Algorithm Aversion in Financial Investing*

Maximilian Germann[†] **and Christoph Merkle**[‡]

August 2018

ABSTRACT

The tendency of humans to shy away from using algorithms—even when algorithms observably outperform their human counterpart—has been referred to as algorithm aversion. We conduct an experiment to test for algorithm aversion in financial decision making. Participants acting as investors can tie their incentives to either a human fund manager or an investment algorithm. We find no sign of algorithm aversion: Investors care about returns, but do not have strong preferences which intermediary obtains these returns. Contrary to what has been suggested, investors are also not quicker to lose confidence in the algorithm after seeing it err. However, we find that investors are unable to fully separate skill and luck when evaluating either intermediary.

Keywords: Algorithm aversion, financial technology, asset management, delegation.

JEL classification: G11, G23, G41, O33.

*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.

†University of Mannheim, Finance Department, L9, 1-2, 68161 Mannheim, Germany. Corresponding author: germann@bank.bwl.uni-mannheim.de.

‡Kühne Logistics University, Grosser Grassbrook 17, 20457 Hamburg, Germany, christoph.merkle@the-klu.org.

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 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,

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

[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.

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 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 2 provides an overview of the literature on algorithm aversion from which we derive hypotheses for the experiment. Section 3 presents the experimental design and participants. In section 4, we report and discuss the main results, before a final section concludes.

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. (2016) 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 perception of algorithms 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 a 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.

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.

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 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.

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

[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."

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

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

[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

amounts to $6 \cdot 3 = 18$€, 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.

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.

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.

compete for the role.

4 Results

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 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 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.

4.2 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 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 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 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 $\times$ 10 blocks), the algorithm outperforms human fund managers by on average 0.58€ per block. As depicted in Figure 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, 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 Online Appendix).

This naturally results in a high number of identical choices between the two intermediaries, which is also documented in Figure 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).

4.3 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

[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%.

(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. 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% ($p = 0.24$). 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% ($p < 0.01$). 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

[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.

[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.

supports 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 Online 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 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 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.

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 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)

[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.

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, investors react to cumulative performance of both intermediaries in the expected direction. The higher the 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 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

[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 Online Appendix).

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 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 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 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 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

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

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 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 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 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.

[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 4.5).

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

A way to corroborate this finding is to look at ex-post errors and ex-ante errors as defined in section 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.

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.

References

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

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

D'Acunto, F., Prabhala, N., Rossi, A. G., 2018. The promises and pitfalls of robo-advising, Working paper.

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

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

Dietvorst, B. J., Simmons, J. P., Massey, C., 2016. Overcoming Algorithm Aversion: People Will Use Imperfect Algorithms If They Can (Even Slightly) Modify Them. *Management Science* .

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

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

Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D., Sunde, U., 2018. Global evidence on economic preferences. *The Quarterly Journal of Economics, forthcoming* .

Fischbacher, U., 2007. z-tree - zurich toolbox for readymade economic experiments. *Experimental Economics* 10, 171–178.

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

Grove, W. M., 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, 293–323.

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

Harvey, C. R., Rattray, S., Sinclair, A., Hemert, O. V., 2017. Man vs. Machine: Comparing Discretionary and Systematic Hedge Fund Performance. *Journal of Portfolio Management* 43, 55–69.

Heuer, J., Merkle, C., Weber, M., 2017. Fooled by Randomness: Investor Perception of Fund Manager Skill. *Review of Finance* 21, 605–635.

Highhouse, S., 2008. Stubborn Reliance on Intuition and Subjectivity in Employee Selection. *Industrial and Organizational Psychology* 1, 333–342.

Hodge, F. D., Mendoza, K., Sinha, R. K., 2018. The Effect of Humanizing Robo-Advisors on Investor Judgments, Working Paper.

Kuhnen, C. M., 2015. Asymmetric Learning from Financial Information. *Journal of Finance* 70, 2029–2062.

Logg, J. M., 2017. Theory of Machine: When Do People Rely on Algorithms?, Working Paper.

Meehl, P. E., 1954. Clinical versus statistical prediction: A theoretical analysis and a review of the evidence. University of Minnesota Press.

Önkal, D., Goodwin, P., Thomson, M., Gönül, S., Pollock, A., 2009. The Relative Influence of Advice From Human Experts and Statistical Methods on Forecast Adjustment. *Journal of Behavioral Decision Making* 22, 390–409.

Promberger, M., Baron, J., 2006. Do patients trust computers? *Journal of Behavioral Decision Making* 19, 455–468.

Shaffer, V. A., Probst, C. A., Merkle, E. C., Arkes, H. R., Medow, M. A., 2013. Why Do Patients Derogate Physicians Who Use a Computer-Based Diagnostic Support System? *Medical Decision Making* 33, 108–118.

van Rooij, M., Lusardi, A., Alessie, R., 2011. Financial literacy and stock market participation. *Journal of Financial Economics* 101, 449–472.

Yeomans, M., Shah, A. K., Mullainathan, S., Kleinberg, J., 2017. Making Sense of Recommendations, Working Paper.

Figure 1: Illustration of experimental design

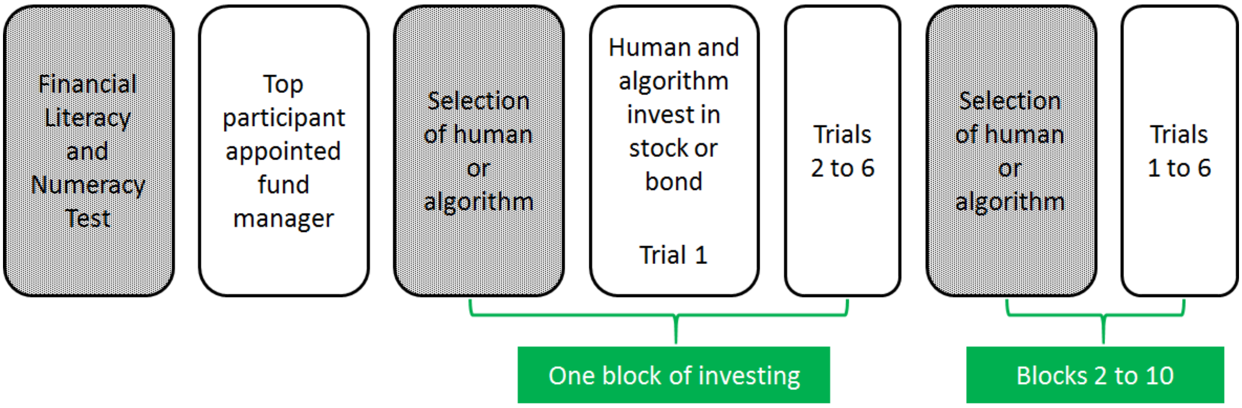


Figure 2: Correct decisions by intermediary

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.

