.8
   (e) −0.4
   (f) 0
   (g) 0.4
   (h) 0.8
   (i) 1.2
   (j) *1.6*
   (k) 2.0

3. Please evaluate the statement: Stock A (+/−4) can only make a return of +4 or −4 per period.
   (a) *True*
   (b) False

B.2 Screenshots of the Experiments

In this section, we display screenshots of the experiments. Figure B.4 to Figure B.7 present screenshots of Experiment 1 and Figure B.8 to Figure B.11 show screenshots of Experiment 3.

Figure B.4: Screen with Satisfaction Question



Performance after one month

The tables show the change in value of each stock (in absolute terms) over one month.

Portfolio X	
Stock A	4
Stock B	10
Stock C	-5
Stock D	-7
Stock E	2
Stock F	5
Stock G	2
Stock H	-9
Stock I	5
Stock J	3
Total	10

Portfolio Y	
Stock K	-2
Stock L	-4
Stock M	-2
Stock N	8
Stock O	-5
Stock P	5
Stock Q	-1
Stock R	-2
Stock S	14
Stock T	-1
Total	10

How satisfied are you with the past performance of your portfolios?

	1 very unsatisfied	2	3	4	5	6	7 very satisfied
Portfolio X	○	○	○	○	○	○	○
Portfolio Y	○	○	○	○	○	○	○

Next

Figure B.5: Screen with Investment Task

Performance after one month

The tables show the change in value of each stock (in absolute terms) over one month.

Portfolio X	
Stock A	4
Stock B	10
Stock C	-5
Stock D	-7
Stock E	2
Stock F	5
Stock G	2
Stock H	-9
Stock I	5
Stock J	3
Total	10

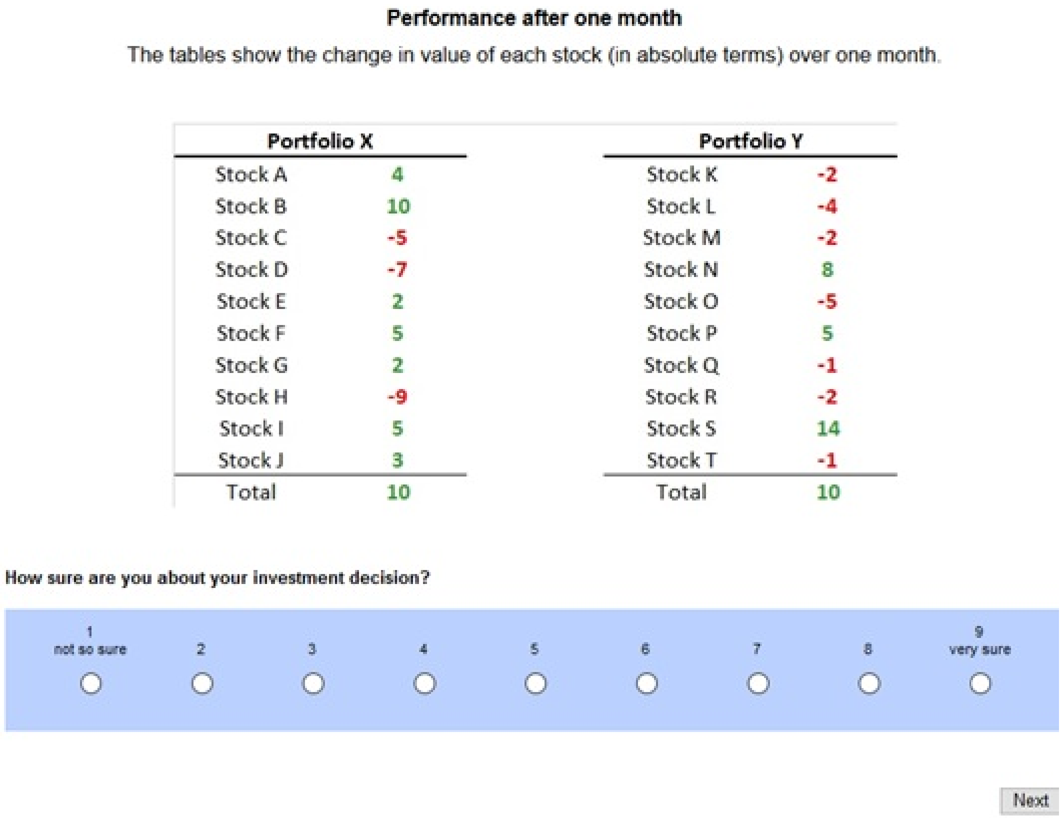
Portfolio Y	
Stock K	-2
Stock L	-4
Stock M	-2
Stock N	8
Stock O	-5
Stock P	5
Stock Q	-1
Stock R	-2
Stock S	14
Stock T	-1
Total	10

If you had to choose, how would you allocate 1000 US Dollar between portfolio X and portfolio Y?

Amount in portfolio X	
Amount in portfolio Y	
Sum	0

Next

Figure B.6: Screen with Confidence Question



Performance after one month

The tables show the change in value of each stock (in absolute terms) over one month.

Portfolio X	
Stock A	4
Stock B	10
Stock C	-5
Stock D	-7
Stock E	2
Stock F	5
Stock G	2
Stock H	-9
Stock I	5
Stock J	3
Total	10

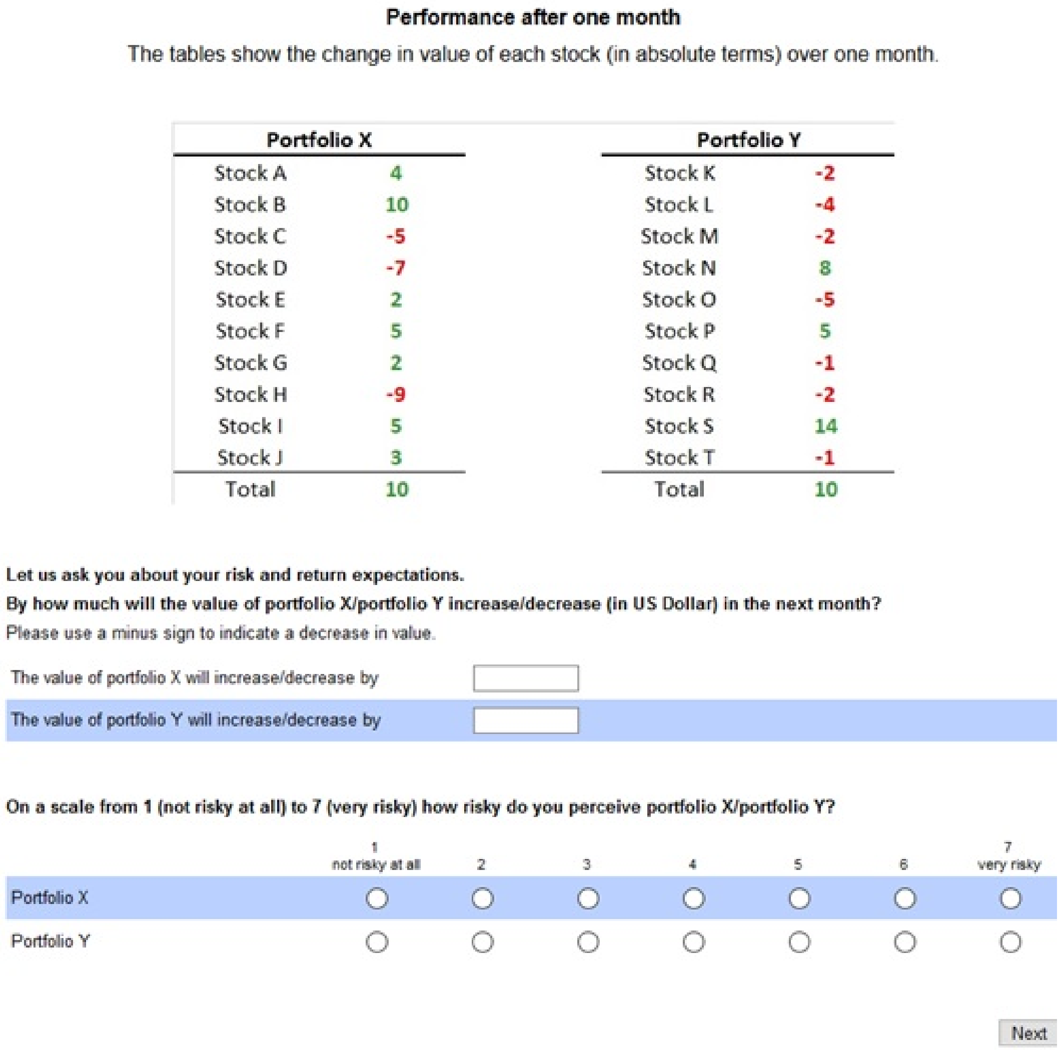
Portfolio Y	
Stock K	-2
Stock L	-4
Stock M	-2
Stock N	8
Stock O	-5
Stock P	5
Stock Q	-1
Stock R	-2
Stock S	14
Stock T	-1
Total	10

How sure are you about your investment decision?

1 not so sure	2	3	4	5	6	7	8	9 very sure
○	○	○	○	○	○	○	○	○

Next

Figure B.7: Screen with Return Expectations and Risk Perception Question



Performance after one month

The tables show the change in value of each stock (in absolute terms) over one month.

Portfolio X		Portfolio Y	
Stock A	4	Stock K	-2
Stock B	10	Stock L	-4
Stock C	-5	Stock M	-2
Stock D	-7	Stock N	8
Stock E	2	Stock O	-5
Stock F	5	Stock P	5
Stock G	2	Stock Q	-1
Stock H	-9	Stock R	-2
Stock I	5	Stock S	14
Stock J	3	Stock T	-1
Total	10	Total	10

Let us ask you about your risk and return expectations.
By how much will the value of portfolio X/portfolio Y increase/decrease (in US Dollar) in the next month?
Please use a minus sign to indicate a decrease in value.

The value of portfolio X will increase/decrease by

The value of portfolio Y will increase/decrease by

On a scale from 1 (not risky at all) to 7 (very risky) how risky do you perceive portfolio X/portfolio Y?

	1 not risky at all	2	3	4	5	6	7 very risky
Portfolio X	○	○	○	○	○	○	○
Portfolio Y	○	○	○	○	○	○	○

Next

Figure B.8: Screen with Assessment of Stock Type

Period 30

Today, you can observe the performance of your investment after one month (from t=0 to t=30).

Portfolio X

Stock		Number of positive return days	Number of negative return days	Total change in value
Stock A	(+/-4)	21	9	48
Stock B	(+/-10)	22	8	140
Stock C	(+/-6)	7	23	-96
Stock D	(+/-7)	8	22	-98
Stock E	(+/-2)	22	8	28
Stock F	(+/-5)	20	10	50
Stock G	(+/-2)	24	6	36
Stock H	(+/-9)	10	20	-90
Stock I	(+/-6)	21	9	72
Stock J	(+/-3)	22	7	42
Total change in portfolio value				132

Portfolio Y

Stock		Number of positive return days	Number of negative return days	Total change in value
Stock K	(+/-2)	3	27	-48
Stock L	(+/-3)	12	18	-18
Stock M	(+/-2)	11	19	-16
Stock N	(+/-8)	24	6	144
Stock O	(+/-5)	8	22	-70
Stock P	(+/-6)	21	9	72
Stock Q	(+/-1)	11	19	-8
Stock R	(+/-2)	11	19	-16
Stock S	(+/-12)	19	11	96
Stock T	(+/-1)	13	17	-4
Total change in portfolio value				132

What are the expected portfolio returns (total expected change in portfolio values) and standard deviation of Portfolio X and Portfolio Y for the next period (t=31)? Based on your evaluation below, the computer will calculate the expected portfolio return and the portfolio standard deviation.

Portfolio X:

Please evaluate for each stock in Portfolio X whether it is more likely to be a good stock (good type) or a bad stock (bad type)?

	Good Type	Bad Type
Stock A	○	○
Stock B	○	○
Stock C	○	○
Stock D	○	○
Stock E	○	○
Stock F	○	○
Stock G	○	○
Stock H	○	○
Stock I	○	○
Stock J	○	○

Portfolio Y:

Please evaluate for each stock in Portfolio Y whether it is more likely to be a good stock (good type) or a bad stock (bad type)?

	Good Type	Bad Type
Stock K	○	○
Stock L	○	○
Stock M	○	○
Stock N	○	○
Stock O	○	○
Stock P	○	○
Stock Q	○	○
Stock R	○	○
Stock S	○	○
Stock T	○	○

Figure B.9: Screen with Return Expectations and Volatility

Period 30

Today, you can observe the performance of your investment after one month (from t=0 to t=30).

