Working Paper 2020:24

Department of Economics
School of Economics and Management

# Delegation Decisions in Finance

Felix Holzmeister
Martin Holmén
Michael Kirchler
Matthias Stefan
Erik Wengström

November 2020

LUND UNIVERSITY

# Delegation Decisions in Finance

Felix Holzmeister[†] Martin Holmén[‡,*] Michael Kirchler[†,‡]
Matthias Stefan[†] Erik Wengström[§,¶]

[†] University of Innsbruck, Department of Banking and Finance
[‡] University of Gothenburg, Department of Economics, Centre for Finance
[§] Lund University, Department of Economics
[¶] Hanken School of Economics, Department of Finance and Economics

[*] *Corresponding author:* martin.holmen@gu.se

**Abstract**

We run an online experiment with finance professionals and subjects from the general population (clients) to examine drivers and implications of clients' delegation decisions. We find that clients favor delegation to investment algorithms, followed by delegation to finance professionals with aligned incentives and lastly to those with fixed incentives. We also show that trust in investment algorithms or money managers (finance professionals), respectively, and clients' propensity to shift blame on others increases the likelihood of delegation, whereas own decision-making quality is associated with a decrease. In measuring the implications of clients' delegation decisions, we report high variability among finance professionals' perceptions of clients' preferred risk levels. We show that this results in overlaps in portfolio risk across risk-levels of clients, indicating problems of risk communication between clients and their money managers.

*JEL:* C93, G11, G41.
*Keywords:* Experimental finance, finance professionals, delegation decisions.

We thank Pol Campos-Mercade, Alain Cohn, Dawei Fang, Håkan J. Holm, Christian König-Kersting, Michael Razen, Markus Walzl, seminar participants at the Center of Social and Economic Behavior in Cologne, the Berlin Behavioral Economics Colloquium and Seminar, the Helsinki GSE Colloquium, as well as conference participants of the Organizations and Society Workshop 2019 at the University of Innsbruck, the NOeG 2019 (Austrian Economic Association) in Graz, the SFB Workshop 2019 in Innsbruck, the Experimental Finance Conference 2019 in Copenhagen, the Decision-Making for Others Workshop in Nijmegen 2018, and the Workshop on Trust and Cooperation in Markets and Organizations in Stavanger 2019 for very valuable comments. We particularly thank Fredrik Bergdahl at *Statistiska centralbyrån* for the excellent collaboration on the project. Financial support from the Austrian Science Fund FWF (START-grant Y617-G11 and SFB F63), the Austrian National Bank (grant 17788) and the Swedish Research Council (grant 2015-01713) is gratefully acknowledged. This study was ethically approved by the review boards at *Statistiska centralbyrån* (*SCB*; Statistics Sweden) and in Gothenburg (Sweden).

## 1. Introduction

Given the complexity of financial products and markets, private investors frequently opt for delegating decisions to finance professionals. This involves decisions about portfolio investments, insurance and pension plans, and seeking advice on various other financial aspects. The economic importance of delegated decision-making in finance is indicated by the large and growing market for financial advice and decision-making on behalf of clients.[1]

While several reasons for private investors' delegation decisions have been accounted for, a general approach of providing causal evidence on their relative merit is widely missing. Some studies focus on agent's financial sophistication (see, e.g., Bhattacharya et al., 2012; Calcagno and Monticone, 2015), but other motives might play an important role too. One straightforward motive for delegating investments is the possibility to blame the agent if the investment does not turn out as expected (see indirect empirical evidence by Shefrin, 2007; Chang et al., 2016).[2] Another motive of delegation decisions seems to be trust in the agent, as outlined theoretically by Gennaioli et al. (2015). In particular, they argue that principals (clients) delegate because they are too anxious to make risky investment decisions. The authors hypothesize that if money managers are trusted, they give confidence to principals to take risks on their behalf.

However, the empirical questions remain open whether trust and blame shifting motives indeed drive delegation propensity. As a related point, it is also unclear whether delegation propensity might vary with different types of agents and with the incentives of the agents. Regarding the latter, misaligned incentives between clients and agents have been portrayed as a driver of excessive risk taking (Rajan, 2006; Diamond and Rajan, 2009; Dewatripont and Freixas, 2012). When this debate has spilled over to the public, trust in the financial industry has been suffering (Sapienza and Zingales, 2012). With regard to the role of the type of agent, robo advisers have gained relevance in recent years. They promise affordable advice (D'Acunto et al., 2019), although algorithm aversion might hinder their acceptance (Germann and Merkle, 2019). It is therefore particularly interesting whether clients prefer human advisers – conditional on their incentive structure – or investment algorithms and which personal characteristics drive delegation choices to either one. Thus, we ask the following research question in this paper: What drives clients' delegation decisions to different types of money managers?[3]

It is our aim to provide an encompassing causal account on the interplay of clients' delegation decisions (moderated by their preferences and characteristics) and the type of "money manager" (investment algorithm or finance professionals with different incentive structures). We thus report from a novel delegation experiment with participants from a sample of Swedish finance professionals and a representative sample of the Swedish general population. We set up six treatments, differing in (i) the pool of subjects enrolled (finance professionals or general population) and in (ii) the type of money manager the principal can delegate to (investment algorithm, finance professional with aligned incentives, and finance professional

[1] For instance, in 2017, the net asset value of US mutual funds equaled 18.8 trillion USD (https://perma.cc/5VUB-U98U), and over 271,000 professionals were employed as personal financial advisers in the United States (https://perma.cc/5RYT-CP6H).

[2] For a general account of shifting blame see, e.g., Bartling and Fischbacher (2012).

[3] Note that we use the terms "principal" and "client" synonymously throughout the paper. Moreover, we also apply the terms "money manager" and "agent" interchangeably for our sample of finance professionals.

with fixed payment). In 25 investment decisions, subjects had to allocate an endowment across two or five investment alternatives that differed in their expected payout, riskiness, and diversification potential. Subjects from the general population were thereafter given the opportunity to delegate their decisions by replacing their own investments with those of a financial professional/investment algorithm for their payout from the experiment. Invitations were sent out via *Statistiska centralbyrån* (*SCB*; Statistics Sweden) and 408 finance professionals and 550 people from the general population completed the experiment. The collaboration with *SCB* allowed accessing a stratified sample of finance professionals, restricted to financial analysts, investment advisers, traders, fund managers, and financial brokers. We are among the first to study delegation with a restricted sample of finance professionals – i.e., skilled employees that take financial decisions for clients in their day-to-day work – and the general population instead of student participants.[4] Additionally, we obtained a set of predefined variables of the subjects' register data for those who completed the experiment from *SCB*.

Our study provides the following main insights: First, we show that clients are most likely to delegate to an investment algorithm, followed by professionals with aligned incentives and professionals with fixed incentives. However, once clients decide to delegate, their willingness to pay does not differ between treatments. Second, on an aggregate level, we observe that trust is a major motive for delegation to agents and that blame shifting motives increase and own decision making quality decrease delegation propensity. In particular, we find that trust in the agent is explanatory for delegation propensity irrespective of the type of agent – this finding is especially pronounced for delegation decisions to investment algorithms. Third, we find that principals tend to ask the agent to take higher levels of risk compared to the perception of risk they took in their own decisions. This is in line with the theoretically postulated explanation of finance professionals as money doctors in Gennaioli et al. (2015). However, we finally show that this delegation channel of trusting the agent to take risky decisions faces substantial communication problems: we find considerable overlaps in the risk of portfolios implemented by professionals on behalf of clients, implying that clients requesting different risk levels may end up with similar portfolios. Finally, our study contains several limitations, which we discuss in detail in the conclusion.

Our study adds to several areas in the literature. First, we contribute to the expanding literature on delegation decisions to investment algorithms and robo advice. This emerging literature offers important insights on the impact of robo advice on investors (see, e.g., D'Acunto et al., 2019; Rossi and Utkus, 2020). Previous research from other domains also shows algorithm aversion, with people distrusting advice from algorithms more than those based on human judgement (see, e.g., Dietvorst et al., 2014; Harvey et al., 2017; Longoni et al., 2019). However, the evidence is mixed, and some studies report no effect or even algorithm appreciation (e.g., Germann and Merkle, 2019; Logg et al., 2017). However, all of these studies remain largely silent about delegation motives like trust and blame shifting. We contribute by showing that investment algorithms are more frequently selected than finance professionals. In addition, we contribute

[4] Note that there is also a growing strand of literature using experiments with student or partly with general population samples to examine risk taking in delegated decision-making (but not with finance professionals acting as money or investment managers). Several studies report a "risky shift" in risk-taking, indicating that decision-makers take more risks or show less loss-averse behavior for others than for themselves (e.g., Sutter, 2009; Chakravarty et al., 2011; Andersson et al., 2016; Vieider et al., 2016). However, several studies also report a "cautious shift" when the money of third parties is invested (Bolton and Ockenfels, 2010; Eriksen and Kvaløy, 2010)—see Füllbrunn and Luhan (2015) and Eriksen et al. (2017) for overviews. Moreover, Andersson et al. (2019) report that decision-makers in their experiments respond to incentives to increase risk-taking on behalf of others, resulting in an increased risk exposure of the principals.

by outlining that trust in investment algorithms and lack of clients' decision making quality are important channels of delegation decisions to investment algorithms.

Second, we add to the literature on incentives of money managers. There is evidence that money managers have an incentive not to correct investors' biased beliefs (Mullainathan et al., 2012). This finding is also related to the credence goods characteristics of financial advice (Dulleck and Kerschbamer, 2006; Inderst and Ottaviani, 2012a,b,c). In addition, misaligned incentives have been portrayed as major contributors to the last financial crisis and as main drivers of excessive financial risk taking in general (Rajan, 2006; Diamond and Rajan, 2009; Bebchuk and Spamann, 2010; Financial Crisis Inquiry Commission, 2011; Dewatripont and Freixas, 2012).[5] Since this debate has spilled over to the public and affected trust in the finance industry (Sapienza and Zingales, 2012), incentives might also play a role for clients' decisions whether to delegate. We contribute by showing that clients prefer delegating to advisers with aligned rather than with fixed incentives.

Third, we add to the literature on trust in the finance industry (Sapienza and Zingales, 2012; Zingales, 2015) and related concepts like blame shifting and risk delegation. For instance, Chang et al. (2016) show that delegation reverses the disposition effect by allowing the investor to blame the money manager, making it easier for the investor to sell losing assets. Moreover, stock market participation and financial development have been shown to be conditional on individuals' trust in general and in the finance sector in particular (Guiso et al., 2004, 2008). There is also evidence for the relevance of trust for financial advice seeking (Lachance and Tang, 2012). In particular, Gennaioli et al. (2015) argue that trust is an important factor for financial delegation in a theoretical framework. Accordingly, when clients trust in money managers, the latter are like "money doctors" and make risky investment decisions the clients are too anxious doing themselves. With our experiment we contribute by providing empirical evidence that trust is a major driver of financial delegation, particularly also for delegation decisions to investment algorithms. We also outline that principals ask the agent to take higher levels of risk than they perceive they implemented themselves, which is in line with Gennaioli et al. (2015). Thus, we summarize that risk communication poses a major issue in this trust-based client-adviser-relationship, which can potentially hinder the trust relationship between principals and agents in the long run.

Fourth, we contribute to the literature on the communication of risk between clients and money managers. For instance, in an experimental study with finance professionals, Kirchler et al. (2020) show that professionals' self-assessed risk attitude in financial matters explains risk-taking on behalf of clients. Moreover, in their empirical study, Foerster et al. (2017) report that adviser fixed effects explain considerably more variation in household portfolio risk than a broad set of investor attributes. Linnainmaa et al. (2019) show that most advisers invest their personal portfolios just as they advise their clients. We contribute by pointing at the problem of communicating risk-levels between principals and agents, resulting in considerable overlaps of portfolio risk across clients' demanded risk classes.

[5] Yet another strand of experimental studies suggests that even strong financial incentives hardly interfere with agents' attempt to adhere to their clients' preferences (see, e.g., Rud et al., 2018; Ifcher and Zarghamee, 2019; Kling et al., 2019).

## 2. Experimental Design

**Allocation decision task.** The workhorse of our experiment is the allocation decision task as used by Banks et al. (2018). The task consisted of 25 items, in which participants were asked to allocate an endowment of 100 SEK on either two or five assets.[6] Participants were informed about the assets' returns per 1 SEK invested, depending on whether a coin toss shows up heads or tails. The returns used in the experiment were adopted from Banks et al. (2018), multiplied by a factor of 1.5, and rounded to one decimal place. The returns for each asset in the 25 investment decisions are depicted in Table 1, and the corresponding opportunity sets are illustrated in Figure B1 in Appendix B.

**Table 1: Parameters of the 25 opportunity sets.** This table indicates returns (in SEK) per 1 SEK invested for the different assets in the 25 opportunity sets, depending on whether the coin toss shows up heads or tails. Within the blocks of two and five assets, the decision problems were randomized in order.

| | **Asset A** | | **Asset B** | | **Asset C** | | **Asset D** | | **Asset E** | |
|---|---|---|---|---|---|---|---|---|---|---|
| *Set* | *Heads* | *Tails* | *Heads* | *Tails* | *Heads* | *Tails* | *Heads* | *Tails* | *Heads* | *Tails* |
| #1 | 0.00 | 1.20 | 3.60 | 0.00 | | | | | | |
| #2 | 3.60 | 0.00 | 0.00 | 1.80 | | | | | | |
| #3 | 4.80 | 0.00 | 0.00 | 1.20 | | | | | | |
| #4 | 2.30 | 0.00 | 0.00 | 4.50 | | | | | | |
| #5 | 0.00 | 2.40 | 2.40 | 0.00 | | | | | | |
| #6 | 1.20 | 0.00 | 0.00 | 4.80 | | | | | | |
| #7 | 0.00 | 2.30 | 4.50 | 0.00 | | | | | | |
| #8 | 0.00 | 3.60 | 1.80 | 0.00 | | | | | | |
| #9 | 0.00 | 2.70 | 3.00 | 0.00 | | | | | | |
| #10 | 1.20 | 0.00 | 0.00 | 3.60 | | | | | | |
| #11 | 0.30 | 2.70 | 0.90 | 0.90 | 1.20 | 0.00 | 0.60 | 1.80 | 0.00 | 3.60 |
| #12 | 0.80 | 1.50 | 2.40 | 0.00 | 0.40 | 2.10 | 1.80 | 0.80 | 0.00 | 3.00 |
| #13 | 2.30 | 0.60 | 0.40 | 1.50 | 0.00 | 2.40 | 1.50 | 0.90 | 3.00 | 0.00 |
| #14 | 0.50 | 4.10 | 1.80 | 0.00 | 0.00 | 5.40 | 0.90 | 2.70 | 0.50 | 0.50 |
| #15 | 2.70 | 0.30 | 3.60 | 0.00 | 0.00 | 1.20 | 0.90 | 0.90 | 1.80 | 0.60 |
| #16 | 2.00 | 1.20 | 3.50 | 0.40 | 4.50 | 0.00 | 0.00 | 3.00 | 1.10 | 2.30 |
| #17 | 1.40 | 0.20 | 0.00 | 1.80 | 0.50 | 1.40 | 0.80 | 0.80 | 1.80 | 0.00 |
| #18 | 2.70 | 0.50 | 3.60 | 0.00 | 0.90 | 1.40 | 0.00 | 1.80 | 1.80 | 0.90 |
| #19 | 0.00 | 2.40 | 2.40 | 0.00 | 1.80 | 0.60 | 0.60 | 1.80 | 1.20 | 1.20 |
| #20 | 0.00 | 4.50 | 3.00 | 0.00 | 2.00 | 0.80 | 0.40 | 3.50 | 1.50 | 2.30 |
| #21 | 0.00 | 3.60 | 2.70 | 0.90 | 3.60 | 0.00 | 1.50 | 1.50 | 0.60 | 2.70 |
| #22 | 2.40 | 0.40 | 1.80 | 0.80 | 0.00 | 2.40 | 3.60 | 0.00 | 0.90 | 1.80 |
| #23 | 0.30 | 2.70 | 1.50 | 0.60 | 1.20 | 1.80 | 2.40 | 0.00 | 0.00 | 3.60 |
| #24 | 5.40 | 0.00 | 2.70 | 0.90 | 0.50 | 0.50 | 0.00 | 1.80 | 4.10 | 0.50 |
| #25 | 0.50 | 2.70 | 1.80 | 0.00 | 1.40 | 0.90 | 0.90 | 1.80 | 0.00 | 3.60 |

The task consists of 10 decisions with two binary assets in a first block, and 15 decisions with five binary assets in a second block. Participants were first presented with the task instructions for the first block. After reading the instructions, participants could only continue once they had correctly answered three

[6] By the end of February 2019, the exchange rate between US dollar and SEK was about 1:9 and between Euro and SEK about 1:10.5.

comprehension questions.[7] After the first ten decisions, participants were informed that five rather than two assets would be available for the remaining 15 decisions. The order of the two blocks was fixed for all subjects, but the order of decisions was randomized in each of the two blocks. Figure A1 in Appendix A shows two screenshots of the main experimental task, i.e., the allocation decision task, with two and five assets, respectively.

