Probabilistic Reasoning in Economic Decisions – Belief Formation, Inference Judgements, and Retirement

Inauguraldissertation

zur Erlangung des akademischen Grades eines

Doktors der Wirtschaftswissenschaften

der Universität Mannheim

vorgelegt von

Pascal Kieren, M.Sc.

im Herbst-/Wintersemester 2020

Dekan: Joachim Lutz

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

Korreferent: Professor Dr. Oliver Spalt

Tag der mündlichen Prüfung: 4. November 2020

Acknowledgements

This thesis is the result of a three years 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 his consistent guidance, many helpful discussions and talks, and his invaluable motivating support. Many thanks are also due to Prof. Dr. Olvier Spalt, my second examiner.

I am very thankful to Jan Müller-Dethard for the great collaboration on our joint projects and for the numerous fruitful discussions we had in our office. I appreciate insightful remarks by Dr. David Becker, Sabine Bernard, Dr. Maximilian Germann, Stefan Greppmair, Prof. Dr. Heiko Jacobs, Dr. Christine Laudenbach, Prof. Dr. Benjamin Loss, Dr. Philipp Marquardt, Prof. Dr. Sebastian Müller, Dr. Christian Schmidt, Yannik Maximilian Schneider, Dr. Peter Severin, Dr. Jiri Tresl, Dr. Maximilian Wimmer and the seminar participants at the University of Mannheim.

Most of all, I would like to thank my family and friends, who were there whenever I needed them. I am deeply indebted to my parents, Annette and Klaus, who always believed in me and gave me the opportunity to pursue my own goals in life. I will never forget what you did for me. I owe my deepest gratitude to Laura. You supported me with all your love, energy, and appreciation over all these years. This thesis would not have been possible without your carrying support and nurturing.

Contents

List of Figures **viii**

List of Tables **x**

1 Introduction **1**

1.1 Relevance and Foundations of Probabilistic Reasoning 1

1.2 Contribution and Main Results of this Dissertation Thesis . . . 13

1.2.1 Why so Negative? Belief Formation and Risk-Taking in Boom and Bust Markets 13

1.2.2 Can Agents Add and Subtract When Forming Beliefs? . 17

1.2.3 Expectation Formation under Uninformative Signals . 20

1.2.4 When Saving is Not Enough – Wealth Decumulation in Retirement 23

2 Why So Negative? Belief Formation and Risk-Taking in Boom and Bust Markets **29**

2.1 Introduction 29

2.2 Experimental Design 37

2.2.1 Detailed Description of the Experiment 37

2.2.2 Discussion of Important Aspects 41

2.3 Hypotheses 42

2.4 Summary Statistics and Randomization Checks 43

2.5 Results 45

2.5.1 Main Result 45

Distorted Belief Formation 45
Belief Formation and Risk-Taking 47
Mechanism 50
2.5.2 Boundaries and External Validity 53
2.5.3 Further Analyses and Robustness Checks 54
2.6 Conclusion 61

3 Can Agents Add and Subtract When Forming Beliefs? 63
3.1 Introduction 63
3.2 Empirical Framework and Hypotheses 69
3.3 Experimental Design 73
3.3.1 Baseline Design 73
3.3.2 Experimental Variations 75
3.3.3 Summary Statistics 77
3.4 Results 77
3.4.1 Main Results 77
Baseline Results 79
Reducing the Diagnosticity of Information Signals . . . 83
Reducing the Uncertainty About the Underlying Distribution 86
Additional Treatments G-1 and G-2 (B-1 and B-2) 90
3.4.2 Signal Ordering 93
3.4.3 Robustness Checks 95
3.5 Conclusion 97

4 Expectation Formation Under Uninformative Signals 101
4.1 Introduction 101
4.2 Conceptual Framework 107
4.2.1 Setup 107
4.2.2 Experimental Design 110

4.2.3 Hypotheses Development . . . 113
4.2.4 Summary Statistics . . . 115
4.3 Results . . . 117
4.3.1 Belief Updating after Uninformative Signals . . . 117
4.3.2 Frequency of Updating Mistakes . . . 120
4.3.3 Mechanism: Reference-dependent Belief Updating . . . 124
4.3.4 Robustness Checks . . . 128
The Role of Memory and Learning . . . 128
Sample Splits . . . 129
4.4 Conclusion . . . 130

5 When Saving is Not Enough – Wealth Decumulation in Retirement 133
5.1 Introduction . . . 133
5.2 Survey Design and Summary Statistics . . . 138
5.2.1 Survey Design . . . 138
5.2.2 Summary Statistics . . . 144
5.3 Results . . . 146
5.3.1 Preferences on the Structure of Phased Withdrawal Accounts . . . 146
5.3.2 Utility Analysis of Phased Withdrawal Accounts . . . 150
Normative Predictions . . . 150
Predicted and Actual Choice . . . 155
5.3.3 Determinants of Participants' Product Choice . . . 157
5.3.4 The Wealth Decumulation Decision . . . 161
Financial Preparedness and the Wealth Decumulation Decision . . . 163
Emotional Preparedness and the Wealth Decumulation Decision . . . 164
5.4 Conclusion . . . 168

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

A.1 Further Analyses I

A.2 Experimental Instructions and Screenshots III

A.3 Experimental Measures VIII

B Can Agents Add and Subtract When Forming Beliefs? **XV**

B.1 Experimental Instructions and Screenshots XV

C Expectation Formation under Uninformative Signals **XXIII**

C.1 Further Analyses XXIII

C.2 Experimental Instructions and Screenshots XXV

C.3 Additional Experimental Measures XXXIII

C.4 Derivations and Proofs XXXV

D When Saving is Not Enough – Wealth Decumulation in Retirement **XXXIX**

D.1 Further Analyses XXXIX

D.2 Survey Instructions and Screenshots XLII

D.3 Construction of the Phased Withdrawal Strategies L

References **LIII**

Eidesstattliche Erklärung **LXVII**

Curriculum Vitae **LXIX**

List of Figures

2.1 **Learning Environments and Treatments** 38

2.2 **Characteristics of Boom and Bust Market Phases** 39

2.3 **Between-subject Measure of Belief- and Preference-based Risk-Taking** 40

2.4 **Structure and Flow of the Experiment** 41

2.5 **Pessimism Bias** 46

2.6 **Dow Jones Estimates and Life Orientation Test** 55

2.7 **Dow Jones Estimates and Life Orientation Test Split by Forecasting Ability** 60

3.1 **Empirical Framework** 70

3.2 **Subjects' Average Updating Behavior in the Good Distribution – Experiment 1** 79

3.3 **Subjects' Average Updating Behavior in the Bad Distribution – Experiment 1** 80

3.4 **Subjects' Average Updating Behavior in the Good Distribution – Experiment 2** 84

3.5 **Subjects' Average Updating Behavior in the Bad Distribution – Experiment 2** 85

3.6 **Subjects' Average Updating Behavior in the Good Distribution – Experiment 3** 87

3.7 **Subjects' Average Updating Behavior in the Bad Distribution – Experiment 3** 88

3.8 **Subjects' Average Updating Behavior in the Good Distribution – Treatments G-1 and G-2** 91

3.9 **Subjects' Average Updating Behavior in the Bad Distribution – Treatments G-1 and G-2** 92

4.1 **Overview of the Updating Task** 112

4.2 **General Updating Tendency** 117

4.3 **Basic Updating Mistakes** 121

4.4 **Prior Dependent Updating** 125

5.1 **Survey Overview and Timing** 142

5.2 **Phased Withdrawal Account Choice** 147

5.3 **Utility Simulations** 152

5.4 **Annuity and Phased Withdrawal Demand** 158

5.5 **Willingness to Decumulate Retirement Savings** 162

A.1 **Payoff Screen** V

A.2 **Probability Estimate Screen** VI

A.3 **Confidence Level Screen** VII

B.1 **Payoff Screen** XVIII

B.2 **Probability Estimate Screen** XIX

B.3 **Confidence Level Screen** XX

C.1 **Investment Screen** XXVII

C.2 **Payoff Screen** XXVIII

C.3 **Accumulated Payoffs Screen** XXIX

C.4 **Probability Estimate Screen** XXX

C.5 **Confidence Level Screen** XXXI

D.1 **Display of Phased Withdrawal Accounts** XLVI

D.2 **Display of Lifelong Annuity** XLVII

List of Tables

2.1 **Summary Statistics on Subjects** 44

2.2 **Pessimism Bias** 47

2.3 **Risk-Taking Across Macroeconomic Cycles I** 48

2.4 **Risk-Taking Across Macroeconomic Cycles II** 50

2.5 **Relation Between Treatment Variable and Probability Estimates** 51

2.6 **Relation Between Beliefs About Success Probability and Investment** 52

2.7 **Risk-Taking Across Macroeconomic Cycles Split by Forecasting Quality** 56

2.8 **Relation Between Treatment and Probability Estimates Split by Forecasting Quality** 57

3.1 **Overview of Treatments** 74

3.2 **Summary Statistics on Subjects** 78

3.3 **Updating Behavior After Disconfirming Signal and Correction – Experiment 1** 83

3.4 **Updating Behavior After Disconfirming Signal and Correction – Experiment 2** 86

3.5 **Updating Behavior After Disconfirming Signal and Correction – Experiment 3** 89

3.6 **Outcome Ordering** 94

3.7 **Forecasting Ability and Extreme Outliers** 96

4.1 **Payoff Distribution** 111
4.2 **Summary Statistics** 116
4.3 **Uninformative Updating** 119
4.4 **Frequency of Directionally Inconsistent Updating Errors** . 123
4.5 **Reference Dependent Belief Updating** 127
4.6 **Robustness Checks** 130

5.1 **Summary Statistics** 145
5.2 **Determinants of the Phased Withdrawal Account Choice** . 149
5.3 **Utility Predictions** 156
5.4 **Determinants of Annuity and Phased Withdrawal Demand** 159
5.5 **Financial Preparedness and Wealth Decumulation** 164
5.6 **Emotional Outlook and Wealth Decumulation** 167

A.1 **Pessimism Bias Split by Forecasting Quality** II

C.1 **Uninformative Updating and Learning** XXIV

D.1 **Utility Parameters and Account Choice** XL
D.2 **Retirement Preparation and Savings** XLI

Für Mama und Papa

Chapter 1

Introduction

1.1 Relevance and Foundations of Probabilistic Reasoning

Most of our everyday decisions are generally made without definite knowledge of their consequences. The decisions to invest in the stock market, to undergo a medical operation, to save for retirement, or to go to court are commonly made without knowing in advance whether the market will go up, the operation will be successful, how much we need to save, or whether the court will decide in one's favor (Tversky and Fox, 1995). More broadly speaking, virtually all social, economic, or technological decisions involve some degree of risk or uncertainty. In some instances (such as games of chance), the probabilities of the alternative consequences can be accurately determined (Machina, 1987). In other cases, individuals have to rely on experience or personal estimates that are usually expressed in statements such as "I think that...", "chances are...", or "it is unlikely that...".

What determines such beliefs and how do people assess the probability of uncertain events? Most traditional theories in economics and psychology assume that agents gather and integrate information in a manner that will result in a relatively accurate representation of reality (e.g. Morgenstern and Von Neumann, 1953; Körding and Wolpert, 2004; or Sharot and Garrett, 2016). Yet, the process of forming accurate beliefs is a challenging cognitive operation as we are constantly flooded with a wealth of information and new stimuli. To manage the constant flow of new information without being overwhelmed, individuals tend to rely on a number of heuristic principles in forming expectations. Generally, such heuristics or "mental shortcuts" are useful tools which reduce the complex task of assessing probabilities to simpler judgmental operations (Tversky and Kahneman, 1974). Despite their usual efficiency, heuristics have also been found to give rise to the occurrence of severe and systematic errors in our judgements. Such predictable errors in thinking can lead to deviations from the normatively expected judgement with potentially negative effects on the decision outcome.

Particularly in financial decision making, people frequently find themselves confronted with decisions that involve uncertainty and which require a probability assessment in order to form accurate beliefs about future outcomes. Examples of such decisions involve consumption and saving choices, investment choices, and financing decisions among many others. Yet, given the complexity that is often associated with financial decisions, it is hardly surprising that individuals make systematic errors in forming accurate expectations. On the individual level for example, biased beliefs have been associated with non-participation in the equity market (e.g. Dimmock et al., 2016), systematic mis-valuation of financial assets (e.g. Shiller, 1981; or De Long et al., 1990), portfolio under-diversification (e.g. Benartzi, 2001; or Coval and Moskowitz, 1999), as well as undersaving (e.g. Han et al., 2019; or Heimer et al., 2019). Importantly, systematic errors in assessing probabilities not only apply to individual consumers or private households, but also to highly trained professionals. Here, biased beliefs have been linked to excessive and value-destroying merger activity (e.g. Malmendier and Tate, 2008), flawed inflation expectations of central bankers (e.g. Malmendier et al., 2017; or Malmendier and Nagel, 2015), or inaccurate analysts' earnings forecasts (e.g. Clement and Tse, 2005) to name just a few.

Why should politicians be alarmed about these findings? First, and perhaps most obviously, the above-mentioned examples clearly demonstrate that the resulting misbehavior can be very costly for households and corporates alike. Ma et al. (2018) for example find that managerial forecast errors may lead to sizable welfare losses for the aggregate economy. Goetzmann and Kumar (2008) conclude that under-diversified households earn up to 3.12 % lower risk-adjusted annual returns compared to diversified households. On top of that, one particular point of concern refers to long-term financial decisions, such as the decision of how to save for retirement. Here, errors in individuals' judgement can accumulate over multiple years and are often irreversible. Due to the structural changes in society and the resulting demographic challenges, less pension contributors will have to pay for an increasing generation of pension receivers. In the light of these developments, many households can no longer exclusively rely on the public pension system. Instead, they will be forced to complement public pension schemes with private savings. In other words, more and more individuals will have to deal with financial decisions when it comes to retirement planning and thus cannot circumvent to form expectations about risk and return characteristics of financial assets. Especially in this domain, systematic errors in our judgement can not only severely impede our own future financial wellbeing, but may also affect our family and children.

Given the far-reaching consequences of biased expectations in many of our everyday decisions, it is and should be an important concern for researchers and politicians alike to understand why and in which way individuals depart from normatively expected judgements. In order to investigate potential discrepancies, it is, however, necessary in a first step to establish how modern decision theory defines choices under risk and uncertainty and then relate observed judgements to normative predictions.

The topic of how individuals make judgements under risk and uncertainty has been the object of theoretical and empirical investigations for centuries. A decision under uncertainty generally requires an evaluation of two attributes: the desirability of possible outcomes and the likelihood with which each outcome may occur (Tversky and Fox, 1995). Economists have developed various models to study how individuals assess each of these attributes individually and how they should be evaluated jointly.

The neoclassical model of decision-making under risk and uncertainty is the expected utility framework. Given the importance of choice under uncertainty in the literature and especially in this thesis, I will briefly describe the main intuition of the model. The idea of maximizing the expected utility of a decision outcome was originally developed by Daniel Bernoulli in 1738 (reprinted and translated in Bernoulli, 1954). In this model, the utility associated with each possible outcome (x_1,. . . ,x_n) is weighted by its probability of occurrence (p(x_1),. . . ,p(x_n)):

$$E(u(x)) = \sum_{i=1}^{n} p(x_i)u(x_i).$$

Under expected utility theory, money and wealth are assumed to diminish in value the more we receive. More precisely, the utility function – which maps actual wealth into utility for wealth -– is generally assumed to be concave. The function's degree of curvature thereby often serves as an index of an individual's degree of risk aversion (Weber and Johnson, 2009). The introduction of a parameter describing individuals risk attitude has intuitive appeal, as some people seem to resolve choices that differ in risk in very cautious ways, while others seem willing to take on greater risks.

Almost two centuries later, von Neumann and Morgenstern (1953) formally axiomatized the expected utility maximization. Under the assumption that preferences satisfy the axioms of completeness, transitivity, continuity, and independence, the concept of expected utility maximization became a normatively attractive decision criterion and served as a foundation for many alternative choice models such as Savage's subjective expected utility theory (1954), regret theory (Bell, 1982; Loomes and Sugden, 1982), or rank-dependent expected utility (Quiggin, 1982).

Based on the traditional economic choice models, the decision-making process can thus be described as follows. First, individuals need to assess the probability of each possible outcome that may result from their initial choice. Then, individuals need to discern the utility to be derived from each outcome and subsequently combine both assessments to make a judgement (Gilovich et al., 2002). Yet, most of our everyday decisions involve situations in which either the consequences are unclear or the probabilities of the consequences are unknown. How do individuals form expectations about the probability of such uncertain events? What assumptions do researchers impose about this process based on the elementary rules of probability theory? Relatedly -– and perhaps most importantly — do individuals behave in a manner that is consistent with these assumptions or do they systematically deviate from normative predictions?

This dissertation thesis aims to provide answers to these questions. A special emphasis is put on whether and when individuals' expectations deviate from rational expectations and on uncovering the psychological roots that drive these processes. In the next paragraphs, I will give a brief introduction to the core component of the neoclassical theory of probabilistic beliefs, *Bayes' Theorem*, and review what the literature already knows about deviations from Bayesian behavior.

Degrees of Uncertainty

The economist Frank Knight (1921) was the first to make a conceptual distinction between risk and uncertainty. A decision under risk typically refers to situations in which the decision-maker knows with certainty the probabilities of possible choice alternatives. Examples of decisions under risk are the toss of a coin, or the roll of a fair die, in which we know the possible outcomes and can assign a unique probability to each outcome. A decision under uncertainty refers to situations in which the decision-maker cannot express the likelihood of possible outcomes with any mathematical precision (Weber and Johnson, 2009). Here, knowledge about the probability distribution of the outcomes of choice alternatives can lie anywhere on a continuum, from complete uncertainty (also referred to as complete ignorance), through various degrees of partial uncertainty, to risk (where the probability distribution of alternative outcomes is fully specified).

Bayesian Inference

In assessing how individuals make judgements in situations in which the probabilities of alternative outcomes are not fully known, it is key to understand how individuals resolve or quantify uncertainty. Generally, uncertainty is reduced by observing, gathering knowledge, and integrating new information to form more accurate beliefs about a particular subject. The accuracy of our probability estimates thus strongly depends both on the quality of our present knowledge and the accuracy and content of the information we acquire. The core component of the neoclassical theory of probabilistic beliefs is the assumption that individuals integrate new information into their prior beliefs according to Bayes' Theorem. Bayes' Theorem, named after the mathematician Thomas Bayes (1763), is an algorithm for combining prior knowledge with current information. Based on past experience, a Bayesian decision maker begins with a *prior* belief that some aspect of the world holds, and then gathers new information that modifies his initial belief to produce a *posterior* belief (Efron, 2013).

In more formal terms, Bayes' Theorem is defined as follows (notation adapted from McNamara et al., 2006): suppose there are n possible states of the world, which are labelled S_1,S_2,...,S_n. The prior probability (based on knowledge or past experience) that state S_i is the underlying objective state is denoted $P(S_i)$. Let A be some event which provides relevant information about the objective state of the world. The probability that event A occurs assuming that the state S_i is the true state of the world is denoted as $P(A|S_i)$. Then, the overall prior probability that event A occurs is:

$$P(A) = \sum_{i=1}^{n} P(A|S_i)P(S_i).$$

Now assume that based on his prior information about the likelihood that event A occurs, a decision maker in fact observes that event A has happened. Given this additional knowledge, Bayes' Theorem prescribes that the posterior probability that S_i is the true state of the world given that event A has happened can now be calculated as:

$$P(S_i|A) = \frac{P(A|S_i)P(S_i)}{P(A)}.$$

In other words, Bayes' Theorem is a way to calculate conditional probabilities based on new information that we integrate into our prior beliefs. As we continue to gather new information, Bayes' Theorem can be applied iteratively. As such, probabilities are updated step by step whenever new information arrives, until uncertainty is reduced to a tolerable level.

Deviations from Bayesian Behavior

On a normative level, Bayes' Theorem provides formidable advice on how individuals should combine prior knowledge with new information to form beliefs relevant for the decision process. Yet, even though Bayes' Theorem is a cornerstone of modern probability theory, it is not free from critique. Contrary to the assumption of traditional models that individuals always follow the elementary rules of probability when calculating the likelihoods of uncertain outcomes, many studies find that individuals are subject to systematic errors in their probabilistic reasoning.

This string of literature -– which is often referred to as "heuristics and biases" -– was introduced by the psychologist Ward Edwards in the 1960s (e.g. Phillips and Edwards, 1966) and presented the starting point of a large research agenda including the seminal papers by Daniel Kahneman and Amos Tversky (1971, 1974). The goal of this field of research is to compare intuitive inferences and probability judgements to the rules and laws of probability theory. Over the years, many different biases have been identified, including the gambler's fallacy (Alberoni, 1962), the conservatism bias (Phillips and Edwards, 1966), base-rate neglect (Phillips and Edwards, 1966; Kahneman and Tversky, 1973), a false belief in the law of small numbers (Tversky and Kahneman, 1971) the representativeness heuristic (Kahneman and Tversky, 1972), and the confirmation bias (Nickerson, 1998).

Especially important to our understanding of how individuals reduce or quantify uncertainty are the biases that affect how individuals revise their prior beliefs upon receipt of new information. Biased belief updating can be identified by comparing people's subjective posteriors with the correct objective posterior belief as implied by Bayes' Theorem.

In this literature, much attention is devoted to two types of potential biases in individuals' updating behavior. The first type of bias refers to the insufficient use of likelihoods in drawing inferences. Whereas Bayes' Theorem prescribes that individuals update their beliefs in proportion to the information they observe, many studies point in the direction that individuals update their beliefs as if the signals provided less information about an objective state than they actually do. This tendency to revise prior beliefs only insufficiently when presented with new evidence was first discussed by Phillips and Edwards (1966) and is referred to as conservatism bias. The second type of bias in drawing inference concerns the use of prior beliefs. Instead of properly combining the inferential impacts of prior knowledge and new diagnostic evidence as prescribed by Bayes' Rule, individuals on average under-use their prior information. This phenomenon -– also labeled as base-rate neglect by Kahneman and Tversky (1973) -– may cause individuals to overinterpret a recent signal indicative of an event that is unlikely given the base rate.

Besides the biased use of prior knowledge and new diagnostic evidence in drawing inference judgements, our belief formation is also affected by our own preferences. In fact, humans tend to form beliefs asymmetrically – we quickly discount bad news but embrace good news (Sharot et al., 2012). For example, studies have shown that people readily adjust their beliefs regarding their level of intelligence and physical attractiveness when they receive favorable information that indicates that they are more intelligent or attractive than they had previously assumed. However, they fail to adjust beliefs after information that suggests otherwise (Köszegi, 2006).

Overview of Chapters

This dissertation thesis contributes to ongoing research in financial economics and psychology, which investigates individuals' belief formation and the mechanisms that underly perception and judgment. Incorporating insights from the finance, psychology, and economics literature, each chapter of this dissertation thesis focuses on one particular aspect of how individuals form expectations relative to Bayesian behavior and relates the findings to explain observed judgements. In the following paragraphs, I will give a brief overview of the main research question covered in each chapter of this dissertation. Afterwards, Section 1.2 contains a more detailed description which focuses on the main findings and the contributions to various strings of literature.

Chapter 2 investigates how biased belief formation may affect investors' willingness to take financial risks across market cycles. One of the major puzzles in the financial economics literature is the fact that investors' risk-taking varies strongly and systematically across market cycles. In particular, investors are generally found to take more risks during boom markets and less risks during bust markets. To account for this pervasive pattern, researchers have proposed rational expectations models which implicitly assume that investors' attitude towards risk (or their *risk preferences*) changes in tandem with market cycles. Alternatively, investors' judgement might not only be affected by their risk preferences — which are often assumed to be a stable construct -– but rather by their expectations about risk and return characteristics. Even after decades of research, the underlying drivers of the observed differences in investors' risk-taking behavior are still not fully understood and subjected of heated debates among researchers. One reason for this long-lasting debate is that even though both the preference as well as the belief channel are observationally equivalent, they nonetheless offer vastly different policy implications. Chapter 2 contributes to this debate by showing that individuals rely on different learning rules when forming their beliefs across market cycles. The resulting systematic deviations from Bayes' Rule can not only explain differences in risk-taking over time, but especially across market cycles, where the underlying learning environments differ.

The following two chapters of this dissertation thesis focus on how individuals revise their prior beliefs upon receipt of new information, which either *disconfirms prior information* (Chapter 3) or which is *non-diagnostic* about the objective state of the world (Chapter 4). An implicit -– albeit often neglected -– implication of Bayes' Theorem is that two informationally equivalent signals of opposite direction cancel each other out so that the total value of the information is much like no information at all. In other words, if individuals' judgements about uncertain events conform to principles of logic, then two opposing signals with the same informational content should not influence their beliefs. Examples of such opposite-directional signals can often be found in our everyday life, such as receiving both positive and negative feedback from two equally trustworthy friends about the quality of a recently opened restaurant. In Chapter 3, we seek to explore how individuals process such pieces of information by testing experimentally how individuals revise their prior beliefs after both same-directional and opposite-directional signals. We contribute to the literature by showing that whenever a sequence of signals that go in the same direction is interrupted by a single signal of opposite direction, individuals tend to strongly overreact to the signal of opposite direction. In other words, individuals appear to process the opposite-directional signal *as if* the signal would carry more weight in the decision process than previous signals.

Non-diagnostic information may not only occur by observing two informationally equivalent signals of opposite direction, but also by observing signals which are plainly uninformative about a particular objective state of the world. An important implication of Bayes' Theorem is that individuals should not differentiate between observing no information signals at all and receiving uninformative signals. Examining whether this is really the case has important consequences for our understanding of whether our judgement can be influenced by irrelevant pieces of information. Chapter 4 aims to add to this research agenda. It shows that individuals revise their prior beliefs even after observing uninformative signals. Importantly, the direction in which individuals tend to update their beliefs depends on the valence of the signal: prior beliefs become more optimistic after desirable uninformative signals and more pessimistic after undesirable uninformative signals. This mechanism implies that there is indeed a distinction between receiving no signals or uninformative signals for drawing inference judgments.

In Chapter 5, the focus shifts from the traditional investigation of biased belief formation towards an application where biased beliefs are often very costly: retirement planning. In many articles on financial planning, retirement planning is often used synonymous with wealth accumulation. However, while wealth accumulation is certainly an important ingredient for successful retirement preparation, it is not sufficient to achieve a targeted steady stream of income during retirement. Individuals close to retirement thus face the following decision problem: out of one's accumulated wealth, one must decide how much to allocate to a savings account (e.g. as protection against unexpected costs) and how much to consume over the course of one's retirement to secure a given standard of living. Chapter 5 of this dissertation thesis adds to this literature by studying how individuals approach this decision problem and which decumulation schemes they find most appealing to transfer wealth into a stream of income.

1.2 Contribution and Main Results of this Dissertation Thesis

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

Chapter 2, coauthored with Jan Müller-Dethard and Martin Weber, presents an experimental study on the role of biased belief formation for investors' risk-taking across macroeconomic cycles. One of the major puzzles in financial economics is the fact that risk premiums of many asset classes vary strongly and systematically over time (Shiller, 1981; Campbell and Shiller, 1988a,1988b): risk-premiums tend to be lower during market cycle booms and higher during market cycle busts. To account for this pervasive finding researchers have proposed rational expectations models which introduce modifications into the representative agent's utility function (Campbell and Cochrane, 1999; Barberis et al., 2001). To generate the empirically observed time-variation in the equity premium, investors in these models are generally assumed to be more risk-averse during bust markets, thus demanding a higher risk premium, and less risk averse during boom markets, thus demanding a lower risk premium. Evidence in favor for this "countercyclical risk-aversion" is found by Cohn et al. (2015) as well as Guiso et al. (2018).

However, the concept that investors exhibit a countercyclical risk aversion is also contested. Rational expectation models typically assume that agents always correctly updated their beliefs as prescribed by Bayes' Theorem. This implies that agents are assumed to know the objective probability distribution in equilibrium and are as such fully aware of the countercyclical nature of the equity risk premium (Nagel and Xu, 2019). In other words, investors in these models should have more pessimistic return expectations during boom markets and more optimistic return expectations during bust markets. This assumption of rational expectations is troublesome for two reasons as recently pointed out by Nagel and Xu (2019). First, conceptually it is unclear how an investor could possess so much knowledge about parameters that even econometricians tend to struggle to estimate with precision. Second, surveys of actual investor return expectations find that investors' return expectations are at odds with the rational expectations assumption. Greenwood and Shleifer (2014) show that the reported return expectations of investors are highly correlated with past returns and as such rather procyclical instead of countercyclical: investors tend to be more optimistic during market booms and more pessimistic during market busts. Similar findings about the procyclicality of investors' return expectations in survey data are presented by Amromin and Sharpe (2014) and Giglio et al. (2019).

Alternatively, the time-varying nature of the equity risk premium might also be caused by changes in investors' beliefs about risk and return characteristics. Whereas this channel is mostly held constant in traditional models due to the rational expectations assumption, it presents an equally valid hypothesis which receives increasing attention in the literature in recent years. In a survey of online-broker customers over the financial crisis in 2008, Weber et al. (2013) show that changes in risk taking are mostly attributable to changes in return expectations and only to a lesser extent to changes in risk preferences. Nagel and Xu (2019) present a model in which investors have heterogenous time-varying beliefs. Their model is able to reconcile asset prices and survey expectations without assuming that investors have unstable risk preferences. Yet, even after decades of research on the time-varying nature of the equity premium, the underlying drivers are still not fully understood. The fact that time-varying beliefs and risk preferences are often observationally equivalent to researchers makes a clean identification very challenging.

Chapter 2 contributes to this ongoing debate by showing that distorted belief formation rules (i.e. systematic deviations from Bayes' Rule) can explain differences in risk-taking across macroeconomic cycles. In an experimental study, we investigate (i) how different learning environments affect the formation of return expectations; (ii) whether systematic differences in the employed learning rules affect risk-taking; and (iii) whether the learning environments only affect beliefs or also investors' risk preferences. While recent survey data on expectations is helpful to establish a link between subjective beliefs and investment decisions, it does 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 subjects' investment choice. In two experiments, we combine an abstract belief formation task (Bayesian learning) 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 make a forecast about the quality of a risky asset. The underlying learning environment of the updating task either resembles key characteristics of a boom market (favorable learning environment) or a bust market (adverse learning environment). Importantly, the underlying probability distribution from which the information is drawn is completely identical in both learning environments. In other words, a Bayesian agent in our setting should make identical forecasts irrespective of whether he learns in the favorable or adverse environment. In the subsequent investment task, we randomly assign subjects to invest either in an ambiguous lottery with unknown success probability, or a risky lottery with known success probabilities. In the risky lottery, we have perfect control over subjects' return and risk expectations since both probabilities and outcomes are known. In the ambiguous lottery, however, we purposefully provide participants room to form subjective beliefs about the underlying probability distribution. 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 their risk preferences. The between-subject comparison allows us to isolate the effect of belief-induced risk-taking caused by different learning environments. Our results can be summarized as follows. First, we find that adverse learning environments which resemble key characteristics of bust markets induce a strong pessimism bias in individuals' belief formation. Second, the induced pessimism not only presents a systematic deviation from Bayesian beliefs, but also translates to lower investments in the ambiguous lottery. In risky lottery, however, we do not find any difference in risk-taking depending on the underlying learning environment. In other words, our results suggest that risk preferences are unaffected by the initial learning environment and stable across treatments. 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.

To conclude, Chapter 2 tests an alternative channel to the countercyclical risk aversion hypothesis which can also explain the empirically observed time-varying changes in risk-taking. Instead of assuming unstable preferences, we investigate whether systematic deviations from Bayesian beliefs can cause similar investment pattern. In our study, we show that individuals tend to employ different learning rules when forming beliefs in boom and bust markets. In adverse learning environments, individuals form overly pessimistic beliefs which subsequently translate to a lower willingness to take risks. This result is also consistent with recent survey evidence reporting pro-cyclical beliefs of investors. Our findings have important policy implications. 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. Such self-reinforcing feedback loops may amplify the intensity and length of market trends.

1.2.2 Can Agents Add and Subtract When Forming Beliefs?

Chapter 3, coauthored with Jan Müller-Dethard and Martin Weber, presents an experimental study on how individuals incorporate information signals which disconfirm prior information into their beliefs. Standard models of economic choice assume that individuals update their prior beliefs upon receipt of new information according to Bayes' Theorem. Besides the prescription of how to calculate 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

Many of our everyday decisions which involve uncertainty require that we collect new pieces of information until uncertainty is reduced to a tolerable level. In the simplest case, we observe only signals which point towards the same conclusion. More often, however, we observe mixed pieces of evidence that sometimes disagree with one another. 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 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? Bayes' Theorem would prescribe that this is in fact the case. However, many studies conclude that individuals are often not perfect Bayesian. Instead, they sometimes under- or overinfer from new information. As recently pointed out by Benjamin (2019), it is an important question to understand when we expect people to update too much and when we expect them to update too little. In this chapter, we take a step in this direction by investigating whether individuals follow this simple counting-based rule, as implied by Bayes' Theorem as well as when and why we may expect them to over- or underinfer.

To examine our research question, we first develop a simple framework to derive hypotheses and to guide the experimental design. In the framework, we define any information signal which confirms the objective state of the world as a *confirming* signal, and any signal which disconfirms the objective state as a *disconfirming* signal. Additionally, we define three phases of how Bayesian beliefs can evolve over a sequence of information signals. Phase 1 is characterized by a sequence of at least two same-directional signals (confirming signals). Phase 2 resembles the moment in which the disconfirming signal occurs, while Phase 3 defines the situation when the previously observed disconfirming signal gets reverted by another confirming signal. The established framework allows us to test (i) how subjects update their priors after a disconfirming signal conditional (i.e. opposite-directional signal) 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). The counting rule implicit in Bayes' Theorem makes clear predictions how individuals 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.

To test this prediction, we embed our empirical framework into the standard incentivized Bayesian updating bookbag-and-poker-chip paradigm by Grether (1980). Participants learn over six periods about the quality of a risky asset from binary signals which are drawn either from a "good distribution" or a "bad distribution". Whereas one signal is more indicative for the good distribution, the other is more indicative of the bad distribution. In the experiment, subjects always observe five confirming signals (depending on the underlying distribution) and a single disconfirming signal. We exogenously manipulate the period in which the single disconfirming signal occurs. This provides us with twelve stratified price paths (six for the good and six for the bad distribution).