Portfolio X

Stock		Number of positive return days	Number of negative return days	Total change in value
Stock A	(+/-4)	21	9	48
Stock B	(+/-10)	22	8	140
Stock C	(+/-6)	7	23	-96
Stock D	(+/-7)	8	22	-98
Stock E	(+/-2)	22	8	28
Stock F	(+/-5)	20	10	50
Stock G	(+/-2)	24	6	36
Stock H	(+/-9)	10	20	-90
Stock I	(+/-6)	21	9	72
Stock J	(+/-3)	22	7	42
Total change in portfolio value				132

Portfolio Y

Stock		Number of positive return days	Number of negative return days	Total change in value
Stock K	(+/-2)	3	27	-48
Stock L	(+/-3)	12	18	-18
Stock M	(+/-2)	11	19	-16
Stock N	(+/-8)	24	6	144
Stock O	(+/-5)	8	22	-70
Stock P	(+/-6)	21	9	72
Stock Q	(+/-1)	11	19	-8
Stock R	(+/-2)	11	19	-16
Stock S	(+/-12)	19	11	96
Stock T	(+/-1)	13	17	-4
Total change in portfolio value				132

Your evaluation of good and bad stocks:

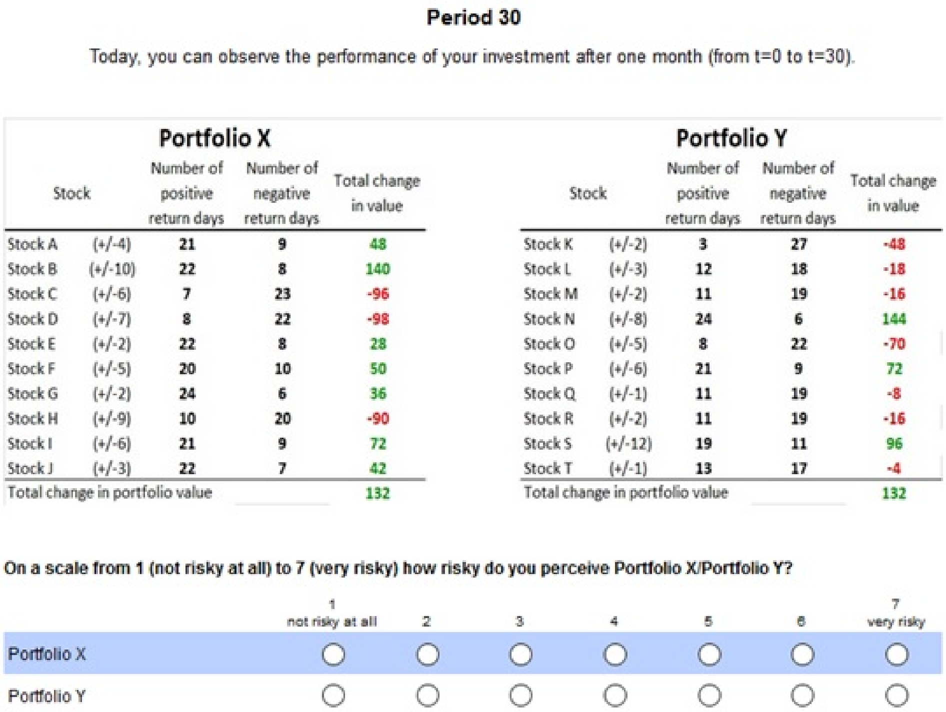
Stock A: Good
Stock B: Good
Stock C: Bad
Stock D: Bad
Stock E: Good
Stock F: Good
Stock G: Good
Stock H: Bad
Stock I: Good
Stock J: Good

Stock K: Bad
Stock L: Bad
Stock M: Bad
Stock N: Good
Stock O: Bad
Stock P: Good
Stock Q: Bad
Stock R: Bad
Stock S: Good
Stock T: Bad

The expected return of Portfolio X for the next period: **4**
The expected standard deviation of Portfolio X for the next period: **24.3**

The expected return of Portfolio Y for the next period: **4**
The expected standard deviation of Portfolio Y for the next period: **24.3**

Figure B.10: Screen with Risk Perception Question



Period 30

Today, you can observe the performance of your investment after one month (from t=0 to t=30).

Portfolio X

Stock		Number of positive return days	Number of negative return days	Total change in value
Stock A	(+/-4)	21	9	48
Stock B	(+/-10)	22	8	140
Stock C	(+/-6)	7	23	-96
Stock D	(+/-7)	8	22	-98
Stock E	(+/-2)	22	8	28
Stock F	(+/-5)	20	10	50
Stock G	(+/-2)	24	6	36
Stock H	(+/-9)	10	20	-90
Stock I	(+/-6)	21	9	72
Stock J	(+/-3)	22	7	42
Total change in portfolio value				132

Portfolio Y

Stock		Number of positive return days	Number of negative return days	Total change in value
Stock K	(+/-2)	3	27	-48
Stock L	(+/-3)	12	18	-18
Stock M	(+/-2)	11	19	-16
Stock N	(+/-8)	24	6	144
Stock O	(+/-5)	8	22	-70
Stock P	(+/-6)	21	9	72
Stock Q	(+/-1)	11	19	-8
Stock R	(+/-2)	11	19	-16
Stock S	(+/-12)	19	11	96
Stock T	(+/-1)	13	17	-4
Total change in portfolio value				132

On a scale from 1 (not risky at all) to 7 (very risky) how risky do you perceive Portfolio X/Portfolio Y?

	1 not risky at all	2	3	4	5	6	7 very risky
Portfolio X	○	○	○	○	○	○	○
Portfolio Y	○	○	○	○	○	○	○

Figure B.11: Screen with Investment Task

Period 30

Today, you can observe the performance of your investment after one month (from t=0 to t=30).

Portfolio X

Stock		Number of positive return days	Number of negative return days	Total change in value
Stock A	(+/-4)	21	9	48
Stock B	(+/-10)	22	8	140
Stock C	(+/-6)	7	23	-96
Stock D	(+/-7)	8	22	-98
Stock E	(+/-2)	22	8	28
Stock F	(+/-5)	20	10	50
Stock G	(+/-2)	24	6	36
Stock H	(+/-9)	10	20	-90
Stock I	(+/-6)	21	9	72
Stock J	(+/-3)	22	7	42
Total change in portfolio value				132

Portfolio Y

Stock		Number of positive return days	Number of negative return days	Total change in value
Stock K	(+/-2)	3	27	-48
Stock L	(+/-3)	12	18	-18
Stock M	(+/-2)	11	19	-16
Stock N	(+/-8)	24	6	144
Stock O	(+/-5)	8	22	-70
Stock P	(+/-6)	21	9	72
Stock Q	(+/-1)	11	19	-8
Stock R	(+/-2)	11	19	-16
Stock S	(+/-12)	19	11	96
Stock T	(+/-1)	13	17	-4
Total change in portfolio value				132

You are endowed with "fresh" money of 10,000 ECU. For a new investment of 30 periods, how do you want to allocate 10,000 ECU between Portfolio X and Portfolio Y?

Portfolio X	
Portfolio Y	
Sum	0

B.3 Portfolio Expected Return and Standard Deviation

Portfolios in Experiment 2 and 3 are designed such that (i) the expected portfolio return and (ii) the standard deviation of portfolio returns are identical. We calculate expected returns and standard deviation using the standard formulas.

$$\mu_p = \sum_{i=1}^{n} w_i \mu_i \tag{B.1}$$

$$\sigma_p^2 = \sum_{i=1}^{n} w_i^2 \sigma_i^2 + \sum_{i=1} \sum_{i \neq j} w_i w_j Cov(i, j) \tag{B.2}$$

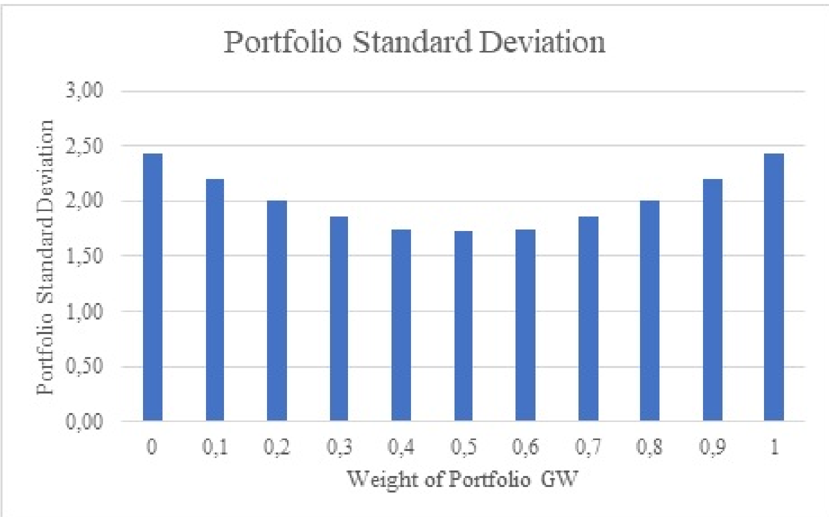
The expected return and the standard deviation of individual stocks are calculated based on these formulas:

$$\mu_i = p_i X_i + (1 - p_i)(-X_i) \tag{B.3}$$

$$\sigma_i^2 = p_i (X_i - \mu)^2 + (1 - p_i)(X_i - \mu)^2 \tag{B.4}$$

Table B.1 show the values for the two portfolios in Experiment 3.

Figure B.12: Portfolio Standard Deviation



The highest Sharpe ratio is achieved by investing 50% in Portfolio GW and 50% in Portfolio GL.

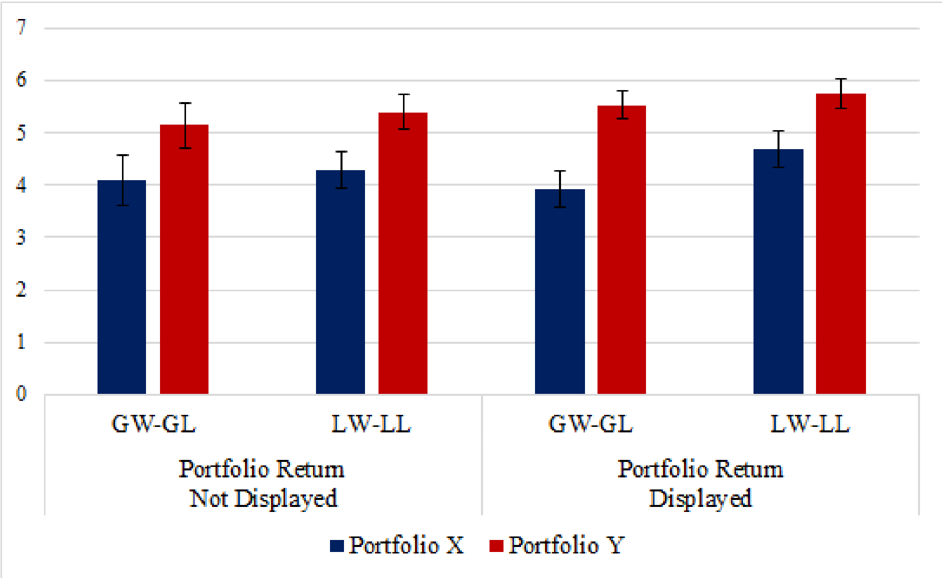
Table B.1: Portfolio Expected Return and Standard Deviation

Portfolio GW							
Stock	High Return	Low Return	P (High Return)	P (Low Return)	E(Return)	Std. Dev.	Weight
A	4	-4	0.7	0.3	1.60	5.97	0.1
B	10	-10	0.7	0.3	4.00	14.88	0.1
C	6	-6	0.3	0.7	-2.40	5.90	0.1
D	7	-7	0.3	0.7	-2.80	6.87	0.1
E	2	-2	0.7	0.3	0.80	3.02	0.1
F	5	-5	0.7	0.3	2.00	7.46	0.1
G	2	-2	0.7	0.3	0.80	3.02	0.1
H	9	-9	0.3	0.7	-3.60	8.80	0.1
I	6	-6	0.7	0.3	2.40	8.94	0.1
J	3	-3	0.7	0.3	1.20	4.50	0.1
Portfolio					4.0	24.30	
Portfolio GL							
Stock	High Return	Low Return	P (High Return)	P (Low Return)	E(Return)	Std. Dev.	Weight
K	2	-2	0.3	0.7	-0.80	2.12	0.1
L	3	-3	0.3	0.7	-1.20	3.04	0.1
M	2	-2	0.3	0.7	-0.80	2.12	0.1
N	8	-8	0.7	0.3	3.20	11.91	0.1
O	5	-5	0.3	0.7	-2.00	4.94	0.1
P	6	-6	0.7	0.3	2.40	8.94	0.1
Q	1	-1	0.3	0.7	-0.40	1.28	0.1
R	2	-2	0.3	0.7	-0.80	2.12	0.1
S	12	-12	0.7	0.3	4.80	17.86	0.1
T	1	-1	0.3	0.7	-0.40	1.28	0.1
Portfolio					4.0	24.30	

B.4 Additional Analyses

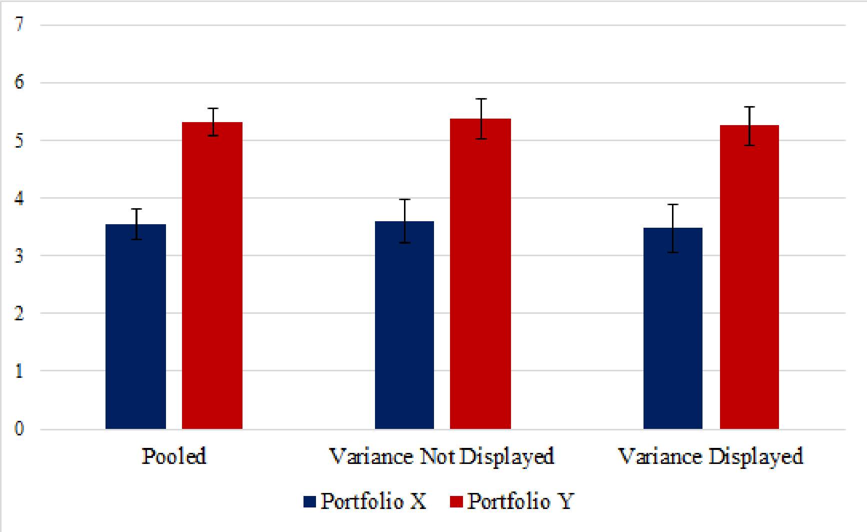
This section presents results of further analyses.

