Delegated Investing and the Role of Investment Outcomes, Quality, and Trust

Inauguraldissertation

zur Erlangung des akademischen Grades eines

Doktors der Wirtschaftswissenschaften

der Universität Mannheim

vorgelegt von

Maximilian Germann, M.Sc.

im Herbst-/Wintersemester 2018

Dekan: Professor Dr. Dieter Truxius

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

Korreferent: Professor Dr. Stefan Ruenzi

Tag der mündlichen Prüfung: 29. November 2018

Contents

List of Figures

List of Tables

For my wife and my family.

Chapter 1

Introduction

1.1 Motivation and Scope of This Dissertation

"Good savers, bad investors" (originally: *"Gute Sparer, schlechte Anleger"*) is how the German magazine *Capital* described Germans and their financial decision making in 2014.[1] This statement captures what both practitioners and academics have been repeating for years: Individuals do not invest sufficiently into the stock market. According to *Deutsche Aktieninstitut*, only 15.7% of the German population invested, directly or indirectly, into stocks in 2017.[2]

In 2008, the financial crisis led to an erosion of trust, a component key to encouraging investors to take financial risks (Guiso et al., 2008; Sapienza and Zingales, 2012). Now that trust has slowly been restored, and of course aided by years of (nominal) interest rates close to zero, the proportion of equity investors has recovered to the level prior to the financial crisis. However, there are various other reasons why investors shy away from the stock market in the first place. Aside from rational reasons such as severe capital and liquidity constraints, investors tend to focus on short-horizon risk when in fact they have long-horizon investment goals, or they do not adequately understand investment risks conveyed through descriptions and illustrations (Benartzi

[1] Editorial in September 2014, see `https://www.capital.de/wirtschaft-politik/gute-sparer-schlechte-anleger`.

[2] `https://www.dai.de/files/dai_usercontent/dokumente/studien/2018-02-19%20Aktieninstitut%20Aktionaerszahlen%202017%20Web.pdf`.

and Thaler, 1995, 1999; Weber et al., 2005; Beshears et al., 2011; Kaufmann et al., 2013).

Unfortunately, convincing oneself to invest into the stock market is only the very first step. Once invested, new “dangers” await. Shall I pick stocks? Shall I try to time the market? Shall I invest into stocks of firms in my region, stocks of firms whose business I know well? In general, the answer to all of these questions is a resounding *No!*. Many investors, however, make such simple investment mistakes (Calvet et al., 2007). Discussing these investor mistakes, their potential psychological roots, and their monetary consequences is far beyond the scope of this dissertation.[3]

Instead, this dissertation focuses on one aspect that can help investors overcome investment mistakes: Delegation of investment decisions. In the last decade, extensive research on two potential solutions to reduce investor mistakes has been produced. The first potential solution is to increase individuals’ financial sophistication or literacy. Pioneered by Annamaria Lusardi (see e.g., Lusardi and Mitchell, 2014), financial literacy has been linked to increased stock market participation (van Rooij et al., 2011), greater awareness for retirement planning (Bucher-Koenen and Lusardi, 2011), and stronger portfolio diversification (Calvet et al., 2007; Gaudecker, 2015; Guiso and Viviano, 2015). There is, however, also skepticism about the benefits of financial literacy. In a meta study of the economics of financial literacy, Fernandes et al. (2014) conclude that the importance of financial literacy may have been vastly overstated due to omitted psychological traits (and other variables).

The second potential solution – and the common theme in this dissertation – is delegation of investment decisions. Instead of investing on their own, investors can easily delegate investment decisions to another party, hereafter summarized by the term money manager. In practice, this party

[3] For an easy-to-read, non-academic handbook on investment mistakes and how they can be avoided, the interested reader is therefore referred to Weber et al. (2007).

is often either an intermediary in the form of a human investment advisor or a robo-advisor, or a fund manager.[4]

Today, the delegation industry is of substantial size. In 2017, the market for financial planning and advice in the U.S. was estimated at US$ 56bn, providing jobs for more than 215.000 employees.[5] U.S. robo-advisors, such as *Betterment* or *Wealthfront*, already report several billion US$ worth of assets under management.[6] In Germany, the largest robo-advisor, *Scalable Capital*, recently surpassed 1bn € of assets under management.[7]

The delegation industry also has major impact on investors' investment decisions. According to the *Investment Company Institute*, U.S. mutual fund investors purchased 50% of their mutual funds through investment sales force and investment professionals in 2017. For mutual funds held outside employer-sponsored retirement plans, investment sales force and investment professionals were the primary purchasing channel.[8] In Canada, 58.5% of all households owning investment funds in 2017 stated to have used financial advice. Wealthy households (> CAD$ 500.000) were especially dependent on financial advice (72.4%).[9]

Reasons for delegating investments can be manifold. Investors, for example, may lack time to manage their investments on their own. Investors may also feel relieved to shift responsibility for investments to someone else (see e.g., Chang et al., 2016). Most notably, investors may simply lack financial knowledge. If aware of this lack of knowledge, investors want to avoid

[4] There are, of course, more types of financial agents, such as brokers, financial planners, etc. In the context of this dissertation, all these agents can be thought of as money managers (see Gennaioli et al., 2015).

[5] `http://www.ibisworld.com/industry/default.aspx?indid=1316`.

[6] US$ 13.5bn for Betterment and US$ 10bn for Wealthfront as of March 2018. Data from *Google Finance*, 22 May 2018.

[7] `https://www.handelsblatt.com/finanzen/banken-versicherungen/scalable-capital-erster-robo-advisor-sammelt-mehr-als-eine-milliarde\-euro-kundengelder-ein/22611308.html`.

[8] `https://www.ici.org/pdf/per23-08.pdf`.

[9] Consultation paper 81-408 of the Canadian Security Administrators from 2017, `https://www.securities-administrators.ca/uploadedFiles/General/RulesPolicyPaper/81-408_Consultation_Paper_09-01-17_EN.pdf`.

making costly investment mistakes by delegating their investments. That is, delegation can be regarded as substitute for financial literacy. In any case, investors must have faith in both the trustworthiness of the other party and the ability of the other party to make better investment decisions than the investor herself.

To the disadvantage of investors, overwhelming evidence suggests that investment decisions made by the other party are far from optimal.[10] In case investors rely on financial advisors,[11] there are two obvious explanations: First, delegation usually creates an agency problem, as the financial advisor (agent) may have monetary incentives to not act in the best interest of the client (principal). This explanation is backed by findings in empirical (Bergstresser et al., 2009; Hackethal et al., 2012; Hoechle et al., 2017, 2018), theoretical (Inderst and Ottaviani, 2009, 2012), and experimental studies (Mullainathan et al., 2012). Second, the financial advisor herself may lack financial skill and knowledge. Thus, the financial advisor does not invest efficiently (or makes efficient recommendations), even if interests are aligned. This explanation is backed by Foerster et al. (2017), who find that financial advisors in Canada recommend portfolios akin to their own to any client, regardless of their clients' needs.

Up to date, the literature on delegation of investments focuses primarily on the parties involved. On the one hand, behavior of the party that the investment is delegated to is analyzed. On the other hand, welfare of investors delegating the investment is investigated. This dissertation is concerned with *how* investors make delegated investment decisions. Drawing from various streams of the finance, psychology, and economics literature, each chapter of

[10] To the best of the author's knowledge, only Gaudecker (2015) finds that investors who make use of advice from family, friends, or professionals, are overall better off.

[11] In order to be concise, literature on fund managers is not discussed here. This strand of literature, however, leads to a similar conclusion for individual investors: Delegating investments to fund managers does not benefit investors, either because there is no investment skill (Fama and French, 2010), or because there is fierce competition and decreasing returns to scale (Berk and Green, 2004; Pastor et al., 2015).

this dissertation identifies a particular factor that influences the delegation of investment decisions.

In chapter 2, the key question is whether investors appreciate the ex ante quality of the delegated investment decision, or the ex post outcome of the delegated investment decision. The underlying motive for this question is the *Outcome Bias* (Baron and Hershey, 1988), which describes the human tendency to judge decisions by their outcome. Because investing in the stock market is risky, quality and outcome of investment decisions can diverge.[12] If investors reward investment outcomes instead of investment quality, their behavior can distort the behavior of the money manager they delegate the investment to. The Outcome Bias can hence lead to suboptimal (i.e., too risky) investment decisions of money managers. Unlike causes studied in the delegation literature, under the Outcome Bias it is flawed judgment of *investors* that leads to suboptimal delegated investment decisions.

In chapter 3, the key question is whether trustworthiness of money managers influences investors' delegation decisions. Specifically, this chapter examines the *Money Doctors* theory by Gennaioli et al. (2015). The intuition of this well-cited[13] theory is that trust lowers the perceived riskiness of investments. That is, trustworthy money managers make investors comfortable to take financial risk by essentially holding their hand. At least two practical implications follow from the Money Doctors theory: First, it provides an explanation why money managers can charge fees for generic services. Second, and more important in the context of this dissertation, it shows that trust enables both money managers and investors to benefit from delegation.

[12] As a simple example, consider an investment into a single stock. According to normative theory, investing into a single stock is not a "good" investment decision ex ante. However, if the stock performs well it is a "good" investment decision ex post.

[13] 225 citations on *Google Scholar*, as of June 2018.

Lastly, in chapter 4 the key question is whether investors are equally willing to delegate their investments to another human or an investment algorithm. The motivation for the last chapter is twofold. On the one hand, literature on humans' interaction with algorithms is indecisive. In some domains humans are found to be averse to the use of algorithms, while in others they are found to rely on algorithms. On the other hand, few industries have been affected by growing digitalization as much as the finance industry. Hence, the fourth chapter sheds light on the presumed importance of a human touch in financial delegation.

1.2 Contribution and Results of This Dissertation

1.2.1 Outcome Bias in Financial Decision Making

Chapter 2, coauthored with Martin Weber, presents an experimental study of the Outcome Bias. The Outcome Bias refers to the human tendency to base judgments about decisions' quality on irrelevant outcome information (Baron and Hershey, 1988). It is well-documented in both experimental and empirical studies in different fields. It also seems to be present in various financial contexts. CEOs, for example, are found to be rewarded for good luck (Bertrand and Mullainathan, 2001) and punished for bad luck (Jenter and Kanaan, 2015). In the mutual fund industry, investors are found to chase returns (Sirri and Tufano, 1998), even though chasing returns does not pay off going forward (Frazzini and Lamont, 2008; Fama and French, 2010).

To study whether investors focus on investment outcomes (i.e., chase returns) or investment quality, we conducted an incentivized online experiment on Amazon Mechanical Turk (AMT). Investigating the Outcome Bias experimentally is warranted, because in real-world data it is virtually impossible to isolate the Outcome Bias from investor beliefs. If, for example, investors believe in fund manager skill, it would be rational to judge fund managers based on past performance. If investors do not believe in fund manager skill, judging fund managers based on past performance would be driven by the Outcome Bias. In an experiment, we can control the information necessary to judge investment quality.

In our experiment, participants have to choose among several investment managers. These investment managers invest either into a stochastically dominant ("good") asset, or a stochastically dominated ("bad") asset. Participants are monetarily incentivized to select an investment manager who invests into the good asset. However, returns are randomly drawn from the respective asset's payoff distribution. Hence, good investments may – ex post – yield worse outcomes than bad investments. In three randomly ordered treatments, we vary the simplicity with which the quality (i.e., the type) of the investment can be inferred. Common to all treatments, quality and outcome can be separated. In the first treatment, participants are shown the asset each investment manager invests into. In the second treatment, participants are not shown the assets of each investment manager. However, the good asset always pays a fixed amount over the bad asset, such that assets have all unique payoffs. Hence, participants can easily infer the quality of the investment from the uniqueness of outcomes. In the third treatment, assets have common payoffs with different probabilities. Again, participants are not shown the assets of each investment manager. Participants are only informed that it is randomly determined (50%) into which asset an investment manager invests, such that participants can choose an investment manager to maximize the chance of investing into the good asset. Our design therefore allows us to pin down the cognitive challenge that is most relevant for the Outcome Bias in financial decision making: Either it is the challenge of using outcomes only to infer investment quality alone, or the additional challenge of dealing with uncertainty about investment quality.

We find that 44% of all investment manager choices are outcome biased even in the first treatment. In the second and third treatment, this fraction increases to approximately 60%. In all treatments, investment manager choices that resulted in investments into the bad asset are almost exclusively driven by the Outcome Bias. Our findings suggest that individuals find it difficult to separate investment quality from investment outcomes. Specifically, individuals mistake high outcomes for good decisions, when all information that proves otherwise is readily available. Using simulations of random choices, we also rule out that results are a product of chance.

Findings from this study may also have implications for policymakers. Today, investment funds, financial advisors, and other financial service providers must inform investors that *"past performance is no reliable indicator for future performance"*. All too often, this crucial disclaimer is hidden in footnotes. In the light of chapter 2, educating investors that investment quality and investment outcomes are often not identical remains a key challenge. Therefore, it may be helpful to stress factors that affect investment quality, such as fees, more prominently.

1.2.2 Trust and Delegated Investing: A Money Doctors Experiment

Chapter 3, coauthored with Benjamin Loos and Martin Weber, presents an experimental study of trust and its role for delegated investing. Trust and its key role as "lubricant" (Arrow, 1972) of economic transactions has been documented in many studies. In particular, overall trust has been linked to stock market participation and greater risk taking (Guiso et al., 2004, 2008). More recently, Gennaioli et al. (2015) propose a theory of delegation that includes trust as its core component. Their *Money Doctors* model explains management fees as a trust premium voluntarily paid by investors. Trusting in money managers reduces investors' anxiety to make risky investments. Ceteris paribus, more trustworthy money managers can set higher fees for generic services, because investors still profit from increased participation in risky investments.

We are the first to test this theory in a laboratory experiment. Investigating this theory experimentally is warranted, because it *a)* permits a a clean quantification of trust, and *b)* allows us to measure the trust-cost relationship. Due to the interactive nature of the experiment, we conducted a laboratory experiment at the Mannheim MLab.

In the experiment, participants first play a trust game (Berg et al., 1995). We exploit variation in the amounts participants return in this game as measure of trustworthiness. In two treatments, participants assuming the role of an investor are then matched to two other participants. These two matched participants represent money managers. By providing the amount these matched participants returned in the trust game, we induce different levels of trustworthiness. In the first treatment, investors are then asked to make two separated, delegated investment decisions. In the second treatment, investors are first asked to make a delegated investment decision with one money manager, and are subsequently asked to indicate costs they are willing to pay to make the same investment decision with a second money manager. Crucially, in both treatments both money managers offer identical before-costs investment opportunities.

In summary, we find that investors take substantially more risk with more trustworthy money managers, even though these are exogenously assigned twice the costs of less trustworthy money managers. Similarly, results from the second treatment show that investors are willing to accept considerably higher costs from a more trustworthy money manager than from a less trustworthy money manager. Both, the willingness to invest more risky and the willingness to pay higher costs, are increasing in the difference in money manager trustworthiness. Furthermore, we show that our results survive if we control for alternative explanations such as biased investor beliefs. Most importantly, our study provides evidence that money managers *and* investors benefit from increased trust: Investors' profits from increased risk taking exceed the additional costs they are charged.

Most of the literature on delegated investing paints a dire picture of the usefulness of money managers. This dissertation's third chapter, on the contrary, highlights a positive aspect of delegated investing. Although investors would be best off with higher risk taking at lower costs, they are still better off with higher risk taking at higher costs, if they trust their money manager. Trust may thus well be the "substantial intangible benefit" Bergstresser et al. (2009, p.4129) suspect but cannot observe.

1.2.3 Algorithm Aversion in Financial Investing

Chapter 4, coauthored with Christoph Merkle, presents an experimental study of *Algorithm Aversion* and its role for delegated investing. The term algorithm aversion was recently coined by Dietvorst et al. (2015). It refers to the tendency to rely on human predictions or recommendations more than on those of an algorithm, even if the latter (observably) performs better. The concept of algorithm aversion, however, is not new. In several contexts, such as medical recommendations, studies have documented aversion towards algorithms (e.g., Promberger and Baron, 2006; Shaffer et al., 2013; Yeomans et al., 2017).

Little attention has been paid to algorithm aversion and its potential role in financial markets. Financial markets in particular, however, have been reshaped by increasing digitalization in recent years. Due to abundant computing power and potent exchange infrastructure, algorithmic traders now try to profit from mispricings within milliseconds. The delegation industry has also been subject to radical change. Robo-advisors, such as *Betterment* or *Wealthfront* in the U.S. or *Scalable Capital* and *Vaamo* in Germany, now compete successfully with human investment advisors. Human investment advisors, to keep their lead, can use computerized financial tools (e.g., *Finametrica*) in the advisory process. These ongoing developments warrant an investigation of algorithm aversion in a financial context.

In our experiment, participants have to delegate an investment decision to an intermediary. This intermediary is either a human fund manager or an investment algorithm. Since we must necessarily pin a human against an algorithm, we conducted a laboratory experiment at the Mannheim MLab. Using a simple market consisting of one risky stock and one riskless bond, we program the investment algorithm to maximize expected return. Human fund managers are appointed from the group of participants after having succeeded in a quiz measuring financial literacy and numeracy skills. Participants assuming the role of investors then have to indicate to whom they want to delegate their investment. To investigate investors' initial algorithm aversion, we measure the strength of investors' preference for either intermediary, and we make use of survey responses collected several weeks before the laboratory experiment.

We find no evidence for algorithm aversion. When both intermediaries charge equal fees, 56% of participants decide to invest with the algorithm in the initial choice. If fees differ, participants mostly ($> 80\%$) choose the intermediary with lower fees. There is also no strong trend in the proportions choosing either intermediary. Once investors learn about investment choices and outcomes of both intermediaries, they focus on performance. In line with Bayesian behavior, choices are strongly influenced by cumulative past performance. Critically, in their reaction to performance, participants do not discriminate between intermediaries. Specifically, they do not respond differently to mistakes by the intermediaries, rejecting the idea of trust in an algorithm eroding more quickly. Hence, we find no support for the two major predictions of algorithm aversion – general preference for human judgment and adverse response to errors by an algorithm – in the domain of financial decision making.

The fourth chapter helps clarify the literature's inconclusive stance on algorithm aversion. When the investment setting is easy to understand, investors do not discriminate between a human fund manager and an investment algorithm. Unlike in previous studies, in our experiment investors can choose a real human. Hence, finding no sign of algorithm aversion is a strong indicator that a human touch is not a necessity for delegated investing. Strong performance of the intermediary, however, is a necessity for delegated investing. Human advisors should therefore consider forming a symbiotic relationship with algorithms in finance: Investors could profit not only from powerful algorithms, but also from the trust relationship human advisors can establish.

Chapter 2

Outcome Bias in Financial Decision Making *

2.1 Introduction

Consider the following scenario: An individual investor with a basic understanding of finance, but no time at hand, wants to invest a portion of her wealth. She therefore asks her trusted financial advisor for help. Knowing about diversification, she asks her financial advisor to set up a broadly diversified portfolio. Instead, the advisor invests her client's wealth exclusively in the oil industry. In the following months, the oil price increases unexpectedly, and the oil industry significantly outperforms the broadly diversified benchmark.

In theory, the random outcome of the investment decision should be irrelevant to the quality of the investment decision (e.g., Vlek, 1984). Thus,

* Authors: Maximilian Germann and Martin Weber. Both authors are at the University of Mannheim. Maximilian Germann: L9, 1-2, 68161 Mannheim. Email: germann@bank.bwl.uni-mannheim.de. Martin Weber: L9, 1-2, 68161 Mannheim. Email: weber@bank.bwl.uni-mannheim.de. We thank seminar participants at the University of Mannheim, participants at the IMEBESS 2017 Barcelona and participants at the SEF 2017 Nice for helpful suggestions. Funding from the Graduate School of Economic and Social Sciences at the University of Mannheim (GESS), the Gesellschaft für experimentelle Wirtschaftsforschung (GfeW), and the Julius-Paul-Stiegler Gedächtnis Stiftung is gratefully acknowledged. We would like to thank Alexander Sandukovskiy for invaluable programming work. All remaining errors are our own.

the financial advisor in this scenario should be held responsible for making a wrong investment decision, as she opted for unnecessary risks ex ante. However, most people would take the (positive) outcome into account when evaluating the advisor's investment decision. As a result, the investment decision would be considered good ex post by the decision evaluator. This tendency to base judgments about decisions' quality on irrelevant outcome information was dubbed *Outcome Bias* by Baron and Hershey (1988). In this paper, we investigate whether individuals exhibit the Outcome Bias in financial decision making. In particular, we examine whether individuals are more prone to the Outcome Bias the more difficult it becomes to separate decision quality from decision outcome. To the best of our knowledge, we are the first to investigate whether and under which circumstances the Outcome Bias exists in financial decision making. For a deeper analysis of the psychological foundations of this bias, the interested reader is referred to Baron and Hershey (1988), Agrawal and Maheswaran (2005), or Savani and King (2015).

The Outcome Bias is well-documented in both experimental and empirical studies in various fields. In an experiment with officers in the Israel Defense Forces, Lipshitz (1989) shows that military decisions are rated better when their outcomes turn out favorable. Participants associate good outcomes with superior decision making processes, even though in Lipshitz' experiment they are independent of another. In a more recent study, Ratner and Herbst (2005) demonstrate that individuals switch from more profitable lotteries to less profitable lotteries after observing unfavorable outcomes, despite being able to correctly recall the probabilities of outcomes. In the study by Ratner and Herbst (2005), however, individuals are not incentivized and do not have to bear the consequences of their choices.

Several empirical studies also provide support for the Outcome Bias. Lefgren et al. (2014), for example, find that the probability of changes made to the starting lineup of NBA teams spikes after games that are lost by a small margin. Losses by a small margin, however, are least informative in a rational Bayesian framework, as they are most likely to occur by pure chance. In politics, electorates appear to punish U.S. presidents and governors for severe natural disasters (Gasper and Reeves, 2011) which are beyond the politicians' control.

In finance, there is evidence that CEOs are rewarded for good luck (Bertrand and Mullainathan, 2001), and punished for bad luck (Jenter and Kanaan, 2015). Linking CEO pay or punishment to a component of luck, however, implies that the decision maker – the CEO – is not only judged for her decision quality. More important and motivating our experiment, the Outcome Bias is closely related to the puzzling finding that investors chase fund returns. The majority of evidence indicates that chasing fund performance and paying fees to active managers does not pay off for investors (Frazzini and Lamont, 2008; Fama and French, 2010). There are two prominent lines of reasoning: First, there may simply be no fund manager skill that justifies higher costs – this is the argument brought forward by Fama and French (2010). Alternatively, there are decreasing returns to scale for capital provided to skilled mutual fund managers. Rationally, capital flows to well-performing funds. In equilibrium, when capital is provided competitively, returns are unpredictable (Berk and Green, 2004; Chen et al., 2004; Pastor et al., 2015). In the first line of reasoning, investors' judgments of the fund's investment decision should not be based on past performance and returns should not be chased. In the second line of reasoning, chasing returns may be rational, but will not pay off going forward.

Several studies suggest that investors indeed chase returns. As documented by Sirri and Tufano (1998), mutual funds with strong recent performance receive disproportionally high fund inflows. A similar performance-flow relationship is observed for private equity partnerships (Kaplan and Schoar, 2005). Heuer et al. (2017) also provide experimental evidence that even sophisticated private investors misperceive luck for skill and hence chase returns. Nonetheless, chasing fund returns is not necessarily equivalent to the Outcome Bias. Investors could believe that fund managers possessed investing skills. Under this premise it would be rational to equate good past performance with investment skill. However, if investors did not believe in investing skills, chasing fund returns would be outcome biased. Since investor beliefs are notoriously difficult to measure, it is virtually impossible to isolate the Outcome Bias in real-world data. Thus we can only observe the result that investors chase returns, but not its precise cause. We therefore investigate the Outcome Bias experimentally, allowing us to set up a clean testing environment.

In our experiment, participants have to choose among investment managers. These investment managers invest into either a good or a bad asset, where good and bad are defined by stochastic dominance. Past performances of all investment managers are provided. Since asset payoffs are randomly distributed and drawn independently, past performances can (and should only) be used to infer the quality of the investment.[1] Because investment managers are fixed to a particular asset, the inference problem becomes one of selecting the appropriate investment manager.[2] While in reality the task of evaluating fund managers' skill and its translation into an investment strategy is complex, this task is simple in our experiment: Either an investment manager follows a good strategy of investing into the good asset, or she does not.

[1] Decision quality and investment quality are essentially the same in our experiment, hence both terms can be used interchangeably.

[2] We believe that fixing investment managers to a particular strategy is reasonably close to reality. First, there are well known examples of fund managers who represent a certain strategy or style. Warren Buffett (value) in the U.S. or Klaus Kaldemorgen (mixed fund with active risk management) in Germany may serve as examples here. There is also a literature in finance that acknowledges this "style investing", see as examples Barberis and Shleifer (2003); Kumar (2009); Cronqvist et al. (2015). Second, fund managers' investment strategies are essentially a translation of their (alleged) investment skill into action. Hence, an investment strategy is necessarily linked to a fund manager, as it reflects the fund manager's best application of her investment skill.

In three randomly ordered treatments, we vary the simplicity with which the quality of the investment can be inferred. Crucially, investment quality and investment outcome – or in finance terms, skill and luck – can be separated in all treatments. Moreover, by the choice of assets and by linking assets to investment managers, we ensure that chasing high outcomes is inconsistent with rational Bayesian behavior. In the first treatment, we explicitly show the assets every investment manager invests into. Thus, it is trivial to identify a good investment. In the second treatment, both assets are constructed such that payoffs uniquely identify the underlying asset. Because the good asset always pays a fixed additional amount over the bad asset, it is state-wise dominant. Hence, a good investment can again be identified – this time from outcomes. In the third treatment we use assets that can yield identical payoffs with different probabilities, such that the good asset first-order stochastically dominates the bad asset. Instead of showing the investment managers' assets, however, we inform participants that investment managers are allocated to either asset with a probability of 0.5. Thus, a good investment can no longer be inferred with certainty, but participants can maximize the probability of making a good investment manager choice. Treatment 1 and Treatment 2 hence share the feature that investment quality can be inferred with certainty, but outcome information is only needed in Treatment 2. Similarly, Treatment 2 and Treatment 3 share the feature that investment quality can be inferred from investment outcomes, but investment quality can be inferred with certainty only in Treatment 2. We can therefore pin down which cognitive challenge is (more) relevant for the Outcome Bias in financial decision making.

We find a substantial Outcome Bias in all treatments. Even in the first treatment, in which the difference between skill and luck is made obvious, approximately 44% of all investment manager choices are outcome biased.

The proportion of outcome biased choices increases by a third in *both* the second and the third treatment. This equal increase is puzzling, since state-wise dominance used in the second treatment allows for a certain identification of investment quality, while first-order stochastic dominance used in the third treatment makes it more complicated to identify investment quality. Our findings suggest that individuals find it difficult to distinguish skill from luck. In particular, participants mistake high outcomes for good decisions, when in fact moderate outcomes are more representative of good decisions. Importantly, this observation holds regardless of the concept of stochastic dominance used. Furthermore, our findings seem to be independent of socio-demographic characteristics. To address sample selection concerns and to establish a benchmark against which our results can be measured, we simulate the experiment with random choices. Comparing observed data to randomly simulated data lends further support to our findings. Observed proportions of outcome biased choices are consistently and significantly larger than predicted by chance.

To the best of our knowledge, there is only one comparable study by König-Kersting et al. (2017). They show that investors reward agents more the better the outcomes of the delegated investment decision. In their setting, however, a good investment cannot be defined ex ante. Our experiment allows us to make such definitions. An important feature of our setting is that we ask participants for their preferred asset in any treatment. Hence, we can account for individual preferences and adjust what constitutes a good investment ex ante accordingly. Contrasting choices of individuals preferring the dominant asset with choices of individuals preferring the dominated asset does not reveal substantial differences in susceptibility to the Outcome Bias. Taken together with evidence from the analysis of simulated data, our main results thus appear to be robust to the objection that participants did not understand the experimental tasks.

Lastly, our study contributes to the list of factors found to mitigate the Outcome Bias (see e.g., Savani and King, 2015; Martin and Cushman, 2016; Sezer et al., 2016). Although the Outcome Bias is prevalent in the first treatment, it is significantly weaker than in the two treatments in which good investments need to be inferred from outcomes. Policymakers and financial institutions should therefore emphasize characteristics that are certain to influence the quality of investment decisions ex ante. Without much doubt, educating investors about the detrimental effects of fund fees and steering them towards low fee funds could be a first step towards improving investment decision quality – that is, before returns materialize (see also Beshears et al., 2011). The remainder of this paper is structured as follows: Section 2.2 gives a detailed overview of the experimental setup, how we analyze participants' choices, and the hypotheses we test. A sample description follows in section 2.3. In section 2.4, general results are presented. In section 2.5, observed and simulated data are compared and alternative explanations are considered. Section 2.6 concludes with potential implications for policymakers and financial institutions.

2.2 Experimental Design and Hypothesis

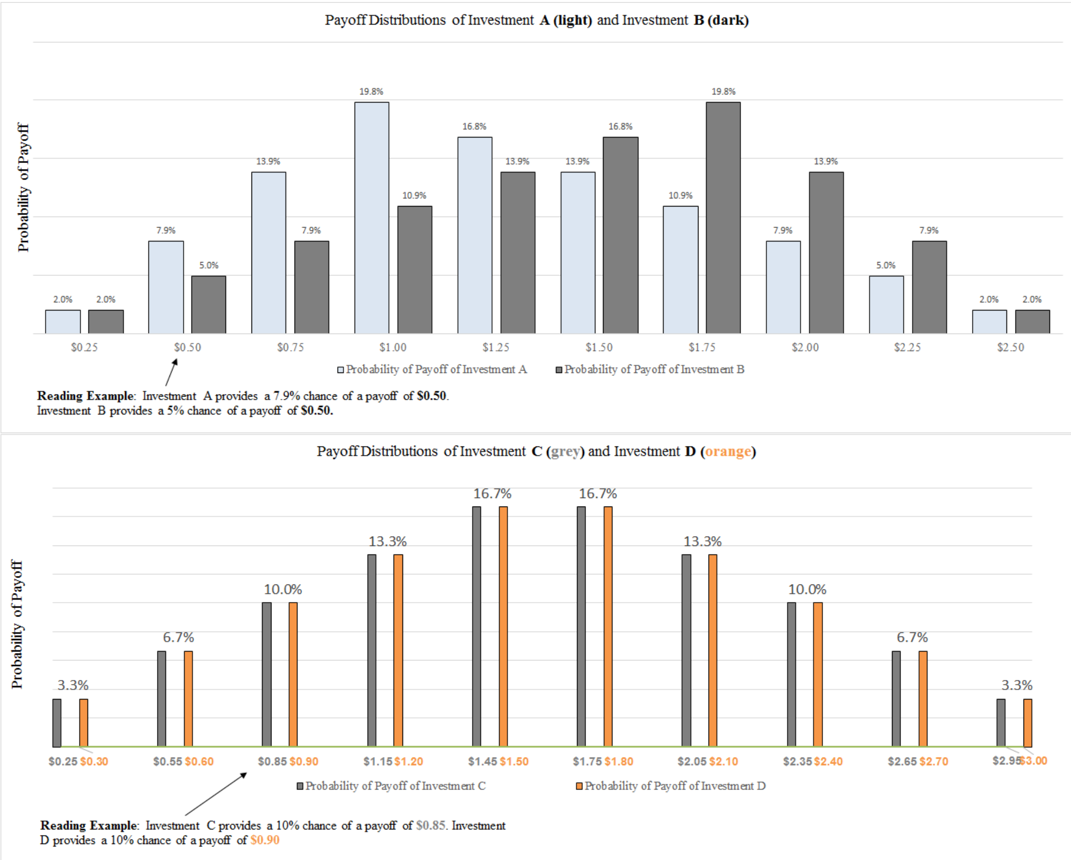
Experimental Design

Under the Outcome Bias, the focus of decision evaluators is on decision outcomes rather than decision quality. The goal of our experimental design is to present participants with the choice to either appreciate investment quality (i.e., skill) or investment outcomes (i.e., luck), while allowing for the two to be distinguished.

In our experiment, participants' task (hereafter called *manager choice*) is to choose one out of five computerized investment managers. This captures the situation investors experience in the real world, in which they have to choose from a pool of financial money managers. Investment managers invest into either a good or a bad asset. Thus all complexity from the otherwise complex task of evaluating fund manager skill is eliminated. In the context of the experiment, skill is equivalent to the strategy of investing into the good asset. Importantly, while the payoff distribution of the good asset dominates that of the bad asset, it is possible that the good asset yields lower payoffs by mere chance. By separating skill from luck, rational investors should choose an investment manager who invests into the good asset. In three randomly ordered within-subject treatments, we vary the simplicity with which skill and luck can be separated. Specifically, these three treatments allow us to pin down whether the challenge of inferring quality from outcomes alone, or only in combination with uncertainty drives the Outcome Bias. In all treatments, the task is as described above. However, characteristics of good and bad assets and the amount of information provided to participants differ across treatments. These differences are described in the following.

In Treatment 1, investment managers are first randomly allocated to either the good asset or the bad asset. These assets are called Investment B and Investment A, respectively, and are constructed such that Investment B first-order stochastically dominates Investment A. Figure 2.1 shows the distribution of the assets' discrete payoffs. A random payoff is then drawn from the allocated asset's distribution. The payoffs of all investment managers' assets are subsequently shown to participants as "Last payoffs realized" (see also Figure A.1). Crucially, in Treatment 1 we also prominently display in which asset each investment manager invests. Inferring the quality of the investment manager's investment is therefore trivial: An investment manager who was allocated to the good asset makes a good investment, while an investment manager who was allocated to the bad asset makes a bad investment.

Figure 2.1: Payoff Distributions of Investments



In Treatment 2, investment managers are again first randomly allocated to either the good or the bad asset. However, in this treatment the good asset is state-wise dominant to the bad asset. The good asset is constructed such that its payoffs in any state are $0.05 higher than payoffs of the bad asset. To make the distinction from the assets used in the first treatment clear, the assets in this treatment are labeled Investment D and Investment C, respectively. Again, a random payoff is then drawn from each investment manager's allocated asset. Payoffs of all investment managers' assets are subsequently shown to participants as "Last payoffs realized". Contrary to the first treatment, we do not provide information about the allocated assets explicitly. However, both assets have all unique payoffs. Hence, participants can easily infer the asset's quality – this time from outcomes. Treatment 2 and Treatment 1 therefore share the feature that investment quality can be inferred with certainty, but outcome information is only needed in Treatment 2. The task in the second treatment thus resembles a situation in which investors have to choose between two investment funds that follow identical strategies, but charge different costs.

Treatment 3 is identical to Treatment 1, except that the investment managers' allocated assets are not shown to participants. However, participants are informed that investment managers are allocated to either asset with a probability of 0.5. Fixing the prior belief about the allocation of assets allows participants to infer the probabilities that last realized payoffs were obtained from either the good or the bad asset. Because assets have simple, discrete distributions, the inference problem can either be solved a) graphically by comparing the ratio of bars from the payoff distributions or b) analytically by calculating probability ratios.[3] Since both assets are skewed into opposite directions, the fixed allocation probability implies that outcomes in the upper tail of the distribution become relatively more likely to stem from the

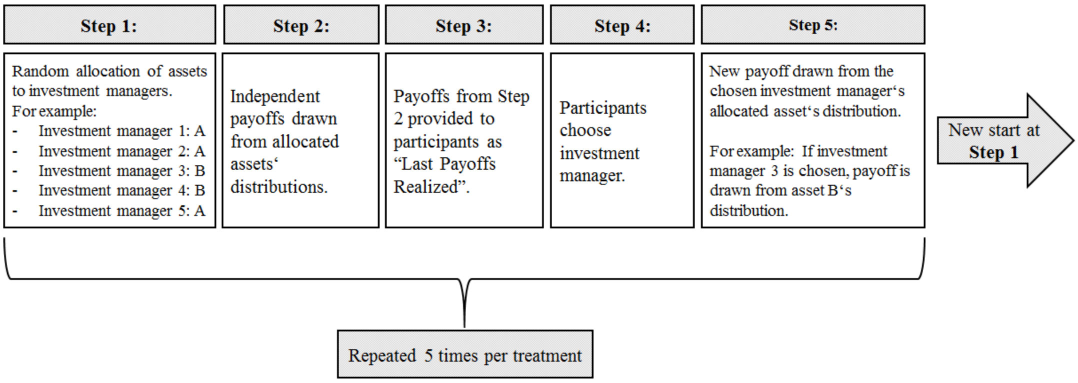
[3] To assist participants we provide a calculator built into the application.

bad asset than from the good asset.[4] Rational investors should therefore opt for an investment manager whose investment yielded a moderate outcome. In other words, chasing extreme outcomes in Treatment 3 does not maximize the probability of investing in the good asset. Treatment 3 and Treatment 2 therefore share the feature that investment quality can be inferred from investment outcomes, but investment quality can be inferred with certainty only in Treatment 2. That is, Treatment 3 adds uncertainty to the task of Bayesian updating, a condition that has proven difficult for individuals (Grether, 1992; Ouwersloot et al., 1998; Charness and Levin, 2005). The task in the third treatment thus resembles a situation in which investors have to choose between investing into a single stock and a diversified fund. The first should not be considered a good investment ex ante and is rather likely to yield extreme outcomes, while latter should be considered a good investment ex ante and is rather unlikely to yield extreme outcomes. The task in the third treatment is also loosely related to Charness and Levin (2005), in whose study high outcomes are a sign of a bad decision.

After participants choose their favored investment manager, a payoff from the investment manager's respective asset is drawn. In each treatment the experimental task is then repeated four additional times with new independent allocations and new random draws of "Last payoffs realized". Thus, we observe a total of 15 manager choices per participant. Participants' earnings are determined by a random draw of the payoff of one of the 15 tasks. Figure 2.2 illustrates the sequence of the tasks. In addition, participants play a practice round in the beginning of each treatment, which does not contribute to their total earnings. The experimental setup, and in particular the assets' distributions, is fully known to participants. Instructions and

[4] Note that the mode of Investment B is at \$1.75, while the highest outcome possible is \$2.50. Assuming that participants are correctly preferring dominant Investment B, the highest chance of ending up with Investment B is obtained for an outcome of \$1.75 ($\Pr(\$1.75|B) \div \Pr(\$1.75|A) = 1.82$), whereas this chance is strictly decreasing for higher outcomes down to $\Pr(\$2.50|B) \div \Pr(\$2.50|A) = 0.5$.

Figure 2.2: Diagram of Experimental Setup



asset distributions are shown at the bottom of the screen during the whole experiment, such that participants do not need to memorize any information. Figure A.1 in the appendix shows an example of the screen participants see in Treatment 1. Because the component of luck is fundamental to our analysis, we require participants to answer two binary questions aimed at testing their understanding of luck at the start of the experiment. The first question asks for predictability of certain payoffs, i.e., whether investment managers can promise a particular payoff. The second question asks for predictability of payoffs based on past payoffs, i.e., whether past payoffs influence the probability distribution of future payoffs. After submitting an answer, an explanation of the correct answer is shown immediately. Furthermore, we ask participants about their experience in understanding the distributions of assets shown. The experiment concludes with a short survey of socio-demographics and the extended 7-points cognitive reflection test (CRT).

Although assets are constructed such that one asset is stochastically dominant to the other, we also ask participants for their preferred asset at the beginning of each treatment. Asking for the preferred asset permits us to always identify a good investment decision ex ante: If a participant indicated a preference for Investment A in Treatment 1, choosing a manager who invests into Investment A is a good manager choice (albeit the asset *itself* is not the good asset). If, as expected, a participant preferred the dominant asset Investment B, choosing a manager who invests into Investment B is a good manager choice. Using participants' indicated asset preferences, we impose a few restrictions on the experimental task that allow us to better analyze the data. First, participants must be able to appreciate a good investment decision. Hence, for every task there must be at least one investment manager who invests into the preferred asset. Second, participants must be able to chase luck instead of skill. Thus, for every task there must be at least one investment manager who invests into the non-preferred asset. Third and last, we rule out cases in which the investment manager with the highest last realized payoff is also an investment manager who invests into the preferred asset. These cases would not allow us to distinguish whether a manager choice was rational or outcome biased and would therefore require us to collect a substantially larger sample. All restrictions are implemented by repeating the allocation of investment managers to assets and the drawing of asset payoffs until all restrictions are met. Participants are then only confronted with tasks that do not violate any of the restrictions above.

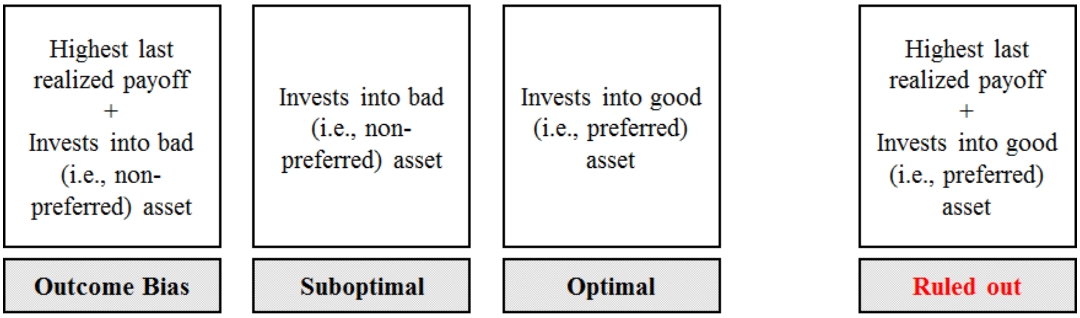
Due to the imposed restrictions, one may object that participants are not presented truly random manager choices. While technically true, we believe that this is not critical to our experiment for two reasons. First, any series of manager choices that complies with our restrictions could also occur if manager choices were truly random. Thus, a rational Bayesian agent would merely conclude that the series he was presented had a low(er) chance of occuring. In fact, imposing the first and second restriction does not change the Bayesian agent's posterior belief about the 50% - 50% allocation of investment managers to assets: Both restrictions eliminate equally likely manager choices and the a-priori probability of investment manager allocations to assets is 50%. Moreover, since participants act in isolation, they cannot learn from series that others are presented. Second, the imposed restrictions apply to all treatments equally. Since there is no reason to believe that they would affect treatments differently, the restrictions may only affect levels of the Outcome Bias. Our conclusion that individuals struggle to separate skill from luck when skill must be inferred from outcomes, regardless of the simplicity of this task, would remain unchanged. Nonetheless, we show basic results of an additional experiment if no restrictions are imposed in the appendix. Results are in line with the hypothesis described in the following section.

Hypothesis

There are two kinds of irrational manager choices that are of interest to us. First, participants can choose the investment manager with the highest last realized payoff. As outlined above, this choice is by construction not a rational choice. Since the Outcome Bias refers to the tendency to focus on decision outcomes instead of decision quality, we classify such manager choices as outcome biased. Second, participants can choose an investment manager who invests into the bad (i.e., non-preferred) asset. However, this investment manager does not necessarily have to be the investment manager with the highest last realized payoff. All such manager choices are classified as suboptimal. Consequently, outcome biased manager choices are included in the set of suboptimal manager choices. If participants choose an investment manager who invests into the good asset, their choice is rational and thus classified as neither of the above. Figure 2.3 depicts the types of manager choices and how they are classified.

Our treatments vary in the degree of simplicity with which skill and luck can be distinguished. Arguably, making a good manager choice is easiest in the first treatment – investment managers' investments are made obvious. In

Figure 2.3: Diagram of Manager Choices and Their Classification



the second treatment, investment managers' investments can also be inferred easily and with certainty. Nonetheless, making a good manager choice requires participants to use outcomes only as indicator of investment quality. Treatment 3 is designed to be most difficult. In this treatment, participants need to discount high outcomes in favor of moderate outcomes and need to compare probabilities of both the good and the bad asset distribution. Therefore, we seek to test the following hypothesis:

Hypothesis 1: The more difficult the distinction between investment quality (i.e., skill) and investment outcome (i.e., luck), the more prone participants are to making outcome biased manager choices. The proportion of outcome biased manager choices should therefore increase from Treatment 1 to Treatment 2 to Treatment 3.

2.3 Sample Description

The experiment was computerized using oTree (Chen et al., 2016). It was then posted on Amazon Mechanical Turk (AMT). We opted for AMT as it is easy to attract participants, and there is ample evidence suggesting that AMT is a valid recruiting source (e.g., Paolacci et al., 2010; Buhrmester et al., 2011; Amir et al., 2012; Goodman et al., 2013; Casler et al., 2013). We recruited

100 participants in total in January 2017. To guarantee a high quality of responses, non-U.S.-AMT workers and those with less than 5,000 completed tasks or an approval rating below 97% were excluded from the experiment.[5] The base payment for participation was US$ 0.50. The variable payment consisted of a randomly drawn realization of one of the 15 manager choices. It could range from US$ 0.25 to US$ 3.00, with an expected value between US$ 1.27 and US$ 1.65, depending on the asset chosen. With a duration of about 15 minutes, the total expected payment for participation thus lay approximately at the level of the current U.S. federal minimum wage of US$ 7.25 per hour.[6]

Study participants were predominantly male (70%). The mean age was $\approx$ 32 years, with the lowest (highest) age being 18 (70) years. Participants' educational background was higher than the national standard,[7] with 46% of participants holding at least a Bachelor degree. The majority of recruited AMT-workers had little investment experience. For example, only 25 participants indicated having ever invested in active funds. Results of the extended cognitive reflection test showed an average number of correct answers of 3.90 out of 7 with considerable variation (*SD*=2.30). The large majority of participants documented they had understood that asset payoffs were independent of one another. The two control questions on payoff independence were answered correctly by 86% and 87%, respectively, and both questions were answered correctly by 78%. Table A.1 in the appendix provides more detailed sample statistics.

[5] We address the concern that high quality workers may work too efficiently later in the paper.
[6] As of October 2016, see `https://www.dol.gov/whd/minwage/america.htm`.
[7] The fraction of U.S. citizens aged 18 or older with a Bachelor degree or higher was $\approx$ 28.4% in 2015 according to the United States Census Bureau, see `https://www.census.gov/hhes/socdemo/education/data/cps/2015/tables.html`.

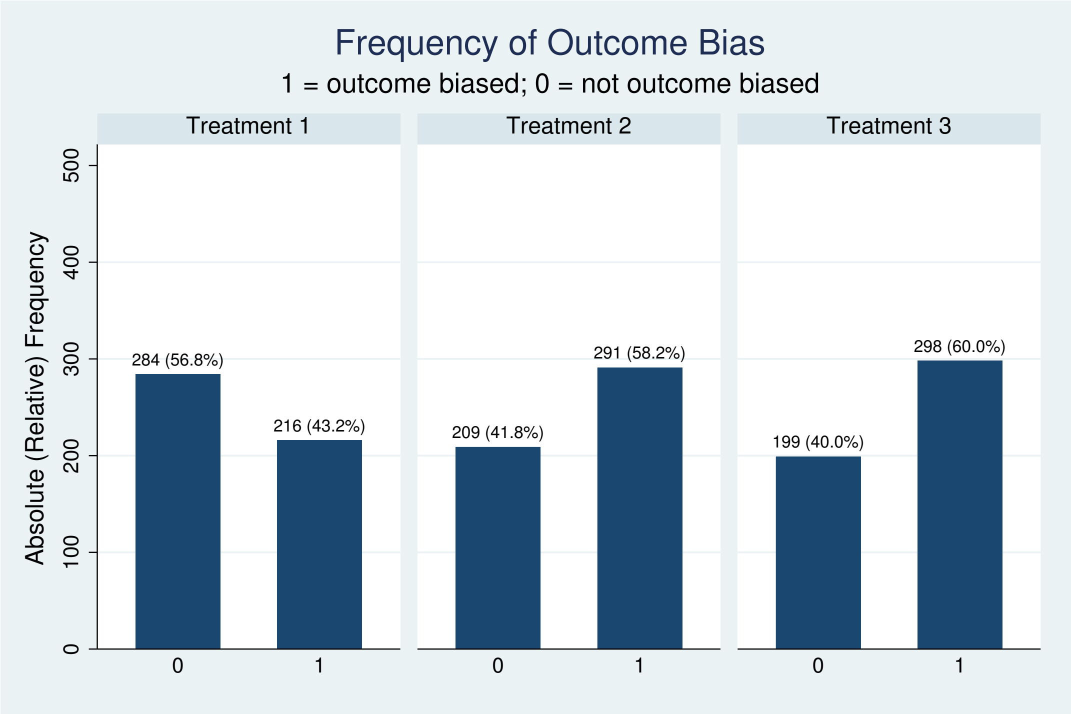
2.4 Results

In the following, we first present results from univariate analyses. We then resort to multivariate analyses to shed light on the drivers of the Outcome Bias. Our central question is whether participants seek high investment outcomes instead of good investment decisions. If they did, they would presumably choose the investment manager with the highest last realized payoff. In this the case, we classify their manager choices as outcome biased. Thus, a dummy variable is coded, taking the value of 1 (=*Outcome Bias*) if participants choose an investment manager with the highest last realized payoff. Figure 2.4 depicts the proportions of outcome biased choices by treatments. In Treatment 1, 43.2% of all manager choices are outcome biased. This fraction increases to 58.2% in Treatment 2, and 60.0% in Treatment 3.[8] To assess differences between treatments we need to account for the fact that observations are not necessarily independent, since each individual is observed multiple times per treatment. Therefore, we run simple regressions of *Outcome Bias* in a given treatment on a constant only and cluster standard errors at the individual level. We then test the constant against the proportion of outcome biased choices observed in the other treatments.[9] This approach effectively performs a test of means with adjusted test statistics. The increase from Treatment 1 to Treatment 2 and 3, respectively, is highly significant (p-values=0.000). The difference between Treatment 2 and 3, however, is small and insignificant (p-value=0.604).