At the end of the experiment, one of a subject's own or—in case a client opts for delegating the decisions—one of the agent's decisions was randomly chosen, and a simulated coin toss determined the participant's payoff. Importantly, returns were paid on top of the endowment, i.e., final payments could not fall below 100 SEK.

**Decision-making quality index:** Similar to Banks et al. (2018), we determine four measures of decision-making quality (*DMQI*) based on the allocation decision task in our experiment: violations of first order stochastic dominance *FOSD*, violations of the general axiom of revealed preferences *GARP*, financial competence *FC*, and failure to minimize risk *FMR*. For each participant, the predicted values of a principal component analysis of the four measures constitute our decision-making quality index (*DMQI*). Detailed descriptions on how each of the decision-making measures is defined are provided in Appendix C.

**Experimental treatments.** Depending on the subject pool, participants were randomly assigned to one of the treatments listed in Table 2. Common to all treatments, both for finance professionals and for the general population sample, is the 25-item allocation decision task, which is described in detail below.

After having completed all items of the allocation decision task, participants from the general population (principals) had the opportunity to delegate their decisions to an agent. If principals opted for delegating their decisions, the experimental payoff depended on the agent's rather than their own decisions. The design choice that principals made the investment decision first, but were informed about the opportunity to delegate the investment decisions only afterwards, warrants further discussion. While potential "clients" that do not want to engage in financial matters at all might have dropped out initially[8], there are clear advantages for this design feature: First, the opportunity to delegate without prior decisions could lead to delegation in order to receive an experimental payment without spending any effort. Second, our experimental design allows studying whether or not delegation pays off for those who delegate and those who stick to their own decisions, since investment choices are observed for all participants, irrespective of the decision whether or not to delegate. Moreover, we can study risk communication by comparing clients' and professionals' investment decisions conditional on risk levels.

Depending on the treatment, the principals' delegation was either to an investment algorithm programmed by the experimenters (*GP-ALGO*), a finance professional with aligned, i.e., linear, incentives (*GP-ALIGNED*), or a finance professional receiving a flat payment of 200 SEK for deciding on behalf of one or more clients (*GP-FIXED*). Note that, compared to the baseline condition *GP-FIXED*, treatment *GP-ALIGNED* modifies the incentive

[7] Note that those subjects who did not answer the comprehension questions correctly, had the opportunity to look at the instructions again until they got the answers right. In addition, they received hints on the correct answers.

[8] For a comprehensive response rate analysis and a discussion of potential self-selection effects, please refer to Appendix E.

**Table 2: Treatment overview.** This table illustrates the randomly assigned between-subjects treatments for both samples, finance professionals and participants from the general population. The sample sizes per condition are indicated in Figure 1.

| | **Finance professionals** *… make decisions …* | | **General population** *… can delegate decisions to …* |
|---|---|---|---|
| *FP-OWN* | on one's own account | *GP-ALGO* | investment algorithm |
| *FP-ALIGNED* | for third party (linear incentives) | *GP-ALIGNED* | finance professional (linear incentives) |
| *FP-FIXED* | for third party (flat payment) | *GP-FIXED* | finance professional (flat payment) |

structure of the agent, while holding the type of agent constant. Treatment *GP-ALGO* modifies the type of agent from a human to an investment algorithm.

If principals chose to delegate, they were asked to specify the risk (on a scale from 1 [no risk] to 4 [maximum return]) they wanted to be taken on their behalf by the agent,[9] as well as their willingness to pay for delegating the investment decisions (between 0 and 50 SEK, in steps of 5 SEK). At the end of the experiment, a "price" for delegating the decision to the agent (between 0 and 50 SEK) was randomly determined: If a participant's willingness to pay was higher than this random number, his/her decisions were delegated to the agent at the randomly determined price (i.e., the agent's decisions were payoff-relevant for the principal); if not, the principal's own decisions were relevant for the payment in the experiment.

Finance professionals were randomly assigned to one of three treatments in which they either made decisions on their own account (*FP-OWN*), or on behalf of subjects from the general population sample. When deciding on principals' account, finance professionals either faced aligned incentives (i.e., they received exactly the same monetary payoff as their client; *FP-ALIGNED*), or were paid a flat fee of 200 SEK (*FP-FIXED*). Moreover, when deciding on behalf of others, finance professionals were asked to comply with a randomly assigned risk level (between 1 [no risk] and 4 [maximum return]). In case a participant from the general population delegated his/her decisions, he/she was matched with a participant from the finance professional sample, based on the particular treatment *and* the stated risk level. All details about the delegation decision itself, the risk levels as a means to communicate the desired riskiness of the allocation decisions, the matching modalities, as well as the payment procedures were common knowledge.

**Questionnaires.** After the allocation decisions (but prior to the choice whether or not to delegate), all participants were asked to self-assess the overall level of risk taken across the 25 items of the allocation

[9] The investment algorithm was programmed to construct investment portfolios, given the particular risk level, as follows: In each investment decision, the minimum variance portfolio and the maximum return portfolio were mapped to the endpoint options of the risk level scale, i.e., 1 and 4, respectively. Thus, risk level 1 was always associated with a sure payoff, whereas risk level 4 always involved a 100% investment in the asset with the highest expected return. For risk levels 2 and 3, portfolio weights were determined in equally sized steps between these fixed endpoints. For instance, if payoffs were 2.40 SEK / 0.00 SEK for asset A and 0.00 SEK / 0.80 SEK for asset B, then the risk-free portfolio was characterized by an investment of 25% in A and 75% in B, whereas the maximum return portfolio corresponded to an investment of 100% in A. Risk levels 2 and 3 were associated with portfolios investing 50% and 75% in A, respectively. To the participants the algorithm was described to be "programmed in such a way that it maximizes your expected profit conditional on the risk level you indicate below". Please note that participants still have to trust that the algorithm operates as described in this statement.

decision task on a scale from 1 to 4, i.e., on the same scale as when choosing the risk level in delegating the risky decisions. In addition, we included the following set of non-incentivized survey items at the end of the experiment: All participants were asked about (i) their self-assessed risk attitude in general and in financial decisions (Dohmen et al., 2011; Falk et al., 2016), (ii) their willingness to abstain from something today for a future benefit (Falk et al., 2016), (iii) their trust in mankind in general, in persons from the finance industry, and in investment algorithms, (iv) their proneness to shift blame on others (Wilson et al., 1990), and (v) their level of prosociality in a hypothetical charitable giving setting (Falk et al., 2018). Furthermore, we included a 5-item questionnaire on delegation and advice-seeking in financial decisions, which was only posed to participants that indicate that they have been active in the financial market. Afterwards, all participants had four minutes to answer an 8-item Rasch-validated numeracy inventory (Weller et al., 2013), including two questions on cognitive reflection. In addition, participants had to provide their self-assessment of the number of correct answers in the numeracy questionnaire as well as of their ranking compared to a random sample of the Swedish population. These assessments allow us constructing two measures of overconfidence, i.e., overestimation and overplacement. Finally, participants had three minutes to answer a 6-item financial literacy questionnaire based on van Rooij et al., 2011. For further details regarding the survey items, please refer to Appendix D.

**Recruitment and data collection.** We conducted an online experiment in Sweden in cooperation with *Statistiska centralbyrån* (*SCB*; Statistics Sweden), who invited subjects for the experiment and provided a set of predefined variables of the subjects' register data for those who completed the experiment. *SCB* sent out invitations (including a hyperlink to the online experiment and a personalized alphanumeric identifier serving as login credentials) to 8,215 finance professionals and a randomly selected representative sample of 8,215 subjects from Sweden's working general population, excluding finance professionals. The sample of finance professionals includes financial analysts and investment advisers, traders and fund managers, and financial brokers. For the general population, following Edin and Fredriksson (2000) and Böhm et al. (2018), we only include people with a declared labor income exceeding the minimum amount that qualifies for the earnings related part of the public pension system. Invitations were sent out in two waves. 20% of the sample were invited in the first week of 2019. Since no technical issues had arisen, the remaining 80% of the sample were invited in the third week of 2019.

Once participants logged in to the online software, programmed in *oTree* (Chen et al., 2016), using their personal identifier, they were presented with a detailed outline of the experiment. In particular, on the first screen, participants were informed that register data provided by *SCB* will be matched with the data collected in the experiment. Moreover, participants were informed that the study has been approved by the ethical review boards in Gothenburg and at *Statistiska centralbyrån*. Participants agreed upon the conditions and were directed to the instructions of the experiment. The data handling procedures ensured full pseudonymity of all participants. Further details and additional information on the recruitment, data collection, and experimental implementation are provided in Appendix A. Complete versions of the experiment and all treatments (in English) are available online via http://hea-2019-01-en.herokuapp.com.

In total, 408 finance professionals and 550 people from the general population completed the experiment. The experiment was conducted in Swedish and took on average 45 minutes to complete. The average payment to participants was 238.9 Swedish Krona (SEK; *SD* = 122.3), which is approximately $30 given the exchange rate at the beginning of 2019.[10] The experimental data was collected between January 4 and February 10, 2019.[11]

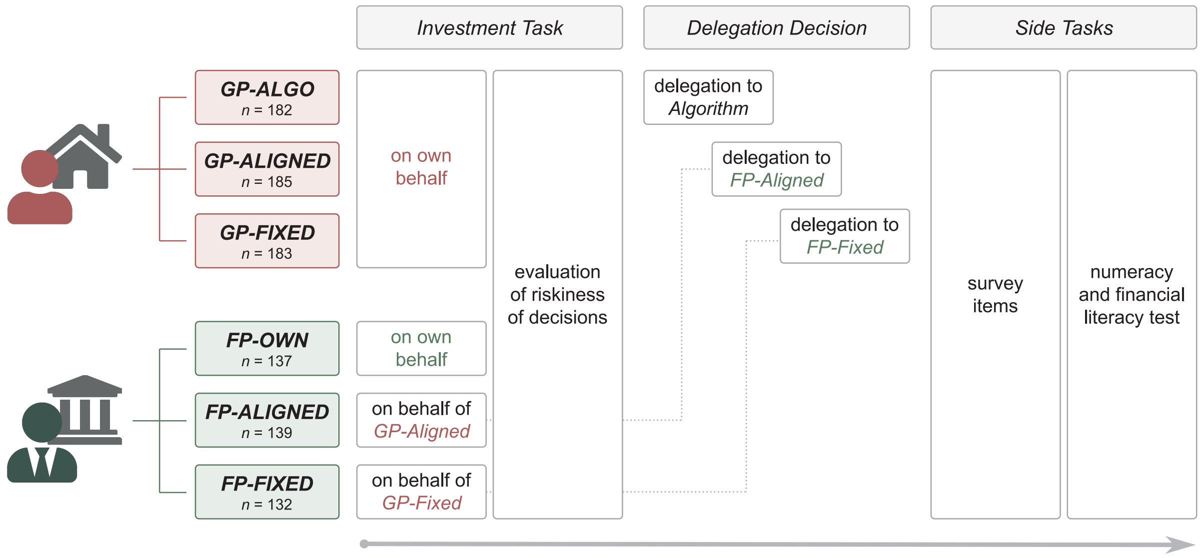


**Figure 1: Flow chart of the experiment.** This figure illustrates the sequence of tasks for subjects in our experiment. First, participants were randomly assigned to one treatment and completed 25 investment decisions. Then, subjects from the general population could delegate their investment decision to an agent in a delegation decision stage. Finally, all subjects completed several side tasks, including self-reported items on economic preferences and supplementary survey questions, a financial literacy test, and a numeracy inventory.

The sequence of tasks within the experiment is graphically summarized in Figure 1. For detailed information on the main task, please refer to Appendix B. Details on the side tasks and questionnaires are provided in Appendix D. Analyses on subjects' decision times across subject pools, treatments, and tasks are summarized in Appendix F, outlining high data quality due to moderate variance across all sub-samples.

**Register data.** In addition to the data collected in the online experiment, we obtained register data from *SCB* for each participant who completed the experiment. In particular, we received data on demographics (e.g., age, gender, income), occupational history (e.g., workplace, firm size), subjects' education, their wealth history, and military records (e.g., scores of the military suitability tests). See Appendix A for further details on these variables. In the analysis of experimental results we only use part of the registry data as control variables – in particular, participants' gender (binary indicator for female), age (in years), net income from major employment in 2017 (in thousand SEK's), and maximum education level (dichoto-

[10] Thus, the average hourly salary for finance professionals and general population subjects amounts to approximately $40. This is comparable to other studies with general population subjects (e.g., Andersson et al., 2016, 2019). This average annual salary is also comparable—although on the lower end—to other studies with financial professionals (Haigh and List, 2005; Kirchler et al., 2018, 2020; Weitzel et al., 2019).

[11] In total, only a relatively small fraction of subjects—especially for an online experiment—dropped out during the experiment. In sum, 68.9% of all subjects that started actually finished the experiment (i.e., 958 out of 1.391). The fraction of completes was 66.3% among the general population and 72.7% among finance professionals, hinting at low and comparable attrition rates across subject pools.

mous indicators for high school education or less, university education smaller or equal to three years, and university education larger than three years).[12]

## 3. Results

In the following, we first examine principals' decisions to delegate across treatments and then identify potential drivers of delegation decisions. Finally, we investigate the communication of requested risk between principals and agents in case of delegation. Note that throughout the presentation of the results, we indicate effect sizes in terms of marginal effects at the means (*MEM*) and/or odds ratios (*OR*) for non-linear models. Descriptive results regarding the samples of finance professionals and the general population are presented in Appendix E and Appendix F.