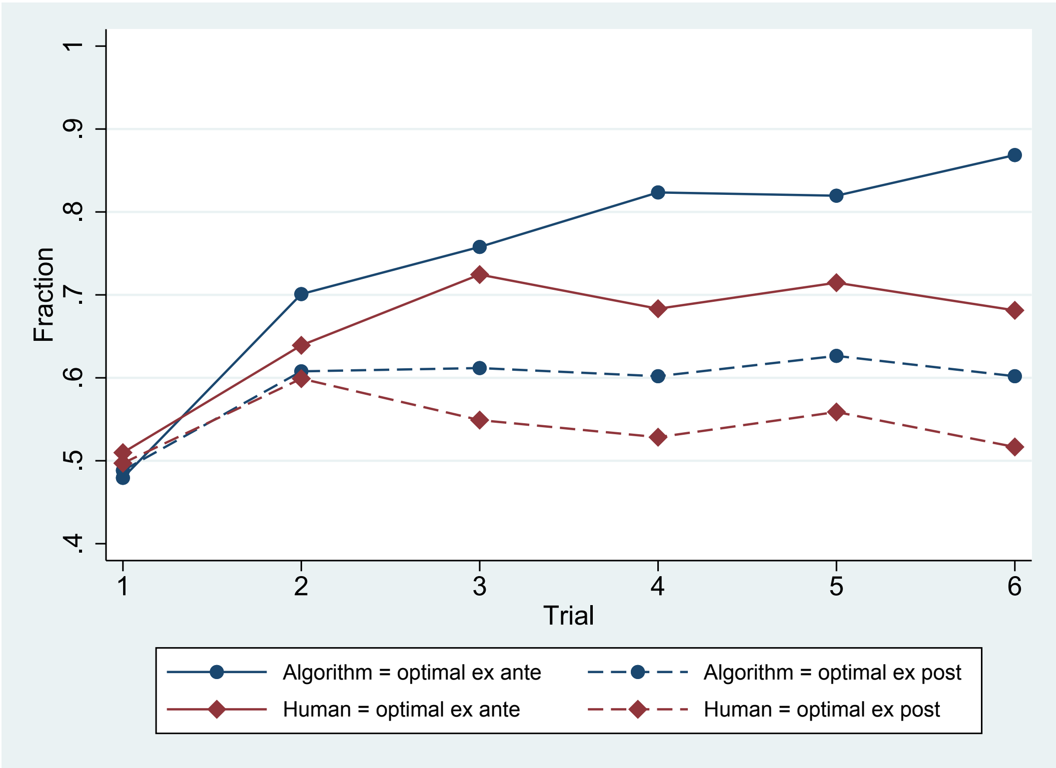


Figure 3: Relative performance investment algorithm and human fund manager

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

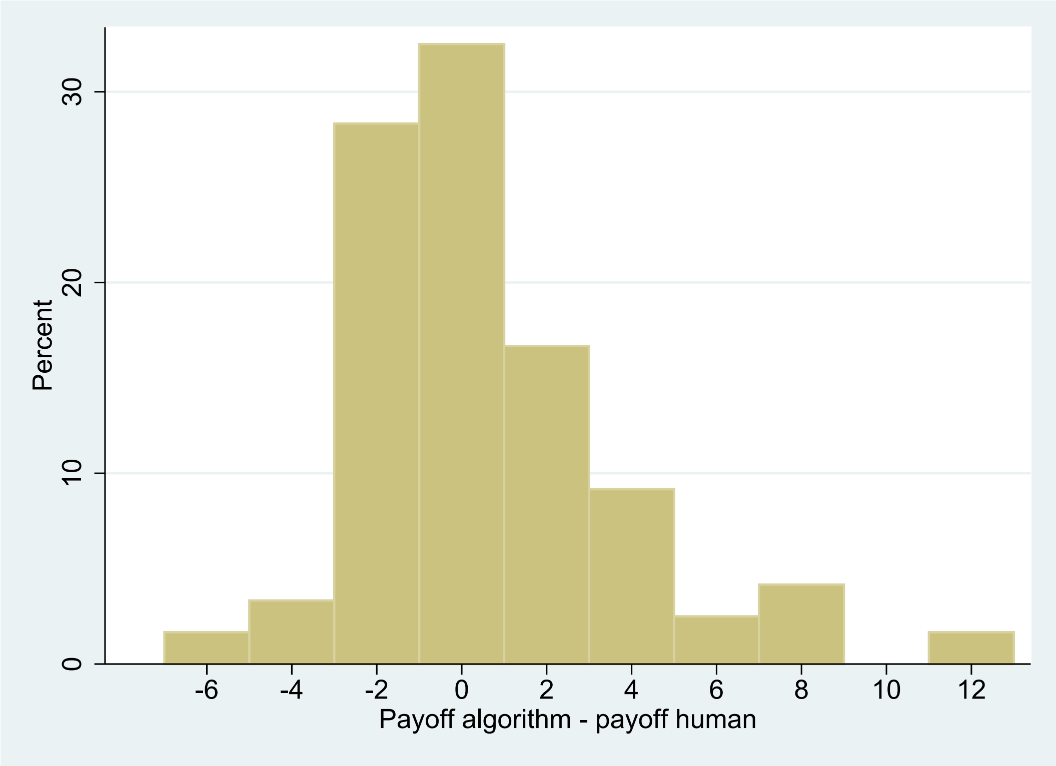


Figure 4: Human fund manager choices

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.