[8] Due to technical difficulties, we lose three (0.6%) observations in Treatment 3.

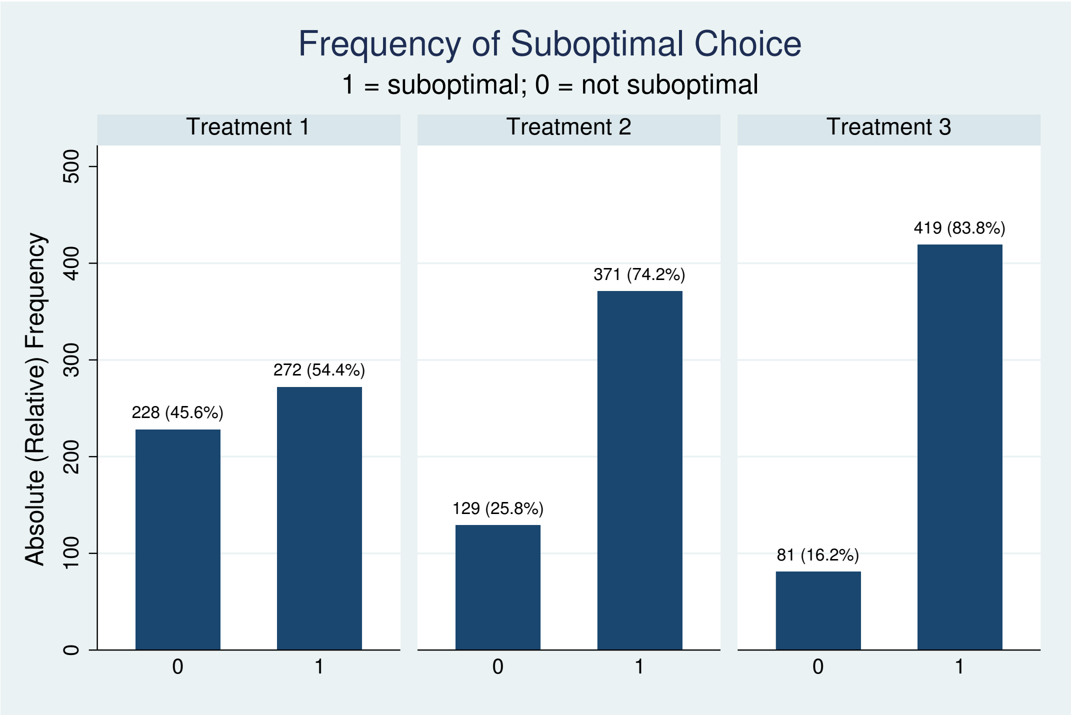
[9] Example: When regressing the dummy variable *Outcome Bias* in Treatment 1 on a constant only, this constant equals 0.432. We can then test this constant against the proportion of outcome biased choices observed in Treatment 2, 58.2%, which translates to a constant of 0.582.

Figure 2.4: Distribution of Outcome Bias



Nonetheless, manager choices need not be outcome biased to be suboptimal. Participants can, in principle, choose an investment manager who invests into the non-preferred asset, but has not obtained the highest last realized payoff (see again Figure 2.3). In this case, we classify manager choices as suboptimal. Thus, a dummy variable is again coded, taking the value of 1 (=*Suboptimal Choice*) if participants make a manager choice which does not maximize the probability of obtaining the preferred asset. Figure 2.5 shows the distribution of suboptimal choices by treatments. In Treatment 1, participants choose suboptimally more than half of the time (54.4%), resulting in not obtaining the preferred asset. The fraction is even larger for Treatment 2, in which 74.2% of manager choices are suboptimal, again resulting in not obtaining the preferred asset. The increase from Treatment 1 to Treatment 2 is also highly significant (p-value=0.000). However, the largest fraction of suboptimal choices can be observed in Treatment 3. In this treatment, in which suboptimal choices correspond to not maximizing the probability of obtaining the preferred asset, 83.8% of manager choices are suboptimal.

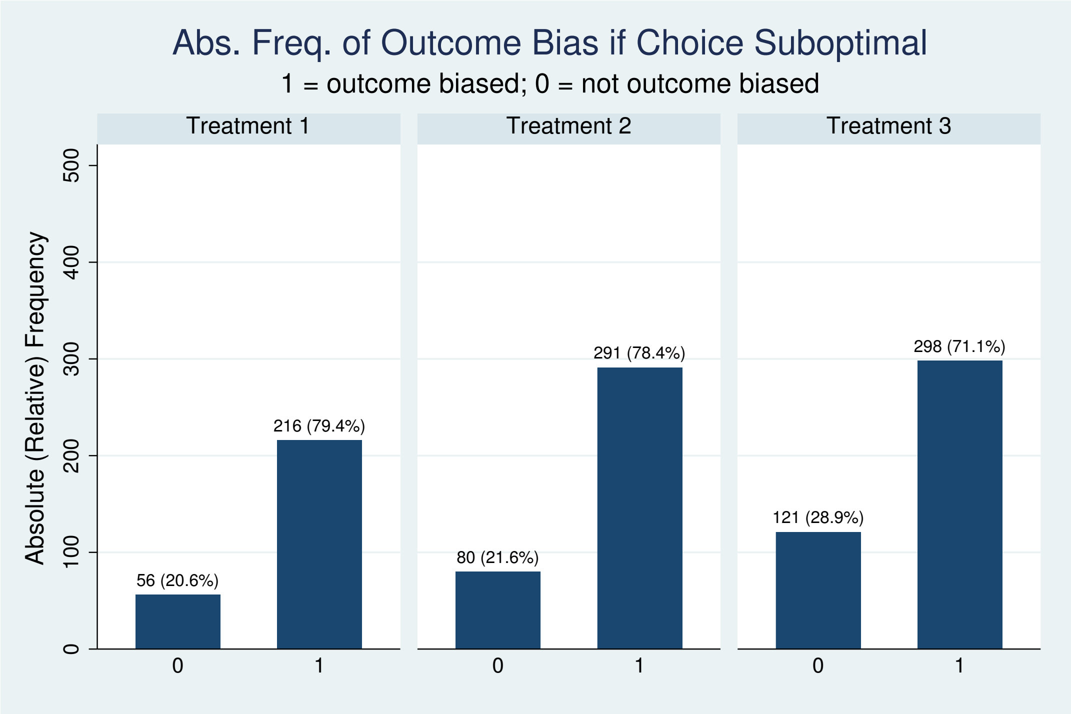
Figure 2.5: Distribution of Suboptimal Choices



Again, differences to both Treatment 1 and Treatment 2 are highly significant (p-values=0.000 and 0.000, respectively).

Suboptimal choices are not necessarily outcome biased choices as well. For this reason, we analyze whether the Outcome Bias is the primary driver of suboptimal choices. We code a dummy variable equal to 1 (=*Outcome Bias* | *Suboptimal*) if the manager choice is classified as outcome biased, conditional on the manager choice being classified as suboptimal. By construction, the total number of observations thus equals the number of suboptimal choices in each treatment. Shown in Figure 2.6, more than 7 in 10 suboptimal manager choices are suboptimal due to the Outcome Bias in all three treatments. Relative to Treatment 1 (79.4%) and Treatment 2 (78.4%, difference to Treatment 1 insignificant), participants in Treatment 3 base suboptimal choices slightly less on outcomes (71.1%). This marginal decrease could be rooted in the fact that the first two treatments differ from Treatment 3 in the way optimal manager choices are made. In Treatment 1 and 2, the assets

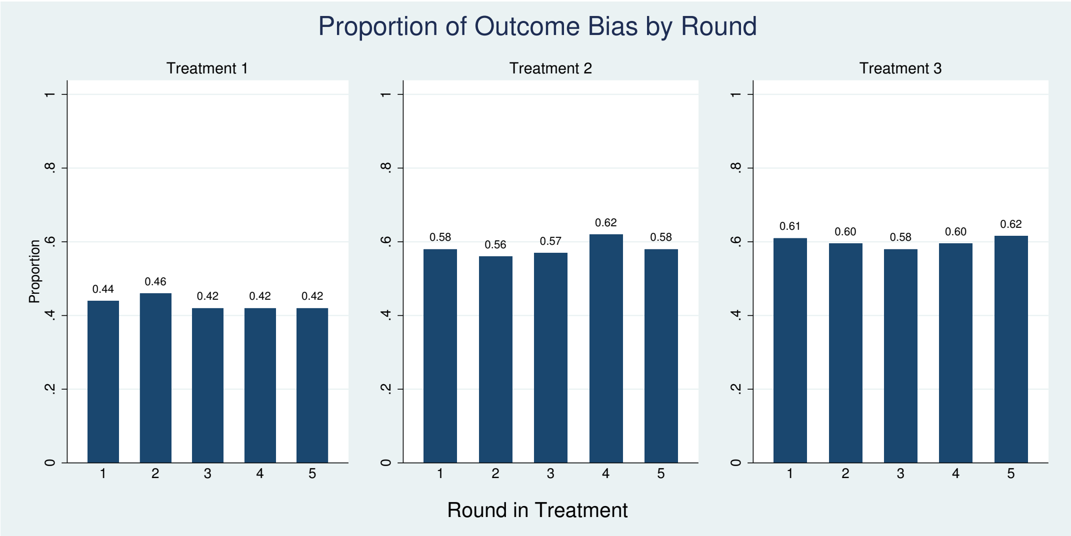
Figure 2.6: Distribution of Outcome Bias | Suboptimal



investment managers invest in can be inferred with certainty, hence an optimal manager choice is defined by the asset type itself. On the contrary, in Treatment 3 the assets investment managers invest in cannot be inferred with certainty, hence an optimal manager choice is defined by the probability that each investment manager invests into the good (i.e., preferred) asset. This probability has to be inferred from the outcomes of all investment managers and hence provides more opportunities for suboptimal manager choices.

Results of univariate analyses clearly indicate that individuals have difficulties in separating investment outcomes from investment quality. A considerable fraction of manager choices is outcome biased in the baseline treatment. This is surprising, given that the quality of the decision is known and given that participants are incentivized to appreciate good decisions. As such, our results can be interpreted as evidence of the Outcome Bias even when decision evaluators bear the consequences of their evaluations. In line with our hypothesis, the tendency to fall prey to the Outcome Bias is even

Figure 2.7: Distribution of Outcome Bias by Treatment and Round



stronger in Treatment 3. Because in this treatment Bayesian updating is required to make optimal manager choices, the result is as expected. More puzzling, however, participants are equally prone to the Outcome Bias in Treatment 2. Since, in this treatment, outcomes provide clear information about the investment's quality, we expected the Outcome Bias to be substantially weaker than in Treatment 3. Furthermore, the Outcome Bias appears to be consistent through rounds within each treatments. As depicted in Figure 2.7, there is no apparent trend along rounds in any treatment.[10] That is, participants do not avoid outcome biased choices the more familiar they become with the manager choice task. Summed up, the Outcome Bias is more pronounced once (moderate) outcomes are indicative of the investment's quality, regardless of how simple the concept of stochastic dominance is. Moreover, participants seem to make erroneous manager choices primarily because they focus on high outcomes.

To provide more details on the Outcome Bias, we now resort to multivariate analyses. To investigate whether the simplicity (i.e., the concept of stochastic dominance) with which skill and luck can be separated is the major driver of the Outcome Bias, the *Outcome Bias* dummy is used as a dependent

[10] In regressions we nonetheless control for potential learning effects over time.

variable in regressions shown in Table 2.1. Although it is a binary variable, linear regression results are reported for ease of interpretation; unreported probit models provide qualitatively similar results. All specifications include dummies for treatments (Treatment 1 as baseline), our main variables of interest. In addition, we control for participants' self-reported difficulty of interpreting the assets' distributions. We also control for the dispersion of last realized payoffs shown to participants. This is motivated by findings of Seta et al. (2015), who document that decisions are rated better the higher their outcomes *and* the lower the outcomes of alternative decisions. In our experiment, this would mean that the investment manager with the highest last realized payoff is more likely to be chosen the worse the other investment managers' outcomes turned out. Furthermore, specification (3) includes controls for age and gender, and a set of controls for education and investment experience. We also control for cognitive ability through the CRT. Given the evidence in the literature (e.g., Ratner and Herbst, 2005), higher performance on the CRT (*Cognitive Score*) is expected to reduce the Outcome Bias. All specifications also account for potential learning effects through round fixed effects. Although we make use of the panel structure of our data, coefficients for treatments do not change substantially if we control for unobserved individual characteristics through either individual fixed effects or individual random effects. We thus only report the more efficient random effects estimates.

Our regression results confirm the pattern from Figure 2.4. Across all specifications, coefficients for Treatment 2 and Treatment 3 are highly statistically significant. Looking at the full specification in column 3, we see that the baseline probability of observing outcome biased choices is approximately 47% (*Constant*=0.471). The coefficients for Treatment 2 and Treatment 3 are similar at 0.160 and 0.171, respectively ($F(1,99)=0.08$, indicates no significant difference). Hence, participants are more than one-third more likely

Table 2.1: Outcome Bias Regressions – Full Sample

This table reports regression results with *Outcome Bias* as dependent variable. *Outcome Bias* is a dummy variable equal to 1 if participants chose the investment manager with the highest last realized payoff. By construction, choosing the investment manager with the highest last realized payoff is not a rational ("good") choice, as it does not maximize the probability of investing into the preferred asset. Specifications (1) and (3) are pooled OLS regressions. (2) is a multi-level panel regression with individual- and treatment-individual random effects. *Treatment 2*, *Treatment 3*, and *Male* are dummy variables. *Age* measures participant's age in years. *Var. of manager payoffs* is calculated as the variance of all five investment managers' last realized payoffs shown to participant in the current choice task. *Cognitive Score* is the number of correct answers on a 7 question cognitive reflection test taken from Toplak et al. (2014). *Understanding Distributions* measures how difficult participants found interpreting the distributions of assets in the experiment. It is calculated from a 5-point Likert scale where 1 refers to "Very easy" and 5 refers to "Very difficult".
Standard errors clustered at the individual level in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

	(1)	(2)	(3)
	POLS	Multi-Level Panel	POLS
Treatment 3	0.170***	0.170***	0.171***
	(0.043)	(0.043)	(0.044)
Treatment 2	0.160***	0.158***	0.160***
	(0.037)	(0.037)	(0.037)
Var. of manager payoffs	-0.056	-0.044	-0.061
	(0.082)	(0.057)	(0.068)
Age			0.001
			(0.003)
Male			-0.068
			(0.064)
Cognitive Score			-0.020
			(0.018)
Understanding Distributions			0.016
			(0.032)
Constant	0.473***	0.470***	0.471**
	(0.059)	(0.058)	(0.237)
Observations	1,497	1,497	1,497
Round FE	YES	YES	YES
Education/Investment Controls			YES
Individual RE		YES	
Treatment-Individual RE		YES	
$R^2_{adjusted}$	0.014		0.093

to make outcome biased choices in these treatments than in Treatment 1. Personal characteristics do not seem to explain the Outcome Bias, as all additional control variables are insignificant at conventional levels. In turn, this suggests that mainly having to infer investment quality from (moderate) outcomes contributes to the Outcome Bias.

For the sake of completeness, we also estimate regressions with both the *Suboptimal Choice* dummy and the *Outcome Bias | Suboptimal* dummy as dependent variables. Regressions are specified exactly as before. Table 2.2 shows the results for Suboptimal Choice.[11] Again, coefficients for Treatment 2 and Treatment 3 are highly statistically significant across all specifications. Focusing on column 3, the coefficient for Treatment 3 (0.297) implies that suboptimal choices are almost 30% more likely to be observed in Treatment 3 than they are in Treatment 1. For Treatment 2, this increase is still approximately 21% (coefficient=0.207). Consistent with Figure 2.5, the difference between both coefficients is statistically significant ($F(1,99)=11.48$). Male participants appear to be slightly less prone to making suboptimal choices, as indicated by a weakly significant coefficient of -0.078. In addition, cognitive ability is negatively related to suboptimal choices, although the coefficient becomes marginally insignificant after controlling for round fixed effects.

In Table 2.3, regression results with *Outcome Bias | Suboptimal* as dependent variable are summarized. The lower proportion of Outcome Bias | Suboptimal in Treatment 3 documented in Figure 2.6 also shows in the regressions. The coefficient for Treatment 3 is negative (-0.083 in column 3) and significant across all specifications, whereas the coefficient for Treatment 2 is not significant. Additionally, all other control variables are

[11] The coefficients of the constants in column 3 is larger than in other specifications since participants whose occupation is "retired" and whose highest educational level is "university entrance qualification" are captured in the constants. Since only one participant reported her occupation as "retired" and since she made more suboptimal choices than the average participant, the effect captured in the constant is large. If other occupations are defined as baseline, the constant drops without qualitatively changing the results of the regressions.

Table 2.2: Suboptimal Choice Regressions

This table reports regression results with *Suboptimal Choice* as dependent variable. *Suboptimal Choice* is a dummy variable equal to 1 if participants chose an investment manager who should not have been chosen rationally, i.e., for whom the probability of investing into the preferred asset is not maximized. Specifications (1) and (3) are pooled OLS regressions. (2) is a multi-level panel regression with individual- and treatment-individual random effects. *Treatment 2*, *Treatment 3*, and *Male* are dummy variables. *Age* measures participant's age in years. *Var. of manager payoffs* is calculated as the variance of all five investment managers' last realized payoffs shown to participant in the current choice task. *Cognitive Score* is the number of correct answers on a 7 question cognitive reflection test taken from Toplak et al. (2014). *Understanding Distributions* measures how difficult participants found interpreting the distributions of assets in the experiment. It is calculated from a 5-point Likert scale where 1 refers to "Very easy" and 5 refers to "Very difficult".
Standard errors clustered at the individual level in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

	(1)	(2)	(3)
	POLS	Multi-Level Panel	POLS
Treatment 3	0.297***	0.295***	0.297***
	(0.039)	(0.039)	(0.040)
Treatment 2	0.208***	0.201***	0.207***
	(0.040)	(0.040)	(0.040)
Var. of manager payoffs	-0.071	-0.019	-0.060
	(0.079)	(0.063)	(0.066)
Age			-0.001
			(0.003)
Male			-0.078*
			(0.045)
Cognitive Score			-0.019
			(0.012)
Understanding Distributions			0.019
			(0.025)
Constant	0.590***	0.576***	0.974***
	(0.055)	(0.054)	(0.213)
Observations	1,500	1,500	1,500
Round FE	YES	YES	YES
Education/Investment Controls			YES
Individual RE		YES	
Treatment-Individual RE		YES	
$R^2_{adjusted}$	0.064		0.104

again insignificant. Taken together, results from the regressions indicate that suboptimal choices are made more frequently the more difficult it becomes to make optimal choices. However, the proportion of outcome biased choices out of suboptimal choices remains relatively similar across treatments.[12] Our regression results therefore point to the Outcome Bias as the primary driver of suboptimal choices. Lastly, multivariate analyses of Suboptimal Choice and Outcome Bias | Suboptimal suggest that personal characteristics do not substantially impact manager choices in our experiment.

2.5 Robustness of Results

Observed vs. Simulated Results

To corroborate our findings, we compare them to randomly simulated results. An advantage of oTree is that it enables testing of experiments through automated bots. This feature can also be modified to simulate experiments. The experiment is simulated 1,000 times (i.e., 1,000 simulated participants) with random manager choices. Thus, any investment manager will be chosen with a probability of one fifth. Such simulation allows us to establish a benchmark for Outcome Bias, Suboptimal Choice, and Outcome Bias | Suboptimal. In other words, it allows us to provide values for the outcome biased and suboptimal choices one would expect from naïve participants. We also touched on issues that potentially come with recruiting "high-quality" AMT-workers. One could argue that, for these workers, the opportunity costs of going through the experiment may outweigh the expected monetary reward. As a consequence, these workers may be inclined to select investment managers arbitrarily in order to rush through the experiment. Simulating the experiment addresses this concern

[12] We already discussed earlier why the slight drop of Outcome Bias | Suboptimal in Treatment 3 is not necessarily surprising.

Table 2.3: Outcome Bias | Suboptimal Regressions

This table reports regression results with *Outcome Bias | Suboptimal* as dependent variable. *Outcome Bias | Suboptimal* is a dummy variable equal to 1 if participants chose the investment manager with the highest last realized payoff, and equal to 0 if participants did not choose the investment manager with the highest last realized payoff but chose an investment manager who should not have been chosen rationally, i.e., for whom the probability of investing into the preferred asset is not maximized. Specifications (1) and (3) are pooled OLS regressions. (2) is a multi-level panel regression with individual- and treatment-individual random effects. *Treatment 2*, *Treatment 3*, and *Male* are dummy variables. *Age* measures participant's age in years. *Var. of manager payoffs* is calculated as the variance of all five investment managers' last realized payoffs shown to participant in the current choice task. *Cognitive Score* is the number of correct answers on a 7 question cognitive reflection test taken from Toplak et al. (2014). *Understanding Distributions* measures how difficult participants found interpreting the distributions of assets in the experiment. It is calculated from a 5-point Likert scale where 1 refers to "Very easy" and 5 refers to "Very difficult".
Standard errors clustered at the individual level in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

	(1)	(2)	(3)
	POLS	Multi-Level Panel	POLS
Treatment 3	-0.084**	-0.079**	-0.083**
	(0.038)	(0.039)	(0.038)
Treatment 2	-0.011	-0.017	-0.011
	(0.037)	(0.033)	(0.034)
Var. of manager payoffs	0.003	-0.030	-0.037
	(0.072)	(0.060)	(0.068)
Age			0.002
			(0.003)
Male			-0.007
			(0.049)
Cognitive Score			-0.008
			(0.015)
Understanding Distributions			0.002
			(0.026)
Constant	0.802***	0.785***	0.484**
	(0.060)	(0.059)	(0.211)
Observations	1,062	1,062	1,062
Round FE	YES	YES	YES
Education/Investment Controls			YES
Individual RE		YES	
Treatment-Individual RE		YES	
$R^2_{adjusted}$	-0.001		0.063

as well. Note that we do not simulate the heuristic of always choosing the investment manager with the highest last realized payoff as, by construction, this would imply simulating only outcome biased (and also suboptimal) choices.

Table 2.4 contrasts simulated with observed results. Across all treatments, random manager choices should only be classified as outcome biased in less than a one out of four cases. Evidently, the Outcome Bias is much more pronounced in the observed data. For Treatment 1, approximately twice as many manager choices are outcome biased in observed data than in simulated data (44.2% to 22.7%). For Treatment 2 and Treatment 3, manager choices are almost three times as likely to be classified as outcome biased (58.2% to 22.9% and 60.0% to 24.0%, respectively). To assess differences, we again use the regression-based approach. The differences between observed and simulated data are highly significant in all treatments (all p-values=0.000).

There are also significant differences between observed and simulated data when investigating suboptimal choices. In Treatment 1, however, participants make insignificantly fewer suboptimal manager choices in the experiment than would be predicted by chance (54.4% to 58.8%, p-value=0.261). On the contrary, participants perform substantially worse in Treatment 2 and Treatment 3. In Treatment 2, the difference between observed (74.2%) and simulated (57.6%) proportions of suboptimal choices is highly significant. Similar holds for Treatment 3, in which the number of suboptimal choices is even larger (83.8%). Although still highly significant, the difference between observed and simulated proportions (74.0%) is smaller. Note that the considerably higher proportion of simulated suboptimal choices in Treatment 3 is not surprising: While in Treatment 1 and Treatment 2 an optimal choice is determined by the initial asset allocation – the asset type can be identified with certainty – an optimal choice in Treatment 3 is determined by the last realized payoffs of all investment managers. Hence, there are fewer cases

Table 2.4: Observed vs. Simulated Data

Outcome Bias

	Outcome Bias = 0	Outcome Bias = 1	Δ p-value
Treatment 1			
Observed	284 (55.8%)	216 (44.2%)	0.000***
Simulated	3867 (77.3%)	1133 (22.7%)	
Treatment 2			
Observed	209 (41.8%)	291 (58.2%)	0.000***
Simulated	3857 (77.1%)	1143 (22.9%)	
Treatment 3			
Observed	199 (40.0%)	298 (60.0%)	0.000***
Simulated	3798 (76.0%)	1180 (24.0%)	

Suboptimal Choice

	Suboptimal Choice = 0	Suboptimal Choice = 1	Δ p-value
Treatment 1			
Observed	228 (45.6%)	272 (54.4%)	0.261
Simulated	2062 (41.2%)	2938 (58.8%)	
Treatment 2			
Observed	129 (25.8%)	371 (74.2%)	0.000***
Simulated	2118 (42.4%)	2882 (57.6%)	
Treatment 3			
Observed	81 (16.2%)	419 (83.8%)	0.000***
Simulated	1299 (26.0%)	3701 (74.0%)	

Outcome Bias | Suboptimal

	Outcome Bias | Suboptimal = 0	Outcome Bias | Suboptimal = 1	Δ p-value
Treatment 1			
Observed	56 (20.6%)	216 (79.4%)	0.000***
Simulated	1805 (61.4%)	1133 (38.6%)	
Treatment 2			
Observed	80 (21.6%)	291 (78.4%)	0.000***
Simulated	1739 (60.3%)	1143 (39.7%)	
Treatment 3			
Observed	121 (28.9%)	298 (71.1%)	0.000***
Simulated	2521 (68.1%)	1180 (31.9%)	

in which more than one investment manager maximizes the probability of obtaining the preferred asset.

Lastly, we can compare how often suboptimal choices are driven by the Outcome Bias. In every treatment, observed suboptimal choices are about 40 percentage points more likely to be outcome biased than simulated suboptimal choices. Due to the definition of suboptimal choices in Treatment 3, the simulated proportion (31.9%) of Outcome Bias | Suboptimal is lower than in the other treatments (38.6% and 39.7%). Nonetheless, all differences between observed and simulated data are again highly significant. Simulated data therefore support our findings. We present clear evidence that participants are considerably more prone to the Outcome Bias than would be expected by chance. Even in our baseline treatment, in which manager choices are least biased, levels of the Outcome Bias (and hence Outcome Bias | Suboptimal) are above what can be accounted for by randomness.

Are Different Manager Choices Comparable?

Two key features of our experiment are the random allocation of investment managers to assets and the random draws of last realized payoffs. Due to this randomness, circumstances under which manager choices are made need not be identical. Taking Treatment 1 as an example, it might be that only one investment manager is allocated to the preferred asset, while all four others are allocated to the non-preferred asset. In this case, there is only one optimal choice to make. It might, however, also be the case that four investment managers are allocated to the preferred asset, while only one investment manager is allocated to the non-preferred asset. In this case, there are four optimal choices, but only one suboptimal choice. Furthermore, investment managers may obtain identical last realized payoffs. Hence, two or more investment managers may tie for the highest last realized payoff.

Table 2.5: Outcome Bias Regression – Number of Wrong Managers and Ties of Highest Last Realized Payoff

This table reports regression results with *Outcome Bias* as dependent variable. *Outcome Bias* is a dummy variable equal to 1 if participants chose the investment manager with the highest last realized payoff. By construction, choosing the investment manager with the highest last realized payoff is not a rational ("good") choice, as it does not maximize the probability of investing into the preferred asset. *Num. of wrong managers = X* are dummy variables indicating the number of investment managers who, for a given manager choice, would have been classified as suboptimal choice. By construction, this number is random randomly distributed between 1 and 4 for any manager choice. *Ties of highest last realized payoff = X* are dummy variables indicating the number of investment managers whose last realized payoff tied for the highest last realized payoff of all five investment managers. By construction, this number is randomly distributed between 1 and 4 for any manager choice. *Treatment 2*, *Treatment 3*, and *Male* are dummy variables. *Age* measures participant's age in years. *Var. of manager payoffs* is calculated as the variance of all five investment managers' last realized payoffs shown to participant in the current choice task. *Cognitive Score* is the number of correct answers on a 7 question cognitive reflection test taken from Toplak et al. (2014). *Understanding Distributions* measures how difficult participants found interpreting the distributions of assets in the experiment. It is calculated from a 5-point Likert scale where 1 refers to "Very easy" and 5 refers to "Very difficult".
Standard errors clustered at the individual level in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	POLS Treatment 1	POLS Treatment 2	POLS Treatment 3	POLS All Treatments	Panel All Treatments	Panel All Treatments	Panel All Treatments	Panel All Treatments
Treatment 3						0.164***		0.163***
						(0.045)		(0.045)
Treatment 2						0.151***		0.152***
						(0.038)		(0.08)
Num. of wrong managers = 2	0.055	-0.089	-0.030	-0.030	0.019	-0.036	0.021	-0.036
	(0.105)	(0.102)	(0.376)	(0.068)	(0.057)	(0.065)	(0.057)	(0.065)
Num. of wrong managers = 3	-0.0160	-0.083	-0.033	-0.045	0.031	-0.056	0.032	-0.056
	(0.108)	(0.095)	(0.349)	(0.068)	(0.060)	(0.065)	(0.060)	(0.065)
Num. of wrong managers = 4	-0.015	-0.065	0.070	0.0130	0.077	-0.030	0.078	-0.030
	(0.109)	(0.100)	(0.344)	(0.068)	(0.061)	(0.067)	(0.061)	(0.067)
Ties of highest last realized payoff = 2	0.083	0.139**	0.169***	0.131***	0.084**	0.106***	0.085**	0.105***
	(0.079)	(0.059)	(0.061)	(0.039)	(0.036)	(0.038)	(0.036)	(0.038)
Ties of highest last realized payoff = 3	0.201	0.303**	0.002	0.195*	0.237***	0.238**	0.239***	0.238**
	(0.227)	(0.122)	(0.222)	(0.105)	(0.086)	(0.102)	(0.086)	(0.102)
Ties of highest last realized payoff = 4							-0.153	-0.104
							(0.313)	(0.245)
Var. of manager payoffs						-0.042		-0.042
						(0.067)		(0.067)
Age						0.002		0.002
						(0.003)		(0.003)
Male						-0.068		-0.067
						(0.064)		(0.064)
Cognitive Score						-0.020		-0.020
						(0.018)		(0.018)
Understanding Distributions						0.016		0.016
						(0.032)		(0.032)
Constant	0.441***	0.670***	0.519	0.556***	0.498***	0.502**	0.495***	0.502**
	(0.115)	(0.097)	(0.347)	(0.072)	(0.068)	(0.241)	(0.068)	(0.241)
Observations	500	499	496	1,495	1,495	1,495	1,497	1,497
Treatment-Round FE	YES	YES	YES	YES	YES		YES	
Round FE						YES		YES
Education/Investment Controls						YES		YES
Individual RE					YES		YES	
$R^2_{adjusted}$	-0.010	0.002	0.001	0.006		0.096		0.096

We address these potential issues in Table 2.5. In particular, we estimate regressions with *Outcome Bias* as a dependent variable but control for both the number of investment managers who are a suboptimal choice and the number of investment managers who tied for the highest last realized payoff.[13] The constant captures the case in which only one investment manager obtained the highest last realized payoff. In all specifications except (7) and (8), we leave out the two observations for which four investment managers tied for highest last realized payoff.[14] We first split the sample by treatments in columns 1 to 3. The number of potentially suboptimal investment managers does not impact the Outcome Bias. None of the coefficients is significant in any of the treatments. However, the number of ties for highest last realized payoff is positively correlated with the dependent variable. In Treatment 2, it increases by 13.9% if two managers tie and by 30.3% if three managers tie. In Treatment 3, if two managers tie for highest historical payoff, the probability of the Outcome Bias increases by 16.9%. Once treatments are pooled, the dummy for two ties for highest last realized payoff still predicts a highly significant 13.1% increase in the probability of observing an outcome biased manager choice. The coefficient for three ties is reduced to 0.195 (19.5%) but remains weakly significant. Column 8 presents a fully specified regression. Specifically, it includes dummies for treatments and socio-demographic controls. Importantly, treatment dummies remain highly significant even after the inclusion of additional controls. Coefficients of 0.152 and 0.163 for Treatment 2 and Treatment 3, respectively, indicate that the Outcome Bias is approximately one third more likely to be observed in these treatments than in the baseline treatment. Socio-demographic controls are

[13] Since there is always one optimal manager choice, the case of five suboptimal manager choices or five ties for highest last realized payoff cannot occur.

[14] The surprisingly negative coefficient for *Ties of highest last realized payoff* $= 4$ is hence only estimated from two observations.

not significant. Nonetheless, dummies for two and three investment managers tying for highest last realized payoff are significant at the 1% and 5% level, respectively. While the coefficient of former dummy is smaller (0.105) than those for Treatment 2 and Treatment 3, the coefficient of latter is larger (0.238).

At a first glance, the sizable effect of the number of ties of highest last realized payoff appears difficult to reconcile with the idea of the Outcome Bias. Individuals should be rather insensitive to the number of ties of highest last realized payoff. At a second glance this effect can be reconciled with the Outcome Bias: The more investment managers tie for highest last realized payoff, the more tempting it may become to believe that high outcomes must stem from good investments. Thus, the more investment managers tie, the more likely participants become to choose one of these managers.

However, the regression coefficients for number of ties for highest last realized payoff may just reflect the mechanical increase in outcome biased choices that would be consistent with random choices. In other words, the more options there are for randomly choosing participants to make outcome biased choices, the more likely it would be to also observe the Outcome Bias. This objection is not backed by the data. To provide more details on this issue, we disaggregate our data further. Tables 2.6 to 2.8 show distributions of the Outcome Bias by treatments. All observations are split by the number of ties for highest last realized payoff (top to bottom) and the number of potentially irrational manager choices (left to right). Values expected from random manager choices (as in the simulations) are shown in italics. Focusing on the cases with one (most likely case) and two (second-most likely case) ties for highest last realized payoff, we observe that the Outcome Bias is more pronounced than expected by random choices. This observation holds true for any treatment. In Treatment 1 the Outcome Bias is observed in 42.3% of the cases in which there is only one investment manager with the highest last realized payoff. Randomly simulated choice making, however, would only account for 20.0% of outcome biased choices in this scenario. The fraction of outcome biased choices also increases to 50.0% when two managers tie, thereby remaining above the 40.0%-fraction predicted by random choices. Expectedly, in Treatment 2 and Treatment 3 outcome biased choices are also more frequent in the observed data than in the simulated data. If there is only one highest last realized payoff, 56.0% are outcome biased choices in Treatment 2 and 58.1% are outcome biased choices in Treatment 3. When there are two highest last realized payoffs, these proportions increase to 68.7% and 75.9%, respectively. Due to small numbers of observations for most specific cases,[15] we refrain from reporting p-values for all differences. However, for all treatments the difference between observed and simulated data is highly significant (p-values=0.000) in the case with only a single highest historical payoff. Hence, regression results are put into perspective: The tendency to make outcome biased choices is positively affected by the number of potentially outcome biased choices and remains above what can be expected from random choices.

Subsample: Understanding Questions

Understanding that past payoffs do not predict future payoffs is crucial in our experiment. If it was not clear to participants that past outcomes do not predict future outcomes, choosing an investment manager with the highest last realized payoff could be perfectly rational. To check this objection, we restrict our sample to those participants who answered both understanding questions correctly. The sample then shrinks moderately to 1,167 observations (or 78 participants). Table 2.9 summarizes regression results with *Outcome Bias* as dependent variable. For ease of comparison of treatment effects,

[15] There is, for example, only one observation with *Number of Ties for Highest Last Realized Payoff* = *3* and *Wrong Manager* = *4* in Treatment 1.

Table 2.6: Outcome Bias by Number of Wrong Managers and Ties of Highest Last Realized Payoff – Treatment 1

This table reports the distribution of outcome-biased choices by the number of investment managers who, for a given manager choice, would have been classified as suboptimal choice (Wrong Managers = X) and by the number of investment managers who tied for the highest last realized payoff (Number of Ties for Highest Last Realized Payoff = X). If only one investment manager obtained the highest last realized payoff, the number of ties for highest last realized payoff is counted as 1. Since in any choice task there is at least one investment manager who should be chosen rationally and because the investment manager with the highest last realized payoff should not be chosen rationally, the maximum (minimum) number of wrong managers is 4 (1).

Treatment 1	Number of Ties for Highest Last Realized Payoff = 1				
	Wrong Managers = 1	Wrong Managers = 2	Wrong Managers = 3	Wrong Managers = 4	
Chance of random choice to be outcome-biased:	*20%*	*20%*	*20%*	*20%*	
Outcome Bias = 0	14 (58.3%)	64 (51.6%)	110 (59.5%)	71 (61.2%)	259 (57.7%)
Outcome Bias = 1	10 (41.7%)	60 (48.4%)	75 (40.5%)	45 (38.8%)	190 (42.3%)
Total	24	124	185	116	449
	Number of Ties for Highest Last Realized Payoff = 2				
Chance of random choice to be outcome-biased:	*40%*	*40%*	*40%*	*40%*	
Outcome Bias = 0	n.a.	4 (57.1%)	8 (53.3%)	11 (45.8%)	23 (50.0%)
Outcome Bias = 1	n.a.	3 (42.9%)	7 (46.7%)	13 (54.2%)	23 (50.0%)
Total	n.a.	7	15	24	46
	Number of Ties for Highest Last Realized Payoff = 3				
Chance of random choice to be outcome-biased:	*60%*	*60%*	*60%*	*60%*	
Outcome Bias = 0	n.a.	n.a.	2 (50.0%)	0 (0.0%)	2 (40.0%)
Outcome Bias = 1	n.a.	n.a.	2 (50.0%)	1 (100.0%)	3 (60.0%)
Total	n.a.	n.a.	4	1	5
	Number of Ties for Highest Last Realized Payoff = 4				
Chance of random choice to be outcome-biased:	*80%*	*80%*	*80%*	*80%*	
Outcome Bias = 0	n.a.	n.a.	n.a.	n.a.	n.a.
Outcome Bias = 1	n.a.	n.a.	n.a.	n.a.	n.a.
Total	n.a.	n.a.	n.a.	n.a.	n.a.

Table 2.7: Outcome Bias by Number of Wrong Managers and Ties of Highest Last Realized Payoff – Treatment 2

This table reports the distribution of outcome-biased choices by the number of investment managers who, for a given manager choice, would have been classified as suboptimal choice (Wrong Managers = X) and by the number of investment managers who tied for the highest last realized payoff (Number of Ties for Highest Last Realized Payoff = X). If only one investment manager obtained the highest last realized payoff, the number of ties for highest last realized payoff is counted as 1. Since in any choice task there is at least one investment manager who should be chosen rationally and because the investment manager with the highest last realized payoff should not be chosen rationally, the maximum (minimum) number of wrong managers is 4 (1).

Treatment 2	Number of Ties for Highest Last Realized Payoff = 1				
	Wrong Managers = 1	Wrong Managers = 2	Wrong Managers = 3	Wrong Managers = 4	
Chance of random choice to be outcome-biased:	*20%*	*20%*	*20%*	*20%*	
Outcome Bias = 0	11 (36.7%)	49 (45.4%)	81 (45.5%)	47 (42.3%)	188 (44.0%)
Outcome Bias = 1	19 (63.3%)	59 (54.6%)	97 (54.5%)	64 (57.7%)	239 (56.0%)
Total	30	108	178	111	427
	Number of Ties for Highest Last Realized Payoff = 2				
Chance of random choice to be outcome-biased:	*40%*	*40%*	*40%*	*40%*	
Outcome Bias = 0	n.a.	3 (30.0%)	8 (29.6%)	9 (33.3%)	20 (31.3%)
Outcome Bias = 1	n.a.	7 (70.0%)	19 (70.4%)	18 (66.7%)	44 (68.7%)
Total	n.a.	10	27 6	27	64
	Number of Ties for Highest Last Realized Payoff = 3				
Chance of random choice to be outcome-biased:	*60%*	*60%*	*60%*	*60%*	
Outcome Bias = 0	n.a.	n.a.	0 (0.0%)	1 (16.7%)	1 (12.5%)
Outcome Bias = 1	n.a.	n.a.	2 (100.0%)	5 (83.3%)	7 (87.5%)
Total	n.a.	n.a.	2	6	8
	Number of Ties for Highest Last Realized Payoff = 4				
Chance of random choice to be outcome-biased:	*80%*	*80%*	*80%*	*80%*	
Outcome Bias = 0	n.a.	n.a.	n.a.	0 (0.0%)	0 (0.0%)
Outcome Bias = 1	n.a.	n.a.	n.a.	1 (100.0%)	1 (100.0%)
Total	n.a.	n.a.	n.a.	1	1

Table 2.8: Outcome Bias by Number of Wrong Managers and Ties of Highest Last Realized Payoff – Treatment 3

This table reports the distribution of outcome-biased choices by the number of investment managers who, for a given manager choice, would have been classified as suboptimal choice (Wrong Managers = X) and by the number of investment managers who tied for the highest last realized payoff (Number of Ties for Highest Last Realized Payoff = X). If only one investment manager obtained the highest last realized payoff, the number of ties for highest last realized payoff is counted as 1. Since in any choice task there is at least one investment manager who should be chosen rationally and because the investment manager with the highest last realized payoff should not be chosen rationally, the maximum (minimum) number of wrong managers is 4 (1).

Treatment 3	Number of Ties for Highest Last Realized Payoff = 1				
	Wrong Managers = 1	Wrong Managers = 2	Wrong Managers = 3	Wrong Managers = 4	
Chance of random choice to be outcome-biased:	*20%*	*20%*	*20%*	*20%*	
Outcome Bias = 0	1 (50.0%)	10 (52.6%)	43 (43.9%)	129 (40.6%)	183 (41.9%)
Outcome Bias = 1	1 (50.0%)	9 (47.4%)	55 (56.1%)	189 (59.4%)	254 (58.1%)
Total	2	19	98	318	437
	Number of Ties for Highest Last Realized Payoff = 2				
Chance of random choice to be outcome-biased:	*40%*	*40%*	*40%*	*40%*	
Outcome Bias = 0	n.a.	0 (0.0%)	3 (50.0%)	10 (21.3%)	13 (24.1%)
Outcome Bias = 1	n.a.	1 (100.0%)	3 (50.0%)	37 (78.7%)	41 (75.9%)
Total	n.a.	1	6	47	54
			6		
	Number of Ties for Highest Last Realized Payoff = 3				
Chance of random choice to be outcome-biased:	*60%*	*60%*	*60%*	*60%*	
Outcome Bias = 0	n.a.	n.a.	0 (0.0%)	2 (50.0%)	2 (40.0%)
Outcome Bias = 1	n.a.	n.a.	1 (100.0%)	2 (50.0%)	3 (60.0%)
Total	n.a.	n.a.	1	4	5
	Number of Ties for Highest Last Realized Payoff = 4				
Chance of random choice to be outcome-biased:	*80%*	*80%*	*80%*	*80%*	
Outcome Bias = 0	n.a.	n.a.	n.a.	1 (100.0%)	1 (100.0%)
Outcome Bias = 1	n.a.	n.a.	n.a.	0 (0.0%)	0 (0.0%)
Total	n.a.	n.a.	n.a.	1	1

Table 2.9: Outcome Bias Regression – Subsamples

This table reports regression results with *Outcome Bias* as dependent variable. *Outcome Bias* is a dummy variable equal to 1 if participants chose the investment manager with the highest last realized payoff. By construction, choosing the investment manager with the highest last realized payoff is not a rational ("good") choice, as it does not maximize the probability of investing into the preferred asset. All specifications are identical to specification (2) of Table 2.1. (1) restricts the sample to those participants who answered both understanding questions correctly. (2) restricts the sample to those participants who chose the dominant asset in all treatments. (3) restricts the sample to those participants who answered both understanding questions correctly and chose the dominant asset in all treatments. *Treatment 2* and *Treatment 3*. *Var. of manager payoffs* is calculated as the variance of all five investment managers' last realized payoffs shown to participant in the current choice task.
Standard errors clustered at the individual level in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

	(1)	(2)	(3)
	Multi-Level Panel only correct answers	Multi-Level Panel only dominant asset	Multi-Level Panel only correct answers and dominant asset
Treatment 3	0.230***	0.320***	0.346***
	(0.049)	(0.063)	(0.067)
Treatment 2	0.194***	0.236***	0.245***
	(0.043)	(0.055)	(0.060)
Var. of manager payoffs	-0.063	-0.053	-0.044
	(0.056)	(0.090)	(0.096)
Constant	0.401***	0.361***	0.326***
	(0.067)	(0.087)	(0.092)
Observations	1,167	674	614
Round FE	YES	YES	YES
Individual RE	YES	YES	YES
Treatment-Individual RE	YES	YES	YES

only random effects regressions are reported. Shown in column 1, findings do not change substantially. If anything, coefficients for Treatment 2 and Treatment 3 increase slightly to 0.194 and 0.230. Results therefore indicate that the impact of requiring good investments to be inferred from moderate outcomes is marginally stronger in the restricted sample than in the full sample.

Subsample: Preferred Asset

We constructed assets such that in any treatment one asset is dominant. In Treatment 1 and 3 assets were characterized by first-order stochastic dominance, in Treatment 2 they were characterized by state-wise dominance. Nonetheless we ask participants for their preferred asset at the beginning of each treatment. In theory, participants should prefer the dominant asset

Table 2.10: Asset Choices by Treatment (in %)

Asset Choice in Treatment 1	
Investment A *(dominated)*	29
Investment B *(dominant)*	71
Asset Choice in Treatment 2	
Investment C *(dominated)*	30
Investment D *(dominant)*	70
Asset Choice in Treatment 3	
Investment A *(dominated)*	23
Investment B *(dominant)*	77

in any treatment. Indeed, shown in Table 2.10, a large majority of participants prefers the dominant asset in a given treatment. 71 and 77 participants prefer Investment B in Treatment 1 and Treatment 3, respectively, and 70 participants prefer Investment D in Treatment 2. Pre-tests with various formats of displaying payoff distributions conducted on Amazon Mechanical Turk showed that at least 30% of participants always preferred the dominated asset.[16] That is, a certain level of "noisy", unexpected asset preferences also remains for the experimental design used here. However, in an unreported regression using participants' characteristics, only cognitive score is positively and significantly correlated with a choice for the dominant asset. Furthermore, the overall stronger preference for Investment B (D) suggests that ordering or alphabetic biases are not tampering with our observations. In short, participants seem to prefer Investment B (D) due to its dominance property.

Nonetheless, only 45 participants prefer the dominant asset in *all* three treatments. To check if our previous results on the Outcome Bias are robust, we condition on those participants who indicated preferences expected from a rational individual. The subsample thus shrinks to 674 observations. The

[16] Data available upon request.

random effects regression is presented in column 2 of Table 2.9. Contrary to our expectations, the coefficients for Treatment 3 (0.320) and Treatment 2 (0.236) are larger than in the full sample. Both coefficients remain significant at the 1%-level. Lastly, column 3 shows regression result for a subsample restricted to both only correct answers to independence questions and rational preferences for assets. Coefficients remain virtually similar to column 2. That is, we find that even those participants best equipped to make good choices because they *a)* understand independence and *b)* prefer the dominant asset are prone to the Outcome Bias. The lower constant in column 3 (compare to column 2 in Table 2.1) suggests that the tendency to make outcome biased choices in our baseline treatment is slightly lower for the subgroup than for the full sample. However, the subgroup is just as likely as the full sample to make outcome biased choices once moderate outcomes correspond to good investment quality.[17]

To comply with the approach from the previous subsection, we again disaggregate manager choices and test differences using the regression-based approach. Manager choices split by preferred asset and treatment are shown in Table 2.11. Regarding the Outcome Bias in Treatment 1, participants preferring Investment B make significantly fewer (39.7% to 51.7%) outcome biased choices than their counterparts. For Treatment 2 and Treatment 3 proportions of outcome biased choices are not significantly different between participants preferring either the dominant or the dominated asset. For suboptimal choices, differences between both groups of participants are generally larger. In Treatment 1, 67.6% of choices are suboptimal for participants preferring Investment A, whereas only 49.0% are suboptimal for those preferring Investment B (p-value=0.000). In Treatment 2, the difference is not

[17] Compare for example the predicted probability of observing outcome biased choices by adding the constant and a treatment dummy in column 2 of Table 2.1 to the predicted probability of observing outcome biased choices by adding the constant and a treatment dummy in column 3 of Table 2.9.

Table 2.11: Manager Choices by Treatment and Preferred Asset

Suboptimal Choice by Preferred Asset

	Suboptimal Choice = 0	Suboptimal Choice = 1	Δ p-value
Treatment 1			
Preferred Asset = A	47 (32.4%)	98 (67.6%)	0.005***
Preferred Asset = B	181 (51.0%)	174 (49.0%)	
Treatment 2			
Preferred Asset = C	35 (23.3%)	115 (76.7%)	0.455
Preferred Asset = D	94 (26.9%)	256 (73.1%)	
Treatment 3			
Preferred Asset = A	11 (9.6%)	104 (90.4%)	0.010**
Preferred Asset = B	70 (18.2%)	315 (81.8%)	

Outcome Bias by Preferred Asset

	Outcome Bias = 0	Outcome Bias = 1	Δ p-value
Treatment 1			
Preferred Asset = A	70 (48.3%)	75 (51.7%)	0.080*
Preferred Asset = B	214 (60.3%)	141 (39.7%)	
Treatment 2			
Preferred Asset = C	60 (40.0%)	90 (60.0%)	0.689
Preferred Asset = D	149 (42.6%)	201 (57.4%)	
Treatment 3			
Preferred Asset = A	39 (34.5%)	74 (65.5%)	0.265
Preferred Asset = B	160 (41.7%)	224 (58.3%)	

Outcome Bias | Suboptimal by Preferred Asset

	Outcome Bias | Suboptimal = 0	Outcome Bias | Suboptimal = 1	Δ p-value
Treatment 1			
Preferred Asset = A	23 (23.5%)	75 (76.5%)	0.421
Preferred Asset = B	33 (19.0%)	141 (81.0%)	
Treatment 2			
Preferred Asset = C	25 (21.7%)	90 (78.3%)	0.959
Preferred Asset = D	55 (21.5%)	201 (78.5%)	
Treatment 3			
Preferred Asset = A	30 (28.8%)	74 (71.2%)	0.993
Preferred Asset = B	91 (28.9%)	224 (71.1%)	

significant (76.7% to 73.1%, p-value=0.437). In Treatment 3, proportions for suboptimal choices are approximately 9% (90.4% to 81.8%, p-value=0.030) lower for participants preferring Investment B. When we turn to Outcome Bias | Suboptimal, differences between both groups of participants are small and insignificant across all treatments. If anything, marginally more suboptimal choices are driven by the Outcome Bias for participants preferring the dominant asset. In conclusion, participants' asset preferences have negligible impact on the most puzzling finding: Once information about the quality of an investment is not provided separately from outcomes, individuals are quick to fall prey to the Outcome Bias. This holds regardless of how simply the investment quality can be inferred from outcomes.