**Result 1** *Delegation rates are highest when principals can delegate to the investment algorithm, followed by the treatments where agents are professionals with aligned incentives and professionals with a flat payment. Principals' willingness to pay for delegation (once they chose to do so) does neither depend on the type of agent nor on the agent's incentives.*

*Support:* Panel (A) of Figure 2 shows the shares of principals delegating their investment decisions to the agent for each of the three treatments. The estimates are based on logit regressions ($n = 550$; robust standard errors) of the binary delegation choice on treatment indicators. We find that delegation rates increase from 16.9% in treatment *GP-FIXED* to 25.9% in treatment *GP-ALIGNED* (*MEM* $= 0.090$, $p = 0.034$), and to 37.9% in treatment *GP-ALGO* (*MEM* $= 0.210$, $p < 0.001$), respectively. The difference in delegation rates between the treatments *GP-ALIGNED* and *GP-ALGO* is statistically significant (*MEM* $= 0.120$, $p = 0.013$). These results suggest that principals take into account both the type of the agent (i.e., whether the agent is an algorithm or a finance professional) and the agents' incentives (fixed or aligned compensation) in their delegation decisions.

Panel (B) of Figure 2 depicts the principals' mean willingness to pay for delegating their investment decisions to the agent, conditional on the treatment. Please note that the willingness to pay is only elicited for participants that chose to delegate ($n = 148$). The estimates are based on an ordinary least squares regression of principals' willingness to pay on treatment indicators (with robust standard errors). Although delegation rates vary considerably across treatments, we report no statistically significant differences between treatments in principals' willingness to pay for delegation. That is, the willingness to pay for delegation (of principals who chose to delegate) does neither differ conditional on whether the agent is a finance professional or an investment algorithm, nor on whether finance professionals face a flat compensation or a linear incentive scheme.

Result 1 provides a clear indication that principals' delegation decisions are affected by both the type of the agent (i.e., finance professionals vs. investment algorithm) and the incentives faced by human agents (i.e.,

[12] Please note several restrictions in the register data: First, the records for wealth data of *SCB* end in 2007 and other potentially relevant data such as portfolio holdings of assets or bank account data is not tracked by *SCB*. Second, adding data from the military suitability tests would lower the sample size by close to 40 percent, as all female subjects would be dropped from analysis. Focusing only on male subjects would lower the generalizability of our findings for both the general population and the finance industry.

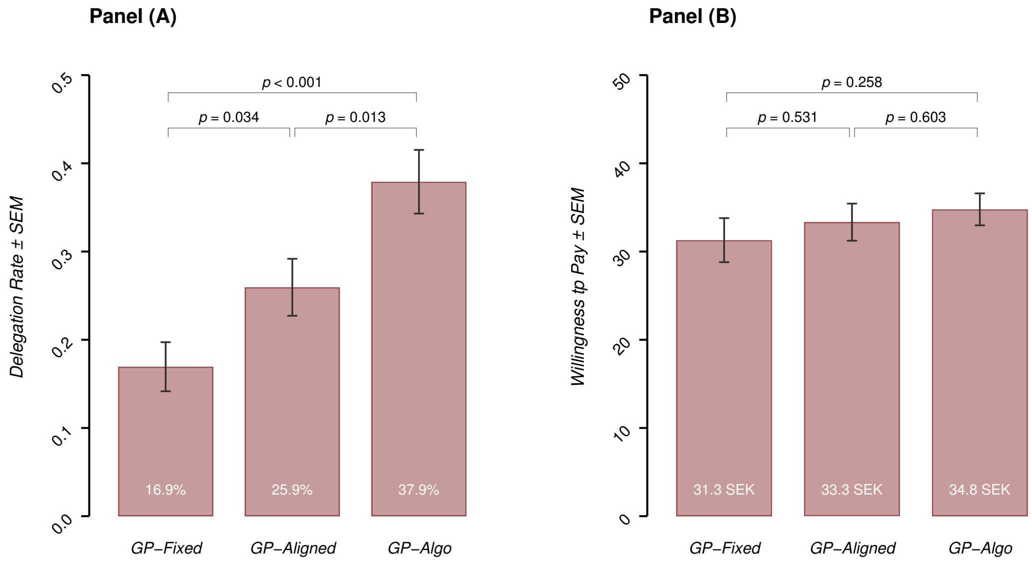


**Figure 2: Delegation frequency.** Panel (A) depicts the share of principals opting for delegating their investment decisions to the agent conditional on the treatment. Error bars indicate standard errors of the mean (*SEM*); $p$-values are based on a logit regression of delegation on treatment indicators (with robust standard errors; $n = 550$). Panel (B) shows the mean willingness to pay for delegating the investment decisions of principals who chose to delegate. Error bars indicate standard errors of the mean (*SEM*); $p$-values are based on an ordinary least squares regression of willingness to pay on treatment indicators (with robust standard errors, $n = 148$).

flat payment vs. linear incentives) – i.e., factors that have been *exogenously* varied in our treatments. In a next step, we investigate whether the variation in *endogenously* varying characteristics of the principals have a systematic effect on the likelihood of delegating financial investment decisions.

**Result 2** *Principals' propensity to delegate their investment decisions increases with trust in agents, both in human agents and investment algorithms. On aggregate, blame shifting motives increase and own decision making quality decrease delegation propensity, but the motives vary across treatments.*

Table 3 reports the estimates from logit regressions of principals' delegation decisions on various experimental and self-reported measures, conditional on the treatment. We find that principals' decision whether to delegate is significantly driven by the principals' trust in the agent (see variable "Trust in Agent"; model 4). Notably, however, we find that the effect of trust varies considerably in terms of the effect size across treatments. The odds of delegating one's decision to a finance professional compensated with a fixed payment (*GP-FIXED*; model 1) are expected to increase by 75.7% (*MEM* $= 0.068$, $p = 0.028$) for a one standard deviation increase in principals' trust in finance professionals. If the agent is a finance professional facing linear incentives (*GP-ALIGNED*; model 2), the odds of delegation, on average, increase by 103.3% (*MEM* $= 0.121$, $p < 0.001$) for a one standard deviation increase in trust. For the treatment in which clients can delegate their decisions to an investment algorithm (*GP-ALGO*; model 3), however, the effect of trust turns out to be largest: a one standard deviation increase in principal's trust in investment algorithms, on average, gives rise to an increase of 170.5% (*MEM* $= 0.174$, $p < 0.001$) in the odds of dele-

**Table 3: Determinants of delegation choices.** This table reports marginal effects estimates from logit regressions of the binary choice whether to delegate the investment decision to the agent on a set of experimental and self-reported measures, conditional on the treatment (models 1–3) and pooled across all treatments (model 4). Robust standard errors are reported in parentheses. * $p < 0.05$, ** $p < 0.005$.

| | (1) *GP-FIXED* | (2) *GP-ALIGNED* | (3) *GP-ALGO* | (4) *Pooled* |
|---|---|---|---|---|
| *Experimental Measures:* | | | | |
| Decision Making Quality Index | $-0.044$ | $-0.006$ | $-0.078$* | $-0.041$* |
| | (0.028) | (0.026) | (0.033) | (0.016) |
| Financial Literacy Score (Std.) | 0.027 | $-0.004$ | 0.003 | 0.003 |
| | (0.055) | (0.062) | (0.058) | (0.034) |
| Numeracy Score (Std.) | $-0.038$ | $-0.030$ | $-0.028$ | $-0.064$ |
| | (0.062) | (0.075) | (0.089) | (0.044) |
| Overestimation (Std.) | $-0.039$ | $-0.017$ | $-0.047$ | $-0.033$ |
| | (0.035) | (0.035) | (0.038) | (0.021) |
| Overplacement (Std.) | 0.003 | $-0.022$ | $-0.015$ | $-0.023$ |
| | (0.039) | (0.042) | (0.052) | (0.026) |
| *Self-Reported Measures:* | | | | |
| Risk Tolerance (Std.) | 0.019 | 0.011 | 0.001 | 0.011 |
| | (0.029) | (0.036) | (0.036) | (0.020) |
| Blame Shifting (Std.) | 0.057* | 0.026 | 0.053 | 0.045* |
| | (0.028) | (0.030) | (0.033) | (0.018) |
| Trust in Agent (Std.) | 0.068* | 0.121** | 0.174** | 0.123** |
| | (0.031) | (0.033) | (0.036) | (0.019) |
| *Controls* | *yes* | *yes* | *yes* | *yes* |
| Wald $\chi^2$ | 16.238 | 18.535 | 42.016 | 54.050 |
| $p > \chi^2$ | 0.236 | 0.138 | 0.000 | 0.000 |
| Pseudo $R^2$ | 0.141 | 0.101 | 0.218 | 0.112 |
| Observations | 183 | 185 | 182 | 550 |

*Notes:* All self-reported measures are standardized scores. "Trust in Agent" refers to a combined variable of trust in finance professionals and financial algorithms, conditional on the treatment. "Blame Shifting" refers to the mean of two standardized survey items on shifting blame on others and resisting the temptation to shift blame on others. "Controls" include gender (binary indicator for female), age (in years), net income from major employment in 2017 (in thousand SEK's), and maximum education level (dichotomous indicators for high school education or less, university education smaller or equal to three years, and university education larger than three years).

To examine whether the effects systematically differ between treatments, we conduct Wald tests on each covariate after a seemingly unrelated regression with robust standard errors in pairwise comparisons of models. Notably, none of the differences are statistically significant with a $p$-value smaller than the 0.05 threshold.

gating the investment decision to the agent (of course, here the variable "Trust in Agent" is composed of answers to the question on trust in investment algorithms).

Apart from trust being a significant driver of delegation decisions, we find that, on aggregate (model 4), blame shifting is significantly positively and clients' decision-making quality is significantly negatively associated with delegation propensity. In particular, we find that the odds of delegating to a finance professional compensated with a fixed payment (*GP-FIXED*), on average, are expected to increase by 60.5% ($me = 0.057, p = 0.041$) for a one standard deviation in blame shifting. While the effect of blame shifting turns out being positive in the other treatments as well, the magnitudes are smaller and the effect is not statistically different from zero (*GP-ALIGNED*: *OR* $= 1.163$, $me = 0.026$, $p = 0.389$; *GP-ALGO*: *OR* $= 1.360$, $me = 0.053$, $p = 0.108$). With respect to the impact of principals' decision-making quality, we report a significant effect for treatment *GP-ALGO*: for a one standard deviation increase in decision-making quality, the odds of delegating to the investment algorithm are expected to decrease by 36.2% (*MEM* $= -0.078$, $p = 0.019$). However, the effect of *DMQI* on the likelihood of delegation turns out being smaller in magnitude and not statistically significant for the other two treatments (*GP-FIXED*: *OR* $= 0.695$, *MEM* $= -0.044$, $p = 0.140$; *GP-ALIGNED*: *OR* $= 0.963$, $me = -0.006$, $p = 0.803$).

Notably, however, neither numeracy skills and financial literacy scores, nor the two measures of overconfidence, nor participants' (self-reported) risk tolerance show any explanatory power with respect to principals' delegation decisions – irrespective of the treatment condition.

Results 1 and 2 identify driving factor of principals' decision whether to delegate to an agent. As a unique feature of our experimental design, the decision to delegate is accompanied with an indication of how much risk the principal wants the agent to take on their account. By this means, our study allows to address a novel and highly relevant aspect of delegated decision-making, summarized by result 3: risk communication.

**Result 3** *Principals, on average, request the agent to take more risk than they perceive to have taken in their own decision. Finance professionals, on average, construct portfolios that come close to principals' desired risk levels. However, portfolios constructed on behalf of clients show considerable overlaps in portfolio risk across risk levels indicated by principals.*

Panel (A) in Figure 3 illustrates the principals' desired levels of risk (x-axis) when delegating their decisions to the agent conditional on the risk perception of their own investment decisions (y-axis). On average, we find that principals tend to ask the agent to take significantly higher levels of risk ($m = 2.84$, $sd = 0.69$) than they perceive they implemented themselves when deciding on their own behalf ($m = 2.58$, $sd = 0.76$; paired-sample $t$-test: $t(147) = 4.081$, $p < 0.001$, $n = 148$).

Comparing the risk levels finance professionals are asked to comply with ($m = 2.45$, $sd = 1.14$, x-axis) with finance professionals' risk perception of the portfolios constructed on behalf of their clients ($m = 2.54$, $sd = 0.97$, y-axis), shows that, on average, finance professionals strive to follow the intentions of potential clients. We find that the average risk perception is slightly higher than the principals' desired risk level (paired-sample $t$-test: $t(270) = 1.994$, $p = 0.0472$; see panel (B) of Figure 3).

Nevertheless, these differences in perceived and desired risk levels translate into risk communication problems. In Figure 4, we show cumulative distributions and boxplots of portfolio risk (i.e., the mean standard deviation of the 25 allocation decisions for each individual in the investment task) conditional on the perceived riskiness of subjects' choices (risk levels 1 to 4). We separate portfolios constructed by the general population (clients) sample (pooled across all treatments) from the sample of finance professionals deciding on behalf of clients (i.e., treatments *FP-FIXED* and *FP-ALIGNED*).

By applying two-sample Kolmogorov-Smirnov tests, we find that the distributions of portfolio risk associated with perceived risk levels differ significantly between the general population (clients) and the finance

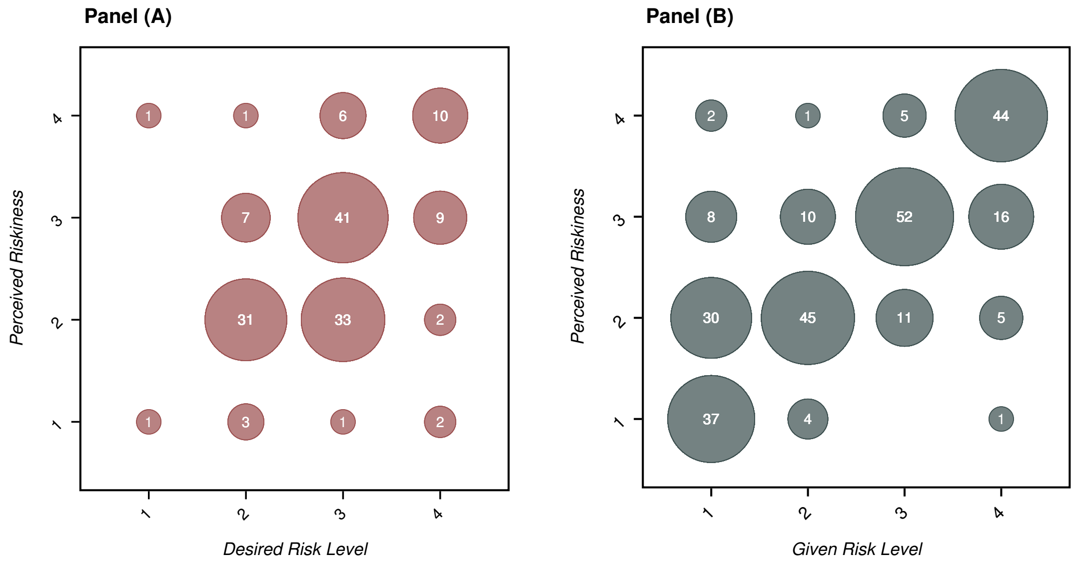


**Figure 3: Risk Communication.** Panel (A) shows principals' perceived riskiness of their own decisions vs. principals' desired risk levels when delegating their investment decisions to the agent ($n = 148$). Panel (B) shows the risk level agents are asked to comply with when deciding on behalf of clients vs. agents' perception of the riskiness of their actual decisions ($n = 271$).