Figure B.13: Risk Perception in Experiment 2 (Baseline Treatments)



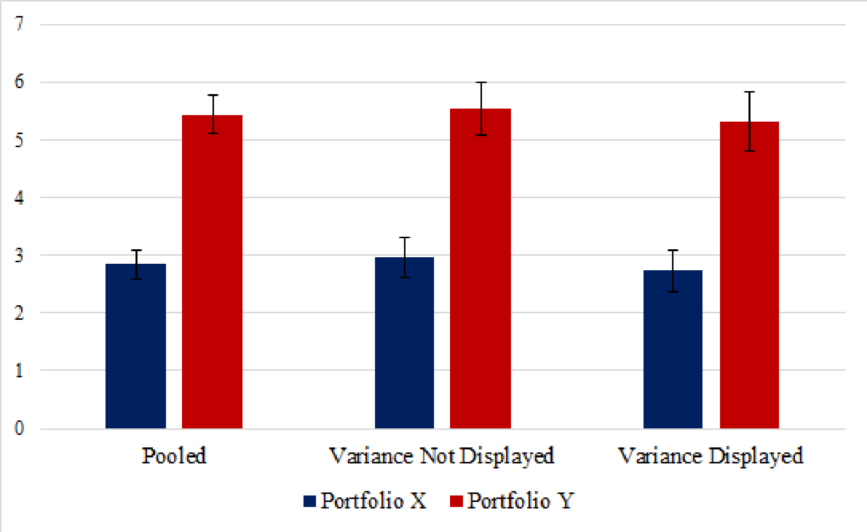
Note: The figure shows participants' mean risk perception for each portfolio elicited on a Likert scale from 1: low to 7: high for the two portfolio pairs $G_pW_S - G_pL_S$ and $L_pW_S - L_pL_S$. The blue bars refer to Portfolio X which corresponds to the first two letters of each portfolio pair (e.g. G_pW_S for the first portfolio pair) and the red bars refer to Portfolio Y which corresponds to the second two letters of each portfolio pair (e.g. G_pL_S for the first portfolio pair). Displayed are 95%confidence intervals.

Figure B.14: Risk Perception in Experiment 3



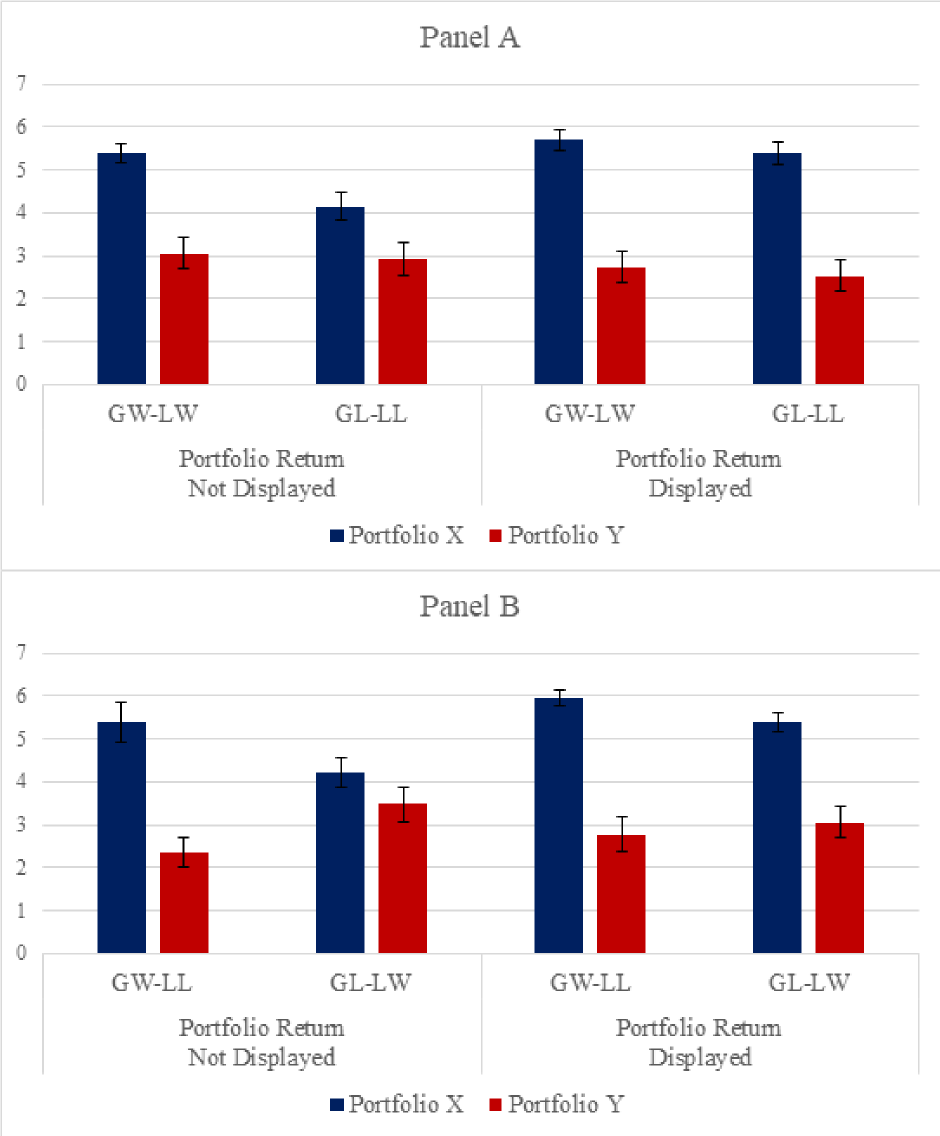
Note: The figure shows participants' mean risk perception for each portfolio elicited on a Likert scale from 1: low to 7: high for the portfolio pair $G_pW_S - G_pL_S$. The blue bar refers to Portfolio X which corresponds to the first two letters of the portfolio pair (e.g. G_pW_S for the first portfolio pair) and the red bar refers to Portfolio Y which corresponds to the second two letters of the portfolio pair (e.g. G_pL_S for the first portfolio pair). Displayed are 95%confidence intervals.

Figure B.15: Risk Perception in Experiment 3 Conditional on Expected Returns



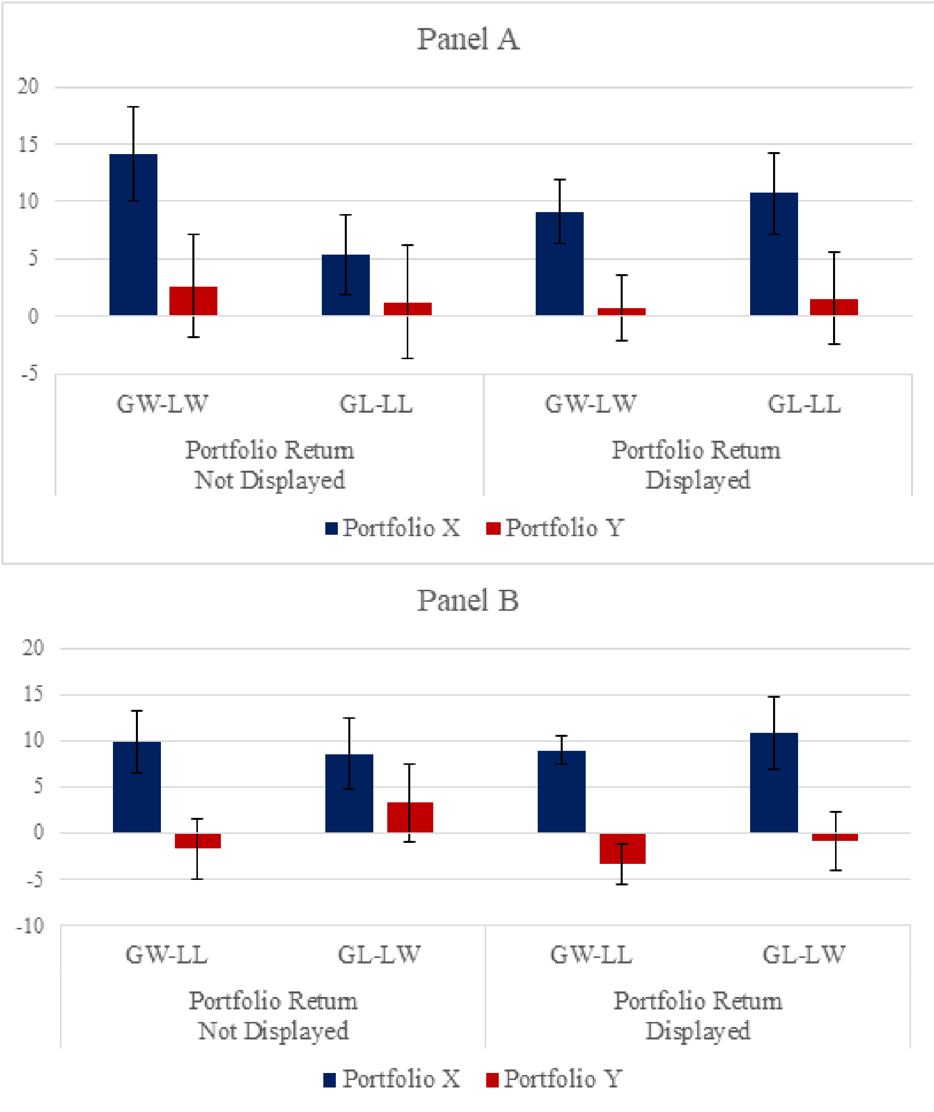
Note: The figure shows participants' mean risk perception for each portfolio elicited on a Likert scale from 1: low to 7: high for the portfolio pair $G_pW_S - G_pL_S$. The blue bar refers to Portfolio X which corresponds to the first two letters of the portfolio pair (e.g. G_pW_S for the first portfolio pair) and the red bar refers to Portfolio Y which corresponds to the second two letters of the portfolio pair (e.g. G_pL_S for the first portfolio pair). Displayed are 95% confidence intervals.

Figure B.16: Satisfaction in Experiment 1 (Additional Treatments)



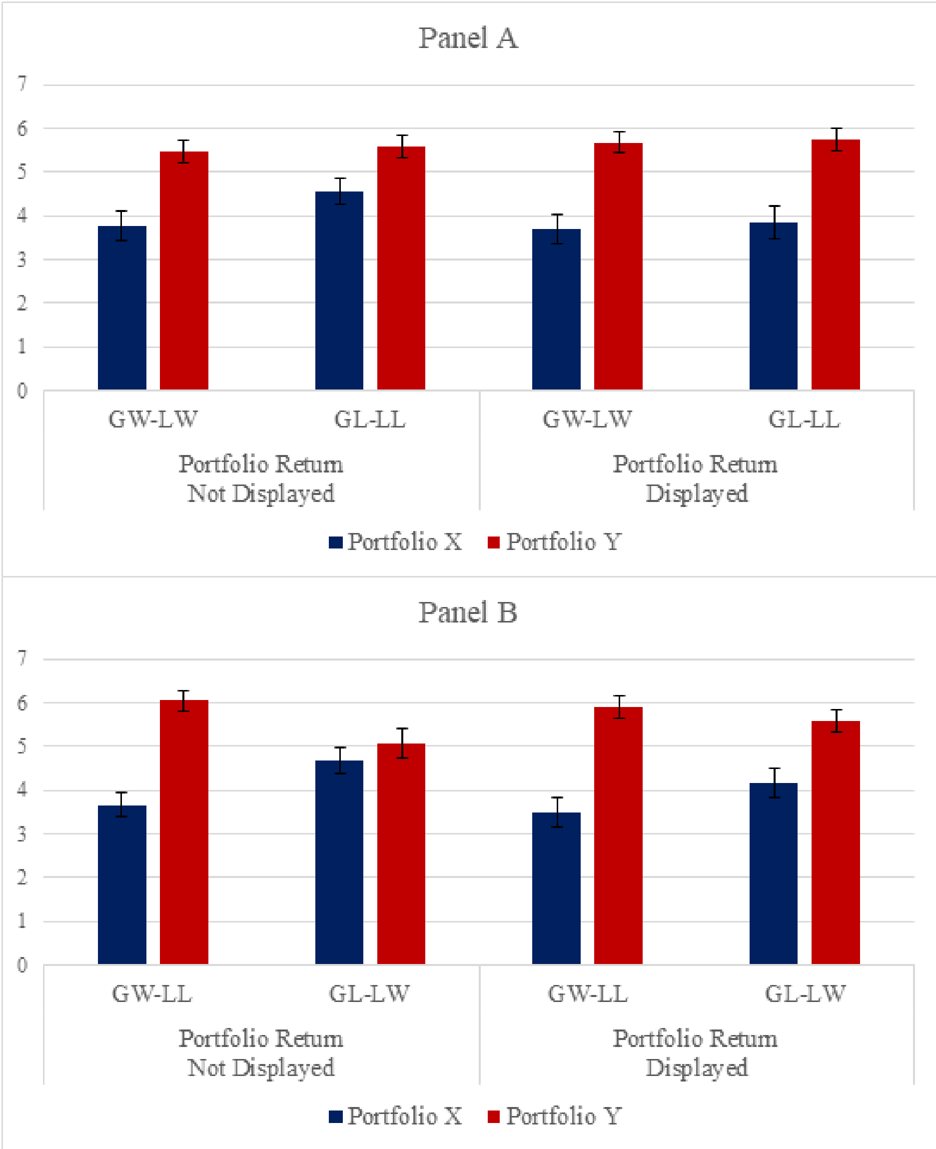
Note: The figure shows participants' mean satisfaction levels for each portfolio elicited on a Likert scale from 1: low to 7: high for the four portfolio pairs $G_pW_S - L_pW_S$ and $G_pL_S - L_pL_S$ (Panel A) and $G_pW_S - L_pL_S$ and $G_pL_S - L_pW_S$ (Panel B). The blue bars refer to Portfolio X which corresponds to the first two letters of each portfolio pair (e.g. GL for the first portfolio pair) and the red bars refer to Portfolio Y which corresponds to the second two letters of each portfolio pair (e.g. LL for the first portfolio pair). Displayed are 95%-confidence intervals.