2.6 Conclusion

Individuals struggle to distinguish skill from luck. In particular, good investment outcomes are erroneously associated with good investment decisions. This experiment presents evidence that individuals follow this flawed logic even when monetarily incentivized not to. A considerable fraction of manager choices are outcome biased, although investment quality is prominently displayed. A puzzling result of our study, however, is that individuals' choices are heavily outcome biased once decision quality has to be inferred from decision outcomes, specifically when moderate outcomes are an indicator of a good decision. Specifically, this observation holds regardless of whether outcomes allow for an easy and certain identification of the investment quality (Treatment 2) or whether outcomes only allow for assessing the probability of obtaining good investment quality (Treatment 3).

Our study may also have implications for policymakers. In a recent blog post,[18] AQR hedge fund manager Cliff Asness points to the danger that outcome biased decision making (calling it "after-the-fact" reasoning) poses for policymakers. In particular, he elaborates on how well-intended fiduciary regulations may entail unintended, negative consequences. To give a concrete example, the much-discussed Fiduciary Rule in the U.S. is going to make it easier to sue financial advisors "after-the-fact", that is after investment returns materialize. However, if investors are susceptible to the Outcome Bias, they might decide to sue when returns were low but the financial advice itself was appropriate. In the light of our findings, leaving the disclaimer that *"[...]performance data shown represent past performance, which is not a guarantee of future results"*[19] to the footnotes may not benefit the individual investor. Instead, it should be highlighted that a good investment decision takes into account the investor's personal preferences and characteristics, and that outcomes are only second to it. Emphasizing the impact that unreasonable fund fees have on investment outcomes could be another way of improving *ex ante* investment quality.

[18] Can be found online at `https://www.aqr.com/cliffs-perspective/caveat-investor`.
[19] As shown in the footnotes on the website of a large U.S. fund provider.

Chapter 3

Trust and Delegated Investing: A Money Doctors Experiment *

3.1 Introduction

A longstanding observation in finance is the inability of fund managers to outperform the market after costs. Since Jensen (1968), research in finance has produced numerous studies questioning the skill of fund managers (e.g., Carhart, 1997; Fama and French, 2010).[1] Because the average mutual fund underperforms the market net of fees, investment managers and advisors also advertise other qualities – one of them being trust (Mullainathan et al., 2008).

Trust is a vital aspect of economic transactions (Arrow, 1972). General trust has been linked to overall economic performance (La Porta et al., 1997;

* Authors: Maximilian Germann, Benjamin Loos, and Martin Weber. Maximilian Germann and Martin Weber are at the University of Mannheim. Maximilian Germann: L9, 1-2, 68161 Mannheim. Email: germann@bank.bwl.uni-mannheim.de. Martin Weber: L9, 1-2, 68161 Mannheim. Email: weber@bank.bwl.uni-mannheim.de. Benjamin Loos is at the University of Technology Sydney. Email: Benjamin.Loos@uts.edu.au. We thank seminar participants at the University of Mannheim, participants at the ESA 2017 Vienna conference , and participants at the SEF AP 2018 Brisbane conference for helpful suggestions. Funding from the Graduate School of Economic and Social Sciences at the University of Mannheim (GESS), the Karin-Islinger-Foundation, and the Julius-Paul-Stiegler Gedächtnis Stiftung is gratefully acknowledged. We would like to thank Alexander Sandukovskiy for invaluable programming work. All remaining errors are our own.

1 We are aware that there exist theories rationalizing low after-costs alphas with decreasing returns to scale, see e.g., Berk and Green (2004), Chen et al. (2004), and Pastor et al. (2015). To make our point in a concise way, we resort to discussing only one strand of the literature concerned with mutual fund returns.

Knack and Keefer, 1997), and in particular to stock market participation (Guiso et al., 2004, 2008). In the absence of trust, financial markets need to be more regulated: Trust leaves agents satisfied with (inevitably) incomplete contracts, so that when trust is lost, laws and regulation must provide additional safety for agents (see e.g., Carlin et al., 2009; Sapienza and Zingales, 2012). Opinions in the *Financial Times* (*"Trustworthiness is key for asset managers"*)[2] and in the blog of the *CFA Institute* (*"How to Win Investors' Trust"*)[3] also support the notion that trust is vital for the finance industry.

In a recent paper, Gennaioli et al. (2015) transfer the importance of trust to delegated investing. They propose a model which explains management fees as a trust premium voluntarily paid by investors. All else equal, more trustworthy money managers[4] can set higher fees for *generic* services. In essence, the value that money managers provide is to hold their clients' hands and make them confident to accept risks.

To our knowledge, we are the first to test this theory in an experiment. Our experiment consists of two parts: First, participants play a trust game in the spirit of Berg et al. (1995). This game allows to measure *trusting* and *trustworthy* behavior (Camerer, 2003; Fehr, 2009; Johnson and Mislin, 2011). We exploit variation in the amounts participants return in this game: Higher returned amounts are considered a signal of higher trustworthiness. Second, participants make investment decisions in two treatments. In both treatments, participants are matched to two other participants, who represent money managers. Participants (i.e., investors) then have to invest separately through both money managers. We induce different levels of money manager trustworthiness by providing the amount each money

[2] `https://www.ft.com/content/fc597c2e-8711-11e2-bde6-00144feabdc0`.

[3] `https://blogs.cfainstitute.org/investor/2014/10/21/how-to-win-investors-trust`.

[4] The idea of Gennaioli et al. applies to various financial intermediaries, such as "families of mutual funds, registered investment advisors, financial planners, brokers, funds of funds, bank trust departments, and others who give investors confidence to take risks." (p.92).

manager returned in the trust game. In particular, we provide the level of money manager trustworthiness because it is the *investor* who needs to place differential trust in money managers: Regardless of an investor's unconditional level of trusting, she will place more trust in a money manager who appears more trustworthy. In the first treatment participants have to specify how risky they want to invest with either money manager. These money managers either charge high or low costs. In the second treatment participants have to specify the costs they are willing to bear from one money manager in order to obtain the same investment as with the other money manager.

We find that investors take substantially more risk when investing through a more trustworthy money manager than when investing through a less trustworthy money manager. On average, the share invested into a risky asset is approximately 16% larger for a more trustworthy money manager than for a less trustworthy manager. This finding is striking, since more trustworthy money managers are exogenously assigned twice the costs (1.5%) of less trustworthy money managers (0.75%). Results from the second treatment show that investors are also willing to bear substantially higher costs for investing with more trustworthy managers. On average, investors indicate acceptable costs of 1.95% for a more trustworthy money manager when the less trustworthy money manager charges only 0.75%. Effect sizes from both the first and the second treatment are increasing in the difference in trustworthiness between money managers. Albeit weakened, our findings do not vanish once we control for alternative explanations, such as biased investor beliefs or rewarding (i.e., reciprocity) as motivation for investing with more trustworthy managers. Our study can not, however, discern the precise influence of these alternative factors on trust and delegated investing. That is, our study only demonstrates and quantifies the impact that trust has on delegated investing through the channel of risk aversion.

Empirical support for the Money Doctor theory comes from Kostovetsky (2016), Dorn and Weber (2017), and Linnainmaa et al. (2018). Kostovetsky (2016) uses announced changes in the ownership of fund management companies as exogenous shock to an existing trust relationship. The study finds that, after controlling for fund characteristics, approximately 7% of assets are withdrawn in the 12-month period following the announcement. Because retail investors and investors in funds with higher expense ratios (i.e., those funds able to extract higher trust premia) are most responsive to ownership changes, Kostovetsky interprets his findings as evidence for the Money Doctor theory. Dorn and Weber (2017) find that retail investors who had delegated all their equity investments to fund managers – money doctors – before the financial crisis, were almost twice as likely to exit the stock market during the crisis than their peers who invested into individual stocks. This finding is consistent with the view of Gennaioli et al. (2015) that those investors relying on a trust relationship to invest into the stock market will be particularly affected by a negative shock to this trust relationship. Linnainmaa et al. (2018) proxy trustworthiness by the length of a client-advisor relationship. They show that, consistent with the Money Doctor hypothesis, investors with a longer-established client-advisor relationship are more willing to take financial risks.

Nonetheless, these empirical studies only reveal the direction in which trust affects mutual fund flows and investor behavior, respectively. Neither empirical setting does allow for a clean quantification of trust or a measurement of the trust-cost-relationship. The assumption that investors balance trust against management fees, however, is critical to the Money Doctors theory. Testing the theory in a controlled experiment allows for both a quantification of trust and a measurement of the trust-cost-relationship. Thus, we contribute to the understanding of the mechanism of the Money Doctors theory.

The rest of the paper is organized as follows: Section 3.2 gives a brief overview of the Money Doctors theory. Section 3.3 outlines the experimental design, in particular the trust game and the two investment treatments, and testable Money Doctors hypotheses are derived. General results follow in section 3.4. In section 3.5, alternative explanations for our results are discussed. Section 3.6 concludes.

3.2 Money Doctors Theory

In the following, we briefly sketch the model of Gennaioli et al. (2015) that we seek to test. Gennaioli et al. think of trust as an ingredient that reduces the perceived riskiness of an investment. In particular, investing through a more trusted money manager is more effective in reducing perceived riskiness of financial investments than is investing through a less trusted money manager. Placing this idea in an economic context, investors' risk aversion is lower when investing with a trusted money manager. Importantly, money managers offer identical investment services and investors have correct beliefs about the investment services provided by money managers.[5] Hence, trustworthiness is *not* mistaken for skill. Formally, assuming a standard quadratic utility function, this translates to

$$u_{i,j}(c) = \mathbb{E}(c) - \frac{a_{i,j}}{2} Var(c), \tag{3.1}$$

where c is the investor's future consumption. Parameter $a_{i,j} \geq 1$ represents investor i's "anxiety" of investing with money manager j. To keep the model simple, Gennaioli et al. (2015) assume that investors do not invest risky themselves, which implies $a_{i,i} = \infty$. From the investor's utility function it becomes

[5] In the latter part of their paper, Gennaioli et al. (2015) also examine implications of their model if investors hold biased beliefs. Our paper, however, focuses on the part of their paper in which investors hold correct beliefs.

evident that placing more trust into a money manager, thereby reducing $a_{i,j}$, decreases the costs of bearing investment risk. However, this also means that more trusted money managers are able to exploit their relative advantage over their less trusted counterparts. Ceteris paribus, more trustworthy money managers can charge higher fees without losing investors to competitors. From the investor's point of view, the investment problem becomes one of weighting the benefits of trust – less perceived risk and thus greater participation in risky investments – against the costs of management fees. Given a riskless asset with return R_f (in which investors can invest on their own) and a risky asset with excess return over the riskless asset of R and variance σ^2,[6] investor i's expected utility of investing with manager j is thus equal to

$$U_{i,j}(x_{i,j}, f_j) \equiv R_f + x_{i,j}(R - f_j) - \frac{a_{i,j}}{2} x_{i,j}^2 \sigma^2, \tag{3.2}$$

where the share of wealth invested into the risky asset is denoted by $x_{i,j}$. Solving for the optimal portfolio composition thus yields

$$\hat{x}_{i,j} = \frac{(R - f_j)}{a_{i,j}\sigma^2}. \tag{3.3}$$

Therefore, the investor will invest a larger proportion of his portfolio into the risky asset when investing with a more trusted money manager. Substituting $\hat{x}_{i,j}$ back gives the utility obtained from investing optimally:

$$U_{i,j}(\hat{x}_{i,j}, f_j) = R_f + \frac{(R - f_j)^2}{2a_{i,j}\sigma^2}. \tag{3.4}$$

Investors still have to choose among money managers. The simplest case is the choice between two money managers (referred to as manager A and manager B), as outlined in the original model. In the simplest case, the investor will prefer manager A over manager B provided that $U(\hat{x}_{i,A}, f_A) \geq$

[6] Gennaioli et al. (2015) denote variance as σ (p.95).

$U(\hat{x}_{i,B}, f_B)$. Rearranging the relationship yields a central prediction of the theory:

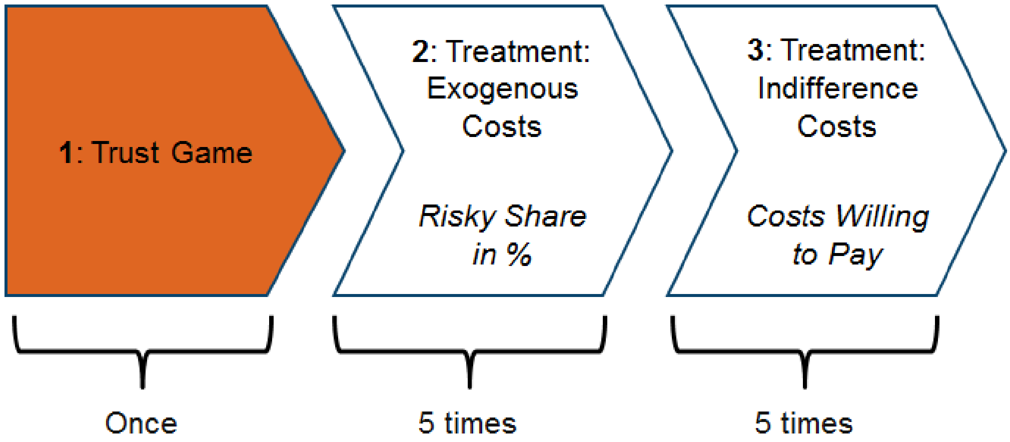
$$\frac{a_{i,B}}{a_{i,A}} \geq \frac{(R - f_B)^2}{(R - f_A)^2}. \tag{3.5}$$

Hence, the investor will choose manager A provided that the benefit of trustworthiness overcompensates for the disutility stemming from higher management fees. The investor's choice thus depends on the difference, but not the level, in trustworthiness of money managers.

3.3 Experimental Design and Hypotheses

The experiment consists of two distinct parts. In the first part, we aim to collect a measure of trustworthiness that is based on human interaction. This step is necessary in order to induce different levels of trustworthiness in the second part. For this purpose, participants first play a trust game. In the second part, participants face two treatments in which they have to make investment decisions. In the first treatment, participants have to make two separate investment decisions with two different money managers, who charge different costs. In the second treatment, participants have to indicate management fees they are willing to pay for one money manager in order to obtain the same investment allocation as with another money manager. Building on the first part, the treatments in the second part allow us to test predictions of the theory. Participants do not know what the second part looks like before completing the first part. Thus, participants have no reason to bias their behavior in the first part in order to obtain more favorable outcomes in the second part. The experiment concludes with control questions and a sociodemographic survey. The sequence of the experiment is shown in Figure 3.1. In the following, the details of the experiment are laid out.

Figure 3.1: Diagram of Experimental Setup



3.3.1 Trust Game

Gennaioli et al. (2015) emphasize that they "do not think of trust as deriving from past performance" (p.92). Since we want to adhere to the original paper, money managers' trustworthiness must also not be induced by past performance in our experiment. We opt for a trust game (Berg et al., 1995) to induce differences in trustworthiness. In the trust game, a sender (trustor) is endowed with an amount X. The sender can transfer any amount between 0 and X to the receiver (trustee). The amount sent to the trustee, S, is then tripled. The trustee has the choice to reciprocate by returning any amount between 0 and $3S$. Because trustees are not obliged to return anything, self-interested trustors should not send anything in the first place. In the trust game, sending is therefore associated with *trusting* behavior, while returning is associated with *trustworthy* behavior (Camerer, 2003; Fehr, 2009; Johnson and Mislin, 2011).

We use the trust game for two reasons: First, results of the trust game are derived from actual human interaction. Second, the trust game is a well-studied game in the economics literature and has been found to predict trusting and trustworthy behavior also outside the lab (see e.g., Baran et al., 2010;

Aksoy et al., 2018). In finance, pro-social behavior in the trust game has recently been linked to real-world propensity to hold socially responsible investments (SRI) (Riedl and Smeets, 2017).

Results from trust games show that trustors usually send part of their endowment, and that trustees usually reciprocate to a certain extent. In a meta study of more than 160 trust games, Johnson and Mislin (2011) find that participants on average send 50% of their endowment, and return 50% of the available amount. Several studies also show that there is variation in the amounts sent and the amounts returned in the trust game (Berg et al., 1995; Croson and Buchan, 1999; Buchan et al., 2002; Keser, 2002; Ashraf et al., 2003; Cox, 2004; Kosfeld et al., 2005; Dubois et al., 2012). These empirical observations are critical for our experiment: Since the amount returned in the trust game represents the level of trustworthiness, we can exploit variation in the amount returned by trustees to induce differences in trustworthiness. Not critical for our experiment is how exactly the trust game is designed. Changing parameters of the trust game (e.g., doubling or quadrupling the amount sent) or having participants play both sender and receiver in the trust game may affect participant behavior (see the meta study by Johnson and Mislin, 2011). These modifications to the trust game, however, affect all participants and thus only affect the level of trust and trustworthiness. Identification in our experiment is based on differences in participants' behavior and is therefore not susceptible to changes in the design of the trust game.

In the first part of the experiment, participants are paired anonymously and randomly. Senders are endowed with 100 ECU and can send any amount of tens between 0 and 100 ECU. The amount sent is then tripled, and receivers can return any amount of tens between 0 and the tripled amount sent. The trust game is played using the strategy method: Participants indicate *a)* how much they would be willing to send if they were playing as sender, and *b)* how much they would be willing to return for any possible amount sent if they were playing as receiver.[7] We incentivize choices in the trust game by randomly picking the roles within each pair and by evaluating the trust game according to the indicated choices. If the trust game is chosen randomly to determine participants' payoff from the experiment, 1 ECU is converted to 0.05€. Based on average levels of trust and trustworthiness found by Johnson and Mislin (2011), the expected payoff for senders is 7.5€ and 3.75€ for receivers.[8]

3.3.2 Treatment: Exogenous Costs

After the trust game, every participant plays the role of an investor. Investors have the choice to invest their endowment (100 ECU) into a riskless asset with return $r = 0$ and a risky asset with normally distributed returns with mean of 6% and volatility of 20%. Because we are interested in the impact of trust on the investment decision, we match every investor with two other participants and their respective decisions in the trust game. These two participants represent money managers. Investors then have to make separate investment decisions with both money managers.

Money managers do not effectively act. In other words, they do not influence the characteristics of the riskless and the risky investment – just as money managers in the real world have no control over the movement of the stock market. For both money managers, the identical expected asset returns before costs are displayed prominently. Nonetheless, both money managers can be associated with a different level of trustworthiness. This level of trustworthiness stems from the money managers' decision to return ECU in the trust game. A money manager who was willing to return more

[7] Using the strategy method for simple economic games such as the trust game has been found to yield similar results as direct (i.e., playing only one role and only once) elicitation approaches, see e.g. Brandts and Charness (2000), Brandts and Charness (2011), and Vyrastekova and Onderstal (2010).

[8] Senders send 50% (50 ECU) of their endowment (100 ECU) and receivers return 50% (75 ECU) of the available amount (150 ECU, 3 times 50 ECU).

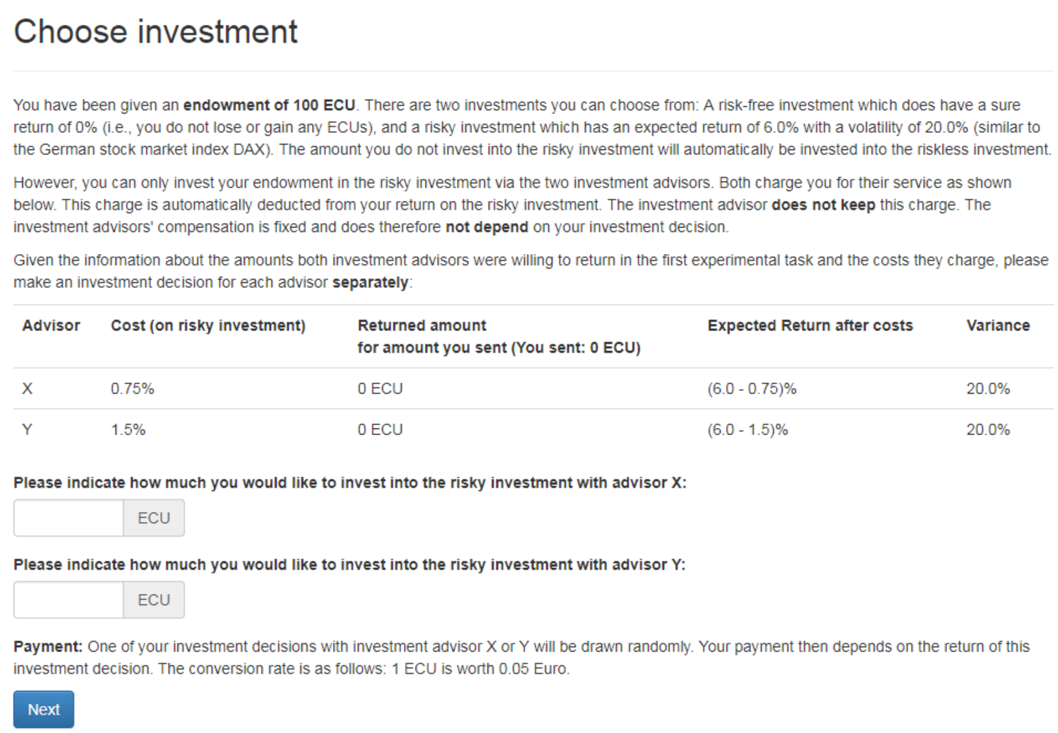
ECU in the trust game is therefore more trustworthy than a money manager who was willing to return fewer ECU. As in the Money Doctors model, risky investments can only be made via money managers. Crucially, both money managers offer identical risky investments *before* costs (mean return of 6% and volatility of 20%). However, money managers charge different costs – specifically, the money manager who returned more in the trust game is assigned high costs ($C_h = 1.5\%$), the money manager who returned less in the trust game is assigned low costs ($C_l = 0.75\%$). In case both money managers returned the same amount in the trust game, one is randomly assigned high costs and one is randomly assigned low costs. We deliberately rule out trivial cases in which more trustworthy managers also charge lower costs. Known to participants, costs are not transferred to managers. Hence, concerns of higher risky investments as means of monetarily "rewarding" more trustworthy managers are alleviated.

Investors receive the following information: *1)* the mean and the volatility of the risky asset, *2)* the costs each money manager charges, and *3)* the amount each money manager was willing to return in the trust game for the amount the investor was willing to send. An exemplary screen of the treatment, also showing the exact wording of the instructions, is shown in Figure 3.2.

Since only one of the two investment decisions is selected randomly for payoff, diversification across money managers is not possible. Thus, rational (risk-averse) investors should invest a greater share of their endowment into the risky asset via the low-cost, low-trust manager than via the high-cost,

Figure 3.2: Exemplary Screen of Treatment: Exogenous Costs

This figure is a screenshot of the instructions and the action screen of Treatment: Exogenous Costs. The level of money manager trustworthiness, as proxied by the amount returned in the trust game, is displayed in the third column as "**Returned amount for amount you sent (You sent: X ECU)**". Because exemplary choices for this screenshot were to send 0 ECU as sender and return 0 ECU as receiver for any amount sent, the level of trustworthiness shows as "0 ECU".



Choose investment

You have been given an **endowment of 100 ECU**. There are two investments you can choose from: A risk-free investment which does have a sure return of 0% (i.e., you do not lose or gain any ECUs), and a risky investment which has an expected return of 6.0% with a volatility of 20.0% (similar to the German stock market index DAX). The amount you do not invest into the risky investment will automatically be invested into the riskless investment.

However, you can only invest your endowment in the risky investment via the two investment advisors. Both charge you for their service as shown below. This charge is automatically deducted from your return on the risky investment. The investment advisor **does not keep** this charge. The investment advisors' compensation is fixed and does therefore **not depend** on your investment decision.

Given the information about the amounts both investment advisors were willing to return in the first experimental task and the costs they charge, please make an investment decision for each advisor **separately**:

Advisor	Cost (on risky investment)	Returned amount for amount you sent (You sent: 0 ECU)	Expected Return after costs	Variance
X	0.75%	0 ECU	(6.0 - 0.75)%	20.0%
Y	1.5%	0 ECU	(6.0 - 1.5)%	20.0%

Please indicate how much you would like to invest into the risky investment with advisor X:

ECU

Please indicate how much you would like to invest into the risky investment with advisor Y:

ECU

Payment: One of your investment decisions with investment advisor X or Y will be drawn randomly. Your payment then depends on the return of this investment decision. The conversion rate is as follows: 1 ECU is worth 0.05 Euro.

Next

high-trust manager.[9] Alternatively, if trustworthy money managers are effective in holding a client's hand, investors could also invest more risky via the high-cost, high-trust manager. In particular, the share invested risky with the high-trust manager relative to the share invested risky with the low-trust manager should increase the larger the difference in trustworthiness.

After one investment decision is chosen randomly, the return of the respective risky investment is drawn and costs are deducted. Participants are then informed which choice was drawn and how their investment decision turned out. As in the trust game, 1 ECU is converted to 0.05€. The expected payoff from this task hence varies between 5€ and 5.3€, depending on how much risk is taken. Afterwards, investors are again independently matched with two new money managers. In total, this investment task is repeated five times with independent matchings of new money managers. If participants' payment for participation is randomly chosen to be determined by this treatment, the outcome of one of the five rounds is chosen randomly. In summary, we test the following hypotheses in the first treatment:

Hypothesis 1 ("Hand-holding"): Investors invest a larger proportion of their wealth into the risky investment via a more trustworthy money manager (higher amount returned in trust game) than via a less trustworthy money manager (lower amount returned in trust game), even if the more trustworthy money manager charges higher costs (twice as much) than the less trustworthy money manager.

Hypothesis 2: The larger the difference in trustworthiness between money managers, the larger the share invested risky with

[9] Risk aversion is assumed in the Money Doctors model. In our experiment, risk-seeking or risk-neutral preferences would imply that investors should invest all their wealth into the risky asset, as it offers a positive expected return as opposed to the riskless asset. From participants' actual choices we can assume that participants do not have such preferences: No participant invested all his wealth into the risky asset in all rounds and only two participants invested all their wealth into the risky asset in four out of five rounds.

the more trustworthy money manager relative to the share invested risky with the less trustworthy money manager.

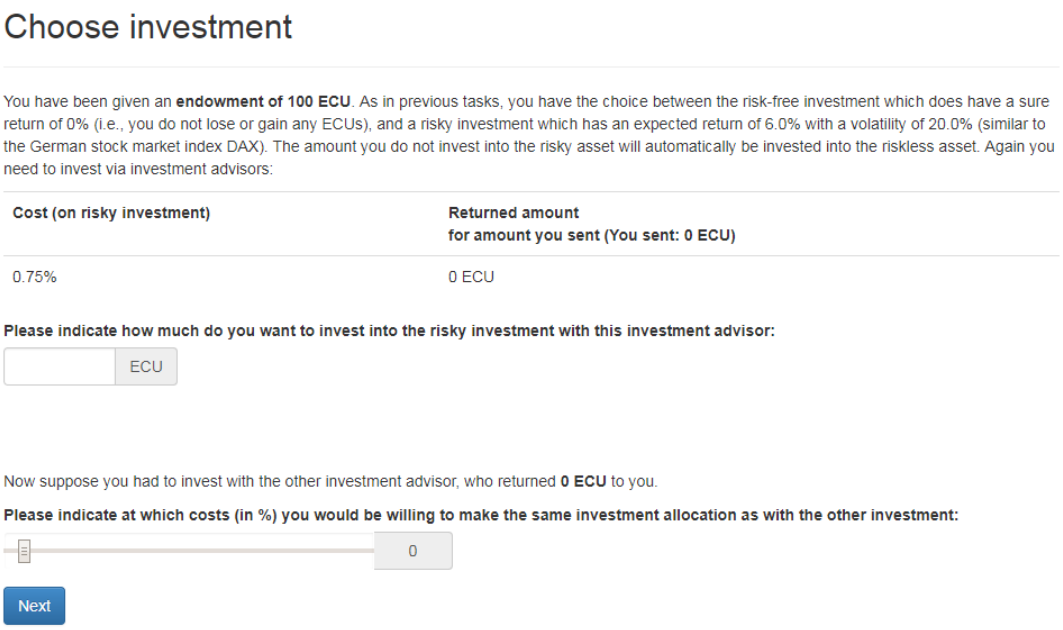
3.3.3 Treatment: Indifference Costs

There are two possible identification strategies in our experiment: One is to fix costs and exploit variation in the share invested risky. The second is to fix the share invested risky, and exploit variation in costs. The previous treatment fixes costs and allows us to elicit investors' risk aversion, which is potentially lowered by trust. In this treatment, we investigate the costs investors are willing to bear to make the same investment decision with a more trustworthy money manager as with a less trustworthy money manager. Again, every participant is matched with two other participants. First, acting as investor, every participant has to indicate how much she would invest risky with the first money manager. Parameters of both assets, riskless and risky, are identical to the previous treatment. By construction, the first money manager always charges fees of $C_l = 0.75\%$ and always returned less than or equal to the second money manager in the trust game. We impose this restrictions to increase the reliability of statistical testing, as costs logically have to be bounded by 0%. Second, investors have to indicate the costs they are willing to accept from the second money manager in order to obtain the same risky investment as with the first money manager.

Participants indicate their indifference costs on a slider with a lower bound of 0% and an upper bound of 10%.[10] The default input is set to 0%, which, if anything, would imply an anchoring bias against our hypothesis. Figure 3.3 shows the setting. As predicted by the theory, investors should indicate higher indifference costs for more trusted money managers. Choices in this treatment are not monetarily incentivized, as indicating indifference

[10] In pretests, participants had trouble entering fees in the correct numerical units when presented with an input box. Thus we opt for the more restrictive slider input.

Figure 3.3: Exemplary Screen of Treatment: Indifference Costs



costs of 0% would be a dominant strategy.[11] Again, this task is repeated five times with new random and independent matchings. In summary, we test the following hypotheses in the second treatment:

Hypothesis 3: Investors are willing to accept higher costs from more trustworthy money managers in order to obtain the same investment allocation as with a less trustworthy money manager.

Hypothesis 4: The larger the difference in trustworthiness, the higher the costs investors are willing to accept from more trustworthy managers in order to obtain the same investment allocation as with less trustworthy money managers.

[11] We refrain from using an incentive-compatible Becker-Degroot-Marschak (Becker et al., 1964) mechanism, as we believe it would considerably complicate the second treatment for participants.

3.4 Results

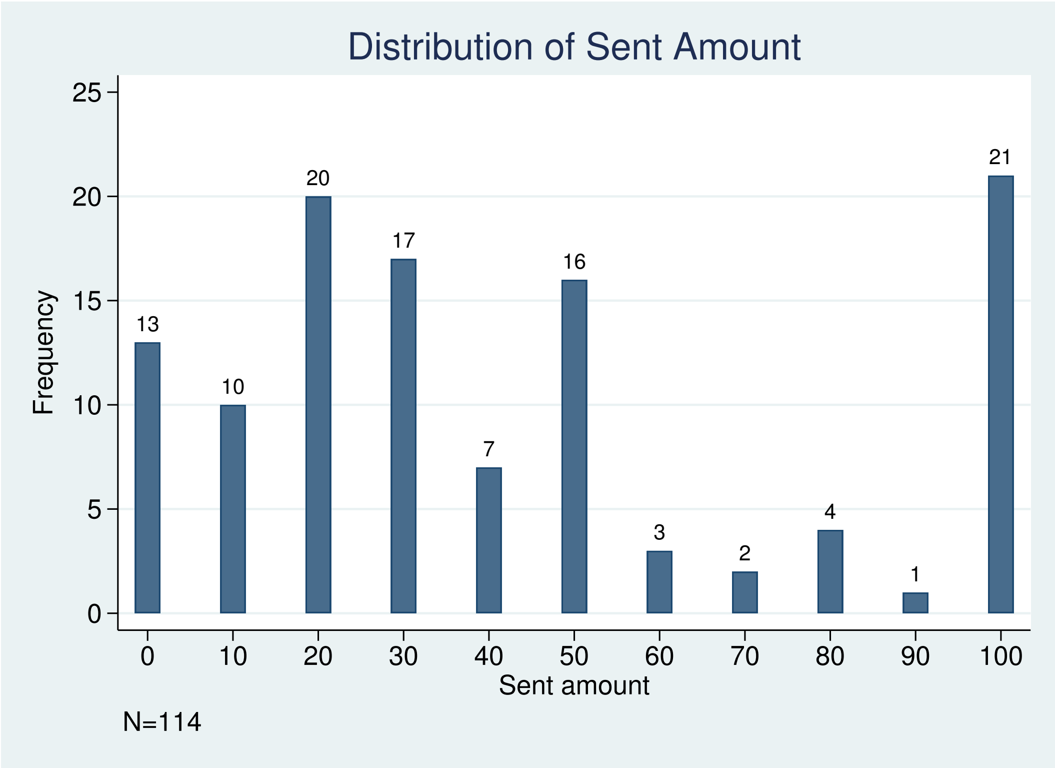
The experiment took place at the University of Mannheim experimental laboratory in July and September 2017. Participants were invited through ORSEE (Greiner, 2015). The experiment was computerized using oTree (Chen et al., 2016). In total, 114 individuals participated in 8 sessions. Participants were predominantly female (58.77%). Almost all participants were students (98.25%). Thus, the mean age was relatively low at 23.35 (*SD*=3.99) years. Furthermore, most participants studied business or economics (71.05%). However, only few participants had any real investment experiences: Only 20.18% and 11.40% of all subjects had invested in passive or active funds, respectively. Sessions lasted approximately 30 minutes and the average payment for participation was 6.16€, including a base payment of 1€. The minimum payment was 2.5€, the maximum payment 16€, and payment variance was 1.86.

Trust Game

In order to induce different levels of trustworthiness, there must be variation in participants' choices in the trust game. Results from our trust game are in line with the literature. Participants usually trust their counterpart. Only 13 (11.4%) participants resorted to the equilibrium strategy of sending 0 ECU in the trust game. On average, 43.16 ECU were sent from trustors. The distribution of sent amounts is depicted in Figure 3.4.

The measure of money manager trustworthiness, however, is the amount the money manager returns in the trust game. Hence, to establish a situation which allows us to test predictions from the Money Doctors theory, there must also be variation in the amounts returned in the trust game. For every possible choice of trustors, we find substantial variation in the choices

Figure 3.4: Distribution of Sent Amount in Trust Game

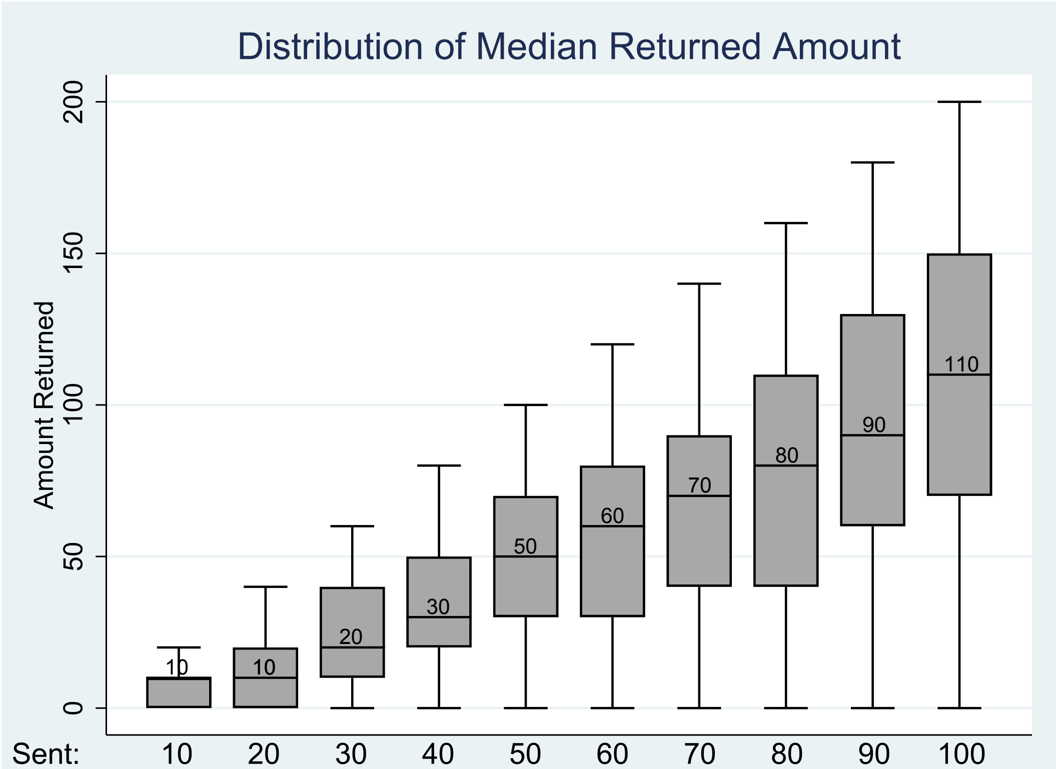


of trustees. The average standard deviation of returned amounts is approximately 29.49. Figure 3.5 shows a boxplot of median returned amounts in the trust game. As expected, the absolute median returned amount increases with the amount sent. Nonetheless, the relative level of median reciprocity (amount returned divided by amount sent) stays relatively constant at one. In summary, results from the trust game offer sufficient variation for our subsequent analysis.

Treatment: Exogenous Costs

In this treatment, we are interested in the share of wealth participants invest risky with both money managers. Specifically, we want to test whether investors are willing to invest more risky with more trustworthy money managers, even if that investment comes at higher costs. For this reason, simple univariate analyses are reported first. To assess differences in the shares of wealth invested risky, we need to account for the fact that observations

Figure 3.5: Distribution of Median Returned Amounts in Trust Game



are not necessarily independent, since participants face multiple choices per treatment. Thus, we regress the difference of the share invested risky with the more trustworthy and the share invested risky with the less trustworthy money manager on a constant only and cluster standard errors at the individual level. Hence, we effectively run a test of means, but adjust for potential non-independence of observations. For all subsequent comparisons of means we use this approach as well. Univariate p-values reported subsequently are therefore adjusted for clustering at the individual level. In Table B.1 in the appendix, we also report results of tests for each round individually.

The order in which more or less trustworthy money managers appear in the investment decision screen is randomized. As results are similar for cases in which the more trustworthy money manager appears on top and for cases in which she appears at the bottom, pooled results are reported throughout this paper. Table 3.1 compares the average amount invested risky with both money managers. When money managers are *not* equally

Table 3.1: Risky Share of Investment

This table shows the share invested into the risky asset, *Risky Share*, for both money managers. *High Trustworthiness, High Costs* corresponds to the more trustworthy (i.e., returned more in the trust game) but more costly money manager. *Low Trustworthiness, Low Costs* corresponds to the less trustworthy (i.e., returned less in the trust game) but less costly money manager. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

		Risky Share in %			
	N	μ	σ	5th Percentile	95th Percentile
High Trustworthiness, High Costs	410	46.15	29.20	0	100
Low Trustworthiness, Low Costs	410	29.27	27.07	0	100
		Δ t-stat = 6.58***			

trustworthy, investors are willing to invest substantially more risky with the trustworthy, but more expensive money manager. The difference of 16.88% is highly statistically significant. Investors profit from investing through a more trustworthy money manager in terms of expected return: The average investment decision with the more trustworthy money manager implies a total expected return on the portfolio of 2.07% (Mean Risky Share$_{HT,HC}$ times 4.75%), whereas the average investment decision with the less trustworthy money manager translates only to a total expected return on the portfolio of 1.54% (Mean Risky Share$_{LT,LC}$ times 5.25%, p-value=0.000). More precise, investors essentially move upwards on the Capital Market Line. While the total investment's Sharpe ratio is unchanged, it is more risky overall and thus offers higher expected return.

By construction, we prohibit trivial cases in which more trustworthy money managers charge lower costs than less trustworthy money managers. However, there are cases in which both money managers are equally trustworthy. If there is no difference in trustworthiness, investors are expected to invest more risky with the money manager who charges lower costs. Results are provided in Table 3.2. On average, investors invest a higher share of their wealth into the risky asset if costs are lower. This difference of 6.98% is also significant at the 10%-level. However, 13 participants

Table 3.2: Risky Share of Investment for Identically Trustworthy Money Managers

This table shows the share invested into the risky asset, *Risky Share*, for both money managers when both money managers are equal in trustworthiness (i.e., returned identical amounts in the trust game). The type of costs, high or low, is indicated by *High Costs* and *Low Costs*, respectively. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

		All Participants			
		Risky Share in %			
	N	μ	σ	5th Percentile	95th Percentile
Low Costs	160	32.71	28.11	0	100
High Costs	160	25.73	27.32	0	100
		Δ t-stat = 1.77*			
		Only Participants Who Sent > 0 ECU			
Low Costs	95	35.84	30.89	0	100
High Costs	95	23.14	25.90	0	100
		Δ t-stat = 3.11***			

chose the Nash equilibrium strategy in the trust game. By default, these participants are always presented cases in which both money managers are equally trustworthy: If senders send 0 ECU, the only choice of receivers is to return 0 ECU. Excluding the choices (13·5=65 choices) of these non-trusting participants results in an increased and highly significant difference of 12.71%. In the Table B.2 in the appendix, we again report results of tests for each round individually. In summary, univariate analyses strongly support our first hypothesis. Investors voluntarily pay a trust premium and are less risk averse when investing with trustworthy money managers. Nonetheless, investors benefit from this increase in risk taking even net of fees.

We use multivariate analyses to test our second hypothesis. Equation (3.5) states that the discrepancy of trustworthiness between money managers is a key factor in the Money Doctors framework by Gennaioli et al. (2015). To analyze whether the difference in trustworthiness is related to the difference of the share invested risky, the following random effects model (RE_i) with round fixed effects ($Round_t$) is estimated:

$$\Delta\ Risky\ Share_{it} = \alpha + \Delta\ Trustworthiness_{it}\beta + RE_i + Round_t + \epsilon_{it}.$$

A random effects model is used because the independent variable, $\Delta\ Trustworthiness_{it}$, is orthogonal to other regressors, as it is obtained by randomly matching investors to money managers. For robustness, fixed effects regressions are reported in the appendix. The dependent variable, Δ *Risky Share*, is calculated as the share of wealth invested risky with the more trustworthy money manager minus the share of wealth invested risky with the less trustworthy manager. In case both managers are equally trustworthy, it is calculated as the share of wealth invested risky with the more costly manager minus the share of wealth invested risky with the less costly manager. Therefore, the constant in the regression is expected to be negative. Because Δ *Risky Share* is technically censored at -100 and +100, we also report random effects tobit regressions in Table B.4 in the appendix.

For the independent variable, Δ *Trustworthiness*, we test three different specifications. In a first specification, it is calculated in absolute terms as the amount the more trustworthy manager returned in the trust game minus the amount the less trustworthy manager returned in the trust game. This absolute difference, however, depends on the amount that was sent in the trust game. For larger amounts sent, the absolute measure may thus be substantially larger. To correct for this mechanical relationship, in a second specification the relative difference in trustworthiness is calculated. It captures the percentage the less trustworthy manager sent less than the more trustworthy manager and is calculated as $(1 - (\frac{Lower\ Returned\ Amount}{Higher\ Returned\ Amount})) * 100$.

Table 3.3: Summary Statistics Δ *Trustworthiness*

Δ *Trustworthiness Absolute* is calculated as the amount the more trustworthy manager returned in the trust game minus the amount the less trustworthy manager returned in the trust game. Δ *Trustworthiness Relative* is calculated as $(1 - (\frac{Lower\ Returned\ Amount}{Higher\ Returned\ Amount})) * 100$. Δ *Trustworthiness Relative to Sent* is calculated as $(\frac{Higher\ Returned\ Amount\ -\ Lower\ Returned\ Amount}{Amount\ Sent}) * 100$.

	N	μ	σ	5th Percentile	95th Percentile
Δ *Trustworthiness Absolute*	570	28.77	34.78	0	160
Δ *Trustworthiness Relative*	570	49.00%	39.86%	0%	100%
Δ *Trustworthiness Relative to Sent*	570	60.00%	56.58%	0%	300%

As a third and last specification, the difference in trustworthiness is calculated adjusting for the amount the investor sent in the trust game. This approach aims at controlling for potentially different sensitivity to differences in trustworthiness depending on investors' own level of trusting. As shown in Table B.3, more trusting investors – not surprisingly – invest more risky and state higher indifference costs. The third measure is calculated as $(\frac{Higher\ Returned\ Amount\ -\ Lower\ Returned\ Amount}{Amount\ Sent}) * 100$. In all three specifications, however, we also implicitly control for the amount sent through random or fixed effects. Summary statistics for all three variable specifications are shown in Table 3.3.

Regression results are summarized in Table 3.4. All regressions account for potential learning effects by including round fixed effects. As hypothesized, coefficients for differences in trustworthiness are positive and significant across all regression specifications. That is, the larger the difference in managers' trustworthiness, the larger the difference of the share invested risky. An absolute difference in trustworthiness of 1 ECU therefore relates to an increase of the amount invested risky with the more trustworthy manager over the amount invested risky with the less trustworthy manager of 0.33 ECU (see column 1). In other words, a third of the absolute difference in trustworthiness translates directly into a difference of the share invested

Table 3.4: Risky Share – Difference in Trustworthiness

This table reports regression results with Δ*Risky Share* as dependent variable. It is calculated as the share of wealth invested risky with the more trustworthy money manager minus the share of wealth invested risky with the less trustworthy money manager. In case both managers are equally trustworthy, it is calculated as the share of wealth invested risky with the more costly manager minus the share of wealth invested risky with the less costly manager. All regressions account for unobserved individual heterogeneity through random effects. Δ *Trustworthiness Absolute* is calculated as the amount the more trustworthy manager returned in the trust game minus the amount the less trustworthy manager returned in the trust game. Δ *Trustworthiness Relative* is calculated as $(1 - (\frac{Lower\ Returned\ Amount}{Higher\ Returned\ Amount})) * 100$. Δ *Trustworthiness Relative to Sent* is calculated as $(\frac{Higher\ Returned\ Amount\ -\ Lower\ Returned\ Amount}{Amount\ Sent}) * 100$.
Standard errors clustered at the individual level are shown in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

	(1)	(2)	(3)
		Random Effects	
Δ*Trustworthiness Absolute*	0.330***		
	(0.067)		
Δ*Trustworthiness Relative*		0.248***	
		(0.042)	
Δ*Trustworthiness Relative to Sent*			0.176***
			(0.031)
Constant	-0.669	-3.801	-1.909
	(3.809)	(4.295)	(3.954)
Observations	570	570	570
Cluster-robust S.E.	YES	YES	YES
Round FE	YES	YES	YES
$R^2_{overall}$	0.082	0.054	0.066

risky. A similar picture remains for relative differences in trustworthiness. Returning 1% less than a more trustworthy manager results in a difference of attracted risky investments of 0.25 percentage points. Scaled by the amount investors sent, a relative difference in trustworthiness of 1% still implies a difference of the share invested risky of 0.18 percentage points. Evidence from three regressions thus is in favor of our second hypothesis. In general, investors are sensitive to differences in trustworthiness. These differences also translate to the risky investment choice: The more trustworthy a money manager is relative to a competitor, the more funds she can attract relative to this competitor.

Treatment: Indifference Costs

Instead of investigating the share invested risky with money managers, one may also look at the costs investors are willing to bear to make risky investments. In this treatment, participants are asked to make an investment decision with one money manager first, and indicate at which costs they are indifferent between making the same risky choice with a second money manager. By construction, the first manager always charges $C_l = 0.75\%$. Thus, to test the third hypothesis, we compare C_l to the average indifference costs investors indicate in cases in which the second money manager is more trustworthy than the first. Results are shown in Table 3.5. On average, investors accept costs of 1.95% when the second money manager is more trustworthy than the first. These costs are 2.6 times the costs the less trustworthy manager charges, or, put differently, almost a third of the return of the risky investment. The difference to the low costs the less trustworthy manager charges is statistically significant at the 1%-level. Table 3.5 also provides the results of a test of those cases in which the second manager and the first manager are equally trustworthy. In this scenario, indifference costs should not be significantly greater than 0.75%. Indeed, indifference costs are only 0.844% on average, and the difference to 0.75% is statistically insignificant. Results are virtually identical if we only include participants who sent a positive amount in the trust game. Again we report results of tests for each round individually in Table B.6 in the appendix.