The main findings from Chapter 3 can be summarized as follows. Whenever individuals observe a single disconfirming signal after a sequence of confirming signals — or in the example above, receive a single critique after a few recommendations -– they violate the counting rule and strongly overreact. This overreaction is relatively independent of the number of previously observed confirming signals and occurs already after a sequence of only two confirming signals. In other words, the overreaction is not triggered by extreme prior beliefs. However, when the disconfirming signal gets revered by another confirming signal (i.e. Phase 3 in the framework), participants on average correctly adhere to the counting rule and almost fully correct their prior overreaction. In contrast to their overreaction when violating the counting rule, we find that individuals generally underinfer whenever they cannot or do not violate the counting rule. This is frequently the case when there are only signals of the same direction (i.e. before observing a disconfirming signal) or if signals alternate (i.e. confirming, disconfirming, confirming).

In summary, Chapter 3 contributes to one important objective in recent research on probabilistic beliefs which is -– according to Benjamin (2019) -– to identify when individuals update too much and when they update too little. Within the common paradigm by Grether (1980), our results coherently suggest that individuals update too much whenever they violate the counting rule implied by Bayes' Theorem, and too little otherwise. This finding has important implications for human decision making in general. In environments with conflicting information, a perfectly Bayesian decision maker would eventually be able to identify the objective state with certainty. However, the overreaction resulting from violating the counting rule implicit in Bayes' Theorem, might prevent individuals to be fully confident about a certain state, causing their beliefs to fluctuate even if the surrounding information environment is fundamentally stationary.

1.2.3 Expectation Formation under Uninformative Signals

Chapter 4, coauthored with Martin Weber, presents an experimental study on how individuals process non-diagnostic information signals when updating their beliefs. The neoclassical theory of probabilistic beliefs assumes that individuals update their prior beliefs according to Bayes' Rule as new (relevant) information arrives. In this model, signals which do not carry relevant information about the objective state of the world play no role and are treated as if no signal occurred. Despite this clear prediction, it cannot always be assumed that individuals' beliefs or inference about uncertain events always conform to principles of logic. In reality, many information structures are complex and generate signals which are often noisy and difficult to ascribe to one particular state of the world. Additionally, new information is rarely processed as being purely informative. Instead, individuals frequently have preferences over which state of the world is true, effectively generating an interaction between beliefs and preferences (Eil and Rao, 2011; Möbius et al., 2014). This interaction may lead to environments, in which information signals carry no information about an underlying state of the world, but which nonetheless appear either desirable or undesirable in the utility they provide.

Correctly identifying the informational value of new pieces of evidence is important in almost all decision which require an assessment of probabilities, including psychologists' interpretation of diagnostic tests, doctors' diagnoses of patients, courts' judgements in trial, or ideological conflicts and political discussions. Errors in probabilistic reasoning in such domains are not only costly, but may also lead to wrong treatments of patients or to mistaken convictions of defendants. In the light of these consequences, it is imperative to obtain a deeper understanding of how individuals process non-diagnostic information and especially whether individuals can correctly discern belief-relevant information from their preferences.

In Chapter 4, we first present a stylized reduced-form model which builds on earlier work by Grether (1980) to guide the design of the experiment and to structure the main part of the empirical analysis. In the model, we formally derive predictions about individuals' updating behavior following both informative and uninformative signals under a Bayesian perspective. This allows us to compare Bayesian beliefs to observed subjective beliefs while also being able to control for under- and overinference as well as base-rate neglect. To test how individuals process non-diagnostic signals, we employ an incentivized bookbag-and-poker-chip experiment, in which we exogenously vary both the informational content and the valence of the observed signals. Over the course of 10 rounds, participants repeatedly have to incorporate a series of information signals into their beliefs to forecast the distribution of a risky asset. The risky asset can generate three outcomes from one of two distributions, a bad distribution and a good distribution. The outcomes can be ranked according to their associated payoff (high, medium, and low). In the good distribution, the high outcome occurs with the highest probability, while the low outcome occurs with the lowest probability. In the bad distribution, probabilities of the high and low outcome are reversed. Following this logic, the high outcome signals that the good distribution is more likely, whereas the low outcome signals that the bad distribution is more likely. Importantly, the medium outcome always occurs with the same probability independent of the underlying distribution. In other words, the medium outcome provides no opportunity to learn about the true state of the risky asset and is thus referred to as an *uninformative* signal or a *non-diagnostic* signal. To investigate how the valence of uninformative signals affects individuals' updating behavior, we also exogenously manipulate the payoff of uninformative signals in a between-subject design. Whereas the uninformative signal provides positive payoffs for some participants, it provides negative payoffs for others. Importantly, the distributions from which information is drawn are constructed in a way, that the medium outcome does not provide any information about whether subjects are currently drawing from the good or the bad distribution. As such, a Bayesian agent in our setting should not update his prior beliefs after observing an uninformative signal, independent of whether the signal is in the positive domain or in the negative domain. This allows us to disentangle the valence from the informational content of a signal and to document systematic errors in the belief formation process.

Results can be summarized as follows. First, we find that individuals strongly and systematically update their prior beliefs after observing signals that are uninformative of the objective state of the world. In contrast to Bayesian behavior, individuals fail to fully extract belief-relevant information. Second, we find that the direction in which individuals update their beliefs strongly depends on the valence of the observed signal. In particular, individuals tend to form more optimistic beliefs about the objective state of the world after observing positive uninformative signals, whereas they form more pessimistic beliefs after observing negative uninformative signals. This effect becomes even more pronounced when individuals observe uninformative signals in an environment in which their beliefs matter for a payoff-relevant decision. Finally, as underlying mechanism we identify that individuals tend to process noisy information signals in a reference-dependent manner dictated by their prior beliefs. They fail to correctly identify that uninformative signals do not carry information about the objective state of the world and update their beliefs based on the valence of the signal relative to their current prior expectations.

To conclude, Chapter 4 suggest that individuals appear to struggle discerning belief-relevant information from their preferences. Such deviations from Bayesian behavior are particularly severe in situations in which the valence of non-diagnostic signals is at odds with the valence of objective pieces of information. Even though decision making frequently involves the accumulation of new pieces of information until uncertainty is reduced to a tolerable level, such a bias may instead lead to a decline in predictive performance. To illustrate this idea, suppose a judge at court is presented with undesirable (albeit irrelevant) information about a person that is at odds with objective information relevant to the legal case. If judges systematically incorporate such facts, they might eventually suppress existing information of possibly greater predictive power, potentially leading to a wrong judgement. In such environments, our results would imply that for the formation of beliefs, more signals are not always superior to less.

1.2.4 When Saving is Not Enough – Wealth Decumulation in Retirement

Chapter 5, coauthored with Martin Weber, studies individuals' decision how to decumulate wealth in retirement. In recent years, the topic of retirement planning experiences growing awareness as individuals have been increasingly expected to take responsibility for their retirement security due to demographic changes in society. However, when conducting a simple Google search on the term "retirement planning" one still finds an overwhelming share of articles which contain recommendations on saving decisions and on how to allocate savings to increase financial wellbeing in retirement. Given this prevailing focus on savings and investment decisions, one could forgive a typical retiree for believing that retirement planning is synonymous with wealth accumulation. Yet, while wealth accumulation is certainly a mandatory condition for successful retirement preparation, it is not a sufficient condition to achieve a targeted steady stream of income during retirement. In essence, retirees have not only to decide how much they want to decumulate, but also how to decumulate their savings. However, determining how to draw down his wealth is not an easy task for a person contemplating retirement, as one cannot rely on experience.

Standard economic choice theory pioneered by Yaari (1965) predicts that individuals should fully convert their accumulated savings into a lifetime annuity to maximize expected utility by smoothing their consumption. A lifetime annuity — as typically offered by insurance companies -– guarantees a certain monthly or yearly payment as long as the policyholder is alive in exchange for a lump sum of money. Despite the attractiveness of annuities as a way to protect against the risk of outliving one's retirement wealth, relatively few of those facing retirement actually annuitize a significant proportion of their wealth, a discrepancy coined the annuity puzzle. Over the past decades, researchers have tried to explain the annuitization puzzle under consideration of both rational and behavioral factors. However, adding some behavioral factors such as self-control problems, inertia, and a lack of financial sophistication only deepened the puzzle (Benartzi et al., 2011), suggesting that many individuals have an inherent aversion against annuitization.

In Chapter 5, we investigate the wealth decumulation decision from the perspective of retirees who are averse to the prospect of fully annuitizing their accumulated savings. Individuals thus face the following decision problem upon entering retirement: out of their non-annuitized wealth, they must decide how much to allocate to a savings account (e.g. as protection against early unexpected costs) and how much (if anything) to decumulate over the course of their retirement. As an alternative to annuities, we investigate consumers' preferences for phased withdrawal accounts. Phased withdrawal accounts typically involve an investment in a balanced retirement fund from which –– according to some prespecified rule –– part of the invested money is withdrawn on a regular basis to fund consumption needs. Whereas phased withdrawal accounts cannot offer protection against longevity risk, they allow retirees to retain control over their accumulated savings. In the light of recent findings, which question the benefit of full annuitization in the presence of stochastic health shocks (e.g. Reichling and Smetters, 2015; or Peijnenburg et al., 2017), such an analysis might not only provide valuable insights for the design of complementary products but also important policy implications.

To examine our research question, we field a large online survey in cooperation with a national German newspaper, *Frankfurter Allgemeine Zeitung* (FAZ), in which we elicit preferences for simple drawdown strategies. The strategies differ across two main dimensions, risky vs. risk-free asset allocation and constant vs. dynamic withdrawal rates. In the Chapter, we address the following questions: 1) what hypothetical decumulation products do individuals find most appealing?; 2) what factors do individuals rate to be most important in their wealth decumulation decision; 3) how does the demand for phased withdrawal products compare to the demand for annuities; and 4) how does retirement preparation affects individuals' willingness to decumulate wealth?

Our findings presented in Chapter 5 can be summarized as follows. First, we find that most participants prefer phased withdrawal accounts with equity-based asset allocation instead of a full risk-free allocation. Additionally, variable payout-schemes –– which adjust to economic conditions — are strongly preferred over constant payout-schemes. Second, the two considerations that respondents report being most important for their withdrawal account choice are sufficient protection against the risk of depleting the capital stock early, while also achieving relatively high returns on the invested assets. Taken together with the actual withdrawal account choice, our results highlight that customers desire flexible payout structures, which dynamically adjust in states of low returns. Most currently offered decumulation products (e.g. lifelong annuities) primarily offer constant income streams, even though there is no economic reason to do so assuming that major expenses (e.g. vacations or health costs) do not occur on a regular basis. While such income streams may allow customers to plan ahead, they could also have detrimental effects on the demand and — relatedly — the generated returns (guaranteed income streams come at the expense of less risky investment options). Third, when offered the choice between phased withdrawal accounts and annuities, a large fraction of individuals decline to annuitize and instead prefer a phased withdrawal account. This result — while surprising — is not only in line with subjects' preference to achieve higher returns on their accumulated savings while being flexible in the way they decumulate wealth but also with general findings on the annuitization puzzle. Finally, we find that participants are willing to decumulate on average 65 % of their liquid savings over the course of their retirement. In contrast to this rather high self-reported willingness to decumulate wealth, actual spending in retirement is still quite low (e.g. Olafsson and Pagel, 2018). Yet, given the low demand for annuities in our sample, we conjecture that part of this discrepancy is driven by the lack of alternative wealth decumulation products.

Taken together, Chapter 5 of this dissertation thesis contributes to the literature by studying how individuals approach the decision how to decumulate wealth and which decumulation schemes they find most appealing to transfer wealth into a stream of income. Our results have several implications for the design and the demand for complementary products. Given the considerably higher demand for phased withdrawal product accounts compared to lifelong annuities in our sample, we conjecture that offering a wider array of phased withdrawal solutions would help retirees to decumulate more of their savings, without being forced to fully convert their wealth. Offering combined solutions of phased withdrawals and partial annuitization could not only help to increase overall retirement welfare but also grant protection against longevity risk.

Chapter 2

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

2.1 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). Implicitly, these models assume 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

* 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, Konrad Stahl, 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).

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 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. bust) 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 a 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 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 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 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

[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).

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 2.2, we outline the experimental design, and briefly discuss the most important design aspects. In Section 2.3, we state our hypotheses, while in Section 2.4 we describe summary statistics of our sample and randomization checks. In Section 2.5, we present our findings, and in Section 2.6 we conclude.

2.2 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).

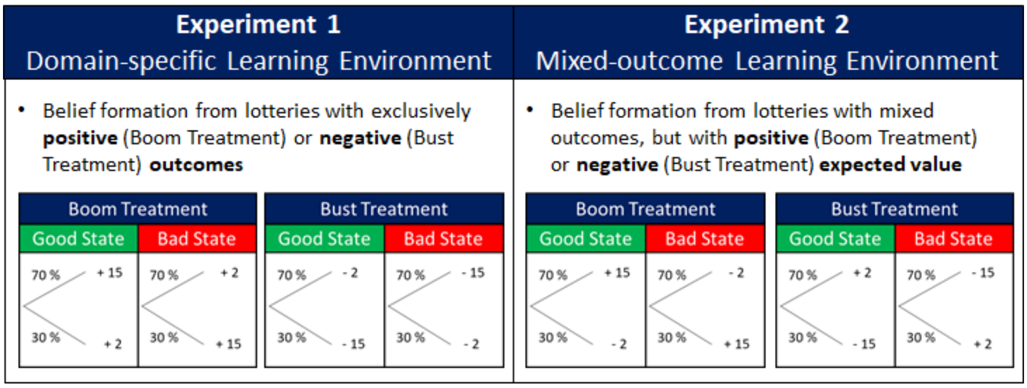
2.2.1 Detailed Description of the Experiment

Both experiments consist of two independent parts, a forecasting task (Bayesian updating) and an investment task. 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 (see Figure 2.1).

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 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 2.2.[2] To account for this fact, we conduct another experiment with a more realistic

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

Figure 2.1: Learning Environments and Treatments

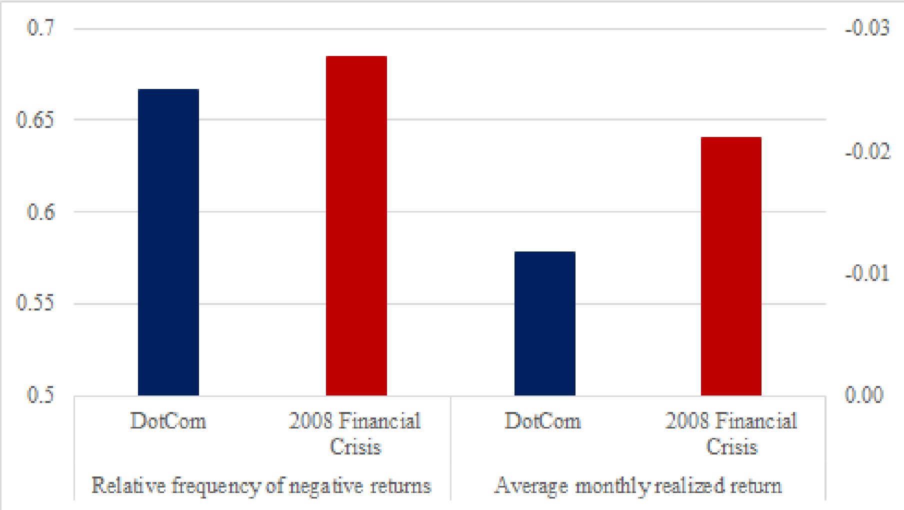


Note: This figure displays the learning environments of the first part, the forecasting task, of our two experiments. In both experiments, subjects are randomly assigned to either a Boom or a Bust Treatment. In Experiment 1, subjects learn from subsequently drawn positive (Boom) or negative (Bust) returns about the underlying state of a lottery (good or bad state). In Experiment 2, subjects learn from subsequently drawn positive and negative returns, but either from a lottery with positive (Boom) or negative (Bust) expected value about the underlying state of the lottery (good or bad state).

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).

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 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

Figure 2.2: Characteristics of Boom and Bust Market Phases

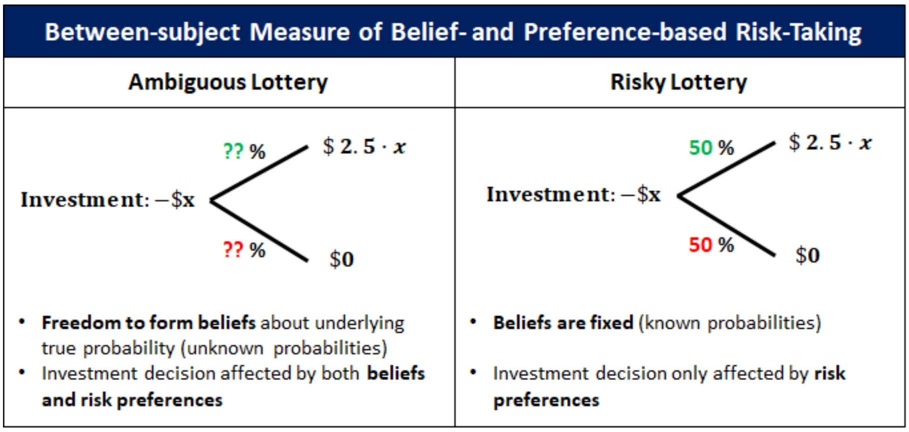


Note: This figure documents both the relative frequency of observing a negative monthly return of the MSCI All Country World Index as well as the average monthly return for the last two financial recessions. Recessions are defined according to the NBER US Business Cycle Contraction classification. The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

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. To ensure that subjects had a sufficient understanding of the forecasting task, they had to correctly answer three comprehension questions before they could continue (see Appendix A).

Figure 2.3: Between-subject Measure of Belief- and Preference-based Risk-Taking

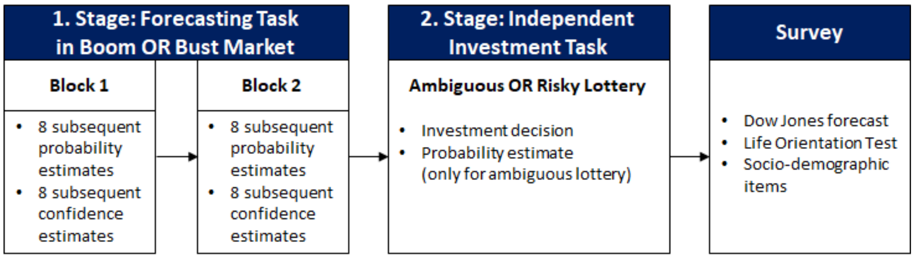


Note: This figure presents the between-subject measure of belief- and preference-based risk-taking used in the investment task of our experiments. The ambiguous lottery is characterized by unknown probabilities, whereas the risky lottery is characterized by known probabilities.

In the second part of each experiment, the investment task, we introduce a between-subject measure of belief- and preference-based risk-taking, presented in Figure 2.3. 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 statistical skills, stock trading experience and whether a participant was invested during the last financial crisis.

Figure 2.4: Structure and Flow of the Experiment



Note: This figure shows a time line of our experiments. Subjects do a forecasting task followed by an independent investment task. The forecasting task consists of two blocks. In each block, subjects have to give eight probability estimates and eight estimates about how confident they are about their forecasts. Both blocks of forecasting are either in a boom market or in a bust market environment. The random assignment of the boom or bust market environment is done at the beginning of the experiment. After the forecasting task, subjects invest either in an ambiguous lottery or in a risky lottery. For the ambiguous lottery, they are in addition asked about an estimate of the underlying success probability. The experiments end with a short survey which consists of a six-month forecast of the Dow Jones Industrial Average, a 10-item Life Orientation Test, and socio-demographic questions.

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. In summary, Figure 2.4 provides a time line of the experiments, including all described stages.

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.

2.2.2 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]

2.3 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

[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.

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.

2.4 Summary Statistics and Randomization Checks

Table 2.1 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 2.1 also reports the mean and standard deviation

Table 2.1: 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.

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.

2.5 Results

2.5.1 Main Result

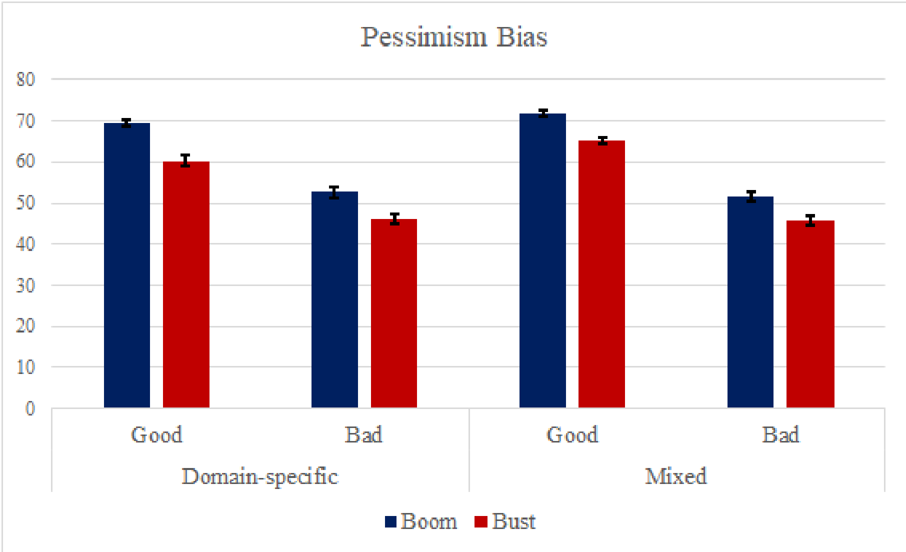
Distorted Belief Formation

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 2.5 displays the average probability estimate over eight rounds for good and bad distributions, separated by treatment and experiment.

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

Figure 2.5: 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.

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 2.2.

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

Table 2.2: 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.

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.

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 2.3 summarizes subjects' average investment in the ambiguous and risky lottery, split by treatment.

Table 2.3: 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.

The results reported in Table 2.3 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:

$$Investment_i = \beta_0 + \beta_1 Bust_i + \beta_2 Ambiguous_i + \beta_3 Bust_i x Ambiguous_i + \sum_{j=1}^{n} \beta_j X_{ij} + \epsilon_i \tag{2.1}$$

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, statistical 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 2.4, 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 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

Table 2.4: 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.

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 2.5.3.

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 in the bust treatments 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:

$$Probability_i = \beta_0 + \beta_1 Bust_i + \sum_{j=1}^{n} \beta_j X_{ij} + \epsilon_i \tag{2.2}$$

$$InvestmentAmbiguous_i = \beta_0 + \beta_1 Probability_i + \sum_{j=1}^{n} \beta_j X_{ij} + \epsilon_i \tag{2.3}$$

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 2.5 and for the second model in Table 2.6.

Table 2.5: 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 2.6: 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 2.6, 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 main 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.

2.5.2 Boundaries and External Validity

In this section, we seek 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 A). 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 A). The results for the Dow Jones return estimate are reported in Figure 2.6 Panel A, whereas the results for the Life Orientation Test are reported in Panel B.

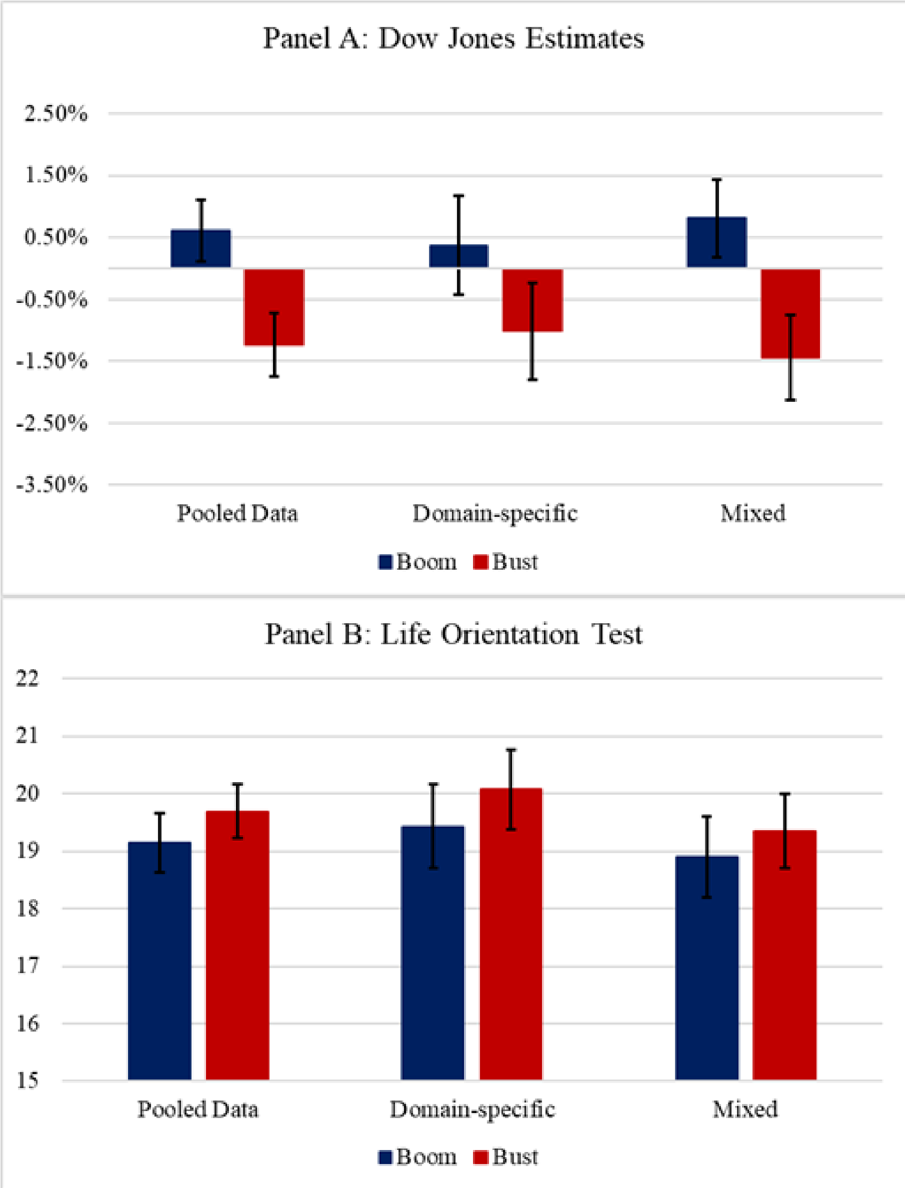
For the Dow Jones return estimates, we consistently find across all learning environments that subjects in the bust treatment are significantly more pessimistic in their return expectations. More strikingly, subjects in the bust treatment provide not only lower return estimates but also negative return 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. 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. Across all experiments 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.

2.5.3 Further Analyses and Robustness Checks

In this section, we seek to establish a more profound understanding of how subjects' forecasting abilities in the first part of the experiments affect their

Figure 2.6: Dow Jones Estimates and Life Orientation Test



Note: Panel A of the 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 B of the 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. Displayed are 95% confidence intervals.

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, $p < 0.001$).

Table 2.7: 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.

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 2.7 reports our main finding.

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 2.8.

Table 2.8: 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 A.1 in the Appendix A).

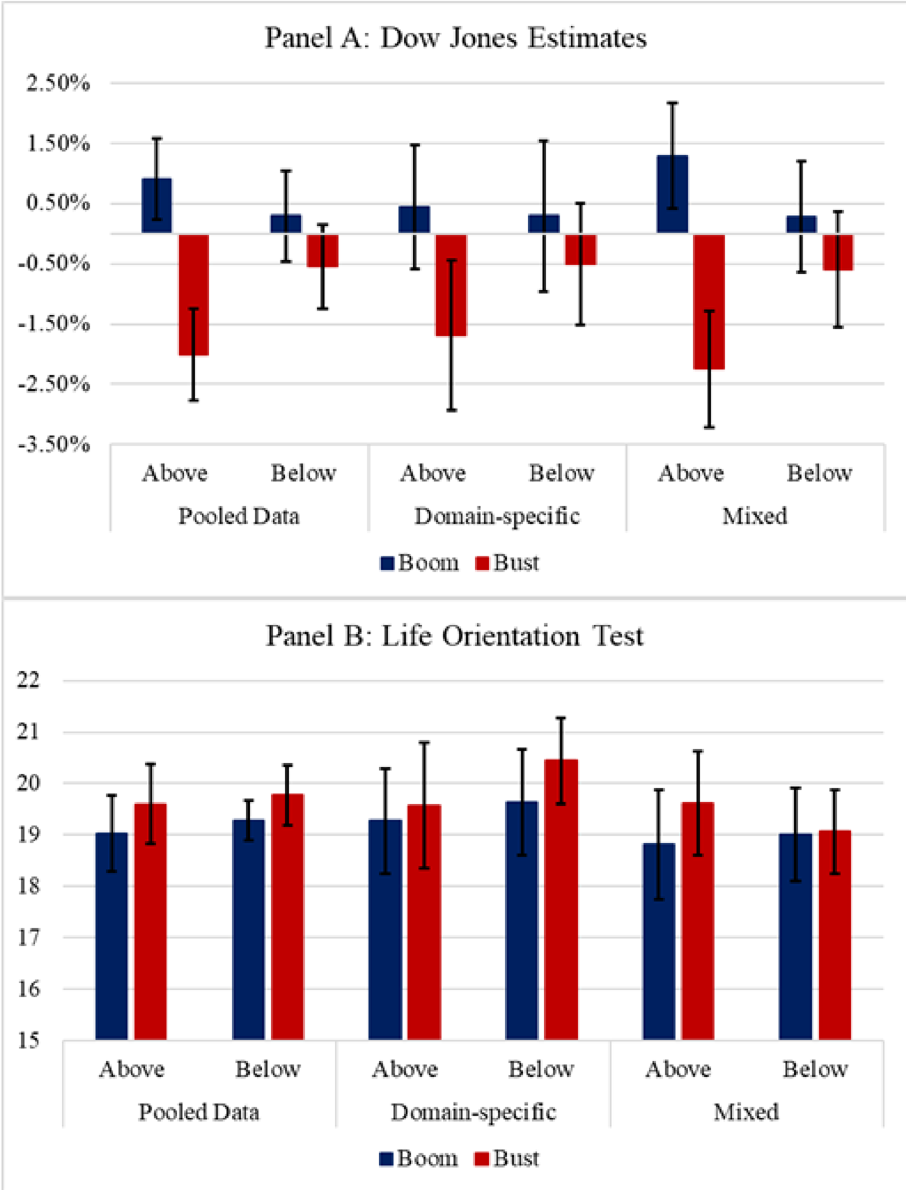
But how is it possible that the risk-taking of the seemingly better performing agents (i.e. the better forecasters) is more affected by the learning environment? One possible explanation could be that our proxy might capture participants' involvement in the experimental task. Effectively, this would 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 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 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 A). 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 A.1. 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.

Finally, we examined whether differences in the forecasting quality also affect our measures of external validity. The results are reported in Figure

Figure 2.7: Dow Jones Estimates and Life Orientation Test Split by Forecasting Ability



Note: Panel A of the figure displays subjects' self-reported return expectations of the Dow Jones Industrial average split by above- and below-forecasting ability. 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 B of the figure displays subjects' answers to a general life orientation test split by above- and below-forecasting ability. 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. Displayed are 95% confidence intervals.

2.7. We first analyze subjects' Dow Jones estimates. When split by forecasting 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, 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. Second, when analyzing subjects' answers to the Life Orientation Test split by forecasting quality, we do not find any significant difference in dispositional optimism/pessimism depending on whether subjects were in the Boom or Bust treatment.

2.6 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, 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 3

Can Agents Add and Subtract When Forming Beliefs? *

3.1 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 equally trustworthy friends who know the restaurant. Suppose, both of them recommend the restaurant, making you rather optimistic about its quality. Yet, since the restaurant is quite expensive, you decide to call two more

* 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).

friends. Assume, the first one did not like the restaurant, whereas the second did like 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.

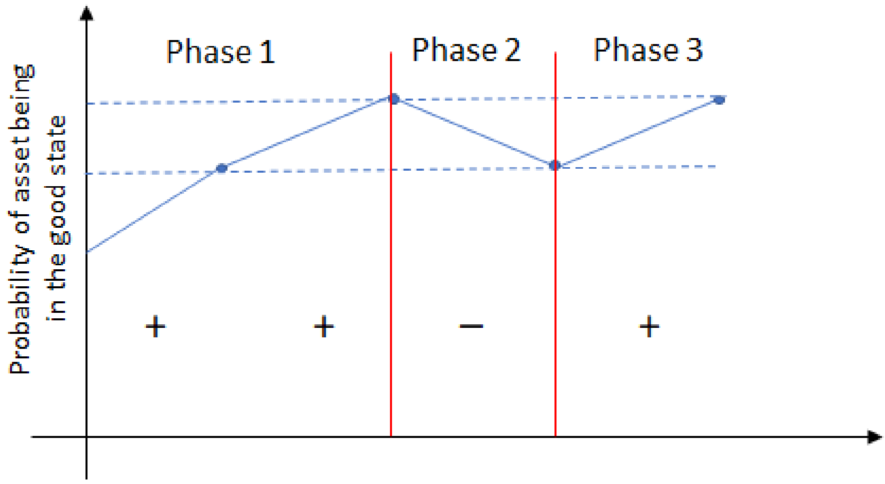
Our paper is structured as follows. In Section 3.2, we present an empirical framework, briefly review the existing literature and state our hypotheses. In Section 3.3, we describe the experimental design and summary statistics. Finally, in Section 3.4 we discuss our results and conclude in Section 3.5.

3.2 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 3.1 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

Figure 3.1: 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.

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:

$$P_t^{Bayes} = P(G|\delta_t)^{Bayes} = \frac{\theta^{\delta_t}}{\theta^{\delta_t} + (1-\theta)^{\delta_t}}, \qquad \delta_t = g_t - b_t \tag{3.1}$$

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 3.1 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 δ_t is relevant. This feature of the described framework has implications which are especially relevant for the second and the third phase in Figure 3.1. 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 3.1, 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

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

cancel the previously observed disconfirming signal and end up with the same probability estimate as he did two signals ago.