Figure B.17: Return Expectations in Experiment 1 (Additional Treatments)



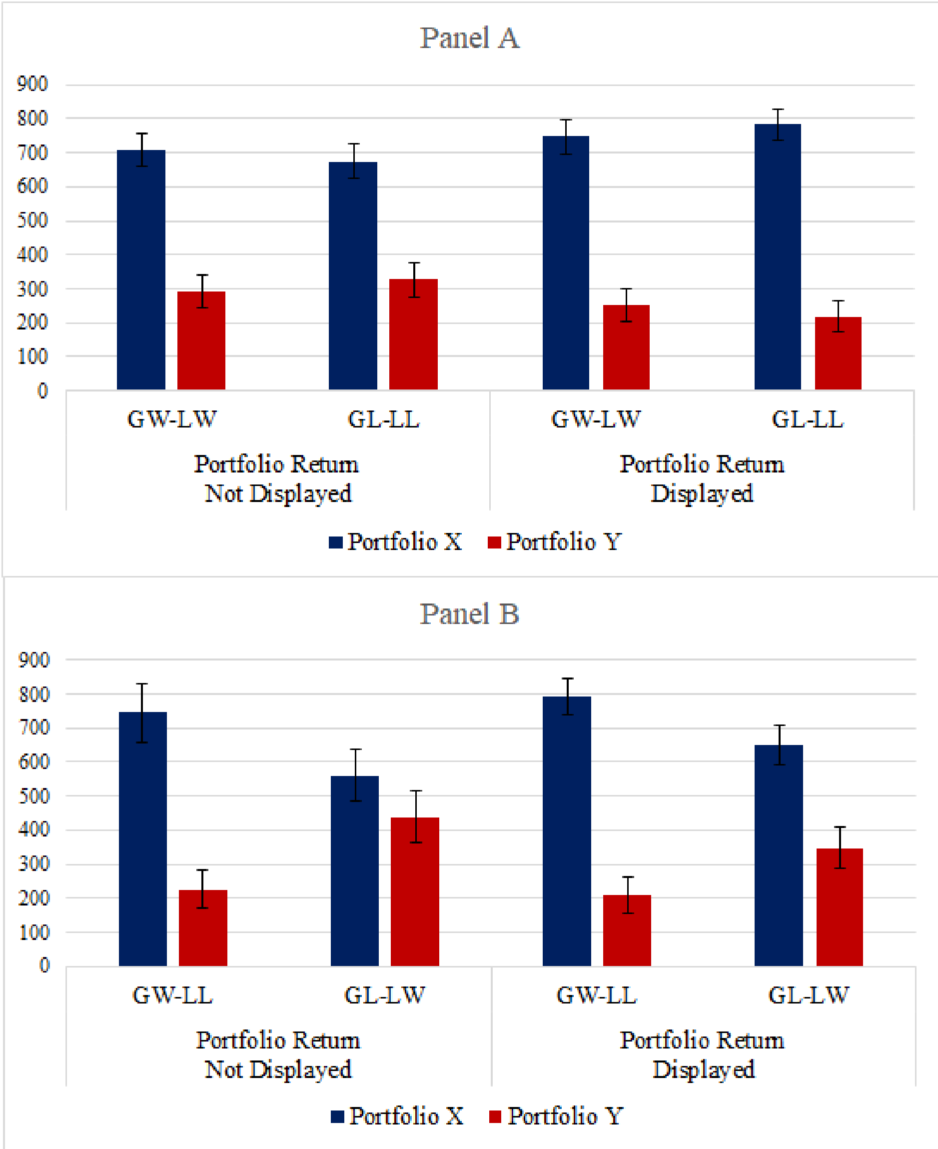
Note: The figure shows participants' mean expected returns for each portfolio for the four portfolio pairs $G_pW_S - L_pW_S$ and $G_pL_S - L_pL_S$ (Panel A) and $G_pW_S - L_pL_S$ and $G_pL_S - L_pW_S$ (Panel B). The blue bars refer to Portfolio X which corresponds to the first two letters of each portfolio pair (e.g. GL for the first portfolio pair) and the red bars refer to Portfolio Y which corresponds to the second two letters of each portfolio pair (e.g. LL for the first portfolio pair). Displayed are 95%-confidence intervals.

Figure B.18: Risk Perception in Experiment 1 (Additional Treatments)



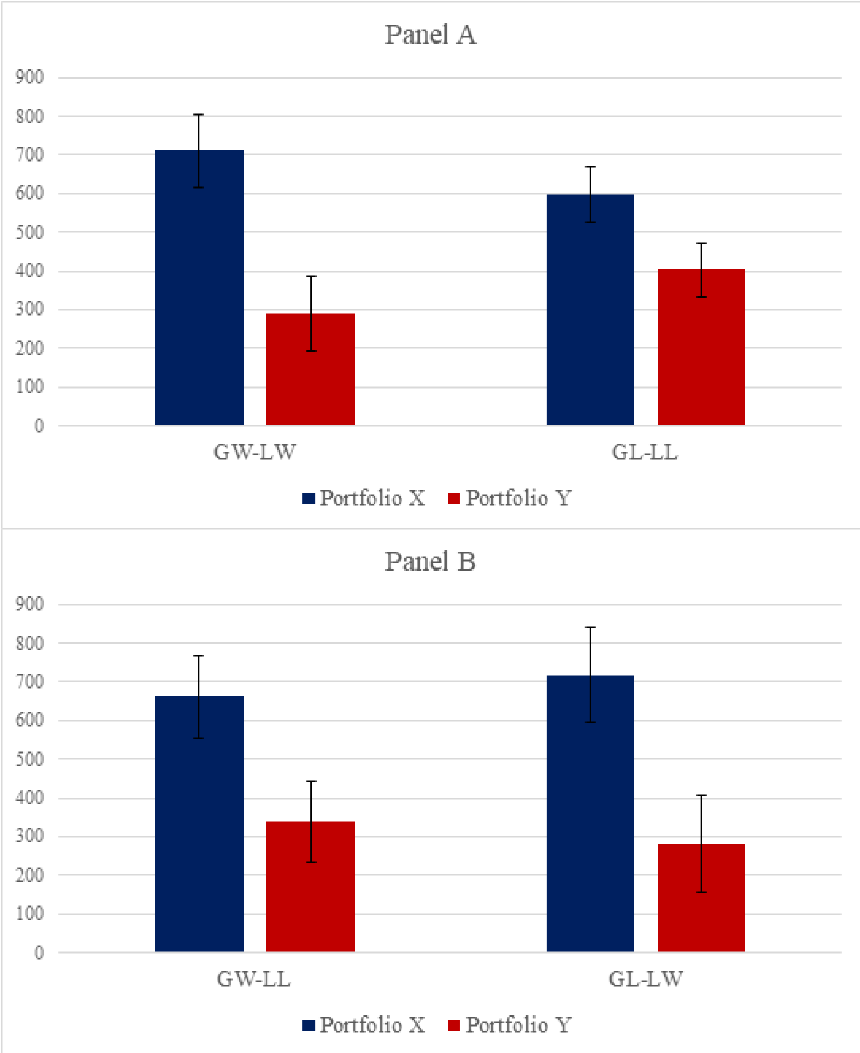
Note: The figure shows participants' mean risk perception for each portfolio elicited on a Likert scale from 1: low to 7: high for the four portfolio pairs $G_pW_S - L_pW_S$ and $G_pL_S - L_pL_S$ (Panel A) and $G_pW_S - L_pL_S$ and $G_pL_S - L_pW_S$ (Panel B). The blue bars refer to Portfolio X which corresponds to the first two letters of each portfolio pair (e.g. GL for the first portfolio pair) and the red bars refer to Portfolio Y which corresponds to the second two letters of each portfolio pair (e.g. LL for the first portfolio pair). Displayed are 95%-confidence intervals.

Figure B.19: Investment in Experiment 2 (Additional Treatments)



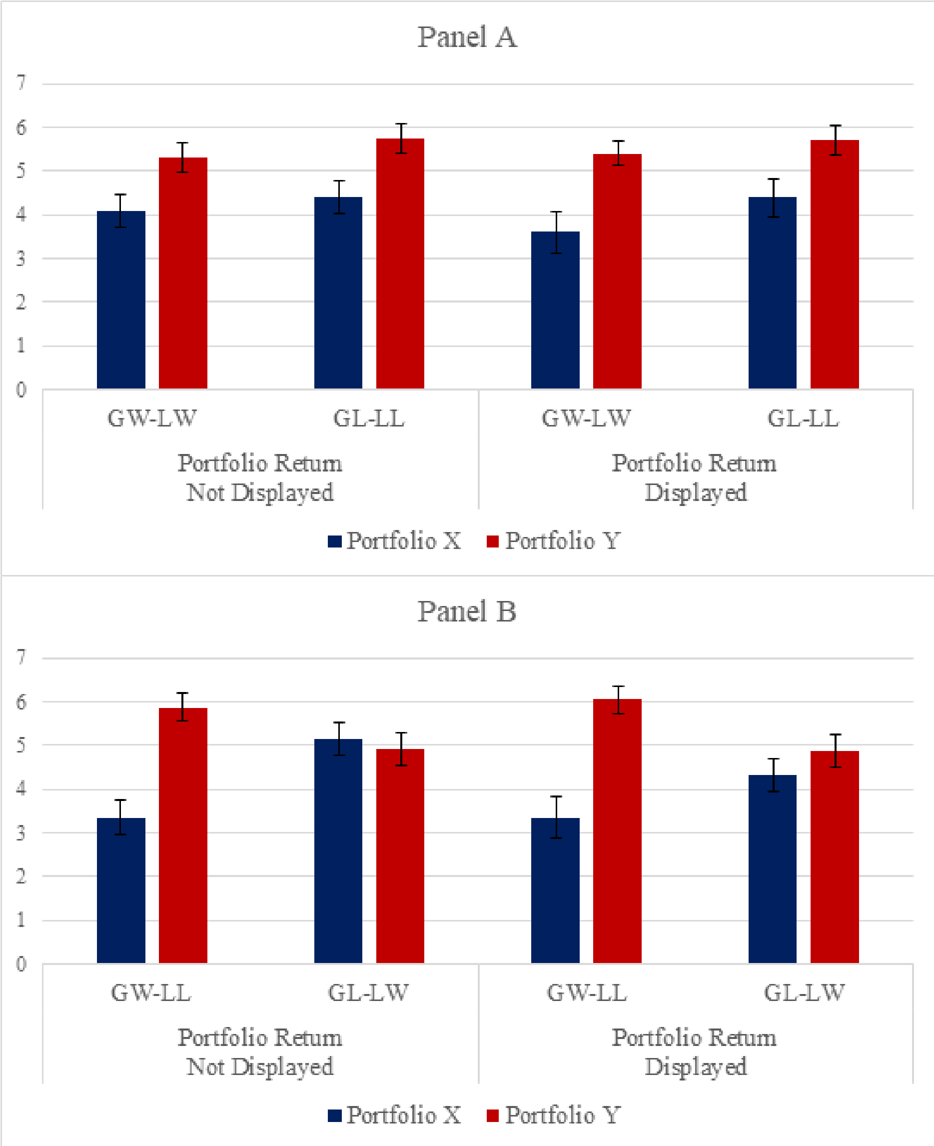
Note: The figure shows participants' mean investments in US dollar in each portfolio for the four portfolio pairs $G_pW_S - L_pW_S$ and $G_pL_S - L_pL_S$ (Panel A) and $G_pW_S - L_pL_S$ and $G_pL_S - L_pW_S$ (Panel B). The blue bars refer to Portfolio X which corresponds to the first two letters of each portfolio pair (e.g. G_pW_S for the first portfolio pair) and the red bars refer to Portfolio Y which corresponds to the second two letters of each portfolio pair (e.g. L_pW_S for the first portfolio pair). Displayed are 95%confidence intervals.

Figure B.20: Investment in Experiment 2 Conditional on Return Expectations (Additional Treatments)



Note: The figure shows participants' mean investments in US dollar in each portfolio of those participants who state the same expected returns for the two portfolios of a pair. The portfolio pairs are $G_pW_S - L_pW_S$ and $G_pL_S - L_pL_S$ (Panel A) and $G_pW_S - L_pL_S$ and $G_pL_S - L_p$ (Panel B). The blue bars refer to Portfolio X which corresponds to the first two letters of each portfolio pair (e.g. GW for the first portfolio pair) and the red bars refer to Portfolio Y which corresponds to the second two letters of each portfolio pair (e.g. LW for the first portfolio pair). Displayed are 95%-confidence intervals.