As in the first treatment, we also test the impact of differences in trustworthiness (Hypothesis 4). For this purpose, indifference costs are used as dependent variable in random effects regressions. Because these costs are technically censored at 0 and +10, we report random effects tobit regressions in Table B.7 in the appendix. The same specifications for Δ *Trustworthiness* as in the previous regressions are used. Table 3.6 shows the results. Coefficients

Table 3.5: Indifference Costs

This table shows indifference costs of investing with the second money manager in Treatment 2. Tests are based against the costs the first money manager charges, which are equal to 0.75%. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Trustworthiness Second Manager > First Manager

	N	μ	σ	5th Percentile	95th Percentile
Indifference Costs	412	1.946	2.243	0	8.02
		Δ *t*-stat = 6.42***			

Trustworthiness Second Manager = First Manager

Indifference Costs	158	0.844	1.174	0	5
		Δ *t*-stat = 0.56			

Trustworthiness Second Manager = First Manager
Only Participants Who Sent > 0 ECU

Indifference Costs	93	0.845	1.021	0	2.99
		Δ *t*-stat = 0.69			

Table 3.6: Indifference Costs – Difference in Trustworthiness

This table reports regression results with *Indifference Costs* as dependent variable. All regressions account for unobserved individual heterogeneity through random effects. Δ *Trustworthiness Absolute* is calculated as the amount the second manager returned in the trust game minus the amount the first manager returned in the trust game. Δ *Trustworthiness Relative* is calculated as $(1 - (\frac{Lower\ Returned\ Amount}{Higher\ Returned\ Amount})) * 100$. Δ *Trustworthiness Relative to Sent* is calculated as $(\frac{Higher\ Returned\ Amount\ -\ Lower\ Returned\ Amount}{Amount\ Sent}) * 100$.
Standard errors clustered at the individual level are shown in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

	(1)	(2)	(3)
	Random Effects		
Δ*Trustworthiness Absolute*	0.0098***		
	(0.0039)		
Δ*Trustworthiness Relative*		0.0063***	
		(0.0022)	
Δ*Trustworthiness Relative to Sent*			0.0069***
			(0.0020)
Constant	1.514***	1.500***	1.390***
	(0.181)	(0.167)	(0.181)
Observations	570	570	570
Cluster-robust S.E.	YES	YES	YES
Round FE	YES	YES	YES
$R^2_{overall}$	0.036	0.031	0.059

are positive and hence point into the hypothesized direction in all specification. All coefficients, that is for absolute and relative differences in trustworthiness, are significant at the 1%-level. In economic terms, investors are willing to accept 0.63 basis points more management fees from a 1% more trustworthy manager. Scaled by the amount investors sent, a relative difference in trustworthiness of 1% translates to 0.69 basis points higher management fees accepted by investors for investing with the more trustworthy money manager. Findings from Treatment 2 therefore provide further evidence of the Money Doctors hypothesis.

3.5 Alternative Explanations

Is Trustworthiness Mistaken for Skill?

Trustworthiness could be mistaken for investment skill. Investors could believe that more trustworthy money managers are able to deliver better investment performance. Beliefs could be such that more trustworthy managers offer an expected return that overcompensates for their higher management fees. In this case, rational investors should invest more risky with the more trustworthy – more skilled – money manager. To control for such biased beliefs, we ask participants whether they believed that more trustworthy money managers could deliver better investment performance after the experiment. Possible answers are "Yes", "No", and "I do not know". We deliberately refrain from asking for participants' beliefs about asset returns during the experiment, as this might tempt them to believe that there was a difference in investment skill, simply because we ask for it explicitly.

The majority of participants (n=66) believes that more trustworthy managers can deliver better investment performance. There may be two explanations for this observation. On the one hand, participants can justify their choices in the experiment ex post. By stating that they (incorrectly) believed that more trustworthy money managers were able to offer better investment performance, participants can rationalize the behavior the experimenter observes. On the other hand, believing in the ability of more trustworthy managers to deliver better investment performance may reflect investors' "wishful thinking". An analogy for this explanation can be a medical diagnosis: Today many diagnoses are automatized by medical hard- and software, and are hence largely independent of the doctor supervising the diagnosis. Nonetheless, patients' trust in a doctor determines how this doctor's abilities are perceived – even if the doctor's diagnosis is based on algorithms (see e.g., Promberger and Baron, 2006; Arkes et al., 2007).

We first contrast choices of participants holding biased beliefs with the choices of those participants holding correct beliefs (n=36, "I do not know" answers excluded) in Treatment 1. As expected, holding biased beliefs increases the difference between the share invested risky with the more trusted money manager and the share invested risky with the less trusted money manager (18.52 to 14.22). However, also for the subgroup of participants holding unbiased beliefs, the difference (14.22) remains highly statistically significant (p-value=0.003).

In Treatment 2, biased beliefs should have a positive impact on stated indifference costs. When the second manager is more trustworthy than the first, investors are willing to accept costs of 2.23% if they hold biased beliefs, but only 1.53% if they do not have biased beliefs. Nonetheless, indifference costs for both groups are significantly different from 0.75% (p-values=0.000 and 0.032, respectively). Note that we excluded – and exclude in the following subsection – observations in which both managers returned equal amounts in the trust game, because the alternative explanations are void in these cases. In summary, biased beliefs amplify the findings in Treatment 1 and Treatment 2. However, evidence in favor of our hypotheses remains if investors hold correct beliefs.

Are More Risky Investments a Means of Rewarding?

A second reason why investors might invest more risky with more trustworthy money managers is that they use the risky investment as a reward. While this reciprocity motivation is interesting in itself, it would describe a different channel than that modeled by Gennaioli et al. (2015). To control for such motivation, we ask participants whether they invested risky with more trustworthy money manager because they wanted to reward them. Possible answers are "Yes", "No", and "I do not know". Half of the participants

(n=58) stated that they wanted to reward more trustworthy managers when investing more risky. On the other hand, 36 participants ("I do not know" answers excluded) stated that rewarding did not motivate their investment choices. Contrasting the choices of both subgroups in Treatment 1 reveals the expected pattern: On average, rewarding investors invest 21.45% more risky with the more trustworthy manager, whereas non-rewarding investors invest only 8.69% more risky with the more trustworthy manager. While the former is significant at the 1%-level, the latter is marginally insignificant (p-value=0.137). In Treatment 2, the reward motivation leads to higher indifference costs for investments with more trustworthy money managers. Rewarding investors indicate indifference costs of 2.22%, while non-rewarding investors indicate indifference costs of 1.58%. Again, however, these costs are significantly higher than the costs (0.75%) charged by less trustworthy managers (p-value=0.000 and 0.022, respectively). Evidence from both treatments points out to "rewarding for trustworthiness" as one of the drivers of investors' investment choices. However, even without this motivation, results are still in line with our hypotheses.

Finally, we investigate whether the difference of the share invested risky with either money manager varies significantly between participants holding correct and biased beliefs, and between participants with and without reward motivation. For this purpose, random effects regressions are estimated.[12] In these regressions, we can also check for any interaction between biased beliefs and reward motivation. For Treatment 1, Table 3.7 shows results of a regression with Δ *Risky More* $-$ *Risky Less* as dependent variable. This variable is calculated as the amount invested risky with the more trustworthy money manager minus the amount invested risky with the less trustworthy money manager. Results of random effects tobit regression can

[12] Here we cannot report fixed effects estimates, as the independent variables of interest – dummies for biased beliefs or reward motivation – are time invariant.

be found in Table B.9 in the appendix. Independent variables are a dummy equal to 1 if investors hold biased beliefs (*Biased Beliefs*), a dummy equal to 1 if investors stated that they were motivated by rewarding for trustworthiness (*Reward Motivation*), and an interaction term of both dummies (*Biased Beliefs*×*Reward Motivation*).[13] If higher trustworthiness were to lower investors' anxiety of investing with a money manager, the constant in this regression should be positive and significant. This is exactly what we find. On average, investments with more trustworthy money managers are 17.5% more risky, despite higher associated costs. On the other hand, neither *Biased Beliefs* nor *Reward Motivation* significantly influence differences in investment choices. Hence, differences in subgroups' investment choices, as observed in univariate tests, seem to be insignificant.

For Treatment 2, Table 3.8 reports results of a regression with *Indifference Costs* as dependent variable. Random effects tobit regression are shown in Table B.10 in the appendix. Independent variables are the same as before. Under our hypotheses, the constant in the regression should be positive and significantly different from the low fees of 0.75%. That is, investors should be willing to accept higher costs for a risky investment made with a more trustworthy money manager. Regression results are as expected: On average, investors are willing to accept costs of 1.79% (p-values of 0.001 and 0.056 when compared to 0 and 0.75, respectively) when investing with a more trustworthy money manager. Coefficients for *Biased Beliefs*, *Reward Motivation*, and the interaction of both are not statistically significant. Thus, even after controlling for potentially confounding factors, evidence from Treatment 2 supports the Money Doctors theory.

[13] Observations with "I do not know" as answer to the control questions are excluded in this analysis.

Table 3.7: Risky Share – Robustness

This table reports regression results with Δ *Risky More* $-$ *Risky Less* as dependent variable. It is calculated as the share of wealth invested risky with the more trustworthy money manager minus the share of wealth invested risky with the less trustworthy money manager. The regression accounts for unobserved individual heterogeneity through random effects. *Biased Beliefs* is an indicator variable equal to 1 if participants stated that they believed that more trustworthy money managers could deliver better investment performance. *Reward Motivation* is an indicator variable equal to 1 if participants stated that they invested more risky with more trustworthy money managers because they wanted to reward them.
Standard errors clustered at the individual level are shown in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

	Random Effects
Biased Beliefs	-12.98
	(11.48)
Reward Motivation	1.289
	(8.862)
Biased Beliefs×*Reward Motivation*	18.46
	(13.27)
Constant	17.46**
	(8.633)
Observations	322
Cluster-robust S.E.	YES
Round FE	YES
$R^2_{overall}$	0.057

Table 3.8: Indifference Costs – Robustness

This table reports regression results with *Indifference Costs* as dependent variable, for cases in which the second money manager is more trustworthy than the first money manager. The regression accounts for unobserved individual heterogeneity through random effects. *Biased Beliefs* is an indicator variable equal to 1 if participants stated that they believed that more trustworthy money managers could deliver better investment performance. *Reward Motivation* is an indicator variable equal to 1 if participants stated that they invested more risky with more trustworthy money managers because they wanted to reward them.
Standard errors clustered at the individual level are shown in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

	Random Effects
Biased Beliefs	0.216
	(0.637)
Reward Motivation	0.002
	(0.688)
Biased Beliefs×*Reward Motivation*	0.591
	(0.844)
Constant	1.787***
	(0.543)
Observations	324
Cluster-robust S.E.	YES
Round FE	YES
$R^2_{overall}$	0.042

Do Participants React to Arbitrary Information?

Lastly, a possible objection to our results is that participants may not interpret higher amounts returned in the trust game as sign of higher trustworthiness. More subtle, this objection would mean that our results could simply be due to participants reacting to arbitrary information. In other words, replacing the amount returned in the trust game with irrelevant information might produce similar results. To control for this objection, we ask whether participants interpreted higher amounts returned in the trust game as signal of higher trustworthiness at the end of the experiment. Possible answers are "Yes", "No", and "I do not know". The manipulation check indicates that only a fifth of participants (20.18%) do *not* associate higher amounts returned in the trust game with higher trustworthiness. For the majority of participants (64.04%), the manipulation through the trust game appears to have been effective. Nonetheless, we check whether both subgroups behave differently. In general, effect sizes should be greater for the subgroup of participants affirming the manipulation question. In both Treatment 1 and Treatment 2, tests of the variable(s) of interest confirm this hypothesis. In Treatment 1, participants answering "Yes" to the manipulation check on average invest 17.93% more risky with the more trusted money manager (p-value=0.000). Participants answering "No" to the manipulation check, on the other hand, only invest 10.94% (p-value=0.056) more risky with the more trusted money manager. In Treatment 2, participants answering "Yes" to the manipulation are on average willing to accept costs of 2.10%. Participants answering "No" to the manipulation check, however, are only willing to accept costs of 1.53%. For both groups, costs are significantly different from 0.75% (p-values=0.000 and 0.054, respectively). In summary, these results do not corroborate the objection that our general findings are driven by participants just reacting to some arbitrary information.

3.6 Conclusion

This experimental study provides a direct test of the Money Doctors theory. Our findings support the notion that trust is an important component for delegated investing. Even at higher costs, investors take more risk when investing through a money manager who can be trusted. Vice versa, investors are willing to accept higher costs for investments made through more trustworthy money managers. The larger the spread between managers' trustworthiness, the larger the observed effects. Collectively, our study highlights a positive aspect of delegated investing. Although investors would be best off with high risk taking at low costs, they are still better off with higher risk taking at higher costs, if they trust their money manager. In short, our study identifies trust as the "substantial intangible benefit" Bergstresser et al. (2009, p.4129) suspect but cannot observe. Trust may thus be the "saving grace" for a delegation and advice industry whose benefits have been severely doubted in several studies (see e.g., Bergstresser et al., 2009; Mullainathan et al., 2012; Hackethal et al., 2012; Hoechle et al., 2017, 2018).

Furthermore, our experiment points to a reward mechanism as another potential channel why trustworthy money managers may be able to charge higher fees *and* attract more funds. While it appears plausible that investors want to reward trustworthiness, it is different from the trust-modified risk aversion mechanism proposed by Gennaioli et al. (2015). Determining the precise influence of the reward mechanism is left for future research..

Chapter 4

Algorithm Aversion in Financial Investing *

4.1 Introduction

Increasing digitalization and automation of processes in all parts of society have sparked a debate on whether people are willing to rely on algorithms. We use the term *algorithm* for any automated formula or rule that is used to process data. In a recent experiment, Dietvorst et al. (2015) show that humans prefer to rely on predictions made by a human rather than an algorithm, even if the latter performs better. Additionally, participants are quicker to lose confidence in the algorithm than the human once they observe an error. Dietvorst et al. (2015) explain this behavior with *algorithm aversion*.

With the emergence and growth of robo-advisors,[1] and with major fund

* Authors: Maximilian Germann and Christoph Merkle. Maximilian Germann is at the University of Mannheim. Maximilian Germann: L9, 1-2, 68161 Mannheim. Email: germann@bank.bwl.uni-mannheim.de. Christoph Merkle is at the Kühne Logistics University. Christoph Merkle: Grosser Grassbrook 17, 20457 Hamburg. Email: christoph.merkle@the-klu.org. For valuable comments, we thank Peter Bossaerts, Noboyuki Hanaki, Martin Weber, participants of the Experimental Finance 2018, the ESA world meeting 2018, and seminar participants at the University of Mannheim. Funding from the Graduate School of Economic and Social Sciences (GESS) at the University of Mannheim is gratefully acknowledged. All remaining errors are our own.

[1] www.ft.com/content/6b2d5490-d9bb-11e6-944b-e7eb37a6aa8e (retrieved July 20, 2018).

companies shifting towards more cost-efficient, automatized products,[2] attitudes towards algorithms become increasingly important for the finance industry. In the words of practitioners "over many years, the fund industry has operated with a false sense of security, assuming that algorithms and computing power would digitize and revolutionize trading, but that the right products would ultimately be selected by humans."[3] Today, robo-advisors like Betterment or Wealthfront already report assets under management worth several billion U.S. dollars.[4] New FinTech companies heavily relying on technology are founded all over the world.

These developments show that the financial industry offers many applications for algorithms, be it in trading, asset management, or financial advice. This makes it vital to understand how algorithm aversion – or the lack thereof – might affect financial decision making. It will determine whether new competitors to traditional financial intermediaries will remain in a niche market and cater the tech-savvy, or will gain wide acceptance in the general population. Our study therefore aims to answer two key questions: 1. Are human investors less likely to invest in a portfolio selected and "managed" by an investment algorithm than in a portfolio selected and managed by a human fund manager? 2. Are they quicker to abandon the investment algorithm than the human fund manager if performance (absolute or relative) is poor?

To answer these questions, we conduct an experimental study consisting of two parts, an online survey and a laboratory experiment. The survey is administered several weeks prior to the experiment and elicits beliefs about the strengths and weaknesses of algorithms relative to human judgment. It

[2] `www.bloomberg.com/news/articles/2017-03-28/blackrock-said-to-cut-jobs-fees-in-revamp-of-active-equity-unit` (retrieved July 20, 2018).

[3] See Bain & Company (2017) report "*Asset-Management: Erfolgsformel gesucht*", p.9 (translated from German by the authors).

[4] US$ 13.5bn for Betterment and US$ 10bn for Wealthfront as of March 2018. Data retrieved from Google Finance as of May 22, 2018.

includes explanations for algorithm aversion suggested in the literature and is designed to contribute to our understanding of participants' preferences for either intermediary. We implement a distance in time between the survey and the main experiment to avoid a direct effect of this task on experimental decisions.

In the laboratory experiment, we ask participants to choose between a human fund manager and an investment algorithm to invest for them. Both financial intermediaries then repeatedly decide to invest in either a risk-free bond or a risky stock. The stock can be in a good state or a bad state, which is slowly revealed by its performance (the design is adapted from Kuhnen, 2015). The pre-programmed algorithm strictly applies Bayes' Law and decides accordingly, while the human fund manager has complete discretion over the decisions to make. In the experiment, the participant with the best financial literacy and numeracy skills takes over the role of the human fund manager. The incentives of participants depend on the payoffs generated by their selected intermediary.

Importantly, the selection of the financial intermediary is repeated ten times, which allows us to study initial preferences without much information, as well as the reaction to the outcomes the intermediaries produce. The experimental design gives rise to frequent (ex-post) mistakes that occur even if a rational strategy is applied. We can thus also examine the consequences of such mistakes on the preferences of participants. To determine how strong these preferences are, we apply several different fee schemes, which render one intermediary more expensive than the other.

We find no evidence for algorithm aversion. In the initial choice with equal fees 56% of participants decide to invest with the algorithm. If fees differ between the intermediaries, participants mostly (>80%) choose the intermediary with the lower fee. They apparently do not believe that one will outperform the other by a high enough margin to justify the higher fee. Indeed, human fund managers perform quite well and register only slight underperformance relative to the Bayesian algorithm.

There is no strong trend in the proportions choosing either intermediary. Once investors learn about investment choices and outcomes of the human fund manager and the algorithm, they focus on performance. Choices are strongly influenced by cumulative past performance whereby the highest weight is given to the most recent performance. In their reaction to performance, participants do not discriminate between intermediaries. In particular, they do not respond differently to (ex-post) mistakes by the intermediaries, rejecting the idea of trust in an algorithm eroding more quickly. We thus do not find support for the two major predictions of algorithm aversion – general preference for human judgment and adverse response to errors by an algorithm – in the domain of financial decision making.

The survey provides some insights into the reasons for this result. Participants believe in the ability of an algorithm to generate higher returns than a human. They also think that an algorithm is better able to learn. On the contrary, they see advantages of the human in using qualitative data and dealing with outliers. Regarding the relationship between the intermediaries they view an algorithm as an aid rather than a competitor to a human fund manager. Of these attributes only the belief about returns has explanatory power for observed choices in the experiment. This is in line with participants paying most attention to returns, as the experimental setting provides little opportunity to play off strengths in analyzing qualitative data or outliers.

We further establish that in focusing mainly on returns, participants fail to distinguish between skill and luck. They take into account the outcome of an investment but not whether an investment decision was reasonable ex ante (outcome bias). They will thus be slow in recognizing true skill, which might explain the absence of a strong trend towards the algorithm over time. The random component in outcomes introduces noise which prevents a small performance difference to be noticeable by participants (consistent with Heuer et al., 2017).

Our results have several implications for the financial industry. First, algorithm aversion is absent in general which suggests that products based on algorithmic strategies should find a large market of interested clients. Second, however, preferences can be quite sticky as the investment proportions in our experiment do not change much. It might need a long performance history or large performance difference to convince people initially in favor of a human fund manager. Third, the expressed view of algorithms serving as an aid suggests that the most preferred intermediary could be a human manager assisted by an algorithm. Even though people are forgiving in case of errors, they might view human intervention in extreme scenarios favorably.

The remainder of this paper is structured as follows: Section 4.2 provides an overview of the literature on algorithm aversion from which we derive hypotheses for the experiment. Section 4.3 presents the experimental design and participants. In section 4.4, we report and discuss the main results, before a final section concludes.

4.2 Literature and Hypotheses

Algorithm aversion is neither a new concept, nor limited to a particular domain. Researchers as early as Meehl (1954) discuss the superior performance of algorithms in various prediction tasks. In comparing statistical and clinical prediction, this line of research pits a statistical algorithm against a human clinician. Dawes (1979) confirms the superiority of even improper specified algorithms and already reports common objections against the use of algorithms. These include technical issues raised against the particular methodology applied, psychological misperceptions of performance, and ethical problems with algorithms deciding in sensitive areas.

In meta studies, Grove and Meehl (1996) and Grove et al. (2000) corroborate the hypothesis that for many forecasting tasks, algorithms are better suited than humans. The tendency to discount algorithms has been documented in a variety of settings as well. In medicine, recommendations coming from a physician are rated higher than recommendations from a computer system or from a physician aided by a computed system (Promberger and Baron, 2006; Shaffer et al., 2013). In matters of personal taste, Yeomans et al. (2017) provide evidence that although an algorithm outperforms humans at recommending jokes that participants rate funny, they still prefer to receive joke recommendations from other humans.

A first hypothesis emerging from this literature is that algorithm aversion exists and that people shy away from using algorithms, most likely also in financial decisions:

Hypothesis 1: A larger fraction of participants will initially select to invest with the human fund manager than with the investment algorithm.

Hypothesis 1a: Participants' willingness-to-pay for the human fund manager (i.e., fees) will initially be higher than their willingness-to-pay for the algorithm.

Hypothesis 1a is added as a measure for the strength of preference for a financial intermediary. By attaching a price to investing, we are able to determine at what price people are indifferent between investing with the human fund manager and the algorithm.

Dietvorst et al. (2015) analyze algorithm aversion in a systematic way. Experiment participants observe predictions of human judges and algorithms in domains such as MBA student performance or U.S. air traffic. In several conditions, the amount of information participants observe is varied. They can either tie their incentives to the performance of an algorithm or to a human judge (which is in some conditions themselves and sometimes another participant). Dietvorst et al. (2015) find that algorithm aversion is most pronounced after seeing the algorithm perform, even if this performance is superior to the human judge. They conclude that people are particularly troubled by seeing the algorithm err and abandon it in response.

We can thus specify the expected reaction to seeing the investment algorithm perform and to mistakes that is makes:

Hypothesis 2: Participants will disregard higher performance of the algorithm and continue to favor the human fund manager after outcomes are observed.

Hypothesis 3: After mistakes by the algorithm, participants will be more prone to switching from the algorithm to the human fund manager than vice versa.

In a follow-up article, Dietvorst et al. (2018) find that allowing participants to modify the forecast of an algorithm makes them considerably more likely to use it. At the same time the modification option increases participants' satisfaction with and belief in the algorithm. There exists further evidence for situations in which humans do rely on algorithms. In a task of evaluating statements and reducing them to a logical problem, participants rely more on algorithms than on other people (Dijkstra et al., 1998), or even themselves (Dijkstra, 1999). As Logg (2017) elaborates, confounding factors in existing studies make it difficult to establish a clear case for or against algorithm aversion. She shows that participants prefer advice from algorithms over advice from other people, and that they particularly prefer advice from algorithms for objective decisions (e.g., estimating air traffic), whereas they prefer advice from humans for subjective decisions (e.g., recommending jokes).

Financial decision making might be perceived as a domain of objective decision making, which would work against Hypotheses 1-3. Little attention has yet been paid to algorithm aversion in a financial context. To our knowledge, there is only a handful of studies on the role of algorithm aversion in finance. In an experiment, Önkal et al. (2009) show that stock price forecasts provided by a statistical forecasting method are more severely discounted than forecasts by a human expert. Based on fund flow data, Harvey et al. (2017) report that algorithm-based ("systematic") hedge funds receive less inflows than actively managed ("discretionary") hedge funds. However, they do not find a performance gap justifying this aversion towards algorithm-based hedge funds.

Most recently, Hodge et al. (2018) provide experimental evidence that investors are more likely to follow the advice of a robo-advisor in an anonymous setting, while they are more likely to follow the advice of a human advisor when advisors are humanized (e.g., by adding a name). Unlike in our study, however, their setting does not feature actual human advisors, nor do the human or the algorithm advisor act in the experiments. D'Acunto et al. (2018) study the characteristics of investors who adopt robo-advising tool and find that they are demographically similar to non-adopters, but have larger portfolios, trade more, and achieve higher risk-adjusted performance. Following their interpretation, more sophisticated investors are more likely to adopt the algorithm.

Our study contributes to this emerging literature on the presence (or absence) of algorithm aversion in financial decision making in multiple ways. To our knowledge, we are the first to use an experimental setting in which both the investment algorithm and human fund manager act and are observed to act. Due to the straightforward design, we are able to exclude many of the confounding factors that make conclusions about algorithm aversion otherwise difficult (Logg, 2017). By presenting the decisions and investment outcomes to participants, we generate rich data on how they respond to performance and to mistakes, which has been described as one of the key elements of algorithm aversion. Finally, we explore the underlying beliefs that shape people's preferences for a human or algorithmic intermediary.

4.3 Experimental Design and Participants

We design an experiment that consists of two parts, an online survey and a laboratory experiment. We need to separate the parts to avoid spill-over effects from the questionnaire to the experiment or vice versa. As for practical reasons the payment of participants takes place at the laboratory stage, the survey is run beforehand. A survey link is sent out to participants about four weeks before the scheduled experiment and the survey closes three weeks before the experiment. Participants are required to complete the survey and receive an individual code in order to partake in the laboratory experiment.

4.3.1 Online Survey

The aim of the survey is to understand perceptions of algorithms and human managers that may affect algorithm aversion in financial decision making. There are several aspects of decision making and data processing for which either an algorithm or a human might be better equipped. We draw on the literature to identify relevant dimensions for which we measure participants' perceptions. Based on this we formulate statements that one intermediary is better than the other in a particular dimension (see Table 4.1 for a list of these statements). To avoid acquiescence bias, there is an inverted version of each statement and one of the two is presented at random. Participants express their agreement on a five-point Likert scale ranging from “strongly disagree” to “strongly agree.”

A straightforward question is whether participants expect an investment fund run by an algorithm or a human fund manager to achieve higher returns (statement one). One objection against algorithms is their supposed inability to learn or to improve through experience (Dawes, 1979; Highhouse, 2008), which we capture in statement two. It has been suggested that algorithms are unable to incorporate qualitative data and to react to unexpected events or outliers (Grove and Meehl, 1996), which we address in statements three and six. There also might be different perception on intermediaries’ ability to identify relevant factors and to integrate this data (statements four and five, Dawes, 1979).

Of specific interest to the industry should be whether algorithms are expected to compete with (and probably replace) human fund managers, or whether they are perceived as an aid to human managers (statement seven). It is unclear whether a combination of the two intermediaries is considered superior to a single one (Shaffer et al., 2013).

In addition, we elicit self-reported measures for trust and risk aversion (Falk et al., 2018), and economic knowledge (van Rooij et al., 2011). Some of these factors might interact with algorithm aversion, as for example more sophisticated investors have been suggested to rely more on algorithm (D’Acunto et al., 2018). The impact of trust and risk-aversion will depend on which intermediary is considered to be more trustworthy and less risky.

Table 4.1: Survey Questions on the Perception of Algorithms in Finance

The table shows the exact wording of the two alternative survey questions for each statement presented to participants, labeled x.1 and x.2, respectively. Which version was presented to participants was randomized. Answers are elicited on a Likert scale ranging from 1 to 5, where 1 was labeled "strongly disagree" and 5 was labeled "strongly agree".

1.1	On average, investment funds run by fund managers achieve higher returns than investment funds that are based on investment algorithms.
1.2	On average, investment funds based on investment algorithms achieve higher returns than investment funds that are run by fund managers.
2.1	Fund managers are better able to adapt their investment approach in response to past success or failure than are investment algorithms.
2.2	Investment algorithms are better able to adapt their investment approach in response to past success or failure than are fund managers.
3.1	Fund managers are better able to interpret qualitative or subjective data than investment algorithms.
3.2	Investment algorithms are better able to interpret qualitative or subjective data than fund managers.
4.1	Fund managers consider a wider range of factors for their investment decisions than investment algorithms.
4.2	Investment algorithms consider a wider range of factors for their investment decisions than fund managers.
5.1	Fund managers are better at correctly assessing the relevance of factors for investment decisions than investment algorithms.
5.2	Investment algorithms are better at correctly assessing the relevance of factors for investment decisions than fund managers.
6.1	Fund managers are better able to react to unexpected events such as a financial crisis than investment algorithms.
6.2	Investment algorithms are better able to react to unexpected events such as a financial crisis than fund managers.
7.1	Investment algorithms are rather a competitor of fund managers than an aid to fund managers.
7.2	Investment algorithms are rather an aid to fund managers than a competitor of fund managers.

4.3.2 Laboratory Experiment

In the laboratory experiment, we simulate financial decisions in the context of delegated investments. This provides a simple setup in which an algorithm can directly compete with a human fund manager. Dietvorst et al. (2015, p.114) define the term algorithm to "encompass any evidence-based forecasting formula or rule. Thus, the term includes statistical methods, decision rules, and all other mechanical procedures that can be used for forecasting." For an investment context, we derive the following criteria for the algorithm: 1) Once constructed, it must act independently of a human, 2) it must be strictly rule-based, and 3) its recommended actions must be executed automatically.

The investment decisions made by the financial intermediaries are repeated choices between a risk-less bond and a risky stock. Our experimental design follows the gain condition in Kuhnen (2015). There are two securities on a market, one of which is a bond paying 3€ for certain. The other is a stock paying either 5€ or 1€. The probability for the high payoff is either 70% (good state) or 30% (bad state). Whether the stock happens to be in a good or bad state is randomly determined with equal probability at the beginning of a block of trials. A trial hereby represents one realization of payoffs for the two securities. The state of the stock is fixed for a block of six trials.

An important difference to the original design is that participants do not decide themselves in which security to invest, but instead choose the intermediary they want to invest with. Intermediaries are presented as investment funds managed either by a human fund manager or by an investment algorithm. The algorithm is programmed to maximize expected return following Bayes' law. In case expected returns are equal for both securities, it chooses randomly. The algorithms' goal of maximizing expected return is disclosed to participants. The exact mechanism, however, is not disclosed.

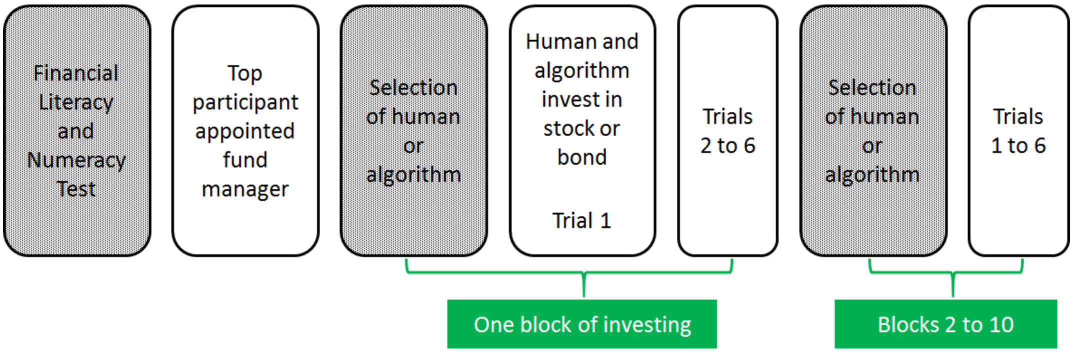
This is consistent with the literature on algorithm aversion which usually does not explain how algorithms work exactly. Likewise, in reality the mechanics of an algorithm would typically not be disclosed by fund companies. Moreover, too much information would be counterproductive to learn about participants' existing dispositions towards investing with an algorithm or a human. While a concern is that participants believe an algorithm constructed by the experimenters must be superior, prior research finds algorithm aversion despite this fact (Dietvorst et al., 2015).[5]

The human fund manager represents an actual human being selected from participants. This avoids simulating human decisions, which would make them appear similar to an algorithmic decision. Participants complete a set of eight (advanced) financial literacy questions (van Rooij et al., 2011) and a four-question numeracy test (Berlin Numeracy Test, see Cokely et al., 2012). Known to participants, the participant with the highest score is anonymously appointed as the human fund manager. This is to ensure that the other participants view this individual as financially competent even though he or she is not a professional fund manager. In case of ties for highest score, one of the tied participants is selected randomly.

After the role of the fund manager is assigned, participants decide whether they want to tie their incentives for the first block of trials to the human or to the algorithm. Their decision is fixed for this block and can be revised only after the block ends. They then observe the choices and the outcomes of the human fund manager and the algorithm. In each trial, the human and the algorithm invest in the stock or the bond and observe the outcomes of both securities. After a block of trial ends, a new state for the stock is drawn and participants can change their preferred intermediary. They are shown a summary of the aggregated payoffs of both intermediaries

[5] Dietvorst et al. (2015) use light deception and tell participants that "the admissions office had created a statistical model that was designed to forecast student performance."

Figure 4.1: Illustration of Experimental Design



for all previous blocks. This is repeated for a total of ten blocks. For an overview of the experimental design see Figure 4.1.

The experimental design allows for (ex-post) mistakes by the human manager and the algorithm, as even perfect information will result in the selection of the asset with the inferior payoff in 30% of the cases. The design thus enables us to study how participants react to mistakes by the intermediaries. It further avoids several of the confounding effects identified in the literature (Logg, 2017).

In addition to the decision for one of the intermediaries, we also measure the strength of participants' preferences. Investing with the human fund manager always costs a fixed fee of 2€. Investing with the algorithm costs a fee of either 0, 1, 2, 3, or 4€. For each of the five fee-combinations we ask which intermediary a participant would prefer (see appendix C for screenshots of the experiment). One fee combination is then randomly drawn and the actual decision for this combination is used. Participants can thus express a preference for either intermediary in the range from -2€ to +2€.[6]

All participants are incentivized based on the outcomes of their decisions. Participants acting as investors receive the payoff generated by their chosen intermediary minus fees. To avoid wealth effects only one of the blocks is randomly drawn for payment. Participants in the role of the fund manager

[6] Fees are *not* transferred to the human fund manager (or the algorithm) to avoid issues of reciprocity.

receive the gross payoff they achieved in a random block. This also provides incentive to become fund manager.[7] The expected payoff for a block of trials amounts to 6·3=18€, the expected fees are 2€. The laboratory experiment concludes with a short questionnaire asking participants how they rate the human fund manager's and the algorithm's investing capability, and an open-ended question regarding participants' primary motivation when choosing between both intermediaries.

4.3.3 Participants

The experiment was implemented using z-tree (Fischbacher, 2007), and the survey was run on the research platform SoSci survey. Participants were invited to the MLab of the University of Mannheim via the recruiting software ORSEE (Greiner, 2015). In total, 114 participants took part in the laboratory experiment, 107 of which could be matched to survey data. To preserve anonymity, the matching was done via an individual code generated in the survey, which some participants could not recall. We nevertheless allowed these participants to enter the main experiment.

We aimed for twelve sessions of ten participants. Due to no-shows some sessions had fewer but never less than eight participants. This means that we ended up with 12 unique human fund managers (one per session) each with seven to nine investors. The small sessions were intended to generate more variation in the human fund manager which implies more independent clusters (i.e., session fixed effects) and reduced risk that results might be driven by extreme strategies of one particular fund manager.

[7] There might be concerns that participants do not want to stand out from their peers and become fund manager. We make it clear in the instructions that the fund manager is appointed anonymously and not revealed to anyone. From the results in the literacy and numeracy tests, we conclude that participants do compete for the role.

Participants were 22.8 years old on average, were predominantly female (58%), and a quarter had already invested in stocks (24.3%). The average payoff from the experiment was 16.79€ for participants in the role of investors and 18.83€ for participants in the role of fund managers. The payoff range was between 4€ and 30€. Considering an average experiment duration of approximately 40 minutes (and an additional 5-10 minutes for the survey), the payoff for participation was substantially higher than the laboratory average and German minimum wage.

4.4 Results

4.4.1 Survey Results

We begin with the analysis of the survey responses on how algorithms are perceived in the financial context. As two reversed versions of each statement are randomly presented, we rescale all answers so that a value of 5 expresses the algorithm is strongly favored, and a value of 1 that the human is strongly favored. Consequently, a value of 3 indicates a neutral perception. We do find a sometimes significant effect of the version of the question shown to participants (acquiescence bias), which we eliminate by counterbalancing versions. As intended, the questions capture different dimensions: Overall, answers have low and mostly insignificant correlations (see appendix Table C.1).

Table 4.2 summarizes participants' perceptions along the dimensions we explained before. On average, investment algorithms are expected to deliver better investment performance than human fund managers. In addition, investment algorithms are viewed to be slightly better at adapting their investment approach. Not surprisingly, however, human fund managers are perceived to make better use of qualitative data. No difference is found for both

Table 4.2: Perceptions of Algorithms in Finance

This table shows how participants perceive algorithms in finance. To avoid acquiescence bias, for each dimension there were two versions of the statement one of which was randomly presented. The exact wording of these questions is shown in Table 4.1. Answers are given on a Likert scale ranging from 1 to 5, where 1 was labeled "strongly disagree" and 5 was labeled "strongly agree." Values shown here are combined values for both versions, with the value of 5 indicating a perception in favor of the algorithm. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively. All tests are two-sided t-tests against a neutral response of 3.

		N	μ	σ	Min	Max
1	Returns	106	3.26***	0.82	1	5
2	Learning	107	3.19*	1.00	1	5
3	Qualitative data	107	2.69***	1.00	1	5
4	Data aggregation	107	3.08	1.05	1	5
5	Data weighting	107	3.06	0.90	1	5
6	Dealing with outliers	107	2.70***	1.08	1	5
7	Aid rather than competitor	107	3.56***	0.81	2	5

data aggregation and data weighting. When it comes do dealing with outliers, such as financial crises, human fund managers are again viewed more capable.

Overall, participants' perceptions of algorithms in finance appear quite reasonable. Some correspond to the views expressed in the literature such as dealing with qualitative data and with outliers. In the domains probably most relevant for the laboratory experiment, the expected return and the ability to learn, participants view algorithms as better than humans. This means their perceptions do not unambiguously support all of the proposed reasons for algorithm aversion. Particularly important for practitioners, we find that participants view algorithms as an aid to instead of a competitor of fund managers.

Lastly, participants are somewhat inclined to take financial risks (5.4 out of 10, *SD*=2.0) and report to have an average level of general trust (4.7 out of 10, *SD*=2.3). As many of the recruited participants are students of business or economics, they rate their economic knowledge above average (4.5 out of 7, *SD*=1.1). As investors in stocks and mutual funds typically represent an economically rather sophisticated group as well, the participant group should be a relevant one even though it lacks investment experience.

4.4.2 Investment Decisions by Financial Intermediaries

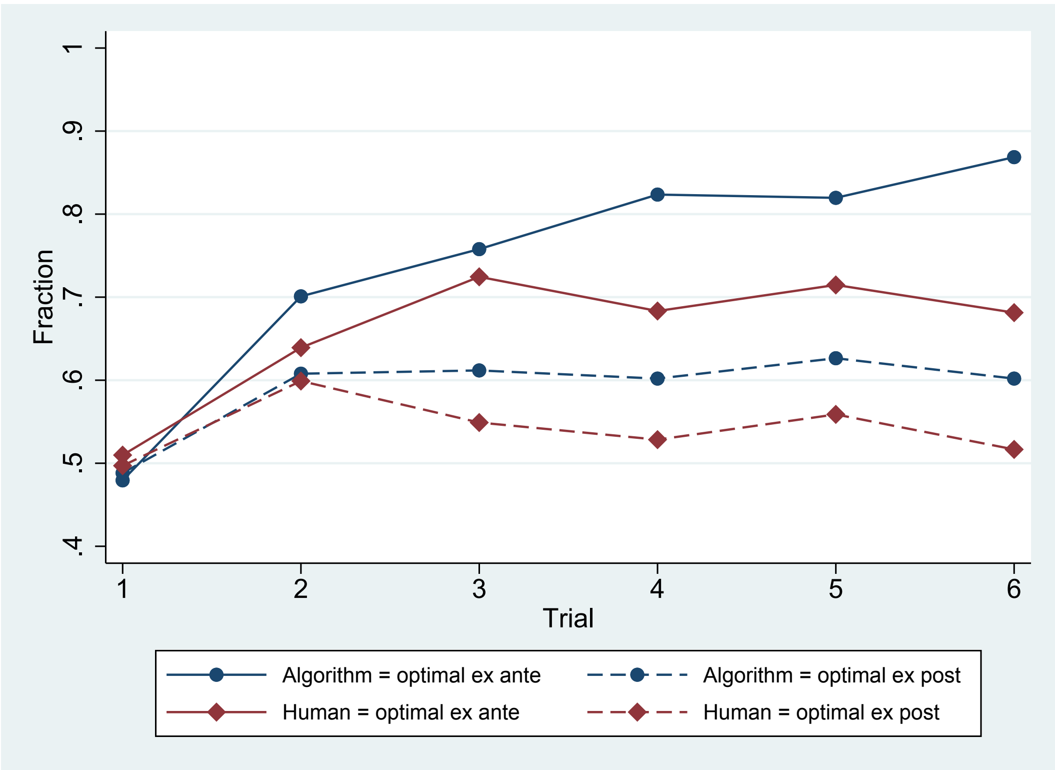
While the selection of an intermediary is in the center of this study, we first report on the investment behavior of the intermediaries. The algorithm maximizes returns following Bayes' law, which is relatively simple in the employed experimental setting. In the first trial without any information, it selects either the stock or the bond at random. In any later trial, it selects the stock if the good outcome of the stock (payoff of five) was observed more often than the bad outcome (payoff of one), and the bond if the bad outcome was observed more often than the good outcome. In case of equal occurrence of both outcomes, the algorithm again selects at random.

Six trials are usually enough to identify the true state of the stock. In the final investment decision, the choice of the algorithm is in line with the true state in 86.9% of the cases. We refer to such a decision as an ex-ante correct decision, because the intermediary selects the asset that is expected to perform better according to the underlying probabilities. Figure 4.2 shows how the fraction of ex-ante correct decisions by the algorithm rises over the course of the trials. However, the fraction of ex-post correct decisions, meaning that the selected asset outperforms the other asset in the following period, is lower, as it is subject to chance. The upper limit for ex-post correct decisions is 70% even for a perfectly informed investor. As the figure shows, the algorithm remains below this limit even the final trials.

The expected investment outcome of a perfectly informed investor would be 22.80€ in the good state (always selecting the stock) and 18€ in the bad state (always selecting the bond). The algorithm on average reaches 21.66€ (good state) and 16.62€ (bad state). These values are the benchmark for the human fund manager. As the algorithm uses the best available strategy, the human fund manager will likely underperform. This is consistent with the literature on algorithm aversion in which the algorithm usually outperforms

Figure 4.2: Correct Decisions by Intermediary

This figure shows the fractions of ex-ante and ex-post optimal decisions by the algorithm and the human. Fractions averaged over all blocks are shown by trial.



the human expert.

By appointing the most financially literate and most numerate participant, the human fund manager should at least be well-equipped to make good investment decisions and might actually come close to the optimal strategy of the algorithm. Participants perform well in the financial literacy quiz. On average, 6.2 out of 8 questions are answered correctly. The numeracy test is harder with on average 1.8 out of 4 correct responses. Adding both scores, participants answer 7.9 questions correctly. As intended, human fund managers perform substantially better with an almost perfect score of 11.7.

Human fund managers' knowledge and abilities also translate into investment decisions. The dashed lines in Figure 4.2 show the fractions of their ex-ante and ex-post correct decisions. From trial two onwards, they are below those of the algorithm but only by on average 11 (ex-ante) and 6 (ex-post) percentage points. Over all 120 blocks (12 human fund managers · 10 blocks), the algorithm outperforms human fund managers by on average 0.58€ per block. As depicted in Figure 4.3, payoff differences between algorithm and human fund manager are skewed to the right. By construction, it is difficult to outperform the algorithm by more than 2€. However, algorithms sometimes significantly outperform the human fund manager.

Human fund managers can only deviate substantially from the algorithm if they do not follow Bayesian logic. A majority of outcomes differing by no more than 2€ suggests that, by and large, human fund managers adopt a Bayesian approach. On average, they make Bayesian investment choices in 4.95 trials per block.[8] As shown in Figure 4.4, in more than 50% of all blocks, human fund managers make Bayesian decisions in every trial. This investment behavior is stable over blocks and does not require initial learning (see Figure C.1 in the appendix).

[8] We generously count any decision as Bayesian in cases where the Bayesian decision is ambiguous (i.e., the algorithm randomizes). If we exclude such decisions, the fraction of Bayesian decisions by human fund managers is reduced to 76%.

Figure 4.3: Relative Performance Investment Algorithm and Human Fund Manager

This figure shows the distribution of cumulated payoff differences (in €) between the algorithm and the human. Payoff differences are cumulated over all trials of one block.

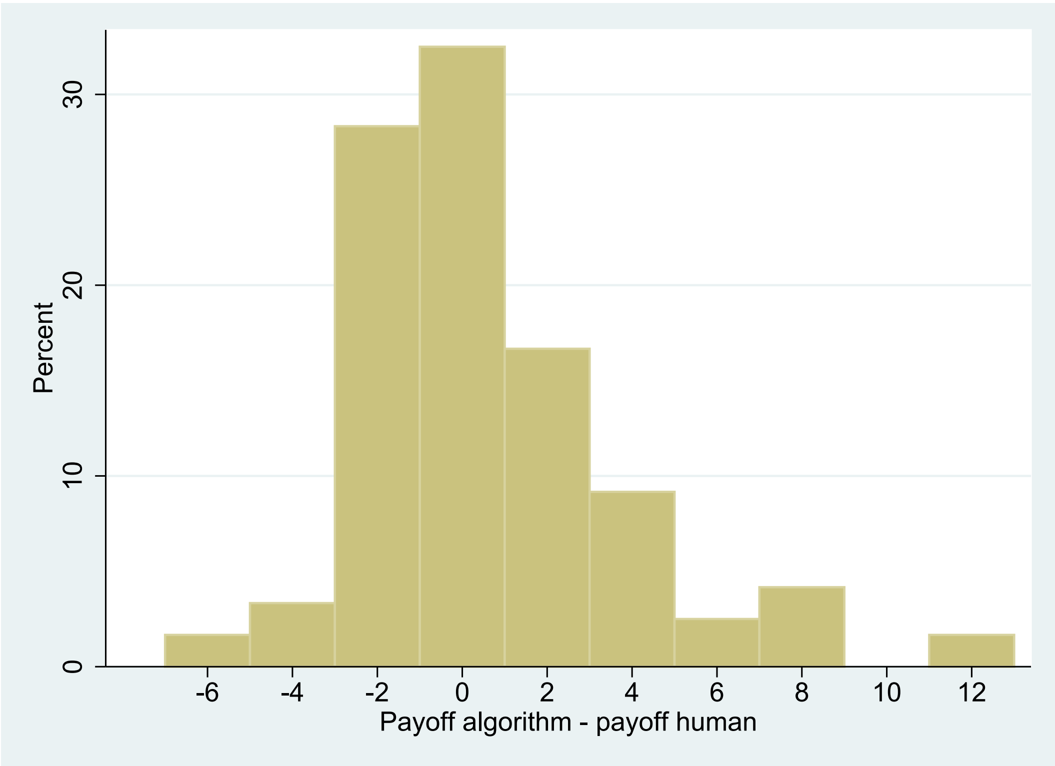
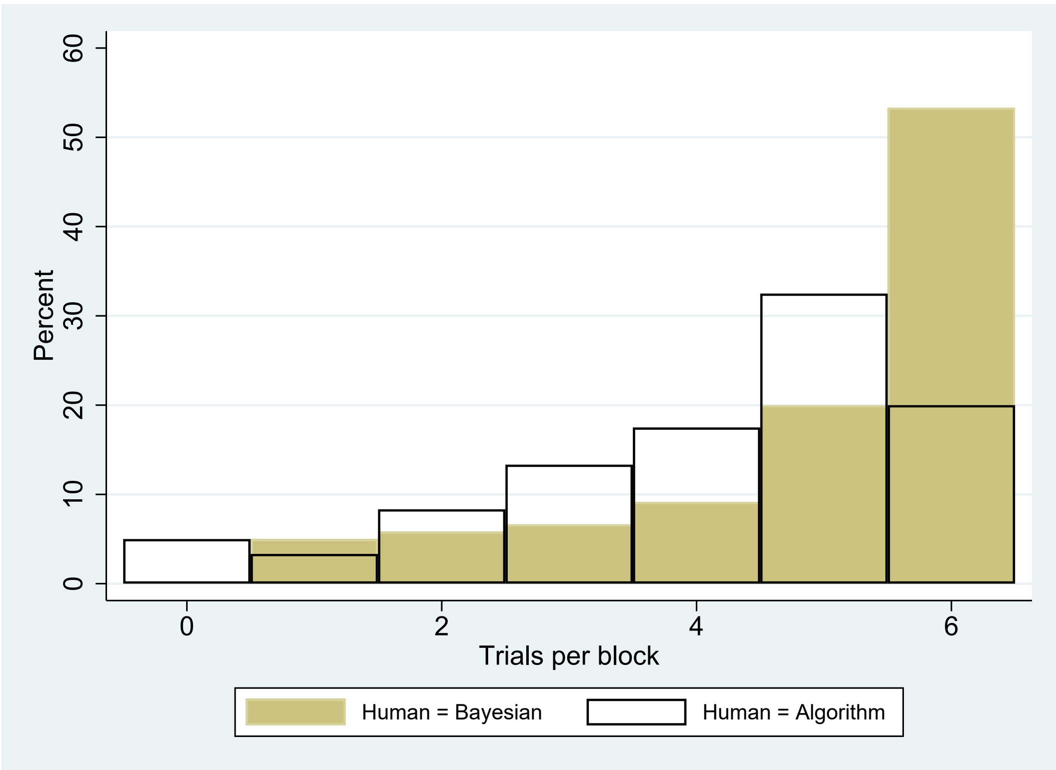


Figure 4.4: Human Fund Manager Choices

This figure shows the distribution of the number of trials in a particular block for which the human invested Bayesian and invested exactly as the algorithm, respectively.