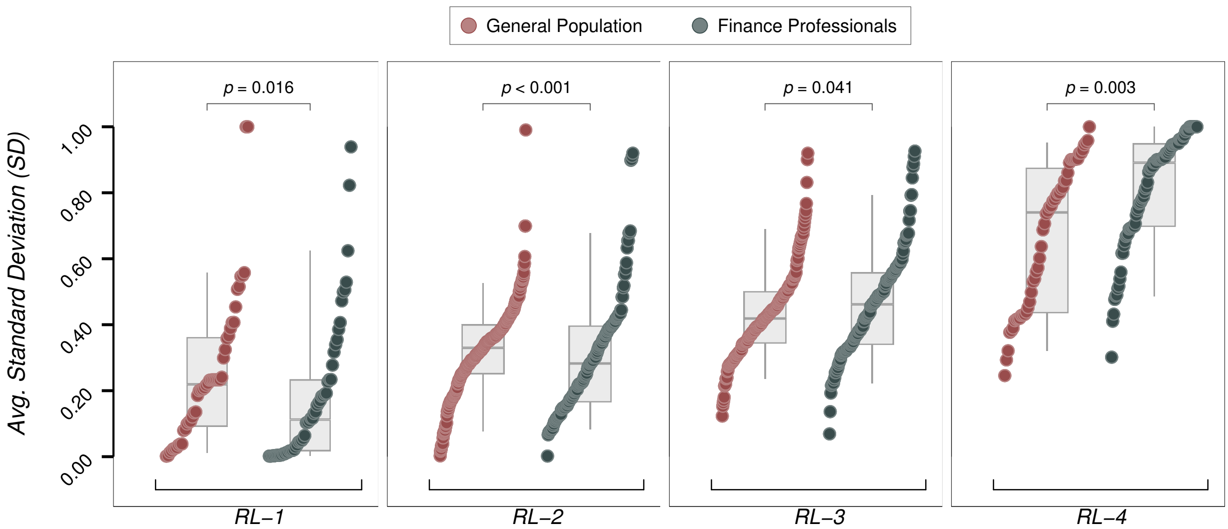


**Figure 4: Portfolio risk conditional on risk perception.** This figure shows cumulative distributions and boxplots of the mean standard deviation of the 25 allocations in the investment task (normalized to 1), conditional on the perceived riskiness of their choices (*RL-1* ... *RL-4*), separated for the general population (clients) sample (pooled across all treatments) and the sample of finance professionals deciding on behalf of clients (i.e., *FP-FIXED* and *FP-ALIGNED*). $p$-values reported for sample comparisons are based on two-sample Kolmogorov-Smirnov tests. Sample sizes for *RL-1* through *RL-4* are $n_1 = 102$, $n_2 = 402$, $n_3 = 318$, and $n_4 = 136$.

professional sample. This finding indicates that both populations perceive risks differently. For risk levels 1 and 2, portfolio risk of clients, on average, significantly exceeds the risk of portfolios implemented by finance professionals deciding on behalf of clients. However, this effect reverses for risk levels 3 and 4:

the risks of portfolios constructed by finance professionals are significantly higher than the risks associated with the allocations of the clients. This result indicates that clients compose more similar portfolios across the risk spectrum than professionals do (see the associated $p$-values of Kolmogorov-Smirnov tests in Figure 4).

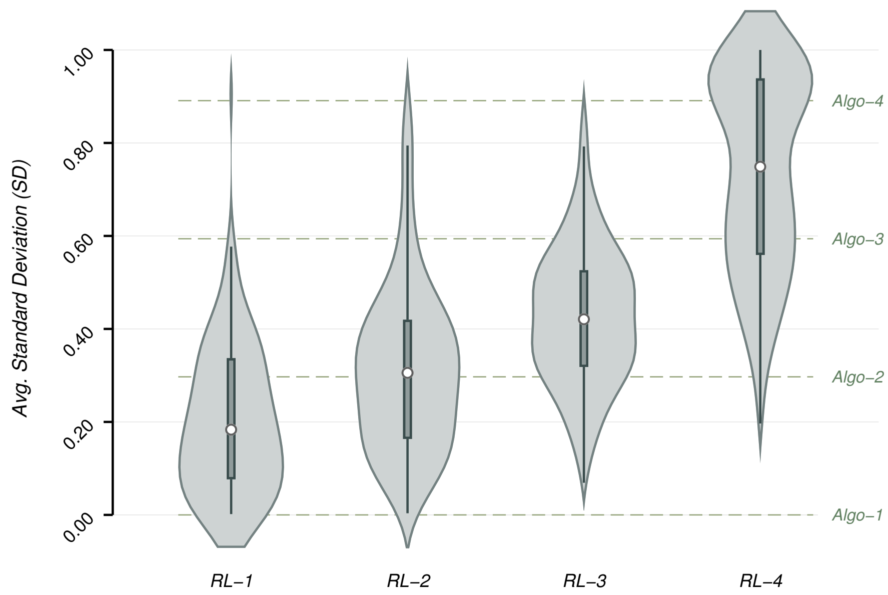


**Figure 5: Portfolio risk conditional on clients' preferred risk level.** This figure shows finance professionals' portfolio risk when deciding on behalf of clients conditional on the risk levels they are asked to comply with. In particular, the figure illustrates the distribution (Gaussian kernel) of the average standard deviation of the 25 allocations in the investment task (normalized to 1), conditional on the risk level (*RL-1* … *RL-4*) indicated by clients for treatments *FP-FIXED* and *FP-ALIGNED*. As a reference, *Algo-1* through *Algo-4* indicate the average standard deviation of portfolios constructed by the investment algorithm for risk levels 1 to 4.

Finally, we investigate the "outcome" of this risk communication problem by examining agents' portfolio risk conditional on delegated (preferred) risk of clients. As illustrated in Figure 5, the distributions of mean portfolio risk (in terms of average standard deviation) vastly overlap for the different risk levels (see also Figure G2 in Appendix G). While the mean portfolio risk increases significantly with the risk level ($p < 0.001$ for all pairwise comparisons; see Table G1 in Appendix G for details), we find that almost the full range of risk might be associated with each level of desired risk. This result implies that clients, indicating different levels of risk when delegating their investment decisions, can eventually end up with similar levels of portfolio risk. For example, 25% of the portfolios constructed by finance professionals for principals indicating risk level 2 / 4 exhibit more risk than 50% of the allocations designed for risk level 3 / 4; even more problematic, about 25% of the portfolios designated for risk level 1 / 4 implies higher risk than 25% of the allocations constructed for risk level 3 / 4.[13]

[13] As indicated in Result 3, principals ask the agents to take more risk than they perceive they took on their own. For those subjects requesting a higher level of risk, delegation, on average, significantly increases portfolio risk – conditional on risk levels (*GP-FIXED*: $d = 0.277$, $t(141) = 3.299$, $p = 0.001$, $n = 3{,}550$; *GP-ALIGNED*: $d = 0.268$, $t(152) = 3.299$, $p = 0.001$, $n = 3{,}825$; *GP-ALGO*: $d = 0.539$, $t(29) = 2.953$, $p = 0.006$, $n = 750$; for comparisons on the subject level, please refer to Figure G1 in Appendix G.

## 4. Conclusion

In this paper, we reported from a novel lab-in-the field (online) experiment with finance professionals serving as money managers and subjects from the general population acting as clients. We examined the motivations and characteristics of clients to delegate investment decisions when facing money managers of different types.

We showed that clients are most likely to delegate to an investment algorithm, followed by professionals with aligned incentives and professionals with fixed incentives. However, we did not find variation in willingness to pay once participants chose to delegate. We also showed that different individual characteristics of clients explained their delegation decisions to various types of agents: on aggregate, principals' propensity to delegate their investment decisions increased with trust in the agent, no matter whether human agent or investment algorithm. Moreover, we reported that blame shifting motives increased and own decision making quality decreased delegation propensity in the full sample. When zooming into the treatments, we showed that blame shifting motives were particularly important for clients' delegation propensity to finance professionals with a fixed payment and that delegation propensity to investment algorithm decreased with client's own decision-making quality. Finally, we observed that principals tend to ask the agent to take more risk compared to their risk perception in their own decisions, which is in line with existing theories (Gennaioli et al., 2015). Moreover, we also found difficulties in communication of desired risk-levels between clients and agents. This resulted in overlaps in the risk of portfolios implemented by professionals over various risk levels, i.e., clients requesting different risk levels ended up with similar portfolios.

Because we chose an experimental setting that allows for thorough control, the design implies several limitations. For this reason, we are careful with generalizing our results. First, other potentially relevant reasons for delegation, such as time constraints or inertia are explicitly ruled out by design, as they are outside the scope of our paper. Particularly, as a first step, we refrained from creating a conflict of interest situation between agents and clients. We leave this topic open for future research, as it might impact clients' delegation decisions and agents' portfolio choices.

Second, our experimental investment task is an abstraction from real-world investment choices and, thus, differs in several aspects: (i) At first sight, there is no option *not* to invest. However, the investment task is designed such that subjects can perfectly hedge all risk, resulting in risk-free allocations. (ii) The investment task is not identical to real-life investment decisions. This is true for almost all finance experiments in the lab and theoretical models for a reason: An abstract investment task with key features of real-world investment situations (e.g., diversification potential, riskiness of alternatives) offers the upside of detaching the decision from a real-world context, thereby allowing to control for subjects' knowledge about real investment products, beliefs of future developments of markets, and reputation. Moreover, with our design, we can measure empirically unobservable variables such as the preference for the type of money manager, trust in the (human or artificial) agent, and motives for blame shifting. Thus, we do not have to rely on indirect proxies, which is a limitation of empirical studies. For the same reasons, our task does not include loss aversion and ambiguity tolerance. (iii) Moreover, we also abstract from introducing tournament components or social comparison – features that finance professionals seem to care about (Kirchler et al., 2020, 2018). In our study, we consciously focus on aligned vs non-aligned incentives as a first step to introduce information about agents' incentives for clients' delegation decisions and trust. We leave variations of aligned incentives like (convex) tournament incentives for future research.

Third, the stake size, compared to salaries of finance professionals, could be considered to be low. However, with the level of monetary stakes applied, we are in line with experimental studies with professionals (Haigh and List, 2005; Alevy et al., 2007; Abdellaoui et al., 2013; Kirchler et al., 2018; Weitzel et al., 2019) and with studies using general population samples (Andersson et al., 2016, 2019). While it is difficult to make a final statement on the level of incentives, the quality of the experimental data is high (in terms of time spent for the experiment and regarding a very low number of outliers in the data), indicating that the subjects took the experiment very seriously.

Despite these limitations, our study has implications for real-world delegation decisions: First, our results highlight the importance of establishing trust in the finance industry in general and in money managers – including investment algorithms – in particular. Beyond clients' decision making quality and shifting blame on others, it appears to be a consistent and major motive for delegating financial decisions.

Second, our results indicate that some clients use delegation as a way of increasing the risk of their portfolio, but the feasibility of this objective seems to be hampered by professionals' troubles to correctly implementing clients' expected portfolio risk-level. The issue of risk communication is particularly relevant for real-world delegation of financial decisions and it also might negatively affect the trust-relationship between clients and agents. Thus, we conclude that a better match of advisers and clients in terms of risk preferences and potentially also regarding risk perception (Holzmeister et al., 2019) might be beneficial both for clients and financial institutions.

## References


Abdellaoui, M., Bleichrodt, H., & Kammoun, H. (2013). Do financial professionals behave according to prospect theory? An experimental study. *Theory and Decision*, *74*(3), 411–429.

Alevy, J. E., Haigh, M. S., & List, J. A. (2007). Information cascades: Evidence from a field experiment with financial market professionals. *Journal of Finance*, *62*(1), 151–180.

Andersson, O., Holm, H. J., Tyran, J.-R., & Wengström, E. (2016). Deciding for others reduces loss aversion. *Management Science*, *62*(1), 29–36.

Andersson, O., Holm, H. J., Tyran, J.-R., & Wengström, E. (2019). Risking other people's money: Experimental evidence on bonus schemes, competition, and altruism. *Scandinavian Journal of Economics, online first*, 1–27.

Banks, J., Carvalho, L., & Perez-Arce, F. (2018). Education, decision-making, and economic rationality. *Review of Economics and Statistics*, *101*(3), 428–441.

Bartling, B. & Fischbacher, U. (2012). Shifting the blame: On delegation and responsibility. *Review of Economic Studies*, *79*(1), 67–87.

Bebchuk, L. & Spamann, H. (2010). Regulating bankers' pay. *Georgetown Law Journal*, *98*(2), 247–287.

Bhattacharya, U., Hackethal, A., Kaesler, S., Loos, B., & Meyer, S. (2012). Is unbiased financial advice to retail investors sufficient? Answers from a large field study. *The Review of Financial Studies*, *25*(4), 975–1032.

Böhm, M., Metzger, D., & Strömberg, P. (2018). Since you are so rich, you must be really smart: Talent and the finance wage premium. *Riksbank Research Paper Series No. 137*.

Bolton, G. & Ockenfels, A. (2010). Betrayal aversion: Evidence from Brazil, China, Oman, Switzerland, Turkey, and the United States: Comment. *American Economic Review*, *100*(1), 628–633.

Calcagno, R. & Monticone, C. (2015). Financial literacy and the demand for financial advice. *Journal of Banking and Finance*, *50*, 363–380.

Chakravarty, S., Harrison, G., Haruvy, E., & Rutström, E. (2011). Are you risk averse over other peoples' money? *Southern Economic Journal*, *77*(4), 901–913.

Chang, T., Solomon, D., & Westerfield, M. (2016). Looking for someone to blame: Delegation, cognitive dissonance, and the disposition effect. *The Journal of Finance*, *71*(1), 267–302.

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

D'Acunto, F., Prabhala, N., & Rossi, A. (2019). The promises and pitfalls of robo-advising. *The Review of Financial Studies*, *32*(5), 1983–2020.

Dewatripont, M. & Freixas, X. (2012). Bank resolution: Lessons from the crisis. In M. Dewatripont & X. Freixas (Eds.), *The crisis aftermath: New regulatory paradigms*. London: Centre for Economic Policy Research.

Diamond, D. W. & Rajan, R. G. (2009). The credit crisis: Conjectures about causes and remedies. *American Economic Review*, *99*(2), 606–610.

Dietvorst, B., Simmons, J., & Massey, C. (2014). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*, *144*(1), 114–126.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., & Wagner, G. (2011). Individual risk attitudes: Measurement, determinants and behavioral consequences. *Journal of the European Economic Association*, *9*(3), 522–550.

Dulleck, U. & Kerschbamer, R. (2006). On doctors, mechanics, and computer specialists: The economics of credence goods. *Journal of Economic Literature*, *44*(1), 5–42.

Echenique, F., Lee, S., & Shum, M. (2011). The Money Pump as a measure of revealed preference violations. *Journal of Political Economy*, *119*(6), 1201–1223.

Edin, P. & Fredriksson, P. (2000). LINDA: Longitudinell individual data for Sweden. *Working Paper.*

Eriksen, K. & Kvaløy, O. (2010). Myopic investment management. *Review of Finance*, *14*(3), 521–542.

Eriksen, K., Kvaløy, O., & Luzuriaga, M. (2017). Risk-taking on behalf of others. *CESifo Working Paper Series No. 6378.*

Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D., & Sunde, U. (2018). Global evidence on economic preferences. *Quarterly Journal of Economics*, *133*(4), 1645–1692.