Figure B.21: Risk Perception in Experiment 2 (Additional Treatments)



Note: The figure shows participants' mean risk perception for each portfolio elicited on a Likert scale from 1: low to 7: high for the four portfolio pairs $G_pW_S - L_pW_S$ and $G_pL_S - L_pL_S$ (Panel A) and $G_pW_S - L_pL_S$ and $G_pL_S - L_pW_S$ (Panel B). The blue bars refer to Portfolio X which corresponds to the first two letters of each portfolio pair (e.g. GL for the first portfolio pair) and the red bars refer to Portfolio Y which corresponds to the second two letters of each portfolio pair (e.g. LL for the first portfolio pair). Displayed are 95%-confidence intervals.

Table B.2: Satisfaction with Portfolios in Experiment 1

Dependent Variable	*Satisfaction*			
	Entire sample	Portfolio return displayed	Portfolio return not displayed	Entire sample
Gain	1.860***	2.455***	1.264***	1.264***
	(0.103)	(0.142)	(0.139)	(0.139)
Winner	0.645***	0.446***	0.843***	0.843***
	(0.0996)	(0.141)	(0.140)	(0.140)
Gain x Winner	0.264**	0.0950	0.434**	0.434**
	(0.124)	(0.170)	(0.172)	(0.171)
Display				−0.124
				(0.156)
Display x Gain				1.190***
				(0.199)
Display x Winner				−0.397**
				(0.199)
Display x Gain x Winner				−0.339
				(0.241)
Constant	2.751***	2.653***	2.777***	2.777***
	(0.0782)	0.112)	(0.109)	(0.109)
Observations	1,936	968	968	1,936
R^2	0.318	0.408	0.263	0.345

Note: The table shows the coefficients of OLS regressions of satisfaction on a gain dummy variable (1 if portfolio trades at a gain), a winner dummy variable (1 if portfolio has more winner than loser assets), the interaction term of gain and winner, a display dummy variable (1 if total portfolio return is displayed) and multiple interaction terms of the display, gain and winner dummy variable. Regression (1) and (2) are run with data from experiment 1, regression (3) and (4) are run with data from experiment 2. We cluster standard errors on the individual investor level and on the portfolio pair level, standard errors are reported in parentheses, and *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

Table B.3: Risk Perception in Experiment 1

Dependent Variable	*Risk Perception*			
	Entire sample	Portfolio return displayed	Portfolio return not displayed	Entire sample
Gain	−1.184***	−1.397***	−0.971***	−0.971***
	(0.0870)	(0.130)	(0.114)	(0.114)
Winner	−0.386***	−0.256**	−0.517***	−0.517***
	(0.0766)	(0.102)	(0.114)	(0.114)
Gain x Winner	−0.355***	−0.314*	−0.397**	−0.397**
	(0.117)	(0.167)	(0.163)	(0.163)
Display				0.0537
				(0.113)
Display x Gain				−0.426**
				(0.173)
Display x Winner				0.260*
				(0.153)
Display x Gain x Winner				0.0826
				(0.233)
Constant	5.676***	5.702***	5.649***	5.649***
	(0.0565)	(0.0810)	(0.0789)	(0.0789)
Observations	1,936	968	968	1,936
R^2	0.221	0.246	0.206	0.228

Note: The table shows the coefficients of OLS regressions of risk perception on a gain dummy variable (1 if portfolio trades at a gain), a winner dummy variable (1 if portfolio has more winner than loser assets), the interaction term of gain and winner, a display dummy variable (1 if total portfolio return is displayed) and multiple interaction terms of the display, gain and winner dummy variable. Regression (1) and (2) are run with data from experiment 1, regression (3) and (4) are run with data from experiment 2. We cluster standard errors on the individual investor level and on the portfolio pair level, standard errors are reported in parentheses, and *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

Table B.4: Return Expectations in Experiment 1

Dependent Variable	*Return Expectations*			
	Entire sample	Portfolio return displayed	Portfolio return not displayed	Entire sample
Gain	7.068***	8.926***	5.188***	5.188***
	(1.086)	(1.371)	(1.685)	(1.685)
Winner	2.654**	1.210	4.116**	4.116**
	(1.098)	(1.409)	(1.677)	(1.677)
Gain x Winner	0.142	−1.024	1.380	1.380
	(1.359)	(1.630)	(2.180)	(2.180)
Display				−1.338
				(1.827)
Display x Gain				3.738*
				(2.172)
Display x Winner				−2.906
				(2.190)
Display x Gain x Winner				−2.404
				(2.721)
Constant	1.000	0.333	1.671	1.671
	(0.913)	(1.203)	(1.375)	(1.375)
Observations	1,533	744	759	1,533
R^2	0.055	0.088	0.044	0.063

Note: The table shows the coefficients of OLS regressions of return expectations on a gain dummy variable (1 if portfolio trades at a gain), a winner dummy variable (1 if portfolio has more winner than loser assets), the interaction term of gain and winner, a display dummy variable (1 if total portfolio return is displayed) and multiple interaction terms of the display, gain and winner dummy variable. Regression (1) and (2) are run with data from experiment 1, regression (3) and (4) are run with data from experiment 2. We cluster standard errors on the individual investor level and on the portfolio pair level, standard errors are reported in parentheses, and *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

Appendix C

Why So Negative? Belief Formation and Risk-Taking in Boom and Bust Markets

C.1 Further Analyses

In this section we present results of further analyses.

Table C.1: Pessimism Bias Split by Forecasting Quality

Dependent Variable	*Probability Estimate (Subjective Posterior)*					
	Pooled Data		**Domain-specific**		**Mixed**	
	Above Median	Below Median	Above Median	Below Median	Above Median	Below Median
Bust	-4.529***	-6.813***	-4.261***	-7.247***	-4.997***	-5.661***
	(-6.13)	(-4.54)	(-3.72)	(-3.11)	(-5.18)	(-2.86)
Objective Posterior	0.671***	0.133***	0.641***	0.165***	0.693***	0.107***
	(48.14)	(7.46)	(34.13)	(6.33)	(35.46)	(4.37)
Constant	20.92***	58.92***	14.88***	66.86***	27.49***	50.78***
	(6.75)	(9.62)	(3.22)	(6.82)	(6.38)	(6.60)
Observations	6032	6016	2704	2896	3328	3120
R^2	0.69	0.10	0.68	0.08	0.70	0.12

Note: This table reports the results of three OLS regressions on how subjective posterior beliefs about the distribution of the lottery depend on the treatment split by above and below median forecasting ability as defined in the text. We report the results of OLS regressions for the whole sample, and for each experiment individually (Experiment 1: Domain-specific; Experiment 2: Mixed). The dependent variable *Probability Estimate* is the subjective posterior belief that the asset is paying from the good distribution. Independent variables include the *Bust* dummy, an indicator variable that equals 1 if participants were in the bust treatment and zero otherwise, as well as *Objective Posterior*, which is the correct Bayesian probability that the stock is good, given the information seen by the participant up to trial t in the learning block. Controls include age, gender, statistical skills, self-reported experience in stock trading, whether subjects were invested in the stock market during the last financial crisis, and the order of outcomes they experienced in the forecasting task. Reported are coefficients and t-statistics (in parentheses) using robust standard errors. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

C.2 Experimental Instructions and Screenshots

Instructions Bayesian Updating (Exemplary for Boom Treatment of Experiment 1)

In this part, we would like to test your forecasting abilities. You will make forecasting decisions in two consecutive blocks each consisting of 8 rounds. Suppose you find yourself in an environment, in which the value of a risky asset can either increase by 2 or by 15. The probability of either outcome (2 or 15) depends on the state in which the asset is (**good** state or **bad** state). If the risky asset is in the **good** state, then the probability that the risky asset increases in value by 15 is 70% and the probability that it increases in value by 2 is 30%. If the risky asset is in the **bad** state, then the probability that the risky asset increases in value by 15 is 30% and the probability that it increases in value by 2 is 70%.
The computer determines the state at the beginning of each block (consisting of 8 rounds). Within a block, the state does not change and remains fixed. At the beginning of each block, you do not know which state the risky asset is in. The risky asset may be in the good state or in the bad state with equal probability.
At the beginning of each round, you will observe the payoff of the risky asset (2 or 15). After that, we will ask you to provide a probability estimate that the risky asset is in the good state and ask you how sure you are about your probability estimate. While answering these questions, you can observe the price development in a chart next to the question.
There is always an objective correct probability that the risky asset is in the good state. This probability depends on the history of payoffs of the risky asset already. As you observe the payoffs of the risky asset, you will update your beliefs whether or not the risky asset is in the good state.
Every time you provide us with a probability estimate that is within 5% of the correct value (e.g., correct probability is 70% and your answer is between 65% and 75%) we will add 10 Cents to your payment.

Objective Bayesian Posterior Probabilities

This table provides all possible values for the objectively correct probability that the asset is in the good state for every possible combination of trials and outcomes. The initial prior for good and bad distribution is set to 50%. The objective Bayesian posterior probability that the asset is in the good state, after observing t high outcomes in n trials so far is given by: $\frac{1}{1+\frac{1-p}{p}\cdot(\frac{q}{1-q})^{n-2t}}$, where p is the initial prior before any outcome is observed that the stock is in the good state (50% here), and q is the probability that the value increase of the asset is the higher one (70% here).

n (number of trials so far)	t (number of high outcomes so far)	Probability [stock is good t high outcomes in n trials]
0	0	50.00%
1	0	30.00%
1	1	70.00%
2	0	15.52%
2	1	50.00%
2	2	84.48%
3	0	7.30%
3	1	30.00%
3	2	70.00%
3	3	92.70%
4	0	3.26%
4	1	15.52%
4	2	50.00%
4	3	84.48%
4	4	96.74%
5	0	1.43%
5	1	7.30%
5	2	30.00%
5	3	70.00%
5	4	92.70%
5	5	98.57%
6	0	0.62%
6	1	3.26%
6	2	15.52%
6	3	50.00%
6	4	84.48%
6	5	96.74%
6	6	99.38%
7	0	0.26%
7	1	1.43%
7	2	7.30%
7	3	30.00%
7	4	70.00%
7	5	92.70%
7	6	98.57%
7	7	99.74%
8	0	0.11%
8	1	0.62%
8	2	3.26%
8	3	15.52%
8	4	50.00%
8	5	84.48%
8	6	96.74%
8	7	99.38%
8	8	99.89%

Screenshots of Experiment 1

Figures C.1 to C.3 present the screens of the forecasting task as seen by subjects in the experiment (example block 1, round 5). One round consists of three sequential screens. First, subjects saw the payoff of the risky asset in the respective round. Second, the cumulated payoffs of the risky asset are shown in a price-line-chart and subjects are asked to provide a probability estimate that the risky asset pays from the good distribution. Finally, subjects are asked on a 9-point Likert scale how confident they are in their probability estimate.