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.

3.3 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.

3.3.1 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 3.1 provides an overview of all twelve treatments.

Table 3.1: 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 B.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.

3.3.2 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 expect 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.

3.3.3 Summary Statistics

Table 3.2 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.

3.4 Results

3.4.1 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.

Table 3.2: Summary Statistics on Subjects

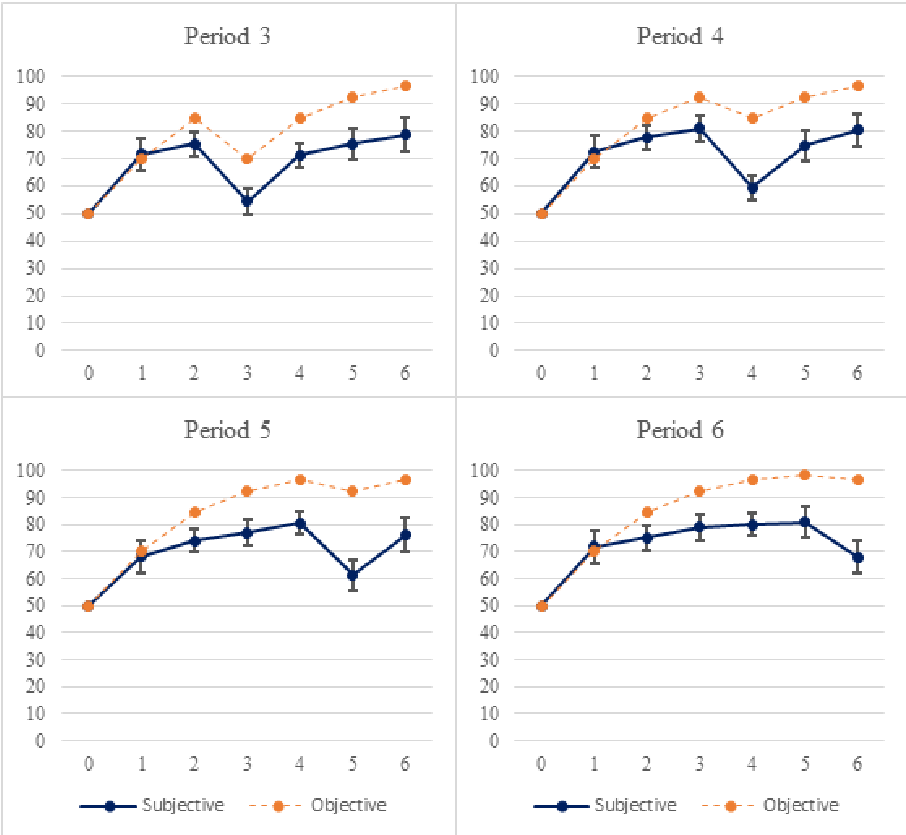
Variable	**Experiment 1** Baseline (N=601)	**Experiment 2** Reduced Diagnosticity (N=602)	**Experiment 3** Reduced Uncertainty (N=604)
Age	33.79	33.59	35.01
	(9.89)	(9.17)	(9.83)
Female	0.35	0.41	0.32
	(0.48)	(0.49)	(0.47)
Statistical Skills (1-7)	4.46	4.42	4.42
	(1.64)	(1.64)	(1.68)
Risk Preferences	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).

Baseline Results

Figure 3.2 and Figure 3.3 present subjects' average updating tendency over all periods for each treatment G-3 to G-6 and B-3 to B-6 of our baseline experiment. Figure 3.2 shows the results of those treatments in which the underlying distribution is *good* and Figure 3.3 shows the results of those treatments in which the underlying distribution is *bad*.

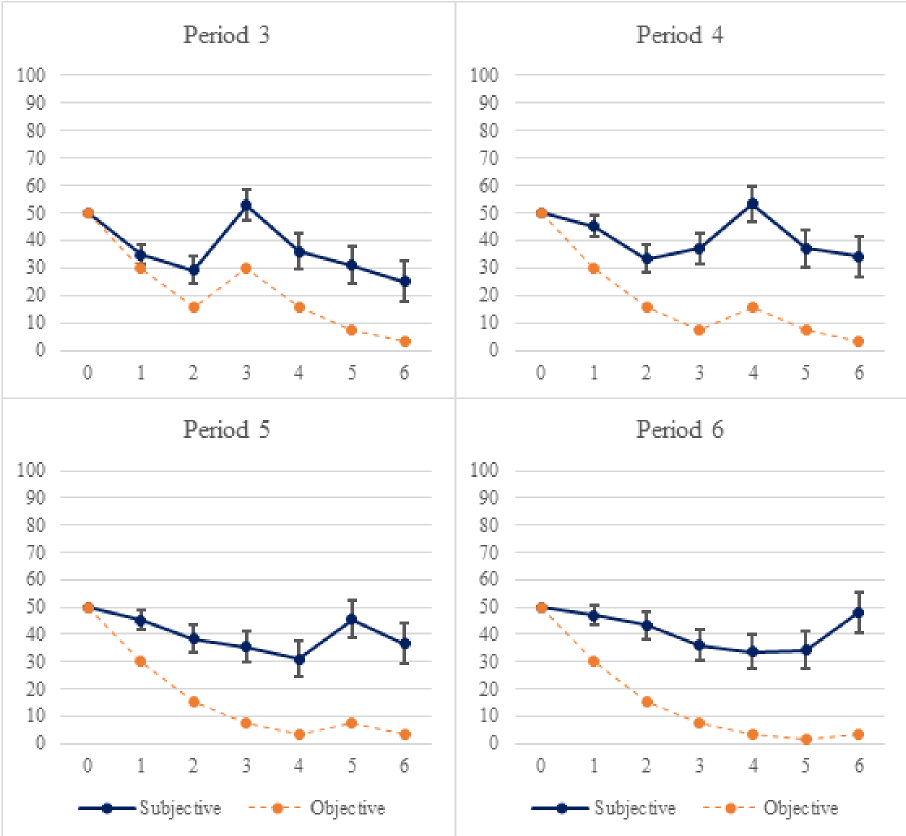
Figure 3.2: Subjects' Average Updating Behavior in the Good Distribution – Experiment 1



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.

To be consistent with our framework in Section 3.2, 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

Figure 3.3: 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 B-3 to B-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.

at the end of this chapter. From Figure 3.2, 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 3.3. 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 independent of the distribution, albeit a little bit stronger in the bad distribution. Given the difference in updating behavior, Figure

3.2 suggests that subjects strongly overreact after a disconfirming signal. In particular, subjects update their beliefs after a disconfirming signal as if they failed to incorporate up to three previously observed 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, 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]:

$$\Delta p_{i,t} = \beta_1 \Delta ObjectivePrior_{i,t} + \beta_2 Disconfirm_{i,t} + \beta_3 Correction_{i,t} + \varepsilon_{i,t}, \quad (3.2)$$

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

[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.

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 3.3.

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.

Next, we examine how our model in which we explicitly control for a disconfirming signal and a subsequent correction performs compared to the standard Bayesian 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 3.3 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.

Table 3.3: 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

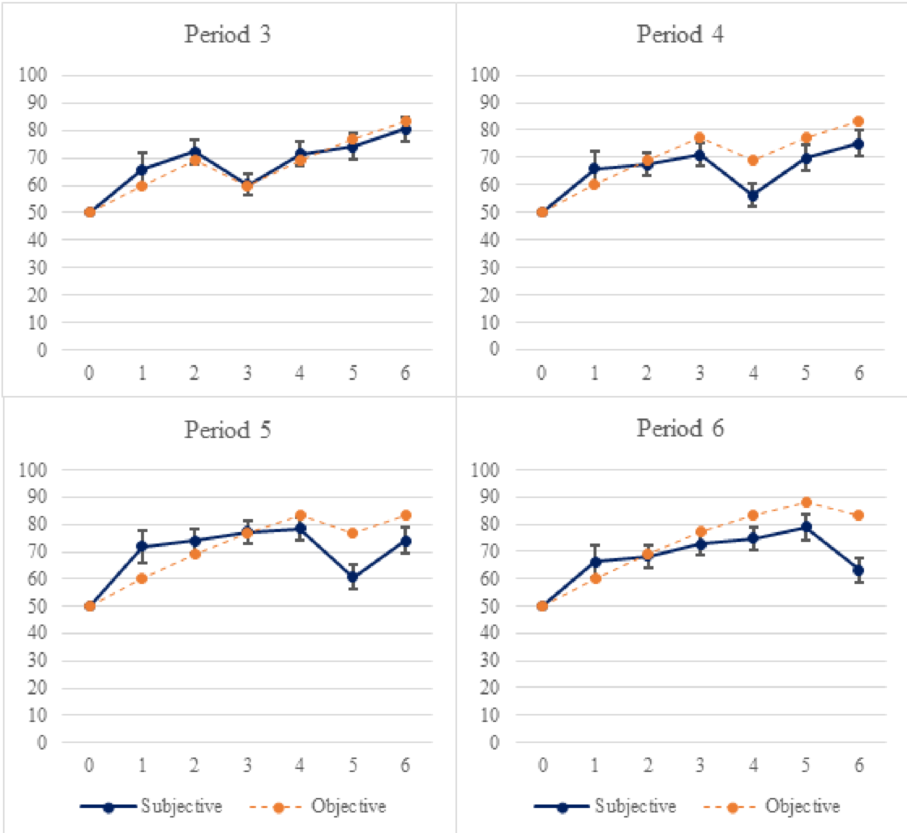
Note: This table reports the results of four OLS regressions on how subjects update their posterior beliefs after a disconfirming signal and a correction in the baseline experiment. We report the results of OLS regressions for each distribution individually (good and bad distribution). The dependent 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.

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 3.4 and Figure 3.5 present subjects' general updating behavior in the *good* and the *bad* distribution, respectively, over all periods for each treatment G-3 to G-6 and B-3 to B-6.

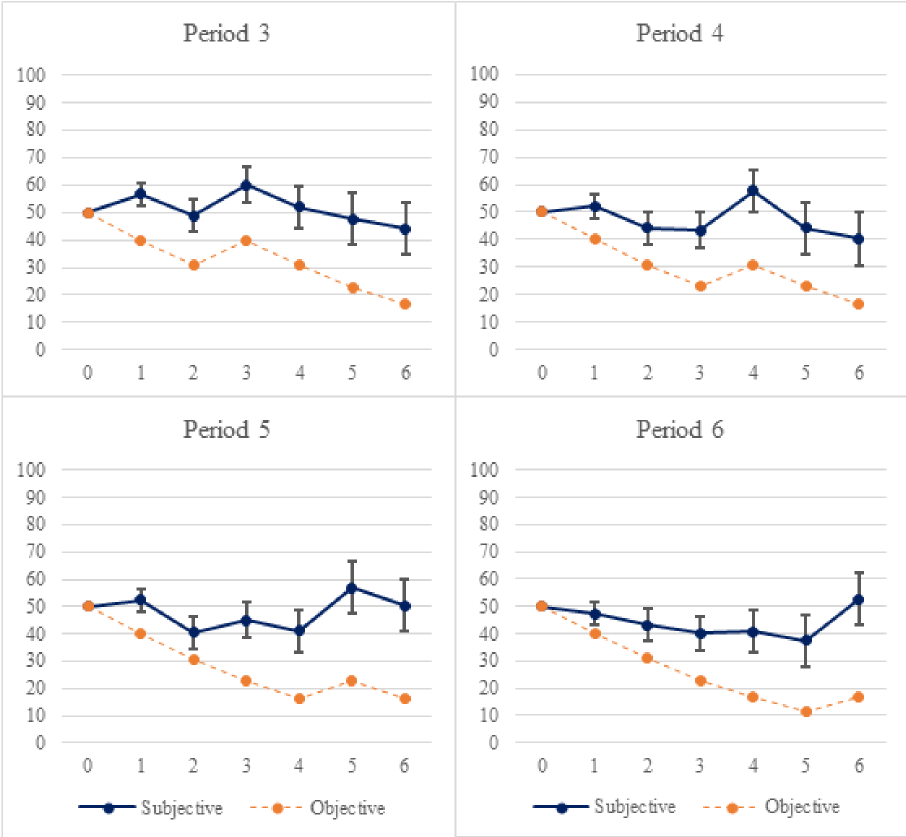
Figure 3.4: Subjects' Average Updating Behavior in the Good Distribution – Experiment 2



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.

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 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

Figure 3.5: 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 B-3 to B-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.

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 3.4. 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 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 3.4: Updating Behavior After Disconfirming Signal and Correction – Experiment 2

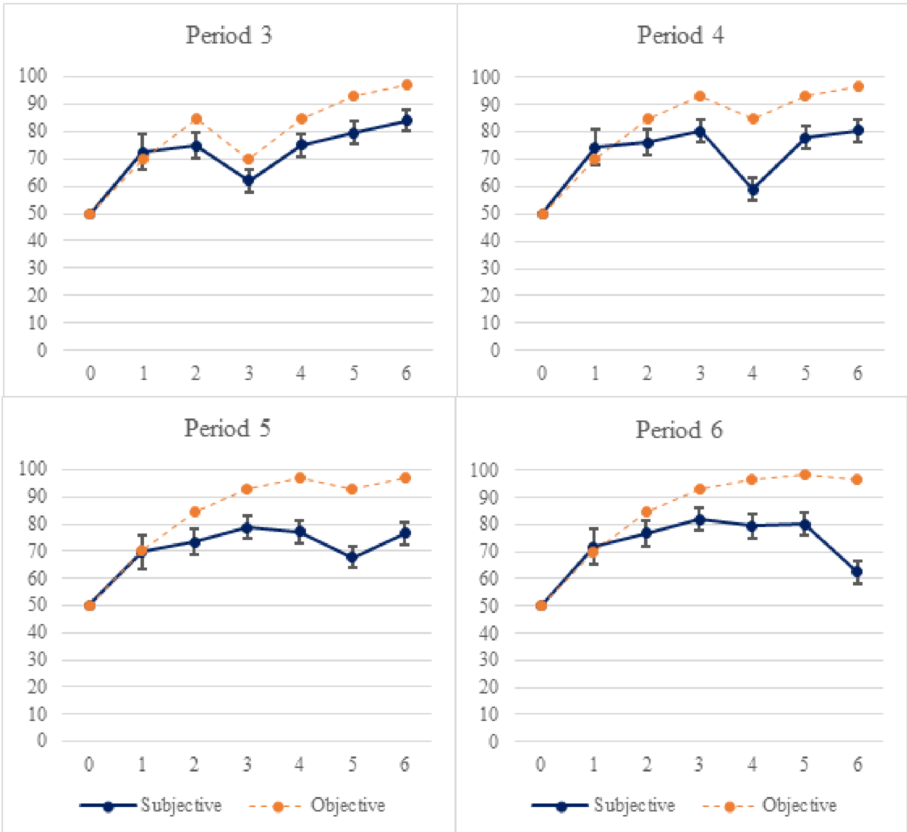
Dependent Variable	*Change in Posterior Probability Estimate*			
	Good Distribution		**Bad Distribution**	
Change in Bayes	0.860***	0.430***	0.877***	0.524***
	(16.47)	(9.08)	(13.98)	(8.84)
Disconfirm		−11.53***		10.06***
		(−8.82)		(6.74)
Correction		9.355***		−6.649***
		(6.57)		(−4.18)
Observations	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 reversals. 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 3.6 and Figure 3.7. Again, Figure 3.6 shows the results of those treatments in which the underlying distribution is *good* and Figure 3.7 shows the results of those treatments in which the underlying distribution is *bad*.

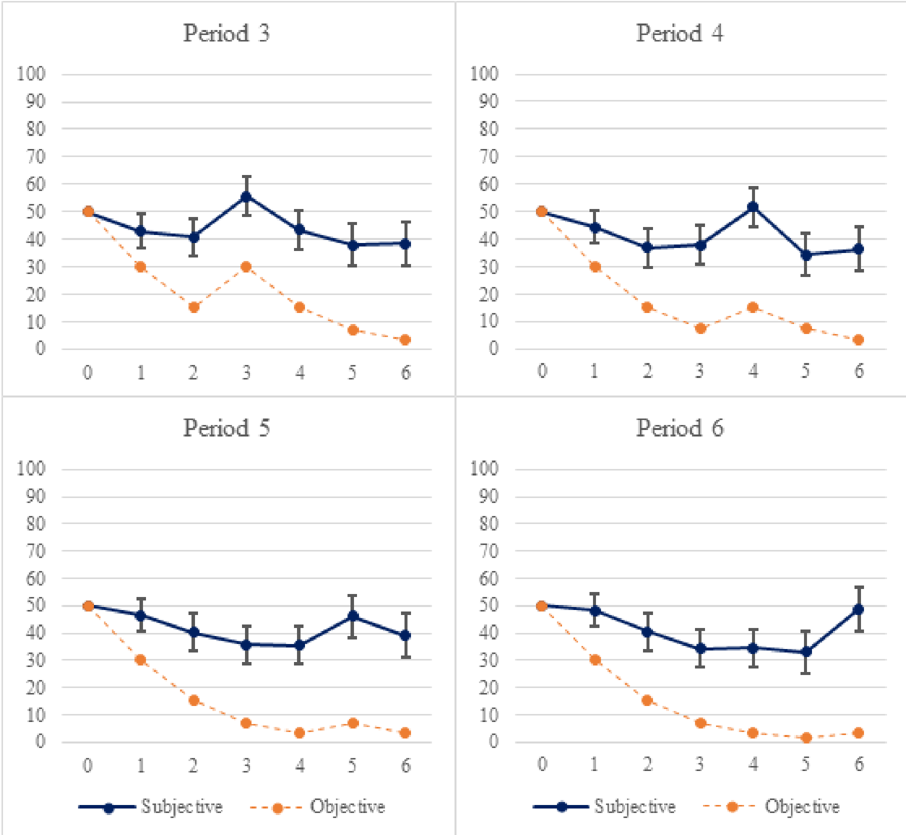
Figure 3.6: Subjects' Average Updating Behavior in the Good Distribution – Experiment 3



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.

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

Figure 3.7: 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 B-3 to B-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.

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 3.5.

Table 3.5: 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)

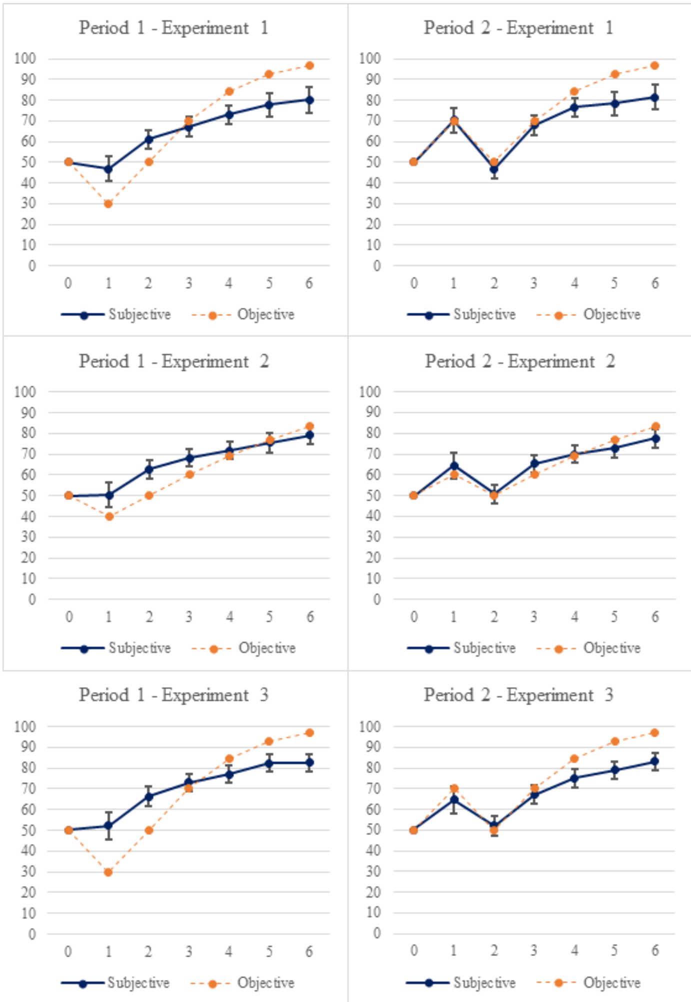
Finally, we analyze the results of treatments G-1 and G-2 (B-1 and B-2) for which – per definition – our empirical framework does not apply. In these treatments, the single opposite-directional signal occurs either directly in the first period or in the second period. As such these treatments describe 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 3.8 reports the results for the *good* distribution split by experiment. Figure 3.9 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, 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

[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.

rule because they ignored or were averse to adjust their beliefs downwards following the first bad signal.

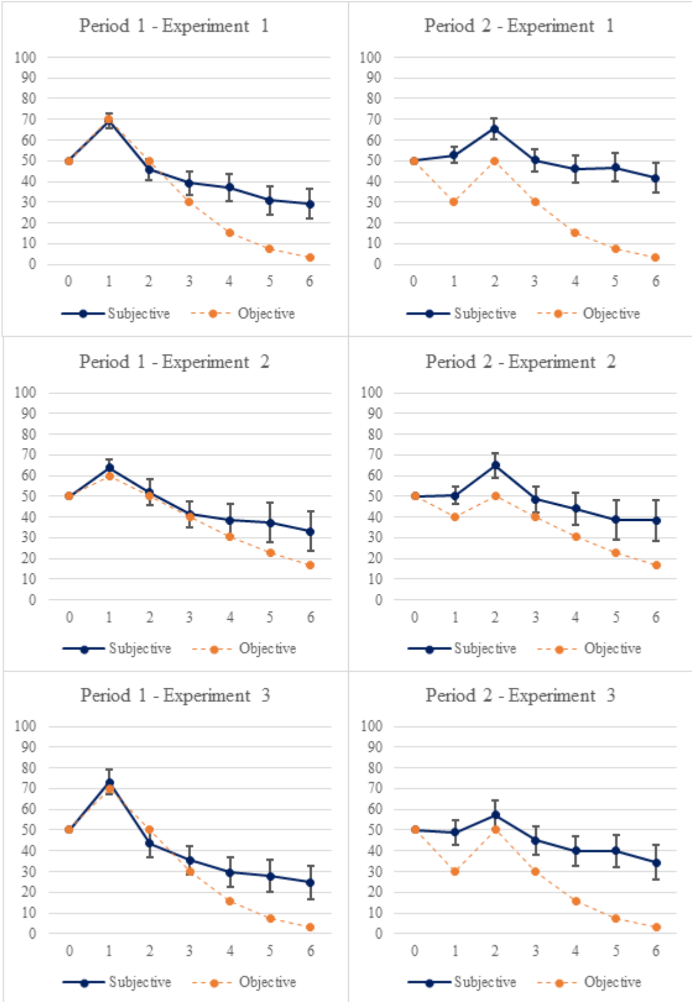
Figure 3.8: Subjects' Average Updating Behavior in the Good Distribution – Treatments G-1 and G-2



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.

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

Figure 3.9: 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 B-1 and B-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.

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.

3.4.2 Signal Ordering

One aspect of the counting heuristic we have not discussed so far is that Equation 3.1 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:

$$P_{i,6} = \beta_0 + \beta_1 D_{i \mid R=2} + \beta_2 D_{i \mid R=3} + \beta_3 D_{i \mid R=4} + \beta_4 D_{i \mid R=5} + \beta_5 D_{i \mid R=6} + \varepsilon_{i,t}, \quad (3.3)$$

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 3.6.

Table 3.6: 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.

3.4.3 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 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

Table 3.7: 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 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. Reported are coefficients and t-statistics (in parentheses) using robust standard errors. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

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 3.7. 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.

3.5 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.

Chapter 4

Expectation Formation Under Uninformative Signals *

4.1 Introduction

Probabilistic judgements are a central feature of any theory that involves decision-making under risk. As such, errors in probabilistic reasoning matter for essentially any economic decision that involves risk, including retirement, investments, purchasing insurance, or attaining various degrees of education. In the textbook model of Bayesian Updating, individuals update their prior beliefs according to Bayes' Rule upon receipt of new information. In this model, signals which do not carry relevant information about the objective state of the world play no role and are treated as if no signal occurred.

In reality however, many information structures are complex, generating signals that are often noisy and difficult to ascribe to one particular state of the world. Additionally, new information is rarely processed as being purely informative. Instead, individuals frequently have preferences over which state of the world is true, effectively generating an interaction between beliefs and preferences (Eil and Rao, 2011; Möbius et al., 2014). This interaction may lead to environments, in which information signals are non-diagnostic about an underlying state of the world, but which nonetheless appear either

* Authors: Pascal Kieren and Martin Weber. All authors are at the University of Mannheim (L9, 1-2, 68161 Mannheim). For valuable comments, we thank seminar participants at the University of Mannheim. The paper is accepted for presentation at the annual meeting of the American Economic Association 2021. We gratefully acknowledge financial support by the Reinhard Selten Scholarship by the GfeW. All remaining errors are our own.

desirable or undesirable. While Bayes' Theorem would prescribe that individuals do not update their prior beliefs in response to such *uninformative* signals, it is unclear whether individuals can correctly discern belief-relevant information from their preferences.

Taking this observation as a point of departure, we conduct an experimental study, in which we investigate how agents process signals which are non-diagnostic about the objective state of the world but which are either desirable or undesirable in the payoffs they generate. In the experiment which partly builds on Grether (1980), subjects have to incorporate a series of information signals into their beliefs to forecast the distribution of a risky asset. The risky asset can generate three outcomes from one of two distributions, a bad distribution and a good distribution. The outcomes can be ranked according to their associated payoff (high, medium, and low). In the good distribution, the high outcome occurs with the highest probability, while the low outcome occurs with the lowest probability. In the bad distribution, probabilities of the high and low outcome are reversed. Following this logic, the high outcome signals that the good distribution is more likely, whereas the low outcome signals that the bad distribution is more likely. Importantly, the medium outcome always occurs with the same probability independent of the underlying distribution. In other words, the medium outcome provides no opportunity to learn about the true state of the risky asset and will subsequently be referred to as an *uninformative* signal.

Over the course of ten rounds, subjects observe random draws from one of the two distributions and have to make a probability forecast about the likelihood that the asset is drawing from the good distribution. In our experiment, we have two key treatment variations which we exogenously vary in a between-subject design. The first treatment variation allows us to investigate how the valence of uninformative signals affects individuals' updating behavior. In the *positive* treatment, the uninformative signal pays a positive payoff, whereas in the *negative* treatment, the uninformative signal pays a negative payoff. The second treatment variation concerns the motivation to provide correct forecasts. In the *passive* treatment, subjects are asked to forecast the distribution of the risky asset after each draw and are thus only motivated to be accurate in their probability forecasts. In the *active* treatment, subjects additionally decide each round between investing in the risky asset or a riskless security which always pays the intermediate outcome. In this condition, subjects are motivated to be accurate in their forecasts and to maximize their payoffs.

In our experimental setting, we have direct control over objective expectations and can compare them to participants' subjective beliefs. Importantly, the distributions from which information is drawn are constructed in a way, that the medium outcome does not provide any information about whether subjects are currently drawing from the good or the bad distribution. As such, a Bayesian agent in our setting would not update his prior beliefs after observing an uninformative signal, independent of whether the signal is in the positive domain or in the negative domain. This allows us to disentangle the valence from the informational content of a signal and to document systematic errors in the belief formation process.

We find that individuals strongly and systematically update their prior beliefs after observing signals that are uninformative of the objective state of the world. In contrast to Bayesian behavior, individuals fail to fully extract belief-relevant information. Whereas they update their priors in on average by about 7.45 percentage points after observing informative signals, they also update their priors by 2.21 percentage points after observing uninformative signals. In relative terms, individuals adjust their priors with about 30 % of the strength as if the observed signal would carry information.

Second, we find that the direction in which individuals update their beliefs strongly depends on the valence of the observed signal. In particular, individuals tend to form more optimistic beliefs about the objective state of the world after observing positive uninformative signals, whereas they form more pessimistic beliefs after observing negative uninformative signals. This effect becomes even more pronounced when individuals observe uninformative signals in an environment in which their beliefs matter for a payoff-relevant decision. To the best of our knowledge, our study is the first test of whether individuals can distinguish their preferences from belief-relevant information in their belief formation. Additionally, we show that the effect is not driven by a few individuals who overreact to the valence of uninformative signals, but rather a general phenomenon. After observing informative signals subjects only occasionally make updating mistakes that are directionally inconsistent with Bayes' Rule (e.g. becoming more optimistic after a bad signal). However, after observing uninformative signals, subjects wrongly update their beliefs in about 68 % of the cases.

Third, as underlying mechanism we identify that individuals tend to process noisy information signals in a reference-dependent manner dictated by their prior beliefs. They fail to correctly identify that uninformative signals do not carry information about the objective state of the world and update their beliefs based on the valence of the signal relative to their current prior expectations. In particular, subjects who hold optimistic prior beliefs about the state of the risky asset (i.e. subjects who belief the good outcome is more likely to occur) only weakly increase their beliefs when the uninformative signal is positive (but in magnitude smaller than the good signal), but strongly decrease their beliefs when the signal is negative. Similarly, subjects who hold pessimistic prior beliefs about the state of the risky asset only weakly increase their beliefs when the uninformative signal is negative (but in magnitude greater than the bad signal), but strongly increase their beliefs when the signal is positive.

Research on errors in probabilistic reasoning has a long-standing tradition (Phillips and Edwards, 1966; Tversky and Kahneman, 1971, 1974). Implications of biased reasoning following new information have been studied in diverse contexts such as in psychologists' interpretation of diagnostic tests (Meehl and Rosen, 1955) , doctors' diagnoses of patients (Eddy, 1982), courts' judgments in trials (Tribe, 1971), or ideological conflicts and political discussions (Kahan, 2013). This article contributes to the literature by identifying an error in how individuals process information signals which provide no relevant learning opportunity about an objective state of the world, but which are nonetheless desirable or undesirable in the payoffs they generate. Our findings most closely relate to earlier studies which investigate base-rate neglect in response to uninformative descriptions of personality sketche (see e.g. Kahneman and Tversky, 1973; Wells and Harvey, 1978; Ginosar and Trope, 1980, 1987; Fischhoff and Bar-Hillel, 1984). These experiments typically consist of instructions which are framed to be irrelevant for judging the likelihood that a person belongs to a particular job group and find that individuals by and large draw inferences from such descriptions. Whereas these studies also examine how uninformative descriptions affect individuals' judgement about underlying probabilities, they are fundamentally different from ours as they do not investigate the influence of uninformative signals in dynamic belief updating problems. Perhaps closest to our study is the study by Troutman and Shanteau (1977), who investigate the effect of different non-diagnostic samples in bookbag-and-poker-chip experiments. They find that non-diagnostic samples result in less extreme probability assessments, as individuals effectively average across all observed signals, thereby mixing both informative and uninformative signals. Yet, different from existing work, our study emphasizes the critical role of preferences in the processing of uninformative signals. We show that, depending on the valence of the signal and individuals' prior beliefs, non-diagnostic signals can also lead to *more extreme* responses. As such, uninformative signals can not only lead to systematically biased beliefs whenever desired or undesired outcomes are non-indicative of the true state of the world, but also reinforce wrongly entertained beliefs.

Our paper also relates and contributes to the literature on prior-biased inference especially in the context of confirmation bias (e.g. Charness and Dave, 2017) and belief-polarization (e.g. Lord et al., 1979; Kahan, 2013 or Benoît and Dubra, 2018). This literature finds that individuals have a tendency to seek, interpret, and use evidence in a manner biased towards current beliefs. In belief polarization experiments, subjects with different priors are typically presented with the same mixed signals, causing their beliefs to move further apart. In these studies, signals are usually informative although noisy, effectively giving room for different interpretations. Our results highlight that even in settings in which signals are non-diagnostic of an objective state of the world, beliefs might drift apart if individuals have different priors and assign a different level of valence to the signal. In the presence of an increasing number of information sources, the mechanism presented here might further reinforce polarized beliefs.

Finally, our paper also contributes to the broad literature on motivated beliefs, which argues that beliefs are adjusted differently depending on the valence of the observed signal. Especially in the context of self-relevant beliefs, individuals appear to asymmetrically process self-serving information, putting more weight on positive than on negative information (see e.g. Eil and Rao, 2011; Sharot et al., 2011 or Zimmermann, 2020). While the beliefs we elicit in our study are not self-relevant, they are nonetheless motivated as participants are motivated to believe that the risky asset is in the good state, because the good state is more likely to result in greater payoffs. For informative signals, we find that individuals update their beliefs regarding the state of the risky asset to a greater extent following positive information than negative information. However, similar conclusions cannot be drawn regarding uninformative signals. Here, individuals appear to process the signals in a rather symmetric manner for priors close to $50-50$, becoming more optimistic after positive uninformative signals and more pessimistic after negative uninformative signals. Yet, once individuals become pessimistic, they also start to asymmetrically update their beliefs, strongly overreacting to positive uninformative signals and mostly neglecting negative uninformative signals. As in the model proposed by Bénabou (2013), this mechanism might suggest that individuals want to quickly revert very pessimistic priors to preserve anticipatory utility from putting a higher probability on the good state.

The remainder of the paper is structured as follows. Section 4.2 offers a stylized formal framework that motivates the experimental design and structures the empirical analysis. Section 4.3 presents evidence that subjects update their prior beliefs even after observing uninformative signals based on the valence of the signal and explores potential mechanisms underlying this phenomenon. Section 4.4 concludes.