This naturally results in a high number of identical choices between the two intermediaries, which is also documented in Figure 4.4. In most blocks between four and six decisions are the same between the intermediaries. This allows us to investigate participants' choices after blocks in which there was virtually no difference between human fund manager and investment algorithm. Interestingly, we find no evidence for risk aversion on the side of the human managers. If information is ambiguous, they invest into the stock 52.6% of the time. Because both intermediaries invest risk neutrally, risk preferences of participants cannot bias our results (e.g., favoring the more risk-averse intermediary).

4.4.3 Investors' Initial Choice of Intermediary

When analyzing the choices of investors, we distinguish between the first decision at the start of the experiment and all subsequent decisions. Entering the first block of investments, participants have to rely on their predispositions towards the investment algorithm and the human fund manager. There is no information yet available on their performance in the task at hand, and the decision might differ from those after seeing the algorithm perform (Dietvorst et al., 2015). To test hypotheses 1 and 1a, we thus examine investors' initial choice of an intermediary before the first block of investments.

The aim of the algorithm to maximize expected return is part of the experimental instructions. Similar to what is observed in reality, the exact mechanism of the investment algorithm is not disclosed. We avoid any particular reference to its quality.[9] It is common knowledge to participants that human fund managers are selected based on financial sophistication and that they are incentivized based on investment performance. It is thus reasonable for participants to assume that they aim at maximizing performance as well.

[9] This is unlike Dietvorst et al. (2015), who explain to participants "that the model was sophisticated, put together by thoughtful analysts (p.117)." If anything, we should observe stronger algorithm aversion in presence of quality uncertainty.

Although human managers as reported act rather risk neutrally, investors might initially believe that they invest more closely to human investors' (potentially risk-averse) preferences.

When participants first select an intermediary, there is no evidence for algorithm aversion. Our baseline is the choice situation with equal fees, in which 56% of investors choose the algorithm. While this is a slight majority, the proportion is not significantly different from 50% (p-value=0.24). To show the absence of algorithm aversion, however, a general preference *for* the algorithm is not necessary. Our data are unlikely to occur in presence of true algorithm aversion of, e.g., 60% (p-value$<$0.01). We interpret the result as evidence against Hypothesis 1.

Under Hypothesis 1a, we should further find a higher willingness-to-pay for the human fund manager when fees are not equal. This means, we should observe relatively more choices in favor of the human fund manager than the algorithm if their respective fees are higher. However, the observed distribution of choices is almost symmetric. If the human fund manager costs 1€ (2€) more than the investment algorithm, 13% (6%) of investors still prefer the human manager in the initial choice. If the investment algorithm costs 1€ (2€) more than the human fund manager, 15% (5%) of investors prefer the algorithm. These revealed preferences imply an on average 7.5 cents higher willingness-to-pay for the algorithm.[10] We therefore cannot confirm Hypothesis 1a.

The low fraction of participants selecting the more expensive intermediary in the unequal fee combinations suggests that they do not believe any intermediary will outperform the other by a Euro or more. This is interesting, as just one more (ex-post) mistake per block loses 2€ relative to the other intermediary. Open-ended feedback at the end of the experiment supports

[10] This is a coarse calculation as exact switching points (maximum willingness-to-pay) cannot be identified. We instead use the mid-point of the fee interval, at which participants switch.

the importance of fees: Out of 95 participants who state their motivation for choosing between intermediaries, 75 (79%) mention costs as a decisive factor. As a consequence, these choices tend to be very stable over the course of the experiment (see Figure C.3 in the appendix). For the following analyses, we thus focus mainly on investors' choices when fees are equal, as they are more sensitive to developments in the experiment.

For participants in the role of investors with a matching survey (n=95), we regress their initial choice for an intermediary on their perceptions of algorithms, demographics and controls. Table 4.3 shows marginal effects of probit regressions. In a first step, we aggregate the perceptions towards the algorithm by taking the mean of questions 1 to 6 as reported in Table 4.1.[11] We find that the general perception of the algorithm is positively correlated with choosing the algorithm in the equal fee condition (column 1). One step on the five-point scale makes it 26.5% more likely to select the algorithm.

Effects for the individual perceptions (columns 2 and 3) are all positive, but only the belief that the algorithm is able to generate higher returns attains significance. This is consistent with participants viewing this ability as the most important attribute in the experimental task. Of the control variables, being male and having invested in stocks have a negative effect. This might be surprising as men are sometimes seen as more affine to technology. On the other hand, male and active investors are prone to overconfidence, and they may believe that the human fund manager can beat the algorithm. Risk tolerance has a positive effect, suggesting that the algorithm is perceived as the riskier alternative. Turning to the choices when fees are unequal (columns 4 and 5), we find only little effects of the independent variables on the choice of intermediary. As assumed before, decisions in these cases seem to be mostly driven by cost considerations.

[11] The simple mean is highly correlated with the first component of a principal component analysis. We exclude the question on perceiving the algorithm as a competitor or an aid, as the direction of this item is unclear.

Table 4.3: Initital Choice of Intermediary

The table reports probit regression results with the initial choice of intermediary as dependent variable. The binary variable takes a value of 1 if an investor chooses to invest with the investment algorithm. Columns (1) to (3) report results for the choice under equal fees, column (4) shows results for the choice when the algorithm demands a 1€ higher fee, and column (5) shows results for the choice when the human manager demands a 1€ higher fee. Independent variables include responses to questions 1 to 6 as reported in Table 4.1, and an aggregated perception of the algorithm which is the mean across question. *Gender* is an indicator variable (male=1), *Invested in stocks* is an indicator whether a participant has invested in stocks (=1). *Risk tolerance* and *Trust* are as defined in Falk et al. (2018) and range from 0 to 10. *Self-reported knowledge* is participants self-reported economic knowledge ranging from 1 (lowest) to 7 (highest). *Knowledge score* is the total score obtained from the financial literacy and numeracy task. Reported are marginal effect with robust standard errors in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

	Equal fees			Fee algo +1	Fee human +1
	(1)	(2)	(3)	(4)	(5)
Perception of Algorithm (aggr.)	0.265				
	(0.123)**				
Returns		0.109	0.126	0.057	–0.027
		(0.060)*	(0.063)**	(0.040)	(0.041)
Learning		0.022	–0.004	0.013	0.056
		(0.052)	(0.050)	(0.036)	(0.030)*
Qualitative Data		0.032	0.037	0.042	0.002
		(0.051)	(0.049)	(0.036)	(0.025)
Data Aggregation		0.041	0.036	0.009	–0.015
		(0.048)	(0.048)	(0.035)	(0.023)
Data Weighting		0.022	0.029	0.055	0.071
		(0.059)	(0.055)	(0.039)	(0.029)**
Outliers		0.030	0.049	–0.050	0.023
		(0.049)	(0.045)	(0.034)	(0.026)
Gender (1=male)			–0.278	–0.002	0.048
			(0.105)***	(0.077)	(0.061)
Age in years			–0.027	–0.004	–0.002
			(0.016)*	(0.011)	(0.007)
Invested in stocks			–0.234	–0.086	0.033
			(0.118)**	(0.090)	(0.063)
Risk tolerance			0.048	0.028	0.005
			(0.024)**	(0.017)	(0.015)
Trust			0.002	–0.011	–0.015
			(0.024)	(0.016)	(0.012)
Self-assessed knowledge			0.060	–0.036	0.006
			(0.046)	(0.029)	(0.018)
Knowledge score			0.019	0.014	–0.024
			(0.019)	(0.015)	(0.011)**
Pseudo-R^2	0.031	0.040	0.149	0.136	0.250
Observations	95	94	94	94	94

4.4.4 Investors' Choices After Seeing Intermediaries Perform

We first consider descriptive evidence to answer the question whether algorithm aversion arises in response to seeing the algorithm perform. Figure 4.5 shows the fraction of participants choosing to invest with the algorithm over the course of the experiment (at equal fees). Indeed, this fraction drops from the initial 56% to a low of 44% in investment blocks 4 and 5. Possibly, participants are disappointed that the algorithm is not perfect and makes (ex-post) mistakes. However, afterwards we observe a strong recovery to above 60% in the final blocks. With accumulating evidence apparently the outperformance of the algorithm becomes harder to ignore.[12] The average after investment block 1 is 51% in favor of the algorithm, which speaks against a general presence of algorithm aversion.

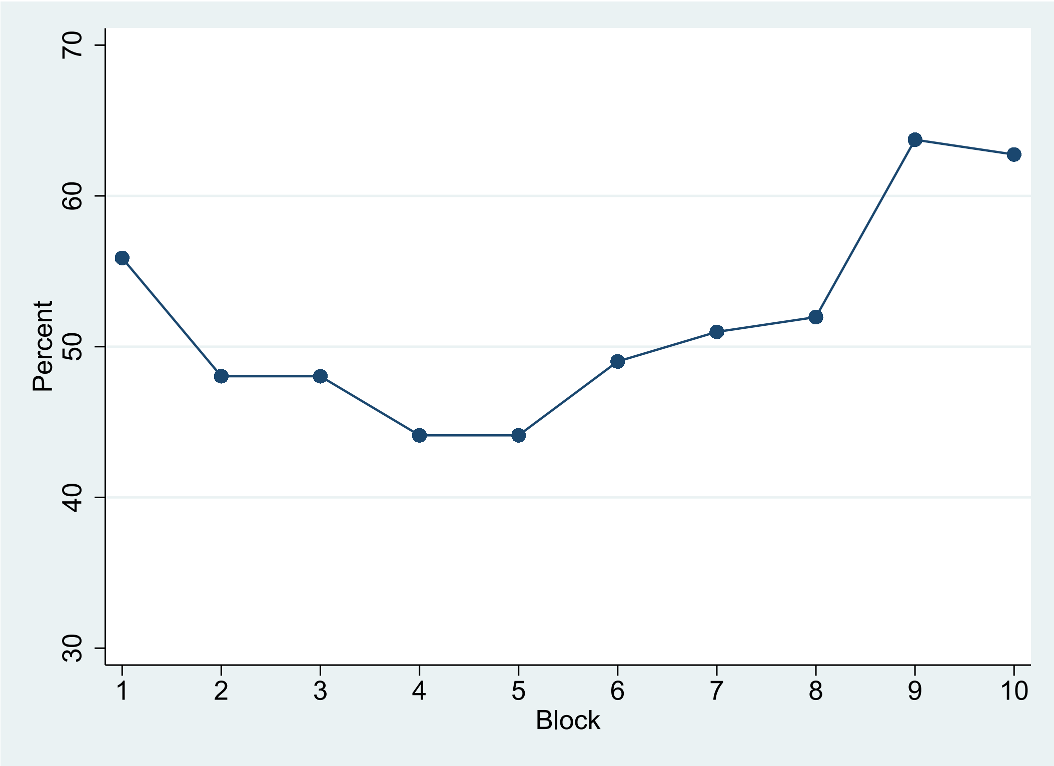
To examine Hypothesis 2 more closely, we treat repeated choices by participants as panel data with investment blocks as time dimension. Rational investors should learn from observing the decisions of intermediaries and their performance. At the end of each block, accumulated payoffs of both intermediaries are prominently displayed (including all previous blocks). We investigate how investors respond to cumulative performance of both intermediaries as well as their performance in individual blocks (e.g., the most recent performance). As before, the dependent variable is whether a participant chooses to invest with the algorithm for the current block. We estimate panel logistic regressions with standard errors clustered by session, as all participants within one session observe the same outcomes.

As displayed in Panel A of Table 4.4, investors react to cumulative performance of both intermediaries in the expected direction. The higher the

[12] In addition, outperformance of the algorithm is not evenly distributed in the experiment but is much stronger in later blocks. This is mainly by chance, as decision quality of the human fund managers remains stable (see also Figures C.1 and C.2 in the appendix).

Figure 4.5: Choice of Investment Algorithm Over Time

This figure shows the percentage of investors choosing to invest with the algorithm in blocks one to ten.



past payoff of the algorithm, the more likely are investors to choose the algorithm in the current block. On the contrary, the higher the payoff of the human fund manager, the less likely are they to choose the algorithm. A one Euro increase in performance of the algorithm implies an about 3.3% increase in the probability of choosing the algorithm. The magnitude of coefficients is very similar for both intermediaries. We cannot reject the null hypothesis that coefficients are of equal size in any of our regression specifications. Hence, we do not find that investors show different sensitivity to the performance of the algorithm.

In further specifications, we add block fixed effects and investor fixed effects (columns 2 and 3). Using investor fixed effects reduces the number of observations, as participants who never change their chosen intermediary (n=30) drop out of the model. Unsurprisingly, the size of the coefficients increases, as we hereby exclude the participants who are most insensitive to

Table 4.4: Choice of Intermediary Depending on Performance

Panel A of this table shows average marginal effects of panel logistic regressions with participants choice of intermediary (algorithm=1) in block t as dependent variable (at equal fees). The cumulative payoff is intermediaries' past payoff, accumulated over all blocks up to t-1. *Algorithm (t-1)* is a dummy variable indicating a participants choice for the previous block. This variable is interacted with the cumulative payoff variables. Regressions include block fixed effects and investor fixed effects as indicated. Panel B shows results for the same dependent variable regressed on up to five lags of payoffs of both intermediaries. Coefficients are average marginal effects of a panel logistic regression estimated with random effects. Clustered standard errors by session are shown in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Panel A	Choice of Algorithm in t				
	(1)	(2)	(3)	(4)	(5)
Cumulative payoff algorithm	0.033	0.036	0.055	0.052	0.047
	(0.003)***	(0.006)***	(0.010)***	(0.009)***	(0.009)***
Cumulative payoff human	–0.033	–0.031	–0.048	–0.046	–0.042
	(0.003)***	(0.004)***	(0.007)***	(0.007)***	(0.006)***
Algorithm (t-1)				0.105	–0.008
				(0.046)**	(0.068)
Algorithm (t-1) × cum. payoff algorithm					0.008
					(0.008)
Algorithm (t-1) × cum. payoff human					–0.007
					(0.008)
Observations	1020	1020	720	720	720
Block FE	NO	YES	YES	YES	YES
Investor FE	NO	NO	YES	YES	YES
Wald test for equal size of coefficients (p-value)	0.71	0.13	0.19	0.25	0.30

Panel B	Choice of Algorithm in t				
	(1)	(2)	(3)	(4)	(5)
Payoff algorithm (t-1)	0.028	0.034	0.036	0.034	0.034
	(0.004)***	(0.005)***	(0.006)***	(0.005)***	(0.006)***
Payoff human (t-1)	–0.028	–0.026	–0.032	–0.033	–0.033
	(0.005)***	(0.004)***	(0.005)***	(0.005)***	(0.006)***
Payoff algorithm (t-2)		0.021	0.027	0.029	0.028
		(0.004)***	(0.003)***	(0.005)***	(0.005)***
Payoff human (t-2)		–0.017	–0.016	–0.023	–0.022
		(0.005)***	(0.003)***	(0.004)***	(0.004)***
Payoff algorithm (t-3)			0.025	0.030	0.033
			(0.004)***	(0.005)***	(0.007)***
Payoff human (t-3)			–0.022	–0.021	–0.027
			(0.004)***	(0.003)***	(0.004)***
Payoff algorithm (t-4)				0.009	0.010
				(0.006)	(0.006)*
Payoff human (t-4)				–0.007	–0.009
				(0.006)	(0.007)
Payoff algorithm (t-5)					0.012
					(0.005)**
Payoff human (t-5)					–0.011
					(0.005)**
Observations	1020	918	816	714	612

performance. We also find an effect of the choice in the previous block, which suggests that having chosen the algorithm in t-1 makes it about 10% more likely to choose the algorithm again (column 4). We interact this variable with past performance to determine whether investors pay different attention to outcomes depending on the intermediary they invested with (column 5). Indeed, those who invested with the algorithm are more sensitive to its performance and less sensitive to the human fund managers' performance (not significant).[13]

Panel B of Table 4.4 reports results for individual lagged payoffs of both intermediaries. Their economic and statistical significance is slightly weaker than that of the cumulative payoffs, as they reflect only part of the observed performance history. There is evidence that more recent payoffs matter more, with the strongest effect of blocks t-1 to t-3. We find mixed evidence on coefficient size, with mostly a larger effect of the algorithm's performance (not significant). In sum, we cannot confirm Hypothesis 2 that participants disregard the performance of the algorithm. Figure 4.6 illustrates the almost monotonous effect of payoff difference in the previous block on the propensity to invest with the algorithm.

It is possible, however, that participants punish the algorithm more severely for bad performance. As the presented results do not condition on good or bad outcomes, the prediction of Hypothesis 3 might still be valid. Table 4.5 shows results of several regression specifications testing for this possibility. With the choice of the algorithm again as dependent variable, we now split past payoff differences into cases when the algorithm outperforms the human fund manager and those when the human outperforms the algorithm (for cumulative payoffs in columns 1 and 2, and for last block payoffs in columns 3 and 4). Coefficients are larger when the algorithm

[13] As interactions in logistic regressions can be misleading, we estimate a linear model for robustness. Magnitude and sign of the coefficients are comparable.

Figure 4.6: Choice of Investment Algorithm by Difference in Last Payoffs

This figure shows the percentage of investors choosing to invest with the algorithm in the current block, depending on the cumulated payoff difference (in €) between the algorithm and the human in the previous block.

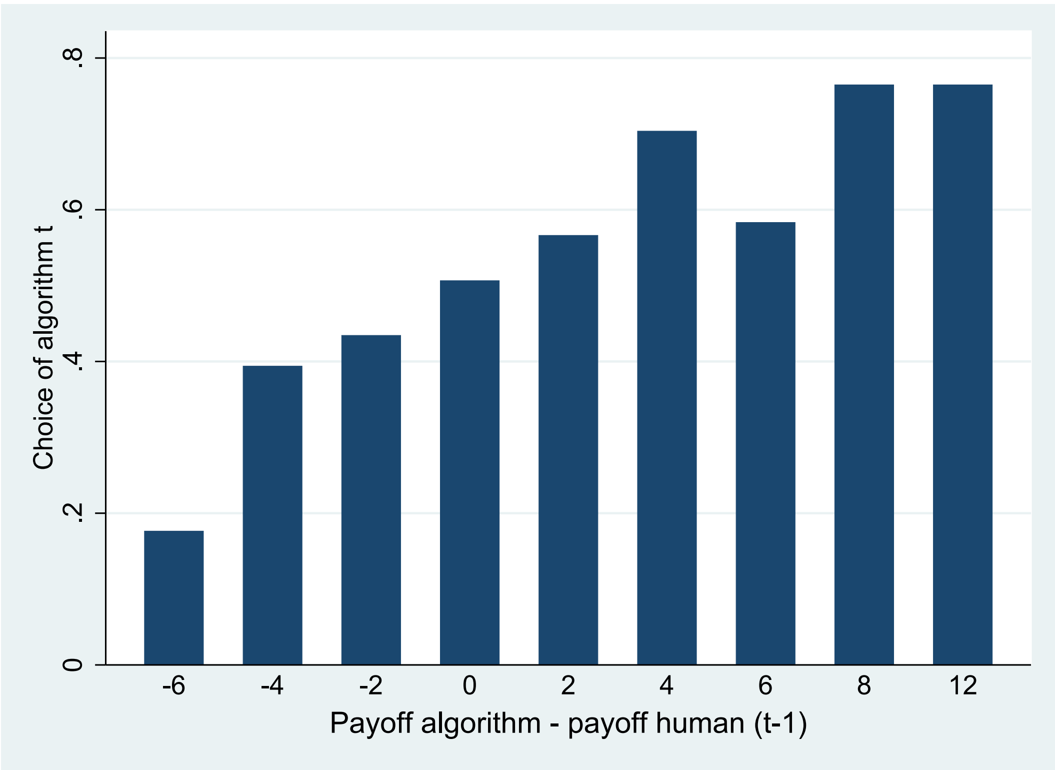


Table 4.5: Choice of Intermediary After Negative Performance

This table shows average marginal effects of panel logistic regressions with participants choice of intermediary (algorithm=1) in block t as dependent variable (at equal fees). The *Cumulative payoff difference* is payoff of the algorithm minus the payoff of the human fund manager, accumulated over all blocks up to t-1. The variable *Last payoff difference* is this difference for (t-1) only. *Number of errors* is the number of ex-post errors (=choosing the intermediary with the lower outcome) for block t-1. Regressions include block fixed effects and investor fixed effects as indicated. Coefficients are average marginal effects of a panel logistic regression estimated with random effects. Clustered standard errors by session are shown in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

	Choice of Algorithm in t					
	(1)	(2)	(3)	(4)	(5)	(6)
Cum. payoff difference (Algorithm>Human)	0.027	0.039				
	(0.004)***	(0.011)***				
Cum. payoff difference (Human>Algorithm)	0.037	0.063				
	(0.010)***	(0.013)***				
Last payoff difference (Algorithm>Human)			0.024	0.041		
			(0.007)***	(0.011)***		
Last payoff difference (Human>Algorithm)			0.034	0.049		
			(0.012)***	(0.020)**		
Number of errors by algorithm					–0.057	–0.098
					(0.013)***	(0.023)***
Number of errors by human					0.052	0.079
					(0.012)***	(0.022)***
Observations	918	585	918	585	918	585
Block FE	NO	YES	NO	YES	NO	YES
Investor FE	NO	YES	NO	YES	NO	YES
Wald test for equal size of coefficients (p-value)	0.47	0.26	0.57	0.76	0.74	0.54

underperforms, suggesting a stronger sensitivity to bad outcomes by the algorithm. The effect size is between 20% and 60% larger than after good outcomes, but does never attain statistical significance.

We earlier defined an ex-post error as choosing the asset with the lower payoff in a given trial. Another way to test whether investors are quicker to abandon the investment algorithm is counting the number of errors per block for both intermediaries. Participants' sensitivity to errors by the algorithm is somewhat higher than to errors by the human manager (see columns 5 and 6; not statistically significant). However, one has to consider that humans make more errors, which renders it quite natural that a single error bears less significance for judging them. A similar argument holds for payoff differences in favor of the human, which are less frequent and on average smaller justifying a stronger reaction on a per Euro basis. These statistics also explain why the found asymmetry does not produce algorithm aversion in the long-run. As the algorithm is the better intermediary, frequency and magnitude of outperformance more than compensates for the slightly lower sensitivity. Evidence for Hypothesis 3 is thus relatively weak.

So far, we treated repeated decisions for an intermediary the same way as a switch between intermediaries. Arguably, a switch has special significance in determining what considerations govern participants' choices. We observe 105 switches to the algorithm and 98 switches to the human fund managers (at equal fees). On average, investors switch intermediaries in 2 out of 9 blocks after their first decision. 30 participants never revise their initial choice, 19 switch once, 15 twice, and 38 switch three or more times. An optimal switching point for a Bayesian would be once the non-chosen intermediary overtakes the chosen one in terms of accumulated payoffs.[14] We identify 109 such situations, which means that participants switch about twice as often as a Bayesian would. However, they seize 57% of the optimal switching opportunities.

In a logistic panel regression with observed switches as dependent variable, we confirm that optimal switching points have strong explanatory power (see Panel A of Table 4.6, column 1). When switching is optimal, we are 29% more likely to observe an actual switch. It could be that participants rather look at the performance of the intermediary they currently invest with (own) or the one they might switch to (target). We find no conclusive evidence in that regard, whether we look at cumulative payoffs, last block payoffs, or number of errors (columns 2-4). Higher own performance always reduces the propensity to switch, while higher target performance increases it, both with very similar effect size. We find that last block results play a relatively large role for switching, consistent with the idea that older

[14] This rule can be refined by considering the decisions and not just the outcomes. For the current purpose optimality based on outcomes is sufficient (see also section 4.4.5).

information could have triggered a switch already before.

More important for algorithm aversion is switching behavior by type of intermediary, which is displayed in Panel B of Table 4.6. When participants switch to the algorithm, they consider the performance of their target as well as the performance of the human fund manager about equally. Likewise, switching to the human fund manager is informed almost symmetrically by the performance of the algorithm and human. Interestingly, coefficients are smaller and significance is weaker for switches to the human, suggesting that participants pay less attention to performance but might have other reasons. We do not find any evidence for more pronounced switching behavior after errors by the algorithm. In fact, errors by the algorithm matter less for switching. We thus do not find support for Hypothesis 3 from switching behavior.

4.4.5 Skill vs. Luck

We present evidence that participants strongly consider performance when selecting a financial intermediary. However, in the used experimental setting as well as in reality, performance is only a noisy signal of true skill (Heuer et al., 2017). We thus break down total performance of both intermediaries into a component of skill and a component of luck. For each trial, we calculate the expected outcome of the intermediary's chosen asset using the information available at that point in time. The expected outcome is the skill component, which we then subtract from the realized payoff of the chosen asset. This difference is the luck component. For bond investments, luck is therefore always zero. For stock investments, luck can either be positive (outcome $>$ expected outcome) or negative (outcome $<$ expected outcome).

Table 4.7 summarizes luck and skill for both intermediaries aggregated by investment block. Average luck is not significantly different from zero

Table 4.6: Analysis of Switching Behavior

Panel A of this table reports results of logistic panel regressions with switch of intermediary as dependent variable. It is a binary variable equal to 1 if an investors switches intermediary from the last block to the current block, and 0 otherwise. *Switch optimal* is a dummy variable equal to 1 if, based on total aggregated performance of both intermediaries, a switch was optimal in a given block, and 0 otherwise. Payoff variables are as defined before, with target indicating that the payoff refers to the (potential) target of a switch, and own indicating that the payoff refers to the intermediary invested with in t-1. Panel B shows the same regression results separately for switches to the algorithm and switches to the human fund manager. Coefficients are average marginal effects of a panel logistic regression estimated with random effects. Clustered standard errors by session are shown in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Panel A	Switch of intermediary in t			
	(1)	(2)	(3)	(4)
Switch optimal	0.286			
	(0.033)***			
Cum. payoff target		0.018		
		(0.003)***		
Cum. payoff own		–0.019		
		(0.003)***		
Payoff target (t-1)			0.035	
			(0.005)***	
Payoff own (t-1)			–0.038	
			(0.005)***	
Number of errors target				–0.077
				(0.019)***
Number of errors own				0.074
				(0.010)***
Observations	918	918	918	918

Panel B	Switch to algorithm			Switch to human		
	(1)	(2)	(3)	(4)	(5)	(6)
Cum. payoff target	0.014			0.004		
	(0.005)***			(0.005)		
Cum. payoff own	–0.014			–0.005		
	(0.005)***			(0.005)		
Payoff target (t-1)		0.023			0.011	
		(0.003)***			(0.005)**	
Payoff own (t-1)		–0.026			–0.010	
		(0.004)***			(0.004)**	
Number of errors target			–0.048			–0.027
			(0.010)***			(0.011)**
Number of errors own			0.052			0.017
			(0.010)***			(0.010)
Observations	918	918	918	918	918	918

Table 4.7: Luck and Skill of Financial Intermediaries

This table shows total payoffs, luck and skill of both financial intermediaries (in €). All outcomes are aggregated by investment block. Skill is calculated as the expected payoff based on the intermediary's asset choices. Luck is calculated as the difference between realized outcomes and expected payoff.

		N	μ	σ	Min	Max
Investment algorithm	Total payoff	120	18.75	3.96	12.00	30.00
	Skill	120	19.02	1.16	18.00	21.08
	Luck	120	-0.27	3.12	-6.64	8.92
Human fund manager	Total payoff	120	18.17	4.55	6.00	30.00
	Skill	120	18.49	1.50	14.92	21.08
	Luck	120	-0.33	3.44	-8.92	8.92

for both intermediaries. This finding is not surprising, as consistent luck would defy the random nature of outcomes. However, luck or bad luck in individual blocks can be large. When either intermediary outperforms the other within a block, luck drives this outperformance in 65% of the cases (due to the larger standard deviation of luck compared to skill). The earlier mentioned payoff difference of 58 cents in favor of the algorithm is almost entirely due to skill.

To disentangle whether skill or luck is appreciated by investors, we include both as variables in a regression of investor choice (Table 4.8). As we have already established, participants respond to overall performance in the previous investment block (column 1). However, the effect of the payoff component produced by skill remains insignificant (column 2). The larger effect size arises from the fact that skill differences are often small. In contrast, participants strongly react to luck, which also is the only relevant payoff component when we include both components simultaneously (columns 3 and 4). Although the assets in the experiment have simple payoff structures and investors possess the same information as intermediaries, they are unable to draw additional inferences from choices. They concentrate on outcomes in line with an outcome bias (Baron and Hershey, 1988).

Table 4.8: Luck Versus Skill in Choice of Intermediary

This table shows average marginal effects of panel logistic regressions with participants choice of intermediary (algorithm=1) in block t as dependent variable (at equal fees). *Payoff difference*, *Skill difference*, and *Luck difference* refer to the differences in payoffs in block t-1 between the investment algorithm and human fund manager as defined in Table 4.7. *Ex-post errors* are the number of instances in which either intermediary chose the asset with the lower payoff in block t-1. Ex-ante errors are the number of instances in which either intermediary chose inferior asset given the true state (good or bad) in block t-1. Clustered standard errors by session are shown in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

	Choice of Algorithm in t						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Payoff difference (t-1)	0.027						
	(0.005)***						
Skill difference (t-1)		0.044		–0.015			
		(0.032)		(0.031)			
Luck difference (t-1)			0.032	0.033			
			(0.005)***	(0.006)***			
Ex-post errors algorithm (t-1)					–0.057		–0.053
					(0.013)***		(0.017)***
Ex-post errors human (t-1)					0.052		0.052
					(0.012)***		(0.015)***
Ex-ante errors algorithm (t-1)						–0.033	–0.007
						(0.011)***	(0.018)
Ex-ante errors human (t-1)						0.027	0.002
						(0.009)***	(0.012)
Observations	918	918	918	918	918	918	918

A way to corroborate this finding is to look at ex-post errors and ex-ante errors as defined in section 4.4.2. Clearly, few ex-ante errors are a better signal of skill as they show how often an intermediary did not identify the superior asset correctly for a given state of the stock (good or bad). Meanwhile, ex-post errors include a major luck component as they depend on the outcome of the payoff draw. Indeed, participants react in expected manner to both types of errors, with a stronger effect of ex-post errors (columns 5 and 6). However, this result may be due to the fact that ex-ante and ex-post errors often coincide. Including both types of errors simultaneously reveals that ex-post errors crowd out the effect of ex-ante errors (column 7). We conclude that participants are unable to distinguish skill and luck in the experimental setting. Contrary to Hypothesis 3, they do not respond more strongly to errors by the algorithm using different types of error definitions.

4.5 Conclusion

While the term "algorithm aversion" has been introduced only recently (Dietvorst et al., 2015), there already exist numerous studies on human preferences for or against using algorithms. However, the literature is still indecisive on the general prevalence of algorithm aversion. Part of this is due to the different contexts in which algorithm aversion is tested, while another part is due to the methodology algorithm aversion is tested. As Logg (2017) points out, it is difficult to assess when and how algorithm aversion matters.

The aim of this study is to provide insights for financial decisions, as finance is a field in which the use of algorithm is not only theoretically promising, but also practically important. We test algorithm aversion in an experimental setting that is (necessarily) simplified, but that contains many features of real-world financial decision making. In particular, a real human fund manager selected by financial knowledge competes with a rule-based investment algorithm. They act as financial intermediaries for investors just as mutual funds would. They operate in a financial market that reveals useful information, but at the same time is driven by chance. This means that they have opportunities to show their skill, but also inevitably will make errors. Investors observe the performance of the intermediaries and their choices and can react by changing their choice of intermediary.

Under these premises, we find no sign of algorithm aversion. Investors initially have a slight preference for the algorithm. After observing outcomes they strongly favor the intermediary who outperforms, but they do so equally for both intermediaries. We do not find support for the assumption that investors abandon algorithms after seeing them err. Instead, better performance by the algorithm over time convinces them to switch to the algorithm. However, investors do not discern luck and skill and mostly rely on investment outcomes without considering the skill revealed by choices.

There are certain ways in which financial decisions differ from decisions typically studied in the algorithm aversion literature. Two prominent examples are university admissions and medical decisions, which are likely perceived as contexts in which human intuition or even human empathy should play a greater role. In these contexts, prospects of academic success and health of humans are judged, while in finance the prospects of (inanimate) financial assets are judged. The contexts might further differ in the weight people place on "soft factors" such as interviews and other direct communication, as opposed to quantitative strategies based on data. Finally, there is a moral dimension which deters people from allowing algorithms to make important life decisions on career or health that is presumably less pronounced for asset allocation. These considerations are in line with lower or absent algorithm aversion as we observe in the experiment.

Lastly, there are also practical implications that follow from our experiment. By collecting a sample of university students, often with background in economics or finance, our sample is likely similar in financial and technological sophistication to the customer base of well-known robo-advisors. In the online survey, sample participants state they believe human fund managers to be better able to deal with outlier events (e.g., financial crisis of 2008). They also state to rather view algorithms as aid to human fund managers. These statements entail that robo-advisors or algorithmic trading funds could highlight that human experts and algorithms form a symbiotic relationship. In other words, human experts could be proclaimed to monitor the complex algorithms, and have the power to ultimately step in in case of extreme events. To a certain extent, this is already done in practice: Both Betterment and Wealthfront frame their services as being delivered by a group of humans ("we").[15] Moreover, both companies also give detailed information about their investment experts and investment committee members.

In addition, as we find that performance but not skill is rewarded, robo-advisors and algorithmic trading funds need to point out to factors that guarantee better performance over their human counterpart ex ante and ex post. One such factor are management fees, which are certain to lower client's returns. In the end, however, businesses based on algorithms will have to prove their success over a long period of time in order to attract convince the skeptics.

[15] See their web presences. For Betterment: "We'll learn a bit about you.", "We'll build you a portfolio.", or "We're on a mission to help you make the most of your money.". For Wealthfront: "Live the life you want. We've got your back.", "Financial planning and investing with Wealthfront couldn't be easier. We do it for you.". As of 8 August 2018.

Appendix A

Outcome Bias

Table A.1: Sample Statistics

This table reports summary statistics for all participants. **CRT** is the number of correct answers on a 7 question cognitive reflection test taken from Toplak et al. (2014). **Age** is the participant's age in years. **Understanding Distributions** measures how difficult participants found interpreting the distributions of assets in the experiment. It is calculated from a 5-point Likert scale where 1 refers to "Very easy" and 5 refers to "Very difficult".

N = 100	μ	σ	Min	Max
CRT	3.90	2.30	0	7
Age	32.08	8.62	18	70
Understanding Distributions	2.58	0.99	1	5

	Frequency
Understanding Question 1	
Correct	86
Wrong	14
Understanding Question 2	
Correct	87
Wrong	13
Both correct	78
Gender	
male	70
female	30
Education	
Bachelor	46
Master (or equivalent)	7
Middle School	5
PhD (or equivalent)	1
University entrance qualification	33
other	7
Occupation	
High school student	1
employee	62
retired	1
self-employed	19
student	9
other	8
Invested in corp. bonds	
No	80
Yes	20
Invested in gov. bonds	
No	72
Yes	28
Invested in passive funds	
No	64
Yes	36
Invested in active funds	
No	75
Yes	25
Invested in derivatives	
No	88
Yes	12

Figure A.1: Examplary Screen for Treatment 1

Choice of investment manager (Round 1/5 in block 1)

Investment manager	Last payoffs realized	Investment that manager has chosen	Ranking of last payoffs realized
Investment manager 1	$1.25	A	2
Investment manager 2	$1.00	A	3
Investment manager 3	$1.25	A	2
Investment manager 4	$1.75	A	1
Investment manager 5	$1.75	B	1

I wish to invest with:

- **Manager 1**
- **Manager 2**
- **Manager 3**
- **Manager 4**
- **Manager 5**

Next

Show calculator

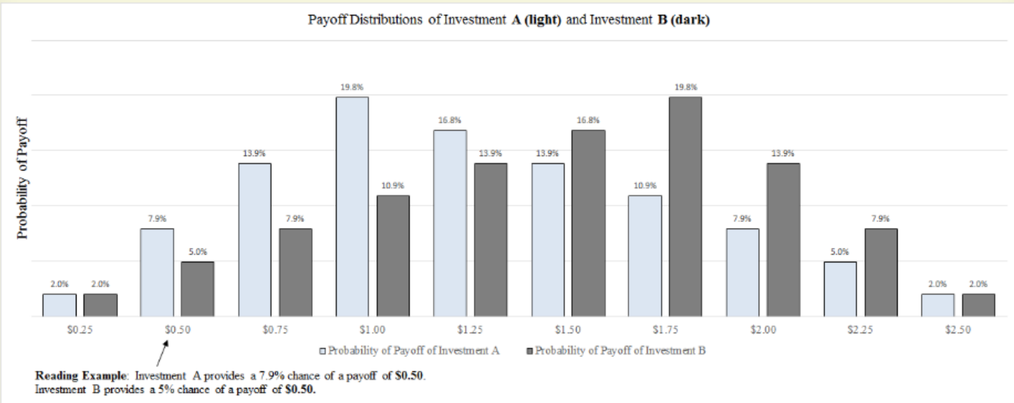
Instructions

You have to choose **one** investment manager who you believe is best for you. You will be able to choose among 5 investment managers. Each investment manager invests either **only** into Investment **A** or **only** into Investment **B**, i.e., investment managers **stick** with their investment **within** the current investment task. After you choose your investment manager, he will invest into the investment he sticks with and you will obtain the payoff from the investment. **The payoffs and their probabilities for both investments are provided to you in the chart below**.

For you to judge the investment manager, the following information is provided to you:

1. **The investment** the investment manager sticks with.
2. **The payoff** obtained by each investment manager the last time he invested into the investment he sticks with.
3. **The ranking** of investment managers based on these payoffs.

Important: This investment task will be repeated 5 times. Each investment task is **independent** from previous and future investment tasks. This means that the investments to which investment managers stick in this **current** investment task have **nothing to do** with the investments they stick with in **previous** or **future** investment tasks.



For your convenience, these instructions will remain available to you on all subsequent screens of this study.

Results of Experiment Without Restrictions

A secondary study without any restrictions to manager choices was run on Amazon Mechanical Turk in March 2018. Qualification criteria of AMT workers were unchanged. In total, 151 participants were recruited. In the following, we show basic results of the secondary study. For a discussion of the disadvantages of running the study without any restrictions to manager choices see section 2.2 of the paper.

To test for *Suboptimal Choice*, we drop *1)* all manager choices for which either every investment manager invests into the preferred asset, as any choice would then be optimal by default, and we drop *2)* all manager choices for which no investment manager invests into the preferred asset, as any choice would then be suboptimal by default. To test for *Outcome Bias* and *Outcome Bias | Suboptimal*, we further drop all manager choices for which the investment manager with the highest historical payoff is also an investment manager who could be chosen by a rational individual. In these cases, we could not distinguish whether the manager choice is driven by the Outcome Bias or by rational decision making.

In Treatment 1, the number of observations decreases from 755 to 716 for tests of *Suboptimal Choice*, and to 373 for tests of *Outcome Bias*. In Treatment 2, the number of observations decreases from 755 to 699 for tests of *Suboptimal Choice* , and to 358 for tests of *Outcome Bias*. In Treatment 3, the number of observations remains at 755 for tests of *Suboptimal Choice*, and decreases to 470 for tests of *Outcome Bias*. Since in Treatment 3 a good investment choice needs to be inferred from payoff probabilities, fewer observations need to be dropped than in the other two treatments.

Results from the secondary study are well in line with Hypothesis 1: The more difficult the separation of skill and luck, the more investors are prone to the Outcome Bias (Figure A.2). Similarly, the more difficult the separation of skill and luck, the more investors tend to make suboptimal manager choices, resulting in not obtaining their preferred investment (Figure A.3). Again, the Outcome Bias seems to be the main driver of suboptimal choices (Figure A.4). Although proportions of *Outcome Bias*, *Suboptimal Choice*, and *Outcome Bias | Suboptimal* are slightly lower throughout all treatments than in the main study, a sizable fraction of manager choices is outcome biased or suboptimal nonetheless.

Results from the secondary study can also not be reconciled with randomly simulated data. Proportions of *Outcome Bias* and *Outcome Bias | Suboptimal* are consistently above proportions expected from random choices. However, compared to randomly simulated data, participants make fewer suboptimal manager choices in all treatments. Nonetheless, as the difficulty of separating skill from luck increases, so does the proportion of suboptimal choices in any treatment. In summary, comparing observed to simulated data in the secondary study presents a similar picture as in the main study (see Table 2.4).

Figure A.2: Distribution of Outcome Bias – No Restrictions

Differences of proportions of *Outcome Bias* between treatments are all significant at the 1%-level. *P*-values were obtained using the regression-based approach as outlined in the main analysis.

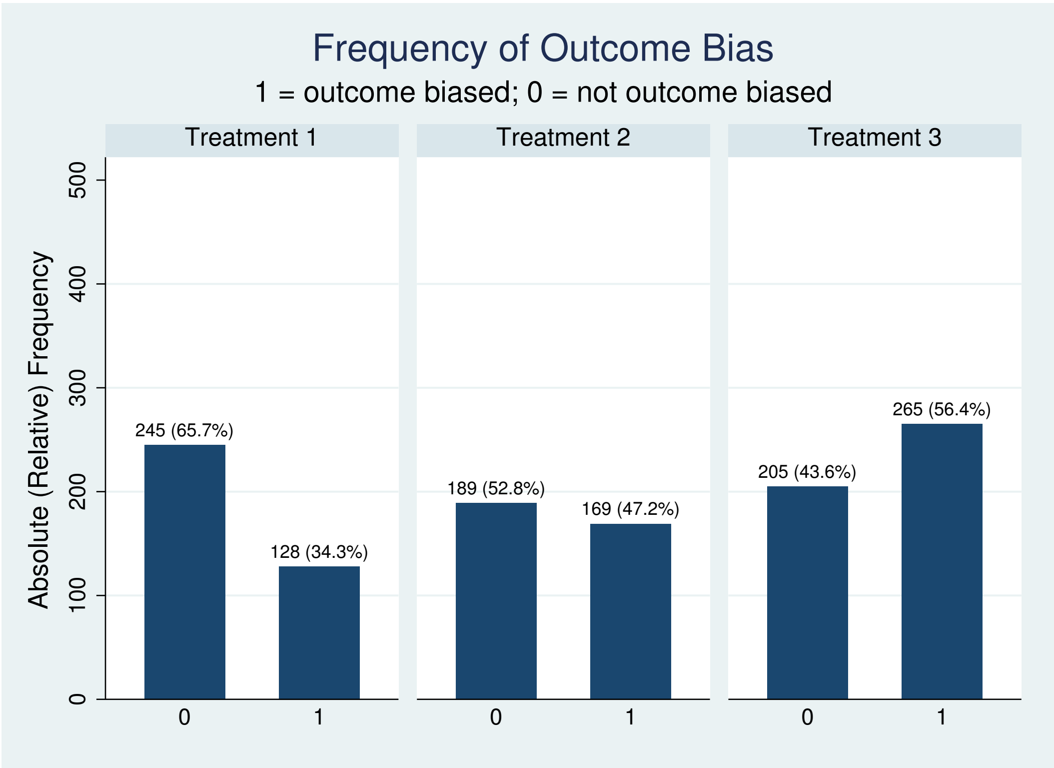


Figure A.3: Distribution of Suboptimal Choices – No Restrictions

Differences of proportions of *Suboptimal Choice* between treatments are all significant at the 1%-level. *P*-values were obtained using the regression-based approach as outlined in the main analysis.

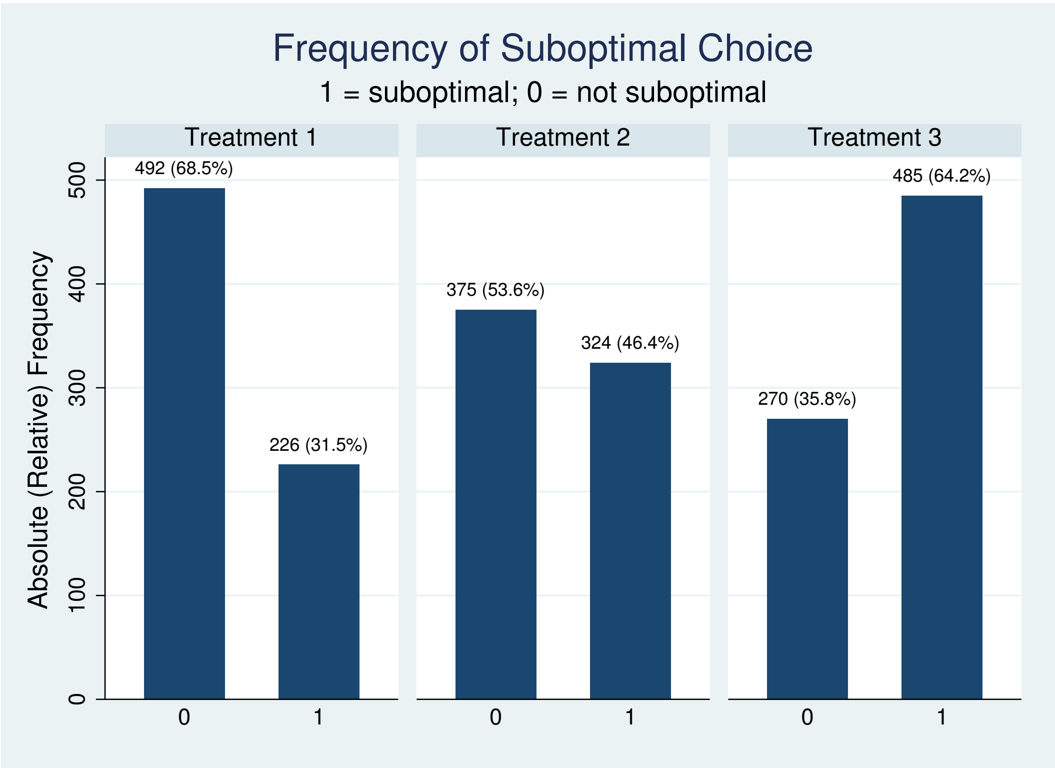


Figure A.4: Distribution of Outcome Bias | Suboptimal – No Restrictions

Differences of proportions of *Outcome Bias | Suboptimal* between treatments are all insignificant. *P*-values were obtained using the regression-based approach as outlined in the main analysis.

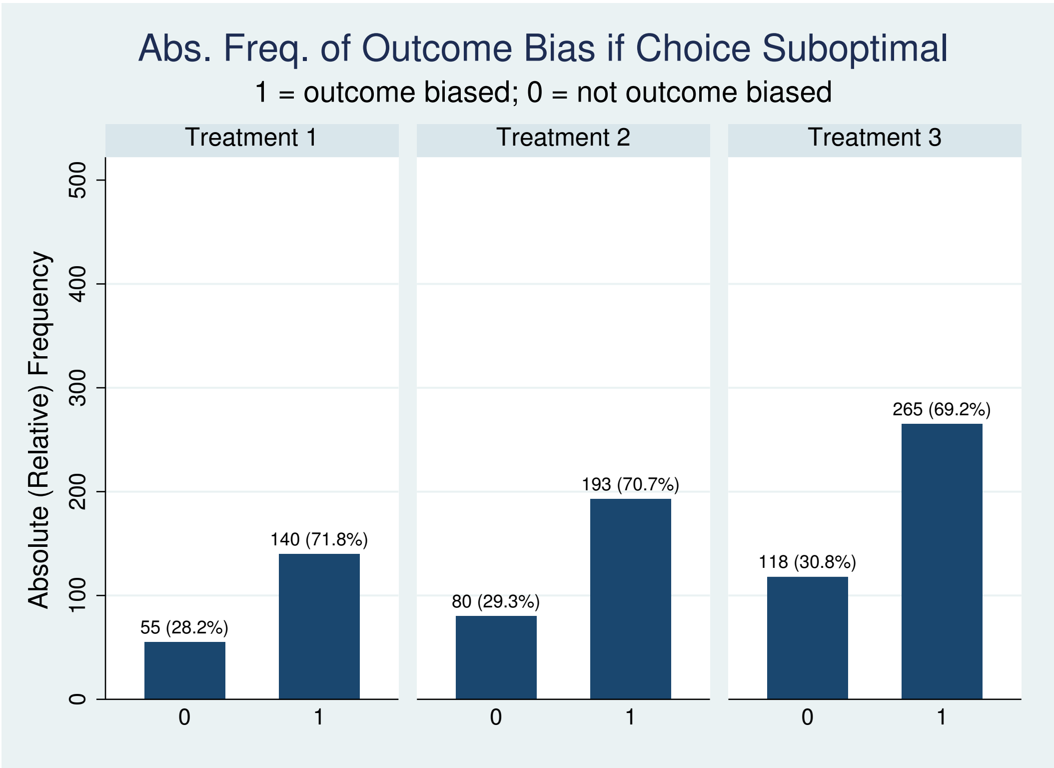


Table A.2: Observed vs. Simulated Data – No Restrictions

Simulated data was obtained by simulating the experiment 1000 times with random choices. *P*-values were obtained using the regression-based approach as outlined in the main analysis.

Outcome Bias

	Outcome Bias = 0	Outcome Bias = 1	Δ *p*-value
Treatment 1			
Observed	245 (65.7%)	128 (34.3%)	0.001***
Simulated	2072 (77.9%)	588 (22.1%)	
Treatment 2			
Observed	189 (52.8%)	169 (47.2%)	0.000***
Simulated	1843 (76.3%)	573 (23.7%)	
Treatment 3			
Observed	205 (43.6%)	265 (56.4%)	0.000***
Simulated	2690 (75.2%)	888 (24.8%)	

Suboptimal Choice

	Suboptimal Choice = 0	Suboptimal Choice = 1	Δ *p*-value
Treatment 1			
Observed	492 (68.5%)	226 (31.5%)	0.000***
Simulated	2365 (50.5%)	2315 (49.5%)	
Treatment 2			
Observed	375 (53.6%)	324 (46.4%)	0.049**
Simulated	2345 (49.7%)	2371 (50.3%)	
Treatment 3			
Observed	270 (35.8%)	485 (64.2%)	0.000***
Simulated	1383 (27.4%)	3671 (72.6%)	

Outcome Bias | Suboptimal

	Outcome Bias | Suboptimal = 0	Outcome Bias | Suboptimal = 1	Δ *p*-value
Treatment 1			
Observed	55 (28.2%)	140 (71.8%)	0.000***
Simulated	992 (61.0%)	634 (39.0%)	
Treatment 2			
Observed	80 (29.3%)	193 (70.7%)	0.000***
Simulated	915 (60.4%)	600 (39.6%)	
Treatment 3			
Observed	118 (30.8%)	265 (69.2%)	0.000***
Simulated	1719 (65.9%)	888 (34.1%)	

Experimental Instructions

The following images show instructions and experimental screens as presented to participants. All realized values shown in the experimental screen are for illustration purposes only.