Falk, A., Becker, A., Dohmen, T., Huffman, D., & Sunde, U. (2016). The Preference Survey Module: A validated instrument for measuring risk, time, and social preferences. *IZA Discussion Paper Series No. 9674.*

Financial Crisis Inquiry Commission. (2011). *The Financial Crisis Inquiry Report: Final report of the National Commission on the causes of the financial and economic crisis in the United States.* Washington, DC: U.S. Government Printing Office.

Foerster, S., Linnaimaa, J. T., Melzer, B. T., & Previtero, A. (2017). Retail financial advice: Does one size fit all? *Journal of Finance*, *72*(4), 1441–1482.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, *19*(4), 25–42.

Füllbrunn, S. & Luhan, W. (2015). Am I my peer's keeper? Social responsibility in financial decision making. *Ruhr Economic Papers #551.*

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

Germann, M. & Merkle, C. (2019). Algorithm Aversion in Delegated Investing. *Working Paper.*

Guiso, L., Sapienza, P., & Zingales, L. (2004). The role of social capital in financial development. *American Economic Review*, *94*(3), 526–556.

Guiso, L., Sapienza, P., & Zingales, L. (2008). Trusting the stock market. *The Journal of Finance*, *63*(6), 2557–2600.

Hadar, J. & Russell, W. (1969). Rules for ordering uncertain prospects. *American Economic Review*, *59*(1), 25–34.

Haigh, M. S. & List, J. A. (2005). Do professional traders exhibit myopic loss aversion? An experimental analysis. *Journal of Finance*, *60*(1), 523–534.

Harvey, C., Rattray, S., Sinclair, A., & Van Hemert, O. (2017). Man vs. machine: Comparing discretionary and systematic hedge fund performance. *Journal of Portfolio Management*, *43*(4), 55–69.

Holzmeister, F., Huber, J., Kirchler, M., Lindner, F., Weitzel, U., & Zeisberger, S. (2019). What drives risk perception? A global survey with financial professionals and lay people. *Management Science*, *forthcoming*.

Ifcher, J. & Zarghamee, H. (2019). Behavioral economic phenomena in decision-making for others. *Journal of Economic Psychology*, *forthcoming*.

Inderst, R. & Ottaviani, M. (2012a). Competition through commissions and kickbacks. *American Economic Review*, *102*(2), 780–809.

Inderst, R. & Ottaviani, M. (2012b). Financial advice. *Journal of Economic Literature*, *50*(2), 494–512.

Inderst, R. & Ottaviani, M. (2012c). How (not) to pay for advice: A framework for consumer financial protection. *Journal of Financial Economics*, *105*(2), 393–411.

Kirchler, M., Lindner, F., & Weitzel, U. (2018). Rankings and risk-taking in the finance industry. *Journal of Finance*, *73*(5), 2271–2302.

Kirchler, M., Lindner, F., & Weitzel, U. (2020). Delegated investment decisions and rankings. *Journal of Banking and Finance*, *forthcoming*.

Kling, L., König-Kersting, C., & Trautmann, S. T. (2019). Investment preferences and risk perception: Financial agents versus clients. *AWI Discussion Paper Series No. 674*.

Lachance, M.-E. & Tang, N. (2012). Financial advice and trust. *Financial Services Review*, *21*(3), 209–226.

Linnainmaa, J. T., Melzer, B., & Previtero, A. (2019). The misguided beliefs of financial advisors. *The Journal of Finance*, *forthcoming*.

Logg, J., Minson, J., & Moore, D. (2017). Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes*, *151*, 90–103.

Longoni, C., Bonezzi, A., & Morewedge, C. K. (2019). Resistance to medical artificial intelligence. *Journal of Consumer Research*, *46*(4), 629–650.

Mullainathan, S., Noeth, M., & Schoar, A. (2012). The market for financial advice: An audit study. *NBER Working Paper No. 17929*.

Rajan, R. G. (2006). Has finance made the world riskier? *European Financial Management*, *12*(4), 499–533.

Rossi, A. G. & Utkus, S. P. (2020). Who benefits from robo-advising? Evidence from machine learning. *Working Paper*.

Rud, O. A., Rabanal, J. P., & Horowitz, J. (2018). Does competition aggravate moral hazard? A multi-principal-agent experiment. *Journal of Financial Intermediation*, *33*, 115–121.

Sapienza, P. & Zingales, L. (2012). A trust crisis. *International Review of Finance*, *12*(2), 123–131.

Shefrin, H. (2007). *Beyond greed and fear: Understanding behavioral finance and the psychology of investing*. Oxford: Oxford University Press.

Sutter, M. (2009). Individual behavior and group membership: Comment. *American Economic Review*, *99*(5), 2247–2257.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2014). Assessing miserly information processing: An expansion of the Cognitive Reflection Test. *Thinking and Reasoning*, *20*(2), 147–168.

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

Vieider, F. M., Villegas-Palacio, C., Martinsson, P., & Mejía, M. (2016). Risk taking for oneself and others: A structural model approach. *Economic Inquiry*, *54*(2), 879–894.

Weitzel, U., Huber, C., Huber, J., Kirchler, M., Lindner, F., & Rose, J. (2019). Bubbles and financial professionals. *Review of Financial Studies*, *forthcoming*.

Weller, J. A., Dieckmann, N. F., Tusler, M., Mertz, C. K., Burns, W. J., & Peters, E. (2013). Development and testing of an abbreviated numeracy scale: A Rasch analysis approach. *Journal of Behavioral Decision Making*, *26*(2), 198–212.

Wilson, G., Gray, J., & Barrett, P. (1990). A factor analysis of the Gray-Wilson personality questionnaire. *Personality and Individual Differences*, *11*(10), 1037–1044.

Zingales, L. (2015). Does finance benefit society? *Journal of Finance*, *70*(4), 1327–1363.

# *Appendices*

*for Online Publication*

## Delegated Decision-Making in Finance

### Contents

## List of Tables



## List of Figures

# A. Data Collection and Recruitment

**Experimental software.** The experimental software—computerized in *oTree* (Chen et al., 2016)— which includes all instructions, treatment variations, as well as the Swedish/English translations has been pre-registered at https://osf.io/ubpr3/. Demo versions of the experiment and all treatments (in English) are available via http://hea-2019-01-en.herokuapp.com. The source code of the experimental software is available at https://osf.io/tfeh5/. Figure A1 shows two screenshots of the main experimental task, i.e., the allocation decision task, with two and five assets, respectively.

**Data availability.** All raw data generated in the online experiments is available at https://osf.io/quxmd/. Moreover, the OSF repository contains all script files used to generate the results presented in the paper and the appendices, together with the processed data files, the figures, and tables. Please note that the register data obtained from *Statistiska centralbyrån* (Statistics Sweden; *SCB*) may not be publicly shared.

**Recruitment.** *Statistiska centralbyrån* (Statistics Sweden; *SCB*) sent out hard copy invitations to participate in the anonymous online experiment. The receivers of the invitations logged in to our experiment using a personalized participant code, which was linked to a key only known to *SCB*. The participant code indicated whether a particular subject was recruited from the finance professional pool or the general population pool. After the data collection has been completed, we sent the identifiers of those participants who completed the experiment to *SCB*, who used their keys to match the experimental data with the requested register data (which is described in detail below). Participants were informed that the data gathered in the experiment is matched with their register data in the invitation letters and on the first screen of the experiment.

**Payments.** To ensure full privacy of the data collected during the experiment, payouts were handled by the third party survey firm *Enkätfabriken*. Once participants completed the online experiment, they were redirected to a dedicated form on the website of *Enkätfabriken*. Participants used the same participant code as in the experiment. For payment purposes, *Enkätfabriken* collected participants' names, email addresses, "personnummer" (personal identity number), and bank account details. The information collected was handled only by *Enkätfabriken* and has been used exclusively for sake of ordering the bank remittances.

**Registry data.** In addition to the data collected in the online experiment, we obtained the following register data from *Statistiska centralbyrån* (Statistics Sweden; *SCB*) for each participant who completed all tasks in the experiment:

- *Demographics*: year born, age, gender, county, municipality, and assembly of residence, marital status, year in marital status, family status, birth country, children living at home age 0–3, 4–6, 7–10, 11–15, 16–17, $\geq 18$, highest finished education level, education orientation, education group, education county, graduation year, primary source of income, work place municipality and county, work place industry 1990–1992, 1993–2001, 2002–2010, and 2007–2014, occupation 2002–2013 and 2014, net income of own business 1991–2003, 2003–2014, and 2004–2014, capital income, disposable income 1990–2004 and 2004–2014, disposable income of family 1990–2004 and 2004–2014, country of birth, date of immigration.

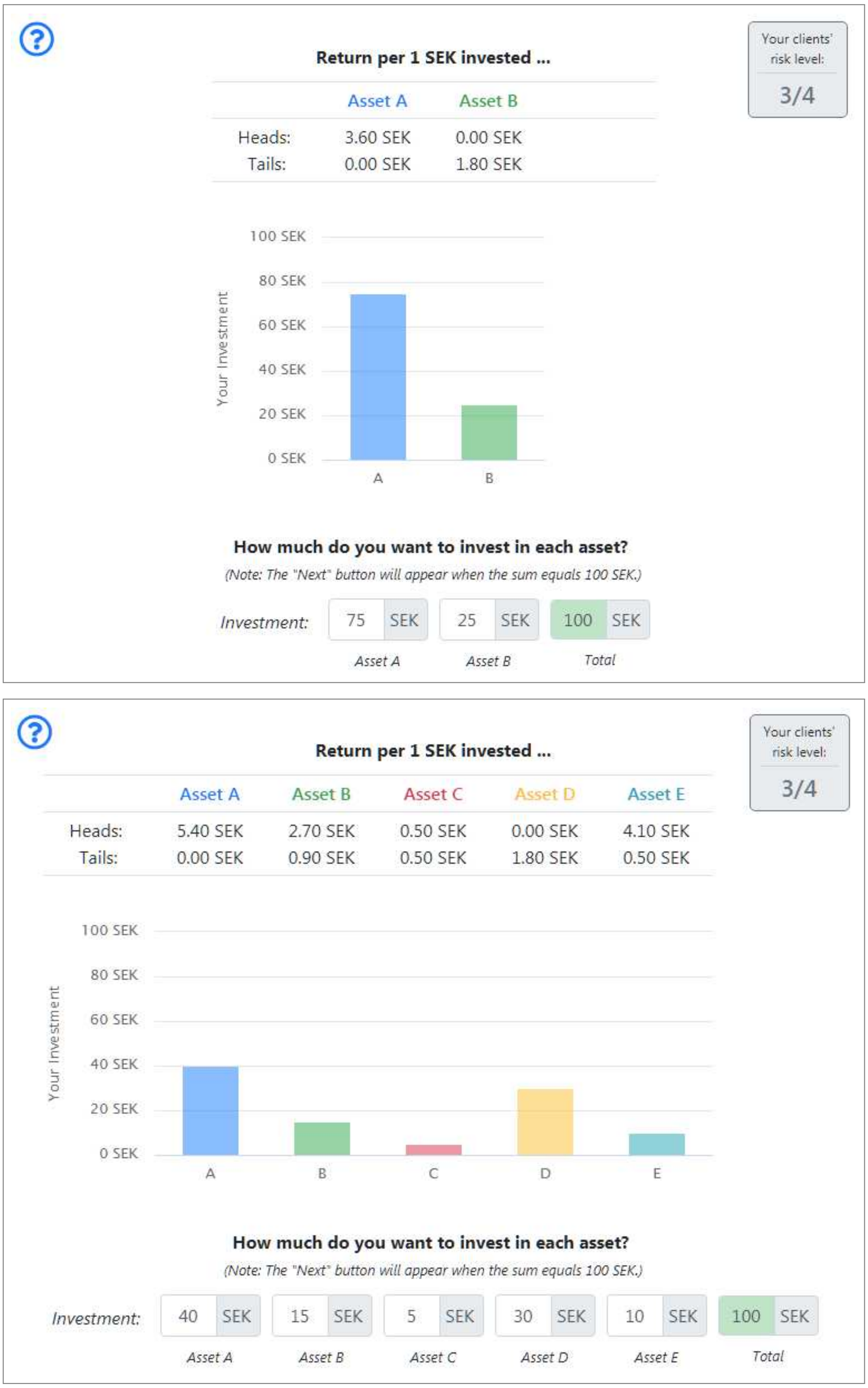


**Figure A1: Screenshots of the allocation decision task.** The figure shows screenshots of the main experimental task as displayed to subjects. Note that the information in the top right corner ("Your clients' risk level") was only displayed to finance professionals in the treatments *FP-FIXED* and *FP-ALIGNED*. By clicking on the question mark icon in the top left corner, participants had the opportunity to reread the instructions at any time. The button to proceed to the next decision was only shown if investments to the available assets summed up to 100.

- *Firm / workplace*: number of employees at firm / workplace, number of men / women at firm / workplace, number of men / women with short / long education at firm / workplace, total salaries paid by firm / workplace.
- *Education*: high school, high school program, high school grades point average, high school graduation year, university, university program, university major, university graduation year.
- *Assets*: net wealth, total debt, bank account, listed equity, fixed income funds, other funds, bonds and other securities, taxable insurances, houses, apartments, holiday homes.
- *Military records*: command suitability, non cognitive abilities score, muscle strength, physical capacity for work, length, weight, cognitive scores 1 and 2 in language and logic, one in spatial understanding, and one in technical understanding.
- *Parents*: adoptive / biological mother / father, occupation mother / father, primary income source mother / father, net income from own business mother / father, net wealth mother / father.

We only use a part of the available registry data as control variables in our analyses of observed behavior, in particular, participants' gender (binary indicator for female), age (in years), net income from major employment in 2017 (in thousand SEK's), and maximum education level (dichotomous indicators for high school education or less, university education smaller or equal to three years, and university education larger than three years). The restricted use of the register data has been pre-registered at the outset (see https://osf.io/ubpr3/ for the pre-registration).

After the experiment reported in this paper, participants were invited to a second, independent experiment for which the obtained registry data plays a more pivotal role. For details about the second experiment, please refer to the respective pre-registration at https://osf.io/6rdp8/.