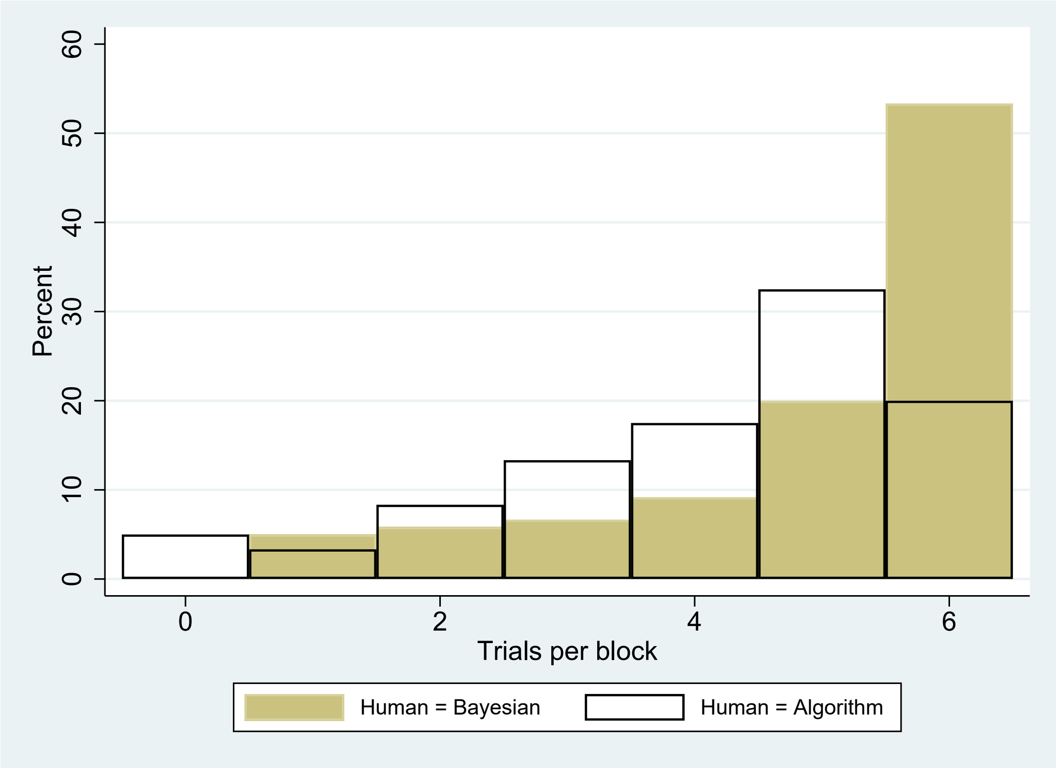


Figure 5: Choice of investment algorithm over time

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

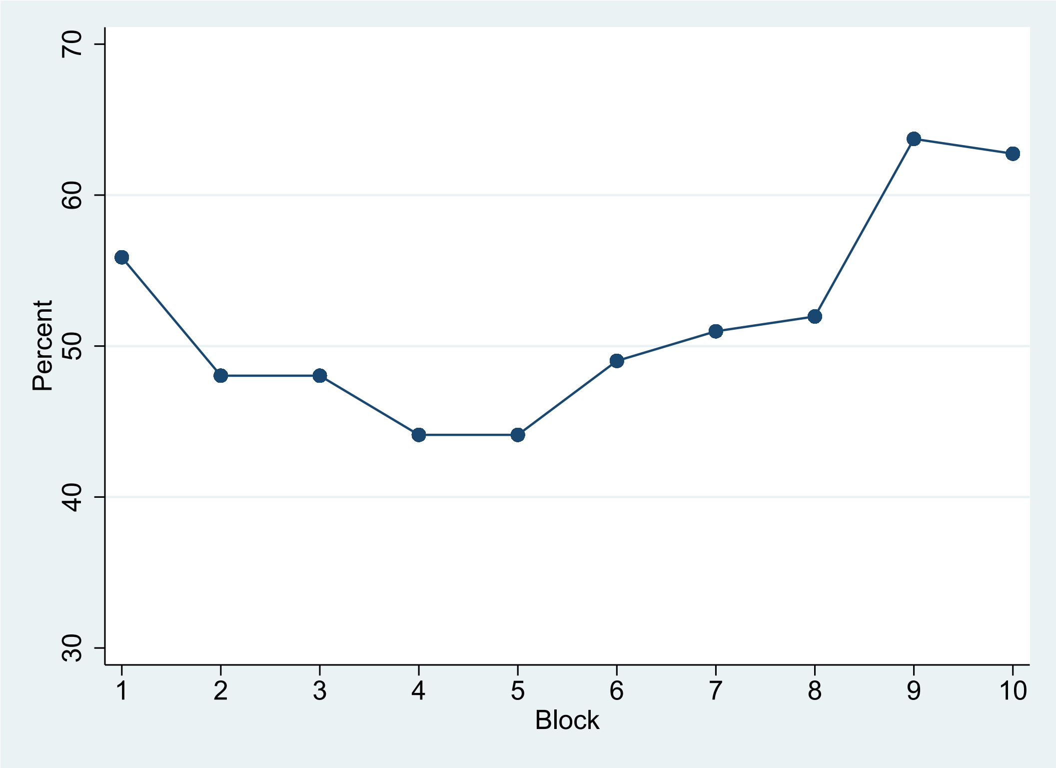


Figure 6: Choice of investment algorithm by difference in last payoffs

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

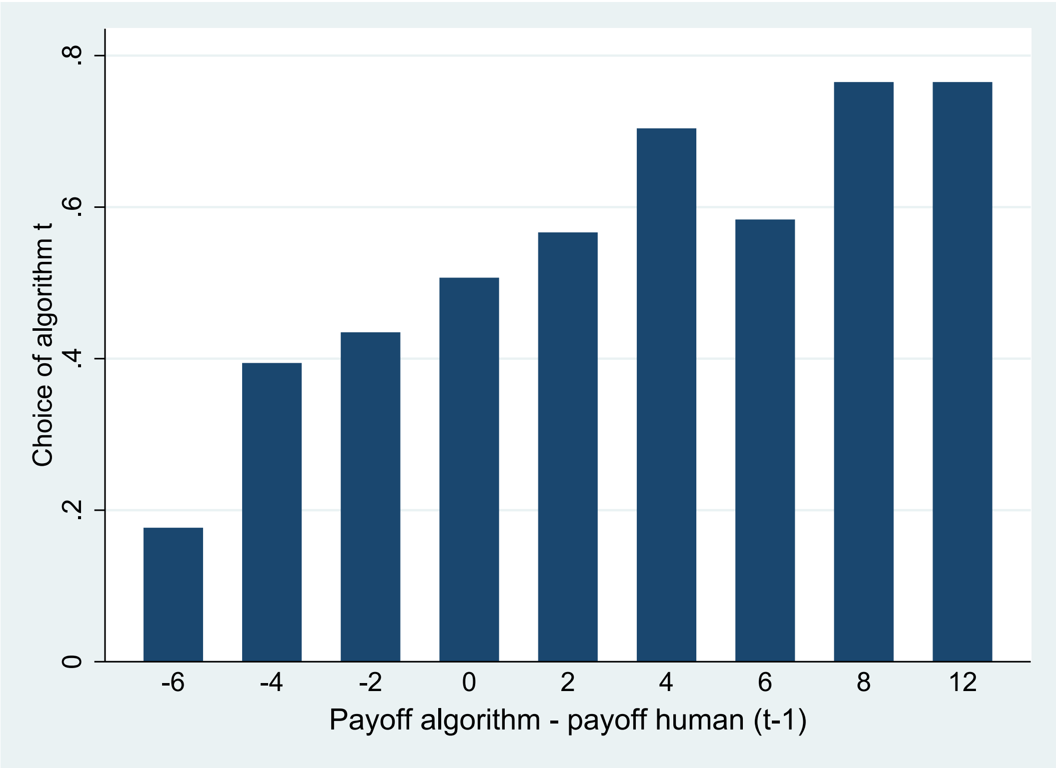


Table 1: Survey questions on the perception of algorithms in finance

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.

Table 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 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

Table 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 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

Table 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) X cum. payoff algorithm					0.008
					(0.008)
Algorithm (t-1) X 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

Table 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

Table 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 7: Luck and skill of financial intermediaries

This table shows total payoffs, luck and skill of both financial intermediaries. 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

Table 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 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

Online Appendix A. 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.

Figure A.1: Experimental instructions. 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:

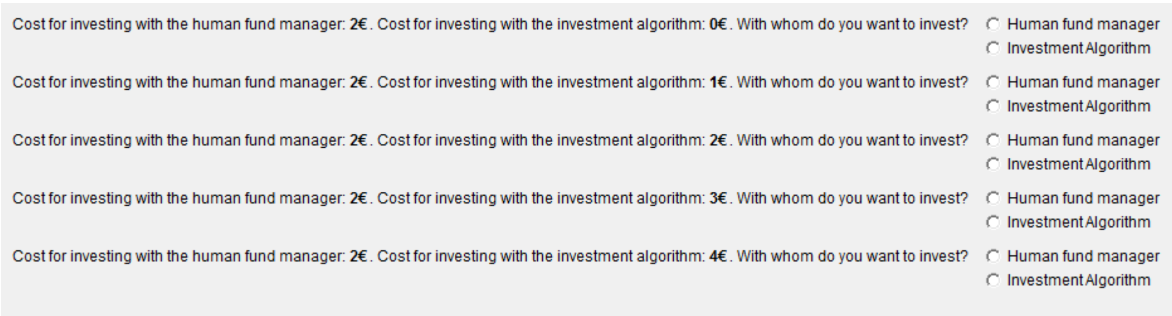
In each trial, the participant appointed as human fund manager is asked to choose to invest into either the stock or the bond. A history of the returns of the stock and the bond, and a history of the investment choices and returns of the investment algorithm is shown in each block.