4.2 Conceptual Framework

4.2.1 Setup

This section presents a stylized framework to guide the design of the experiment and to structure the main part of the empirical analysis. The underlying mechanics of the framework directly build on a reduced-form model originally introduced by Grether (1980). To keep the focus on the processing of uninformative signals, we assume only two possible states of the world, a good state (denoted as G) and a bad state (denoted as B). Consider a decision-maker (DM) who wants to learn about the current state of the world. The agent's prior beliefs are denoted by $p(g)$ and $p(b)$. To decide which state of the world is more likely, the DM receives a number of signals S, in which each signal s_t can either be informative of a good state (signal g) or of a bad state (signal b). Additionally, the DM may also receive uninformative signals (signal u), which are neither indicative of a good state nor of a bad state. As the DM observes a new signal, she updates her prior beliefs according to the following function:

$$\pi\left(G|S\right) = \frac{p(S|G)^c p\left(G\right)^d}{p(S|G)^c p\left(G\right)^d + p(S|B)^c p\left(B\right)^d} \tag{4.1}$$

$$\pi\left(B|S\right) = \frac{p(S|B)^c p\left(B\right)^d}{p(S|G)^c p\left(G\right)^d + p(S|B)^c p\left(B\right)^d} \tag{4.2}$$

where $p(.)$ refers to a true conditional probability, $\pi(.)$ refers to an agent's (potentially biased) belief, and $c, d \geq 0$. The parameter c governs the (biased) use of likelihoods, while the parameter d governs the (biased) use of prior beliefs. To interpret the magnitudes of c and d, we follow Benjamin (2019), and write the model in the posterior-odds form, dividing (4.1) by (4.2):

$$\frac{\pi\left(G|S\right)}{\pi\left(B|S\right)} = \left[\frac{p\left(S|G\right)}{p\left(S|B\right)}\right]^{c}\left[\frac{p\left(G\right)}{p\left(B\right)}\right]^{d}. \tag{4.3}$$

In this equation, $c < 1$ corresponds to updating as if the signals provided less information about the state (underinference), while $c > 1$ corresponds to updating as if the signals provided more informative than they do (overinference). Similarly, $d < 1$ corresponds to treating the priors as less informative than they are (also referred to as base-rate neglect), while $d > 1$ corresponds to the opposite. The model nests Bayes' Theorem as a special case, in which $c = d = 1$.

To infer the underlying state of the world, consider that a DM receives each period t a new signal, which can either be good, bad, or uninformative. In a signal structure where only two signals carry information about the underlying state of the world, the conditional probability of being in the good state given the signal history at time t $(\pi\left(G|S\right))$ can be calculated as follows:

$$\pi^{Bayes}\left(G|S\right) = \frac{\theta^{z_t}}{\theta^{z_t} + \left(1-\theta\right)^{z_t}}, \qquad z_t = g_t - b_t \tag{4.4}$$

where g_t (b_t) denotes the number of good (bad) signals that have been observed until period t and z_t is the difference between both. The parameter $\theta \in [0,1]$ captures the diagnosticity of an informative signal. Since a Bayesian DM would neglect uninformative signals, only the difference of good and bad signals is of relevance. Additionally, note that he is indifferent regarding the order in which the signals occur.

Following Charness and Dave (2017), we make use of the fact that the natural logarithm of the odds ratio within a round $(\ln\left(\frac{\pi^{Bayes}(G|z_t)}{\pi^{Bayes}(B|z_t)}\right))$ for a Bayesian is given by[1]:

$$\pi_t^{Bayes} = \ln\left(\frac{\pi\left(G|s_1,\ldots,s_t\right)}{\pi\left(B|s_1,\ldots,s_t\right)}\right) = \ln\left(\frac{\theta}{1-\theta}\right)\cdot z_t \tag{4.5}$$

[1] A detailed explanation and derivation is provided in Appendix C.4.

As such, the Bayesian log-odds ratio is updated by $\pm\theta \cdot z_t$ after each new signal and − in contrast to Bayes probability − linear in the number of signals. To make the interpretation easier, we take the first-difference of both sides of the equation, yielding:

$$\Delta\pi_t = \pi_t - \pi_{t-1} = \ln\left(\frac{\theta}{1-\theta}\right) \cdot \Delta z_t, \tag{4.6}$$

where $\Delta\pi_t \in \{-\theta, \theta\}$ and $\Delta z_t \in \{-1, 0, 1\}$. The interpretation of $\Delta\pi_t$ is straightforward. If the DM observes a new good (bad) signal, then the Bayesian log-odds ratio is updated by $\theta(-\theta)$. If the DM observes an uninformative signal (i.e. $\Delta z_t = 0$), then the Bayesian log-odds ratio remains constant.

To incorporate that a non-Bayesian DM may falsely incorporate an uninformative signal in his belief updating process, we consider the possibility that c may not only depend on the information of a signal, but also on the valence. In our setting, valence can be broadly defined as any signal that does not help the DM to learn about the current state of the world but which provides either positive or negative utility (e.g. through a payoff or other factors that might be relevant for the DM):

$$\frac{\pi(G|S)}{\pi(B|S)} = \left[\frac{p(S|G)}{p(S|B)}\right]^{c_0 + I\{u|\ desirable\}c_1 + I\{u|\ undesirable\}c_2} \left[\frac{p(G)}{p(B)}\right]^{d}, \tag{4.7}$$

where $I\{u|\ desirable\}$ equals 1 if $s = u$ and the signal is perceived as desirable and $I\{u|\ undesirable\}$ equals 1 if $s = u$ and the signal is perceived as undesirable. In this case, we obtain three reduced-form parameters which describe biased inference: c_0 captures inference of informative signals, while c_1 and c_2 capture uninformative signals which are desirable or undesirable, respectively.

Equation (4.7) will be the core expression to investigate individuals' propensity to update after uninformative signals. It nests the Bayesian prediction that priors are fully incorporated in the belief formation process (i.e. $d = 1$) and that individuals respond with a coefficient of $c_0 = 1$ to informative signals. Finally, since a Bayesian DM would not update his prior beliefs after an uninformative signal (i.e. $\Delta z_t = 0$), we would expect that $c_1 = c_2 = 0$.

4.2.2 Experimental Design

To study the degree to which individuals' belief formation process is affected by uninformative signals, we require an environment in which (i) individuals repeatedly incorporate signals with varying degrees of information into their beliefs; (ii) Bayesian beliefs can be clearly identified; (iii) treatment variations allow the exogenous variation of the desirability of uninformative signals; (iv) holding positive/negative beliefs has a value in and of itself; and (v) the belief elicitation is incentive-compatible. The design of our experimental study was built to accommodate these features.

The experiment consists of two parts, the main task (a forecasting task in the spirit of Grether, 1980) and a brief survey. In the forecasting task, subjects receive information about a risky asset, whose payoffs are either drawn from a "good distribution" or from a "bad distribution". Both distributions have three outcomes, which are identical across distributions but differ with respect to the probability with which they occur. All three outcomes can be ranked based on the payoff they generate and are thus labeled high, medium, or low. In the good distribution, the high payoff occurs with a 50 % probability, while the low payoff occurs with a 20 % probability. In the bad distribution, probabilities are reversed, i.e. the low payoff occurs with a 50 % probability, while the high payoff occurs with only 20 % probability. Importantly, and crucial for the experimental design, the medium payoff always occurs with 30 % probability, irrespective of whether the distribution is good or bad. This ensures that the medium outcome does not provide any information about the underlying distribution, from which outcomes are drawn.

We introduce a 2x2 between-subject design with respect to the forecasting task. The first treatment dimension to which subjects are assigned depicts the potential payoffs of the two distributions. In particular, subjects are

Table 4.1: Payoff Distribution

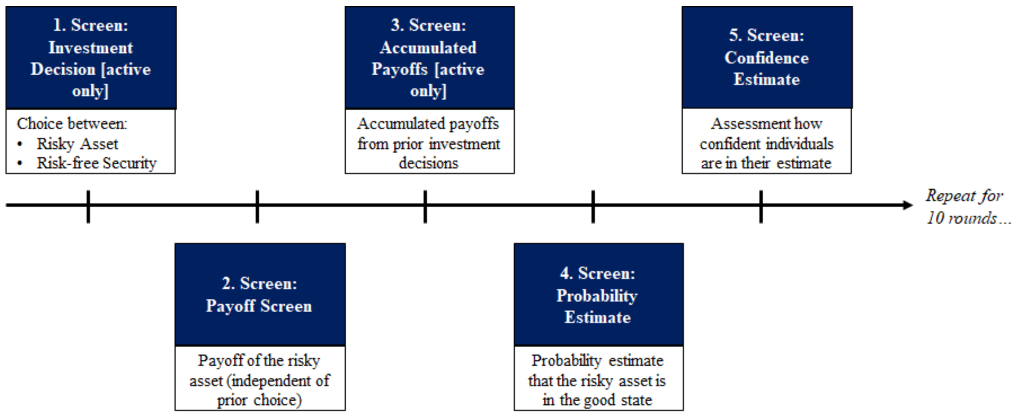
Payoffs	**Good Signal**	**Uninformative Signal**	**Bad Signal**
Positive Treatment	+5	+1	−3
Negative Treatment	+3	−1	−5

Note: This table reports the payoffs associated with good, uninformative, and bad signals, split by positive and negative treatment.

randomly assigned to either a *"positive"* condition or a *"negative"* condition. The three possible payoffs in the positive condition are +5 (high outcome), +1 (medium outcome), or -3 (low outcome). In the negative condition, all outcomes are shifted by -2 to keep the higher moments of the distribution constant while reducing the mean. As such, the three possible payoffs in the negative condition are +3 (high outcome), -1 (medium outcome), or -5 (low outcome). Table 4.1 provides a brief overview of possible outcomes across treatments.

The second treatment dimension relates to the set of questions subjects have to answer in the forecasting task. Subjects can be assigned to an *"active"* or a *"passive"* condition. In both conditions, subjects observe a payoff of the risky asset in ten consecutive rounds. Before the first round, the computer randomly determines the distribution of the risky asset (which can be good or bad). In the active condition, subjects decide before the beginning of each round whether they want to invest in the risky asset (whose distribution they have to forecast) or a bond, which always pays the medium outcome (-1, or +1; depending on the treatment) for sure. The payoff of the bond is thus equal to the expected value of the risky asset when no information about the underlying distribution is available. If the good distribution becomes more likely (i.e. occurs with a probability of greater than 50 %), the expected value of the risky asset is greater then the expected value of the bond, and vice versa. After their decision to invest in one of the two securities, subjects observe the payoff of the risky asset (irrespective of their choice) and are reminded of how much they have earned so far given their prior choices. Finally, subjects are asked to provide an estimate of the probability that the risky asset was paying from the good distribution and to rate their confidence in this estimate (assessed on a seven-point Likert scale). In the passive condition, subjects do not make any investment decision and start each round by observing the payoff of the risky asset in that trial. Afterwards, they are immediately asked to provide an estimate of the probability that the risky asset was paying from the good distribution and to rate their confidence in this estimate (also assessed on a seven-point Likert scale). An overview of all questions and the order in which subjects answer the questions is provided in Figure 4.1.

Figure 4.1: Overview of the Updating Task



Note: This figure provides an overview of the questions subjects have to answer in the forecasting task. Subjects in the passive treatment have to answer three questions in each round, whereas subjects in the active treatment have to answer five questions in each round (denoted with [active only]). Subjects make forecasting decisions in 10 consecutive rounds.

To avoid potentially confounding factors resulting from biased memories (see Gödker et al., 2019), we explicitly display the prior outcomes in a table next to the questions. Additionally, we recognize that belief updating is an abstract task for many individuals. To ensure that subjects have a sufficient understanding of the forecasting task, they had to correctly answer three comprehension questions before they could continue (see Appendix C.2 for the exact wording).

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). The latter two measures were obtained by asking subjects two questions regarding a portfolio allocation problem. In the first question, participants had to allocate \$10,000 between a broadly diversified index fund and a savings account. This answer provides a proxy for their risk preferences. The second question asked subjects to identify the correct formula for calculating the expected value of the portfolio they selected. Through multiple-choice answers, we can detect whether people lacked an understanding of probabilities, of the difference between net and gross returns, or of the difference between stocks and savings accounts. This yielded a financial knowledge score between zero to three (exact wording of questions is provided in the Appendix).

In the active condition participants were paid based on their investment payoffs and the accuracy of the probability estimates provided. Specifically, they received one twentieth of their accumulated payoffs (in the negative condition, all outcomes were shifted by +2 for the final calculation to make payment equivalent), plus 10 Cents for each probability estimate within 5 % of the objective Bayesian value. As such, subjects were motivated to be accurate in their forecasts and to maximize their payoffs. In the passive condition participants were paid based on the accuracy of the provided probability estimates, with the same rules as in the active condition.[2]

4.2.3 Hypotheses Development

To obtain parameter estimates for our main specification in the conceptual framework, we estimate a regression based on the natural logarithm of Equation (4.7).[3]

[2] While the resulting payment for the passive condition was lower on average, participants also completed the experiment faster.

[3] The detailed derivation can be found in Appendix C.4.

$$\ln \frac{\lambda(G|s_1,\dots,s_t)}{\lambda(B|s_1,\dots,s_t)} = \widehat{\beta_1}\cdot D_{informative} + \widehat{\beta_2}\cdot \ln \frac{\lambda(G|s_1,\dots,s_{t-1})}{\lambda(B|s_1,\dots,s_{t-1})} + \widehat{\beta_3}\cdot D_{uninformative|desirable} + \widehat{\beta_4}\cdot D_{uninformative|undesirable} + \varepsilon_t \quad (4.8)$$

Note that within a round t, the natural logarithm of subject's i odds ratio, based on her stated probability $P_{it}\left(G|s_1,\dots,s_t\right)$ that the asset is paying dividends from the good state is:

$$\lambda_{it} = \ln\left(\frac{\lambda\left(G|s_1,\dots,s_t\right)}{\lambda\left(B|s_1,\dots,s_t\right)}\right) = \ln\left(\frac{P_{it}\left(G|s_1,\dots,s_t\right)}{1-P_{it}\left(G|s_1,\dots,s_t\right)}\right) \quad (4.9)$$

which may differ from the objective Bayesian probability (π_{it}). To make sure that the above ratio is defined for all observations, we truncated the data to lie in the [0.01, 0.99] interval. The interpretation of λ_{it} is straightforward. If λ_{it} is greater than (less than) zero, then person i beliefs in round t that the asset is more (less) likely to draw from the good state.

In the regression specification, we replaced $\ln\frac{p(S|G)}{p(S|B)}$ with a dummy $D_{informative}$ taking the value 1 if the tth signal is g, 0 if the tth signal is u, and -1 if the tth signal is b (see Appendix C.4). While this specification is equivalent (see Benjamin, 2019), we need to interpret the coefficient $\widehat{\beta_1}$ relative to $\left(\frac{\theta}{1-\theta}\right)$ instead of 1. Even though we have three possible outcomes in our experimental environment, which occur with 50 % (signal g or b), 30 % (signal u), and 20 % (signal g or b), only two of them are informative about the objective state of the world (signal g and signal b). Thus, the diagnosticity of an informative signal is set to $\theta = \frac{0.5}{0.5+0.2} = 0.714$ in our experiment and we need to interpret the coefficient $\widehat{\beta_1}$ relative to $\left(\frac{\theta}{1-\theta}\right) = \frac{0.714}{1-0.714} \approx 0.916$.

The regression specification allows us to control for several deviations from Bayesian behavior simultaneously, while testing whether individuals systematically incorporate uninformative signals into their beliefs. More precisely, if individuals are subject to conservatism (overreaction), one would expect $\widehat{\beta_1} < (>)\ 0.916$. Similarly, if individuals put too little (much) weight on their priors, one would expect $\widehat{\beta_2} < (>)\ 1$. Importantly, if people falsely incorporate uninformative signals in their belief formation process, one would observe that both $\widehat{\beta_3}$ and $\widehat{\beta_4}$ predict log-odds. In contrast, a test of Bayesian behavior would be:

$$\widehat{\beta_1} = \ln\left(\frac{\theta}{1-\theta}\right) = 0.916;\ \ \widehat{\beta_2} = 1;\ \ \widehat{\beta_3} = \widehat{\beta_4} = 0$$

4.2.4 Summary Statistics

Table 4.2 presents summary statistics. Overall, six-hundred forty-one individuals (420 males, 221 females, mean age 33 years, 8.8 years standard deviation) were recruited from Amazon Mechanical Turk (MTurk) to participate in an online experiment. MTurk advanced to a widely used and accepted recruiting platform for economic experiments. Not only does it offer a large and diverse subject pool 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). Participants reported average statistic skills of 4.71 out of 7 and would invest roughly 49 percent of their hypothetical endowment into a risky fund, which will serve as a proxy of risk aversion. Moreover, participants achieved a financial literacy score of approximately 1.34 out of 3.

Additionally, we tested whether the randomization from our between-subject design successfully resulted in a balanced sample. Table 1 also reports the mean and standard deviation for each control variable split by whether the uninformative signal was positive or negative (Panel A) and whether participants played the active or passive version of the forecasting task (Panel B). Differences were tested using rank-sum tests, or χ^2-tests for binary variables. Generally, we barely find any significant difference between our treatments, suggesting that our randomization was successful. The only exceptions are

Table 4.2: Summary Statistics

Panel A Variable	Full Sample (N=641)	Negative Treatment (N=321)	Positive Treatment (N=320)	Difference	p-value
Age	33.00 (8.79)	33.00 (9.04)	33.00 (8.55)	0.00	1.00
Female (1 = Yes)	0.34 (0.48)	0.36 (0.48)	0.33 (0.47)	0.03	0.38
Statistic Skills (1-7)	4.71 (1.76)	4.76 (1.73)	4.66 (1.78)	0.10	0.47
Risk Preferences (% invested in risky asset)	0.49 (0.31)	0.49 (0.30)	0.50 (0.31)	0.01	0.51
Financial Literacy	1.34 (0.91)	1.43 (0.88)	1.26 (0.93)	0.17	0.01

Panel B Variable	Full Sample (N=641)	Passive Treatment (N=330)	Active Treatment (N=311)	Difference	p-value
Age	33.00 (8.79)	33.49 (8.47)	32.48 (9.11)	1.01	0.15
Female (1 = Yes)	0.34 (0.48)	0.36 (0.48)	0.32 (0.47)	0.04	0.30
Statistic Skills (1-7)	4.71 (1.76)	4.55 (1.83)	4.87 (1.67)	0.31	0.02
Risk Preferences (% invested in risky asset)	0.49 (0.31)	0.48 (0.28)	0.51 (0.33)	0.03	0.18
Financial Literacy	1.34 (0.91)	1.35 (0.92)	1.34 (0.89)	0.01	0.85

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 (Panel A for Positive/Negative; Panel B for Active/Passive). 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. *Statistic skills* denotes participants' self-assessed statistical skills on a 7-point Likert scale. *Risk Preferences* is the percentage of their initial endowment that subjects invested in a risky investment option. *Financial Literacy* is a score between zero (lacking basic understanding) to three (correctly answered each question).

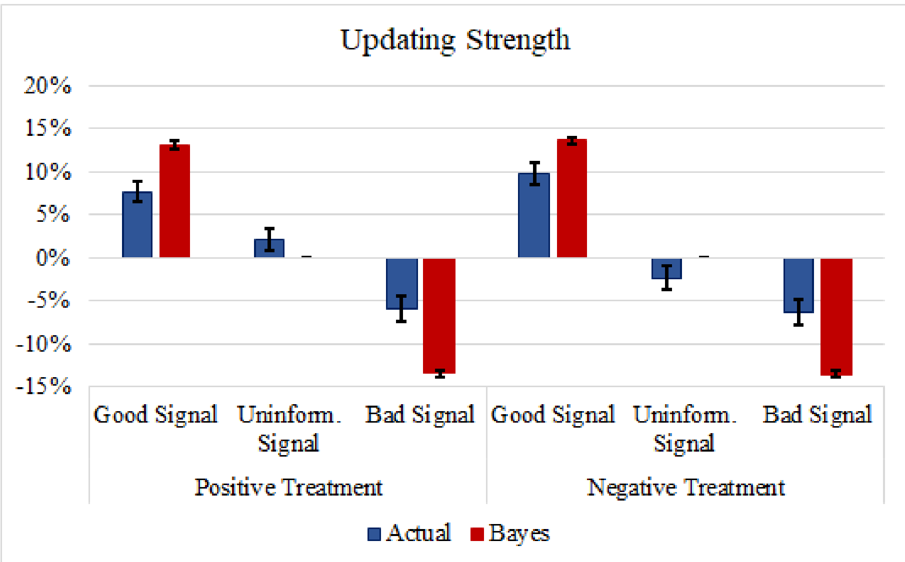
minor differences in financial literacy for the first treatment dimension (Panel A) and minor differences in self-reported statistic skills for the second treatment dimension (Panel B). Due to the random allocation across treatments, these differences arise most likely due to chance. Nevertheless, we control for these factors in all our further analyses. For the remaining variables, we cannot reject the null hypothesis that the socio-economic background of the subjects is balanced between our treatments.

4.3 Results

4.3.1 Belief Updating after Uninformative Signals

Before we delve into the statistical analysis, Figure 4.2 visualizes participants general updating tendency after good, bad, and uninformative signals and compares it to Bayesian behavior. The figure displays results separately by whether the uninformative signal was positive (positive treatment) or negative (negative treatment).

Figure 4.2: General Updating Tendency



Note: This figure illustrates subjects' general belief updating after observing good, uninformative, and bad signals about the state of the risky asset. Displayed are actual changes in prior beliefs as well as the correct Bayesian change in probability. Results are displayed separately by whether subjects encountered the uninformative signal in the positive or negative domain. Displayed are 95% confidence intervals.

As can be inferred, subjects' beliefs adjust in the appropriate direction after both good and bad (informative) signals. Relative to Bayesian beliefs, we observe conservatism on average as subjects generally update too little both after good and after bad news. However, even after uninformative signals, subjects' beliefs adjust substantially and in the direction of the domain of the uninformative signal. While a Bayesian decision maker would not update his prior beliefs at all, subjects increase their priors after observing a positive uninformative signal, whereas they decrease their priors after observing a negative uninformative signal. Additionally, the strength with which they update their priors is symmetric for positive and negative uninformative signals.

While the pattern in Figure 4.2 provides first insights into subjects' updating behavior it is, of course, insufficient to justify a causal interpretation. To provide more formal evidence of how individuals update their priors after observing uninformative signals, we estimate OLS regressions of Equation (4.8):

$$\begin{aligned} \lambda_{i,t} = \widehat{\beta_1} \cdot D_{informative;\, i,t} + \widehat{\beta_2} \cdot \lambda_{i,t-1} \\ + \widehat{\beta_3} \cdot D_{uninformative;\, i,t} + \widehat{\beta_4} \cdot D_{uninformative;\, i,t}\ x\ negative_i + \varepsilon_t \end{aligned} \tag{4.10}$$

where participants' subjective log-odds ratio is the dependent variable, and $D_{informative;i,t}$ is a variable taking the value 1 if the tth signal of subject i is $good$, 0 if the tth signal is $uninformative$, and -1 if the tth signal is bad. $\lambda_{i,t-1}$ denotes the use of priors (i.e. the base-rate) and is defined as $\ln \frac{\lambda(G|s_1,\ldots,s_{t-1})}{\lambda(B|s_1,\ldots,s_{t-1})}$. $D_{uninformative;\, i,t}$ is a dummy if the tth signal of subject i is uninformative, whereas $negative_i$ is a dummy if participant i is in the negative treatment (and zero otherwise). The interaction term thus displays whether participant i encountered a negative uninformative signal in round t. Finally, X_{ij} is a set of control variables including age, gender, statistic skills, risk-aversion, and financial literacy. Results for the full sample and split by active and passive treatment are reported in Table 4.3.[4]

Table 4.3: Uninformative Updating

Dependent Variable	Log Odds Ratio (Subjective) $\lambda_{i,t}$		
	Full Sample	Active	Passive
$D_{informative;\ i,t}$ *(Inference)*	0.472***	0.369***	0.569***
	(15.22)	(9.06)	(12.39)
$\lambda_{i,t-1}$ *(Use of Priors)*	0.697***	0.758***	0.624***
	(32.99)	(27.91)	(19.60)
$D_{uninformative;\ i,t}$	0.256***	0.329***	0.152**
	(5.89)	(5.16)	(2.56)
$D_{uninformative;\ i,t}$ x	−0.514***	−0.563***	−0.443***
$negative_i$	(−7.19)	(−5.40)	(−4.54)
Observations	5769	2799	2970
R^2	0.538	0.611	0.468

Note: This table reports the results of three OLS regressions on how information signals and their valence affect individuals' beliefs. We report results for the full sample and split by active and passive treatment. The dependent variable is participants' subjective log-odds ratio as defined in Section 2. $D_{informative;\ i,t}$ is a variable taking the value 1 if the tth signal of subject i is *good*, 0 if the tth signal is *uninformative*, and -1 if the tth signal is *bad*. $D_{uninformative;\ i,t}$ is a dummy if the tth signal of subject i is uninformative, whereas $negative_i$ is a dummy if participant i is in the negative treatment (and zero otherwise). The interaction term thus displays whether participant i encountered a negative uninformative signal in round t. Controls include age, gender, statistical skills, risk aversion, and participants' financial literacy. Reported are coefficients and t-statistics (in parentheses). All standard errors are clustered at the individual level. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Our results suggest that Bayesian behavior is not predominant in the data. Specifically, we observe that $\widehat{\beta_1} \neq 0.916$, $\widehat{\beta_2} \neq 1$, and that both $\widehat{\beta_3}$ and $\widehat{\beta_4} \neq 0$. To interpret in which way individuals depart from Bayesian behavior, it is instructive to review what it would mean for individual coefficient estimates to vary from their Bayesian counterparts. Since $\widehat{\beta_1} < 0.916$ and $\widehat{\beta_2} < 1$, individuals suffer both from conservatism (i.e. they underinfer) and base-rate neglect (i.e. they under-use prior information). Most importantly however, we find that $\widehat{\beta_3} > 0$, whereas $\widehat{\beta_4} < 0$. In other words, even though a Bayesian would not update his prior beliefs after observing an uninformative

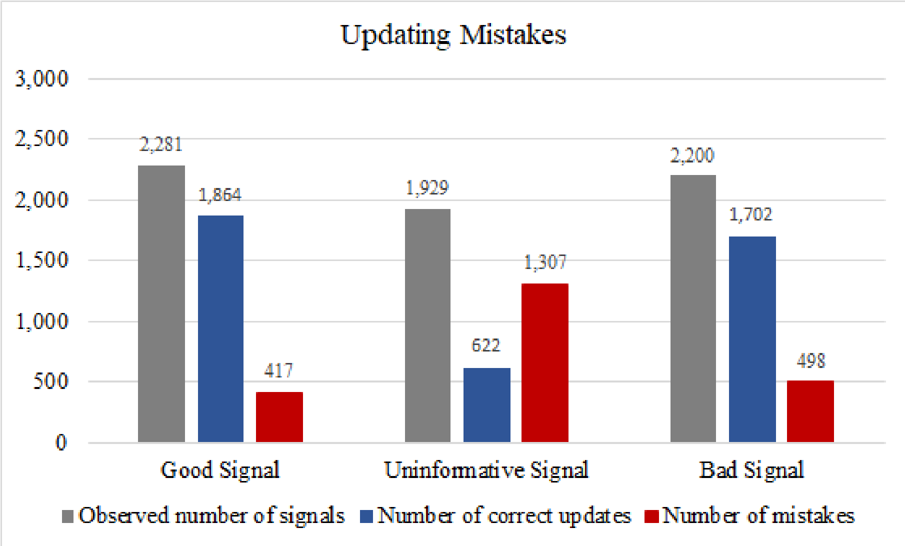
[4] Regression specifications are chosen to be identical to theory (i.e. estimated without a constant). However, other specifications yield similar results. We opt to present the simplest possible evidence.

signal, both positive and negative uninformative signals predict log-odds ratios. More precisely, controlling for both conservatism and base-rate neglect, individuals on average increase their priors after observing a positive uninformative signal with about half the strength as if the signal would contain information. Given the magnitude of both $\widehat{\beta_3}$ and $\widehat{\beta_4}$, this effect is mostly symmetric, suggesting that individuals also decrease their priors with about half the strength as if a negative uninformative signal would contain information. Lastly, we observe stronger effects when individuals have the opportunity to invest in the asset compared to when they simply state their beliefs. Taken together, this suggests that having stakes in the task exacerbates the bias resulting from uninformative signals, potentially because individuals *hope* to observe positive payoffs to maximize their earnings.

4.3.2 Frequency of Updating Mistakes

Thus far, we have established that individuals *on average* incorporate even uninformative signals into their beliefs based on the valence of the signal. However, these average patterns may mask a substantial amount of heterogeneity. In particular, it is not clear whether our results are driven by a few individuals who neglect the informational content but strongly focus on the valence or whether the here reported updating tendency applies to a large share of individuals, thus being a rather general phenomenon. To draw inference about the relation between the informational content and the valence of signals and to determine which aspect is most prevalent when processing uninformative signals, we examine how frequently individuals falsely update their beliefs after observing uninformative signals. To investigate the frequency, we define any belief update that is directionally inconsistent with the observed signal as an updating mistake. In the case of informative signals, an updating error is thus defined as a decrease (increase) in prior beliefs that the asset is drawing from the *good state* after subjects observed a good (bad) signal. Similarly, for uninformative signals, an updating error is defined as *any* update in prior beliefs after having observed an uninformative

Figure 4.3: Basic Updating Mistakes



Note: This figure illustrates the number of directionally inconsistent belief updates relative to the overall number of observed signals. Results are displayed separately for good, uninformative, and bad signals.

signal. Importantly, the definition above does not rely on the magnitude of the error, but only on the occurrence of such an error.

Figure 4.3 plots the absolute number of good, uninformative, and bad signals, as well as the number of mistakes after observing any of the three signals.

Across all rounds and subjects, there are a total of 2,281 good signals, 1,929 uninformative signals, and 2,200 bad signals. Looking at informative signals, subjects only made basic errors in about 20 % of the cases (18 % and 23 %, for good and bad signals, respectively). However, the rate at which subjects perform basic errors is substantially higher for uninformative signals. Here subjects updated their beliefs in 68 % of the cases, even though the signal did not provide any learning opportunity about the underlying distribution. While Figure 4.3 already shows that the frequency of errors is substantially different for informative and uninformative signals, we further validate the robustness of the finding in a linear probability model. To do so, we estimate the following model[5]:

[5] While we estimate the model using OLS, results remain unchanged if we use a logit model instead.

$$
\begin{aligned}
Error_{i,t} = {} & \beta_0 + \beta_1 D_{uninformative;\, i,t} + \beta_2 Objective\ Prior_{i,t} \\
& + \beta_3 Subjective\ Prior_{i,t} + \beta_4 Confidence_{i,t} + \sum_{j=1}^{n} \beta_j X_{ij} + \varepsilon_{i,t}, \quad (4.11)
\end{aligned}
$$

where $Error_{i,t}$ is defined as individual i performing an updating error that is directionally inconsistent with the observed signal in round t. $Objective\ Prior_{i,t}$ is the rational prior for individual i as prescribed by Bayes' Theorem given the observed outcome history in round t, while $Subjective\ Prior_{i,t}$ is subjects' probability estimate in round t. Finally, $Confidence_{i,t}$ is subjects' self-reported confidence in their ability to provide correct probability forecasts. Results are reported in Table 4.4.

Consistent with our prior conjecture, we find that observing an uninformative signal in a given round substantially increases the likelihood of conducting an updating error. More precisely, we find that the likelihood of conducting an error is roughly 50 percentage points higher after observing an uninformative signal compared to observing a signal that does carry information about the underlying distribution. Interestingly, while this effect does not largely depend on the valence of the signal (Columns 3 and 4) it is less pronounced in the active treatment and more pronounced in the passive treatment (Column 2). The latter difference might be driven by the fact that subjects in the active treatment can derive payoff-relevant information from inferring the correct state of the underlying asset. As such, they might pay more attention to the information structure of the signals, thereby reducing their propensity to update their beliefs in response to uninformative signals. Besides the treatment, we find that the probability of updating one's beliefs in the wrong direction also correlates to participants' confidence in their own forecasts. Those individuals who are more confident that their forecast is correct are also less likely to update their beliefs in response to uninformative signals, suggesting that individuals are mindful about their own ability to provide correct forecasts.

Table 4.4: Frequency of Directionally Inconsistent Updating Errors

Dependent Variable	*Updating Error*			
	Full Sample	Full Sample	Positive Treatment	Negative Treatment
$D_{uninformative;\, i,t}$	0.475***	0.517***	0.449***	0.506***
	(33.42)	(26.54)	(22.35)	(25.08)
Objective Posterior	−0.00046	−0.00046*	−0.00012	−0.00079**
	(−1.63)	(−1.67)	(−0.30)	(−2.10)
Subjective Probability Estimate	0.00067**	0.00077**	0.00010	0.0014***
	(1.99)	(2.31)	(−0.20)	(3.17)
Confidence Estimate	−0.0147***	−0.0153***	−0.0111*	−0.0166***
	(−3.71)	(−3.84)	(−1.81)	(−3.15)
Active		−0.00983		
		(−0.59)		
$D_{uninformative;\, i,t}$ *x Active*		−0.0850***		
		(−3.03)		
Constant	0.255***	0.256***	0.275***	0.241***
	(5.34)	(5.37)	(3.90)	(3.63)
Observations	6410	6410	3210	3200
R^2	0.224	0.227	0.204	0.250

Note: This table reports the results of four OLS regressions on how frequently individuals perform directionally inconsistent updating mistakes. We report results for the full sample and split by positive and negative treatment. The dependent variable is *Updating Error*, a dummy that equals 1 if participants perform a updating mistake that is directionally inconsistent with Bayes' Rule. $D_{uninformative;\, i,t}$ is a dummy taking the value 1 if the tth signal is *uninformative*, and 0 otherwise. *Objective Posterior* is the correct Bayesian probability that the risky asset is in the good state, given the information seen by the participant up to trial t in the learning block. *Subjective Probability Estimate* and *Confidence Estimate* are participants' estimates of the probability that the risky asset is in the good state and their assessed confidence, respectively. Controls include age, gender, statistical skills, risk aversion, and participants' financial literacy. Reported are coefficients and t-statistics (in parentheses). All standard errors are clustered at the individual level. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Taken together, the analysis reinforces our prior evidence that individuals face difficulties in discerning the informational content of a signal from its valence. Additionally, the effect appears to be a general and quite robust phenomenon, as individuals more frequently update their beliefs after informationally irrelevant signals than they do not.