Figure C.1: Payoff Screen

Figure C.2: Probability Estimate Screen

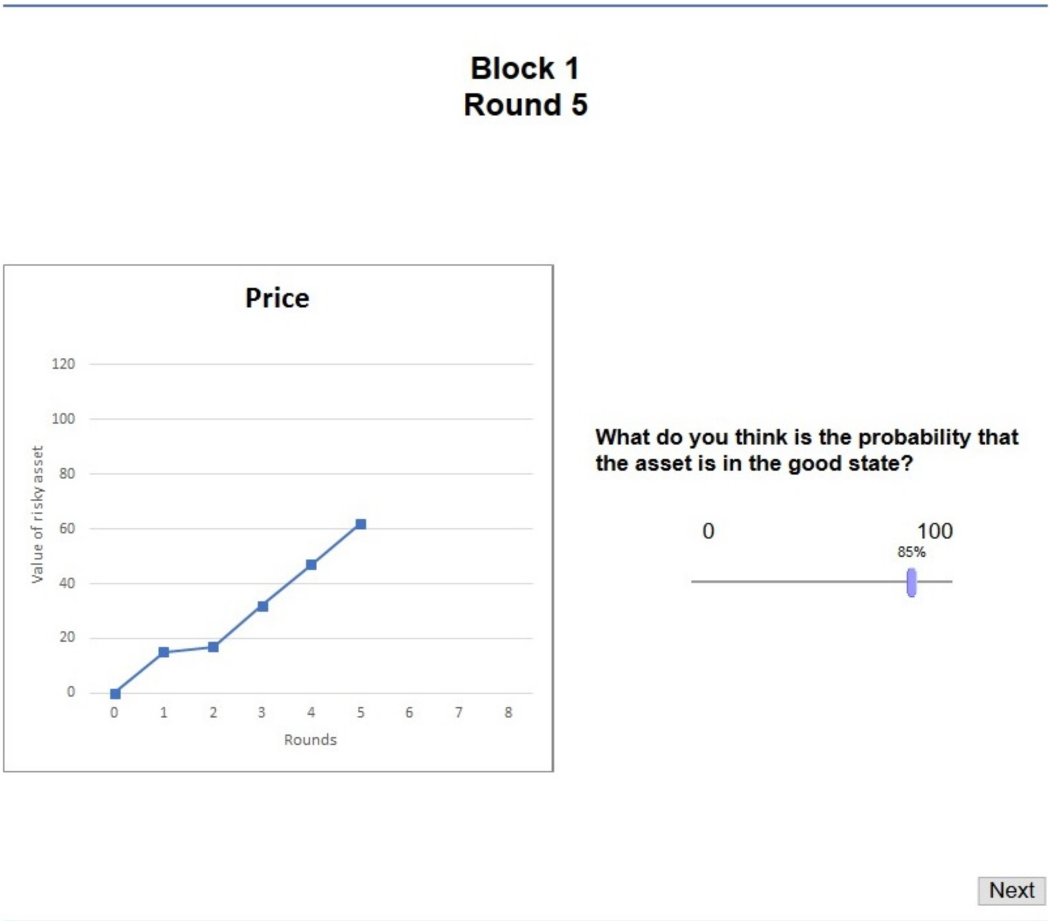
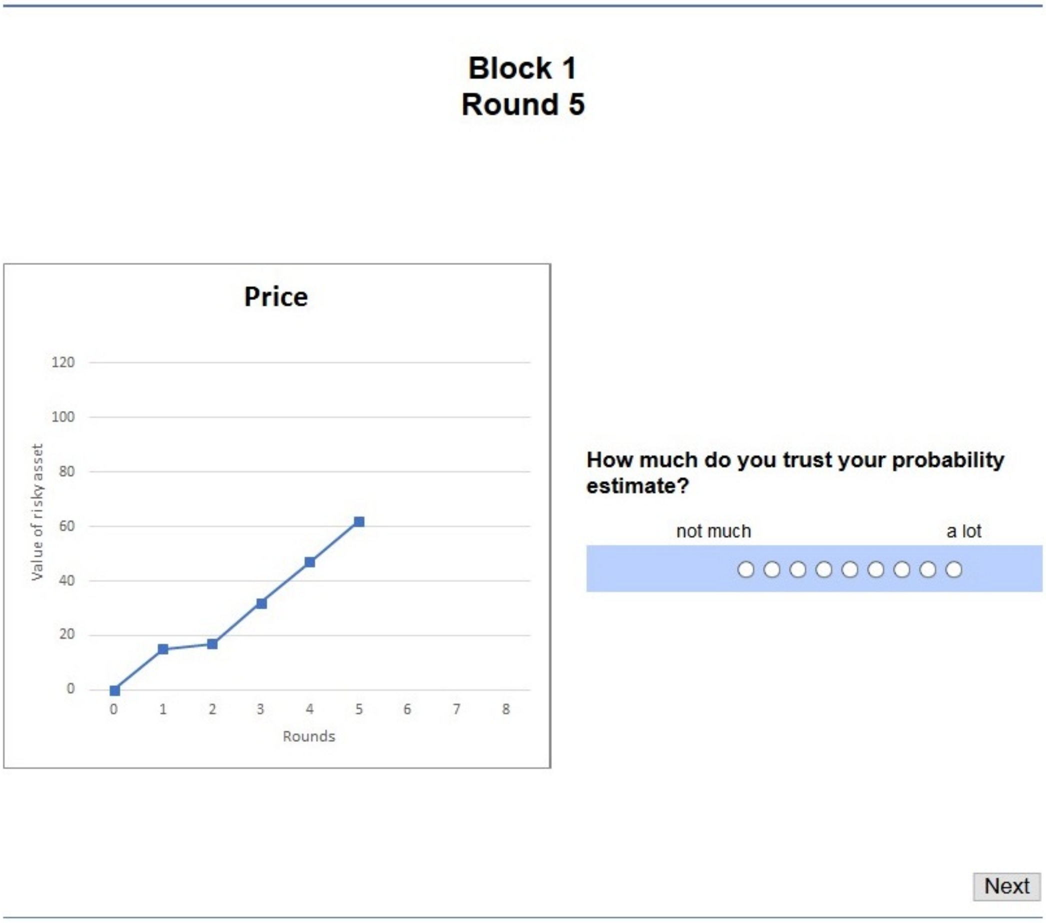


Figure C.3: Confidence Level Screen



C.3 Experimental Measures

Risky Lottery

Imagine in the stock market there is a risky asset, in which you can invest 100 Cent now. The asset pays you either 2.5 times the amount you invest or it becomes valueless, i.e. your invested amount is lost. **The probability of either outcome is exactly 50%.**

You can keep whatever amount you decide not to invest in the risky asset.

How much of your endowment do you want to invest in the risky asset?

[Dropdown Menu of all possible combinations in 5 Cent steps]

Ambiguous Lottery

Imagine in the stock market there is a risky asset, in which you can invest 100 Cent now. The asset pays you either 2.5 times the amount you invest or it becomes valueless, i.e. your invested amount is lost. **However, the probability of either outcome is unknown.**
You can keep whatever amount you decide not to invest in the risky asset.
How much of your endowment do you want to invest in the risky asset?

[Dropdown Menu of all possible combinations in 5 Cent steps]

Life Orientation Test

Below we report the questions used in the revised version of the Life Orientation Test developed by Scheier et al. (1994). All questions were answered on a 5-point Likert scale from "do not agree at all" to "fully agree". Reverse-coded items are indicated by [R]. Filler-items are indicated by [F]. The non-filler items were added to a final score.

1. In uncertain times, I usually expect the best.
2. It's easy for me to relax. [F]
3. If something can go wrong, it will. [R]
4. I'm always optimistic about my future.
5. I enjoy my friends a lot. [F]
6. It's important for me to keep busy. [F]
7. I hardly ever expect things to go my way. [R]
8. I don't get upset too easily. [F]
9. I rarely count on good things happening to me. [R]
10. Overall, I expect more good things to happen to me than bad.

Comprehension Questions for Bayesian Updating Task

Below we report the comprehension questions that participants had to answer correctly after reading the instructions to proceed to the Bayesian Updating task. Correct responses are displayed in italic.

1. If you see a series of +15 [−2 for Bust treatment], what is more likely?
   (a) *The risky asset is in the good state.*
   (b) The risky asset is in the bad state.

2. The correct probability estimate is let's say 0.70. Which probability estimate(s) would be in the range such that you earn 10 cents? [Note: You can check multiple boxes.]
   (a) 0.55
   (b) *0.67*
   (c) *0.75*
   (d) 0.85
   (e) 0.87

3. At the beginning of each block, the probability that the risky asset is in the good state is 50%.
   (a) *True*
   (b) False

Dow Jones Return Expectations Question in Experiment 1

The Dow Jones Industrial Average (Stock Market Index of the 30 largest US companies) is currently trading at around 25,343.
In which price range would you expect this index to trade in 6 months from now? [Dropdown]

- $<$ 23,000
- 23,000 - 23,500
- 23,501 - 24,000
- 24,001 - 24,500
- 24,501 - 25,000
- 25,001 - 25,500
- 25,501 - 26,000
- 26,001 - 26,500
- 26,501 - 27,000
- 27,001 - 27,500
- 27,501 - 28,000
- $>$ 28,000

Dow Jones Return Expectations Question in Experiment 2

The Dow Jones Industrial Average (Stock Market Index of the 30 largest US companies) is currently trading at around 26,770.
In which price range would you expect this index to trade in 6 months from now? [Dropdown]

- $<$ 24,500
- 23,000 - 23,500
- 24,500 - 25,000
- 25,001 - 25,500
- 25,501 - 26,000
- 26,001 - 26,500
- 26,501 - 27,000
- 27,001 - 27,500
- 27,501 - 28,000
- 28,001 - 28,500
- 28,501 - 29,000
- 29,001 - 29,500
- $>$ 29,500

Appendix D

Can Agents Add and Subtract When Forming Beliefs?

D.1 Experimental Instructions and Screenshots

Instructions Bayesian Updating (Exemplary for Experiment 1)

In this part we would like to test your forecasting abilities. You will make forecasting decisions in one block consisting of 6 rounds.

Suppose you find yourself in an environment, in which a risky asset with an initial value of 50 can either increase by 5 or decrease by 5. The probability of either outcome (5 or -5) depends on the state in which the asset is (**good** state or **bad** state). If the risky asset is in the **good** state, then the probability that the risky asset increases in value by 5 is 70% and the probability that it decreases in value by 5 is 30%. If the risky asset is in the **bad** state, then the probability that the risky asset increases in value by 5 is 30% and the probability that it decreases in value by 5 is 70%.

The computer determines the state at the beginning of the block (consisting of 6 rounds). Within a block, the state does not change and remains fixed.

At the beginning of the block, you do not know which state the risky asset is in. The risky asset may be in the good state or in the bad state with equal probability.

At the beginning of each round, you will observe the payoff of the risky asset (5 or -5). After that, we will ask you to provide a probability estimate that the risky asset is in the good state and ask you how sure you are about your probability estimate. While answering these questions, you can observe the price development in a chart next to the question.

There is always an objective correct probability that the risky asset is in the good state. This probability depends on the history of payoffs of the risky asset already. As you observe the payoffs of the risky asset, you will update your beliefs whether or not the risky asset is in the good state.

Objective Bayesian Posterior Probabilities

This table provides all possible values for the objectively correct probability that the asset is in the good state for every possible combination of trials and outcomes. The initial prior for good and bad distribution is set to 50%. The objective Bayesian posterior probability that the asset is in the good state, after observing t high outcomes in n trials so far is given by:

$$\frac{1}{1+\frac{1-p}{p}\cdot\left(\frac{q}{1-q}\right)^{n-2t}},$$

where p is the initial prior before any outcome is observed that the stock is in the good state (50% here), and q is the probability that the value increase of the asset is the higher one (70% in Experiment 1 & 3, and 60% in Experiment 2).

		Experiment 1 and 3 (q = 70%)	Experiment 2 (q = 60 %)
n (number of trials so far)	t (number of high outcomes so far)	Probability [stock is good t high outcomes in n trials]	Probability [stock is good t high outcomes in n trials]
0	0	50.00%	50.00%
1	0	30.00%	40.00%
1	1	70.00%	60.00%
2	0	15.52%	30.77%
2	1	50.00%	50.00%
2	2	84.48%	69.23%
3	0	7.30%	22.86%
3	1	30.00%	40.00%
3	2	70.00%	60.00%
3	3	92.70%	77.14%
4	0	3.26%	16.49%
4	1	15.52%	30.77%
4	2	50.00%	50.00%
4	3	84.48%	69.23%
4	4	96.74%	83.51%
5	0	1.43%	11.64%
5	1	7.30%	22.86%
5	2	30.00%	40.00%
5	3	70.00%	60.00%
5	4	92.70%	77.14%
5	5	98.57%	88.36%
6	0	0.62%	8.7%
6	1	3.26%	16.49%
6	2	15.52%	30.77%
6	3	50.00%	50.00%
6	4	84.48%	69.23%
6	5	96.74%	83.51%
6	6	99.38%	91.93%

Screenshots of Experiment 1

Figures D.1 to D.3 present the screens of the forecasting task as seen by subjects in the experiment (example round 4). One round consists of three sequential screens. First, subjects saw the payoff of the risky asset in the respective round. Second, the cumulated payoffs of the risky asset are shown in a price-line-chart and subjects are asked to provide a probability estimate that the risky asset pays from the good distribution. Finally, subjects are asked on a 9-point Likert scale how confident they are in their probability estimate.

Figure D.1: Payoff Screen

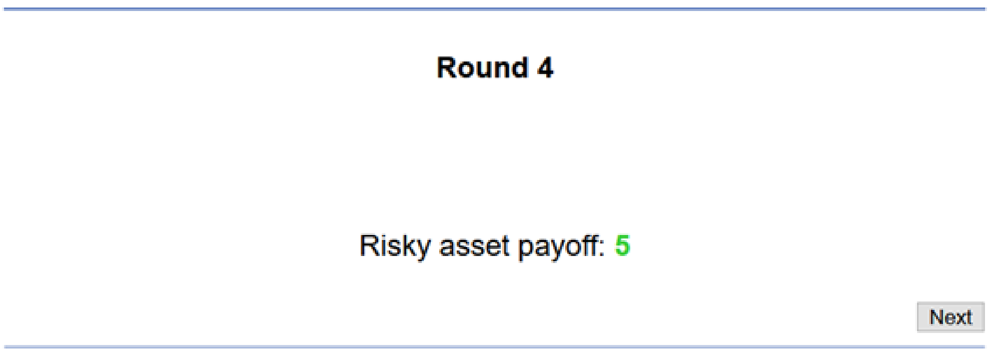
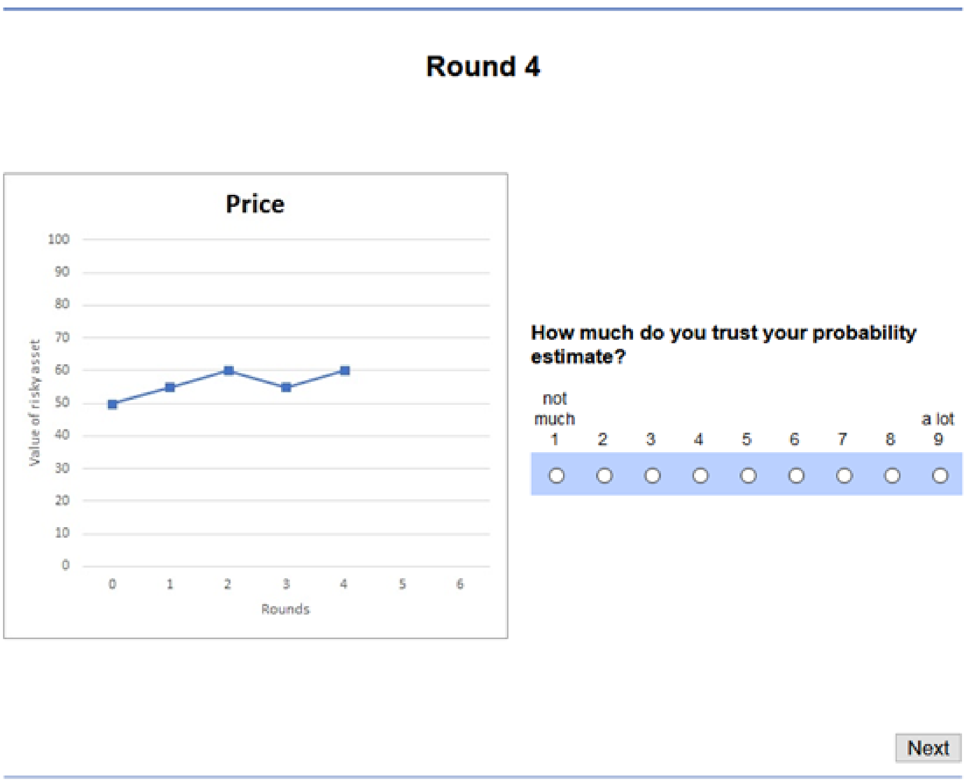


Figure D.2: Probability Estimate Screen

Figure D.3: Confidence Level Screen



Comprehension Questions for Bayesian Updating Task

Below we report the comprehension questions that participants had to answer correctly after reading the instructions to proceed to the Bayesian Updating task. Correct responses are displayed in *italic*.