Screen 1:

Introduction (1/2)

Dear participant,
thank you for taking part in our study.

Purpose: In this study, we would like to elicit your preferences regarding investment opportunities. The main part of the study consists of 3 blocks à 5 identical investment choices, for a total of 15 investment decisions to make. In each block you will be able to play 1 practice round which does not count towards your payment. A short quiz and a short survey conclude the study. Instructions and necessary information will be provided to you at each step of the study. Please read them carefully.

Next

Screen 2:

Introduction (2/2)

Variable Payment: Your variable payment for participation depends on the results of your investment choices. After each of your investment choices you will be shown the payoff from this respective choice. Out of all the investment choices you make, one will be selected randomly. The payoff from this randomly selected investment choice thus is your variable payoff for participating in the experiment. Your variable payment can range from a minimum of $0.25 to a maximum of $3.00.

Next

0

Screen 3 (based on ***Treatment 3****):*

Instructions for block 1

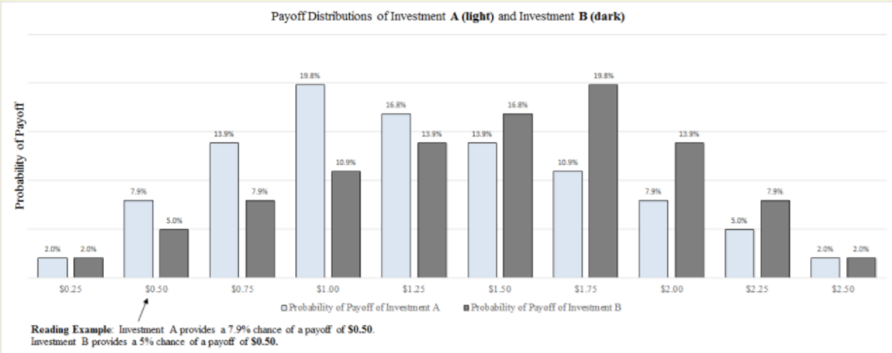
Instructions

You have to choose **one** investment manager who you believe is best for you. You will be able to choose among 5 investment managers. Each investment manager invests either **only** into Investment **A** or **only** into Investment **B**, i.e., investment managers **stick** with their investment **within** the current investment task. Whether investment managers stick with investment **A** or investment **B** is determined **randomly** with **50%** chance at the beginning of each investment task. After you choose your investment manager, he will invest into the investment he sticks with and you will obtain the payoff from the investment. **The payoffs and their probabilities for both investments are provided to you in the chart below.**

For you to judge the investment manager, the following information is provided to you:

1. **The payoff** obtained by each investment manager the last time he invested into the investment he sticks with.
2. **The ranking** of investment managers based on these payoffs.

Important: This investment task will be repeated 5 times. Each investment task is **independent** from previous and future investment tasks. This means that the investments to which investment managers stick in this **current** investment task have **nothing to do** with the investments they stick with in **previous** or **future** investment tasks.



For your convenience, these instructions will remain available to you on all subsequent screens of this study.

Next

*Screen 4 (based on **Treatment 3**): If answer "yes"*

Your answer "yes" is incorrect. Investment managers can not promise you an exact payoff once they have chosen any investment. Payoffs are randomly drawn from the investment's respective payoff distribution as shown to you.

Are investment managers able to guarantee you a payoff of, for example, $2 if they invest into Investment A or Investment B?

- yes
- no

Next

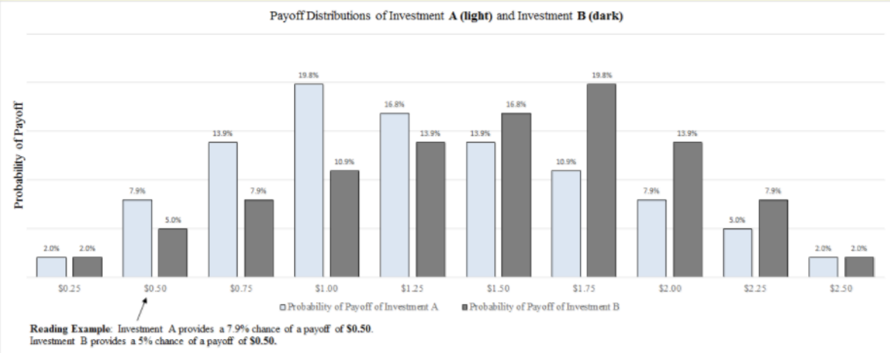
Instructions

You have to choose **one** investment manager who you believe is best for you. You will be able to choose among 5 investment managers. Each investment manager invests either **only** into Investment **A** or **only** into Investment **B**, i.e., investment managers **stick** with their investment **within** the current investment task. Whether investment managers stick with investment **A** or investment **B** is determined **randomly** with **50%** chance at the beginning of each investment task. After you choose your investment manager, he will invest into the investment he sticks with and you will obtain the payoff from the investment. **The payoffs and their probabilities for both investments are provided to you in the chart below**.

For you to judge the investment manager, the following information is provided to you:

1. **The payoff** obtained by each investment manager the last time he invested into the investment he sticks with.
2. **The ranking** of investment managers based on these payoffs.

Important: This investment task will be repeated 5 times. Each investment task is **independent** from previous and future investment tasks. This means that the investments to which investment managers stick in this **current** investment task have **nothing to do** with the investments they stick with in **previous** or **future** investment tasks.



For your convenience, these instructions will remain available to you on all subsequent screens of this study.

Screen 4 (based on ***Treatment 3****): If answer "no"*

Your answer "no" is correct. Investment managers can not promise you an exact payoff once they have chosen any investment. Payoffs are randomly drawn from the investment's respective payoff distribution as shown to you.

Are investment managers able to guarantee you a payoff of, for example, $2 if they invest into Investment A or Investment B?

- yes
- no

Next

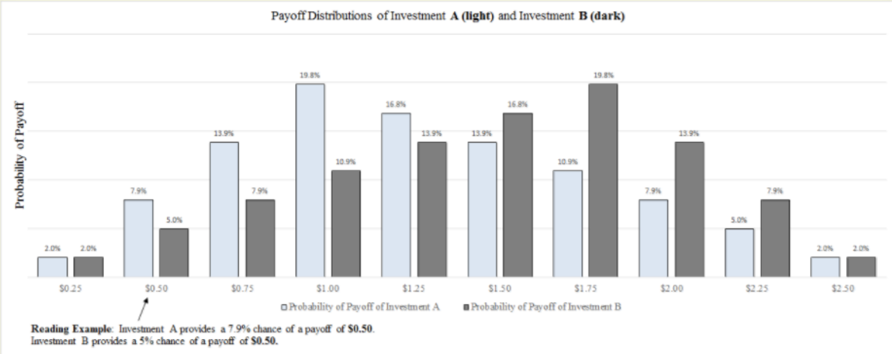
Instructions

You have to choose **one** investment manager who you believe is best for you. You will be able to choose among 5 investment managers. Each investment manager invests either **only** into Investment **A** or **only** into Investment **B**, i.e., investment managers **stick** with their investment **within** the current investment task. Whether investment managers stick with investment **A** or investment **B** is determined **randomly** with **50%** chance at the beginning of each investment task. After you choose your investment manager, he will invest into the investment he sticks with and you will obtain the payoff from the investment. **The payoffs and their probabilities for both investments are provided to you in the chart below**.

For you to judge the investment manager, the following information is provided to you:

1. **The payoff** obtained by each investment manager the last time he invested into the investment he sticks with.
2. **The ranking** of investment managers based on these payoffs.

Important: This investment task will be repeated 5 times. Each investment task is **independent** from previous and future investment tasks. This means that the investments to which investment managers stick in this **current** investment task have **nothing to do** with the investments they stick with in **previous** or **future** investment tasks.



For your convenience, these instructions will remain available to you on all subsequent screens of this study.

*Screen 5 (based on **Treatment 3**): If answer "yes"*

Your answer "yes" is incorrect. The chances of realizing a certain payoff always stay the same, regardless of previous payoff realiziations. Payoffs are randomly drawn from the investment's respective payoff distribution.

Is it true that if an investment yielded a payoff of for example $1.50 the last time, the chances of realizing a payoff of $1.50 in the future increase?

- yes
- no

Next

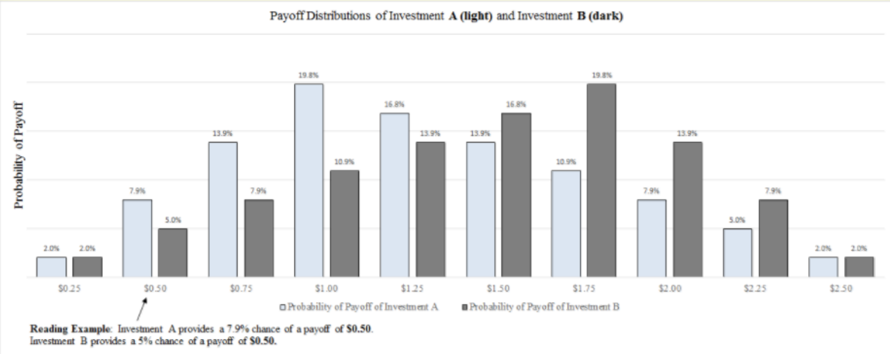
Instructions

You have to choose **one** investment manager who you believe is best for you. You will be able to choose among 5 investment managers. Each investment manager invests either **only** into Investment **A** or **only** into Investment **B**, i.e., investment managers **stick** with their investment **within** the current investment task. Whether investment managers stick with investment **A** or investment **B** is determined **randomly** with **50%** chance at the beginning of each investment task. After you choose your investment manager, he will invest into the investment he sticks with and you will obtain the payoff from the investment. **The payoffs and their probabilities for both investments are provided to you in the chart below**.

For you to judge the investment manager, the following information is provided to you:

1. **The payoff** obtained by each investment manager the last time he invested into the investment he sticks with.
2. **The ranking** of investment managers based on these payoffs.

Important: This investment task will be repeated 5 times. Each investment task is **independent** from previous and future investment tasks. This means that the investments to which investment managers stick in this **current** investment task have **nothing to do** with the investments they stick with in **previous** or **future** investment tasks.



For your convenience, these instructions will remain available to you on all subsequent screens of this study.

Screen 5 (based on **Treatment 3**): If answer "no"

Your answer "no" is correct. The chances of realizing a certain payoff always stay the same, regardless of previous payoff realiziations. Payoffs are randomly drawn from the investment's respective payoff distribution.

Is it true that if an investment yielded a payoff of for example $1.50 the last time, the chances of realizing a payoff of $1.50 in the future increase?

- yes
- no

Next

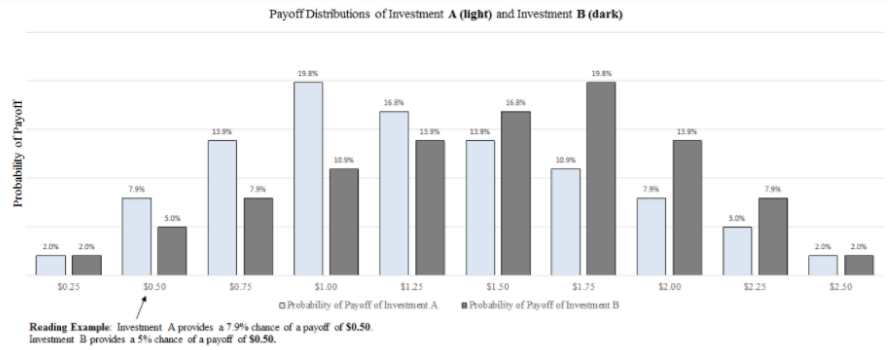
Instructions

You have to choose **one** investment manager who you believe is best for you. You will be able to choose among 5 investment managers. Each investment manager invests either **only** into Investment **A** or **only** into Investment **B**, i.e., investment managers **stick** with their investment **within** the current investment task. Whether investment managers stick with investment **A** or investment **B** is determined **randomly** with **50%** chance at the beginning of each investment task. After you choose your investment manager, he will invest into the investment he sticks with and you will obtain the payoff from the investment. **The payoffs and their probabilities for both investments are provided to you in the chart below.**

For you to judge the investment manager, the following information is provided to you:

1. **The payoff** obtained by each investment manager the last time he invested into the investment he sticks with.
2. **The ranking** of investment managers based on these payoffs.

Important: This investment task will be repeated 5 times. Each investment task is **independent** from previous and future investment tasks. This means that the investments to which investment managers stick in this **current** investment task have **nothing to do** with the investments they stick with in **previous** or **future** investment tasks.



For your convenience, these instructions will remain available to you on all subsequent screens of this study.

*Screen 6 (based on **Treatment 3**):*

Investment preferences

Please indicate into which investment you would invest if you could invest yourself:

- Investment A
- Investment B

Next

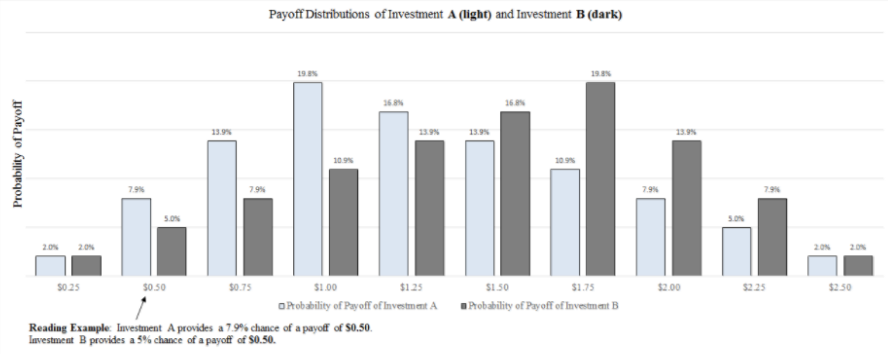
Instructions

You have to choose **one** investment manager who you believe is best for you. You will be able to choose among 5 investment managers. Each investment manager invests either **only** into Investment **A** or **only** into Investment **B**, i.e., investment managers **stick** with their investment **within** the current investment task. Whether investment managers stick with investment **A** or investment **B** is determined **randomly** with **50%** chance at the beginning of each investment task. After you choose your investment manager, he will invest into the investment he sticks with and you will obtain the payoff from the investment. **The payoffs and their probabilities for both investments are provided to you in the chart below.**

For you to judge the investment manager, the following information is provided to you:

1. **The payoff** obtained by each investment manager the last time he invested into the investment he sticks with.
2. **The ranking** of investment managers based on these payoffs.

Important: This investment task will be repeated 5 times. Each investment task is **independent** from previous and future investment tasks. This means that the investments to which investment managers stick in this **current** investment task have **nothing to do** with the investments they stick with in **previous** or **future** investment tasks.



For your convenience, these instructions will remain available to you on all subsequent screens of this study.

*Screen 7 (based on **Treatment 3**):*

Choice of investment manager (**Practice round** in block 1)

Investment manager	**Last payoffs realized**	**Ranking of last payoffs realized**
Investment manager 1	$1.25	4
Investment manager 2	$2.50	1
Investment manager 3	$1.75	3
Investment manager 4	$0.50	5
Investment manager 5	$2.00	2

I wish to invest with:

- **Manager 1**
- **Manager 2**
- **Manager 3**
- **Manager 4**
- **Manager 5**

Next

Show calculator

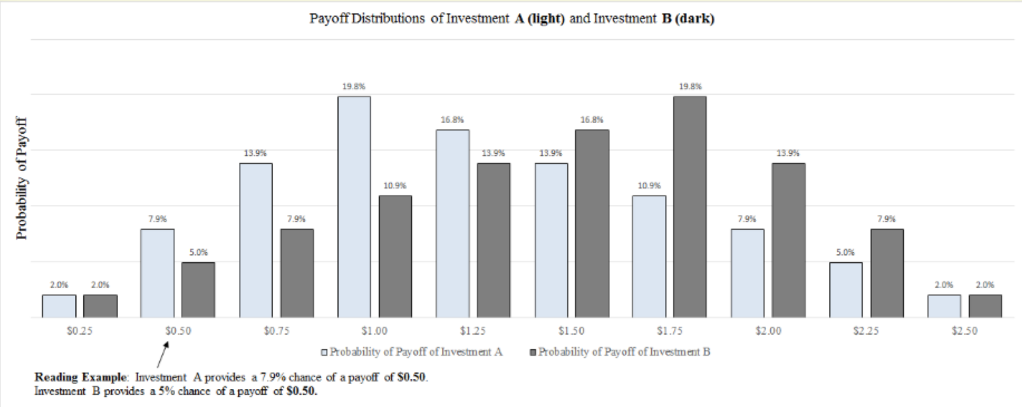
Instructions

You have to choose **one** investment manager who you believe is best for you. You will be able to choose among 5 investment managers. Each investment manager invests either **only** into Investment **A** or **only** into Investment **B**, i.e., investment managers **stick** with their investment **within** the current investment task. Whether investment managers stick with investment **A** or investment **B** is determined **randomly** with **50%** chance at the beginning of each investment task. After you choose your investment manager, he will invest into the investment he sticks with and you will obtain the payoff from the investment. **The payoffs and their probabilities for both investments are provided to you in the chart below**.

For you to judge the investment manager, the following information is provided to you:

1. **The payoff** obtained by each investment manager the last time he invested into the investment he sticks with.
2. **The ranking** of investment managers based on these payoffs.

Important: This investment task will be repeated 5 times. Each investment task is **independent** from previous and future investment tasks. This means that the investments to which investment managers stick in this **current** investment task have **nothing to do** with the investments they stick with in **previous** or **future** investment tasks.



For your convenience, these instructions will remain available to you on all subsequent screens of this study.

*Screen 8 (based on **Treatment 3**):*

Results (**Practice round** in block 1)

You have chosen to invest with **investment manager 2**, who sticks with investment **B**.
Your realized payoff with this investment is **$0.50**.

Next

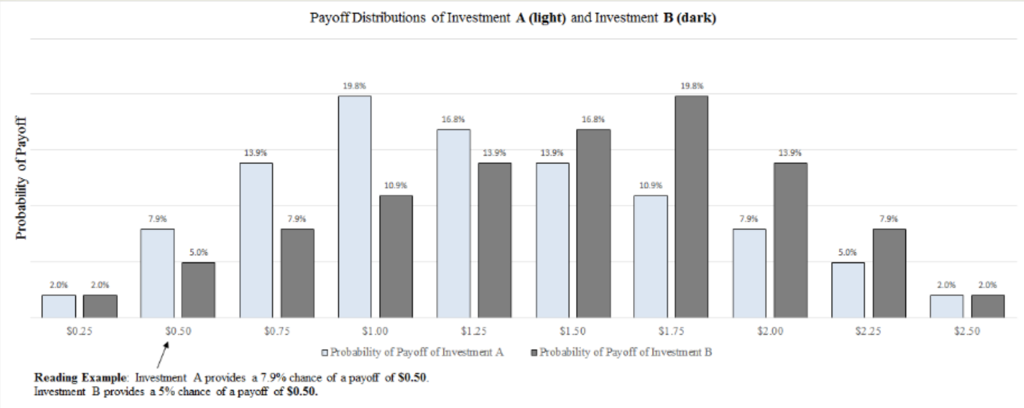
Instructions

You have to choose **one** investment manager who you believe is best for you. You will be able to choose among 5 investment managers. Each investment manager invests either **only** into Investment **A** or **only** into Investment **B**, i.e., investment managers **stick** with their investment **within** the current investment task. Whether investment managers stick with investment **A** or investment **B** is determined **randomly** with **50%** chance at the beginning of each investment task. After you choose your investment manager, he will invest into the investment he sticks with and you will obtain the payoff from the investment. **The payoffs and their probabilities for both investments are provided to you in the chart below**.

For you to judge the investment manager, the following information is provided to you:

1. **The payoff** obtained by each investment manager the last time he invested into the investment he sticks with.
2. **The ranking** of investment managers based on these payoffs.

Important: This investment task will be repeated 5 times. Each investment task is **independent** from previous and future investment tasks. This means that the investments to which investment managers stick in this **current** investment task have **nothing to do** with the investments they stick with in **previous** or **future** investment tasks.



For your convenience, these instructions will remain available to you on all subsequent screens of this study.

*Screen 9 (based on **Treatment 3**): Repeated*

Choice of investment manager (Round 1/5 in block 1)

Investment manager	Last payoffs realized	Ranking of last payoffs realized
Investment manager 1	$0.75	4
Investment manager 2	$1.00	3
Investment manager 3	$1.25	2
Investment manager 4	$1.50	1
Investment manager 5	$1.25	2

I wish to invest with:

- **Manager 1**
- **Manager 2**
- **Manager 3**
- **Manager 4**
- **Manager 5**

Next

Show calculator

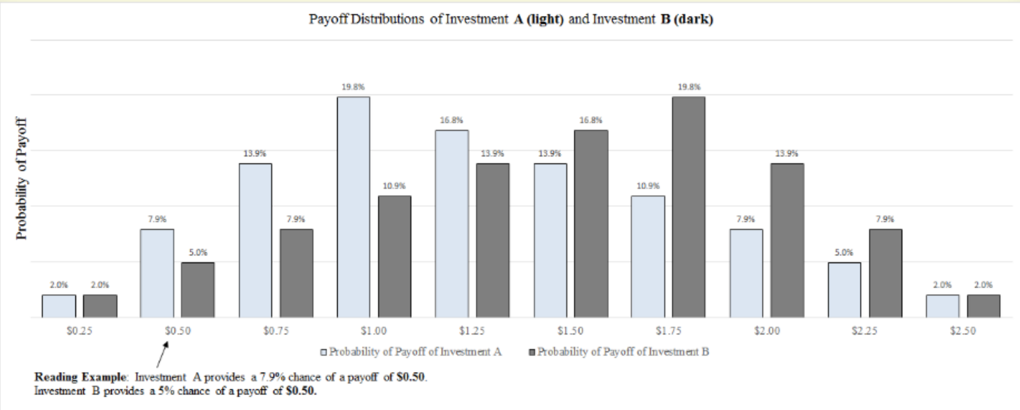
Instructions

You have to choose **one** investment manager who you believe is best for you. You will be able to choose among 5 investment managers. Each investment manager invests either **only** into Investment **A** or **only** into Investment **B**, i.e., investment managers **stick** with their investment **within** the current investment task. Whether investment managers stick with investment **A** or investment **B** is determined **randomly** with **50%** chance at the beginning of each investment task. After you choose your investment manager, he will invest into the investment he sticks with and you will obtain the payoff from the investment. **The payoffs and their probabilities for both investments are provided to you in the chart below**.

For you to judge the investment manager, the following information is provided to you:

1. **The payoff** obtained by each investment manager the last time he invested into the investment he sticks with.
2. **The ranking** of investment managers based on these payoffs.

Important: This investment task will be repeated 5 times. Each investment task is **independent** from previous and future investment tasks. This means that the investments to which investment managers stick in this **current** investment task have **nothing to do** with the investments they stick with in **previous** or **future** investment tasks.



For your convenience, these instructions will remain available to you on all subsequent screens of this study.

*Screen 10 (based on **Treatment 3**): Repeated*

Results (Round 1/5 in block 1)

You have chosen to invest with **investment manager 2**, who sticks with investment **A**.
Your realized payoff with this investment is **$1.00**.

Next

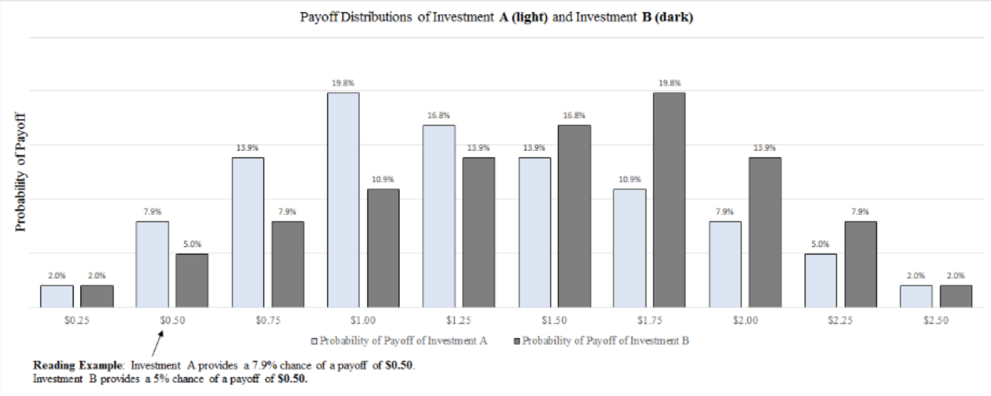
Instructions

You have to choose **one** investment manager who you believe is best for you. You will be able to choose among 5 investment managers. Each investment manager invests either **only** into Investment **A** or **only** into Investment **B**, i.e., investment managers **stick** with their investment **within** the current investment task. Whether investment managers stick with investment **A** or investment **B** is determined **randomly** with **50%** chance at the beginning of each investment task. After you choose your investment manager, he will invest into the investment he sticks with and you will obtain the payoff from the investment. **The payoffs and their probabilities for both investments are provided to you in the chart below**.

For you to judge the investment manager, the following information is provided to you:

1. **The payoff** obtained by each investment manager the last time he invested into the investment he sticks with.
2. **The ranking** of investment managers based on these payoffs.

Important: This investment task will be repeated 5 times. Each investment task is **independent** from previous and future investment tasks. This means that the investments to which investment managers stick in this **current** investment task have **nothing to do** with the investments they stick with in **previous** or **future** investment tasks.



For your convenience, these instructions will remain available to you on all subsequent screens of this study.

*Screen 11 (based on **Treatment 1**):*

Instructions for block 2

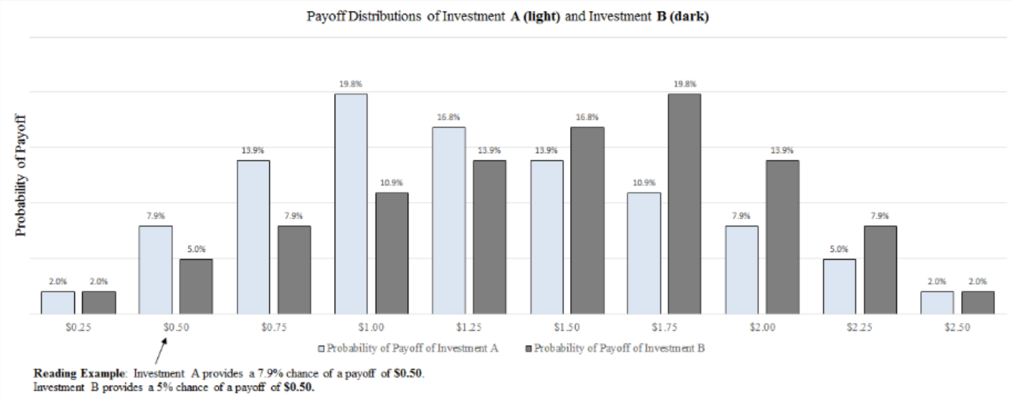
Instructions

You have to choose **one** investment manager who you believe is best for you. You will be able to choose among 5 investment managers. Each investment manager invests either **only** into Investment **A** or **only** into Investment **B**, i.e., investment managers **stick** with their investment **within** the current investment task. After you choose your investment manager, he will invest into the investment he sticks with and you will obtain the payoff from the investment. **The payoffs and their probabilities for both investments are provided to you in the chart below**.

For you to judge the investment manager, the following information is provided to you:

1. **The investment** the investment manager sticks with.
2. **The payoff** obtained by each investment manager the last time he invested into the investment he sticks with.
3. **The ranking** of investment managers based on these payoffs.

Important: This investment task will be repeated 5 times. Each investment task is **independent** from previous and future investment tasks. This means that the investments to which investment managers stick in this **current** investment task have **nothing to do** with the investments they stick with in **previous** or **future** investment tasks.



For your convenience, these instructions will remain available to you on all subsequent screens of this study.

Next

*Screen 12 (based on **Treatment 1**):*

Investment preferences

Please indicate into which investment you would invest if you could invest yourself:

- Investment A
- Investment B

Next

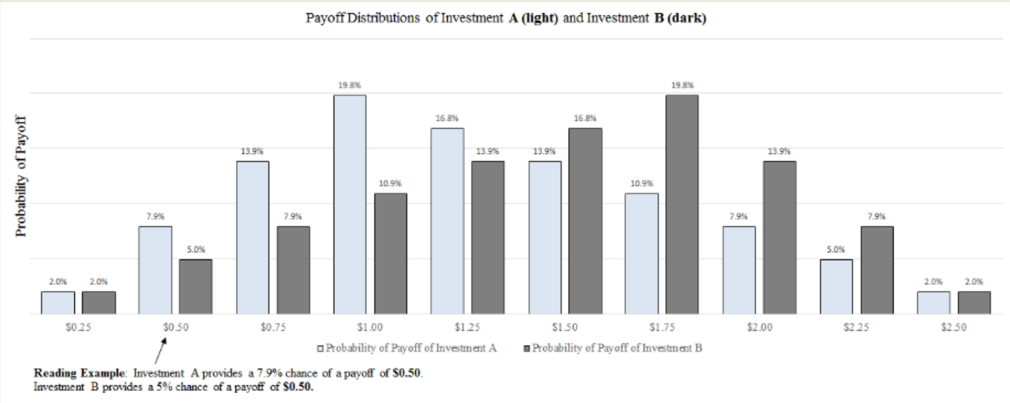
Instructions

You have to choose **one** investment manager who you believe is best for you. You will be able to choose among 5 investment managers. Each investment manager invests either **only** into Investment **A** or **only** into Investment **B**, i.e., investment managers **stick** with their investment **within** the current investment task. After you choose your investment manager, he will invest into the investment he sticks with and you will obtain the payoff from the investment. **The payoffs and their probabilities for both investments are provided to you in the chart below**.

For you to judge the investment manager, the following information is provided to you:

1. **The investment** the investment manager sticks with.
2. **The payoff** obtained by each investment manager the last time he invested into the investment he sticks with.
3. **The ranking** of investment managers based on these payoffs.

Important: This investment task will be repeated 5 times. Each investment task is **independent** from previous and future investment tasks. This means that the investments to which investment managers stick in this **current** investment task have **nothing to do** with the investments they stick with in **previous** or **future** investment tasks.



For your convenience, these instructions will remain available to you on all subsequent screens of this study.

*Screen 13 (based on **Treatment 1**):*

Choice of investment manager (**Practice round** in block 2)

Investment manager	**Last payoffs realized**	**Investment that manager has chosen**	**Ranking of last payoffs realized**
Investment manager 1	$2.50	A	1
Investment manager 2	$1.00	B	4
Investment manager 3	$2.25	B	2
Investment manager 4	$1.00	A	4
Investment manager 5	$1.75	A	3

I wish to invest with:

- **Manager 1**
- **Manager 2**
- **Manager 3**
- **Manager 4**
- **Manager 5**

Next

Show calculator

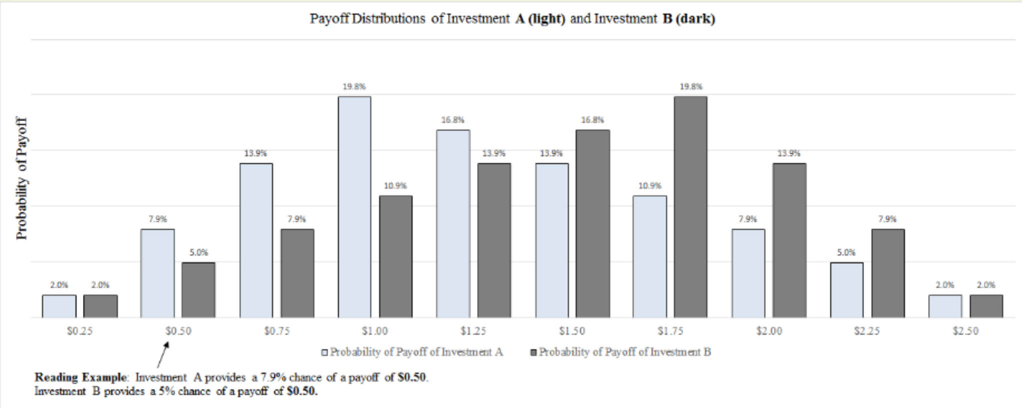
Instructions

You have to choose **one** investment manager who you believe is best for you. You will be able to choose among 5 investment managers. Each investment manager invests either **only** into Investment **A** or **only** into Investment **B**, i.e., investment managers **stick** with their investment **within** the current investment task. After you choose your investment manager, he will invest into the investment he sticks with and you will obtain the payoff from the investment. **The payoffs and their probabilities for both investments are provided to you in the chart below**.

For you to judge the investment manager, the following information is provided to you:

1. **The investment** the investment manager sticks with.
2. **The payoff** obtained by each investment manager the last time he invested into the investment he sticks with.
3. **The ranking** of investment managers based on these payoffs.

Important: This investment task will be repeated 5 times. Each investment task is **independent** from previous and future investment tasks. This means that the investments to which investment managers stick in this **current** investment task have **nothing to do** with the investments they stick with in **previous** or **future** investment tasks.



For your convenience, these instructions will remain available to you on all subsequent screens of this study.

*Screen 14 (based on **Treatment 1**):*

Results (**Practice round** in block 2)

You have chosen to invest with **investment manager 4**, who sticks with investment **A**.
Your realized payoff with this investment is **$1.50**.

Next

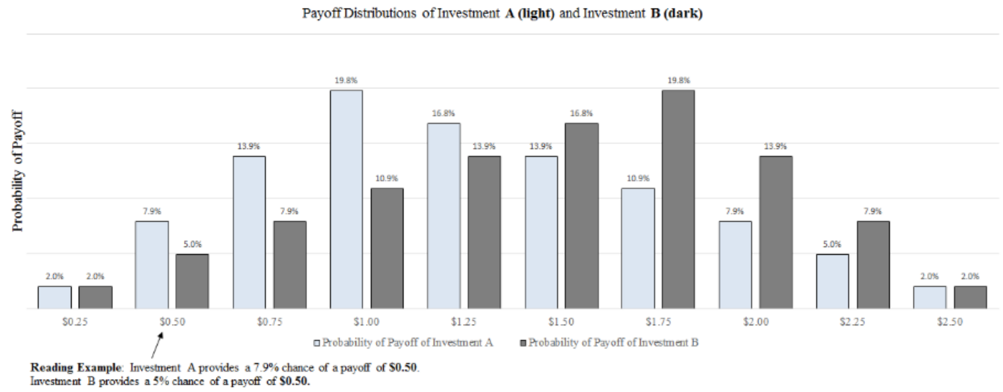
Instructions

You have to choose **one** investment manager who you believe is best for you. You will be able to choose among 5 investment managers. Each investment manager invests either **only** into Investment **A** or **only** into Investment **B**, i.e., investment managers **stick** with their investment **within** the current investment task. After you choose your investment manager, he will invest into the investment he sticks with and you will obtain the payoff from the investment. **The payoffs and their probabilities for both investments are provided to you in the chart below**.

For you to judge the investment manager, the following information is provided to you:

1. **The investment** the investment manager sticks with.
2. **The payoff** obtained by each investment manager the last time he invested into the investment he sticks with.
3. **The ranking** of investment managers based on these payoffs.

Important: This investment task will be repeated 5 times. Each investment task is **independent** from previous and future investment tasks. This means that the investments to which investment managers stick in this **current** investment task have **nothing to do** with the investments they stick with in **previous** or **future** investment tasks.



For your convenience, these instructions will remain available to you on all subsequent screens of this study.

Screen 15 (based on ***Treatment 1****): Repeated*

Choice of investment manager (Round 1/5 in block 2)

Investment manager	Last payoffs realized	Investment that manager has chosen	Ranking of last payoffs realized
Investment manager 1	$2.00	A	2
Investment manager 2	$1.75	B	3
Investment manager 3	$2.25	B	1
Investment manager 4	$1.25	B	4
Investment manager 5	$0.50	A	5

I wish to invest with:

- **Manager 1**
- **Manager 2**
- **Manager 3**
- **Manager 4**
- **Manager 5**

Next

Show calculator

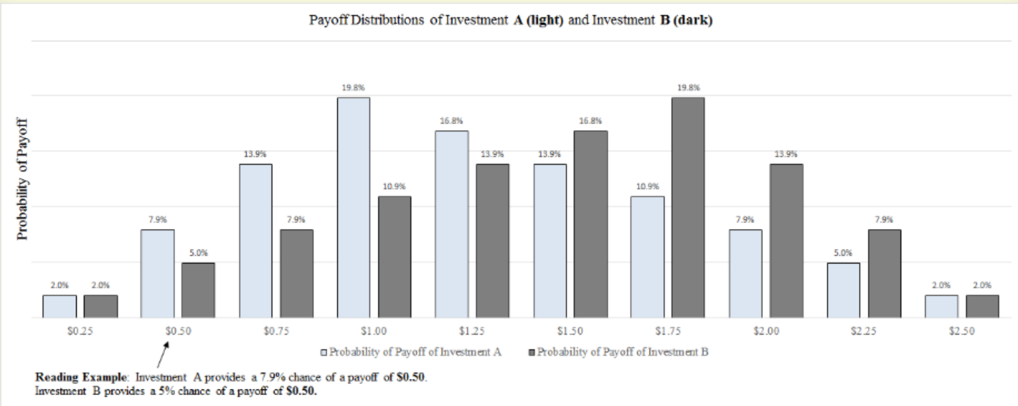
Instructions

You have to choose **one** investment manager who you believe is best for you. You will be able to choose among 5 investment managers. Each investment manager invests either **only** into Investment **A** or **only** into Investment **B**, i.e., investment managers **stick** with their investment **within** the current investment task. After you choose your investment manager, he will invest into the investment he sticks with and you will obtain the payoff from the investment. **The payoffs and their probabilities for both investments are provided to you in the chart below**.

For you to judge the investment manager, the following information is provided to you:

1. **The investment** the investment manager sticks with.
2. **The payoff** obtained by each investment manager the last time he invested into the investment he sticks with.
3. **The ranking** of investment managers based on these payoffs.

Important: This investment task will be repeated 5 times. Each investment task is **independent** from previous and future investment tasks. This means that the investments to which investment managers stick in this **current** investment task have **nothing to do** with the investments they stick with in **previous** or **future** investment tasks.



For your convenience, these instructions will remain available to you on all subsequent screens of this study.

*Screen 16 (based on **Treatment 1**): Repeated*

Results (Round 1/5 in block 2)

You have chosen to invest with **investment manager 5**, who sticks with investment **A**.
Your realized payoff with this investment is **$0.75**.

Next

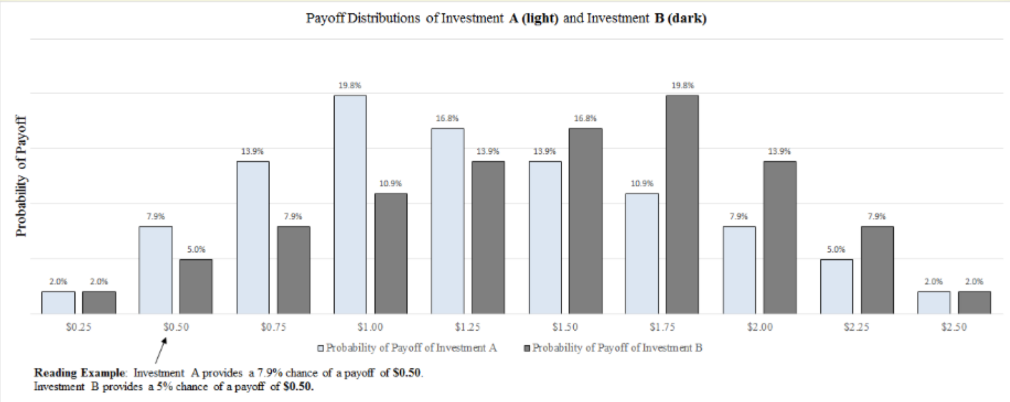
Instructions

You have to choose **one** investment manager who you believe is best for you. You will be able to choose among 5 investment managers. Each investment manager invests either **only** into Investment **A** or **only** into Investment **B**, i.e., investment managers **stick** with their investment **within** the current investment task. After you choose your investment manager, he will invest into the investment he sticks with and you will obtain the payoff from the investment. **The payoffs and their probabilities for both investments are provided to you in the chart below**.

For you to judge the investment manager, the following information is provided to you:

1. **The investment** the investment manager sticks with.
2. **The payoff** obtained by each investment manager the last time he invested into the investment he sticks with.
3. **The ranking** of investment managers based on these payoffs.

Important: This investment task will be repeated 5 times. Each investment task is **independent** from previous and future investment tasks. This means that the investments to which investment managers stick in this **current** investment task have **nothing to do** with the investments they stick with in **previous** or **future** investment tasks.



For your convenience, these instructions will remain available to you on all subsequent screens of this study.

*Screen 17 (based on **Treatment 2**):*

Instructions for block 3

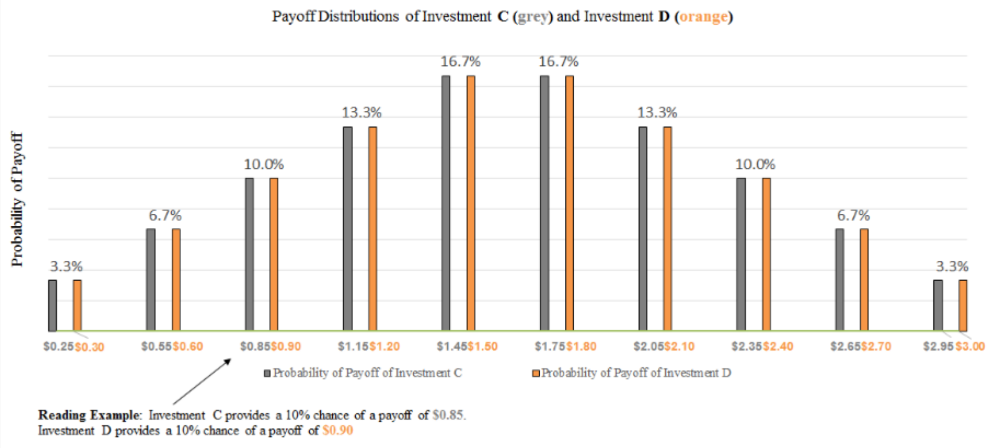
Instructions

You have to choose **one** investment manager who you believe is best for you. You will be able to choose among 5 investment managers. Each investment manager invests either **only** into Investment **C** or **only** into Investment **D**, i.e., investment managers **stick** with their investment **within** the current investment task. Whether investment managers stick with investment **C** or investment **D** is determined **randomly** with **50%** chance at the beginning of each investment task. After you choose your investment manager, he will invest into the investment he sticks with and you will obtain the payoff from the investment. **The payoffs and their probabilities for both investments are provided to you in the chart below**.

For you to judge the investment manager, the following information is provided to you:

1. **The payoff** obtained by each investment manager the last time he invested into the investment he sticks with.
2. **The ranking** of investment managers based on these payoffs.

Important: This investment task will be repeated 5 times. Each investment task is **independent** from previous and future investment tasks. This means that the investments to which investment managers stick in this **current** investment task have **nothing to do** with the investments they stick with in **previous** or **future** investment tasks.



For your convenience, these instructions will remain available to you on all subsequent screens of this study.

Next

*Screen 18 (based on **Treatment 2**):*

Investment preferences

Please indicate into which investment you would invest if you could invest yourself:

- Investment C
- Investment D

Next

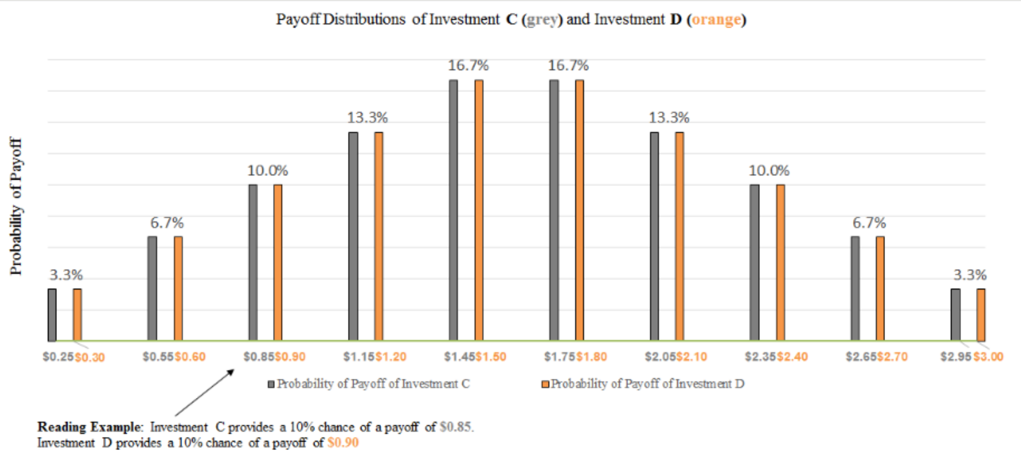
Instructions

You have to choose **one** investment manager who you believe is best for you. You will be able to choose among 5 investment managers. Each investment manager invests either **only** into Investment **C** or **only** into Investment **D**, i.e., investment managers **stick** with their investment **within** the current investment task. Whether investment managers stick with investment **C** or investment **D** is determined **randomly** with **50%** chance at the beginning of each investment task. After you choose your investment manager, he will invest into the investment he sticks with and you will obtain the payoff from the investment. **The payoffs and their probabilities for both investments are provided to you in the chart below.**

For you to judge the investment manager, the following information is provided to you:

1. **The payoff** obtained by each investment manager the last time he invested into the investment he sticks with.
2. **The ranking** of investment managers based on these payoffs.

Important: This investment task will be repeated 5 times. Each investment task is **independent** from previous and future investment tasks. This means that the investments to which investment managers stick in this **current** investment task have **nothing to do** with the investments they stick with in **previous** or **future** investment tasks.



For your convenience, these instructions will remain available to you on all subsequent screens of this study.

*Screen 19 (based on **Treatment 2**):*

Choice of investment manager (**Practice round** in block 3)

Investment manager	Last payoffs realized	Ranking of last payoffs realized
Investment manager 1	$1.15	2
Investment manager 2	$0.55	4
Investment manager 3	$1.80	1
Investment manager 4	$0.85	3
Investment manager 5	$1.80	1

I wish to invest with:

- **Manager 1**
- **Manager 2**
- **Manager 3**
- **Manager 4**
- **Manager 5**

Next

Show calculator

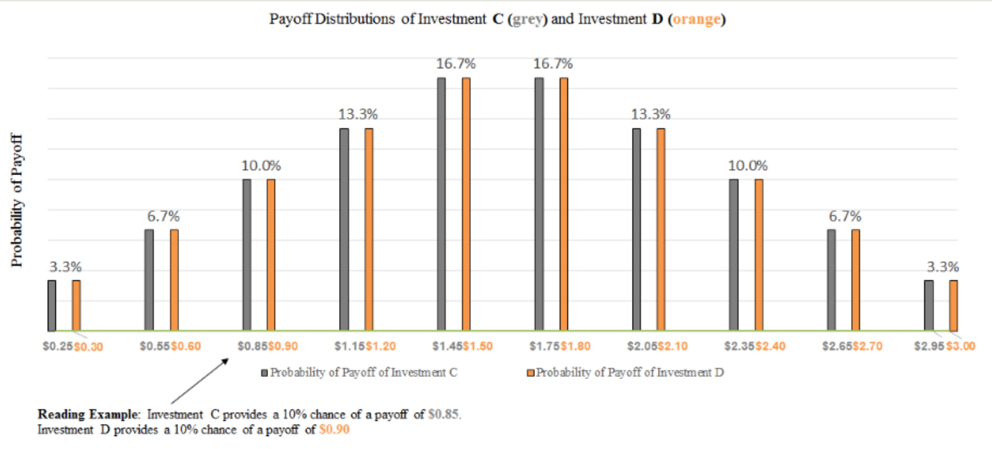
Instructions

You have to choose **one** investment manager who you believe is best for you. You will be able to choose among 5 investment managers. Each investment manager invests either **only** into Investment **C** or **only** into Investment **D**, i.e., investment managers **stick** with their investment **within** the current investment task. Whether investment managers stick with investment **C** or investment **D** is determined **randomly** with **50%** chance at the beginning of each investment task. After you choose your investment manager, he will invest into the investment he sticks with and you will obtain the payoff from the investment. **The payoffs and their probabilities for both investments are provided to you in the chart below**.

For you to judge the investment manager, the following information is provided to you:

1. **The payoff** obtained by each investment manager the last time he invested into the investment he sticks with.
2. **The ranking** of investment managers based on these payoffs.

Important: This investment task will be repeated 5 times. Each investment task is **independent** from previous and future investment tasks. This means that the investments to which investment managers stick in this **current** investment task have **nothing to do** with the investments they stick with in **previous** or **future** investment tasks.



For your convenience, these instructions will remain available to you on all subsequent screens of this study.

*Screen 20 (based on **Treatment 2**):*

Results (**Practice round** in block 3)

You have chosen to invest with **investment manager 4**, who sticks with investment **C**.
Your realized payoff with this investment is **$2.95**.