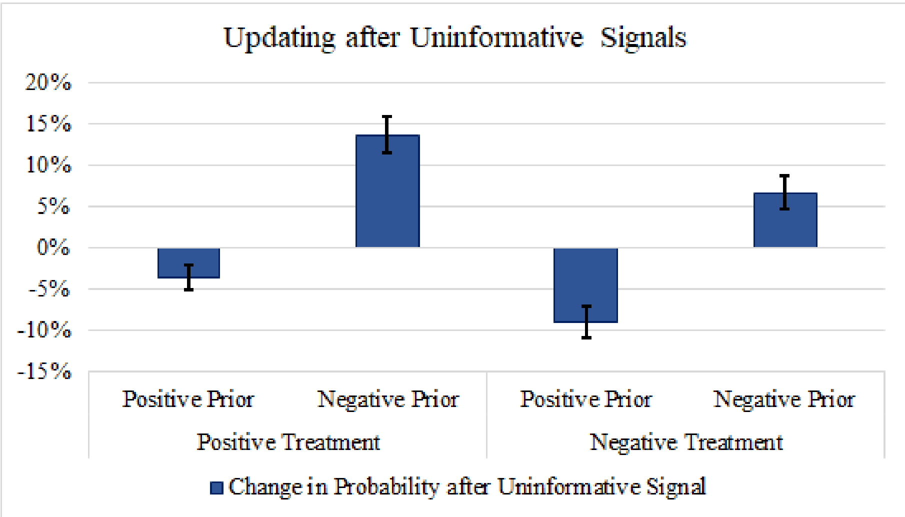
4.3.3 Mechanism: Reference-dependent Belief Updating

One common and persistent finding so far is that individuals not only face difficulties in correctly identifying the informational content of signals but also that they struggle to discern it from the valence of the signal. In this section, we explore potential mechanisms underlying this pattern.

To do so, we focus on the influence of participants' prior beliefs about the state of the risky asset. Prior beliefs have previously been shown to affect updating mistakes and biased inference in multiple ways and thus serve as a natural starting point for our analysis. Testing the implications of a model by Rabin and Schrag (1999) both Charness and Dave (2017) and Pouget et al. (2017), find evidence that individuals draw inference in a manner that is biased in favor of current beliefs about the objective state of the risky asset. A related body of research documents that people update their beliefs about future outcomes in an asymmetric manner: they tend to neglect undesirable information, and overweight desirable information (Eil and Rao, 2011; Möbius et al., 2014; Sharot and Garrett, 2016). In our experiment, prior beliefs are important for two reasons. First, when deciding between investing in the risky asset or choosing the risk-free alternative, holding a particular belief has direct consequences for the investment decision. As such, beliefs have a value in and of themselves, as positive beliefs about the state of the risky asset are related to higher potential payoffs. Second, and perhaps even more important, extreme priors (both optimistic and pessimistic) are usually the result of observing one particular signal more frequently than the other signals (i.e. very optimistic beliefs usually develop in response to observing many *good* signals). In our environment, good signals are always associated with the highest payoff, whereas bad signals are associated with the lowest payoff and uninformative signals with a medium payoff (as illustrated in Table 4.1). Thus, extreme priors (which develop in tandem with the associated high, medium, or low payoffs) might shift participants' reference point. To illustrate this idea, consider a participant who frequently observes the *good* signal with the respective high payoff. Such a participant might react differently when observing a positive uninformative signal (with medium payoffs) compared to someone who frequently observes *bad* signals.

To differentiate optimistic from pessimistic priors, we define a prior that the asset is drawing from the good distribution greater than 50 % as positive prior and a prior that the asset is drawing from the good distribution smaller than 50 % as negative prior[6]. Figure 4.4 visualizes participants updating behavior after non-diagnostic signals, split by treatment (positive vs. negative) and by prior.

Figure 4.4: Prior Dependent Updating



Note: This figure illustrates the change in prior beliefs after observing uninformative signals split by positive and negative prior beliefs and by treatment. Results are displayed separately for the positive and negative treatment. Displayed are 95% confidence intervals.

Figure 4.4 reveals that priors appear to play an important role in processing uninformative signals. Those subjects who hold positive priors only update their beliefs weakly in response to positive uninformative signals, whereas they update strongly in response to negative uninformative signals. Symmetrically, subjects who hold negative priors substantially increase their priors after observing positive uninformative signals, whereas they only weakly update their priors after observing negative uninformative signals.

6 This definition is consistent with the point where one of the two assets has a higher expected value. For priors greater (smaller) than 50 percent, the expected value of the risky asset is greater (smaller) than the expected value of the riskless security.

This pattern suggests that not only the domain of the uninformative signal is important but also how desirable the signal is in relation to what subjects expect to observe.

To more rigorously investigate how prior beliefs affect our previously reported results, we estimate OLS regressions of our main specification (Equation 4.8) split by prior beliefs and by whether participants are invested in the risky asset or not. Importantly, we only include participants from the active treatment in the analysis, as the decision to be invested in the asset or not is most likely a deliberate choice that depends on prior beliefs [7]. Table 4.5 reports results.

Coefficient estimates for actively invested participants reveal a fundamental asymmetry in how the processing of uninformative signal depends on prior beliefs. Those subjects who hold optimistic prior beliefs about the state of the risky asset (i.e. subjects who belief the good outcome is more likely to occur) only weakly increase their beliefs when the uninformative signal is positive (but in magnitude smaller than the good signal), but strongly decrease their beliefs when the signal is negative. Similarly, subjects who hold pessimistic prior beliefs about the state of the risky asset only weakly increase their beliefs when the uninformative signal is negative (but in magnitude greater than the bad signal), but strongly increase their beliefs when the signal is positive. As such, subjects appear to process uninformative signals not exclusively on the basis of the valence of the signal, but rather relative to some reference point which is dictated by their prior beliefs.

Importantly, this finding cannot be explained by prior-biased inference (or confirmation bias) as tested by Charness and Dave (2017) and Pouget et al. (2017) as individuals show stronger reactions to the valence uninformative signals that contradict their prior beliefs. Additionally, this finding is also different from preference-biased inference (Eil and Rao, 2011; Möbius et al., 2014) as individuals also overreact to undesirable signals. Subjects both

[7] Results for participants in the passive treatment are directionally consistent. However, we decide to present the results that are undoubtedly affected by participants' prior beliefs.

Table 4.5: Reference Dependent Belief Updating

Dependent Variable	*Log Odds Ratio (Subjective)* $\lambda_{i,t}$			
	Actively Invested		Not Invested	
	Positive Prior	Negative Prior	Positive Prior	Negative Prior
$D_{informative;\, i,t}$ *(Inference)*	0.354***	0.459***	0.214***	0.592***
	(7.13)	(6.01)	(2.81)	(4.85)
$\lambda_{i,t-1}$ *(Use of Priors)*	0.797***	0.755***	0.748***	0.814***
	(22.55)	(11.04)	(14.88)	(9.77)
$D_{uninformative;\, i,t}$	0.213**	0.740***	0.307**	0.221
	(2.38)	(3.52)	(2.51)	(1.00)
$D_{uninformative;\, i,t}$ x	−0.579***	−0.842***	−0.545**	−0.302
$negative_i$	(−3.58)	(−3.21)	(−2.52)	(−0.79)
Observations	1371	530	533	287
R^2	0.698	0.454	0.580	0.538

Note: This table reports the results of four OLS regressions on how information signals and their valence affect individuals' beliefs. We report results split by whether participants are actively invested in the risky asset or not and by participants' prior beliefs about the state of the risky asset as defined in Section 3.3. The dependent variable is participants' subjective log-odds ratio as defined in Section 2. $D_{informative;\, i,t}$ is a variable taking the value 1 if the tth signal of subject i is *good*, 0 if the tth signal is *uninformative*, and -1 if the tth signal is *bad*. $D_{uninformative;\, i,t}$ is a dummy if the tth signal of subject i is uninformative, whereas $negative_i$ is a dummy if participant i is in the negative treatment (and zero otherwise). The interaction term thus displays whether participant i encountered a negative uninformative signal in round t. Controls include age, gender, statistical skills, risk aversion, and participants' financial literacy. Reported are coefficients and t-statistics (in parentheses). All standard errors are clustered at the individual level. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

strongly react to positive uninformative signals that contradict pessimistic priors as well as to negative uninformative signals that contradict optimistic priors. Instead, it appears that subjects incorporate uninformative signals in a reference-dependent manner, dictated by their prior beliefs. They fail to correctly identify that uninformative signals do not carry information about the objective state of the world and update their beliefs based on the valence of the signal relative to their current prior expectations. Yet, given that we find that individuals react strongest to desirable uninformative signals when their priors are pessimistic, it appears that they seek to revert very pessimistic priors most quickly, consistent with the model of Bénabou (2013).

However, similar conclusions cannot be drawn for subjects who are not actively invested in the risky asset. While those who hold positive priors update their beliefs rather symmetrically in the direction of the domain of the uninformative signal, individuals who hold negative priors do not appear to update their beliefs at all after uninformative signals. One potential explanation for this behavior is that subjects who are rather optimistic about the state of the risky asset (while not being invested) still follow the outcomes to invest in a future round once they become more certain of the state. Finally, individuals who are not invested and hold pessimistic beliefs might simply not pay enough attention to the outcomes, as they continue to collect the risk-free payoff.

Taken together, our results suggest that being invested in the risky asset appears to be a necessary condition for subjects to engage in reference-dependent updating following uninformative signals. More generally, there has to be some intrinsic or extrinsic advantage for holding a particular belief such as making a payoff-relevant investment decision.

4.3.4 Robustness Checks

The Role of Memory and Learning

Two important concepts related to the formation of probabilistic beliefs are the role of memory and learning effects. However, our experiment was constructed in a way to ensure that neither of the effects can account for our findings. First, subjects are always provided with the full outcome history of prior signals. In particular, as can be seen in the Appendix, the history of prior signals is clearly displayed next to the forecasting question. Additionally, our experimental design does not provide feedback and hence little scope for learning. Moreover, it is highly doubtful that subjects would learn within the course of ten experimental periods even in the presence of feedback. To verify that the effect is not driven by initial forecasting errors when subjects lack the experience and potentially less pronounced the more signals individuals observe, we separately estimate our main specification for the first five and the final five signals that individuals observe. Table C.1 in the Appendix reveals that coefficient estimates for positive and negative uninformative signals remain relatively stable throughout the experiment. Together with the fact that we did not provide any feedback, we conclude that the effect appears to be stable over time.

Sample Splits

Finally, we replicate our main analyses on different subsamples to validate its robustness. In particular, we conduct splits regarding (i) extreme outliers; (ii) "speeders"; and (iii) forecasting performance. Extreme outliers are individuals whose subjective priors largely deviate from the Bayesian benchmark and who frequently update in the wrong direction. Similar to the exclusion criteria 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 quintile of the response time distribution. Finally, we also conduct splits regarding how subjects overall performed in the forecasting task. To verify that the effect does not capture those individuals who showed difficulties in understanding the task, 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. Results are reported in Table 4.6.

Overall, results are very similar across all subsamples and confirm our previously drawn conclusions. First, we consistently find that uninformative signals predict log-odds ratios in every subsample. Second, similar to our main analysis, positive uninformative signals predict an increase in the log-odds ratio, whereas negative uninformative signals predict a decrease, with the effect being of similar strength. Lastly, we also find differences between the different subgroups. In particular, extreme outliers, speeders and individuals with below-median forecasting performance show more pronounced effects both for positive and negative uninformative signals.

Table 4.6: Robustness Checks

Dependent Variable	*Log Odds Ratio (Subjective)* $\lambda_{i,t}$					
	(1) Outlier?		(2) Speeder?		(3) Above Median Forecaster?	
	No	Yes	No	Yes	No	Yes
$D_{informative;\ i,t}$ *(Inference)*	0.561*** (17.54)	0.168** (2.41)	0.493*** (15.69)	0.259** (2.02)	0.263*** (5.48)	0.610*** (16.71)
$\lambda_{i,t-1}$ *(Use of Priors)*	0.734*** (35.29)	0.533*** (10.32)	0.715*** (33.48)	0.547*** (6.83)	0.545*** (14.74)	0.800*** (41.65)
$D_{uninformative;\ i,t}$	0.238*** (4.88)	0.383*** (4.70)	0.226*** (5.23)	0.702*** (3.41)	0.418*** (5.99)	0.161*** (3.48)
$D_{uninformative;\ i,t}$ *x* $negative_i$	-0.444*** (-6.09)	-0.701*** (-4.18)	-0.436*** (-6.15)	-1.296*** (-4.31)	-0.716*** (-6.43)	-0.343*** (-4.47)
Observations	4275	1494	5193	576	2880	2889
R^2	0.638	0.340	0.557	0.477	0.747	0.344

Note: This table reports the results of six OLS regressions to investigate the robustness of our main finding. We report sample splits based on three measures as defined in Section 3.4: (1) strong outliers; (2) speeder; and (3) forecast quality. The dependent variable is participants' subjective log-odds ratio as defined in Section 2. $Prior\ Signal_{i,t}$ is a variable taking the value 1 if the tth signal of subject i is *good*, 0 if the tth signal is *uninformative*, and -1 if the tth signal is *bad*. $Uninformative_{i,t}$ is a dummy if the tth signal of subject i is uninformative, whereas $negative_i$ is a dummy if participant i is in the negative treatment (and zero otherwise). The interaction term thus displays whether participant i encountered a negative uninformative signal in round t. Controls include age, gender, statistical skills, risk aversion, and participants' financial literacy. Reported are coefficients and t-statistics (in parentheses). All standard errors are clustered at the individual level. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

4.4 Conclusion

This article experimentally studies how individuals update their beliefs after observing non-diagnostic information signals with varying degrees of desirability. Whereas Bayes' Rule predicts that such *uninformative* signals do not influence inference judgements, we find that individuals systematically incorporate them in their belief formation process. Importantly, the direction in which individuals update their beliefs strongly depends on the valence of the observed signal. Individuals tend to form more optimistic beliefs about the objective state of the world after observing desirable uninformative signals, whereas they form more pessimistic beliefs after observing undesirable uninformative signals. As mechanism we identify that individuals process the valence of new signals in a reference-dependent manner, dictated by their prior beliefs. Whenever they observe non-diagnostic outcomes which are close to their prior expectations, they only weakly update their beliefs, whereas when they observe non-diagnostic outcomes which are at odds with their prior expectations, they strongly overreact.

Taken together, our findings suggest that individuals appear to struggle discerning belief-relevant information from their preferences. Such deviations from Bayesian behavior are particularly severe in situations in which the valence of non-diagnostic signals is at odds with the valence of objective pieces of information. In such an environment, uninformative signals can not only lead to systematically biased beliefs whenever desired or undesired outcomes are non-indicative of the true state of the world. Instead, they may also reinforce wrongly entertained beliefs based on individuals' preferences. Even though decision making frequently involves the accumulation of new pieces of information until uncertainty is reduced to a tolerable level, such a bias may instead lead to a decline in predictive performance.

Chapter 5

When Saving is Not Enough – Wealth Decumulation in Retirement *

5.1 Introduction

When conducting a simple Google search on the term 'retirement planning' one finds an overwhelming share of articles which contain recommendations on saving decisions and on how to allocate savings to increase financial well-being in retirement. Given this prevailing focus on savings and investment decisions, one could forgive a typical retiree for believing that retirement planning is synonymous with wealth accumulation. Yet, while wealth accumulation is certainly a mandatory condition for successful retirement preparation, it is not a sufficient condition to achieve a targeted steady stream of income during retirement. However, determining how to draw down his wealth is not an easy task for a person contemplating retirement, as one cannot rely on experience.

Rational choice theory predicts that, in the absence of a bequest motive, households will fully convert their savings into a lifetime annuity (Yaari, 1965). Yet, despite the attractiveness of annuities as a way to protect against

* Authors: Pascal Kieren and Martin Weber. All authors are at the University of Mannheim (L9, 1-2, 68161 Mannheim). For valuable comments, we thank participants at the Experimental Finance Conference 2019, and seminar participants at the University of Mannheim. All remaining errors are our own.

the risk of outliving one's retirement wealth, relatively few of those facing retirement actually annuitize a significant proportion of their wealth, a discrepancy coined the annuity puzzle.[1]

In this paper, we seek to investigate the wealth decumulation decision from the perspective of a retiree who is averse to the prospect of fully annuitizing his savings. Such an individual faces the following decision problem. Out of one's non-annuitized wealth, one must decide how much to allocate to a savings account (e.g. as protection against early unexpected costs) and how much (if anything) to decumulate over the course of retirement. As an alternative to annuities, we investigate consumers' preferences for phased withdrawal accounts. In the light of recent findings, which question the benefit of full annuitization in the presence of stochastic health shocks (e.g. Reichling and Smetters, 2015, or Peijnenburg et al., 2017), such an analysis might not only provide valuable insights for the design of complementary products but also important policy implications.[2]

To study these issues, we field a large online survey in cooperation with a national German newspaper, *Frankfurter Allgemeine Zeitung* (FAZ), in which we elicit preferences for simple drawdown strategies. The strategies differ across two main dimensions, risky vs. risk-free asset allocation and constant vs. dynamic withdrawal rates. We examine 1) what hypothetical products individuals find most appealing, 2) what factors people say are most important in their wealth decumulation decision, 3) whether standard utility functions to study consumption decisions adequately capture observed preferences for phased-withdrawals, 4) how the demand for phased withdrawal products compares to the demand for annuities, and 5) how retirement preparation affects individuals willingness to decumulate wealth.

[1] Over the past decades, economists have focused on explaining the annuity puzzle under consideration of both behavioral and rational factors. For a review, see Brown (2007) or Benartzi et al. (2011).

[2] We do not attempt to claim that phased withdrawals are superior to annuities, as they cannot eliminate longevity risk. Instead, we seek to obtain a more holistic understanding of the wealth decumulation decision by investigating preferences for phased withdrawals. In our view, phased withdrawals should be seen as a complement, rather than a substitute, for those individuals who want to retain control over their wealth and are averse to full-annuitization.

Using a survey to investigate our research question has both advantages and disadvantages. On the positive side, we can use hypothetic choice questions to measure preferences for specific (non-existing) products, which are unobservable in field data. Another advantage is the sample from which we can draw the survey data. While readers of the FAZ are not representative of the general population (they are on average more educated and have higher income), they are highly representative of those most affected by the decision of how much wealth to decumulate. On the negative side, the choices individuals make do not translate to their actual life outcomes. As a consequence, the results may not correspond to the choices people would make in a real-life situation. However, even though the resulting choice behavior might be noisy, it would be surprising if it leads to systematic patterns that are absent in actual behavior.

From our survey, five main findings emerge. First, we find that most participants prefer phased withdrawal accounts with equity-based asset allocation and dynamic withdrawal rates, which smooth the withdrawal amount across market phases. Overall, roughly 81 % of our respondents select a drawdown strategy with an equity-based asset allocation, while only 19 % prefer a strict risk-free asset allocation. Additionally, out of those participants who prefer an equity-based decumulation strategy, only 35 % prefer constant withdrawal rates, which cannot offer protection against depleting the capital stock early, as withdrawal rates do not adjust for periods of low returns. Conversely, 65 % prefer dynamic withdrawal rates, which adjust the withdrawal amount based on realized returns in order to avoid depleting the capital stock too fast. This choice pattern suggests that while retirees are highly averse to some risks (namely having to live on a permanently lower income) they are less risk-averse when it comes to equity investments with long planning horizons.

Second, the self-reported importance of various withdrawal characteristics is closely in line with participants' actual choice. The two considerations that respondents report being most important for their withdrawal account choice are sufficient protection against the risk of depleting the capital stock early, while also achieving relatively high returns on the invested assets. Taken together with the actual withdrawal account choice, our results highlight that customers desire flexible payout structures, which dynamically adjust in states of low returns. Most currently offered decumulation products (e.g. lifelong annuities) primarily offer constant income streams, even though there is no economic reason to do so assuming that major expenses (e.g. vacations or health costs) do not occur on a regular basis. While such income streams may allow customers to plan ahead, they could also have detrimental effects on the demand and – relatedly – the generated returns (guaranteed income streams come at the expense of less risky investment options). We provide new evidence that the latter effect is of importance.

Third, a time-separable power utility function with bequest motives as frequently employed in life cycle models predicts that a decumulation strategy with equity-based asset allocation and dynamic withdrawal rates is the utility-maximizing choice for a large number of preference parameter combinations. As the predictions of the utility function are closely in line with participants' actual choice, our results provide evidence for the suitability of such utility functions to study not only consumption and savings decisions but also wealth decumulation topics.

Fourth, we find that only 12 % of all respondents would choose an annuity product to decumulate their wealth while 88 % would rather select a phased withdrawal solution. This result – while surprising – is not only in line with subjects' preference to achieve higher returns on their accumulated savings while being flexible in the way they decumulate wealth but also with general findings on the annuitization puzzle. According to a survey conducted by Beshears et al. (2014), many subjects report that "flexibility in the timing of my spending" is an important factor in their annuitization decision. Yet, many consumers still seem to neglect that while phased withdrawals provide more flexibility in the timing of the spending, they cannot offer protection against longevity risk. In the light of current regulatory efforts, which discuss the benefits and drawbacks of forced annuitization of defined contribution payments, our results suggest that policymakers should consider offering combined solutions. Distributing wealth among annuities and phased withdrawals could help retirees who are averse to full annuitization to insure against longevity risk, while also preserving liquid wealth and making use of the equity premium.

Finally, we find that participants are willing to decumulate on average 65 % of their liquid savings over the course of their retirement. In contrast to this rather high self-reported willingness to decumulate wealth, actual spending in retirement is still quite low (e.g. Olafsson and Pagel, 2018). Yet, given the low demand for annuities in our sample, we conjecture that part of this discrepancy is driven by the lack of alternative wealth decumulation products. Additionally, we find two opposing effects of how retirement preparation affects individuals' willingness to decumulate wealth. First, individuals who successfully prepare for retirement by consulting financial planners or by sticking to saving plans do not show an increased propensity to draw down a greater fraction of their savings, even though they accumulated more wealth on average. Thus, while wealth accumulation is certainly an important ingredient for retirement preparation, it does not predict subsequent decumulation. However, we do find that individuals' attitude towards retirement affects their willingness to decumulate wealth. To capture the fact that individuals cannot rely on their experience in deciding how much wealth to decumulate, we investigate the impact of optimism, as research has shown that these are the decisions most likely to be influenced by emotional dispositions (Puri and Robinson, 2007). We find that while moderate optimism is positively related to the wealth participants are willing to decumulate, extreme optimism leads to a strong negative effect. Consistent with the model of Brunnermeier and Parker (2005), it appears that moderately optimistic individuals are more inclined to take small risks to increase their wellbeing, while extreme optimists reduce their spending possibly to protect against longevity, thereby overestimating their income from non-annuitized wealth.

The remainder of the paper is organized as follows. In Section 5.2 we describe the design of our online survey and outline how we elicit preferences about the properties of retirement products. We then present our key empirical results on respondents' product choice followed by a utility-analysis of income drawdown offerings and an analysis of the wealth decumulation decision in Section 5.3. Finally, in Section 5.4, we conclude with a discussion of possible policy implications and future research questions.

5.2 Survey Design and Summary Statistics

5.2.1 Survey Design

To investigate the wealth decumulation decision and to derive predictions about the design of phased withdrawal strategies, we conduct an online survey in cooperation with the newspaper *Frankfurter Allgemeine Zeitung* (FAZ). The survey was promoted to cover retirement decisions and planning and was accessible through a link that was posted on their online portal on August 16, 2018. The survey and related material can be found in Appendix D.

Overall, 3598 participants with an age ranging from 18 to 93 completed the survey. Participants answered hypothetical questions about different retirement products, their willingness to decumulate wealth in retirement, and rate how the payout structure of a hypothetical income drawdown offering should look like. Moreover, they answered questions about demographics and household characteristics, risk preferences, financial literacy, and numeracy.

Preferences regarding the payout structure of phased withdrawal products were elicited in two different ways in a within-subject design, which will be described subsequently. In both elicitation strategies, *product-based* and *self-reported*, we ask respondents to rate the importance of four characteristics related to the shape of the stream of payouts. The first characteristic resembles participants' attitude about the size of the payouts. The second characteristic is what we refer to as the variance in the payout stream. Many currently offered retirement products (e.g. most annuities) feature constant payout streams, which allow consumers to plan ahead with a given budget. Yet, from an economic perspective, there is no reason to primarily offer constant payouts, as fluctuating payouts can dynamically adjust to economic conditions. That is, in states of high returns, consumers can either increase consumption or increase savings (e.g. by capping the maximal withdrawal amount) to shift more consumption to states with adverse market conditions. The third characteristic we assess is the uncertainty in the payout stream. As phased withdrawals can invest in equities, they are necessarily subject to capital market risks, which − depending on the payout policy − can lead to default risk. In our context, we use the term default risk to refer to the probability of exhausting the capital stock before the end of the planning horizon. Finally, we also assess to what extent participants view wealth that is not consumed before they die as an inefficient way of allocating resources or as an opportunity to benefit future generations. In other words, the last characteristic resembles bequest motives.

Product-based elicitation

In the product-based elicitation, we seek to measure the importance of the payout characteristics by presenting participants with three different options to draw down their retirement savings. As the aforementioned characteristics are not mutually exclusive, we construct phased withdrawal strategies that differ across two dimensions, constant vs. dynamic withdrawal amounts and risky vs. risk-free asset allocation. We label the resulting drawdown strategies as (1) *risk-free − constant consumption*, (2) *risky − constant consumption*, and (3) *risky − dynamic consumption*.[3] To avoid too much complexity in the decumulation strategies and to ensure that the characteristics are still clearly differentiable for participants, we use simple heuristics to construct the strategies (a precise definition is provided in the Appendix):

[3] Note that while constructing strategies which differ across two dimensions (2x2) would result in four different strategies, we only use three of them as the combination fluctuating withdrawals and risk-free asset allocation would not make sense in a hypothetical choice scenario.

1. *Risk-free - Constant Consumption*. For the first strategy, calculating the constant amount that can be withdrawn over a fixed number of years assuming deterministic returns is straightforward and only depends on participants' planning horizon and their accumulated wealth.[4] For the risk-free asset allocation, we use the historical inflation-adjusted average of 1-year German government bonds, which amounts to roughly 1.22 % for the past 30 years (German Federal Bank, 2018).

2. *Risky - Constant Consumption*. The second strategy combines a constant yearly withdrawal amount with a risky investment strategy. We implement these features by withdrawing each year a fixed percentage of the original retirement wealth (adjusted for inflation), which is invested in a well-diversified portfolio described subsequently.[5] Note that by combining constant withdrawal rates with stochastic returns, the strategy can neither guarantee that the capital stock is sustained until the end of the planning horizon (i.e. it can default), nor that the initial wealth will be fully exhausted in the decumulation process (i.e. it could also end up with a large terminal wealth). To ensure comparability across different planning horizons, we selected the fixed percentage such that the default probability remains constant at 10 % (i.e. a higher withdrawal amount for shorter horizons). The resulting withdrawal rates for different horizons are displayed in the Appendix.

3. *Risky - Dynamic Consumption*. The third strategy features dynamic withdrawal rates paired with a risky investment strategy. It can be implemented in a similar fashion as the first strategy with one exception. Once return expectations are stochastic, the realized return will most likely not equal the expected return. As a consequence, the actual withdrawal amount for each year has to be recalculated each period, taking

[4] The present value of constant income stream that pays a yearly amount y conditional on an expected return r, a planning horizon of T years, and an initial portfolio value V is calculated using the following formula: $y = V \cdot \frac{(1+r)^{T-1} \cdot r}{(1+r)^{T} - 1}$.

[5] Besides its simplicity, a similar decumulation strategy was originally developed by Bengen (1994) and Cooley et al. (1998).

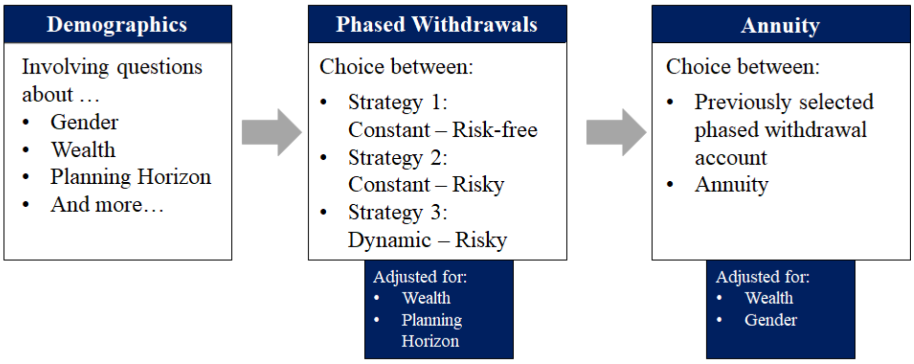
the realized return (and as such the actual portfolio value) of the previous period into account. The resulting relatively high withdrawal rates (the portfolio value will be fully exhausted at the end of the final period) come at the expense of uncertainty about the actual withdrawal amount.

To allow participants to compare risk and benefit characteristics of each strategy conditional on survival, we conduct Monte Carlo simulations. To do so, we assume that retirees decumulate their wealth over a period of a minimum of 20 years and a maximum of 30 years (i.e. to the age of 95 assuming a retirement age of 65). The risky investment strategies assume that retirees hold their non-annuitized assets in a 60 % stock, 40 % bond portfolio, as typically offered by balanced funds (e.g. Gomes et al., 2008). The equity component in our study is represented by the MSCI World Index, while the bond component is represented by monthly U.S. treasury bills. The plan assets are rebalanced annually within a buy-and-hold approach and returns are adjusted for inflation. Portfolios are constructed for the period between February 1970 to February 2018. Return data for the MSCI component was obtained from Datastream, while the risk-free rate was downloaded from the union of the CRSP/Compustat database.

To simulate outcomes, we employ a bootstrapping algorithm, which randomly draws (with replacement) 360 return observations (twelve months over 30 years) from our portfolio data to generate one scenario with 30 years of data. This process is then repeated 10,000-times to obtain a sufficient number of scenarios.

Figure 5.1 depicts the order in which questions on the phased withdrawal choice are presented. The exact wording of the strategies is reported in Appendix D.2. Before participants observe the withdrawal strategies, they answer general demographic questions including a forecast of their wealth level at retirement (assessed by five categories or an exact number) and the time over which they would want to decumulate their assets (choice between 20, 25, or 30 years). Afterward, participants can choose one of the

Figure 5.1: Survey Overview and Timing



Note: This figure illustrates is the order in which the questions regarding demographics, the withdrawal plan choice, and the annuity choice were presented. Both decumulation strategies and annuities were adjusted for previously reported demographics, including gender, wealth, and planning horizon.

three decumulation strategies, each tailored to participants' personal wealth and their desired planning horizon.

Each withdrawal strategy is described by four key financial variables (average consumption, default probability, consumption fluctuation, and consumption in the worst 5 % of the cases) and a brief overview of advantages and disadvantages.[6] In a consecutive question, participants decide whether they prefer a decumulation strategy or a life-long annuity. The annuity is presented in a similar fashion compared to the withdrawal strategies. The annuity values are calculated assuming a real interest rate of 1.22 % (as for the risk-free decumulation strategy) and using the latest life tables for Germany. Moreover, due to adverse selection in the annuity market, we made a downward adjustment to the expected present discounted value of the fair annuity following Mitchell et al. (1999). This downward adjustment amounts to 15 % and 10 % of the fair value for male and female participants, respectively. To avoid potentially confounding framing effects as discussed by Brown et al. (2008), the variables for the phased withdrawals and the annuity were both framed in terms of consumption and described periods in terms of participants' age in retirement. Finally, after subjects decided which

[6] Advantages and disadvantages were chosen to highlight participants the difference between both constant vs. fluctuating consumption streams as well as a risky vs. risk-free asset allocation.

product best suits their preferences, they are asked to answer a question about how much of their overall wealth they would be willing to decumulate over the course of their retirement.

Self-reported elicitation

In the self-reported elicitation strategy, we directly ask subjects to assess the importance of the four payout characteristics on a seven-point Likert scale. The exact wording is reported in the Appendix. Participants have to answer these questions after they chose their preferred decumulation strategy. While this was done to ensure that subjects have a profound understanding of what the statements mean, the increased knowledge comes at the expense of individuals potentially ex-post rationalizing their initial choice. To avoid that the order in which questions are presented affects our results, we focus the subsequent analyses on participants' preferred strategy and use the self-reported measures as consistency check.

In addition, we also ask participants to provide an estimate of their life expectancy and health status (adopted from the Survey of Consumer Finances and Mirowsky, 1999), to indicate which tools they use to prepare for retirement and whether they have tried to figure out how much their household would need to save for retirement. Self-reported life expectancy and health status have been found to be amongst the most important factors influencing the annuitization decision, while the latter factors are important determinants for successful retirement planning (Lusardi and Mitchell, 2011).

Controls

We elicit a financial literacy score based on participants' answers to six questions, of which three are pure knowledge questions and another three are related to financial numeracy. We select one of the basic questions from Lusardi and Mitchell (2007), two advanced questions from Van Rooij et al. (2011), one question from Schreiber and Weber (2016), one question from Lusardi and Tufano (2009), and one question from Ensthaler et al. (2018).

The exact wording of the questions can be found in Appendix D.2. Following the suggestion of Behrman et al. (2012) and Fernandes et al. (2014), we also collect information on parents' and siblings' highest level of education and assess a scale of need for cognition (Epstein et al., 1996) as instruments for financial literacy not caused by financial behaviors.

To control for risk and loss aversion, we ask participants to rate their risk and loss attitude on a seven-point Likert scale. Earlier studies on risk-taking find that self-reported risk attitude is a good predictor of actual risk-taking (e.g. Nosić and Weber, 2010; Van Rooij et al., 2011). Moreover, we also assess participants' trust in financial markets on a seven-point Likert scale as a proxy of participants' general willingness to invest in financial products.

5.2.2 Summary Statistics

Table 5.1 presents summary statistics on participant demographics, life expectancy, financial literacy, and risk aversion.

The average age in our sample is 52.1 years (median 54). While this is higher compared to similar surveys of our kind (e.g. Merkle et al., 2017 or Müller and Weber, 2014), it is well suited to study hypothetical retirement choices. Men are overrepresented in our study (85 %), which reflects the fact that the majority of *FAZ* readers are male. Participants report a relatively high after-tax income of about 5440 €, compared to the German average of about 3300 €(German Federal Statistical Office, 2018). Additionally, participants report having Social Security benefits of roughly 3556 €(retired participants only), and an average net worth of roughly 455,357 €[7]. They are well educated with about 78 % having obtained a university degree. Around 21 % are already retired, 63 % are married, and participants have on average 1.15 children. We also asked respondents about their average life expectancy. While female participants expect to live on average 86.5 years, male participants expect to live on average 85.7 years. While these estimates are slightly higher

[7] In the survey, participants could either provide interval responses for net wealth or an exact value. To calculate net wealth for participants with interval responses, we map each interval to its midpoint except in the case of the final interval without an upper bound, which we map to a value equal to the lower bound.