## B. Allocation Decision Task

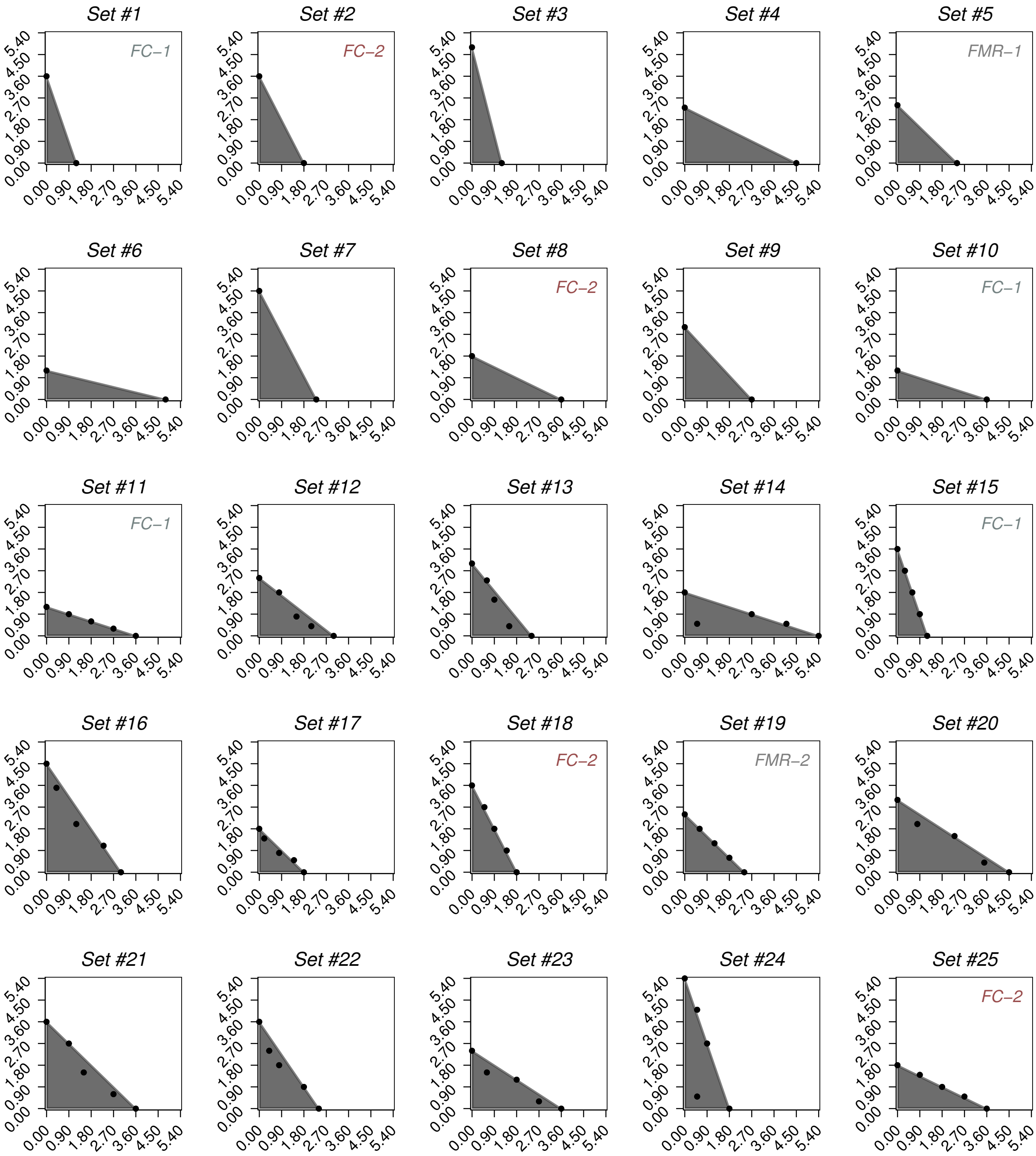


**Figure B1: Opportunity sets in the allocation decision task.** In each panel of this figure, the vertical (horizontal) axis indicates the return per 1 SEK invested if the coin shows up heads (tails). Each dot indicates a single asset. The labels *FC-1*, *FC-2*, *FMR-1*, and *FMR-2* denote particular opportunity sets used for constructing the decision-making quality measures "financial competence" (*FC*) and "failure to minimize risk" (*FMR*).

## C. Decision-Making Quality Measures

In each opportunity set $j \in \{1, 2, \ldots, 25\}$, each participant $i$ is endowed with 100 SEK to allocate on assets $k \in \{1, 2, \ldots, 5\}$. Let $a_{i,j,k}$ denote the fraction of the endowment allocated on asset $k$ such that $\sum_k a_{i,j,k} = 1$.

The return per SEK invested in asset $k$ if the coin comes up heads is denoted as $h_{j,k}$; the return per SEK invested if it comes up tails is denoted as $t_{j,k}$. Thus, the return of participant $i$'s allocation in opportunity set $j$ will either be

$$\begin{aligned} H_{i,j} &= \sum_k a_{i,j,k} \cdot h_{j,k} && \text{if the coin comes up heads, or} \\ T_{i,j} &= \sum_k a_{i,j,k} \cdot t_{j,k} && \text{if the coin comes up tails.} \end{aligned}$$

Let the tuple $\boldsymbol{x}_{i,j} = (H_{i,j}, T_{i,j})$ denote the portfolio of participant $i$ in opportunity set $j$. In addition to the measures of expected return and standard deviation, following Banks et al. (2018) we also define four measures of decision-making quality: (i) violations of first order stochastic dominance (*FOSD*), (ii) violations of the generalized axiom of revealed preferences (*GARP*), (iii) financial competence (*FC*), and (iv) failure to minimize risk (*FMR*). Each of these measures is defined in detail below.

**Expected Return.** The expected portfolio return of participant $i$'s investment in opportunity set $j$, i.e., the expected return from allocating the endowment on the available assets, is given by

$$ER_{i,j} = \frac{H_{i,j} + T_{i,j}}{2}.$$

Participant $i$'s mean expected return, $ER_i$, is calculated as the average of $ER_{i,j}$ across 23 of the 25 opportunity sets, as the expected returns are identical for all portfolios in the two remaining opportunity sets (set #5 and #19; see Table 1 and Figure B1), i.e., $ER_i = 1/23 \cdot \sum_{j=1}^{23} ER_{i,j}$.

**Standard Deviation.** As a measure of portfolio risk, we calculate the standard deviation of participant $i$'s portfolio in opportunity set $j$, i.e., the standard deviation of $H_{i,j}$ and $T_{i,j}$ occurring with a probability of 50% each:

$$SD_{i,j} = \sqrt{\frac{H_{i,j}^2 + T_{i,j}^2}{2} - \left(\frac{H_{i,j} + T_{i,j}}{2}\right)^2}.$$

The average portfolio risk for individual $i$, $SD_i$, is defined as the mean standard deviation across all 25 opportunity sets, i.e., $SD_i = 1/25 \cdot \sum_{j=1}^{25} SD_{i,j}$.

**Violations of First Order Stochastic Dominance (*FOSD*).** Following Banks et al. (2018), we use the difference between the maximum expected return of a portfolio that provides the same minimum payoff as the chosen portfolio and the expected return of the chosen portfolio as a measure of how closely participant $i$'s choice in opportunity set $j$ complies with the principle of *FOSD* (Hadar and Russell, 1969).

Given a chosen portfolio $\boldsymbol{x}_{i,j} = (H_{i,j}, T_{i,j})$, let $h_j^* = \max_k h_{j,k}$ be the maximum return across all assets $k$ if the coin comes up heads and $t_j^* = \max_k t_{j,k}$ if the the coin comes up tails. By investing the fraction

$$w = \frac{\min(H_{i,j}, T_{i,j})}{\min(h_j^*, t_j^*)}$$

on the asset paying $\min(h_j^*, t_j^*)$ and 0 SEK otherwise, and investing the fraction $(1 - w)$ on the asset paying $\max(h_j^*, t_j^*)$ and 0 SEK otherwise, participant $i$ maximizes the expected return but still guarantees a minimum return of $\min(H_{i,j}, T_{i,j})$. Thus, our measure of *FOSD* is:

$$\mathit{FOSD}_{i,j} = \left( w \cdot \frac{\min(h_j^*, t_j^*)}{2} + (1 - w) \cdot \frac{\max(h_j^*, t_j^*)}{2} \right) - \frac{(H_{i,j} + T_{i,j})}{2}.$$

To assess participant $i$'s average violations of *FOSD*, we average the measure over all choices, except for the two opportunity sets for which any portfolio will yield the same expected returns (set #5 and #19; see Table 1 and Figure B1), i.e., $\mathit{FOSD}_i = 1/23 \cdot \sum_{j=1}^{23} \mathit{FOSD}_{i,j}$.

**Violations of the General Axiom of Revealed Preferences (*GARP*).** According to the Generalized Axiom of Revealed Preferences, for any two opportunity sets $m$ and $n$ ($m \neq n$), if participant $i$ reveals to prefer $\boldsymbol{x}_{i,m}$ over $\boldsymbol{x}_{i,n}$, then $\boldsymbol{x}_{i,n}$ is not strictly preferred to $\boldsymbol{x}_{i,m}$. An instance of a *GARP* violation occurs when a participant $i$ chooses $\boldsymbol{x}_{i,m}$ in opportunity set $m$ when $\boldsymbol{x}_{i,n}$ is affordable, and also chooses $\boldsymbol{x}_{i,n}$ in opportunity set $n$ when $\boldsymbol{x}_{i,m}$ is affordable.

Let $p_j$ denote the ratio of maximum returns for heads and tails in opportunity set $j$, respectively, i.e., $p_j = h_j^*/t_j^*$. The extent of violations of *GARP* is measured with the Money Pump Index (*MPI*), which is based on the idea that an arbitrageur can exploit violations in revealed preferences (Echenique et al., 2011): The arbitrageur could make profit by buying portfolio $\boldsymbol{x}_{i,m}$ at price $p_n$ and then selling it at price $p_m$; likewise, the arbitrageur could buy portfolio $\boldsymbol{x}_{i,n}$ at price $p_m$ and sell it at price $p_n$. The Money Pump Index is the total profit the arbitrageur could make, i.e.,

$$\begin{aligned} \mathit{MPI}_{i,m,n} &= \alpha_{i,m,n} + \beta_{i,m,n} \\ \mathit{MPI}_{i,m,n} &= p_m \cdot (\boldsymbol{x}_{i,m} - \boldsymbol{x}_{i,n}) + p_n \cdot (\boldsymbol{x}_{i,n} - \boldsymbol{x}_{i,m}). \end{aligned}$$

We calculate the money pump for each violation of *GARP*, i.e., for $25 \cdot (25 - 1) \cdot 1/2 = 300$ pairwise combinations of opportunity sets. For each participant $i$, we determine the average money pump index, over all pairwise combinations, i.e., $\mathit{MPI}_i = 1/300 \cdot \sum_{m=1}^{25} \sum_{n=1}^{25} \mathit{MPI}_{i,m,n} \; \forall \; m > n$.

**Financial competence (*FC*).** Four opportunity sets were identical in the two-asset- and the five-asset-frame, i.e., four sets were presented in both the two-asset- (sets #1, #2, #8, and #10) and the five-asset-frame (sets #11, #15, #18, and #25). Moreover, two of the four opportunity sets presented in the two-asset- and five-asset-frame, respectively, were constructed as mirror images of one another, i.e., only the payoffs for heads and tails were interchanged. Thus, two opportunity sets (denoted as $\mathit{FC}_1$ and $\mathit{FC}_2$ in Figure B1) were effectively presented four times each (#1 = #10 = #11 = #15 and #2 = #8 = #18 = #25).

Let $J_1 = \{\#1, \#10, \#11, \#15\}$ and $J_2 = \{\#2, \#8, \#18, \#25\}$. The financial competence of individual $i$ is defined as the average absolute differences between the expected returns across the identical opportunity sets in $J_1$ and $J_2$, i.e.,

$$FC_i = \frac{1}{12} \cdot \left( \sum_{k,l \in J_1} |ER_{i,k} - ER_{i,l}| + \sum_{m,n \in J_2} |ER_{i,m} - ER_{i,n}| \right) \quad \forall\, k > l \;\; \& \;\; m > n.$$

Note that our definition of $FC_i$ differs from the measure used by Banks et al. (2018), who average the absolute differences in expected returns across the two frames, but not across the mirrored versions of the sets.

**Failure to minimize risk (*FMR*).** In two opportunity sets (#5 and #19; see Figure B1), the expected return per 1 SEK invested was the same for all assets $k$, such that all feasible portfolios will share the same expected return. Choosing a fully-hedged portfolio (i.e., a zero-risk portfolio), thus, (second-order) dominates all other feasible portfolios in these two opportunity sets. The failure to minimize risk for subject $i$ in opportunity set $j$, $FMR_{i,j}$, is measured as the standard deviation $SD_{i,j}$ of the particular portfolio allocation, which is then averaged over the two opportunity sets, i.e.,

$$FMR_i = \frac{1}{2} \cdot \sum_{j=1}^{2} SD_{i,j}.$$

**Decision-making quality index (*DMQI*).** We utilize the predicted values of a principal component analysis of the four measures to constitute the *DMQI* index for each participant (please note that this approach differs from the one in Banks et al., 2018). The predictions of the principal component analysis serve as a unified decision-making quality index, denoted as *DMQI*, on the participant level. Note that, in theory, the predicted values have a mean of zero and a standard deviation of unity. Thus, positive values can be interpreted as above average while negative values indicate that a participants' decision-making quality is below average.

# D. Questionnaires and Side Tasks

After the main experiment, participants were asked to answer a set of Likert items—all scaled from 0 (minimum) to 10 (maximum)—which are summarized in Table D1 below. The questions on risk tolerance and patience are based on Dohmen et al. (2011) and Falk et al. (2016, 2018); and the two statements addressing the proneness to shift blame are based on the inventory introduced by Wilson et al. (1990). In addition to the survey items reported in Table D1, participants were exposed to a hypothetical charitable giving setting based on Falk et al. (2018), asking how much they would donate to a good cause if they had unexpectedly received 10,000 SEK.[14] The 14 items were displayed on five separate screens: the first screen contained the questions regarding risk preferences in general and their willingness to give up something today in order to benefit more in the future; the second screen included the item of risk tolerance in financial matters as well as the three statements on trust; the third screen showed the hypothetical charitable giving task; the fourth screen comprised the two questions on blame shifting; and the fifth screen involved the five questions related to financial investments and the use of expertise.

Once the above questionnaires had been completed, participants answered eight questions allowing to determine their numeracy skills. As a measure of subjects' numeracy, we use the number of correct answers. The numeracy task is based on the Rasch-validated inventory proposed by Weller et al. (2013). Two of the eight questions in the original set are well-known items from the Cognitive Reflection Test (CRT) introduced by Frederick (2005). Since this three-item test has been widely spread on the Internet, many people will know the questions and the corresponding answers. Therefore, the two items on cognitive reflection skills have been replaced by items from the CRT proposed by Toplak et al. (2014). For answering the eight questions, participants faced a time constraint of four minutes. Since the items, by construction of the test, differ considerably in difficulty, the order of the questions has been randomized to avoid systematic effects arising from the time constraint. The questions used in the numeracy task are listed in Table D2.

After submitting their answers to the numeracy questions, participants were asked to self-assess their performance in the task in two different ways. The respective questions read as follows: (i) "How many of the eight questions you answered on the previous screen did you answer correctly?" (0 to 8), and (ii) "Compared to a random sample of the Swedish population, how did you score in terms of correct answers? Please estimate your position in the ranking." (Top 10%, Top 20%, …, Bottom 20%, Bottom 10%). While the first question allows for determining participants' overestimation of their own skills (as the difference between their estimates and actual performance), the second question allows for quantifying subjects' tendency to "overplace" their performance relative to others. Question (ii) asks participants to evaluate their performance relative to a *random* sample of the Swedish population. However, our sample is not representative with respect to the level of education due to self-selection effects. For this reason we take a detour to derive a sensible measure of overplacement: The validated inventory proposed by Weller et al. (2013) is constructed in such a way that scores are approximately normally distributed among a general population sample. The fact that the numeracy scores in our general population sample are significantly different from a normal distribution (Shapiro-Wilk-Test; $W = 0.987$, $p < 0.001$, $n = 550$) somewhat confirms our conjecture of a self-selection effect in our sample. Thus, in a first step, we draw random integers from a normal distribution with a mean of 4.07 and a standard deviation of 1.83, the first and second moment reported for Study 2 in Weller et al. (2013), validating their Rasch-based measure. In a second step, we determine the percentiles associated with each possible score between 0 and 8. Finally,

[14] The question was presented to participants as follows: "Imagine the following situation: Today you unexpectedly received 10,000 SEK. How much of this amount would you donate to a good cause?"

we deduct the percentile (from the sampled normal distribution) corresponding to participants' numeracy score from their estimated decile, i.e., their answer to question (ii), to assess the degree of participants' overplacement.[15] As a final task of the experiment, participants were asked to answer six single-choice questions based on van Rooij et al. (2011), allowing to determine their financial literacy. In particular, three of the questions stem from their basic literacy inventory (*Q1–Q3*, corresponding to (2), (3), and (5) in van Rooij et al. (2011)), and three questions are based on the advanced literacy inventory (*Q4–Q6*, corresponding to (12), (16), and (7) in van Rooij et al. (2011)). As an index of financial literacy, we use the sum of participants' correct answers. The questions used in the financial literacy task are depicted in Table D3.