Next

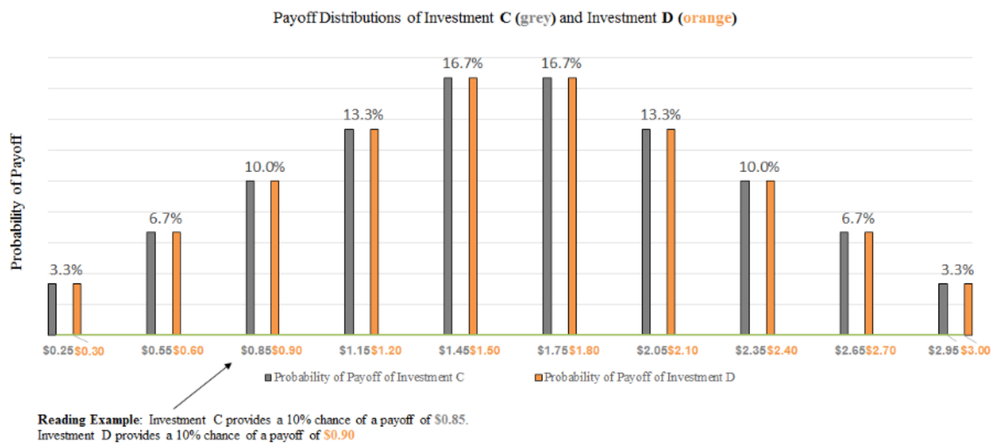
Instructions

You have to choose **one** investment manager who you believe is best for you. You will be able to choose among 5 investment managers. Each investment manager invests either **only** into Investment **C** or **only** into Investment **D**, i.e., investment managers **stick** with their investment **within** the current investment task. Whether investment managers stick with investment **C** or investment **D** is determined **randomly** with **50%** chance at the beginning of each investment task. After you choose your investment manager, he will invest into the investment he sticks with and you will obtain the payoff from the investment. **The payoffs and their probabilities for both investments are provided to you in the chart below.**

For you to judge the investment manager, the following information is provided to you:

1. **The payoff** obtained by each investment manager the last time he invested into the investment he sticks with.
2. **The ranking** of investment managers based on these payoffs.

Important: This investment task will be repeated 5 times. Each investment task is **independent** from previous and future investment tasks. This means that the investments to which investment managers stick in this **current** investment task have **nothing to do** with the investments they stick with in **previous** or **future** investment tasks.



For your convenience, these instructions will remain available to you on all subsequent screens of this study.

*Screen 21 (based on **Treatment 2**): Repeated*

Choice of investment manager (Round 1/5 in block 3)

Investment manager	Last payoffs realized	Ranking of last payoffs realized
Investment manager 1	$0.60	5
Investment manager 2	$2.70	2
Investment manager 3	$2.40	3
Investment manager 4	$1.50	4
Investment manager 5	$2.95	1

I wish to invest with:

- **Manager 1**
- **Manager 2**
- **Manager 3**
- **Manager 4**
- **Manager 5**

Next

Show calculator

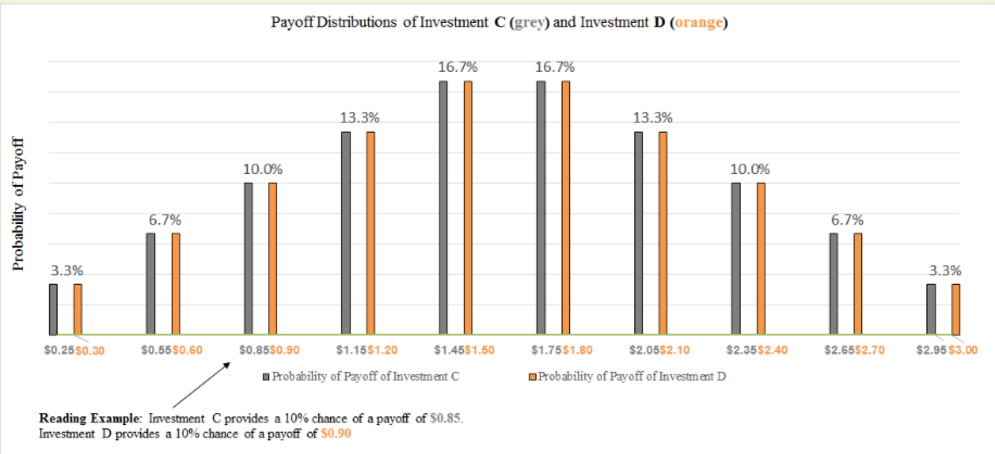
Instructions

You have to choose **one** investment manager who you believe is best for you. You will be able to choose among 5 investment managers. Each investment manager invests either **only** into Investment **C** or **only** into Investment **D**, i.e., investment managers **stick** with their investment **within** the current investment task. Whether investment managers stick with investment **C** or investment **D** is determined **randomly** with **50%** chance at the beginning of each investment task. After you choose your investment manager, he will invest into the investment he sticks with and you will obtain the payoff from the investment. **The payoffs and their probabilities for both investments are provided to you in the chart below**.

For you to judge the investment manager, the following information is provided to you:

1. **The payoff** obtained by each investment manager the last time he invested into the investment he sticks with.
2. **The ranking** of investment managers based on these payoffs.

Important: This investment task will be repeated 5 times. Each investment task is **independent** from previous and future investment tasks. This means that the investments to which investment managers stick in this **current** investment task have **nothing to do** with the investments they stick with in **previous** or **future** investment tasks.



For your convenience, these instructions will remain available to you on all subsequent screens of this study.

*Screen 22 (based on **Treatment 2**): Repeated*

Results (Round 1/5 in block 3)

You have chosen to invest with **investment manager 2**, who sticks with investment **C**.
Your realized payoff with this investment is **$2.65**.

Next

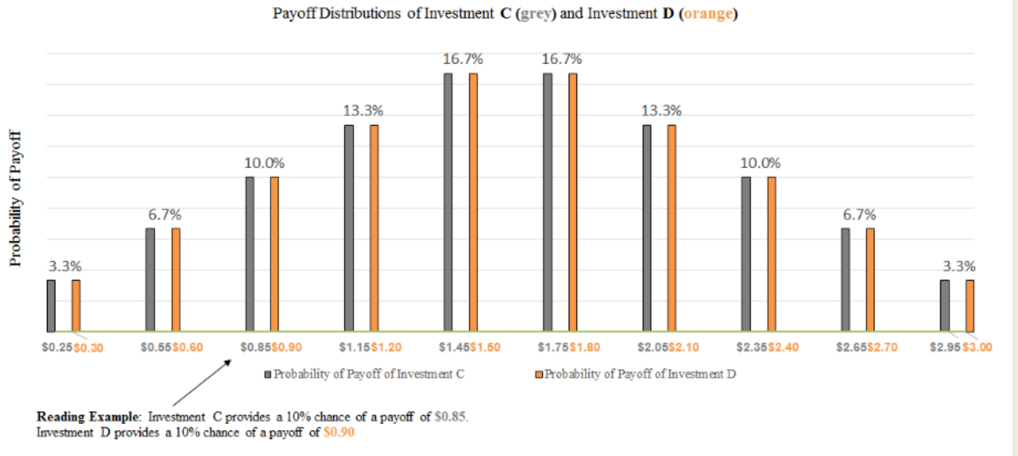
Instructions

You have to choose **one** investment manager who you believe is best for you. You will be able to choose among 5 investment managers. Each investment manager invests either **only** into Investment **C** or **only** into Investment **D**, i.e., investment managers **stick** with their investment **within** the current investment task. Whether investment managers stick with investment **C** or investment **D** is determined **randomly** with **50%** chance at the beginning of each investment task. After you choose your investment manager, he will invest into the investment he sticks with and you will obtain the payoff from the investment. **The payoffs and their probabilities for both investments are provided to you in the chart below**.

For you to judge the investment manager, the following information is provided to you:

1. **The payoff** obtained by each investment manager the last time he invested into the investment he sticks with.
2. **The ranking** of investment managers based on these payoffs.

Important: This investment task will be repeated 5 times. Each investment task is **independent** from previous and future investment tasks. This means that the investments to which investment managers stick in this **current** investment task have **nothing to do** with the investments they stick with in **previous** or **future** investment tasks.



For your convenience, these instructions will remain available to you on all subsequent screens of this study.

Screen 23:

Quiz (Part 1/2)

A bat and a ball costs $1.10 in total. The bat costs a dollar more than the ball. How much does the ball cost?

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

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

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

Next

Screen 24:

Quiz (Part 2/2)

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

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

ECU

Simon decided to invest \$8,000 in the investment market one day early in 2008. Six months after he invested, on July 17, the investments he had purchased were down 50%. Fortunately for Simon, from July 17 to October 17, the investments he had purchased went up 75%.

- broken even in the investment market
- is ahead of where he began
- has lost money

Next

0

Screen 25:

Socio-demographic information

Please enter your age in years:

Please indicate your sex:
- male
- female

Please indicate your nationality:

Please indicate the highest education qualification you have obtained:
- Middle School
- University entrance qualification
- Bachelor
- Master (or equivalent)
- PhD (or equivalent)
- other

0

Please indicate your current occupational status:
- High school student
- student
- employee
- self-employed
- retired
- other

Have you at some point in your life invested into corporate bonds:
- Yes
- No

Have you at some point in your life invested into government bonds:
- Yes
- No

Have you at some point in your life invested into index funds:
- Yes
- No

Have you at some point in your life invested into active/strategic funds:
- Yes
- No

Have you at some point in your life invested into options and/or other derivatives:
- Yes
- No

Please rate how difficult you found using the distributions of the assets shown to you in this experiment:

Very easy   Rather easy   Neutral   Rather difficult   Very difficult

Next

Screen 26:

Final results

Round 1	You have chosen to invest with investment manager 3, who sticks to investment B	Your realized payoff is 2.00
Round 2	You have chosen to invest with investment manager 3, who sticks to investment A	Your realized payoff is 1.75
Round 3	You have chosen to invest with investment manager 4, who sticks to investment B	Your realized payoff is 1.50
Round 4	You have chosen to invest with investment manager 5, who sticks to investment B	Your realized payoff is 1.75
Round 5	You have chosen to invest with investment manager 2, who sticks to investment B	Your realized payoff is 1.00
Round 6	You have chosen to invest with investment manager 3, who sticks to investment B	Your realized payoff is 2.25
Round 7	You have chosen to invest with investment manager 5, who sticks to investment B	Your realized payoff is 1.75
Round 8	You have chosen to invest with investment manager 5, who sticks to investment B	Your realized payoff is 1.75
Round 9	You have chosen to invest with investment manager 1, who sticks to investment B	Your realized payoff is 1.25
Round 10	You have chosen to invest with investment manager 5, who sticks to investment A	Your realized payoff is 1.00
Round 11	You have chosen to invest with investment manager 2, who sticks to investment C	Your realized payoff is 2.65
Round 12	You have chosen to invest with investment manager 3, who sticks to investment D	Your realized payoff is 1.50
Round 13	You have chosen to invest with investment manager 2, who sticks to investment C	Your realized payoff is 0.85
Round 14	You have chosen to invest with investment manager 2, who sticks to investment D	Your realized payoff is 1.80
Round 15	You have chosen to invest with investment manager 2, who sticks to investment D	Your realized payoff is 0.60
Total Experiment Payment	The chosen round was 5	Your realized payoff for the experiment is therefore 1.00

Next

Appendix B

Trust and Delegated Investing

Table B.1: Risky Share of Investment (By Round)

This table shows the share of wealth invested into the risky asset, *Risky Share*, for both money managers. HT, HC corresponds to the more trustworthy (i.e., returned more in the trust game) but more expensive money manager. LT, LC corresponds to the less trustworthy (i.e., returned less in the trust game) but less expensive money manager. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

	N	Mean Risky Share$_{HT,HC}$	Mean Risky Share$_{LT,LC}$	Paired t-test	Wilcoxon signed-rank
Round 1	86	42.97	26.05	0.000***	0.000***
Round 2	81	46.48	28.64	0.000***	0.000***
Round 3	85	48.06	28.94	0.000***	0.000***
Round 4	77	42.66	31.95	0.003***	0.001***
Round 5	81	50.49	31.11	0.000***	0.000***

Table B.2: Risky Share of Investment for Identically Trustworthy Money Managers (By Round)

This table shows the share of wealth invested into the risky asset, *Risky Share*, for both money managers when both money managers are equal in trustworthiness (i.e., returned the same amounts in the trust game). The type of costs, high or low, is indicated by subscripts $_{HC}$ and $_{LC}$, respectively. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

		All Participants			
	N	Mean Risky Share$_{LC}$	Mean Risky Share$_{HC}$	Paired t-test	Wilcoxon signed-rank
Round 1	28	31.96	18.21	0.033**	0.013**
Round 2	33	33.33	24.94	0.152	0.013**
Round 3	29	30.17	27.76	0.736	0.111
Round 4	37	36.27	28.05	0.181	0.045**
Round 5	33	30.94	28.52	0.696	0.108
		Only Participants Who Sent > 0 ECU			
Round 1	15	31.67	15.67	0.050*	0.033**
Round 2	20	37.25	22.65	0.100*	0.012**
Round 3	16	31.88	20.94	0.291	0.090*
Round 4	24	42.71	25.83	0.024**	0.025**
Round 5	20	32.50	27.75	0.538	0.279

Table B.3: Amount Sent – Risky Share and Indifference Costs

This table reports regression results with the amount participants sent in the trust game, *Amount Sent*, as independent variable. For this regression all observations of one treatment were pooled. The value of *Amount Sent* is fixed for an individual for all five rounds of a treatment. In column (1) the dependent variable is the share investors invested risky with the first money manager in Treatment 1. In column (2) the dependent variable is the share investors invested risky with the second money manager in Treatment 1. In column (3) the dependent variable is the indifference costs investors specified.
Standard errors clustered at the individual level are shown in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

	(1)	(2)	(3)
	Random Effects	Random Effects	Random Effects
	Risky Share (in %) *1st Money Manager*	Risky Share (in %) *2nd Money Manager*	Indifference Costs
Amount Sent	0.237***	0.216***	0.0087**
	(0.060)	(0.060)	(0.0044)
Constant	27.13***	17.69***	1.413***
	(3.721)	(3.447)	(0.272)
Observations	570	570	570
Cluster-robust S.E.	YES	YES	YES
Round FE	YES	YES	YES
$R^2_{overall}$	0.073	0.080	0.024

Table B.4: Risky Share – Difference in Trustworthiness (Tobit)

This table reports random effects tobit regression results with Δ*Risky Share* as dependent variable. It is calculated as the share of wealth invested risky with the more trustworthy money manager minus the share of wealth invested risky with the less trustworthy money manager. In case both managers are equally trustworthy, it is calculated as the share of wealth invested risky with the more costly manager minus the share of wealth invested risky with the less costly manager. All regressions account for unobserved individual heterogeneity through random effects. Δ *Trustworthiness Absolute* is calculated as the amount the more trustworthy manager returned in the trust game minus the amount the less trustworthy manager returned in the trust game. Δ *Trustworthiness Relative* is calculated as $(1 - (\frac{Lower\ Returned\ Amount}{Higher\ Returned\ Amount})) * 100$. Δ *Trustworthiness Relative to Sent* is calculated as $(\frac{Higher\ Returned\ Amount\ -\ Lower\ Returned\ Amount}{Amount\ Sent}) * 100$.
Δ*Risky Share* is censored at -100 and +100.
Bootstrapped standard errors (100 repetitions) are shown in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

	(1)	(2)	(3)
	Random Effects Tobit		
Δ*Trustworthiness Absolute*	0.341***		
	(0.069)		
Δ*Trustworthiness Relative*		0.252***	
		(0.041)	
Δ*Trustworthiness Relative to Sent*			0.179***
			(0.032)
Constant	-0.357	-3.423	-1.531
	(4.108)	(4.630)	(4.559)
Observations	570	570	570
Bootstrapped S.E.	YES	YES	YES
Round FE	YES	YES	YES
Log-likelihood	-2735	-2739	-2737

Table B.5: Risky Share – Difference in Trustworthiness (FE)

This table reports regression results with Δ*Risky Share* as dependent variable. It is calculated as the share of wealth invested risky with the more trustworthy money manager minus the share of wealth invested risky with the less trustworthy money manager. In case both managers are equally trustworthy, it is calculated as the share of wealth invested risky with the more costly manager minus the share of wealth invested risky with the less costly manager. All regressions account for unobserved individual heterogeneity through fixed effects. Δ *Trustworthiness Absolute* is calculated as the amount the more trustworthy manager returned in the trust game minus the amount the less trustworthy manager returned in the trust game. Δ *Trustworthiness Relative* is calculated as $(1 - (\frac{Lower\ Returned\ Amount}{Higher\ Returned\ Amount})) * 100$. Δ *Trustworthiness Relative to Sent* is calculated as $(\frac{Higher\ Returned\ Amount\ -\ Lower\ Returned\ Amount}{Amount\ Sent}) * 100$.
Standard errors clustered at the individual level are shown in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

	(1)	(2)	(3)
		Fixed Effects	
Δ*Trustworthiness Absolute*	0.384***		
	(0.081)		
Δ*Trustworthiness Relative*		0.290***	
		(0.047)	
Δ*Trustworthiness Relative to Sent*			0.190***
			(0.033)
Constant	-2.307	-6.044	-2.815
	(3.651)	(3.813)	(3.370)
Observations	570	570	570
Cluster-robust S.E.	YES	YES	YES
Round FE	YES	YES	YES
$R^2_{adjusted}$	0.090	0.095	0.092

Table B.6: Indifference Costs (By Round)

This table shows indifference costs of investing with the second money manager in Treatment 2. Tests are based against the costs the first money manager charges, which are equal to 0.75%. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

	N	Mean Indifference Costs	Exogenous Costs	Paired t-test	Wilcoxon signed-rank
Trustworthiness Second Manager > First Manager					
Round 1	80	2.24	0.75	0.000***	0.000***
Round 2	86	1.98	0.75	0.000***	0.000***
Round 3	79	1.71	0.75	0.000***	0.000***
Round 4	81	2.02	0.75	0.000***	0.000***
Round 5	86	1.78	0.75	0.000***	0.000***
Trustworthiness Second Manager = First Manager					
Round 1	34	0.71	0.75	0.835	0.037**
Round 2	28	0.91	0.75	0.545	0.087*
Round 3	35	0.81	0.75	0.715	0.375
Round 4	33	0.92	0.75	0.444	0.180
Round 5	28	0.89	0.75	0.598	0.194
Trustworthiness Second Manager = First Manager Only Participants Who Sent > 0 ECU					
Round 1	21	0.61	0.75	0.141	0.091*
Round 2	15	1.13	0.75	0.307	0.477
Round 3	22	0.73	0.75	0.834	0.625
Round 4	20	1.00	0.75	0.373	0.478
Round 5	15	0.86	0.75	0.765	0.393

Table B.7: Indifference Costs – Difference in Trustworthiness (Tobit)

This table reports random effects tobit regression results with *Indifference Costs* as dependent variable. All regressions account for unobserved individual heterogeneity through random effects. Δ *Trustworthiness Absolute* is calculated as the amount the second manager returned in the trust game minus the amount the first manager returned in the trust game. Δ *Trustworthiness Relative* is calculated as $(1 - (\frac{Lower\ Returned\ Amount}{Higher\ Returned\ Amount})) * 100$. Δ *Trustworthiness Relative to Sent* is calculated as $(\frac{Higher\ Returned\ Amount\ -\ Lower\ Returned\ Amount}{Amount\ Sent}) * 100$.
Indifference Costs is censored at 0 and +10.
Bootstrapped standard errors (100 repetitions) are shown in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

	(1)	(2)	(3)
	Random Effects Tobit		
ΔTrustworthiness Absolute	0.0102**		
	(0.0045)		
ΔTrustworthiness Relative		0.0073***	
		(0.0025)	
ΔTrustworthiness Relative to Sent			0.0075***
			(0.0022)
Constant	1.358***	1.309***	1.208***
	(0.196)	(0.184)	(0.188)
Observations	570	570	570
Bootstrapped S.E.	YES	YES	YES
Round FE	YES	YES	YES
Log-likelihood	-1074	-1075	-1067

Table B.8: Indifference Costs – Difference in Trustworthiness (FE)

This table reports regression results with *Indifference Costs* as dependent variable. All regressions account for unobserved individual heterogeneity through fixed effects. Δ *Trustworthiness Absolute* is calculated as the amount the second manager returned in the trust game minus the amount the first manager returned in the trust game. Δ *Trustworthiness Relative* is calculated as $(1 - (\frac{Lower\ Returned\ Amount}{Higher\ Returned\ Amount})) * 100$. Δ *Trustworthiness Relative to Sent* is calculated as $(\frac{Higher\ Returned\ Amount\ -\ Lower\ Returned\ Amount}{Amount\ Sent}) * 100$.
Standard errors clustered at the individual level are shown in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

	(1)	(2)	(3)
		Fixed Effects	
Δ*Trustworthiness Absolute*	0.0094**		
	(0.0039)		
Δ*Trustworthiness Relative*		0.0056***	
		(0.0022)	
Δ*Trustworthiness Relative to Sent*			0.0064***
			(0.0019)
Constant	1.527***	1.531***	1.417***
	(0.125)	(0.115)	(0.135)
Observations	570	570	570
Cluster-robust S.E.	YES	YES	YES
Round FE	YES	YES	YES
$R^2_{adjusted}$	0.031	0.020	0.053

Table B.9: Risky Share – Robustness (Tobit)

This table reports random effects tobit regression results with Δ *Risky More* − *Risky Less* as dependent variable. It is calculated as the share of wealth invested risky with the more trustworthy money manager minus the share of wealth invested risky with the less trustworthy money manager. The regression accounts for unobserved individual heterogeneity through random effects. *Biased Beliefs* is an indicator variable equal to 1 if participants stated that they believed that more trustworthy money managers could deliver better investment performance. *Reward Motivation* is an indicator variable equal to 1 if participants stated that they invested more risky with more trustworthy money managers because they wanted to reward them.
Δ *Risky More* − *Risky Less* is censored at -100 and +100.
Bootstrapped standard errors (100 repetitions) are shown in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

	Random Effects Tobit
Biased Beliefs	-12.67
	(14.47)
Reward Motivation	1.170
	(9.357)
Biased Beliefs×*Reward Motivation*	18.21
	(16.84)
Constant	18.39**
	(8.825)
Observations	322
Bootstrapped S.E.	YES
Round FE	YES
Log-likelihood	-1525

Table B.10: Indifference Costs – Robustness (Tobit)

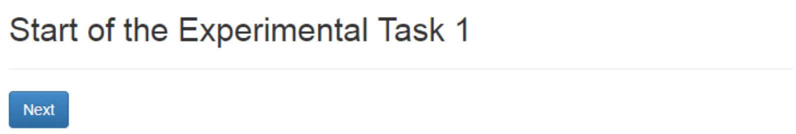
This table reports regression results with *Indifference Costs* as dependent variable, for cases in which the second money manager is more trustworthy than the first money manager. The regression accounts for unobserved individual heterogeneity through random effects. *Biased Beliefs* is an indicator variable equal to 1 if participants stated that they believed that more trustworthy money managers could deliver better investment performance. *Reward Motivation* is an indicator variable equal to 1 if participants stated that they invested more risky with more trustworthy money managers because they wanted to reward them.
Indifference Costs is censored at 0 and +10.
Bootstrapped standard errors (100 repetitions) are shown in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

	Random Effects Tobit
Biased Beliefs	0.214
	(0.652)
Reward Motivation	0.010
	(0.656)
Biased Beliefs × *Reward Motivation*	0.644
	(0.857)
Constant	1.751***
	(0.527)
Observations	324
Bootstrapped S.E.	YES
Round FE	YES
Log-likelihood	-659

Experimental Instructions

The following images show instructions and experimental screens as presented to participants. All realized values shown in the experimental screens are for illustration purposes only.

Screen 1:



Screen 2:

Introduction

Instructions

You have been randomly and anonymously paired with another participant. One of you will be randomly selected to be participant A, the other will be selected to be participant B. To start with, participant A receives 100 experimental currency units (ECU), participant B receives nothing. Participant A can send any integer amount of tens between 0 and 100 (i.e., 10, 20, 30, etc.) ECU to participant B. The amount participant B receives will be **tripled**. In the next step participant B can decide to return any integer amount of tens between 0 and the tripled amount.

Payment: Your payment from the experiment depends on your and your counterpart's choice in this game and whether you were selected as participant A or participant B. The conversaion rate is as follows: 1 experimental currency unit is worth 0.05 Euro.

For your convenience, these instructions will remain available to you on all subsequent screens of this study.

Next

Screen 3:

Indicate your choices

Suppose you were participant A and had the choice to send an integer amount of tens between 0 and 100:

Please indicate the amount you would like to send:

ECU

Now suppose you were participant B. Please indicate how much of the tripled amount you would like to return. Again you can only send integer amounts of tens.

If participant A had sent 10 ECU, I would like to return:

ECU

If participant A had sent 20 ECU, I would like to return:

ECU

If participant A had sent 30 ECU, I would like to return:

ECU

If participant A had sent 40 ECU, I would like to return:

ECU

If participant A had sent 50 ECU, I would like to return:

ECU

If participant A had sent 60 ECU, I would like to return:

ECU

If participant A had sent 70 ECU, I would like to return:

ECU

If participant A had sent 80 ECU, I would like to return:

ECU

If participant A had sent 90 ECU, I would like to return:

ECU

If participant A had sent 100 ECU, I would like to return:

ECU

Next

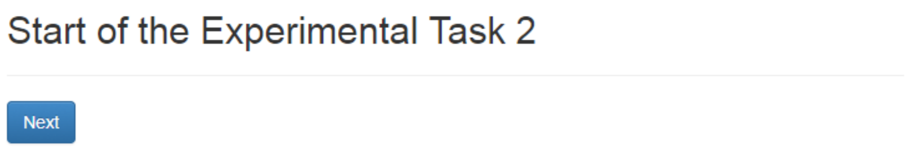
Instructions

You have been randomly and anonymously paired with another participant. One of you will be randomly selected to be participant A, the other will be selected to be participant B. To start with, participant A receives 100 experimental currency units (ECU), participant B receives nothing. Participant A can send any integer amount of tens between 0 and 100 (i.e., 10, 20, 30, etc.) ECU to participant B. The amount participant B receives will be **tripled**. In the next step participant B can decide to return any integer amount of tens between 0 and the tripled amount.

Payment: Your payment from the experiment depends on your and your counterpart's choice in this game and whether you were selected as participant A or participant B. The conversaion rate is as follows: 1 experimental currency unit is worth 0.05 Euro.

For your convenience, these instructions will remain available to you on all subsequent screens of this study.

Screen 4:



Screen 5:

Introduction (Experimental Task 2)

You have been randomly and anonymously paired with two other participants in the room. You are an investor having to make an investment decision, the two participants you are paired with can be considered investment advisors offering you investment opportunities. They will be called investment advisor X and investment advisor Y. The only information available to you is the amount the two investment advisors **would have returned** to you in the first task.

You will have to make several investment decisions in this second experimental task. Investment decisions are **independent**. In other words, every investment decision will start with a **new random and anonymous** pairing with two other participants (investment advisors) in the room.

Next

Screen 6: Repeated

Choose investment

You have been given an **endowment of 100 ECU**. There are two investments you can choose from: A risk-free investment which does have a sure return of 0% (i.e., you do not lose or gain any ECUs), and a risky investment which has an expected return of 6.0% with a volatility of 20.0% (similar to the German stock market index DAX). The amount you do not invest into the risky investment will automatically be invested into the riskless investment.

However, you can only invest your endowment in the risky investment via the two investment advisors. Both charge you for their service as shown below. This charge is automatically deducted from your return on the risky investment. The investment advisor **does not keep** this charge. The investment advisors' compensation is fixed and does therefore **not depend** on your investment decision.

Given the information about the amounts both investment advisors were willing to return in the first experimental task and the costs they charge, please make an investment decision for each advisor **separately**:

Advisor	Cost (on risky investment)	Returned amount for amount you sent (You sent: 0 ECU)	Expected Return after costs	Variance
X	0.75%	0 ECU	(6.0 - 0.75)%	20.0%
Y	1.5%	0 ECU	(6.0 - 1.5)%	20.0%

Please indicate how much you would like to invest into the risky investment with advisor X:

ECU

Please indicate how much you would like to invest into the risky investment with advisor Y:

ECU

Payment: One of your investment decisions with investment advisor X or Y will be drawn randomly. Your payment then depends on the return of this investment decision. The conversion rate is as follows: 1 ECU is worth 0.05 Euro.

Next

Screen 7: Repeated

Results

You chose to invest (100 ECU - 0 ECU) into the riskless investment and 0 ECU into the risky investment with investment advisor X. The return from the risky investment turned out to be 8.66%.

You chose to invest (100 ECU - 0 ECU) into the riskless investment and 0 ECU into the risky investment with investment advisor Y. The return from the risky investment turned out to be -9.68%.

For this round, your investment decision with investment advisor Y was selected randomly. Your total total payoff from this round is thus **100.0** = (100 ECU - 0 ECU)+ (100% -9.68%)*0 ECU) .

Please remember: Investment decisions are independent and you will be matched with new investment advisors in the next round. Therefore your result in this round will not affect your result in the next round.

Next

Screen 8: Repeated

Choose investment

You have been given an **endowment of 100 ECU**. As in previous tasks, you have the choice between the risk-free investment which does have a sure return of 0% (i.e., you do not lose or gain any ECUs), and a risky investment which has an expected return of 6.0% with a volatility of 20.0% (similar to the German stock market index DAX). The amount you do not invest into the risky asset will automatically be invested into the riskless asset. Again you need to invest via investment advisors:

Cost (on risky investment)	Returned amount for amount you sent (You sent: 0 ECU)
0.75%	0 ECU

Please indicate how much do you want to invest into the risky investment with this investment advisor:

ECU

Now suppose you had to invest with the other investment advisor, who returned **0 ECU** to you.

Please indicate at which costs (in %) you would be willing to make the same investment allocation as with the other investment:

0

Next

Screen 9:

Feedback on the Experiment

Did you perceive advisors who returned more to you in the trust game (Experimental Task 1) as more trustworthy?

- Yes
- No
- I do not know

Did you expect investment decisions with advisors who returned more in the trust game (Experimental Task 1) to give you higher returns than investment decisions with advisors who returned less in the trust game?

- Yes
- No
- I do not know

Did you invest more risky with the advisor who returned more in the trust game (Experimental Task 1) because you wanted to reward him?

- Yes
- No
- I do not know

Next

Screen 10:

Sociodemographics

Please indicate your gender:

- Male
- Female

Please enter your age in years:

Please indicate your educational background (select highest):

- Primary School
- High School
- College degree (e.g., Bachelor, Master)
- Ph.D.

Please indicate your current profession:

- Student
- Employee
- Self-Employed
- Unemployed
- Retired
- I do not want to answer

If applicable, please indicate in which field of study you have studied or studying:

- Business
- Economics
- Mathematics, Physics, Engineering
- Other
- Not applicable

Have you ever invested into active funds:

- Yes
- No
- I do not want to answer

Have you ever invested into passive funds:

- Yes
- No
- I do not want to answer

In general, I am willing to take financial risks (1=Not at all willing, 5=Very willing):

1 2 3 4 5

In general, I trust others (1=Not at all, 5=Very much):

1 2 3 4 5

Next

Screen 11:

Quiz (Part 1/2)

A bat and a ball costs $1.10 in total. The bat costs a dollar more than the ball. How much does the ball cost?

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

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

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

Next

Screen 12:

Quiz (Part 2/2)

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

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

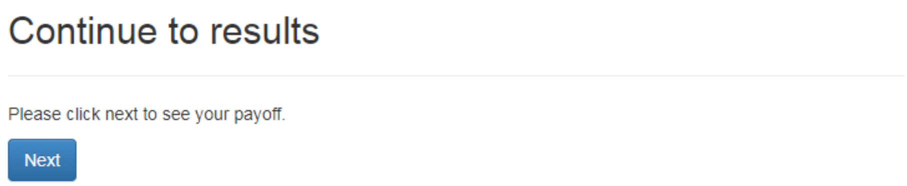
ECU

Simon decided to invest $8,000 in the investment market one day early in 2008. Six months after he invested, on July 17, the investments he had purchased were down 50%. Fortunately for Simon, from July 17 to October 17, the investments he had purchased went up 75%.

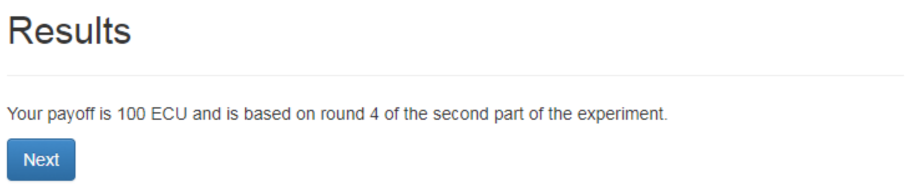
- broken even in the investment market
- is ahead of where he began
- has lost money

Next

Screen 13:


Continue to results
Please click next to see your payoff.
Next

Screen 14:


Results
Your payoff is 100 ECU and is based on round 4 of the second part of the experiment.
Next

Appendix C

Algorithm Aversion

Experimental Instructions

This appendix contains the experimental instructions as they were used in the experiment. Instructions were in English and were handed out on paper.

Written Instructions MLab. Page 1:

Instructions

Overview

This laboratory experiment consists of two parts. The first part will determine which participants will be selected for the role of a "**human fund manager**". There is one fund manager for each randomly assigned group of ten participants. In the second part, the human fund manager will have to make active choices between investing into either a stock or a bond. All participants **not** selected as human fund manager are investors and have to make an investment choice between investing either with the human fund manager or an **investment algorithm**. The goal of the investment algorithm is to maximize expected terminal wealth. The human fund manager will obtain his/her terminal wealth according to his/her investment decisions. Investors will obtain the terminal wealth of their selected investment intermediary (human or algorithm) minus a fee. As participant, please make your decisions carefully as these decisions determine your payoff for participation.

Part 1

You will be asked to answer 8 questions on financial matters. Time is limited to 1 minute per financial question. You will then be asked to answer 4 numeracy questions. Time is limited to 2 minutes per numeracy question. The participant with the highest number of correctly answered questions will be **anonymously** appointed as "**human fund manager**", his/her identity will not be revealed to the other participants. In case there are ties for the highest number of correctly answered questions, a random number draw will resolve the tie.

As an incentive to assume the role of the fund manager, this participant can collect his/her investment outcomes without deduction of any fees.

Part 2

General structure:

There are two securities on a market, one of which is a **bond** paying **3€** for **certain**. The other is a **stock** paying either **5€** or **1€**. The probability for the high payoff is either **70% (good state)** or **30% (bad state)**. Whether the good or bad state applies is **randomly** determined (50%/50%) at the beginning of a block of trials. A trial represents one draw of returns for the stock and the bond. Each **block** contains 6 trials for which the **state** of the stock **is fixed**. There is a total of 10 blocks.

As Human Fund Manager:

In each trial, the participant appointed as human fund manager is asked to choose to invest into either the stock or the bond. A history of the returns of the stock and the bond, and a history of the investment choices and returns of the investment algorithm is shown in each block.

Written Instructions MLab. Page 2:

As Investor:

At the start of **each block** you have to choose whether to **invest** with the **human fund manager** or the **investment algorithm**. Investing with the human fund manager always costs a fee of 2€ per block. Investing with the investment algorithm costs a fee of either 0€, 1€, 2€, 3€ or 4€. The respective fee is subtracted from the final investment outcome of each block (but not given to the fund manager or algorithm). For each of the 5 possible cost-combinations you will be asked to choose with which intermediary you would like to invest in this block (see Figure 1 below).

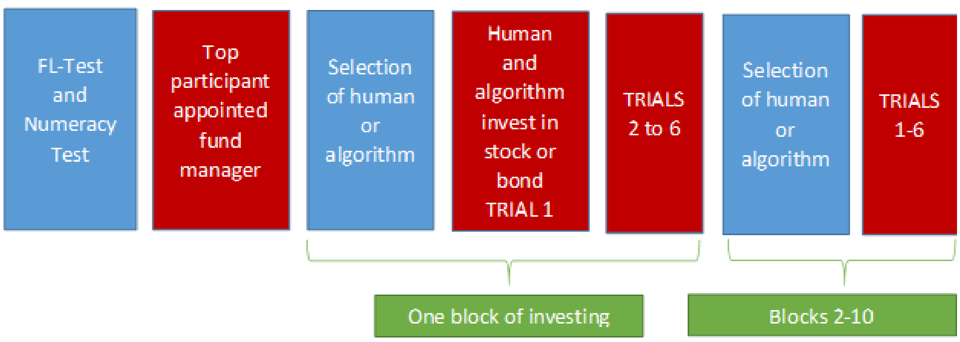
Figure 1: Example of Cost-Combination Choice

Cost for investing with the human fund manager: **2€**. Cost for investing with the investment algorithm: **0€**. With whom do you want to invest?	◯ Human fund manager ◯ Investment Algorithm
Cost for investing with the human fund manager: **2€**. Cost for investing with the investment algorithm: **1€**. With whom do you want to invest?	◯ Human fund manager ◯ Investment Algorithm
Cost for investing with the human fund manager: **2€**. Cost for investing with the investment algorithm: **2€**. With whom do you want to invest?	◯ Human fund manager ◯ Investment Algorithm
Cost for investing with the human fund manager: **2€**. Cost for investing with the investment algorithm: **3€**. With whom do you want to invest?	◯ Human fund manager ◯ Investment Algorithm
Cost for investing with the human fund manager: **2€**. Cost for investing with the investment algorithm: **4€**. With whom do you want to invest?	◯ Human fund manager ◯ Investment Algorithm

A random draw then determines which cost-combination applies, and your indicated choice for this combination will be implemented. You then observe the choices of the human fund manager and the algorithm for the six trials within the block. You cannot change your chosen intermediary within a block. For the next block, however, you make a new decision on with whom to invest.

Figure 2 shows the structure of the experiment.

Figure 2: Sequence of Experiment



Written Instructions MLab. Page 3:

Test block:

There will be a test block to familiarize participants with the screens and steps of the experiment. In this test block all choices are randomly determined: They do not stem from the human fund manager or the investment algorithm. The test block is purely for illustration purposes. The test block also does not count towards your payoff for participation.

Payoffs:

As Human Fund Manager:

You will receive the outcome for one block of your investment choices. At the end of the experiment, 1 out of the 10 blocks will be chosen randomly. The accumulated terminal wealth for this block will be paid out to you.

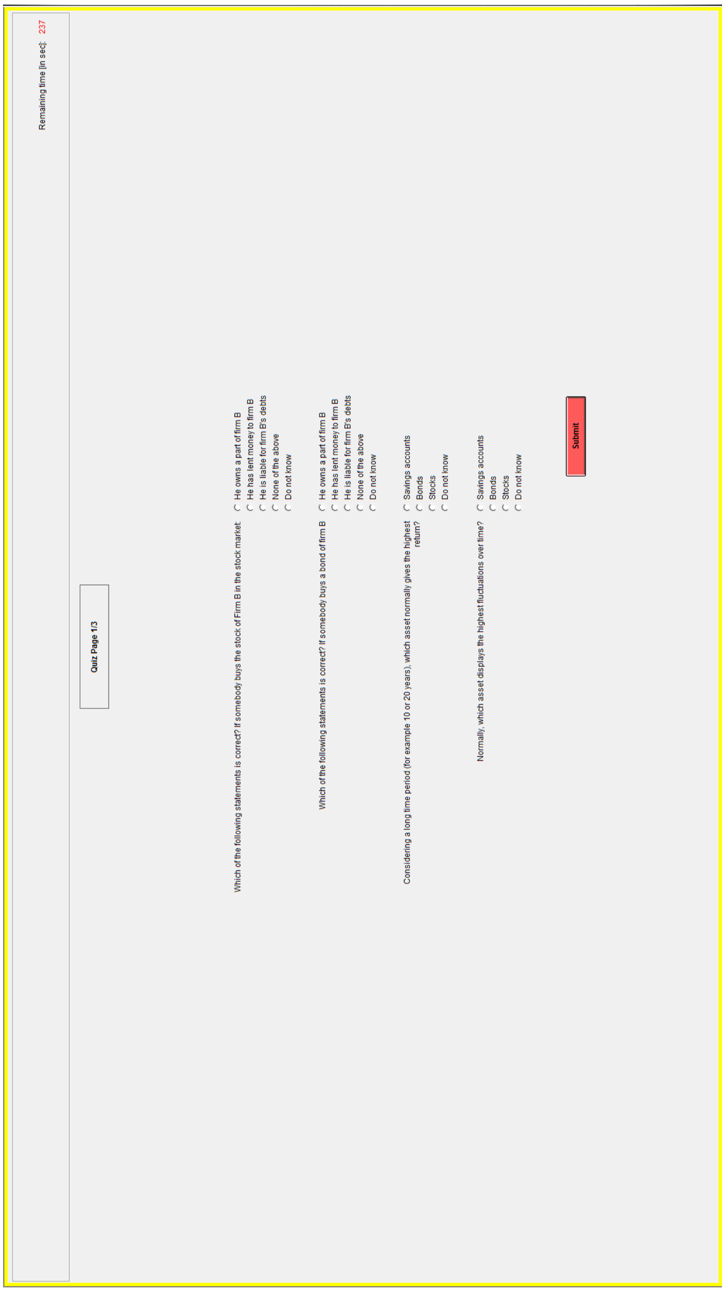
As Investor:

You will receive the outcome for one block of your investment choices, minus costs. At the end of the experiment, 1 out of the 10 blocks will be chosen randomly. The accumulated terminal wealth of the chosen investment intermediary for this block minus the respective costs will be paid out to you.

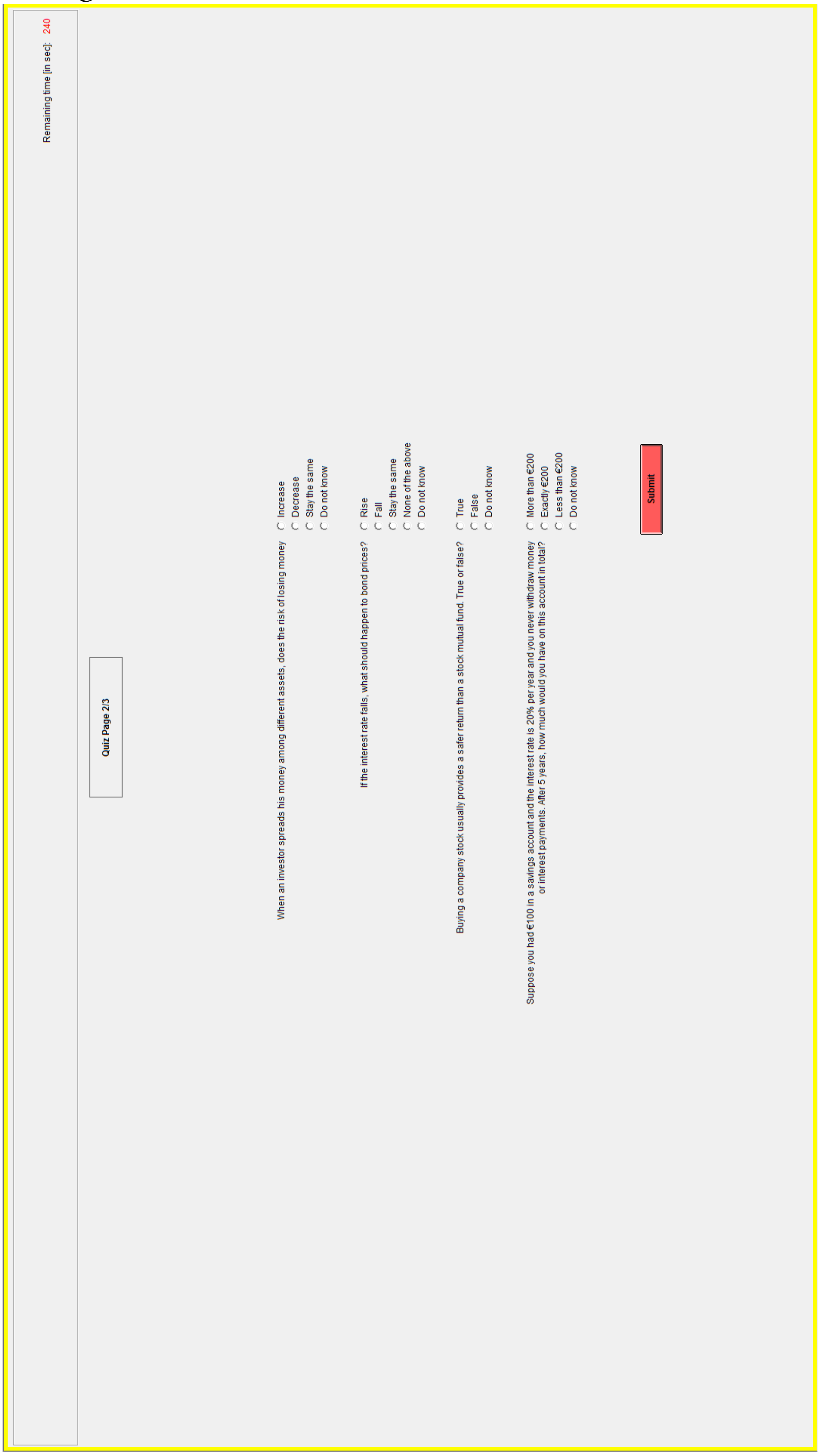
Experimental Screens

This appendix shows screenshots from the experiment as seen by participants. All realized values shown in the experimental screens are for illustration purposes only.