Figure A.2: Experimental instructions. 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

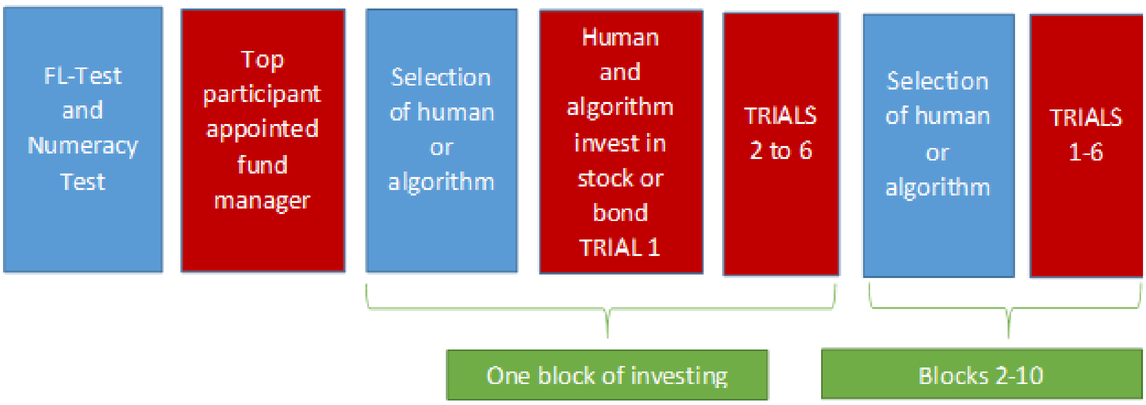


Figure A.3: Experimental instructions. Page 3

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.

Online Appendix B. 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.

Figure B.1: Experimental screens. Financial literacy and numeracy test, page 1/3.

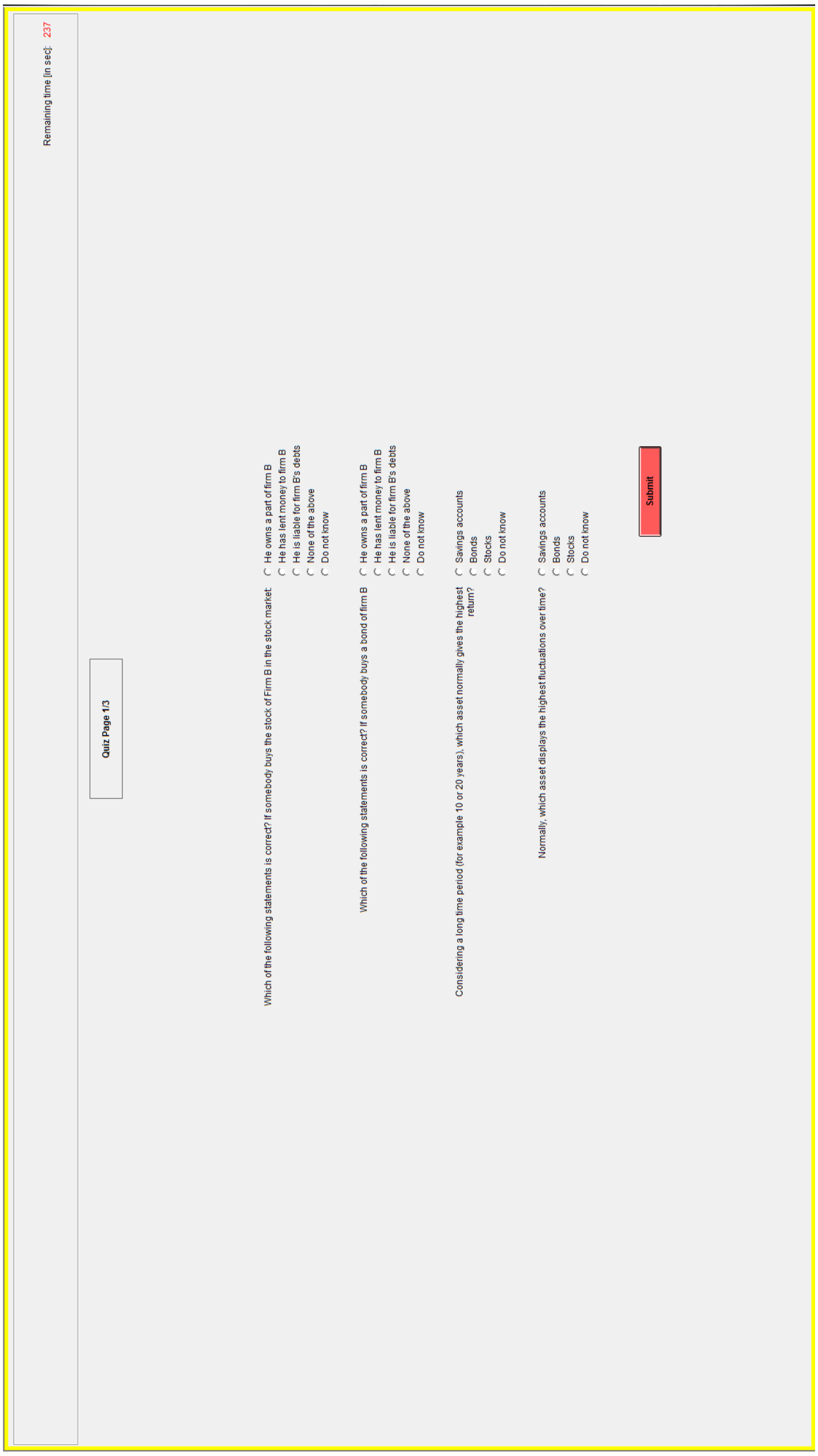


Figure B.2: Experimental screens. Financial literacy and numeracy test, page 2/3.