Table 5.1: Summary Statistics

Variable	Mean	Median	Std. dev.
Demographics			
Age	52.12	54	13.6
Female	15%		0.36
Income in €(after tax)	5,441	6,000	2011
Social Security in €(only retired participants)	3,556	4,500	1580
Liquid wealth in €	455,357	375,000	924,041
Retired	21%		0.41
Number of children	1.15	1	1.21
Married	63%		0.48
Highest education attained			
No high school diploma	11%		
High school diploma	11%		
College degree	59%		
Graduate degree	19%		
Controls			
Health status (1-5)	4.17	4	0.7
Life expectancy (in years)	85.82	85	6.73
Saving plan	62%		0.49
Knowledge score (0-3)	1.77	2	0.73
Numeracy score (0-3)	1.67	2	0.66
Need for cognition	26.22	27	4.77
Risk aversion (1-7)	3.77	4	1.54
Loss aversion (1-7)	4.16	4	1.63

Note: This table presents summary statistics of our survey. Included are all 3598 participants. Statistics are split across two categories: demographics, and controls. Reported are mean, median (whenever applicable), and standard deviation. Female, retired, married, and saving plan are dummy variables.

than the average life expectancy implied by recent life tables for the respective cohort, this does not necessarily present evidence for an optimism bias, since our participants are on average wealthier and more educated which is positively correlated to life expectancy (Meara et al., 2008).

Participants correctly answer on average 1.77 of the knowledge questions and roughly 1.67 of the numeracy questions out of 3. Given the level of complexity of the questions, participants do quite good. Need for cognition score is on average 26.22 out of 35. Asking participants for their risk and loss aversion leads to an average of 3.77 and 4.16, respectively. Overall, one should emphasize that our sample is most likely not representative of the general

German population. However, it is highly representative to study wealth decumulation preferences for individuals who have the choice to decumulate a significant proportion of their accumulated savings.

5.3 Results

We present four set of findings: 1) observed choices and demographic correlates of phased withdrawal accounts, 2) a utility analysis for the demand of phased withdrawal accounts under a standard time-separable power utility function, 3) factors differentiating the demand for annuities versus phased withdrawals, and 4) factors related to the decision of how much wealth one is willing to decumulate. In all subsequent analyses, we present results for our full sample. However, the conclusions we draw do not critically depend on this, as results are similar when restricting our analyses to individuals close to retirement age.

5.3.1 Preferences on the Structure of Phased Withdrawal Accounts

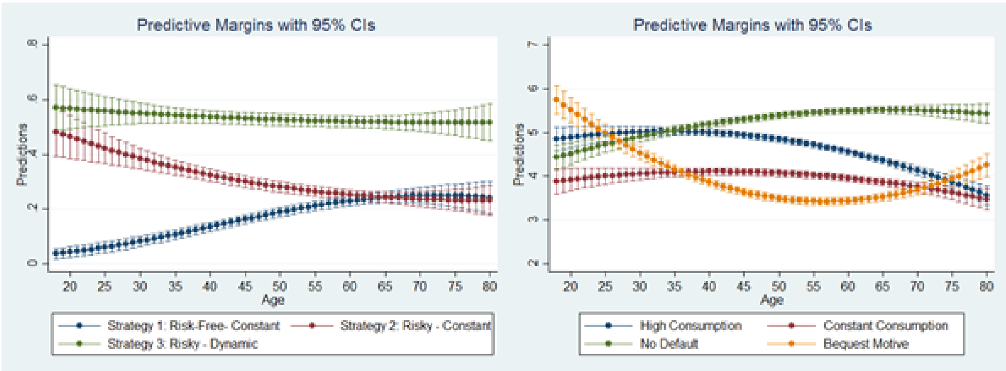
Figure 5.2 displays both the withdrawal strategies that participants prefer and the characteristics they deem important on an aggregate level (Panel A) and across age (Panel B).

From Figure 5.2, it becomes evident that independent of age, the majority of participants prefer a withdrawal strategy with risky asset allocation and dynamic withdrawal rates. Given that many currently offered decumulation products (such as most lifelong annuities) involve constant income streams, this finding is quite surprising. However, it indicates that consumers are by no means averse to fluctuating income streams, assuming that they receive sufficient compensation in return. Regarding the strategies with constant withdrawals, we observe two opposing effects across different age groups.

Figure 5.2: Phased Withdrawal Account Choice

Panel A

Panel B



Note: This figure displays participants' preferred withdrawal strategies and their rating of various decumulation attributes. Panel A displays the strategies and attributes for the whole sample, while Panel B displays the predictive margins with 95% confidence intervals from logit regressions on the plan choice or the respective characteristic for age.

While the strategy with risk-free asset allocation is mostly preferred by participants who are close to retirement, it is hardly chosen by younger participants. Conversely, the strategy with constant withdrawals and risky asset allocation is popular among younger participants while it loses in popularity among those close to retirement.

When looking at the assessed withdrawal characteristics, this pattern becomes even clearer. In particular, avoiding the risk of depleting the capital stock early gradually becomes the most desired characteristic as individuals approach retirement followed by the desire for a high average consumption. However, quite the reverse appears to hold for bequest motives. While younger participants list bequest motives among the most important characteristics, they are hardly relevant for those close to retirement and become more important again as participants approach later stages of their retirement. The latter relation is consistent with participants' declining demand for the strategy with risky asset allocation and constant withdrawals, which is the only strategy that might result in unintended bequests after the final period. Finally, a constant consumption stream in retirement is deemed rather unimportant by most participants, which is not only consistent with the dynamic withdrawal strategies participants choose but also further evidence for the demand for flexible decumulation options.

Besides the descriptive analysis, we also examine how cognitive abilities and demographic characteristics relate to the choice using multinomial logistic regressions. The dependent variable is *strategy* that takes on three values, which capture participants' preferred withdrawal strategy. The independent variables are knowledge, numeracy, the log of wealth and various demographics. Results are reported in Table 5.2.

We can draw several conclusions from Table 5.2. First, it appears that the more financially savvy individuals are, the more likely they are to select an equity-based withdrawal strategy (i.e. Strategy 2 or 3). Similarly, the more trust individuals have in financial markets and the more educated they are (i.e. having at least a university degree), the more likely they are to select either the second or the third strategy. Yet, we also observe differences within the strategies, which invest in equities. That is, more statistically numerate individuals show a higher propensity to choose withdrawal strategies with dynamic withdrawal rates, both relative to the risk-free alternative and relative to the risky strategy with constant withdrawal rates. Taken together, these results imply that while financial education is positively related to a return-oriented investment behavior in retirement, it cannot explain whether investors prefer dynamic or constant payoff streams. Those individuals, however, who show – ceteris paribus – also a deeper understanding of financial mathematics and compound interest, are significantly more likely to choose dynamic payoff streams. Consistent with the findings from Bateman et al. (2018), basic financial literacy helps retirees to manage decumulation, but it is not sufficient for effective ruin risk management.

Table 5.2: Determinants of the Phased Withdrawal Account Choice

	(1) Baseline: Strategy 1		(2) Baseline: Strategy 2	
Strategy 1				
Knowledge			−0.211***	−0.0776
Numeracy			−0.0219	−0.083
Trust			−0.381***	−0.0379
Log(Wealth)			−0.360***	−0.071
18<Age<35			−0.992***	−0.21
50<Age<65			0.328**	−0.14
Age>65			0.443**	−0.177
Female			0.277*	−0.151
Married			0.115	−0.121
University			−0.338***	−0.123
Kids			−0.235***	−0.0473
Strategy 2				
Knowledge	0.211***	−0.0776		
Numeracy	0.0219	−0.083		
Trust	0.381***	−0.0379		
Log(Wealth)	0.360***	−0.071		
18<Age<35	0.992***	−0.21		
50<Age<65	−0.328**	−0.14		
Age>65	−0.443**	−0.177		
Female	−0.277*	−0.151		
Married	−0.115	−0.121		
University	0.338***	−0.123		
Kids	0.235***	−0.0473		
Strategy 3				
Knowledge	0.266***	−0.0698	0.0542	−0.0572
Numeracy	0.256***	−0.0751	0.234***	−0.0639
Trust	0.382***	−0.0347	0.000879	−0.0269
Log(Wealth)	0.273***	−0.06	−0.0866	−0.054
18<Age<35	0.624***	−0.2	−0.368***	−0.126
50<Age<65	−0.271**	−0.127	0.0561	−0.102
Age>65	−0.298*	−0.158	0.145	−0.136
Female	−0.0469	−0.131	0.230*	−0.122
Married	−0.108	−0.11	0.00701	−0.0901
University	0.352***	−0.109	0.014	−0.0998
Kids	0.106**	−0.0447	−0.129 * **	−0.0356
N	3573		3573	
R^2	0.062		0.062	

Note: This table reports results of two multinomial logit regressions with varying baseline values. Dependent variable is *Strategy*, a categorial variable, which denotes participants' preferred withdrawal strategy (1 – 3). Age is captured by clustering participants in four age groups, with the medium category (between 36 and 49) as baseline. Results for Strategy 3 as baseline are suppressed as the table is symmetric. Reported are coefficients and robust standard errors. ***, **, and * indicate significance at the 1%, 5% and 10%-level, respectively.

Regarding the impact of demographics and household characteristics, we find a positive relationship between participants' accumulated wealth and their willingness to invest in equity products, with no significant difference between dynamic and constant withdrawal rates. Earlier studies find that wealthier individuals usually have a higher exposure to financial markets, are on average more sophisticated, and hold better diversified stock portfolios (e.g. Goetzmann and Kumar, 2008). Our results suggest that their familiarity with equity investments also makes them more likely to favor a return-oriented asset allocation for wealth decumulation post retirement.

5.3.2 Utility Analysis of Phased Withdrawal Accounts

Normative Predictions

To derive normative predictions, we start by assuming that an exemplary agent enters retirement at the age 65[8] (t=1) with an accumulated initial wealth $W_0 > 0$. To decumulate his wealth, the retiree has access to the three different withdrawal strategies described previously, which ultimately define the amount C_t he is able to consume at the beginning of each period. Moreover, we assume that the retiree survives every year with a positive probability $p_{t,g} > 0$ ($g \in \{male, female\}$) until the last year of the planning horizon is reached. If the retiree either dies before the final period or does not consume all of his wealth before the plan ends, we assume that the remaining wealth will be transferred to an heir, yielding a (dis-)utility in the form of a bequest B. Note that this analysis only captures a fixed period of years and neglects the period after the planning horizon. While simplifying, the assumption is not unjustified as non-insurance products cannot offer longevity protection. As such, retirees face both capital markets risk and longevity risk.

We assume that retirees' preferences are described by a time-separable power utility function proposed by Cocco et al. (2005):

[8] The average retirement age in Germany is around 65. However, our results do not depend on this assumption.

$$U_0\,(C,B) = \sum_{t=1}^{T} \delta^{t-1} \left(\prod_{j=0}^{t-2} p_j \right) \left\{ p_{t-1} \frac{C_t^{1-\gamma}}{1-\gamma} + b(1-p_{t-1}) \frac{B_t^{1-\gamma}}{1-\gamma} \right\} \tag{5.1}$$

where $\delta < 1$ is the discount factor and $\gamma > 0$ is the coefficient of relative risk aversion. The parameter b controls the intensity of the bequest motive. While positive values of b translate to retirees' desire to benefit future generations, negative values of b correspond to a view that bequests are an inefficient resource allocation. For simplicity, we assume that the utility function applied to the bequest is identical to the utility function of the retiree's own consumption.

We begin our analysis by restricting our attention to preference parameter tuples (γ, b) for which $\gamma \in [1,\ 10]$, and $b \in [-0.5, 2]$. We focus on values of γ that are below 10, as this is the upper bound for risk aversion considered reasonable by Mehra and Prescott (1985), and restrict the intensity of the bequest motive to not exceed the benefit of own consumption by a factor of two. We then discretize each of the two intervals (γ, b) into a set of 40 and 25 equally spaced points and study parameter tuples where each parameter takes a value that corresponds to one of the discrete points. As we repeat the analysis for four different planning horizons $T \in \{20, 25, 30, 35\}$, we study $40 \cdot 25 \cdot 4 = 4,000$ different scenarios for each simulated consumption path.

Figure 5.3 presents results. First, we see that the withdrawal strategy with risky asset allocation and dynamic consumption is the utility-maximizing choice for the majority of more realistic preference parameter combinations. Considering all combinations, this strategy is optimal in 2,533 out of the 4,000 parameter tuples. In particular, for medium positive values of the intensity of the bequest motive ($0 \leq b \leq 0.5$), we find the third strategy is the utility-maximizing choice for all levels of risk aversion. Yet, as risk aversion increases, the floor between the third and the other two strategies is decreasing. This finding is not surprising. As relative risk aversion increases, the benefit of an additional unit of consumption is strictly decreasing. As such, the high average consumption of the third strategy becomes relatively

Figure 5.3: Utility Simulations

Note: This figure displays the utility-maximizing withdrawal plan for the given preference parameter tuple, as indicated by the colored dots. The y-axis captures in intensity of the bequest motive for parameter values $b \in [-0.5, 2]$. The x-axis depicts the parameter of relative risk aversion for values $\gamma \in [1, 10]$. Each figure displays one planning horizon for $T = \{20, 25, 30, 35\}$

less important.

Observation 1: A withdrawal strategy with risky asset allocation and fluctuating consumption is the utility-maximizing choice for most realistic parameter tuples. Its utility is decreasing in the parameter of relative risk aversion and decreasing the further the intensity of the bequest motive is away from zero.

Second, we find that a withdrawal strategy with risk-free asset allocation and constant withdrawals is never optimal as long as $b \geq 0$. As risk aversion increases and the bequest intensity decreases, this strategy becomes gradually the optimal choice until it is optimal for any $b < 0$. In other words, as long as a retired investor with preferences as described here is at least indifferent to the prospect of leaving a bequest, he would never choose to decumulate his wealth using a completely risk-free asset allocation. For those investors, however, who are both highly risk-averse (and as such do not value high consumption) and who view bequests as an "inefficient" way of allocating their retirement resources, such an allocation would be the utility-maximizing choice.

Observation 2: A withdrawal strategy with risk-free asset allocation is never the optimal choice unless investors are both highly risk-averse and averse to the prospect of leaving bequests.

The remaining withdrawal strategy features a risky asset allocation paired with constant withdrawal amounts. In contrast to the risk-free strategy, we observe that as the intensity of the bequest motive increases, this strategy becomes the utility-maximizing choice for both low and high parameters of relative risk aversion. For medium values of relative risk aversion however, the strategy with dynamic withdrawal amounts remains the utility-maximizing choice. The intuition for this finding is as follows. Both strategies follow the same asset allocation, and as such, generate the same returns over the respective time horizon. As the consumption of the constant withdrawal strategy is on average lower compared to the consumption of the dynamic strategy, more overall wealth is generated. As bequests rise in importance, so does overall wealth, which explains the positive relation with the intensity of the bequest motive. The relation with the parameter of relative risk aversion is a little more subtle. For low levels of risk aversion, more consumption (or wealth) is always better due to the low concavity of the utility function. As average wealth is much higher than average consumption, the lower consumption of Strategy 2 is outweighed by the high average bequeathable wealth and as such, Strategy 2 is the optimal choice. For intermediate values of risk aversion however, the more balanced relation between consumption and wealth of Strategy 3 eventually becomes superior. Yet, for high levels of risk aversion, this relation shifts once again. Now, the utility function has a fairly high concavity and as such, even great differences in consumption and wealth translate to only marginal increases in utility. At this point, the difference in consumption between both strategies is no longer enough to offset the difference in wealth at later stages of the planning horizon. In particular, while the wealth profile of Strategy 3 is decreasing, it is increasing for Strategy 2. As a consequence, Strategy 2 becomes optimal once again, given a relatively high intensity of the bequest motive.

Observation 3: A withdrawal strategy with risky asset allocation and constant withdrawals is the utility-maximizing choice for investors with strong bequest motives who show either a relatively low level of risk aversion or a relatively high level of risk aversion.

Finally, we can also compare the utility of the withdrawal strategies across different time horizons. Most notably, we find that while Strategy(3 becomes the utility-maximizing choice for an even greater range of preference parameters, Strategy 2 vanishes nearly entirely for very long planning horizons (T = 35). Only for very low values of relative risk aversion and a high intensity of bequests, Strategy 2 is still utility-maximizing. This is partially related to how the second strategy was constructed for different time horizons. To make the strategy comparable, we adjusted the withdrawal amounts under the constraint that the probability of default remains constant across all horizons. As such, average yearly consumption declines for longer planning horizons while average wealth levels increase. For higher parameters of risk aversion, the increase in wealth is, however, not enough to outweigh the drop in consumption compared to other strategies.

Observation 4: Withdrawal strategies whose withdrawal rates adjust to market conditions are increasingly optimal the longer individuals' planning horizons are compared to strategies who do not adjust withdrawals.

Predicted and Actual Choice

To test how well a time-separable power-utility function with bequest motives describes participants' actual choice behavior, we construct three dummy variables that indicate which of the three withdrawal strategies a participant prefers, which will be the dependent variables for our subsequent analyses. To match participants' actual choices with the previously generated predictions, we discretize their self-reported risk-aversion $(1-7)$ and their self-reported bequest motive $(1-7)$ into seven equally spaced points to fit the described intervals used for the utility simulations, i.e. $\gamma \in [1, 10]$ and $b \in [-0.5, 2]$. Since self-reported risk-aversion and bequest intentions are only noisy measures of the true parameter values, we also consider alternative limits for both intervals. To assign each participant a "best-choice"-prediction, we match the simulated utility-maximizing choices with our survey data based on the two described intervals and based on participants' chosen planning horizon. As a result, each participant is matched with a unique utility-maximizing choice that corresponds to her parameter triple (γ, b, T), which will serve as the main independent variable. Table 5.3 reports the marginal effects of five sets of probit regressions with participants' chosen decumulation strategy as dependent variable. Each specification represents a different interval over which parameters were linearized while the last specification represents a placebo-test, where participants were matched with random recommendations.

Across all specifications, we reliably find that a time-separable power utility function with bequest motives successfully predicts preferences for the first and the third withdrawal strategy. For the risk-free strategy (Strategy 1), we find that when the utility functions predict the first strategy to be the optimal choice for a given participant, participants are on average between 11

Table 5.3: Utility Predictions

Dependent Variable	Strategy 1: Risk-free – Constant	Strategy 2: Risky – Constant	Strategy 3: Risky – Dynamic
	Specific. 1: $b \in [-0.5, 2]$& $\gamma \in [1, 10]$		
CRRA Strategy 1	0.1181***		
	(6.83)		
CRRA Strategy 2		0.0175	
		(1.02)	
CRRA Strategy 3			0.1081***
			(6.36)
	Specific. 2: $b \in [-0.5, 1]$& $\gamma \in [1, 7]$		
CRRA Strategy 1	0.1958***		
	(9.75)		
CRRA Strategy 2		0.0968**	
		(2.22)	
CRRA Strategy 3			0.1478***
			(5.32)
	Specific. 3: $b \in [-0.5, 2]$& $\gamma \in [1, 7]$		
CRRA Strategy 1	0.1933***		
	(8.18)		
CRRA Strategy 2		0.0308	
		(1.60)	
CRRA Strategy 3			0.1276***
			(6.61)
	Specific. 4: $b \in [-0.5, 1]$& $\gamma \in [1, 10]$		
CRRA Strategy 1	0.1137***		
	(7.32)		
CRRA Strategy 2		0.038	
		(1.52)	
CRRA Strategy 3			0.0953***
			(5.02)
	Specific. 5: Placebo-test with random allocation		
CRRA Strategy 1	0.0077		
	(0.57)		
CRRA Strategy 2		0.0188	
		(1.18)	
CRRA Strategy 3			-0.0175
			(-1.00)
N	3553	3553	3553

Note: This table reports the marginal effects of probit regressions. Dependent variables are indicator variables of participants' chosen withdrawal strategies. The main independent variables are indicator variables that denote whether our utility specification would recommend a participant to choose a specific strategy, based on self-reported risk-aversion, bequest intensity and planning horizon. Self-reported risk-aversion and bequest are linearized into a set of seven equally spaced points on the intervals denoted in the four specifications. Our full set of controls is included in every regression. Reported are coefficients and t-statistics. ***, **, and * indicate significance at the 1%, 5% and 10%-level, respectively.

% and 19 % more likely to also select this strategy. Similarly, for the strategy with dynamic withdrawal rates (Strategy 3), we find that participants are on average between 10 % and 15 % more likely to also select the strategy if the utility function would predict it to be the optimal choice. Yet, despite these consistent results, the employed utility function appears to struggle in predicting preferences for the strategy with constant withdrawal rates (Strategy 2). While the coefficients are not only economically small, they are also not statistically different from zero. One potential driver for this inconsistency is the relation with risk aversion. Instead of the positive relation implied by Figure 5.3, our data suggests the reverse.[9] In our sample, it appears that the more risk-averse participants are, the less likely they are to select a decumulation strategy with risky asset allocation and constant withdrawals. This is not entirely unexpected. Due to the nature of how the second strategy is constructed (constant withdrawals paired with stochastic returns), it cannot guarantee that wealth levels are always sufficient to sustain the withdrawal rate. While this strategy only defaults in about 1.2 % of the time five years before the planning horizon, this risk increases to roughly 10 % until the final year. Considering that most participants are highly averse to the prospect of defaulting before the planning horizon (most important characteristic across all age groups), this fear appears to be reflected in participants' self-reported risk aversion.

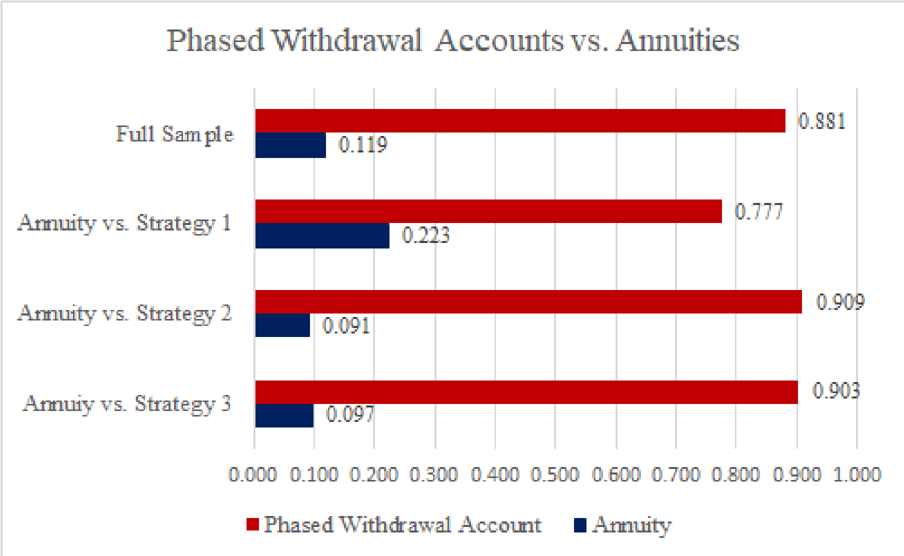
5.3.3 Determinants of Participants' Product Choice

In this section, we investigate whether participants rather prefer a lifelong annuity or a phased withdrawal account. A summary of observed choices is provided in Figure 5.4.

Overall, we observe that only a small fraction of participants preferred an annuity over a phased withdrawal plan (12 %). This result is more or less independent of the actual properties of the phased withdrawal account,

[9] Instead of regressing on the individual predictions, we could also regress on participants' self-reported risk-aversion, their bequest intentions, and on their planning horizon. Results of these regressions are reported in the Appendix (Table D.1 in the Appendix).

Figure 5.4: Annuity and Phased Withdrawal Demand



Note: This figure displays participants' choice between their preferred phased decumulation option and a lifelong annuity. Displayed are comparisons for the full sample and for each option individually. All differences are significant at the 1%-level.

although it is smaller for those participants who initially preferred the risk-free decumulation option (Strategy 1). While consistent with earlier results on the annuitization puzzle (in which participants rather prefer a lump sum payment), this finding provides a new perspective to the discussion. More precisely, it appears that individuals do not exhibit a general aversion to wealth decumulation but rather a specific aversion to invest in annuities. Looking at the difference between the risk-free phased withdrawal account and the annuity, this aversion seems even more surprising. For any wealth level and any planning horizon except 20 years, the annuity in our study provides higher monthly payments than the risk-free decumulation option, which are not only indefinitely, but also guaranteed by an insurance provider. Conversely, the only benefit of selecting a completely risk-free phased withdrawal account comes in the form of retaining control over one's financial assets, which – depending on the situation – can fulfil both precautionary and bequest motives. Yet, even for this comparison, roughly 78 % appear to favor a phased decumulation option.

Next, we investigate potential drivers of this discrepancy. Table 5.4 shows

Table 5.4: Determinants of Annuity and Phased Withdrawal Demand

Dependent Variable	*Annuity*			
	Full Sample	Strategy 1	Strategy 2	Strategy 3
Life Expectancy	0.00306***	0.00491**	0.00382**	0.00246**
	(3.57)	(2.16)	(2.29)	(2.25)
Trust in Financial Markets	−0.0138***	−0.0108	−0.00743	−0.00346
	(3.73)	(0.97)	(1.08)	(0.74)
Financial Literacy	0.00334	0.0290*	0.0151*	0.00659
	(0.65)	(1.93)	(1.86)	(1.01)
Age	−0.00147***	−0.00157	−0.00140*	−0.00235***
	(3.20)	(1.10)	(1.71)	(3.74)
Married	−0.0324**	−0.0347	−0.0485**	−0.0243
	(2.38)	(0.91)	(2.16)	(1.43)
Kids	−0.0166***	−0.0166	0.00941	−0.00247
	(3.61)	(1.08)	(1.02)	(0.44)
Log(Income)	0.0151	0.0118	0.0173	0.00256
	(1.29)	(0.42)	(0.83)	(0.16)
Log(Wealth)	−0.0417***	−0.0683***	−0.0167	−0.0234**
	(5.56)	(3.62)	(1.25)	(2.28)
Female	0.00807	0.0053	0.0527	−0.0292
	(0.48)	(0.13)	(1.58)	(1.46)
University	−0.0105	0.0173	0.0169	−0.0109
	(0.77)	(0.52)	(0.72)	(0.61)
High Consumption		0.0266**	−0.0117*	−0.00746*
		(2.28)	(1.84)	(1.75)
Constant Consumption		0.0231**	0.0237***	0.0351***
		(2.08)	(3.48)	(6.51)
Low Default Risk		−0.0133	0.00814	0.00186
		(1.00)	(1.49)	(0.47)
Bequest		−0.0152**	−0.0237***	−0.0113***
		(1.97)	(4.58)	(3.29)
N	3593	665	1029	1897
R^2	0.035	0.084	0.077	0.069

Note: This table reports results of four OLS regressions. Dependent variable is *Annuity*, a dummy variable which equals one if a participant prefers an annuity over a phased withdrawal account. Column (1) looks at the full sample independent of the previously chosen phased withdrawal account, while columns (2) to (4) condition the analysis on the chosen phased withdrawal account. Reported are coefficients and t-statistics (in parentheses). All standard errors are robust. ***, **, and * indicate significance at the 1%, 5% and 10%-level, respectively.

the results of regressing annuity (a dummy that equals one if participants would prefer an annuity to a phased decumulation product) on participants' self-reported life expectancy, trust in financial markets, financial literacy and a set of demographic variables. In columns (2) to (4), we shift the focus on the trade-off between choosing an annuity versus a specific phased withdrawal account, which a participant previously preferred. To differentiate the characteristics of the phased withdrawal accounts, we also include the self-reported importance of the four self-assessed withdrawal characteristics as additional explanatory variables.

Across all regressions, we collectively find that higher life expectancy is positively related to choosing an annuity over a phased withdrawal account. Given the fact that insurance against longevity risk is one of the primary benefits of annuities, this finding is not surprising and consistent with earlier studies (e.g. Beshears et al., 2014; Schreiber and Weber, 2016). Relatedly, results in column (1) suggest that married individuals and the number of children are negative predictors of the annuitization decision. Looking at the individual differences between annuities and specific phased withdrawal strategies as reported in columns (2) to (4) provides further insights. First, we consistently find that bequest intentions are negatively related to the annuitization decision, independent of the phased withdrawal account participants have selected. The importance of bequest motives for the selection of wealth decumulation products does not come unexpectedly. Kotlikoff and Summers (1981) estimated that a large fraction of the U.S. capital stock was attributable to intergenerational transfer. Similarly, Gale and Scholz (1994) showed that both bequests and inter vivos transfers are common and can be sizeable. Second, we find that consumers who value constant income streams are significantly more likely to select an annuity despite having the option to choose a phased withdrawal account with constant withdrawals. Finally, and perhaps most interestingly, we find no effect of participants' desire for safe income streams on their product choice, independent of their preferred phased withdrawal account. In other words, even though "avoiding the risk of depleting the capital stock early" ranks as the most important characteristic for phased withdrawals, investors seem to neglect that it is one of the primary benefits of annuities. While not entirely surprising for the first and the second phased withdrawal option, this is especially troublesome for the decumulation strategy with constant withdrawal amounts and risky asset allocation which exhibits a rather high risk of exhausting the withdrawal account before the end of life.

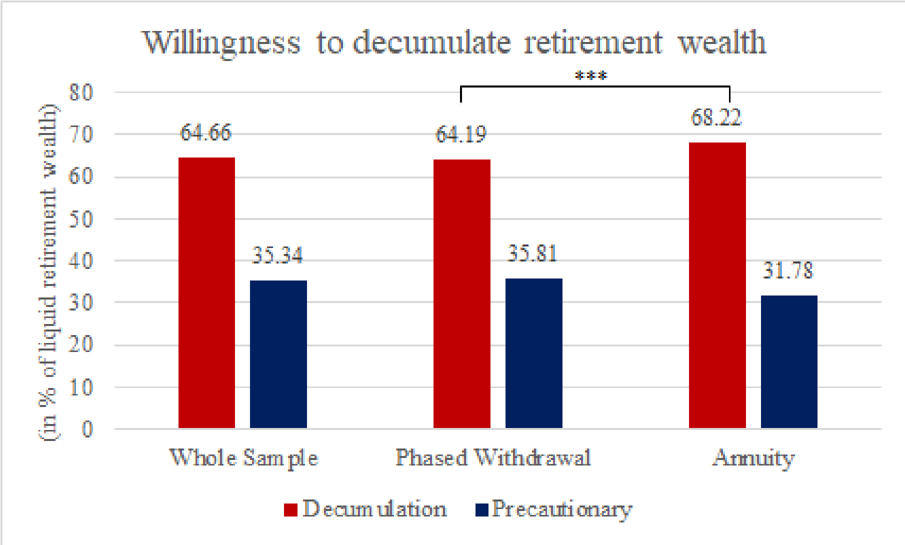
Finally, we also consider the possibility that investor sentiment might drive the low demand for annuities in our sample. Chalmers and Reuter (2012) and Previtero (2014) both document that past stock market returns have a strong effect on the demand for annuities, with high stock market returns generally reducing the demand for annuities. To account for the fact that past market conditions might drive the demand for equity investments, we investigate the 3-month and 12-month return of the DAX index prior to the release of our survey. Yet, even though both returns are positive (approx. 1.6 % and 4 %, respectively), they are below their recent historic averages (2.1 % and 8 %, respectively). While we cannot fully rule out the possibility that participants falsely use recent market returns as a proxy for future returns, the mild economic conditions during our sample period would make the capital market investment appear rather less attractive instead of more attractive.

5.3.4 The Wealth Decumulation Decision

In this section, we analyze respondents' general willingness to decumulate wealth. Figure 5.5 displays the average self-reported amount (in % of total liquid wealth) that participants would be willing to decumulate over the course of their retirement.

Overall, participants would be willing to draw down roughly 65 % of their retirement savings while they would keep 35 % as precautionary savings. Participants who prefer a lifelong annuity to a phased withdrawal would decumulate 68 % of their savings, while those preferring a phased withdrawal account would decumulate 64 % (t-statistic: 2.91, two-sided test).

Figure 5.5: Willingness to Decumulate Retirement Savings



Note: This figure displays participants' willingness how much of their retirement wealth they would be willing to decumulate and how much they would keep as precautionary savings. Displayed are comparisons for the full sample and split by the product (phased withdrawal or annuity) they have previously chosen. *** indicates significance at the 1%-level.

Yet, despite the rather high self-reported willingness to decumulate wealth, actual spending in retirement appears to be still moderately low (Olafsson and Pagel, 2018). Given that only 12 % of the respondents in our sample would select an annuity to begin with, our results suggest that the difference between observed spending and self-reported willingness to spend is driven by the lack of demand for annuities and the desire for flexibility.

Next, we seek to obtain a more pronounced understanding of the factors that affect individuals' decision of how much wealth to decumulate. In analyzing this decision, we differentiate two sets of variables related to participants' retirement preparedness. First, we look at factors that are related so successful wealth accumulation (i.e. individuals' "financial" preparedness). Second, we investigate factors which capture participants' attitude towards retirement (i.e. individuals' "emotional" preparedness).

Financial Preparedness and the Wealth Decumulation Decision

To identify factors related to successful wealth accumulation, we follow Lusardi and Mitchell (2007, 2011). The authors find that both financial literacy and planning are strong predictors for financial wellbeing in retirement. In particular, it appears that those individuals who successfully develop and stick to a saving plan not only accumulate more wealth but also make better investment decisions. Additionally, the authors find that these individuals are also more likely to follow sound financial advice and less likely to follow investment recommendations from friends and family members. As such, we include a dummy variable that equals 1 if participants report sticking to a saving plan for their retirement, a measure of financial literacy, and dummy variables that indicate the source of financial advice participants use for their retirement planning.[10] Results of 4 OLS regressions with dissave as dependent variable are reported in Table 5.5.