Descriptive results relating to the questionnaires are provided in Table E2 ; summary results of the side experiments on numeracy skills, financial literacy, and the two measures of overconfidence are provided in Table E3 in Appendix E.

[15] As we ask participants to estimate their performance relative to the general population in deciles rather than percentiles, we use the *minimum* difference to either of the bounds of the interval they implicitly provide as our measure of overestimation. That is, if the percentile (from the sampled normal distribution) lies within the interval subjects estimate, the measure takes value 0; if the percentile is smaller than the lower bound (upper bound) of the estimated interval, we evaluate the percentile to the lower bound (upper bound) of the interval.

**Table D1: Survey questions.** This table summarizes the Likert items, all participants answered after the main experimental task. In particular, the table depicts the variable description as referred to in the main text, the wording of the question/statement, and the corresponding labelling of the minimum and maximum values for each item. The three items indicated with an asterisk were only displayed if the question "Frequent Investments" was not answered with 0 ("does not describe me at all").

| | | ***Likert Scale*** | |
|---|---|---|---|
| ***Variable*** | ***Question / Statement*** | *min (0)* | *max (10)* |
| *Risk Tolerance (in General)* | Are you generally a person who is willing to take risks or do you try to avoid taking risks? | not at all willing to take risks | very willing to take risks |
| *Patience (in General)* | How willing are you to give up something that is beneficial for you today in order to benefit more from that in the future? | not at all willing to give up something today | very willing to give up something today |
| *Risk Tolerance in Financial Matters* | I am generally willing to take risks in financial matters. | does not describe me at all | describes me perfectly |
| *Trust (in General)* | I generally trust other people. | does not describe me at all | describes me perfectly |
| *Trust in Finance Professionals* | I generally trust employees from the finance industry. | does not describe me at all | describes me perfectly |
| *Trust in Investment Algorithms* | I generally trust robo-advisors (i.e. computer programs) in financial matters. | does not describe me at all | describes me perfectly |
| *Blame Shifting (Others)* | If you hurt yourself accidentally, do you sometimes blame somebody who happens to be nearby even though you realize, on reflection, that they were not responsible? | I never blame others | I often blame others |
| *Blame Shifting (Temptation)* | Can you easily resist the temptation to blame others for the accidents that happen to you? | I can resist easily | I cannot resist at all |
| *Frequent Investments* | I frequently invest in stocks and mutual funds myself (not through the national pension system). | does not describe me at all | describes me perfectly |
| *Delegate to Fin. Profs.** | I delegate my investment decisions (e.g., purchase of stocks, bonds, investment funds, real estate) to financial advisors at banks or other institutions and refrain from taking decisions myself. | does not describe me at all | describes me perfectly |
| *Delegate to Inv. Algos.** | I delegate my investment decisions (e.g., purchase of stocks, bonds, investment funds, real estate) to robo-advisors at banks or other institutions and refrain from taking decisions myself. | does not describe me at all | describes me perfectly |
| *Use Expertise of Fin. Profs.** | I use the expertise of financial advisers for my investments/pension savings. | does not describe me at all | describes me perfectly |
| *Responsibility in Financial Matters* | I am solely responsible for financial decisions in my household. | does not describe me at all | describes me perfectly |

**Table D2: Numeracy inventory based on Weller et al. (2013).** This table summarizes the questions used to assess participants' numeracy and the correct answers to each of the questions. For answering all items, participants were given a maximum of four minutes. The inventory proposed by Weller et al. (2013) includes two questions from Frederick (2005). As these are likely to be known by many people, items *Q2* and *Q3* have been replaced by questions from Toplak et al. (2014).

| ***ID*** | ***Question*** | ***Correct Answer*** |
|---|---|---|
| *Q1* | Suppose you have a close friend who has a lump in her breast and must have a mammogram. Of 100 women like her, 10 of them actually have a malignant tumor and 90 of them do not. Of the 10 women who actually have a tumor, the mammogram indicates correctly that 9 of them have a tumor and indicates incorrectly that 1 of them does not. Of the 90 women who do not have a tumor, the mammogram indicates correctly that 81 of them do not have a tumor and indicates incorrectly that 9 of them do have a tumor. Imagine that your friend tests positive (as if she had a tumor), what is the likelihood that she actually has a tumor? | 50 percent |
| *Q2* | 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? | 4 days |
| *Q3* | A man buys a pig for 600 SEK, sells it for 700 SEK, buys it back for 800 SEK, and sells it finally for 900 SEK. How much has he made? | 200 SEK |
| *Q4* | In a lottery, the chance of winning a car is 1 in 1000. What percent of lottery tickets win a car? | 0.1 percent |
| *Q5* | In a lottery, the chances of winning a 10.000 SEK prize are 1%. What is your best guess about how many people would win a 10.000 SEK prize if 1000 people each buy a single lottery ticket? | 10 people |
| *Q6* | Imagine that we roll a fair, six-sided die 1000 times. Out of 1000 rolls, how many times do you think the die would come up as an even number? | 500 times |
| *Q7* | If the chance of getting a disease is 20 out of 100, this would be the same as having a ... chance of getting the disease. | 20 percent |
| *Q8* | If the chance of getting a disease is 10%, how many people would be expected to get the disease out of 1000? | 100 people |

**Table D3: Financial literacy inventory based on van Rooij et al. (2011).** This table summarizes the questions used to assess participants' literacy in financial matters and the corresponding choice options to each of the questions. Correct answers are highlighted in *italics*. For answering all items, participants were given a maximum of three minutes.

| *ID* | *Question* | *Choices* |
|---|---|---|
| *Q1* | Suppose you had 1,000 SEK in a savings account and the interest rate is 20% per year and you never withdraw money or interest payments. After 5 years, how much would you have on this account in total? | ○ *more than 2,000 SEK*<br>○ exactly 2,000 SEK<br>○ less than 2,000 SEK<br>○ do not know |
| *Q2* | Imagine that the interest rate on your savings account was 1% per year and inflation was 2% per year. After 1 year, how much would you be able to buy with the money in this account? | ○ more than today<br>○ exactly the same<br>○ *less than today*<br>○ do not know |
| *Q3* | Suppose that in the year 2025, your income after tax has doubled and prices of all goods have doubled too. In 2025, how much will you be able to buy with your income? | ○ more than today<br>○ *exactly the same*<br>○ less than today<br>○ do not know |
| *Q4* | When an investor spreads his money among different assets, does the risk of losing money in general: | ○ increase<br>○ *decrease*<br>○ stay the same<br>○ do not know |
| *Q5* | If the interest rate falls, what should happen to bond prices? | ○ *rise*<br>○ fall<br>○ stay the same<br>○ none of the above<br>○ do not know |
| *Q6* | Which of the following statements is correct? If somebody buys the stock of firm B in the stock market: | ○ *he owns a part of firm B*<br>○ he has lent money to firm B<br>○ he is liable for the firm B's debt<br>○ none of the above<br>○ do not know |

## E. Descriptive Results

In the following, we present a set of descriptive results for all measures elicited in the experiment. Many of these variables only enter our analyses as controls. Yet, while several results presented below back up our main findings, we also deem it interesting to compare our two subject pools—participants from the general population and professionals from the finance industry—along these measures.

**Response rate analysis and self-selection.** For the finance professionals group, the response rate analysis shows that men responded to a greater extent than women, and that finance individuals in the age group 45–59 years responded to a slightly lesser extent than other ages. Furthermore, the non-response analysis shows that those with the lowest income responded to a somewhat higher degree compared to the others, and that those with a post-secondary education level of three years or more responded to greater extent than others. In the case of country of birth, the response rate was slightly higher for those born in Sweden compared to other countries. In the finance group there was a certain difference between the different job codes where the response frequency was slightly lower (5%) in the group of traders and portfolio managers (job code “2414”) compared with analysts and advisers (code “2413”) and brokers (code“3311”) (6.4%).

For the general population group, the response rate analysis shows similar patterns regarding gender, i.e., men responded to a greater extent than women. The response rate was lowest among the elderly. Furthermore, the response rate analysis indicates that those with the lowest and highest income responded to a somewhat higher extent compared to other income groups. When it comes to the level of education, those with a post-secondary education of three years or more tend to be over-represented in our sample. In the case of country of birth, the response rate was slightly higher for the ones born in Sweden compared to other countries.

A detailed summary of participants demographics compared to the characteristics of the sample invited is presented in Table E1. In particular, Table E1 reports the number of respondents and non-respondents per category of several socio-demographic characteristics, separated for both samples, as reported by *SCB*. Moreover, we report $\chi^2$-tests comparing whether participants in our samples differ significantly from those who have been invited by *SCB* but did not participate in the experiment. We report self-selection effects in terms of gender, age, country of birth, income, and education for the general population sample, and self-selection effects with respect to gender, age, and education for the finance professionals sample.

**Table E1: Sample characteristics by subject pools.** This table depicts the number (in %) of respondents (*"Resp."*), i.e., those who participated in our experiment, and non-respondents (*"No Resp."*), i.e., those who were invited but did not participate, for a number of different characteristics, separated for the general population and the finance profession sample. $\chi^2$-tests (with $k$-1 degrees of freedom) and the corresponding $p$-values are reported.

| | *General Population* | | | *Finance Professionals* | | |
|---|---|---|---|---|---|---|
| | *Resp.* | *No Resp.* | $\chi^2$ / $p$ | *Resp.* | *No Resp.* | $\chi^2$ / $p$ |
| *Gender:* | | | | | | |
| *Male* | 55.35 | 49.36 | 9.322 | 75.30 | 68.47 | 10.169 |
| *Female* | 44.65 | 50.64 | (0.002) | 24.70 | 31.53 | (0.001) |
| *Age:* | | | | | | |
| *20 – 29 years* | 11.55 | 10.28 | 37.789 | 11.85 | 8.73 | 14.062 |
| *30 – 39 years* | 31.69 | 23.18 | (< 0.001) | 31.12 | 28.79 | (0.015) |
| *40 – 49 years* | 26.62 | 26.39 | | 28.51 | 30.04 | |
| *50 – 59 years* | 20.99 | 26.74 | | 17.27 | 22.83 | |
| *60 – 69 years* | 9.15 | 13.41 | | 10.04 | 8.60 | |
| *70 – 79 years* | 0.00 | 0.00 | | 1.20 | 1.00 | |
| *Country of Birth:* | | | | | | |
| *Sweden* | 88.17 | 82.84 | 13.248 | 89.76 | 88.95 | 0.311 |
| *Abroad* | 11.83 | 17.16 | (< 0.001) | 10.24 | 11.05 | (0.577) |
| *Citizenship:* | | | | | | |
| *Swedish* | 97.04 | 95.64 | 3.132 | 97.59 | 96.53 | 1.604 |
| *Foreign* | 2.96 | 4.36 | (0.077) | 2.41 | 3.47 | (0.205) |
| *Marital Status:* | | | | | | |
| *Married* | 46.90 | 46.26 | 2.247 | 52.21 | 56.31 | 4.910 |
| *Unmarried* | 41.41 | 40.49 | (0.523) | 40.36 | 35.46 | (0.179) |
| *Divorced* | 11.27 | 12.42 | | 7.03 | 7.79 | |
| *Widowed* | 0.42 | 0.83 | | 0.40 | 0.45 | |
| *Income:* | | | | | | |
| *< 124,999* SEK | 3.24 | 2.70 | 25.646 | 2.01 | 1.53 | 2.985 |
| *125,000 – 199,999* SEK | 5.63 | 5.79 | (< 0.001) | 2.41 | 2.16 | (0.560) |
| *200,000 – 279,999* SEK | 12.82 | 15.25 | | 3.01 | 3.41 | |
| *280,000 – 369,999* SEK | 24.08 | 31.16 | | 5.22 | 6.85 | |
| *> 370,000* SEK | 54.23 | 45.11 | | 87.35 | 86.06 | |
| *Education:* | | | | | | |
| *No High School* | 1.83 | 8.89 | 198.587 | 0.80 | 1.08 | 32.058 |
| *High School* | 28.45 | 46.89 | (< 0.001) | 7.83 | 17.06 | (< 0.001) |
| *University (< 3 years)* | 19.86 | 14.95 | | 11.45 | 11.32 | |
| *University (> 3 years)* | 49.86 | 28.61 | | 79.72 | 69.95 | |
| *Unknown, n/a* | 0.00 | 0.66 | | 0.20 | 0.59 | |

**Table E2: Descriptive statistics and comparisons between pools for the survey items.** This table reports the means and standard deviations (in parentheses) for all survey items included in the experiment, separated for the general population and the finance professionals subject pool. The column "$t$-test" reports the differences in means and the $t$-values (in brackets) from two-sample $t$-tests based on $n = 958$. * $p < 0.05$, ** $p < 0.005$.

| | *Gen. Pop.* | *Fin. Prof.* | *$t$-Test* |
|---|---|---|---|
| Altruism/Charitable Giving | 0.79 | 0.85 | −0.061 |
| | (1.37) | (1.69) | [0.099] |
| Blame Shifting - Others | 1.12 | 1.19 | −0.065 |
| | (1.56) | (1.59) | [0.103] |
| Blame Shifting - Temptation | 1.55 | 1.78 | −0.239 |
| | (2.11) | (2.20) | [0.140] |
| Risk Tolerance in General | 4.79 | 5.80 | −1.017** |
| | (2.14) | (1.94) | [0.135] |
| Risk Tolerance | 4.28 | 6.09 | −1.806** |
| | (2.34) | (2.13) | [0.147] |
| Patience in General | 6.03 | 7.21 | −1.179** |
| | (2.00) | (1.81) | [0.125] |
| Trust in General | 5.71 | 5.79 | −0.080 |
| | (2.36) | (2.21) | [0.150] |
| Trust in Finance Professionals | 4.16 | 4.63 | −0.478** |
| | (2.33) | (2.37) | [0.154] |
| Trust in Investment Algorithms | 4.02 | 4.04 | −0.024 |
| | (2.25) | (2.45) | [0.153] |
| Frequent Investments | 3.54 | 6.69 | −3.149** |
| | (3.31) | (3.25) | [0.215] |
| Responsibility in Fin. Matters | 5.60 | 6.85 | −1.249** |
| | (3.67) | (3.33) | [0.231] |
| Use Expertise of Fin. Profs. | 3.58 | 2.21 | 1.376** |
| | (3.19) | (2.81) | [0.214] |
| Delegate to Fin. Profs. | 3.36 | 1.32 | 2.039** |
| | (3.11) | (2.16) | [0.192] |
| Delegate to Inv. Algorithms | 1.71 | 0.85 | 0.865** |
| | (2.37) | (1.67) | [0.147] |
| Observations | 550 | 408 | 958 |

*Notes:* All items, except for "Altruism," were answered on Likert scales ranging from 0 (minimum) to 10 (maximum). The variable "Altruism" refers to the amount transferred (up to 10,000 SEK) in a hypothetical charitable giving setting. For reasons of comparison, the variable is re-scaled to thousands SEK.