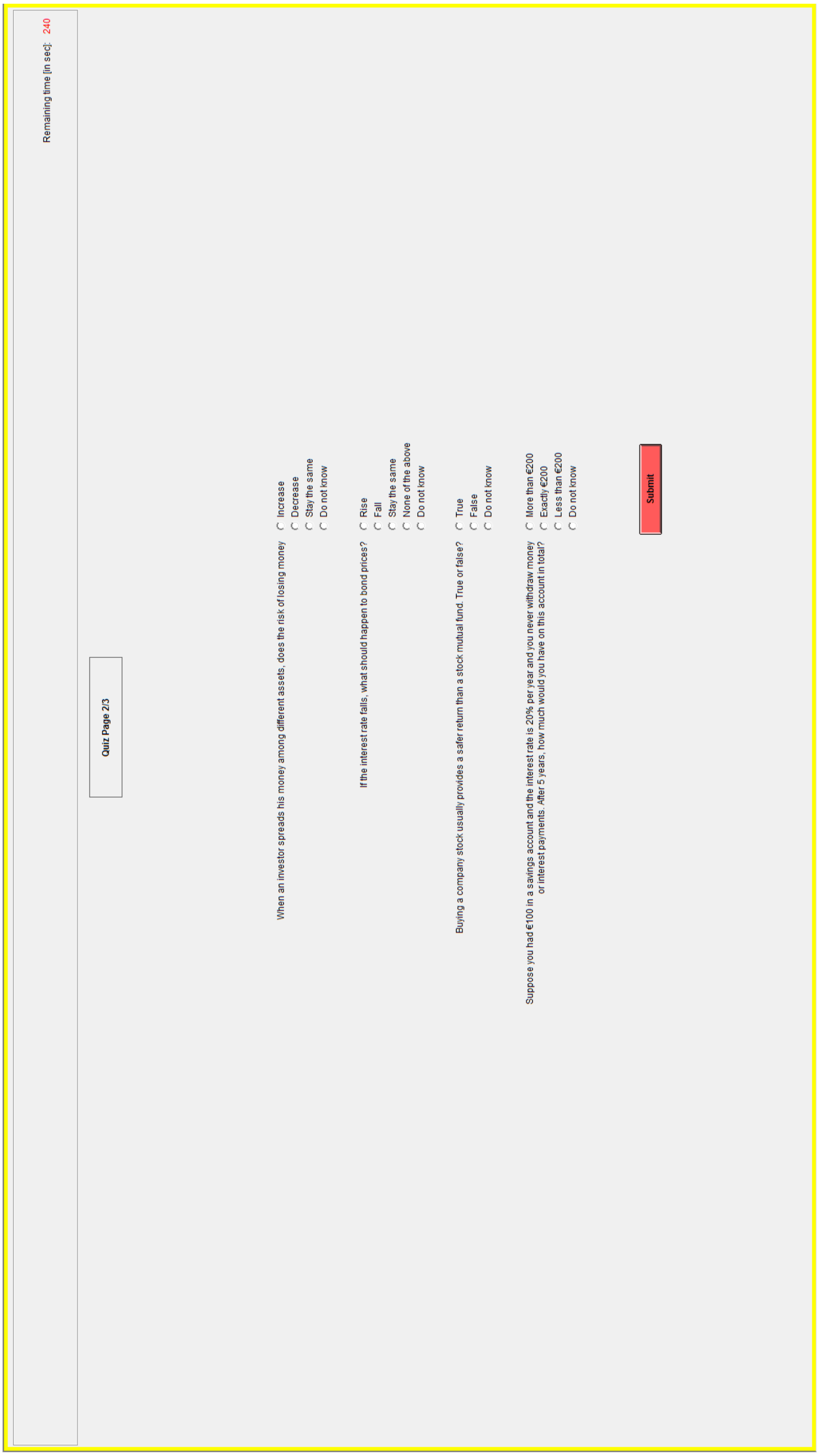


Figure B.3: Experimental screens. Financial literacy and numeracy test, page 3/3.

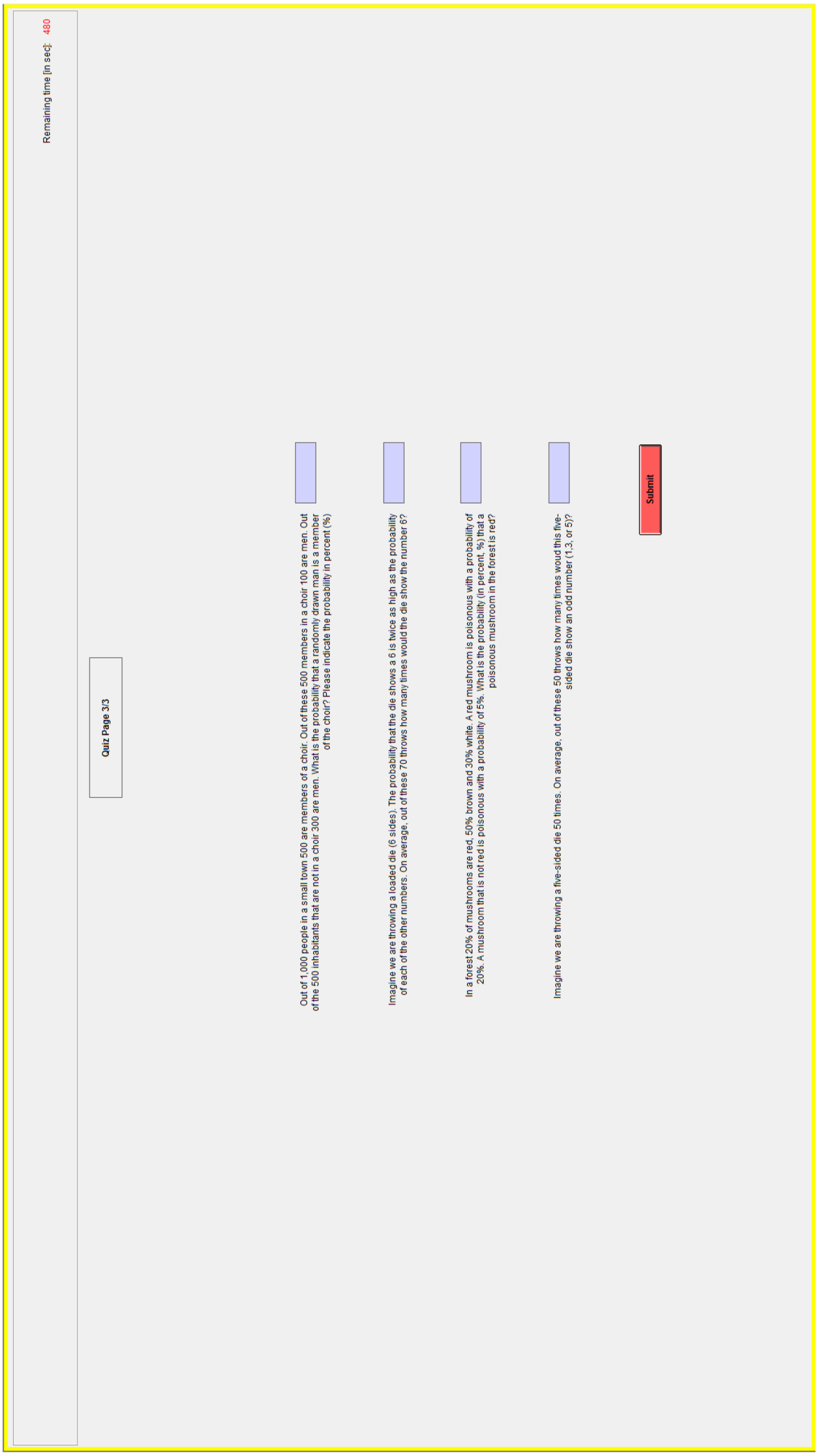


Figure B.4: Experimental screens. Investor choice of financial intermediary.

Figure B.5: Experimental screens. Outcome of random draw of fee combination.