Based on the results of Table 5.5, we can draw several conclusions. First, those participants who report sticking to a saving plan to save for retirement show a weak tendency to decumulate a greater fraction of their accumulated savings, even after control for wealth. However, the effect is both economically small and statistically only marginally significant. Looking at other indicators of successful wealth accumulation, this relation even appears to vanish entirely. While financial literacy is a highly significant and positive predictor of wealth accumulation (both in our sample and in previous studies; see for example Behrman et al., 2012), it does not appear to be related to wealth decumulation. Similar results can be found by looking at the source of financial advice participants take. While we observe a weak negative effect for those people who take advice from their family and a weak positive effect for those who use spreadsheets and similar planning tools, none of the effects is statistically significant at the 10 %-level. One potential reason for this seemingly non-existent relationship might be that wealth decumulation

[10] To ensure the suitability of our proxies for successful wealth accumulation, we test the implications of Lusardi and Mitchell (2007, 2011) in the Appendix (Table D.2). Consistent with their study, our results leave no doubt that financial literacy and the ability to develop and stick to a saving plan are important determinants for effective retirement preparation.

Table 5.5: Financial Preparedness and Wealth Decumulation

Dependent Variable		*Dissave*		
Saving Plan	1.811**			1.572*
	(-1.96)			(-1.66)
Financial Literacy		0.286		0.0932
		(-0.67)		(-0.21)
Advice_Family			(-1.063)	(-0.908)
			(-1.08)	(-0.91)
Advice_Work			(-0.246)	(-0.218)
			(-0.17)	(-0.15)
Advice_Tool			0.81	0.48
			(-0.89)	(-0.52)
Advice_Media			0.752	0.612
			(-0.72)	(-0.58)
Advice_Advisor			-0.703	-0.781
			(-0.73)	(-0.80)
Controls?	Yes	Yes	Yes	Yes
N	3508	3508	3508	3508
R^2	0.063	0.063	0.064	0.064

Note: This table reports results of four OLS regressions. Dependent variable is *Dissave*, which can take values between 0% and 100%. Main independent variables are *Saving Plan* (1 = participant follows a saving plan for retirement), *Financial Literacy*, and five dummy variables that indicate whether a participant follows financial advice of family members, work colleagues, planning tools, the media, or from financial planners. Reported are coefficients and t-statistics (in parentheses). All standard errors are robust. ***, **, and * indicate significance at the 1%, 5% and 10%-level, respectively. We control for socio-demographic variables and household composition whenever indicated.

– in contrast to saving and investment decisions – is still a relatively new and unexplored topic for the broad population and even for most financial institutions besides insurance companies. As such, there is relatively little guidance for consumers about how much wealth should be decumulated.

Emotional Preparedness and the Wealth Decumulation Decision

To capture participants' attitudes towards retirement (i.e. their "emotional" preparedness), we include both a measure of optimism and the planning horizon over which they would want to decumulate their wealth. With our measure of optimism, we seek to capture the fact that consumers cannot rely on their experience when evaluating how much of their wealth they would be willing to decumulate. According to Puri and Robinson (2007), these are the decisions which are most affected by attitudes and emotional dispositions as there is no available data on which to base an opinion. Following Puri and Robinson (2007), optimism was measured as the difference between participants' self-reported life expectancy and that implied by statistical tables, adjusted for gender. To differentiate moderate optimism from extreme optimism, we take the right-most 5 % of optimists to be extreme optimists.[11] Including participants' planning horizon allows us to control for participants' outlook on their retirement, which is not caused by optimism (i.e. information about their health status, general longevity in their family, or aversion to the prospect of outliving their retirement resources). Yet, longer planning horizons are likely to have diverse implications for individuals who prefer phased withdrawals over annuities or vice versa. For individuals who prefer phased withdrawals, longer planning horizons should be associated with a decreased willingness to decumulate greater amounts, as there is no protection against longevity risk. Conversely, for those preferring annuities, one should expect that longer planning horizons increase the willingness to decumulate greater amounts, as predicted by Yaari (1965) or Davidoff et al. (2005). Additionally, both optimism and participants' planning horizon are likely correlated to their bequest motive (e.g. Ameriks et al., 2011). Following this logic, we include both our regular measure of optimism, a dummy variable for extreme optimists, participants' bequest motive, their planning horizon, a dummy for preferring annuities over phased withdrawals and an interaction between the last two. Results are reported in Table 5.6.

We find that the tendency to plan for a longer retirement is negatively related to the wealth decumulation decision for participants preferring phased withdrawals while having a positive impact for those preferring annuities.

[11] In our study, extreme optimists overestimate their life expectancy by roughly 20 to 30 years, similar to the 20 years reported by Puri and Robinson (2007).

In the absence of a significant proportion of annuitized wealth, households must decide which fraction of their savings they decumulate and how much they keep as precautionary savings. The longer they expect to live (or the more averse they are to the prospect of outliving their resources) the more precautionary savings they should build. Conversely, participants interested in annuities have no incentive to keep large precautionary savings in response to longer planning horizons. In particular, the longer one expects to live, the higher are the benefits from livelong pension payments and the more beneficial it becomes to drawdown a greater fraction of one's savings (while keeping smaller precautionary savings as protection against adverse health shocks). Additionally, participants' self-reported bequest intentions are a strong negative predictor of the wealth they would be willing to decumulate. Interestingly however, this relation strongly depends on their planning horizon and is most pronounced for long horizons. While our survey does not allow us to strictly disentangle strategic bequest motives (e.g. caused by public care aversion) from intentional bequest motives, this interaction rather points towards the former explanation, supporting the findings of Ameriks et al. (2011).

Regarding the impact of optimism, we find that moderate optimism is positively related to wealth decumulation. This finding is in line with the view that optimism is generally correlated with positive beliefs about future economic conditions, as postulated by Puri and Robinson (2007). As such, more optimistic individuals appear to be attracted by the prospect of a higher consumption during retirement without worrying too much about the state of their precautionary savings. In our sample, more optimistic individuals are willing to decumulate roughly 4 % more of their savings compared to rather pessimistic individuals (as defined by being one standard deviation away from the mean). While this difference is hardly decisive for living retirement in luxury or in poverty, it might benefit those retirees who systematically overestimate their life expectancy (Heimer et al., 2019). Yet, similar

Table 5.6: Emotional Outlook and Wealth Decumulation

Dependent Variable			*Dissave*		
Optimism	0.136*	0.0743	0.141*	0.137*	0.133*
	(1.76)	(1.05)	(1.82)	(1.76)	(1.70)
Extreme Optimism	-8.918***	-8.986***	-8.967***	-8.875***	-8.893***
	(-3.80)	(-3.83)	(-3.81)	(-3.78)	(-3.80)
Bequest	-5.851***	-5.939***	-5.895***	-5.875***	-1.654
	(-26.69)	(-27.08)	(-26.86)	(-26.80)	(-1.31)
Horizon	-0.346***		-0.333***	-0.414***	0.250
	(-3.00)		(-2.88)	(-3.52)	(1.08)
Annuity		-2.808**	-2.593*	-21.21**	-18.31**
		(-2.13)	(-1.96)	(-2.41)	(-2.09)
Horizon x Annuity				0.739**	0.614*
				(2.15)	(1.79)
Horizon x Bequest					-0.170***
					(-3.34)
Controls?	Yes	Yes	Yes	Yes	Yes
N	3525	3525	3525	3525	3525
R^2	0.254	0.253	0.255	0.256	0.259

Note: This table reports results of four OLS regressions. Dependent variable is Dissave, a categorial variable, which can take values between 0% and 100%. Main independent variables are *Optimism* and *Extreme_optimism*, which were constructed following Puri and Robinson (2007). Reported are coefficients and t-statistics (in parentheses). All standard errors are robust. ***, **, and * indicate significance at the 1%, 5% and 10%-level, respectively. We control for socio-demographic variables and household composition whenever indicated.

to earlier findings on optimism, we report strikingly different results for individuals who are overly optimistic. In particular, instead of moderately increasing their consumption, extreme optimists decumulate between 2 % and 5 % less than the average participant. The implications of this finding might hint at an inherent misunderstanding of how to protect against longevity risk. Given that extremely optimistic individuals overestimate their life expectancy by roughly 20 to 30 years, a 2 % to 5 % increase in precautionary savings is barely relevant to sustain their financial needs until the age of 105. Instead, those individuals would benefit most by annuitizing an even greater fraction of their accumulated savings.

Taken together, our results on optimism are largely consistent with findings from Brunnermeier and Parker (2005). In their model, forward-looking agents who believe that better outcomes are more likely, are more inclined to take ex-post optimal actions that others might find too costly ex-ante. Overly optimistic agents, however, neglect the benefit of those actions and instead perceive future income as too certain. Similarly, our findings imply that moderately optimistic individuals decumulate a greater fraction of their wealth to either increase their wellbeing over a shorter horizon (through phased withdrawals or immediate consumption) or over a longer horizon (through annuities). Extreme optimists – however – decumulate a smaller fraction of their wealth possibly to protect against longevity risk and as such neglect the uncertainty in their life expectancy and overestimate their income from non-annuitized wealth.

5.4 Conclusion

The goal of this study has been to obtain a more holistic understanding of the wealth decumulation decision from the perspective of an individual who is averse to fully annuitizing her accumulated savings in retirement. Such an individual faces the following decision problem. Out of her non-annuitized wealth, she must decide how much to allocate to a savings account (i.e. as a protection against unexpected costs) and how much (if anything) to decumulate over the course of her retirement. As an alternative way to decumulate savings, we investigate preferences for phased withdrawal products by fielding a large online survey.

Our results have several implications for the design and the demand for complementary products. In our sample, annuity demand is still relatively low, as only 12 % of respondents would prefer a lifelong annuity to decumulate savings, while 88 % would prefer some form of phased withdrawal. Offering a wider array of phased withdrawal solutions would help retirees to decumulate more of their savings, without being forced to fully convert their wealth. As flexibility in the timing of spending is among the most important factors of the annuitization decision (Beshears et al., 2014), offering combined solutions of phased withdrawals and partial annuitization could help to increase overall retirement welfare while protecting retirees against longevity risk. Yet, finding the optimal mix of phased withdrawals and annuitization remains a significant challenge.

Regarding the concrete design of phased withdrawal products, our results suggest that even in retirement most people are willing to invest in equities to sustain higher withdrawal rates. Additionally, the majority of respondents would choose a product with dynamic withdrawal rates. Given the current standard of constant payout policies (e.g. as offered by most annuity contracts), our findings highlight once more the importance and the demand for more flexible retirement solutions. Whereas similar proposals exist in the variable annuity market (e.g. penalty-free early withdrawals, or flexible payout streams), annuities face much higher hurdles to implement such suggestions due to adverse selection, which eventually increases product complexity for consumers.

Appendix A

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

A.1 Further Analyses

In this section we present results of further analyses.

Table A.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.

A.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 A.1 to A.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 A.1: Payoff Screen

Figure A.2: Probability Estimate Screen

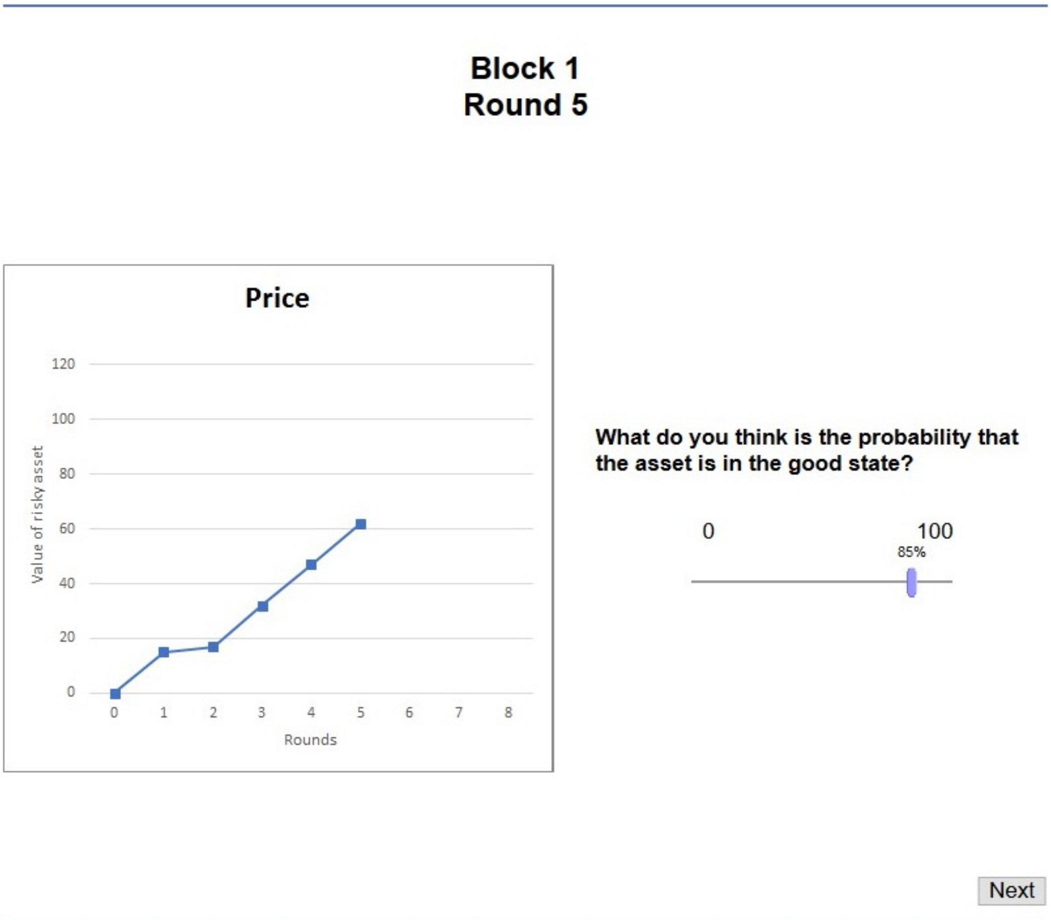
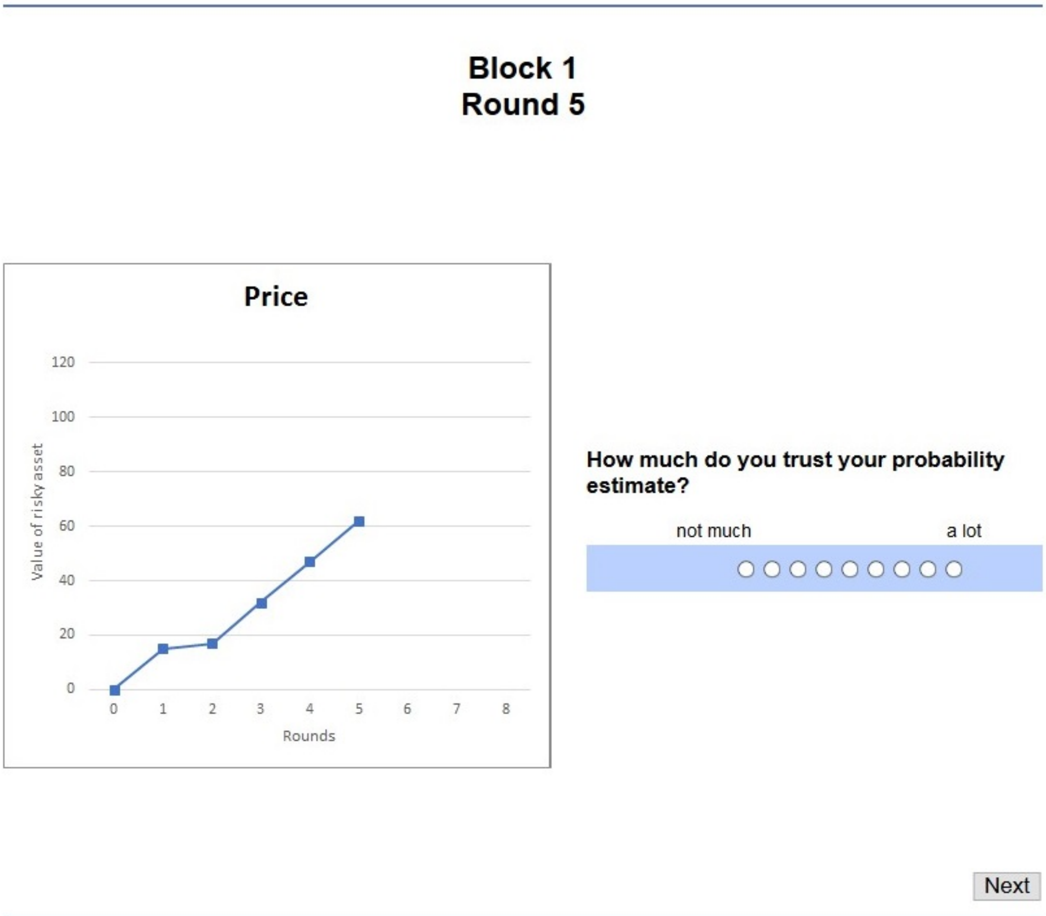


Figure A.3: Confidence Level Screen



A.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 B

Can Agents Add and Subtract When Forming Beliefs?

B.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 B.1 to B.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 B.1: Payoff Screen

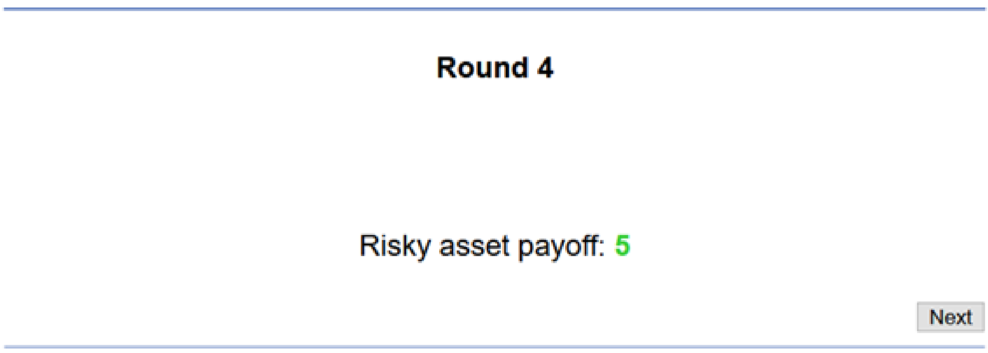
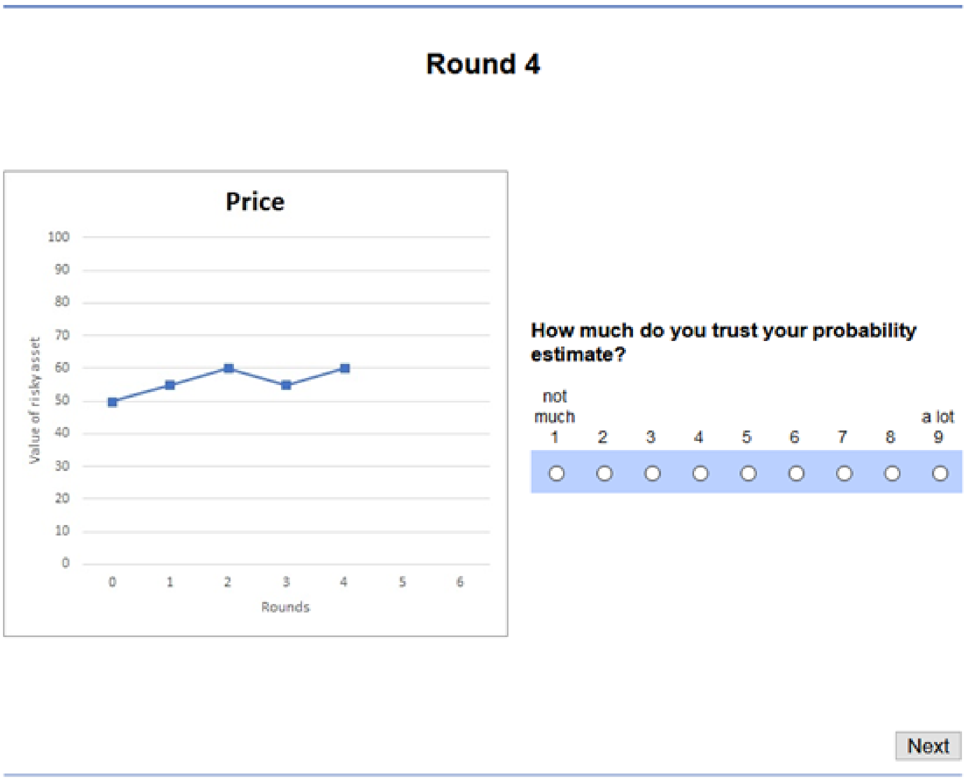


Figure B.2: Probability Estimate Screen

Figure B.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

Appendix C

Expectation Formation under Uninformative Signals

C.1 Further Analyses

In this section we present results of further analyses.

Table C.1: Uninformative Updating and Learning

Dependent Variable	*Log Odds Ratio (Subjective)* $\lambda_{i,t}$			
	Active Treatment		Passive Treatment	
	Round 1 to 5	Round 6 to 10	Round 1 to 5	Round 6 to 10
$D_{informative;\ i,t}$ *(Inference)*	0.412***	0.340***	0.604***	0.536***
	(8.39)	(6.93)	(10.58)	(10.36)
$\lambda_{i,t-1}$ *(Use of Priors)*	0.701***	0.789***	0.520***	0.689***
	(19.51)	(26.23)	(11.92)	(21.88)
$D_{uninformative;\ i,t}$	0.367***	0.310***	0.257***	0.0831
	(4.16)	(3.93)	(2.98)	(1.19)
$D_{uninformative;\ i,t}$ x	-0.544***	-0.589***	-0.596***	-0.348***
$negative_i$	(-4.04)	(-4.11)	(-4.24)	(-2.88)
Observations	1371	530	533	287
R^2	0.698	0.454	0.580	0.538

Note: This table reports the results of four OLS regressions on how information signals and their valence affect individuals' beliefs. We report results split by the first five and the last five rounds of the experiment. The dependent variable is participants' subjective log-odds ratio. *Prior Signal* is a variable taking the value 1 if the tth signal of subject i is good, 0 if the tth signal is uninformative, and -1 if the tth signal is bad. *Uninformative* is a dummy if the tth signal of subject i is uninformative, whereas *Negative* is a dummy if participant i is in the negative treatment (and zero otherwise). The interaction term thus displays whether participant i encountered a negative uninformative signal in round t. Controls include age, gender, statistical skills, risk aversion, and participants' financial literacy. Reported are coefficients and t-statistics (in parentheses). All standard errors are clustered at the individual level. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

C.2 Experimental Instructions and Screenshots

Instructions Bayesian Updating (Exemplary for Positive Treatment)

In this part we would like to test your forecasting abilities. You will make forecasting decisions in ten consecutive rounds.

Suppose you find yourself in an environment, in which a risky asset can pay a dividend of either -3, $+1$, or $+5$. The probability of each outcome 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 it pays a dividend of $+5$ is 50%, the probability that it pays a dividend of -1 is 30% and the probability that it pays a dividend of -3 is 20%. If the risky asset is in the **bad** state, then the probability that it pays a dividend of $+5$ is 20%, the probability that it pays a dividend of -1 is 30% and the probability that it pays a dividend of -3 is 50%.

The computer determines the state of the risky asset before the first round. Afterwards, the state does not change and remains fixed. At first, 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 a dividend payment of the risky asset (-3, $+1$, 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 previous dividend payments 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:

$$\frac{1}{1+\frac{1-p}{p}\cdot\left(\frac{q}{1-q}\right)^{(n-u)-2t}},$$

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 (71.43% here). Finally, t and u are the number of observed high signals and uninformative signals until trial n, respectively. Displayed are results for 8 rounds, as probabilities converge quickly to 1.

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	28.57%
1	1	71.43%
2	0	13.79%
2	1	50.00%
2	2	86.21%
3	0	6.02%
3	1	28.57%
3	2	71.43%
3	3	93.98%
4	0	2.50%
4	1	13.79%
4	2	50.00%
4	3	86.21%
4	4	97.50%
5	0	1.01%
5	1	6.02%
5	2	28.57%
5	3	71.43%
5	4	93.98%
5	5	98.99%
6	0	0.41%
6	1	2.50%
6	2	13.79%
6	3	50.00%
6	4	86.21%
6	5	97.50%
6	6	99.59%
7	0	0.16%
7	1	1.01%
7	2	6.02%
7	3	28.57%
7	4	71.43%
7	5	93.98%
7	6	98.99%
7	7	99.84%
8	0	0.07%
8	1	0.41%
8	2	2.50%
8	3	13.79%
8	4	50.00%
8	5	86.21%
8	6	97.50%
8	7	99.59%
8	8	99.93%

Screenshots of the Experiment

Figures C.1 to C.5 present the screens of the forecasting task (screens that only belong to the active treatment are marked as [active]) as seen by subjects in the experiment. One round consists of three [five] sequential screens.

Figure C.1: Investment Screen

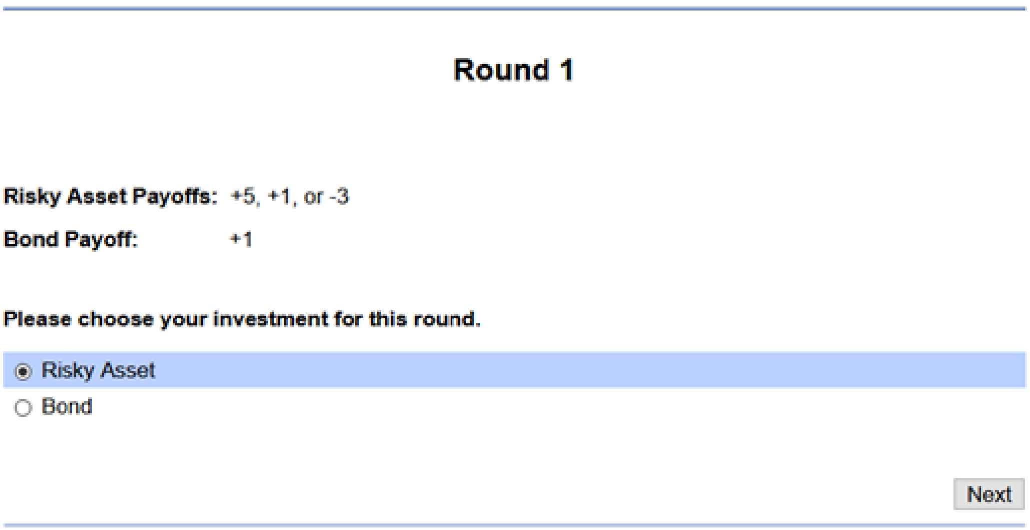


Figure C.2: Payoff Screen

Figure C.3: Accumulated Payoffs Screen

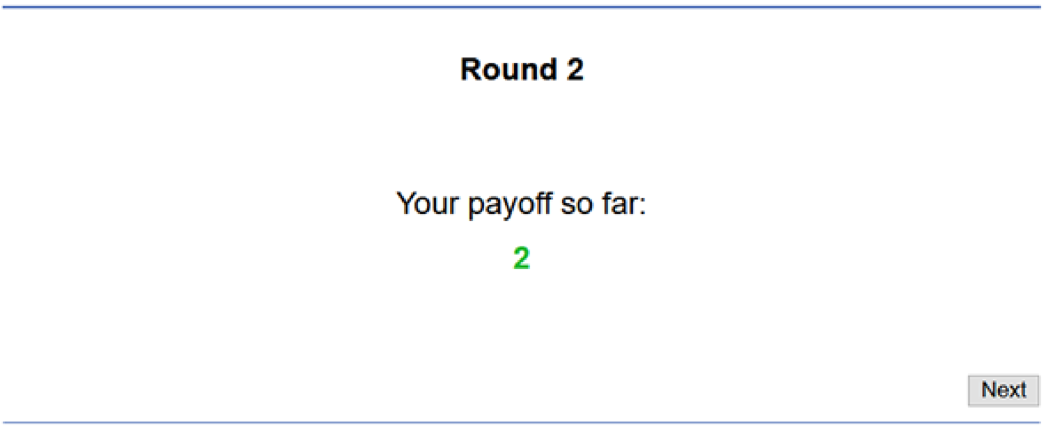


Figure C.4: Probability Estimate Screen

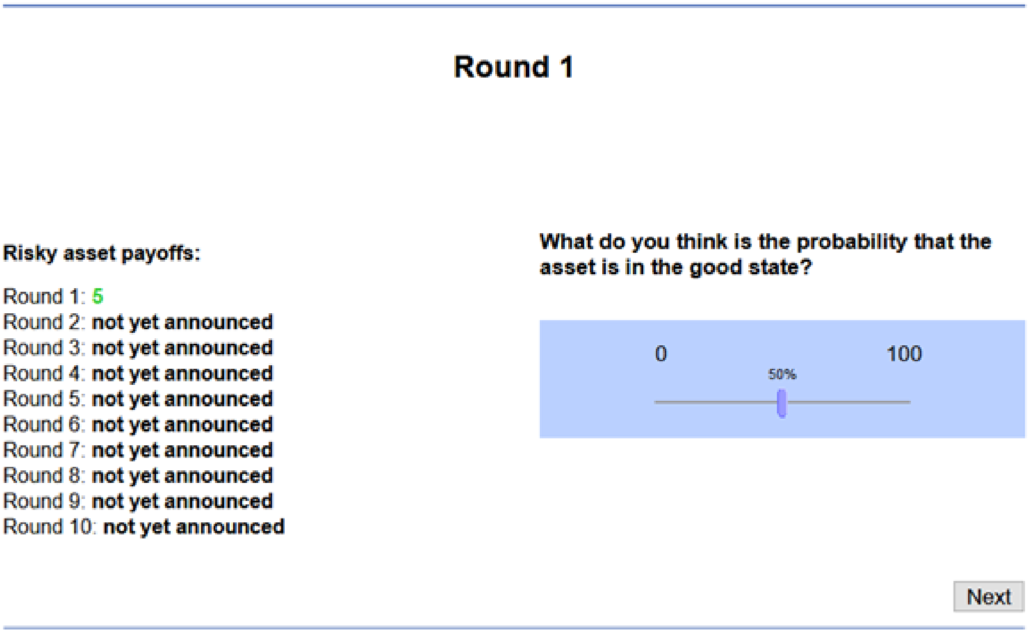
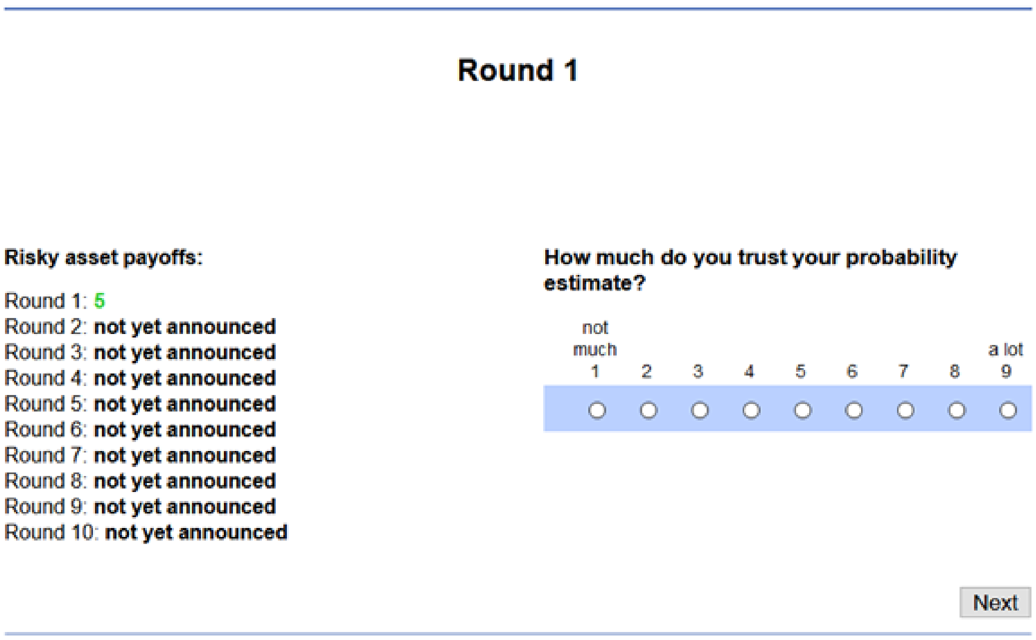


Figure C.5: Confidence Level Screen



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

C.3 Additional Experimental Measures

Risk Aversion

Below we report the risk aversion question adopted from Kuhnen (2015):

Imagine you have saved $\$10,000$. You can now invest this money over the next year using two investment options: a U.S. stock index mutual fund, which tracks the performance of the U.S. stock market, and a savings account. The annual return per dollar invested in the stock index fund will be either +40% or -20%, with equal probability. In other words, it is equally likely that for each dollar you invest in the stock market, at the end of the one year investment period, you will have either gained 40 cents, or lost 20 cents. For the savings account, the known and certain rate of return for a one year investment is 5%. In other words, for each dollar you put in the savings account today, for sure you will gain 5 cents at the end of the one year investment period. We assume that whatever amount you do not invest in stocks will be invested in the savings account and will earn the risk-free rate of return. Given this information, how much of the $\$10,000$ will you invest in the U.S. stock index fund? Choose an answer that you would be comfortable with if this was a real-life investment decision.

[Please enter a value between 0 and 10,000 here]

Financial Literacy

Below we report the financial literacy question adopted from Kuhnen (2015). Correct responses are displayed in *italic*.

Let's say that when you answered the prior question you decided to invest x dollars out of the $\$10,000$ amount in the U.S. stock index fund, and therefore you put (10,000 - x) dollars in the savings account. Recall that over the next year the rate of return of the stock index fund will be +40% or -20%, with equal probability. For the savings account, the rate of return is 5% for sure. What is the amount of money you expect to have at the end of this one year investment period?