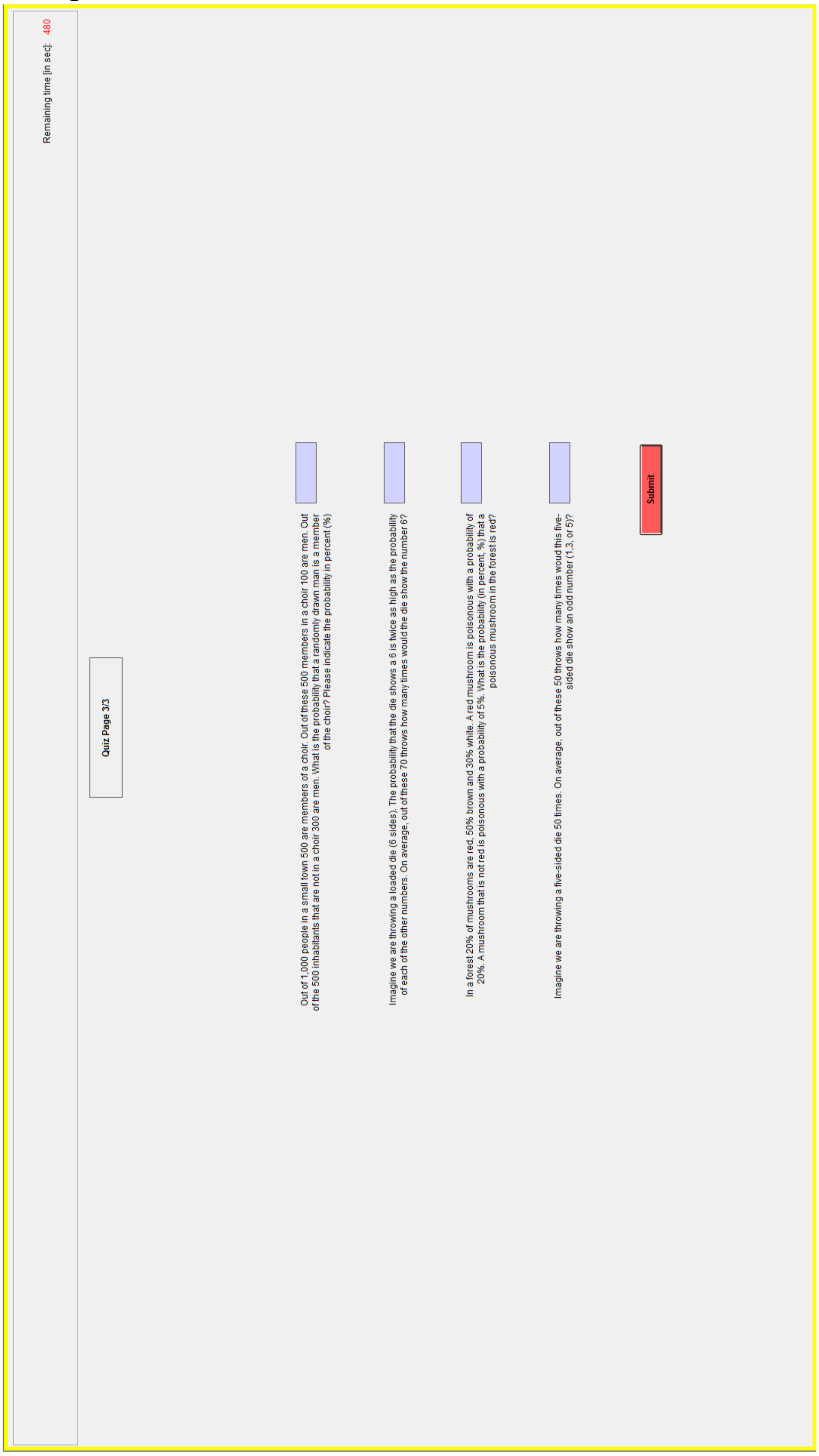
Experimental Screens. Financial Literacy and Numeracy Test, Page 1/3:



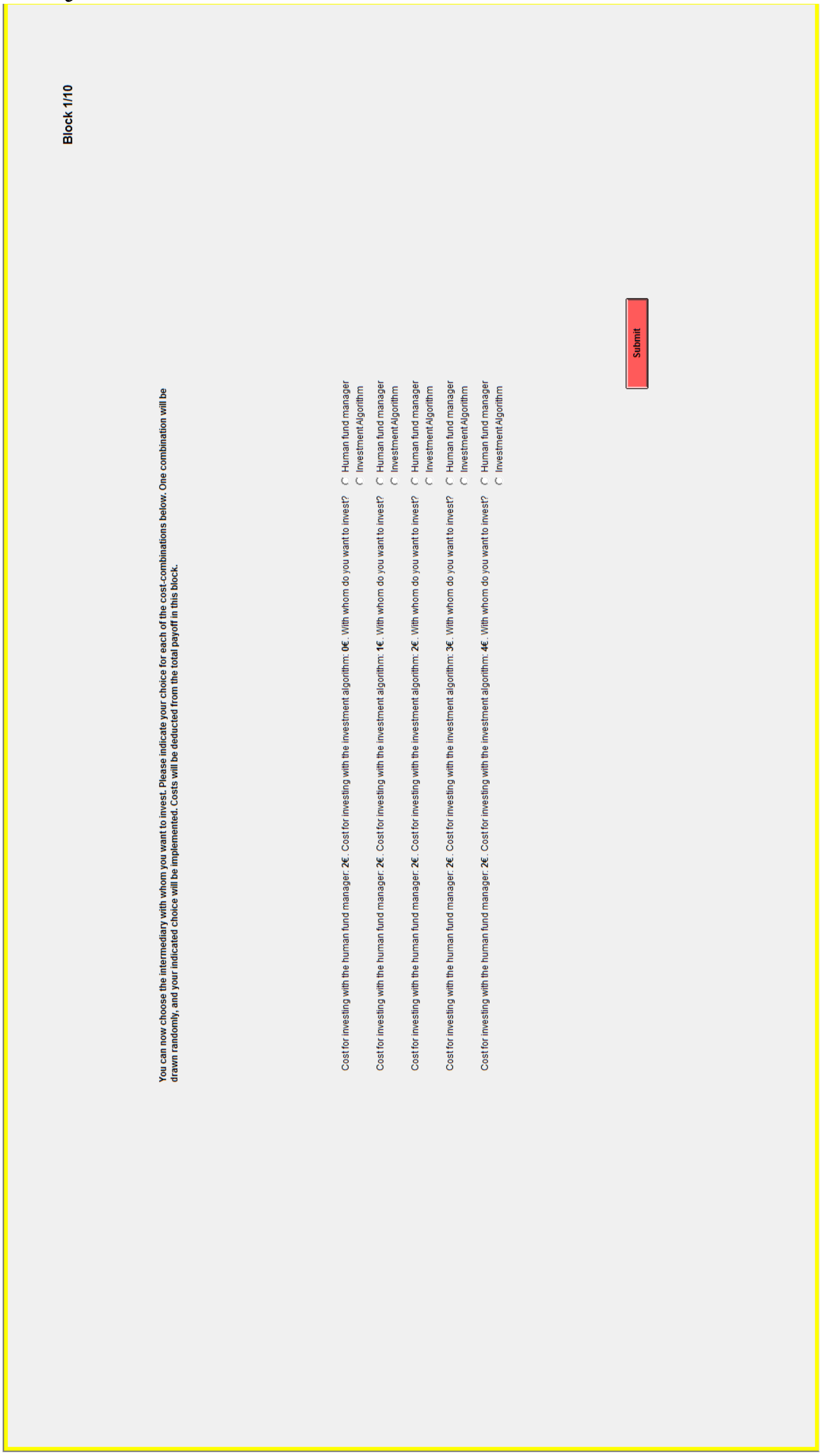
Experimental Screens. Financial Literacy and Numeracy Test, Page 2/3:



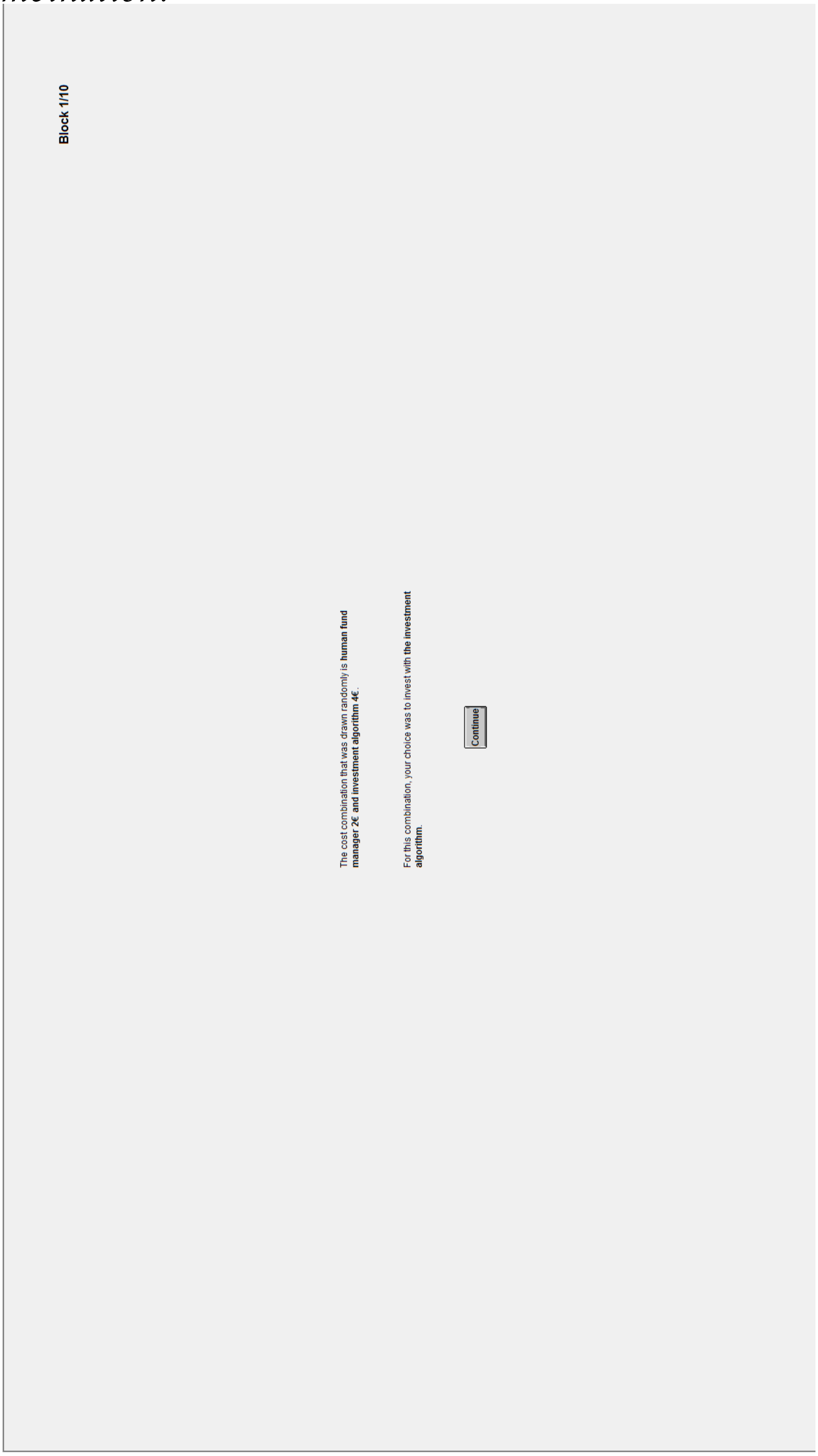
Experimental Screens. Financial Literacy and Numeracy Test, Page 3/3:



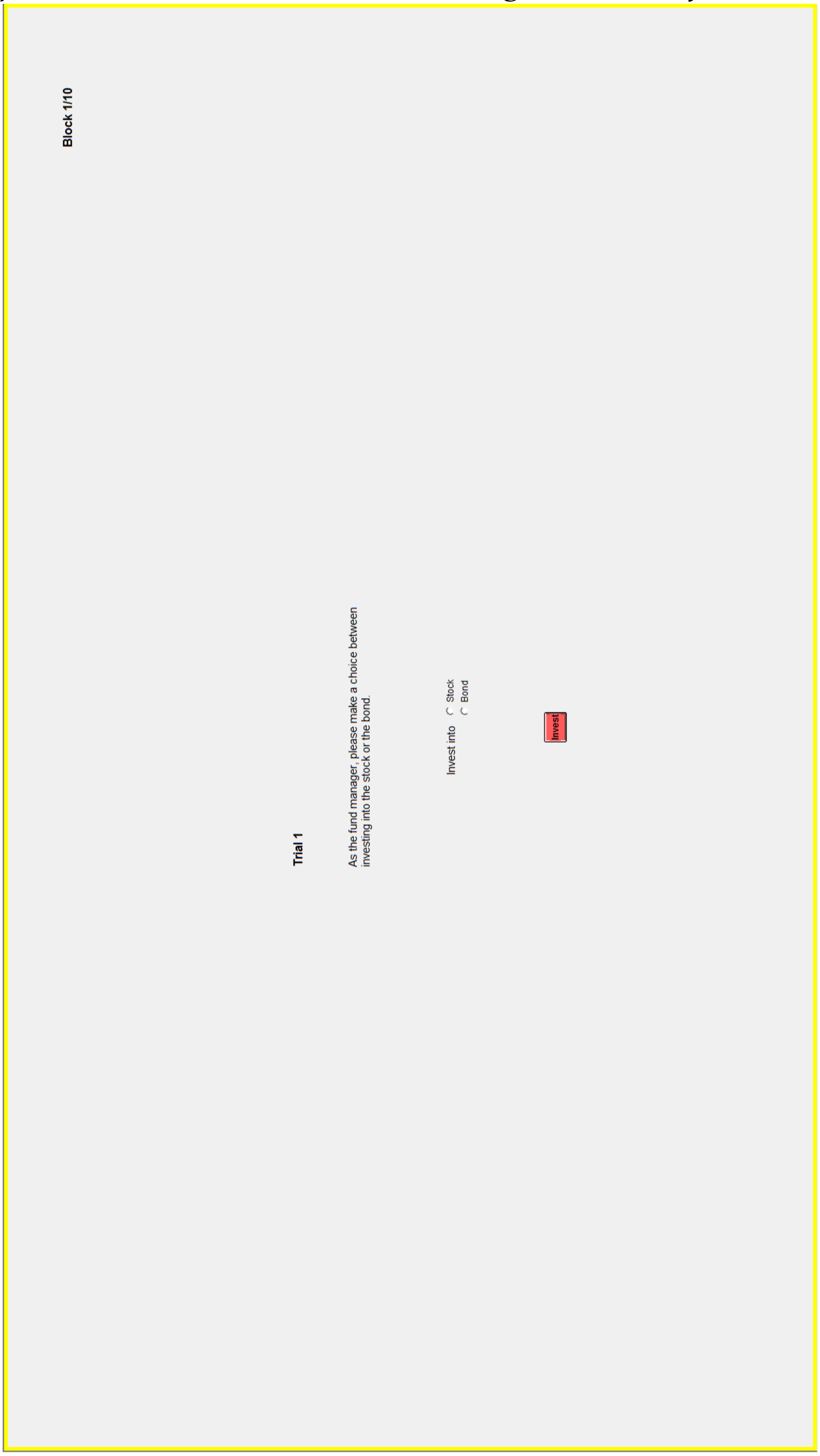
Experimental Screens. Investor Choice of Financial Intermediary:



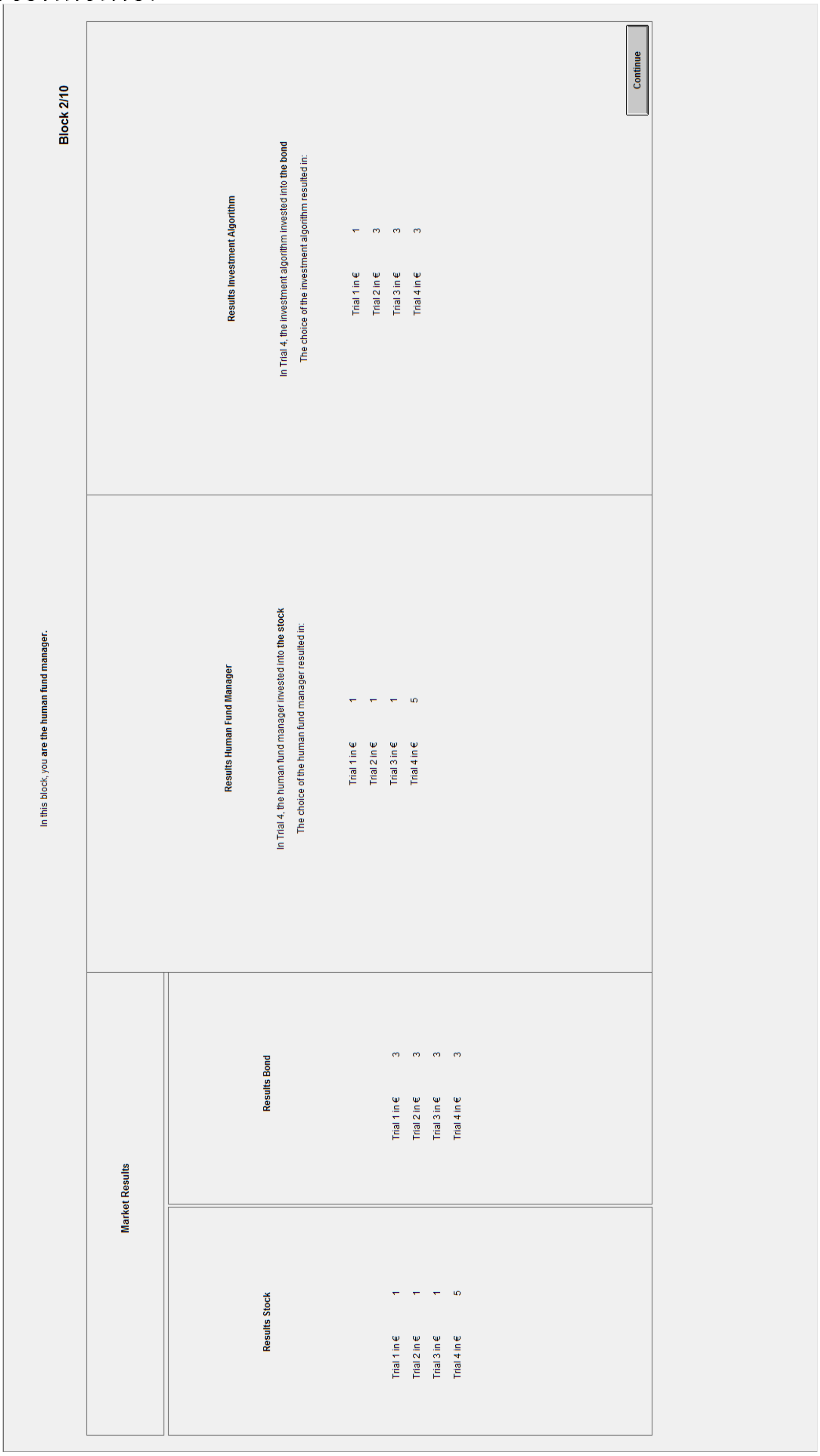
Experimental Screens. Outcome of Random Draw of Fee Combination:



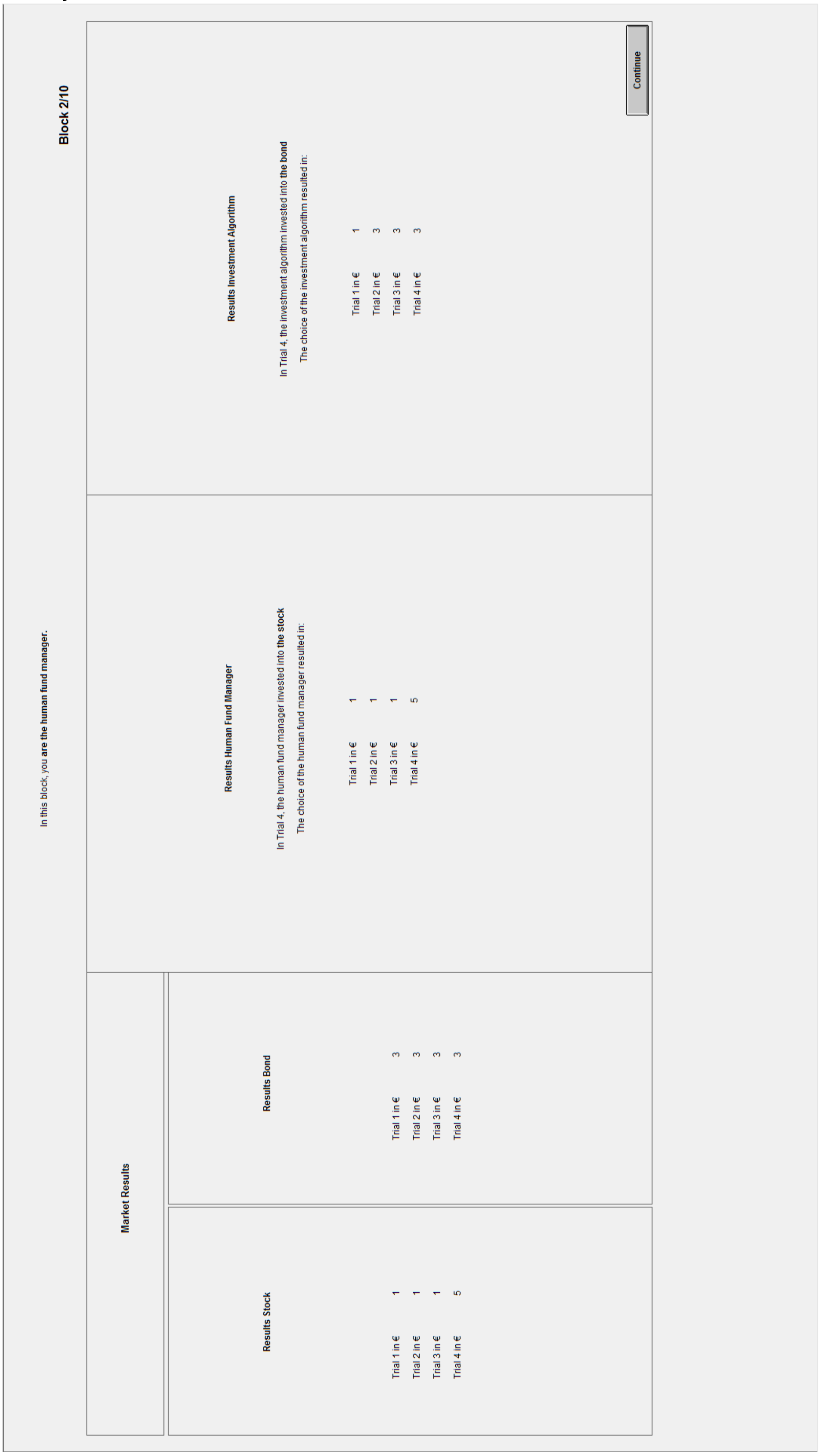
Experimental Screens. Fund Manager Choice of Asset:



Experimental Screens. Outcome Screen Within Block of Investments:



Experimental Screens. Outcome Screen After Finished Block of Investments:



Additional Results

This appendix contains additional results from the experiment.

Figure C.1: Human Equal Bayesian in Block, Over Blocks

This figure shows the average number of Bayesian investment choices by the human in a block.

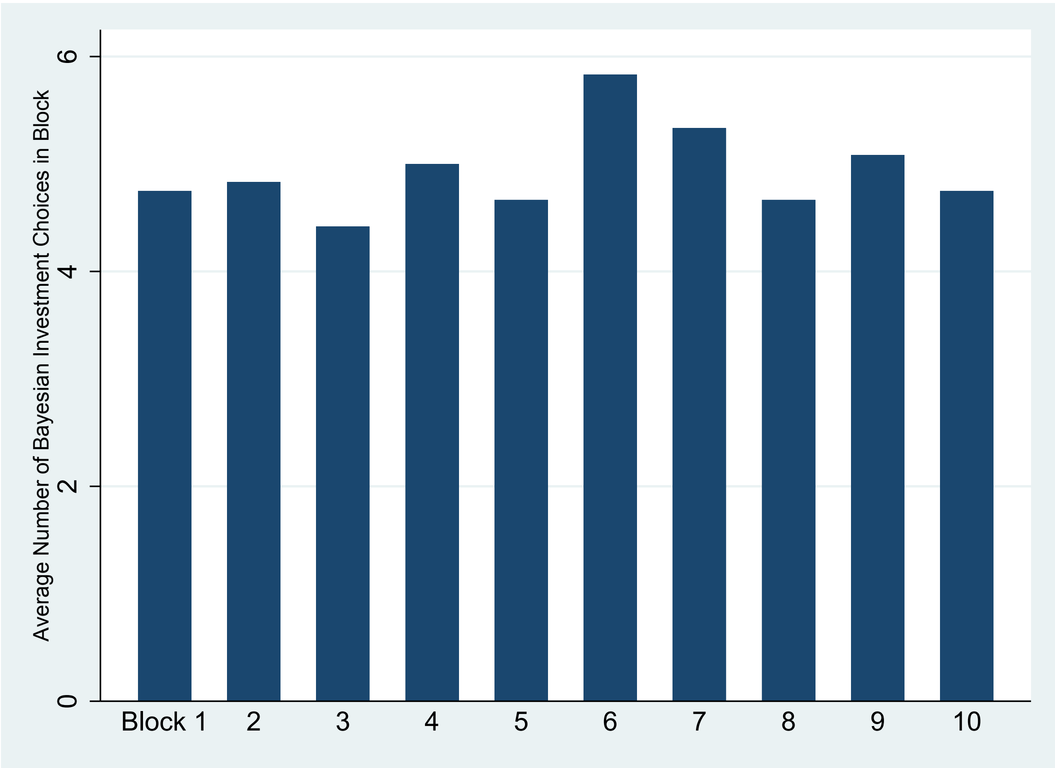


Figure C.2: Performance Human and Algorithm, by Block

This figure shows the average cumulated performance (in €) of the algorithm and the human in a block.

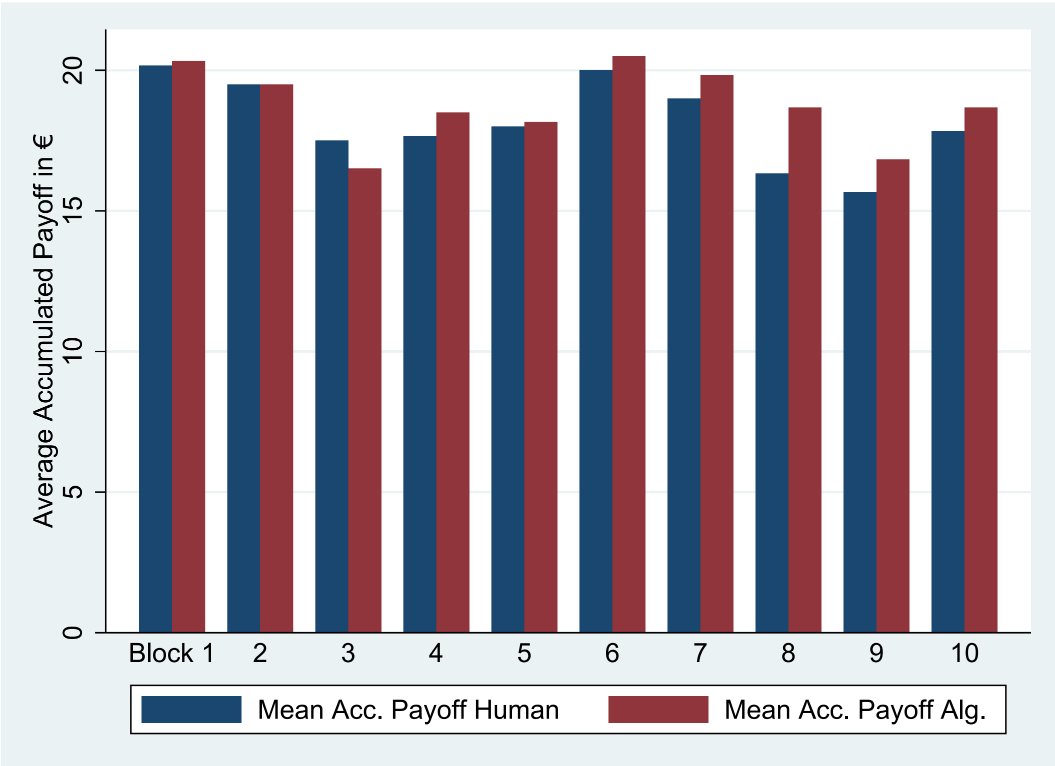


Figure C.3: Choices by Block and Costs

This figure shows the percentage of investors choosing to invest with the human in a block if fees for both intermediaries are not equal. "H: 2€, A: 0€" refers to fees of 2€ for the human and fees of 0€ for the algorithm.

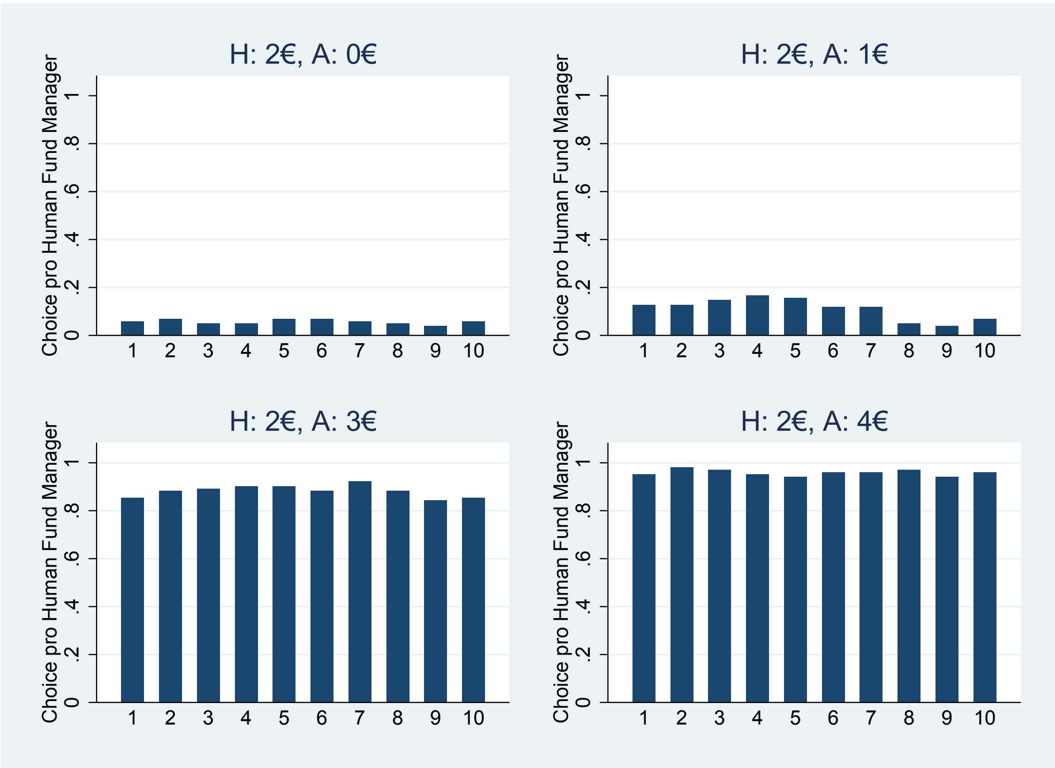


Table C.1: Correlation Matrix: Perception of Algorithms in Finance

This table shows correlations of how participants perceive algorithms in finance. To avoid acquiescence bias, for each dimension there were two questions whose order was randomized. The exact wording of these questions is shown in Table 4.1. Answers could be given on a likert scale ranging from 1 to 5, where 1 was labeled *"Strongly disagree"* and 5 was labeled *"Strongly agree"*. Values shown here are combined values for both randomized types of questions. *P*-values for significance of correlations are shown in parentheses.

	Returns	Learning	Qualitative Data	Data Aggregation	Data Weighting	Dealing With Outliers	Competitor
Returns	1.00						
Learning	0.26	1.00					
	(0.01)						
Qualitative Data	0.07	0.02	1.00				
	(0.48)	(0.84)					
Data Aggregation	0.14	-0.02	-0.06	1.00			
	(0.14)	(0.81)	(0.51)				
Data Weighting	0.13	0.06	-0.04	0.19	1.00		
	(0.2)	(0.53)	(0.66)	(0.06)			
Dealing With Outliers	0.00	-0.04	-0.10	0.13	0.06	1.00	
	(0.98)	(0.65)	(0.29)	(0.18)	(0.56)		
Competitor	0.08	-0.04	-0.10	-0.22	-0.11	-0.04	1.00
	(0.43)	(0.70)	(0.32)	(0.02)	(0.27)	(0.65)	

Table C.2: Choice Intermediary – Lags of Performance Difference

This table reports panel regressions with *Choice Human$_t$* as dependent variable. It is a dummy equal to 1 if investors choose to invest with the human fund manager if costs are equal at 2€ per intermediary, and 0 otherwise. *Performance Difference$_{t-x}$* measures the performance difference of the human fund manager minus the investment algorithm, accumulated over all trials of block t-x.
Cluster-robust standard errors are shown in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

	(1)	(2)	(3)	(4)	(5)
			Fixed Effects		
Performance Difference$_{t-1}$	0.023***	0.025***	0.026***	0.028***	0.025***
	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)
Performance Difference$_{t-2}$		0.014**	0.018***	0.017***	0.019***
		(0.006)	(0.004)	(0.004)	(0.003)
Performance Difference$_{t-3}$			0.020***	0.024***	0.022***
			(0.003)	(0.004)	(0.005)
Performance Difference$_{t-4}$				0.007	0.008
				(0.006)	(0.008)
Performance Difference$_{t-5}$					0.007
					(0.007)
Constant	0.522***	0.521***	0.532***	0.562***	0.508***
	(0.026)	(0.061)	(0.043)	(0.035)	(0.034)
Observations	918	816	714	612	510
Investor FE	YES	YES	YES	YES	YES
Block FE	YES	YES	YES	YES	YES
$R^2_{adjusted}$	0.063	0.082	0.114	0.116	0.088

Table C.3: Choice Intermediary – Total Performance, Split by Gain and Loss

This table reports panel regressions with *Choice Human$_t$* as dependent variable. It is a dummy equal to 1 if investors choose to invest with the human fund manager if costs are equal at 2€ per intermediary, and 0 otherwise. *Performance Difference$_{t=0\ to\ t-1}$* measures the performance difference of the human fund manager minus the investment algorithm, accumulated over all blocks to t-1. *Positive Performance Difference$_{t=0\ to\ t-1}$* is a dummy that takes the value of 1 if the total aggregated performance of the human fund manager up to block t-1 is greater than the total aggregated performance of the investment algorithm, and 0 otherwise. Observations for which total aggregated performance of both intermediaries up to block t-1 is equal are dropped.
Cluster-robust standard errors are shown in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

	Fixed Effects
Performance Difference$_{t=0\ to\ t-1}$	0.020**
	(0.009)
Positive Performance Difference$_{t=0\ to\ t-1}$	0.104
	(0.076)
Performance Difference$_{t=0\ to\ t-1}$ × Positive Performance Difference$_{t=0\ to\ t-1}$	0.004
	(0.019)
Constant	0.477***
	(0.059)
Observations	734
Investor FE	YES
Block FE	YES
$R^2_{adjusted}$	0.085

Table C.4: Choice Intermediary – Identical Investments in Previous Block

This table reports panel regressions with *Choice Human$_t$* as dependent variable. It is a dummy equal to 1 if investors choose to invest with the human fund manager if costs are equal at 2€ per intermediary, and 0 otherwise. *Human Equal Algorithm$_{t-1}$* is a dummy equal to 1 if human fund manager and investment algorithm had all identical investment outcomes in block t-1.
Cluster-robust standard errors are shown in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

	Fixed Effects
Human Equal Algorithm$_{t-1}$	-0.004
	(0.050)
Constant	0.521***
	(0.042)
Observations	918
Investor FE	YES
Block FE	YES
$R^2_{adjusted}$	0.026

References

Agrawal, N. and Maheswaran, D. (2005). Motivated Reasoning in Outcome-Bias Effects. *Journal of Consumer Research*, 31(4):798–805.

Aksoy, B., Harwell, H., Kovaliukaite, A., and Eckel, C. (2018). Measuring trust: A reinvestigation. *Southern Economic Journal*, 84(4):992–1000.

Amir, O., Rand, D. G., and Kobi Gal, Y. (2012). Economic Games on the Internet: The Effect of 1$ Stakes. *PLoS ONE*, 7(2):1–4.

Arkes, H. R., Shaffer, V. A., and Medow, M. A. (2007). Patients Derogate Physicians Who Use a Computer-Assisted Diagnostic Aid. *Medical Decision Making*, 27(2):189–202.

Arrow, K. J. (1972). Gifts and Exchanges. *Philosophy & Public Affairs*, 1(4):343–362.

Ashraf, N., Bohnet, I., and Piankov, N. (2003). Is Trust a Bad Investment?

Baran, N. M., Sapienza, P., and Zingales, L. (2010). Can We Infer Social Preferences From the Lab? Evidence from The Trust Game.

Barberis, N. and Shleifer, A. (2003). Style investing. *Journal of Financial Economics*, 68(2):161–199.

Baron, J. and Hershey, J. C. (1988). Outcome bias in decision evaluation. *Journal of Personality and Social Psychology*, 54(4):569–579.

Becker, G., DeGroot, M. H., and Marschak, J. (1964). Measuring Utility by a Single-Response Sequential Method. *Journal of the International Society for the Systems Sciences*, 9(3):226–232.

Benartzi, S. and Thaler, R. H. (1995). Myopic loss aversion and the equity premium puzzle. *Quarterly Journal of Economics*, 110(1):73–92.

Benartzi, S. and Thaler, R. H. (1999). Risk Aversion or Myopia? Choices in Repeated Gambles and Retirement Investments. *Management Science*, 45(3):364–381.

Berg, J., Dickhaut, J., and McCabe, K. (1995). Trust, Reciprocity, and Social History. *Games and Economic Behavior*, 10(1):122–142.

Bergstresser, D., Chalmers, J. M. R., and Tufano, P. (2009). Assessing the costs and benefits of brokers in the mutual fund industry. *Review of Financial Studies*, 22(10):4129–4156.

Berk, J. B. and Green, R. C. (2004). Mutual Fund Flows and Performance in Rational Markets. *Journal of Political Economy*, 112(6):1269–1295.

Bertrand, M. and Mullainathan, S. (2001). Are CEOs Rewarded for Luck? The Ones Without Principals Are. *Quarterly Journal of Economics*, 116(3):901–932.

Beshears, J., Choi, J., Laibson, D., and Madrian, B. (2011). How Does Simplified Disclosure Affect Individuals' Mutual Fund Choices? In Wise, D. A., editor, *Explorations in the Economics of Aging*, chapter 2, pages 75–96. University of Chicago Press, Chicago.

Brandts, J. and Charness, G. (2000). Hot vs. Cold: Sequential Responses and Preference Stability in Experimental Games. *Experimental Economics*, 2(3):227–238.

Brandts, J. and Charness, G. (2011). The strategy versus the direct-response method: A first survey of experimental comparisons. *Experimental Economics*, 14(3):375–398.

Buchan, N. R., Croson, R. T. A., and Johnson, E. J. (2002). Trust and reciprocity: An international experiment.

Bucher-Koenen, T. and Lusardi, A. (2011). Financial literacy and retirement planning in Germany. *Journal of Pension Economics and Finance*, 10(4):565–584.

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(1):3–5.

Calvet, L. E., Campbell, J. Y., and Sodini, P. (2007). Down or Out: Assessing the Welfare Costs of Household Investment Mistakes. *Journal of Political Economy*, 115(5):707–747.

Camerer, C. F. (2003). *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton University Press, Princeton.

Carhart, M. M. (1997). On Persistence in Mutual Fund Performance. *Journal of Finance*, 52(1):57–82.

Carlin, B. I., Dorobantu, F., and Viswanathan, S. (2009). Public trust, the law, and financial investment. *Journal of Financial Economics*, 92(3):321–341.

Casler, K., Bickel, L., and Hackett, E. (2013). Separate but equal? A comparison of participants and data gathered via Amazon's MTurk, social media, and face-to-face behavioral testing. *Computers in Human Behavior*, 29(6):2156–2160.

Chang, T., Solomon, D. H., and Westerfield, M. M. (2016). Looking for Someone to Blame: Delegation, Cognitive Dissonance, and the Disposition Effect. *Journal of Finance*, 71(1):267–302.

Charness, G. and Levin, D. (2005). When Optimal Choices Feel Wrong: A Laboratory Study of Bayesian Updating, Complexity, and Affect. *American Economic Review*, 95(4):1300–1309.

Chen, D. L., Schonger, M., and Wickens, C. (2016). oTree-An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, 9(1):88–97.

Chen, J., Hong, H., Huang, M., and Kubik, J. D. (2004). Does Fund Size Erode Mutual Fund Performance ? The Role of Liquidity and Organization. *American Economic Review*, 94(5):1276–1302.

Cokely, E. T., Galesic, M., Schulz, E., Ghazal, S., and Garcia-Retamero, R. (2012). Measuring Risk Literacy: The Berlin Numeracy Test. *Judgment and Decision Making*, 7(1):25–47.

Cox, J. C. (2004). How to identify trust and reciprocity. *Games and Economic Behavior*, 46(2):260–281.

Cronqvist, H., Siegel, S., and Yu, F. (2015). Value versus growth investing: Why do different investors have different styles? *Journal of Financial Economics*, 117(2):333–349.

Croson, R. and Buchan, N. (1999). Gender and culture: International experimental evidence from trust games. *American Economic Review*, 89(2):386–391.

D'Acunto, F., Prabhala, N., and Rossi, A. G. (2018). The Promises and Pitfalls of Robo-advising.

Dawes, R. M. (1979). The Robust Beauty of Improper Linear Models in Decision Making. *American Psychologist*, 34(7):571–582.

Dietvorst, B. J., Simmons, J. P., and Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*, 144(1):114–126.

Dietvorst, B. J., Simmons, J. P., and Massey, C. (2018). Overcoming Algorithm Aversion: People Will Use Imperfect Algorithms If They Can (Even Slightly) Modify Them. *Management Science*, 64(3):983–1476.

Dijkstra, J. J. (1999). User agreement with incorrect expert system advice. *Behaviour & Information Technology*, 18(6):399–411.

Dijkstra, J. J., Liebrand, W. B. G., and Timminga, E. (1998). Persuasiveness of expert systems. *Behaviour & Information Technology*, 17(3):155–163.

Dorn, D. and Weber, M. (2017). Losing Trust in Money Doctors.

Dubois, D., Willinger, M., and Blayac, T. (2012). Does players' identification affect trust and reciprocity in the lab? *Journal of Economic Psychology*, 33(1):303–317.

Falk, A., Enke, B., and Huffman, D. (2018). Global Evidence on Economic Preferences. *Quarterly Journal of Economics, forthcoming*.

Fama, E. F. and French, K. R. (2010). Luck versus Skill in the cross-section of mutual fund returns. *Journal of Finance*, 65(5):1915–1947.

Fehr, E. (2009). On the Economics and Biology of Trust. *Journal of the European Economic Association*, 7(2-3):235–266.

Fernandes, D., Lynch, J. G., and Netemeyer, R. G. (2014). Financial Literacy, Financial Education, and Downstream Financial Behaviors. *Management Science*, 60(8):1861–1883.

Fischbacher, U. (2007). Z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2):171–178.

Foerster, S., Linnainmaa, J. T., Melzer, B. T., and Previtero, A. (2017). Retail Financial Advice : Does One Size Fit All ? *Journal of Finance*, 72(4):1441–1482.

Frazzini, A. and Lamont, O. A. (2008). Dumb money: Mutual fund flows and the cross-section of stock returns. *Journal of Financial Economics*, 88(2):299–322.

Gasper, J. T. and Reeves, A. (2011). Make It Rain? Retrospection and the Attentive Electorate in the Context of Natural Disasters. *American Journal of Political Science*, 55(2):340–355.

Gaudecker, H.-M. V. (2015). How Does Household Portfolio Diversification Vary with Financial Literacy and Financial Advice? *Journal of Finance*, 70(2):489–507.

Gennaioli, N., Shleifer, A., and Vishny, R. (2015). Money Doctors. *Journal of Finance*, 70(1):91–114.

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.

Greiner, B. (2015). Subject pool recruitment procedures: organizing experiments with ORSEE. *Journal of the Economic Science Association*, 1(1):114–125.

Grether, D. M. (1992). Testing bayes rule and the representativeness heuristic: Some experimental evidence. *Journal of Economic Behavior and Organization*, 17(1):31–57.

Grove, W. M. and Meehl, P. E. (1996). Comparative Efficiency of Informal (Subjective, Impressionistic) and Formal (Mechanical, Algorithmic) Prediction Procedures: The Clinical-Statistical Controversy. *Psychology, Public Policy, and Law*, 2(2):293–323.

Grove, W. M., Zald, D. H., Lebow, B. S., Snitz, B. E., and Nelson, C. (2000). Clinical versus mechanical prediction: A meta-analysis. *Psychological Assessment*, 12(1):19–30.

Guiso, L., Sapienza, P., and Zingales, L. (2004). The Role Of Social Capital In Financial Development. *American Economic Review*, 94(3):526–556.

Guiso, L., Sapienza, P., and Zingales, L. (2008). Trusting the Stock Market. *Journal of Finance*, 63(6):2557–2600.

Guiso, L. and Viviano, E. (2015). How Much Can Financial Literacy Help?

Hackethal, A., Haliassos, M., and Jappelli, T. (2012). Financial advisors: A case of babysitters? *Journal of Banking and Finance*, 36(2):509–524.

Harvey, C. R., Rattray, S., Sinclair, A., and Hemert, O. V. (2017). Man vs. Machine: Comparing Discretionary and Systematic Hedge Fund Performance. *Journal of Portfolio Management*, 43(4):55–69.

Heuer, J., Merkle, C., and Weber, M. (2017). Fooled by Randomness: Investor Perception of Fund Manager Skill. *Review of Finance*, 21(2):605–635.

Highhouse, S. (2008). Stubborn Reliance on Intuition and Subjectivity in Employee Selection. *Industrial and Organizational Psychology*, 1(3):333–342.

Hodge, F. D., Mendoza, K., and Sinha, R. K. (2018). The Effect of Humanizing Robo-Advisors on Investor Judgments.

Hoechle, D., Ruenzi, S., Schaub, N., and Schmid, M. (2017). The Impact of Financial Advice on Trade Performance and Behavioral Biases. *Review of Finance*, 21(2):871–910.

Hoechle, D., Ruenzi, S., Schaub, N., and Schmid, M. (2018). Financial Advice and Bank Profits. *Review of Financial Studies, forthcoming*.

Inderst, R. and Ottaviani, M. (2009). Misseling through Agents. *American Economic Review*, 99(3):883–908.

Inderst, R. and Ottaviani, M. (2012). How (not) to pay for advice: A framework for consumer financial protection. *Journal of Financial Economics*, 105(2):393–411.

Jensen, M. (1968). Problems in Selection of Security Portfolios: The Performance of Mutual Funds in the Period 1945-1964. *Journal of Finance*, 23(2):389–416.

Jenter, D. and Kanaan, F. (2015). CEO Turnover and Relative Performance Evaluation. *Journal of Finance*, 70(5):2155–2184.

Johnson, N. D. and Mislin, A. A. (2011). Trust games: A meta-analysis. *Journal of Economic Psychology*, 32(5):865–889.

Kaplan, S. N. and Schoar, A. (2005). Private Equity Performance : Returns, Persistence, and Capital Flows. *Journal of Finance*, 60(4):1791–1824.

Kaufmann, C., Weber, M., and Haisley, E. (2013). The Role of Experience Sampling and Graphical Displays on One's Investment Risk Appetite. *Management Science*, 59(2):323–340.

Keser, C. (2002). Trust and reputation building in E-Commerce.

Knack, S. and Keefer, P. (1997). Does Social Capital Have an Economic Payoff? *Quarterly Journal of Economics*, 112(4):1251–1288.

König-Kersting, C., Pollmann, M., Potters, J., and Trautmann, S. T. (2017). Good decision vs. good results: Outcome bias in financial agents' rewards.

Kosfeld, M., Heinrichs, M., Zak, P. J., Fischbacher, U., and Fehr, E. (2005). Oxytocin increases trust in humans. *Nature*, 435(7042):673–677.

Kostovetsky, L. (2016). Whom Do You Trust?: Investor-Advisor Relationships and Mutual Fund Flows. *Review of Financial Studies*, 29(4):898–936.

Kuhnen, C. M. (2015). Asymmetric Learning from Financial Information. *Journal of Finance*, 70(5):2029–2062.

Kumar, A. (2009). Dynamic style preferences of individual investors and stock returns. *Journal of Financial and Quantitative Analysis*, 44(3):607–640.

La Porta, R., Lopez-de Silanes, F., Shleifer, A., and Vishny, R. W. (1997). Trust in Large Organizations. *American Economic Review*, 87(2):333–338.

Lefgren, L., Platt, B., and Price, J. (2014). Sticking with What (Barely) Worked: A Test of Outcome Bias. *Management Science*, 61(5):1121–1136.

Linnainmaa, J., Melzer, B., Previtero, A., and Foerster, S. (2018). Financial Advisors and Risk-Taking.

Lipshitz, R. (1989). "Either a medal or a corporal": The effects of success and failure on the evaluation of decision making and decision makers. *Organizational Behavior and Human Decision Processes*, 44(3):380–395.

Logg, J. M. (2017). Theory of Machine: When Do People Rely on Algorithms?

Lusardi, A. and Mitchell, O. S. (2014). The Economic Importance of Financial Literacy. *Journal of Economic Literature*, 52(1):5–44.

Martin, J. W. and Cushman, F. (2016). Why we forgive what can't be controlled. *Cognition*, 147:133–143.

Meehl, P. E. (1954). *Clinical versus statistical prediction: A theoretical analysis and a review of the evidence*. University of Minnesota Press, Minneapolis, MN, US.

Mullainathan, S., Noeth, M., and Schoar, A. (2012). The market for financial advice: An audit study.

Mullainathan, S., Schwartzstein, J., and Shleifer, A. (2008). Coarse Thinking and Persuasion. *Quarterly Journal of Economics*, 123(2):577–619.

Önkal, D., Goodwin, P., Thomson, M., Gönül, S., and Pollock, A. (2009). The Relative Influence of Advice From Human Experts and Statistical Methods on Forecast Adjustment. *Journal of Behavioral Decision Making*, 22(4):390–409.

Ouwersloot, H., Nijkamp, P., and Rietveld, P. (1998). Errors in probability updating behaviour. *Journal of Economic Psychology*, 19(5):535–563.

Paolacci, G., Chandler, J., and Stern, L. N. (2010). Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making*, 5(5):411–419.

Pastor, L., Stambaugh, R. F., and Taylor, L. A. (2015). Scale and skill in active management. *Journal of Financial Economics*, 116(1):23–45.

Promberger, M. and Baron, J. (2006). Do patients trust computers? *Journal of Behavioral Decision Making*, 19(5):455–468.

Ratner, R. K. and Herbst, K. C. (2005). When good decisions have bad outcomes: The impact of affect on switching behavior. *Organizational Behavior and Human Decision Processes*, 96(1):23–37.

Riedl, A. and Smeets, P. (2017). Why Do Investors Hold Socially Responsible Mutual Funds? *Journal of Finance*, 72(6):2505–2550.

Sapienza, P. and Zingales, L. (2012). A Trust Crisis. *International Review of Finance*, 12(2):123–131.

Savani, K. and King, D. (2015). Perceiving outcomes as determined by external forces: The role of event construal in attenuating the outcome bias. *Organizational Behavior and Human Decision Processes*, 130:136–146.

Seta, C. E., Seta, J. J., Petrocelli, J. V., and McCormick, M. (2015). Even better than the real thing: Alternative outcome bias affects decision judgements and decision regret. *Thinking & Reasoning*, 21(4):446–472.

Sezer, O., Zhang, T., Gino, F., and Bazerman, M. H. (2016). Overcoming the outcome bias: Making intentions matter. *Organizational Behavior and Human Decision Processes*, 137:13–26.

Shaffer, V. A., Probst, C. A., Merkle, E. C., Arkes, H. R., and Medow, M. A. (2013). Why Do Patients Derogate Physicians Who Use a Computer-Based Diagnostic Support System? *Medical Decision Making*, 33(1):108–118.

Sirri, E. R. and Tufano, P. (1998). Costly Search and Mutual Fund Flows. *Journal of Finance*, 53(5):1589–1622.

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

Vlek, C. (1984). What constitutes 'a good decision'? *Acta Psychologica*, 56(1-3):5–27.

Vyrastekova, J. and Onderstal, S. (2010). The Trust Game Behind the Veil of Ignorance: A Note on Gender Differences.

Weber, E. U., Siebenmorgen, N., and Weber, M. (2005). Communicating asset risk: How name recognition and the format of historic volatility information affect risk perception and investment decisions. *Risk Analysis*, 25(3):597–609.

Weber, M., Borgsen, S., Glaser, M., Norden, L., Nosic, A., Savov, S., Schmitz, P., and Welfens, F. (2007). *Genial Einfach Investieren*. Campus Verlag, Frankfurt a. M.

Yeomans, M., Shah, A. K., Mullainathan, S., and Kleinberg, J. (2017). Making Sense of Recommendations.

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 "Delegated Investing and the Role of Investment Outcomes, Quality, and Trust" 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

Maximilian Germann

Since 2016	Research Assistant Chair of Finance and Banking Supervisor: Prof. Dr. Dr. h.c. Martin Weber University of Mannheim, Germany
Since 2014	Doctoral Student in Finance Center for Doctoral Studies in Business (CDSB) Graduate School of Economic and Social Sciences (GESS) University of Mannheim, Germany
2013 – 2014	Master in Accounting and Finance (M.Sc.) University of Edinburgh, United Kingdom
2010 – 2013	Bachelor in General Management (B.Sc.) EBS Universität für Wirtschaft und Recht, Germany
02/2012 – 06/2012	Visiting Student Seoul National University, South Korea