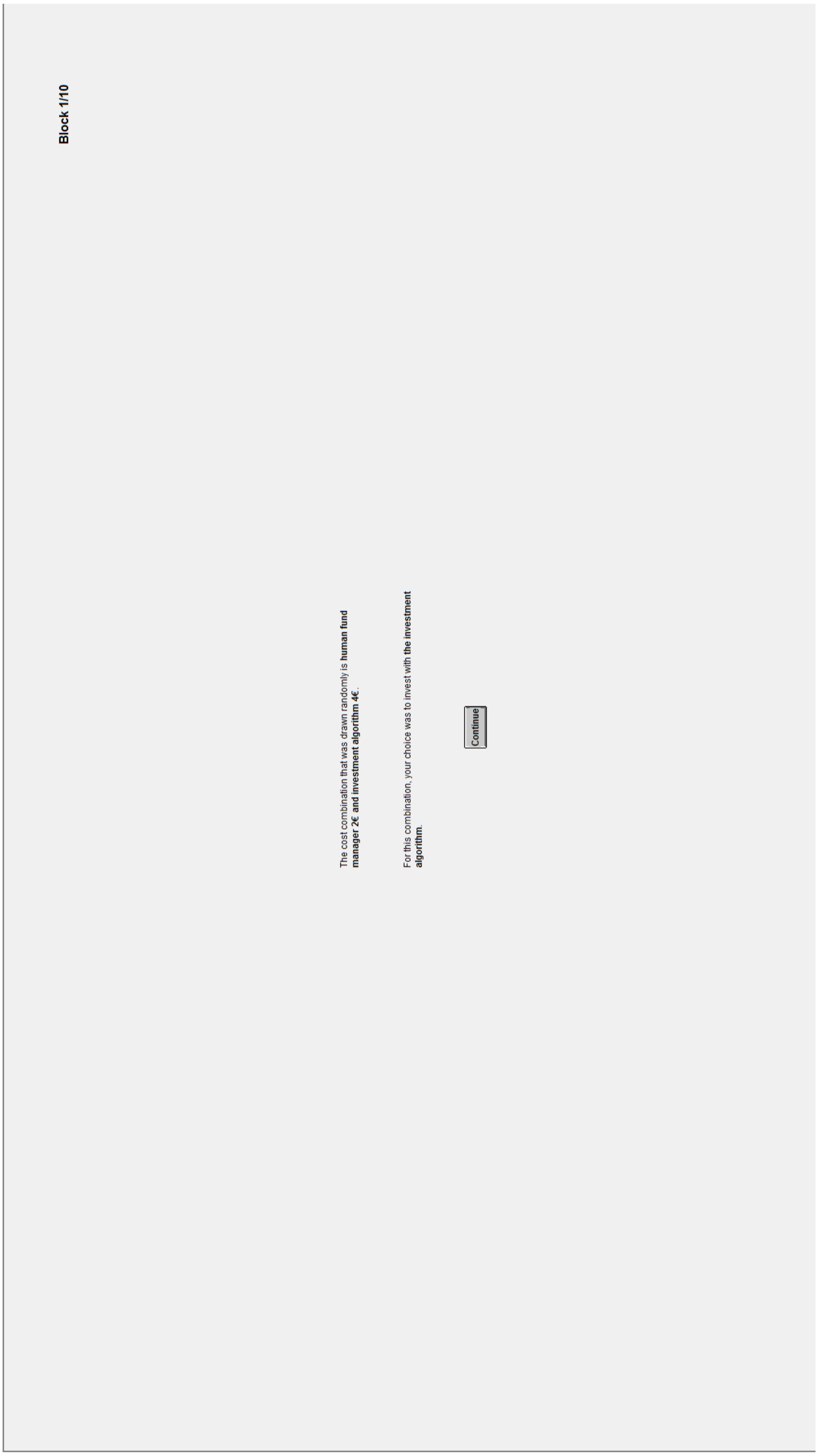


Figure B.6: Experimental screens. Fund manager choice of asset.

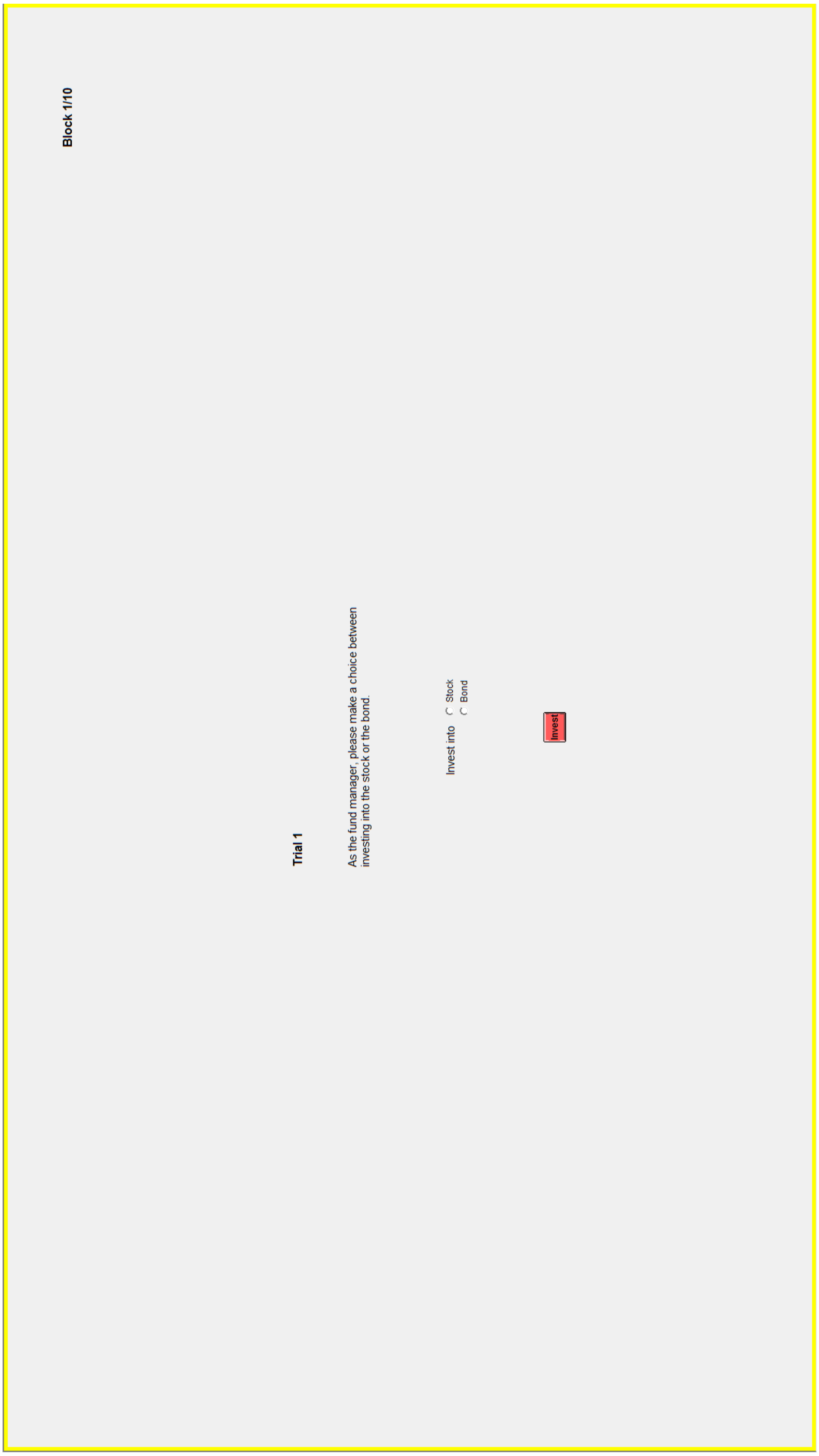


Figure B.7: Experimental screens. Outcome screen within block of investments.