Please choose one of the answers below

[A] $0.5(0.4x - 0.2x) + 0.05(10,000 - x)$

[B] $1.4x + 0.8x + 1.05(10,000 - x)$

[C] $0.4(10,000 - x) - 0.2(10,000 - x) + 0.05x$

[D] $0.5(0.4(10,000 - x) - 0.2(10,000 - x)) + 0.05x$

[E] $0.4x - 0.2x + 0.05(10,000 - x)$

[F] $0.5(1.4x + 0.8x) + 1.05(10,000 - x)$

[G] $1.4(10,000 - x) + 0.8(10,000 - x) + 1.05x$

[H] $0.5(1.4(10,000 - x) + 0.8(10,000 - x)) + 1.05x$

C.4 Derivations and Proofs

Sequential Bayesian Updating Behavior

Following Dave and Wolfe (2003) and Charness and Dave (2017), we briefly sketch individuals' sequential updating behavior as prescribed by Bayes' Law:

Suppose there are two possible states of the world, denoted 'G' (for good) and 'B' (for bad). Additionally, over the course of t rounds, individuals may observe signals that are either indicative of a good state g, or of a bad state b, or which are non-diagnostic about the underlying state u.
Within a given round t, Bayes' Rule assumes that individuals posterior logs π_{1t} are formed as a function of their prior logs π_0 and some likelihood L_k:

$$\pi_{1k} = L_k \pi_0$$

Given that only two signals are indicative about the possibles states 'G' and 'B', the likelihood L_k takes the following form:

$$L_k = \left(\frac{\theta}{1-\theta}\right)^{z_t},$$

where θ is the proportion of g signals to b signals and z_t is the difference between the number of g signals and b signals as of the t^{th} round. Note that only the difference of informative signals is important for the likelihood function, as the non-diagnostic signal u does not provide any relevant information for the decision maker. Combining the above two equations yields the following.

$$\pi_{1k} = \left(\frac{\theta}{1-\theta}\right)^{z_t} \pi_0$$

Taking logs now yields:

$$\ln \pi_{1k} - \ln \pi_0 = z_t \cdot \left(\frac{\theta}{1-\theta}\right)$$

Finally, first differencing the above equation yields,

$$\Delta \ln \pi_{1k} = \Delta z_t \cdot \left(\frac{\theta}{1-\theta} \right)$$

where $\Delta z_t \in \{-1, 0, 1\}$. The first differenced equation demonstrates that – in absolute terms – a Bayesian agent updates, in log odds terms, at a constant of $\left(\frac{\theta}{1-\theta} \right)$.

Deriving the Regression Equation

Equation 4.8 is obtained by taking the natural logarithm of equation 4.7:

$$\ln \frac{\pi(G|S)}{\pi(B|S)} = \ln \left(\left[\frac{p(S|G)}{p(S|B)} \right]^{c_0 + I\{u|\ desirable\}c_1 + I\{u|\ undesirable\}c_2} \left[\frac{p(G)}{p(B)} \right]^{d} \right)$$

$$\Leftrightarrow \ln \frac{\pi(G|S)}{\pi(B|S)} = (c_0 + I\{u|\ desirable\}c_1 + I\{u|\ undesirable\}c_2) \cdot \ln \frac{p(S|G)}{p(S|B)} + d \cdot \ln \frac{p(G)}{p(B)}$$

Next, note that an agents' prior belief about the objective state of the world $\frac{p(G)}{p(B)}$ is equal to the agents' posterior belief from last period $\frac{\pi(G|s_1,\ldots,s_{t-1})}{\pi(B|s_1,\ldots,s_{t-1})}$.

$$\Leftrightarrow \ln \frac{\pi(G|S)}{\pi(B|S)} = (c_0 + I\{u|\ desirable\}c_1 + I\{u|\ undesirable\}c_2) \cdot \ln \frac{p(S|G)}{p(S|B)} + d \cdot \ln \frac{\pi(G|s_1,\ldots,s_{t-1})}{\pi(B|s_1,\ldots,s_{t-1})}$$

Additionally, we accommodate the fact that a Bayesian agent updates his beliefs, in log odds terms, at a constant of $\Delta z_t \cdot \left(\frac{\theta}{1-\theta}\right)$. To do so, we follow Charness and Dave (2017) and Benjamin (2019), and replace $\ln \frac{p(S|G)}{p(S|B)}$ with a dummy $D_{informative}$ taking the value 1 if the tth signal is g, 0 if the tth signal is u, and -1 if the tth signal is b. This alternate specification is equivalent, but the coefficient c_0 needs to be interpreted relative to $\left(\frac{\theta}{1-\theta}\right)$ instead of 1. To test whether individuals update their prior beliefs in response to non-diagnostic signals (i.e. when the tth signal is u), we additionally add two dummies, $D_{uninformative|desirable}$ and $D_{uninformative|undesirable}$ which equal 1 if the tth signal is u and if the signal is either in the positive domain, or in the negative domain, respectively:

$$
\begin{aligned}
\Rightarrow \ln \frac{\pi(G|S)}{\pi(B|S)} \\
&= c_0 \cdot D_{informative} + c_1 \cdot D_{uninformative|desirable} \\
&\quad + c_2 \cdot D_{uninformative|undesirable} + d \cdot \ln \frac{\pi(G|s_1,\ldots,s_{t-1})}{\pi(B|s_1,\ldots,s_{t-1})}
\end{aligned}
$$

Finally, note that the natural logarithm of subject's i odds ratio, based on her stated probability $P_{it}\left(G|s_1,\ldots,s_t\right)$ that the asset is paying dividends from the good state is:

$$
\lambda_{it} = \ln\left(\frac{\lambda\left(G|s_1,\ldots,s_t\right)}{\lambda\left(B|s_1,\ldots,s_t\right)}\right) = \ln\left(\frac{P_{it}\left(G|s_1,\ldots,s_t\right)}{1-P_{it}\left(G|s_1,\ldots,s_t\right)}\right)
$$

which may differ from the objective Bayesian probability. Incorporating this into the above equation, the final regression equation that we seek to estimate is as follows:

$$
\begin{aligned}
\Rightarrow \ln \frac{\lambda(G|s_1,\ldots,s_t)}{\lambda(B|s_1,\ldots,s_t)} \\
&= \widehat{\beta_1} \cdot D_{informative} + \widehat{\beta_2} \cdot \ln \frac{\lambda(G|s_1,\ldots,s_{t-1})}{\lambda(B|s_1,\ldots,s_{t-1})} \\
&\quad + \widehat{\beta_3} \cdot D_{uninformative|desirable} + \widehat{\beta_4} \cdot D_{uninformative|undesirable} + \varepsilon_t
\end{aligned}
$$

Appendix D

When Saving is Not Enough – Wealth Decumulation in Retirement

D.1 Further Analyses

In this section we present results of further analyses.

Table D.1: Utility Parameters and Account Choice

Dep. Variable	Strategy 1: Risk-free – Constant		Strategy 2: Risky – Constant		Strategy 3: Risky – Dynamic	
Risk Aversion	0.067***	0.061***	−0.029***	−0.025***	−0.040***	−0.037***
	(16.67)	(15.18)	(6.09)	(5.19)	(7.45)	(6.79)
Bequest	−0.007**	−0.002	0.029***	0.026***	−0.023***	−0.025***
	(2.13)	(0.53)	(8.37)	(6.85)	(5.83)	(5.93)
Planning Horizon	−0.009***	−0.004**	0.001	−0.001	0.007***	0.005**
	(5.63)	(2.40)	(0.68)	(0.70)	(3.45)	(2.11)
Controls?	No	Yes	No	Yes	No	Yes
N	3596	3592	3596	3592	3596	3592
pseudo R^2	0.10	0.13	0.02	0.03	0.02	0.02

Note: This table reports the marginal effect of six probit regressions. Dependent variables are indicator variables which equal 1 if participants have selected a given phased withdrawal strategy. Reported are coefficients and z-statistics (in parentheses). All standard errors are robust. ***, **, and * indicate significance at the 1%, 5% and 10%-level, respectively. We have between 3596 and 3592 observations due to missing answers. We control for socio-demographic variables and household composition whenever indicated.

Table D.2: Retirement Preparation and Savings

Dependent Variable	(1) Saving Plan	(2) Saving Plan	(3) Wealth (log)	(4) Wealth (log)
Fin. Literacy	0.0538***	0.0389***	0.186***	0.145***
	(7.51)	(5.12)	(14.76)	(11.36)
Advice_Family	−0.0682***	−0.0791***	−0.114***	−0.108***
	(3.73)	(4.34)	(3.69)	(3.48)
Advice_Work	−0.0148	−0.0308	−0.150***	−0.157***
	(0.58)	(1.19)	(3.69)	(3.96)
Advice_Tool	0.207***	0.179***	0.0916***	0.0762***
	(13.04)	(11.15)	(3.34)	(2.83)
Advice_Expert	0.0519***	0.0507***	0.0752**	0.0779**
	(2.95)	(2.91)	(2.39)	(2.53)
Advice_Media	0.0712***	0.0650***	0.0624**	0.0495
	(3.85)	(3.54)	(1.97)	(1.60)
Saving Plan			0.144***	0.127***
			(4.91)	(4.39)
Controls?	No	Yes	No	Yes
N	3.578	3.578	3.578	3.578
R^2	0.08	0.11	0.1	0.15

Note: This table reports results of four OLS regressions. Dependent variable in specification (1) and (2) is Saving plan, a dummy variable, which equals 1 if participants report to stick to a saving plan in preparing for retirement. Dependent variable in specification (2) is the log of wealth. Reported are coefficients and t-statistics (in parentheses). All standard errors are robust. ***, **, and * indicate significance at the 1%, 5% and 10%-level, respectively. We have 3578 observations due to missing answers. We control for socio-demographic variables and household composition whenever indicated.

D.2 Survey Instructions and Screenshots

Welcome Insturctions

Dear participant,

on the following pages, you will find a survey of the **University of Mannheim in cooperation with FAZ.NET**. The survey covers topics on retirement planning. In particular, it deals with the questions of how to convert our savings into a stream of income once we enter retirement, in order to increase our standard of living.

Please note that the survey takes some time (approx. 15 minutes). Also, note that you might encounter questions that require some time to answer (just as your retirement planning!). In return, we will present you different ways on how to convert your savings into a steady stream of income. Should you be interested in further results, we will gladly send you a summary of the main results after the completion of the study via email.

In addition, we are giving away ten Behavioral Finance volumes on the subject "Entsparen im Alter — Portfolioentnahmestrategien in der Rentenphase" by Prof. Dr. Dr. h.c. Martin Weber from the University of Mannheim. **All data collected here is anonymous and exclusively used for research purposes**.

We are looking forward to your participation!

Financial Literacy Questions

Below we report the financial literacy questions that were used to calculate participants' financial literacy score. Correct responses are displayed in *italic*.

1. Do you think the following statement is true of false? "Buying a single company's stock usually provides a safer return than a stock mutual fund."
   (a) The state is true
   (b) *The statement is false*
   (c) Do not know / Refuse to answer

2. If the interest rate falls, what should happen to bond prices?
   (a) *Rise*
   (b) Stay the same
   (c) Fall
   (d) None of the above
   (e) Do not know / Refuse to answer

3. Consider a call-option with a stock as underlying. Please judge the following statement: "The price of the call-option should increase if the volatility of the underlying stock increases."
   (a) *The state is true*
   (b) The statement is false
   (c) Do not know / Refuse to answer

4. Suppose you have 100 € in a savings account and the interest rate is 4% per year and you never withdraw money or interest payments. After 10 years, how much would you have in this account in total?
   (a) *More than 140 €*
   (b) Exactly 140 €
   (c) Less than 140 €
   (d) Do not know / Refuse to answer

5. Suppose you owe 3,000 € on your credit card. You pay a minimum payment of 30 € each month. At an annual percentage rate of 12% (or 1% per month), how many years would it take to eliminate your credit card debt if you made no additional new charges?
   (a) Less than 5 years
   (b) Between 5 and 10 years
   (c) Between 10 and 15 years
   (d) *Never*
   (e) Do not know / Refuse to answer

6. A very volatile asset either increases in value by 70% or decreases in value by 60% in every period, each growth rate realizing with a change of one half. If the investor buys the asset she must hold it for 12 periods. With an initial value of 10,000 what would the asset likely be worth at the end of period 12?
   (a) *Up to 6,400*
   (b) Between 6,400 and 12,800
   (c) Between 12,800 and 19,200
   (d) Between 19,200 and 25,600
   (e) Above 25,600
   (f) Do not know / Refuse to answer

Screenshots of the Phased Withdrawal Accounts and the Annuity

Figures D.1 and D.2 present the screen of the phased withdrawal account choice and the annuity choice as seen in the survey.

Figure D.1: Display of Phased Withdrawal Accounts

Below you will find 3 potential strategies for you, which are based on the following information you provided:

- **Saved assets at retirement: about 500.000 €** [Angabe: über 500.000 €]
- **Planning horizon (as from retirement): 25 years**

The strageties are used to distribute your wealth over the course of your retirement and they differ with regard to the way **your wealth is invested** (riskfree vs. risky) and the way **your wealth is consumed annually** (constant vs. fluctuating consumption).

Please examine the strategies thoroughly and decide which one you would choose **at the beginning of your retirement.**

(Note: By clicking on the information boxes, you will receive more information or explanations regarding the terms.)

Strategy 1: Constant withdrawals- risk-free allocation

Advantages	Disadvantages
✓ Constant withdrawals	× Relatively low withdrawal amounts due to riskfree investment

Annual withdrawals over 25 years	23.050 €
Withdrawal fluctuation	± 0 %
Withdrawal in worst 5% of the cases	23.050 €
Default risk	0 %
Average possible bequest	0 €

Strategy 2: Constant Consumption - risky

Advantages	Disadvantages
✓ Constant Consumption ✓ Higher consumption due to risky investment	× Default risk due to constant consumption

Annual consumption over 25 years	27.500 €
Consumption fluctuations	± 0 %
Consumption in worst 5% of the cases	0 € (Default)
Default risk	10 %
Average possible bequest	1.110.300 €

Strategy 3: Fluctuating Consumption - risky

Advantages	Disadvantages
✓ Highest average consumption ✓ No default risk	× Relatively high consumption fluctuations

Annual Consumption over 25 years	37.250 €
Consumption fluctuations	± 33 %
Consumption in worst 5% of the cases	18.230 €
Default risk	0 %
Average possible bequest	0 €

Figure D.2: Display of Lifelong Annuity

As an alternative to a **drawdown strategy** with given planning horizon, you could exchange your saved wealth for an **immediate annuity**.

In the case of an immediate annuity an insurance provider guarantees **save pension payments every year for the rest of your life** in return for a lump-sum payment.

For your wealth of 500.000 € a potential offer might look the following:

Option 4: Lifelong immediate annuity

Advantages	Disadvantages
✓ Constant payments	× No control over accumulated wealth
✓ Pension payments guaranteed until the day of death	× No financial bequest possible

Annutal pension payments	27.979 €
Fluctuatings in pension payments	± 0 %
Payments in worst 5% of the cases	27.979 €
Default risk	0 %
Average possible bequest	0 €

20. Suppose you had to decide between a decumulation strategy and a lifelong annuity. Which would you prefer?

- ◯ I would prefer a decumulation strategy
- ◯ I would prefer a lifelong annuity

Self-reported Importance of Various Retirement Characteristics

1. **Please rate following statements on a scale from 1 (not very important) to 7 (very important)**

   (a) "How important are high withdrawal rates for you?"

   (b) "How important is it for you that the withdrawal amount remains constant over time?"

   (c) "How important are withdrawals which cannot deplete the capital stock early?"

   (d) "How important are bequests for you?"

2. **Which factors not previously mentioned affected your choice?**

3. **Would you say your current health status is...**

   (a) Very Good

   (b) Good

   (c) Medium

   (d) Rather Bad

   (e) Very Bad

4. **If you think about it, to what age do you expect to live?**

   Please tell us about the ways you tried to figure out how much your household would need for retirement.

   (a) Did you talk to family and relatives?

   (b) Did you talk to co-workers or friends?

   (c) Did you use calculators or worksheets that are computer- or internet-based?

   (d) Did you consult a financial planner or advisor or an accountant?

   (e) Did you follow advice received from the media?

5. **Have you ever tried to figure out how much your household would need to save for retirement?**

6. **Do you work in the financial industry or do you have an education in a financial domain?**

Need for Cognition Inventory (adopted from Epstein et al., 1996)

Please rate the following statements on a scale from 1 (do not agree at all) to 7 (do fully agree).

1. I don't like to have to do a lot of thinking (R)
2. I try to avoid situations that require thinking in depth about something (R)
3. I prefer to do something that challenges my thinking rather than something that requires little thought
4. I prefer complex to simple problems
5. Thinking hard and for a long time about something gives me little satisfaction (R)

D.3 Construction of the Phased Withdrawal Strategies

Strategy 1: Risk-Free – Constant

The withdrawal amount C_t in every period t is defined using the following formula:

$$C_t = \frac{(1 + r_f)^{H-1} \cdot r_f}{(1 + r_f)^{H} \cdot -1} \cdot W_0,$$

where r_f is the real risk-free rate of return, H is the planning horizon in years, and W_0 is the initial wealth that an agent wants to decumulate.

Strategy 2: Risky – Constant

Strategy 2 was constructed such that the real withdrawal amount is a fixed percentage of the initial wealth level. The percentage was chosen in a way such that the default probability remains constant at 10% (determined with simulations). In other words, as long as there is enough wealth, the real withdrawal amount remains constant every year. If there is not enough wealth, the remaining wealth is withdrawn and the strategy ends prematurely (i.e. it defaults in our terminology).

$$C_t = w \cdot W_0 \mid W_t > C_t, \quad else$$

$$C_t = W_t$$

Given the underlying asset allocation, the resulting withdrawal rates are as follows:

Planning Horizon	Withdrawal Rate
20 Years	6.27%
25 Years	5.50%
30 Years	5.13%

Strategy 3: Risky – Dynamic

The real withdrawal amount C_t in every period t is defined using the following formula:

$$C_t = \frac{(1 + E[r])^{H-t} \cdot E[r]}{(1 + E[r])^{H+1-t} - 1} \cdot W_t,$$

where $E[r]$ is the real expected return of the underlying investment strategy, H is the planning horizon, and W_t is the current wealth level in period t that the agent wants to decumulate. Whenever the expected return in any given period does not equal the realized return, the consumption does not equal to the consumption in the previous period. In other words, the resulting consumption pattern fluctuates and directly depends on the realized return and the number of periods that are left ($H - t$).

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. i. *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.

Ameriks, J., Caplin, A., Laufer, S., and Van Nieuwerburgh, S. (2011). The joy of giving or assisted living? using strategic surveys to separate public care aversion from bequest motives. *The Journal of Finance*, 66(2):519–561.

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.

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., 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.

Bateman, H., Eckert, C., Iskhakov, F., Louviere, J., Satchell, S., and Thorp, S. (2018). Individual capability and effort in retirement benefit choice. *Journal of Risk and Insurance*, 85(2):483–512.

Bayes, T. (1763). Lii. an essay towards solving a problem in the doctrine of chances. by the late rev. mr. bayes, frs communicated by mr. price, in a letter to john canton, amfr s. *Philosophical transactions of the Royal Society of London*, (53):370–418.

Behrman, J. R., Mitchell, O. S., Soo, C. K., and Bravo, D. (2012). How financial literacy affects household wealth accumulation. *American Economic Review*, 102(3):300–304.

Bell, D. E. (1982). Regret in decision making under uncertainty. *Operations Research*, 30(5):961–981.

Bénabou, R. (2013). Groupthink: Collective delusions in organizations and markets. *Review of Economic Studies*, 80(2):429–462.

Benartzi, S. (2001). Excessive extrapolation and the allocation of 401 (k) accounts to company stock. *The Journal of Finance*, 56(5):1747–1764.

Benartzi, S., Previtero, A., and Thaler, R. H. (2011). Annuitization puzzles. *Journal of Economic Perspectives*, 25(4):143–164.

Bengen, W. P. (1994). Determining withdrawal rates using historical data. *Journal of Financial Planning*, 7(4):171–181.

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.

Benoît, J.-P. and Dubra, J. (2018). When do populations polarize? an explanation. Working Paper.

Bernoulli, D. (1954). A new theory on the measurement of risk (l. sommer, trans.). *Econometrica*, 22(1):23–36.

Beshears, J., Choi, J. J., Laibson, D., Madrian, B. C., and Zeldes, S. P. (2014). What makes annuitization more appealing? *Journal of Public Economics*, 116:2–16.

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

Brown, J. R. (2007). Rational and behavioral perspectives on the role of annuities in retirement planning. NBER Working Paper Series No. 13537.

Brown, J. R., Kling, J. R., Mullainathan, S., and Wrobel, M. V. (2008). Why don't people insure late-life consumption? a framing explanation of the under-annuitization puzzle. *American Economic Review*, 98(2):304–309.

Brunnermeier, M. K. and Parker, J. A. (2005). Optimal expectations. *American Economic Review*, 95(4):1092–1118.

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.

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.

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.

Campbell, J. Y. and Shiller, R. J. (1988a). The dividend-price ratio and expectations of future dividends and discount factors. *The Review of Financial Studies*, 1(3):195–228.

Campbell, J. Y. and Shiller, R. J. (1988b). Stock prices, earnings, and expected dividends. *The Journal of Finance*, 43(3):661–676.

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.

Chalmers, J. and Reuter, J. (2012). How do retirees value life annuities? evidence from public employees. *The Review of Financial Studies*, 25(8):2601–2634.

Charness, G. and Dave, C. (2017). Confirmation bias with motivated beliefs. *Games and Economic Behavior*, 104:1–23.

Clement, M. B. and Tse, S. Y. (2005). Financial analyst characteristics and herding behavior in forecasting. *The Journal of Finance*, 60(1):307–341.

Cocco, J. F., Gomes, F. J., and Maenhout, P. J. (2005). Consumption and portfolio choice over the life cycle. *The Review of Financial Studies*, 18(2):491–533.

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.

Cooley, P. L., Hubbard, C. M., and Walz, D. T. (1998). Retirement savings: Choosing a withdrawal rate that is sustainable. *AAII Journal*, 20(2):16–21.

Coval, J. D. and Moskowitz, T. J. (1999). Home bias at home: Local equity preference in domestic portfolios. *The Journal of Finance*, 54(6):2045–2073.

Daniel, K., Hirshleifer, D., and Subrahmanyam, A. (1998). Investor psychology and security market under- and overreactions. *The Journal of Finance*, 53(6):1838–1885.

Dave, C. and Wolfe, K. W. (2003). On confirmation bias and deviations from bayesian updating. Working Paper.

Davidoff, T., Brown, J. R., and Diamond, P. A. (2005). Annuities and individual welfare. *American Economic Review*, 95(5):1573–1590.

De Long, J. B., Shleifer, A., Summers, L. H., and Waldmann, R. J. (1990). Noise trader risk in financial markets. *Journal of political Economy*, 98(4):703–738.

Dimmock, S. G., Kouwenberg, R., Mitchell, O. S., and Peijnenburg, K. (2016). Ambiguity aversion and household portfolio choice puzzles: Empirical evidence. *Journal of Financial Economics*, 119(3):559–577.

DuCharme, W. M. and Peterson, C. R. (1968). Intuitive inference about normally distributed populations. *Journal of Experimental Psychology*, 78(2p1):269.

Eddy, D. M. (1982). Probabilistic reasoning in clinical medicine: Problems and opportunities. In *Judgment Under Uncertainty: Heuristics and Biases*, pages 249–267. Kahneman, D.; Slovic, P.; Tversky, A. (Eds.), University Press, New York.

Efron, B. (2013). Bayes' theorem in the 21st century. *Science*, 340(6137):1177–1178.

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.

Enke, B. and Graeber, T. (2019). Cognitive uncertainty. Working Paper.

Ensthaler, L., Nottmeyer, O., Weizsäcker, G., and Zankiewicz, C. (2018). Hidden skewness: On the difficulty of multiplicative compounding under random shocks. *Management Science*, 64(4):1693–1706.

Epstein, S., Pacini, R., Denes-Raj, V., and Heier, H. (1996). Individual differences in intuitive–experiential and analytical–rational thinking styles. *Journal of Personality and Social Psychology*, 71(2):390–405.

Fernandes, D., Lynch Jr, J. G., and Netemeyer, R. G. (2014). Financial literacy, financial education, and downstream financial behaviors. *Management Science*, 60(8):1861–1883.

Fischhoff, B. and Bar-Hillel, M. (1984). Diagnosticity and the base-rate effect. *Memory & Cognition*, 12(4):402–410.

Gale, W. G. and Scholz, J. K. (1994). Intergenerational transfers and the accumulation of wealth. *Journal of Economic Perspectives*, 8(4):145–160.

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.

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.

Gilovich, T., Griffin, D., and Kahneman, D. (2002). *Heuristics and biases: The psychology of intuitive judgment*. Cambridge University Press.

Ginosar, Z. and Trope, Y. (1980). The effects of base rates and individuating information on judgments about another person. *Journal of Experimental Social Psychology*, 16(3):228–242.

Ginosar, Z. and Trope, Y. (1987). Problem solving in judgment under uncertainty. *Journal of Personality and Social Psychology*, 52(3):464–474.

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.

Gödker, K., Jiao, P., and Smeets, P. (2019). Investor memory. Working Paper.

Goetzmann, W. N. and Kumar, A. (2008). Equity portfolio diversification. *Review of Finance*, 12(3):433–463.

Gomes, F. J., Kotlikoff, L. J., and Viceira, L. M. (2008). Optimal life-cycle investing with flexible labor supply: A welfare analysis of life-cycle funds. *American Economic Review*, 98(2):297–303.

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.

Grether, D. M. (1980). Bayes rule as a descriptive model: The representativeness heuristic. *The Quarterly Journal of Economics*, 95(3):537–557.

Grossman, S. J. and Shiller, R. J. (1981). The determinants of the variability of stock market prices. *American Economic Review*, 71(2):222–227.

Guiso, L., Sapienza, P., and Zingales, L. (2018). Time varying risk aversion. *Journal of Financial Economics*, 128(3):403–421.

Han, B., Hirshleifer, D., and Walden, J. (2019). Visibility bias in the transmission of consumption beliefs and undersaving. Technical report, National Bureau of Economic Research.

Heimer, R. Z., Myrseth, K. O. R., and Schoenle, R. S. (2019). Yolo: Mortality beliefs and household finance puzzles. *The Journal of Finance*, 74(6):2957–2996.

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.

Kahan, D. M. (2013). Ideology, motivated reasoning, and cognitive reflection: An experimental study. *Judgment and Decision Making*, 8:407–424.

Kahneman, D. and Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, 3(3):430–454.

Kahneman, D. and Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, 80(4):237.

Klibanoff, P., Marinacci, M., and Mukerji, S. (2005). A smooth model of decision making under ambiguity. *Econometrica*, 73(6):1849–1892.

Knight, F. H. (1921). *Risk, uncertainty and profit*, volume 31. Houghton Mifflin.

König-Kersting, C. and Trautmann, S. T. (2018). Countercyclical risk aversion: Beyond financial professionals. *Journal of Behavioral and Experimental Finance*, 18:94–101.

Körding, K. P. and Wolpert, D. M. (2004). Bayesian integration in sensorimotor learning. *Nature*, 427(6971):244–247.

Köszegi, B. (2006). Ego utility, overconfidence, and task choice. *Journal of the European Economic Association*, 4(4):673–707.

Kotlikoff, L. J. and Summers, L. H. (1981). The role of intergenerational transfers in aggregate capital accumulation. *Journal of Political Economy*, 89(4):706–732.

Kuhnen, C. M. (2015). Asymmetric learning from financial information. *The Journal of Finance*, 70(5):2029–2062.

Loomes, G. and Sugden, R. (1982). Regret theory: An alternative theory of rational choice under uncertainty. *The Economic Journal*, 92(368):805–824.

Lord, C. G., Ross, L., and Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, 37(11):2098.

Lusardi, A. and Mitchell, O. S. (2007). Baby boomer retirement security: The roles of planning, financial literacy, and housing wealth. *Journal of Monetary Economics*, 54(1):205–224.

Lusardi, A. and Mitchell, O. S. (2011). Financial literacy and planning: Implications for retirement wellbeing. NBER Working Paper Series No. 17078.

Lusardi, A. and Tufano, P. (2009). Debt literacy, financial experiences, and overindebtedness. Working Paper, National Bureau of Economic Research, Cambridge, MA.

Ma, Y., Sraer, D., and Thesmar, D. (2018). The aggregate cost of systematic forecast errors. Technical report, 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., Nagel, S., and Yan, Z. (2017). The making of hawks and doves: Inflation experiences on the fomc. Technical report, National Bureau of Economic Research.

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

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

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.

McNamara, J. M., Green, R. F., and Olsson, O. (2006). Bayes' theorem and its applications in animal behaviour. *Oikos*, 112(2):243–251.

Meara, E. R., Richards, S., and Cutler, D. M. (2008). The gap gets bigger: changes in mortality and life expectancy, by education, 1981–2000. *Health Affairs*, 27(2):350–360.

Meehl, P. E. and Rosen, A. (1955). Antecedent probability and the efficiency of psychometric signs, patterns, or cutting scores. *Psychological Bulletin*, 52(3):194–216.

Mehra, R. and Prescott, E. C. (1985). The equity premium: A puzzle. *Journal of Monetary Economics*, 15(2):145–161.

Merkle, C., Schreiber, P., and Weber, M. (2017). Framing and retirement age: The gap between willingness-to-accept and willingness-to-pay. *Economic Policy*, 32(92):757–809.

Mirowsky, J. (1999). Subjective life expectancy in the us: correspondence to actuarial estimates by age, sex and race. *Social science & Medicine*, 49(7):967–979.

Mitchell, O. S., Poterba, J. M., Warshawsky, M. J., and Brown, J. R. (1999). New evidence on the money's worth of individual annuities. *American Economic Review*, 89(5):1299–1318.

Möbius, M. M., Niederle, M., Niehaus, P., and Rosenblat, T. S. (2014). Managing self-confidence. NBER Working Paper.

Morgenstern, O. and Von Neumann, J. (1953). *Theory of games and economic behavior*. Princeton university press.

Müller, S. and Weber, M. (2014). Evaluating the rating of stiftung warentest: How good are mutual fund ratings and can they be improved? *European Financial Management*, 20(2):207–235.

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.

Nosić, A. and Weber, M. (2010). How riskily do i invest? the role of risk attitudes, risk perceptions, and overconfidence. *Decision Analysis*, 7(3):282–301.

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.

Olafsson, A. and Pagel, M. (2018). The retirement-consumption puzzle: New evidence from personal finances. NBER Working Paper Series No. 24405.

Peijnenburg, K., Nijman, T., and Werker, B. J. (2017). Health cost risk: A potential solution to the annuity puzzle. *The Economic Journal*, 127(603):1598–1625.

Phillips, L. D. and Edwards, W. (1966). Conservatism in a simple probability inference task. *Journal of Experimental Psychology*, 72(3):346–354.

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.

Pouget, S., Sauvagnat, J., and Villeneuve, S. (2017). A mind is a terrible thing to change: confirmatory bias in financial markets. *The Review of Financial Studies*, 30(6):2066–2109.

Previtero, A. (2014). Stock market returns and annuitization. *Journal of Financial Economics*, 113(2):202–214.

Puri, M. and Robinson, D. T. (2007). Optimism and economic choice. *Journal of Financial Economics*, 86(1):71–99.

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior & Organization*, 3(4):323–343.

Rabin, M. and Schrag, J. L. (1999). First impressions matter: A model of confirmatory bias. *The Quarterly Journal of Economics*, 114(1):37–82.

Reichling, F. and Smetters, K. (2015). Optimal annuitization with stochastic mortality and correlated medical costs. *American Economic Review*, 105(11):3273–3320.

Savage, L. J. (1954). *The foundations of statistics*. Wiley Online Library.

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.

Schreiber, P. and Weber, M. (2016). Time inconsistent preferences and the annuitization decision. *Journal of Economic Behavior and Organization*, 129:37–55.

Sharot, T. and Garrett, N. (2016). Forming beliefs: Why valence matters. *Trends in Cognitive Sciences*, 20(1):25–33.

Sharot, T., Kanai, R., Marston, D., Korn, C. W., Rees, G., and Dolan, R. J. (2012). Selectively altering belief formation in the human brain. *Proceedings of the National Academy of Sciences*, 109(42):17058–17062.

Sharot, T., Korn, C. W., and Dolan, R. J. (2011). How unrealistic optimism is maintained in the face of reality. *Nature Neuroscience*, 14(11):1475–1479.

Shiller, R. (1981). Do stock prices move too much to be justified by subsequent changes in dividends? *American Economic Review*, 71:421–436.

Tribe, L. H. (1971). Trial by mathematics: Precision and ritual in the legal process. *Harvard Law Review*, 84(6):1329–1393.

Troutman, C. M. and Shanteau, J. (1977). Inferences based on nondiagnostic information. *Organizational Behavior and Human Performance*, 19(1):43–55.

Tversky, A. and Fox, C. R. (1995). Weighing risk and uncertainty. *Psychological Review*, 102(2):269.

Tversky, A. and Kahneman, D. (1971). Belief in the law of small numbers. *Psychological Bulletin*, 76(2):105–110.

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

Van Rooij, M., Lusardi, A., and Alessie, R. (2011). Financial literacy and stock market participation. *Journal of Financial Economics*, 101(2):449–472.

Weber, E. U. and Johnson, E. J. (2009). Decisions under uncertainty: Psychological, economic, and neuroeconomic explanations of risk preference. In *Neuroeconomics*, pages 127–144. Elsevier.

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.

Wells, G. L. and Harvey, J. H. (1978). Naive attributors' attributions and predictions: What is informative and when is an effect an effect? *Journal of Personality and Social Psychology*, 36(5):483–490.

Yaari, M. E. (1965). Uncertain lifetime, life insurance, and the theory of the consumer. *The Review of Economic Studies*, 32(2):137–150.

Zimmermann, F. (2020). The dynamics of motivated beliefs. *American Economic Review*, 110(2):337–361.

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 *"Probabilistic Reasoning in Economic Decisions – Belief Formation, Inference Judgements, and Retirement"* 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.

Unterschrift:

Mannheim, den:

Curriculum Vitae

Pascal Kieren

Since 2017	Doctoral Student in Finance Chair of Senior Professor Weber Supervisor: Prof. Dr. Dr. h.c. Martin Weber University of Mannheim, Germany
Since 2016	Doctoral Student in Finance Center for Doctoral Studies in Business (CDSB) Graduate School of Economic and Social Sciences (GESS) University of Mannheim, Germany
2015 – 2017	Master in Business Research – Finance Track (M.Sc.) University of Mannheim, Germany
2011 – 2014	Bachelor in Business Administration (B.Sc.) Saarland University, Germany