1. If you see a series of −5, what is more likely?
   (a) *The risky asset is in the good state.*
   (b) The risky asset is in the bad state.

2. You observe a −5, how do you have to update your probability estimate that the asset draws from the good distribution??
   (a) *I reduce the probability estimate that the asset is in the good distribution.*
   (b) In increase the probability estimate that the asset is in the good distribution.

3. The correct probability estimate is let's say 0.70. Which probability estimate(s) would be in the range such that you earn 10 cents? [Note: You can check multiple boxes.]
   (a) 0.55
   (b) *0.67*
   (c) *0.75*
   (d) 0.85
   (e) 0.87

4. At the beginning of each block, the probability that the risky asset is in the good state is 50%.
   (a) *True*
   (b) False

References

Adam, K., Matveev, D., and Nagel, S. (2020). Do survey expectations of stock returns reflect risk adjustments? *Journal of Monetary Economics*. in press.

Alberoni, F. (1962). Contribution to the study of subjective probability. *The Journal of General Psychology*, 66(2):241–264.

Alempaki, D., Starmer, C., and Tufano, F. (2019). On the priming of risk preferences: The role of fear and general affect. *Journal of Economic Psychology*, 75:102–137.

Allport, G. W., Clark, K., and Pettigrew, T. (1954). The nature of prejudice.

Amromin, G. and Sharpe, S. A. (2014). From the horse's mouth: Economic conditions and investor expectations of risk and return. *Management Science*, 60(4):845–866.

An, L., Engelberg, J., Henriksson, M., Wang, B., and Williams, J. (2019). The portfolio-driven disposition effect. Working Paper.

Andrade, E. B. and Iyer, G. (2009). Planned versus actual betting in sequential gambles. *Journal of Marketing Research*, 46(3):372–383.

Anzoni, L. and Zeisberger, S. (2016). What is risk? understanding how investors perceive risk in return distribution. Working Paper.

Arkes, H. R., Joyner, C. A., Pezzo, M. V., Nash, J. G., Siegel-Jacobs, K., and Stone, E. (1994). The psychology of windfall gains. *Organizational Behavior and Human Decision Processes*, 59(3):331–347.

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

Barber, B. M. and 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*, 21:785–818.

Barber, B. M. and Odean, T. (2013). The behavior of individual investors. In *Handbook of the Economics of Finance*, volume 2, pages 1533–1570. Elsevier.

Barberis, N. (2012). A model of casino gambling. *Management Science*, 58(1):35–51.

Barberis, N. (2018). Psychology-based models of asset prices and trading volume. In *Handbook of Behavioral Economics: Applications and Foundations 1*, volume 1, pages 79–175. Elsevier.

Barberis, N., Greenwood, R., Jin, L., and Shleifer, A. (2015). X-capm: An extrapolative capital asset pricing model. *Journal of Financial Economics*, 115(1):1–24.

Barberis, N., Greenwood, R., Jin, L., and Shleifer, A. (2018). Extrapolation and bubbles. *Journal of Financial Economics*, 129(2):203–227.

Barberis, N. and Huang, M. (2001). Mental accounting, loss aversion, and individual stock returns. *The Journal of Finance*, 56(4):1247–1292.

Barberis, N., Huang, M., and Santos, T. (2001). Prospect theory and asset prices. *The Quarterly Journal of Economics*, 116(1):1–53.

Barberis, N. and Shleifer, A. (2003). Style investing. *Journal of Financial Economics*, 68(2):161–199.

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

Barberis, N. and Xiong, W. (2009). What drives the disposition effect? An analysis of a long-standing preference-based explanation. *The Journal of Finance*, 64(2):751–784.

Barberis, N. and Xiong, W. (2012). Realization utility. *Journal of Financial Economics*, 104(2):251–271.

Ben-Rephael, A., Kandel, S., and Wohl, A. (2011). The price pressure of aggregate mutual fund flows. *Journal of Financial and Quantitative Analysis*, 46(2):585–603.

Benartzi, S. and Thaler, R. H. (2001). Naive diversification strategies in defined contribution saving plans. *American Economic Review*, 91(1):79–98.

Benjamin, D. J. (2019). Errors in probabilistic reasoning and judgment biases. In *Handbook of Behavioral Economics: Applications and Foundations 1*, volume 2, pages 69–186. Elsevier.

Bernoulli, D. (1954). Exposition of a new theory on the measurement of risk. *Econometrica*, 22:23–36.

Bhattacharya, U., Loos, B., Meyer, S., and Hackethal, A. (2017). Abusing etfs. *Review of Finance*, 21(3):1217–1250.

Boldin, M. and Cici, G. (2010). The index fund rationality paradox. *Journal of Banking & Finance*, 34(1):33–43.

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

Buhrmester, M., Kwang, T., and Gosling, S. D. (2011). Amazon's mechanical turk: A new source of inexpensive, yet high-quality data? *Perspectives on Psychological Science*, 6(3):3–5.

Calvet, L. E., Campbell, J. Y., and Sodini, P. (2009). Measuring the financial sophistication of households. *American Economic Review*, 99(2):393–98.

Camerer, C. (1995). Individual decision making. *Handbook of Experimental Economics*.

Camerer, C. F. (1987). Do biases in probability judgment matter in markets? experimental evidence. *The American Economic Review*, 77(5):981–997.

Camerer, C. F., Dreber, A., Forsell, E., Ho, T.-H., Huber, J., Johannesson, M., Kirchler, M., Almenberg, J., Altmejd, A., Chan, T., Heikensten, E., Holzmeister, F., Imai, T., Isaksson, S., Nave, G., Pfeiffer, T., Razen, M., and Wu, H. (2016). Evaluating replicability of laboratory experiments in economics. *Science*, 351(6280):1433–1436.

Campbell, J. Y. and Cochrane, J. H. (1999). By force of habit: A consumption-based explanation of aggregate stock market behavior. *Journal of Political Economy*, 107(2):205–251.

Cassella, S. and Gulen, H. (2018). Extrapolation bias and the predictability of stock returns by price-scaled variables. *The Review of Financial Studies*, 31(11):4345–4397.

Chang, T. Y., Solomon, D. H., and Westerfield, M. M. (2016). Looking for someone to blame: Delegation, cognitive dissonance, and the disposition effect. *The Journal of Finance*, 71(1):267–302.

Charness, G. and Dave, C. (2017). Confirmation bias with motivated beliefs. *Games and Economic Behavior*, 104:1–23.

Christensen, G. and Miguel, E. (2018). Transparency, reproducibility, and the credibility of economics research. *Journal of Economic Literature*, 56(3):920–80.

Clark, J. (2002). House money effects in public good experiments. *Experimental Economics*, 5(3):223–231.

Clifford, C. P., Fulkerson, J. A., and Jordan, B. D. (2014). What drives ETF flows? *Financial Review*, 49(3):619–642.

Cohn, A., Engelmann, J., Fehr, E., and Maréchal, M. A. (2015). Evidence for countercyclical risk aversion: An experiment with financial professionals. *American Economic Review*, 105(2):860–85.

Coval, J. and Stafford, E. (2007). Asset fire sales (and purchases) in equity markets. *Journal of Financial Economics*, 86(2):479–512.

Coval, J. D. and Shumway, T. (2005). Do behavioral biases affect prices? *The Journal of Finance*, 60(1):1–34.

Daniel, K., Hirshleifer, D., and Subrahmanyam, A. (1998). Investor psychology and security market under- and overreactions. *The Journal of Finance*, 53(6):1838–1885.

Dawes, R. M. (1979). The robust beauty of improper linear models in decision making. *American Psychologist*, 34(7):571.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., and Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, 9(3):522–550.

DuCharme, W. M. and Peterson, C. R. (1968). Intuitive inference about normally distributed populations. *Journal of Experimental Psychology*, 78(2p1):269.

Edelen, R. M. and Warner, J. B. (2001). Aggregate price effects of institutional trading: a study of mutual fund flow and market returns. *Journal of Financial Economics*, 59(2):195–220.

Eil, D. and Rao, J. M. (2011). The good news-bad news effect: asymmetric processing of objective information about yourself. *American Economic Journal: Microeconomics*, 3(2):114–38.

Elton, E. J., Gruber, M. J., and Busse, J. A. (2004). Are investors rational? Choices among index funds. *The Journal of Finance*, 59:261–288.

Enke, B. and Graeber, T. (2019). Cognitive uncertainty. Working Paper.

Falk, A., Becker, A., Dohmen, T., Huffman, D., and Sunde, U. (2016). The preference survey module: A validated instrument for measuring risk, time, and social preferences. Working Paper.

Feng, L. and Seasholes, M. S. (2005). Do investor sophistication and trading experience eliminate behavioral biases in financial markets? *Review of Finance*, 9(3):305–351.

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

Franken, I. H., Georgieva, I., and Muris, P. (2006). The rich get richer and the poor get poorer: On risk aversion in behavioral decision-making. *Judgment and Decision Making*, 1(2):153–158.

Frino, A., Grant, J., and Johnstone, D. (2008). The house money effect and local traders on the Sydney Futures Exchange. *Pacific-Basin Finance Journal*, 16(1-2):8–25.

Frydman, C., Barberis, N., Camerer, C., Bossaerts, P., and Rangel, A. (2014). Using neural data to test a theory of investor behavior: An application to realization utility. *The Journal of Finance*, 69(2):907–946.

Frydman, C., Hartzmark, S. M., and Solomon, D. H. (2017). Rolling mental accounts. *The Review of Financial Studies*, 31(1):362–397.

Gächter, S., Johnson, E. J., and Herrmann, A. (2007). Individual-level loss aversion in riskless and risky choices. Working Paper.

Garrett, T. A. and Sobel, R. S. (1999). Gamblers favor skewness, not risk: Further evidence from united states' lottery games. *Economics Letters*, 63(1):85–90.

Geller, E. S. and Pitz, G. F. (1968). Confidence and decision speed in the revision of opinion. *Organizational Behavior and Human Performance*, 3(2):190–201.

Genesove, D. and Mayer, C. (2001). Loss aversion and seller behavior: Evidence from the housing market. *The Quarterly Journal of Economics*, 116(4):1233–1260.

Gennaioli, N. and Shleifer, A. (2010). What comes to mind. *The Quarterly Journal of Economics*, 125(4):1399–1433.

Gennaioli, N., Shleifer, A., and Vishny, R. (2012). Neglected risks, financial innovation, and financial fragility. *Journal of Financial Economics*, 104(3):452–468.

Gennaioli, N., Shleifer, A., and Vishny, R. (2015). Neglected risks: The psychology of financial crises. *American Economic Review*, 105(5):310–14.

Giglio, S., Maggiori, M., Stroebel, J., and Utkus, S. (2019). Five facts about beliefs and portfolios. Working Paper.

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

Gneezy, U. and Potters, J. (1997). An experiment on risk taking and evaluation periods. *The Quarterly Journal of Economics*, 112(2):631–645.

Golec, J. and Tamarkin, M. (1998). Bettors love skewness, not risk, at the horse track. *Journal of Political Economy*, 106(1):205–225.

Goodman, J. K., Cryder, C. E., and Cheema, A. (2013). Data collection in a flat world: The strengths and weaknesses of mechanical turk samples. *Journal of Behavioral Decision Making*, 26(3):213–224.

Graham, J. R. and Narasimhan, K. (2004). Corporate survival and managerial experiences during the great depression. Working Paper.

Greenwood, R. and Shleifer, A. (2014). Expectations of returns and expected returns. *The Review of Financial Studies*, 27(3):714–746.

Greiner, B. (2015). Subject pool recruitment procedures: Organizing experiments with ORSEE. *Journal of the Economic Science Association*, 1(1):114–125.

Grether, D. M. (1980). Bayes rule as a descriptive model: The representativeness heuristic. *The Quarterly Journal of Economics*, 95(3):537–557.

Grinblatt, M. and 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.

Grossman, P. J. and Eckel, C. C. (2015). Loving the long shot: Risk taking with skewed lotteries. *Journal of Risk and Uncertainty*, 51(3):195–217.

Grossman, S. J. and Shiller, R. J. (1981). The determinants of the variability of stock market prices. *American Economic Review*, 71(2):222–227.

Gruber, M. J. (1996). Another puzzle: The growth in actively managed mutual funds. *The Journal of Finance*, 51:783–810.

Guiso, L., Sapienza, P., and Zingales, L. (2018). Time varying risk aversion. *Journal of Financial Economics*, 128(3):403–421.

Haliassos, M. and Bertaut, C. C. (1995). Why do so few hold stocks? *The Economic Journal*, 105(432):1110–1129.