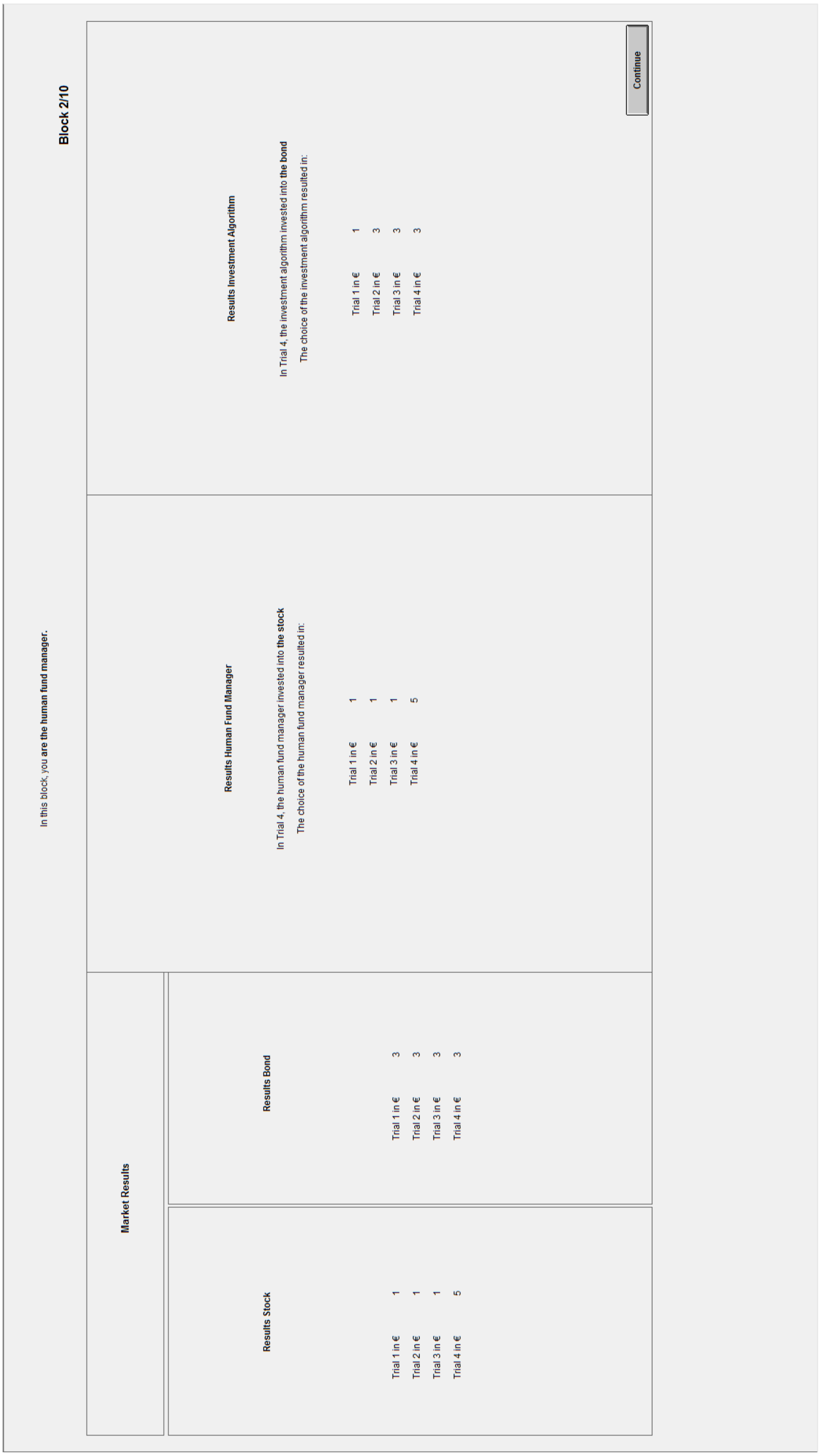
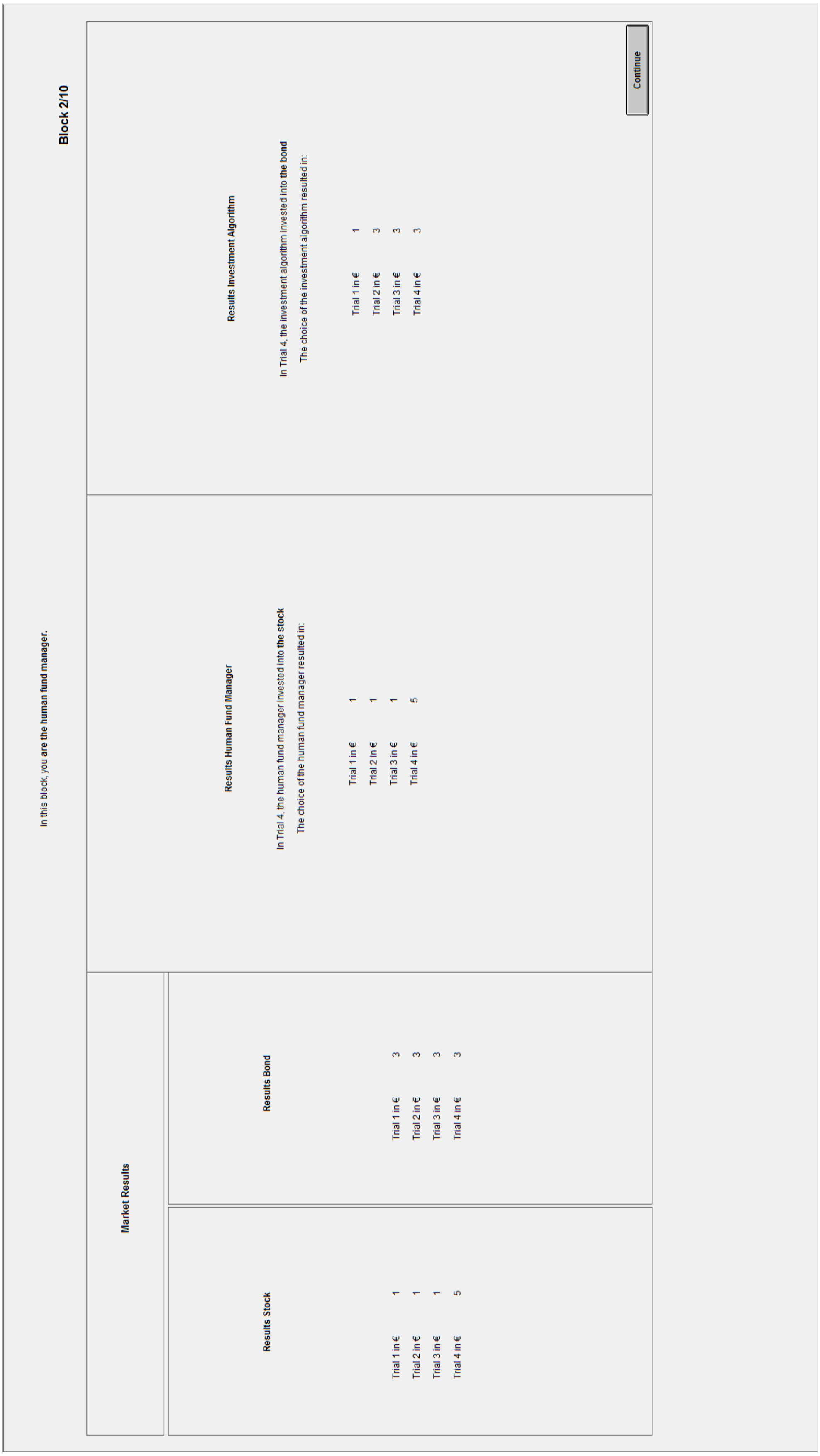


Figure B.8: Experimental screens. Outcome screen after finished block of investments.



Online Appendix C. 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.

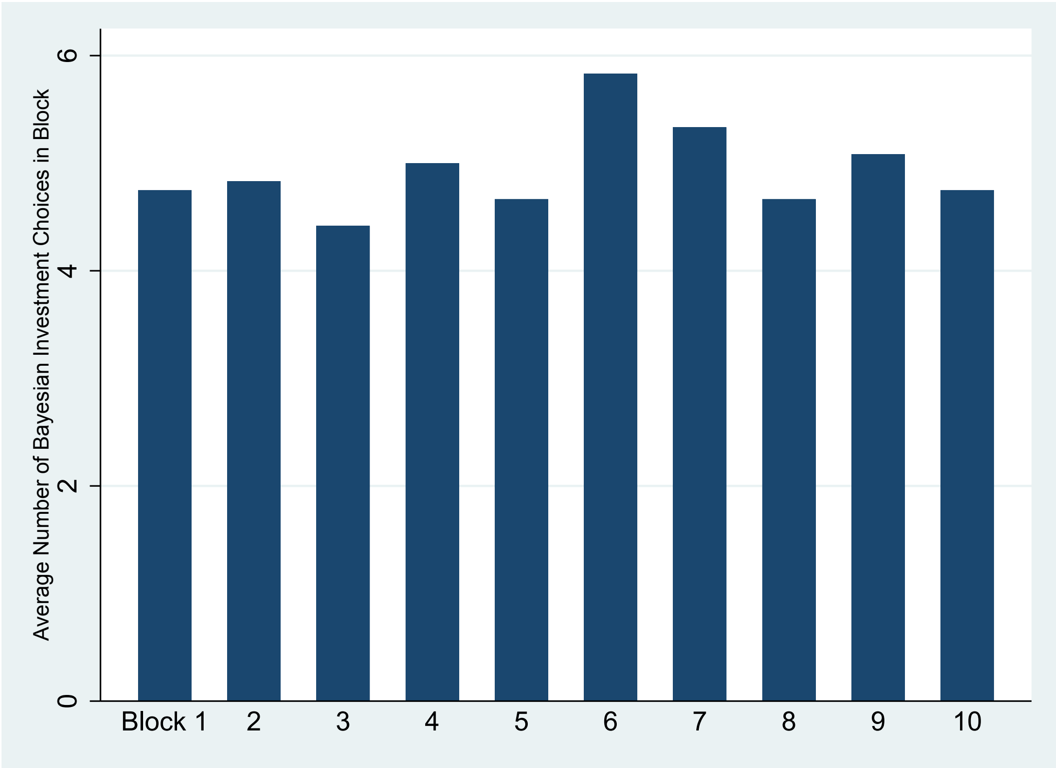


Figure C.2: Performance human and algorithm, by block

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

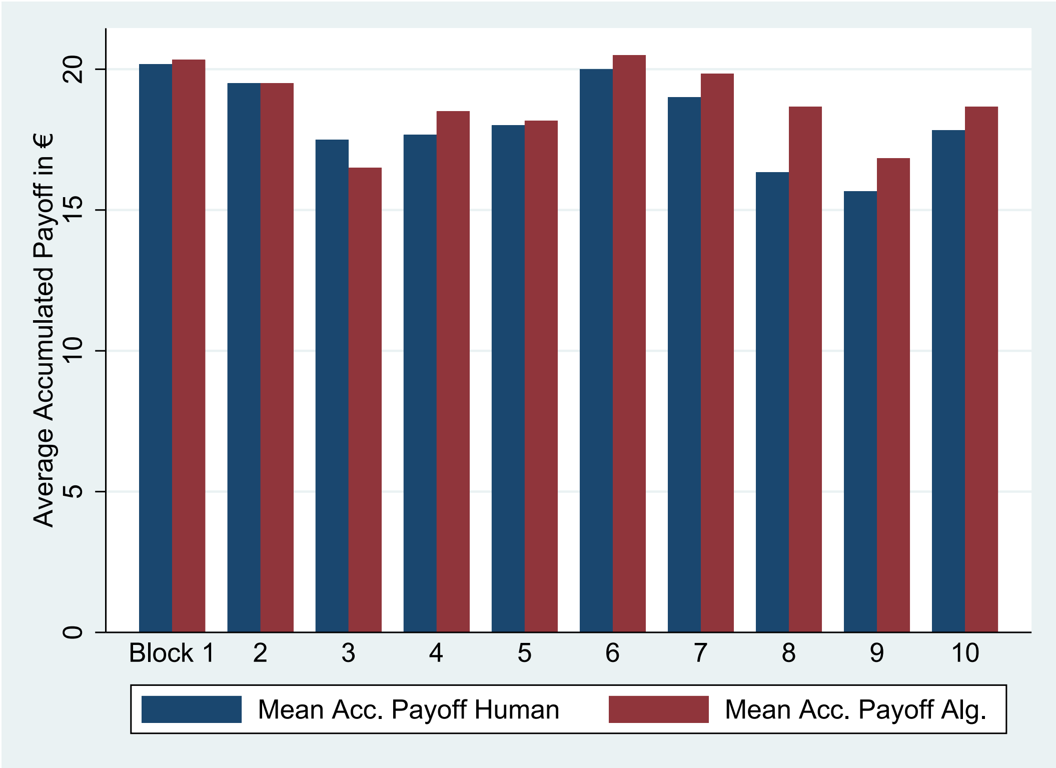


Figure C.3: Choices by block and costs

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.

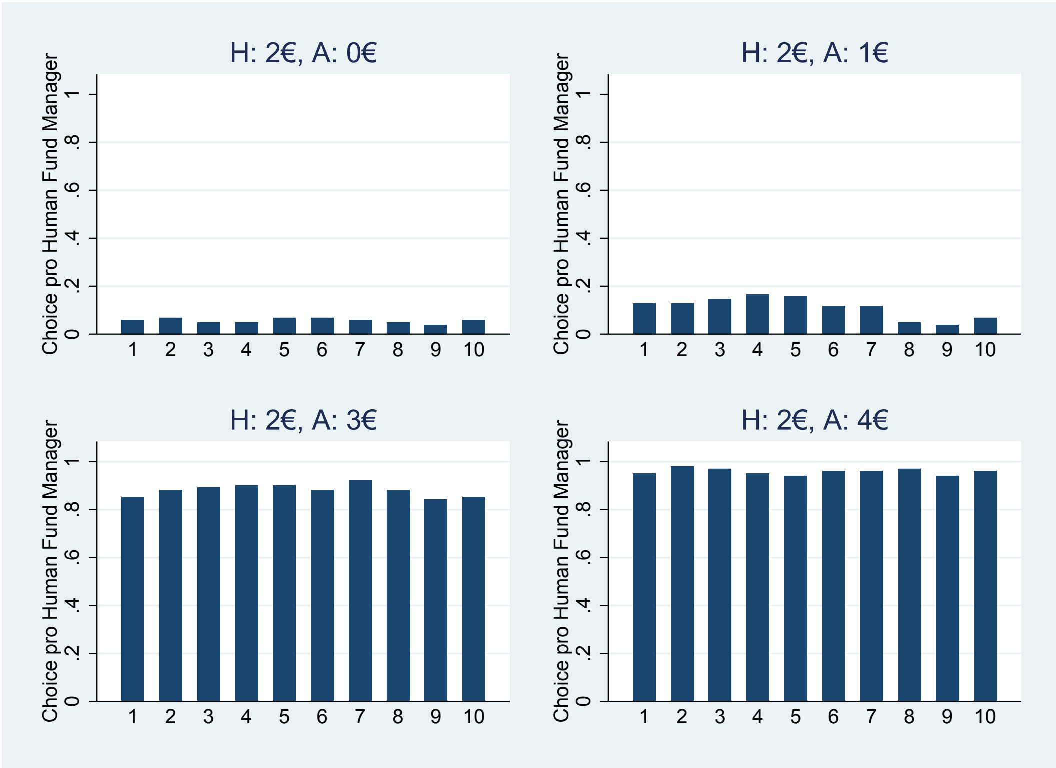


Table C.1: Correlation matrix: Perception of algorithms in finance

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 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***
	(0.0262	(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 $\text{Difference}_{t=0\ to\ t-1}$ measures the performance difference of the human fund manager minus the investment algorithm, accumulated over all blocks to t-1. Positive Performance $\text{Difference}_{t=0\ to\ t-1}$ 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 $\text{Difference}_{t=0\ to\ t-1}$	0.020**
	(0.009)
Positive Performance $\text{Difference}_{t=0\ to\ t-1}$	0.104
	(0.076)
Performance $\text{Difference}_{t=0\ to\ t-1}\times$Positive Performance $\text{Difference}_{t=0\ to\ t-1}$	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