Hartzmark, S. M. (2015). The worst, the best, ignoring all the rest: The rank effect and trading behavior. *The Review of Financial Studies*, 28(4):1024–1059.

Hartzmark, S. M. and Solomon, D. H. (2012). Efficiency and the disposition effect in nfl prediction markets. *The Quarterly Journal of Finance*, 2(03):1250013.

Heath, C., Huddart, S., and Lang, M. (1999). Psychological factors and stock option exercise. *The Quarterly Journal of Economics*, 114(2):601–627.

Heath, C. and Soll, J. B. (1996). Mental budgeting and consumer decisions. *Journal of Consumer Research*, 23(1):40–52.

Henderson, P. W. and Peterson, R. A. (1992). Mental accounting and categorization. *Organizational Behavior and Human Decision Processes*, 51(1):92–117.

Holzmeister, F., Huber, J., Kirchler, M., Lindner, F., Weitzel, U., and Zeisberger, S. (2020). What drives risk perception? a global survey with financial professionals and laypeople. *Management Science*. in press.

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

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

Hsee, C. K. (1996). The evaluability hypothesis: An explanation for preference reversals between joint and separate evaluations of alternatives. *Organizational Behavior and Human Decision Processes*, 67(3).

Hsu, Y.-L. and Chow, E. H. (2013). The house money effect on investment risk taking: Evidence from Taiwan. *Pacific-Basin Finance Journal*, 21(1):1102–1115.

Imas, A. (2016). The realization effect: Risk-taking after realized versus paper losses. *American Economic Review*, 106(8):2086–2109.

Ingersoll, J. E. and Jin, L. J. (2013). Realization utility with reference-dependent preferences. *The Review of Financial Studies*, 26(3):723–767.

Ippolito, R. A. (1992). Consumer reaction to measures of poor quality: Evidence from the mutual fund industry. *The Journal of Law and Economics*, 35(1):45–70.

Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *American Economic Review*, 93(5):1449–1475.

Kahneman, D. and Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, 3(3):430–454.

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47:263–291.

Kalaycıoğlu, S. (2004). Exchange traded fund flows. Working Paper.

Klibanoff, P., Marinacci, M., and Mukerji, S. (2005). A smooth model of decision making under ambiguity. *Econometrica*, 73(6):1849–1892.

König-Kersting, C. and Trautmann, S. T. (2018). Countercyclical risk aversion: Beyond financial professionals. *Journal of Behavioral and Experimental Finance*, 18:94–101.

Kuhnen, C. M. (2015). Asymmetric learning from financial information. *The Journal of Finance*, 70(5):2029–2062.

Kumar, A. (2009). Who gambles in the stock market? *The Journal of Finance*, 64(4):1889–1933.

Langer, T. and Weber, M. (2008). Does commitment or feedback influence myopic loss aversion?: An experimental analysis. *Journal of Economic Behavior & Organization*, 67(3-4):810–819.

Lintner, J. (1965). Security prices, risk, and maximal gains from diversification. *The Journal of Finance*, 20(4):587–615.

Lippi, A., Barbieri, L., Piva, M., and De Bondt, W. (2018). Time-varying risk behavior and prior investment outcomes: Evidence from Italy. *Judgment and Decision Making*, 13(5):471–483.

List, J. A. (2002). Preference reversals of a different kind: The" more is less" phenomenon. *American Economic Review*, 92(5):1636–1643.

Liu, W., Kolari, J. W., and Huang, J. (2019). Return dispersion and the cross-section of stock returns. Working Paper.

Machina, M. J. (1987). Decision-making in the presence of risk. *Science*, 236(4801):537–543.

Malmendier, U. and Nagel, S. (2011). Depression babies: do macroeconomic experiences affect risk taking? *The Quarterly Journal of Economics*, 126(1):373–416.

Malmendier, U. and Nagel, S. (2015). Learning from inflation experiences. *The Quarterly Journal of Economics*, 131(1):53–87.

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

Malmendier, U., Tate, G., and Yan, J. (2011). Overconfidence and early-life experiences: the effect of managerial traits on corporate financial policies. *The Journal of Finance*, 66(5):1687–1733.

Maniadis, Z., Tufano, F., and List, J. A. (2014). One swallow doesn't make a summer: New evidence on anchoring effects. *American Economic Review*, 104(1):277–90.

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

Markowitz, H. (1959). Portfolio selection: efficient diversification of investments.

Massa, M. and Simonov, A. (2005). Behavioral biases and investment. *Review of Finance*, 9(4):483–507.

Meier, P., Flepp, R., Rüdisser, M., and Franck, E. P. (2020). The effect of paper versus realized losses on subsequent risktaking: Field evidence from casino gambling. UZH Business Working Paper No. 385.

Merkle, C. (2020). Financial loss aversion illusion. *Review of Finance*, 24(2):381–413.

Möbius, M. M., Niederle, M., Niehaus, P., and Rosenblat, T. S. (2014). Managing self-confidence. *NBER Working Paper*.

Nagel, S. and Xu, Z. (2019). Asset pricing with fading memory. Working Paper.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, 2(2):175–220.

Nielsen, K. (2019). Dynamic risk preferences under realized and paper outcomes. *Journal of Economic Behavior and Organization*, 161:68–78.

O'Brien, E. (2019). When small signs of change add up: The psychology of tipping points. *Current Directions in Psychological Science*, 29(1):55–62.

O'Brien, E. and Klein, N. (2017). The tipping point of perceived change: Asymmetric thresholds in diagnosing improvement versus decline. *Journal of Personality and Social Psychology*, 112(2):161.

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

Peng, J., Miao, D., and Xiao, W. (2013). Why are gainers more risk seeking. *Judgment and Decision Making*, 8(2):150.

Phillips, L. D. and Edwards, W. (1966). Conservatism in a simple probability inference task. *Journal of Experimental Psychology*, 72(3):346.

Pitz, G. F. (1969). The influence of prior probabilities on information seeking and decision-making. *Organizational Behavior and Human Performance*, 4(3):213–226.

Pitz, G. F., Downing, L., and Reinhold, H. (1967). Sequential effects in the revision of subjective probabilities. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, 21(5):381.

Rabin, M. and Weizsäcker, G. (2009). Narrow bracketing and dominated choices. *American Economic Review*, 99(4):1508–43.

Raghubir, P. and Srivastava, J. (2008). Monopoly money: The effect of payment coupling and form on spending behavior. *Journal of Experimental Psychology: Applied*, 14(3):213–225.

Read, D., Loewenstein, G., and Rabin, M. (1999). Choice bracketing. In Fischhoff, B. and Manski, C. F., editors, *Elicitation of Preferences*, pages 171–202. Springer, New York.

Rieskamp, J. and Hoffrage, U. (1999). When do people use simple heuristics, and how can we tell? In *Simple heuristics that make us smart*, pages 141–167. Oxford University Press.

Roll, R. and Ross, S. A. (1980). An empirical investigation of the arbitrage pricing theory. *The Journal of Finance*, 35(5):1073–1103.

Rosch, E. and Lloyd, B. (1978). Cognition and categorization. *Hillsdale, New Jersey*.

Ross, S. A. (1976). The arbitrage theory of capital asset pricing. pages 341–360.

Sarin, R. K. and Weber, M. (1993). Risk-value models. *European Journal of Operational Research*, 70(2):135–149.

Scheier, M. F., Carver, C. S., and Bridges, M. W. (1994). Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): a reevaluation of the life orientation test. *Journal of Personality and Social Psychology*, 67(6):1063.

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

Shefrin, H. and 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.

Shefrin, H. and Statman, M. (1987). A behavioral finance solution to the noise trading puzzle. *Santa Clara U. Working Paper*.

Shefrin, H. M. and Thaler, R. H. (1988). The behavioral life-cycle hypothesis. *Economic Inquiry*, 26(4):609–643.

Shiller, R. (1981). Do stock prices move too much to be justified by subsequent changes in dividends? *American Economic Review*, 71:421–436.

Shiv, B., Loewenstein, G., Bechara, A., Damasio, H., and Damasio, A. R. (2005). Investment behavior and the negative side of emotion. *Psychological Science*, 16(6):435–439.

Sirri, E. R. and Tufano, P. (1998). Costly search and mutual fund flows. *The Journal of Finance*, 53:1589–1622.

Slovic, P. and Lichtenstein, S. (1968). Importance of variance preferences in gambling decisions. *Journal of Experimental Psychology*, 78(4p1):646.

Staer, A. (2017). Fund flows and underlying returns: The case of etfs. *International Journal of Business*, 22(4).

Stivers, A., Tsang, M., Deaves, R., and Hoffer, A. (2020). Behavior when the chips are down: An experimental study of wealth effects and exchange media. *Journal of Behavioral and Experimental Finance*. in press.

Stivers, C. and Sun, L. (2010). Cross-sectional return dispersion and time variation in value and momentum premiums. *Journal of Financial and Quantitative Analysis*, 45(4):987–1014.

Suhonen, N. and Saastamoinen, J. (2018). How do prior gains and losses affect subsequent risk taking? New evidence from individual-level horse race bets. *Management Science*, 64(6):2473–2972.

Thaler, R. (1985). Mental accounting and consumer choice. *Marketing Science*, 4(3):199–214.

Thaler, R. (1999). Mental accounting matters. *Journal of Behavioral Decision Making*, 12(3):183–206.

Thaler, R. and Johnson, E. (1990). Gambling with the house money and trying to break even: The effects of prior outcomes on risky choice. *Management Science*, 36(6):643–660.

Toplak, M. E., West, R. F., and Stanovich, K. E. (2014). Assessing miserly information processing: An expansion of the cognitive reflection test. *Thinking & Reasoning*, 20(2):147–168.

Tversky, A. and Kahneman, D. (1971). Belief in the law of small numbers. *Psychological Bulletin*, 76(2):105.

Tversky, A. and Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5(2):207–232.

Tversky, A. and Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4):297–323.

Ungeheuer, M. and Weber, M. (2020). The perception of dependence, investment decisions, and stock returns. *The Journal of Finance*. forthcoming.

Van Rooij, M., Lusardi, A., and Alessie, R. (2011). Financial literacy and stock market participation. *Journal of Financial Economics*, 101(2):449–472.

Verbruggen, F., Chambers, C. D., Lawrence, N. S., and McLaren, I. P. (2017). Winning and losing: Effects on impulsive action. *Journal of Experimental Psychology: Human Perception and Performance*, 43(1):147–168.

Von Neumann, J. and Morgenstern, O. (1947). Theory of games and economic behavior (2nd rev). Princeton University Press.

Warther, V. (1995). Aggregate mutual fund flows and security returns. *Journal of Financial Economics*, 39:209–235.

Weber, M. and 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., and Nosić, A. (2013). Who takes risks when and why: Determinants of changes in investor risk taking. *Review of Finance*, 17(3):847–883.

Weber, M. and Zuchel, H. (2005). How do prior outcomes affect risk attitude? Comparing escalation of commitment and the house-money effect. *Decision Analysis*, 2(1):30–43.

Wood, W. S. and Clapham, M. M. (2005). Development of the drake beliefs about chance inventory. *Journal of Gambling Studies*, 21(4):411–430.

Zhang, K. and Clark, L. (2020). Loss-chasing in gambling behaviour: neurocognitive and behavioural economic perspectives. *Current Opinion in Behavioral Sciences*, 31:1–7.

Eidesstattliche Erklärung

Eidesstattliche Versicherung gemäß Paragraph 8 Absatz 2 Satz 1 Buchstabe b) der Promotionsordnung der Universität Mannheim zur Erlangung des Doktorgrades der Betriebswirtschaftslehre:

Bei der eingereichten Dissertation zum Thema "Financial Decision Making – The Role of Realization, Portfolio Composition and Biased Belief Formation" handelt es sich um mein eigenständig erstelltes Werk, das den Regeln guter wissenschaftlicher Praxis entspricht. Ich habe nur die angegebenen Quellen und Hilfsmittel benutzt und mich keiner unzulässigen Hilfe Dritter bedient. Insbesondere habe ich wörtliche und nicht wörtliche Zitate aus anderen Werken als solche kenntlich gemacht. Die Arbeit oder Teile davon habe ich bislang nicht an einer Hochschule des In- oder Auslands als Bestandteil einer Prüfungs- oder Qualifikationsleistung vorgelegt. Die Richtigkeit der vorstehenden Erklärung bestätige ich. Die Bedeutung der eidesstattlichen Versicherung und die strafrechtlichen Folgen einer unrichtigen oder unvollständigen eidesstattlichen Versicherung sind mir bekannt. Ich versichere an Eides statt, dass ich nach bestem Wissen die reine Wahrheit erklärt und nichts verschwiegen habe. Ich bin damit einverstanden, dass die Arbeit zum Zwecke des Plagiatsabgleichs in elektronischer Form versendet, gespeichert und verarbeitet wird.

Mannheim, den

Curriculum Vitae

Jan Müller-Dethard

Since 01/2017	Doctoral Student in Finance Chair of Finance and Banking Chair of Senior Professor Weber Supervisor: Prof. Dr. Dr. h.c. Martin Weber University of Mannheim, Germany
Since 09/2016	Teaching Assistant Chair of Finance and Chair of Corporate Finance University of Mannheim, Germany
2014 – 2016	Master in Management (M.Sc.) University of Mannheim, Germany
08/2015 – 12/2015	Visiting Student Sciences Po, Paris, France
2011 – 2014	Bachelor in General Management (B.Sc.) University of Mannheim, Germany
08/2013 – 12/2013	Visiting Student McGill University, Montréal, Canada