**Table E3: Descriptive statistics and comparisons between pools for numeracy, financial literacy, and overconfidence.** This table reports the means and standard deviations (in parentheses) for participants' numeracy and financial literacy scores, their self-estimates regarding their numeracy scores (in terms of estimates of the score and their relative performance compared to the Swedish general population), and the two measures of overconfidence (overestimation and overplacement), separated for the general population and the finance professional subject pool. The column "$t$-test" reports the differences in means and the $t$-values (in brackets) from two-sample $t$-tests based on $n = 958$. * $p < 0.05$, ** $p < 0.005$.

| | *Gen. Pop.* | *Fin. Prof.* | *$t$-Test* |
|---|---|---|---|
| *Skills:* | | | |
| *Numeracy Score* | 4.44 | 5.31 | −0.865** |
| | (1.63) | (1.59) | [0.106] |
| *Financial Literacy Score* | 4.29 | 5.39 | −1.099** |
| | (1.20) | (0.94) | [0.071] |
| *Self-Assessment:* | | | |
| *Estimated Numeracy Score* | 5.34 | 6.17 | −0.825** |
| | (1.81) | (1.59) | [0.112] |
| *Estimated Decile* | 0.56 | 0.68 | −0.121** |
| | (0.20) | (0.18) | [0.013] |
| *Overconfidence:* | | | |
| *Overestimation* | 0.90 | 0.86 | 0.040 |
| | (1.57) | (1.34) | [0.096] |
| *Overplacement* | −0.03 | −0.03 | 0.005 |
| | (0.23) | (0.19) | [0.014] |

*Notes:* Overestimation refers to the difference between participants' estimate of their numeracy and their actual numeracy score. Overplacement refers to the (minimum) difference between participants' estimate of the decile, their performance in the numeracy task belongs to, and the percentiles of the numeracy scores evaluated based on a normal distribution (see Appendix D for further details).

# F. Descriptives and Analyses of Time Spent

In the following, we examine the time spent per experimental task in the online experiment. Throughout the analysis, we truncate the time spent per task at the 99% percentile to avoid that outliers distort the results. In particular, for each task, durations exceeding this threshold are replaced by the value of the 99% percentile. Descriptive statistics of the time spent per task, separated for the general population and the finance professionals subject pools, are presented in Table F1. On average, the times spent in the experimental tasks appear to be sufficiently long to be confident that participants in both samples took the experiment very seriously, which is also confirmed by the high levels of decision-making quality (see Appendix C for details). Differences in the time spent between the two pools are reported in Table F2.

With respect to the main task, we examine learning effects by means of ordinary least squares regressions of the time spent on the 25 decisions on a linear time trend (with standard errors clustered at the subject level). The regressions reveal that the time spent per decision decreases with the progressing round numbers, in the decisions with both two and five assets, respectively. For the first two-asset item, participants from the general population take, on average, 57.1 seconds; for the subsequent decisions, the time spent, on average, decreases by 5.1 seconds per item ($t(548) = 13.916$, $p < 0.001$, $n = 5{,}500$). Finance professionals take, on average, 72.7 seconds for the first two-asset decision; for the following nine decisions with two assets, the time spent, on average, decreases by 6.5 seconds per item ($t(406) = 8.776$, $p < 0.001$, $n = 6{,}120$). Likewise, learning is observed for consecutive investment decisions with five assets. For the first five-asset item, participants from the general population take, on average, 3.5 minutes; for the subsequent decisions, the time spent, on average, decreases by 13.1 seconds per item ($t(548) = 2.065$, $p = 0.039$, $n = 5{,}500$). Finance professionals take, on average, 2.6 minutes for the first five-asset decision; for the following fourteen decisions with five assets, the time spent, on average, decreases by 7.7 seconds per item ($t(406) = 2.844$, $p = 0.005$, $n = 6{,}120$).

In addition, we investigate whether decision-making quality is systematically affected by time participants take to decide on the 25 investment decisions. Notably, ordinary least squares regression of *DMQI* on the time spent on the investment task (i.e., the sum of the time spent in the investment task with two and five assets) reveal that subjects' proneness to poor investment decisions is not significantly driven by the time they spend on each decision, neither in the general population sample ($\beta = 0.005$, $t(548) = 1.314$, $p = 0.189$, $n = 550$), nor in the finance professionals sample ($\beta = 0.003$, $t(406) = 1.261$, $p = 0.208$, $n = 408$).

**Table F1: Descriptive statistics of time spent per task.** This table reports the means and standard deviations (in parentheses) as well as the median and interquartile ranges (*IQR*; in brackets) for the time spent per experimental task (measured in minutes), separated for the general population sample (all treatments) as well as the three treatments conducted among finance professionals.

| | *GP-** | | *FP-FIXED* | | *FP-ALIGNED* | | *FP-OWN* | |
|---|---|---|---|---|---|---|---|---|
| | *m / sd* | *q50 / iqr* | *m / sd* | *q50 / iqr* | *m / sd* | *q50 / iqr* | *m / sd* | *q50 / iqr* |
| Investment Task w/ Two Assets | 5.53 | 4.30 | 7.00 | 5.13 | 7.26 | 5.07 | 5.16 | 4.25 |
| | (4.17) | [3.10] | (5.37) | [4.56] | (5.23) | [5.98] | (3.94) | [3.08] |
| Investment Task w/ Five Assets | 15.24 | 11.99 | 18.40 | 13.57 | 19.11 | 13.73 | 15.65 | 11.77 |
| | (10.73) | [10.52] | (14.85) | [14.41] | (15.44) | [14.02] | (12.07) | [10.30] |
| Questionnaires (Self-Reported) | 2.67 | 2.30 | 2.49 | 2.18 | 2.55 | 2.20 | 2.52 | 2.15 |
| | (1.54) | [1.20] | (1.29) | [1.01] | (1.53) | [1.08] | (1.16) | [1.27] |
| Numeracy Inventory (8 Items) | 3.65 | 4.00 | 3.68 | 4.00 | 3.58 | 4.00 | 3.60 | 4.00 |
| | (0.55) | [0.62] | (0.57) | [0.58] | (0.64) | [0.92] | (0.62) | [0.77] |
| Financial Literacy Test (6 Items) | 2.05 | 1.99 | 1.74 | 1.63 | 1.76 | 1.63 | 1.77 | 1.65 |
| | (0.56) | [0.85] | (0.59) | [0.77] | (0.52) | [0.72] | (0.60) | [0.90] |
| Observations | 550 | | 132 | | 139 | | 137 | |

**Table F2: Differences in time spent.** This table reports the $t$-statistics from two-sample $t$-tests between the general population sample (pooled across all treatments) and the finance professionals sample separated for the treatment conditions for the time spent per experimental task (measured in minutes). Standard errors ($se$) are reported in parentheses. Means, standard deviations, medians, and interquartile ranges for the time spent per experimental task in all treatments are reported in Table F1. * $p < 0.05$, ** $p < 0.005$.

| | *GP-** vs. *FP-FIXED* | *GP-** vs. *FP-ALIGNED* | *GP-** vs. *FP-OWN* | *FP-FIXED* vs. *FP-ALIGNED* | *FP-FIXED* vs. *FP-OWN* | *FP-ALIGNED* vs. *FP-OWN* |
|---|---|---|---|---|---|---|
| | *t / se* | *t / se* | *t / se* | *t / se* | *t / se* | *t / se* |
| Investment Task w/ Two Assets | −2.447* | −2.975** | 0.988 | −0.390 | 3.218** | 3.754** |
| | (0.558) | (0.544) | (0.483) | (0.644) | (0.573) | (0.558) |
| Investment Task w/ Five Assets | −1.218 | −1.676 | 0.676 | −0.388 | 1.670 | 2.076* |
| | (1.509) | (1.523) | (1.353) | (1.842) | (1.648) | (1.670) |
| Obs. | 315 | 322 | 320 | 271 | 269 | 276 |

## G. Supplementary Results

**Table G1: Finance professionals' risk taking when deciding on behalf of clients.** This table reports estimates from ordinary least squares regressions of the mean standard deviation across the 25 items of the investment task (*SD*) on an indicator variable for the treatment *FP-ALIGNED*, indicators for the risk level professionals are asked to take into consideration when deciding on clients' behalf, a set of experimental measures, and self-reported measures. Robust standard errors are reported in parentheses. * $p < 0.05$, ** $p < 0.005$.

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| *Treatment Indicator:* | | | | |
| *FP-ALIGNED* | 2.387 | 2.284 | 2.360 | 2.671 |
| | (4.403) | (4.351) | (4.447) | (4.412) |
| *Given Risk Level:* | | | | |
| Risk Level=2 | 18.951** | 18.572** | 19.722** | 19.241** |
| | (6.323) | (6.316) | (6.236) | (6.241) |
| Risk Level=3 | 38.613** | 37.865** | 39.499** | 39.502** |
| | (5.582) | (5.827) | (5.598) | (5.818) |
| Risk Level=4 | 94.786** | 93.585** | 95.742** | 95.082** |
| | (6.609) | (6.544) | (6.632) | (6.585) |
| *Experimental Measures:* | | | | |
| Decision Making Quality Index | | | 2.465 | 2.500 |
| | | | (2.134) | (1.930) |
| Financial Literacy Score (Std.) | | | 9.341 | 6.044 |
| | | | (4.824) | (5.146) |
| Numeracy Score (Std.) | | | −2.299 | −5.122 |
| | | | (5.311) | (5.518) |
| Overestimation (Std.) | | | −5.235* | −5.824* |
| | | | (2.605) | (2.571) |
| Overplacement (Std.) | | | 2.239 | 1.385 |
| | | | (3.243) | (3.353) |
| *Self-Reported Measures:* | | | | |
| Risk Tolerance (Std.) | | | 1.945 | 2.875 |
| | | | (2.630) | (2.761) |
| Blame Shifting (Std.) | | | −1.814 | −2.401 |
| | | | (2.580) | (2.565) |
| *Constant:* | | | | |
| *FP-FIXED* | 40.894** | 33.745* | 22.090* | 15.431 |
| | (4.779) | (15.283) | (10.186) | (19.686) |
| Controls | *no* | *yes* | *no* | *yes* |
| $F$ | 56.676 | 33.283 | 23.477 | 19.787 |
| $p > F$ | 0.000 | 0.000 | 0.000 | 0.000 |
| Adj. $R^2$ | 0.486 | 0.498 | 0.497 | 0.507 |
| Observations | 271 | 271 | 271 | 271 |

*Notes:* Treatment *GP-FIXED* serves as reference condition. All self-reported measures are standardized scores. "Blame Shifting" refers to the mean of two standardized survey items on shifting blame on others and resisting the temptation to shift blame on others. "Controls" include gender (binary indicator for female), age (in years), net income from major employment in 2017 (in thousand SEK's), and maximum education level (dichotomous indicators for high school education or less, university education smaller or equal to three years, and university education larger than three years).

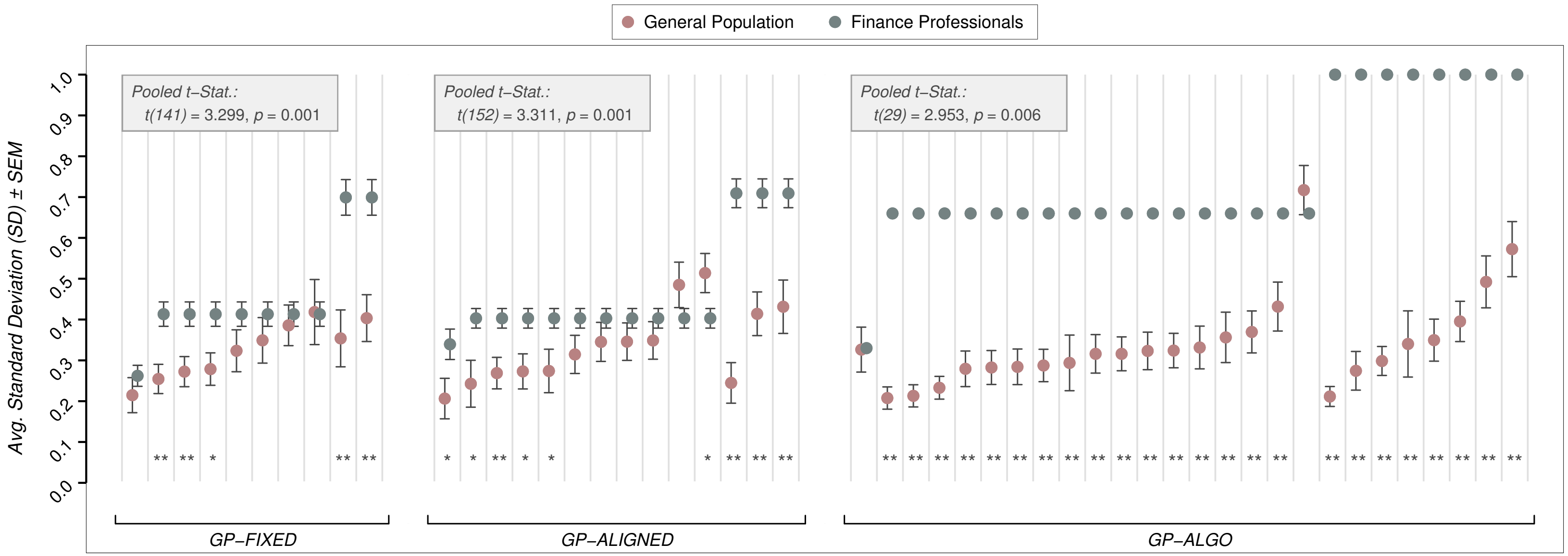


**Figure G1: Portfolio risk of clients' asking the agent to take more risk when delegating.** This figure shows the average portfolio risk across the 25 investment decisions (*SD*; normalized to 1) of those clients that choose to delegate *and* ask the agent to take more risk than they think they took in their own decision, separated for the treatments *GP-FIXED*, *GP-ALIGNED*, and *GP-ALGO* (red dots). The blue dots indicate the mean portfolio risk across the 25 investment decisions (*SD*) of agents that serve as potential matching partners, i.e., those in the corresponding treatment deciding for clients with the risk level that matches their desired risk level when delegating. Error bars indicate standard errors of the mean (*SEM*) and are clustered on the individual level for agents. Asterisks indicate significant differences per principal-agent level and are based on two-sample $t$-tests (with clustered standard errors); * $p < 0.05$, ** $p < 0.005$. Aggregate comparisons between clients' and agents' portfolio risk per treatment are reported in the gray boxes. $t$-statistics are based on ordinary least squares regressions of portfolio risk on an indicator variable for "agent", controlling for risk level indicators, with standard errors being clustered on the individual level.

| | RL–1 | RL–2 | RL–3 | RL–4 | HHI |
|---|---|---|---|---|---|
| 0.00 – 0.10 | 0.786 | 0.179 | 0.036 | 0.000 | 0.651 |
| 0.10 – 0.20 | 0.529 | 0.324 | 0.118 | 0.029 | 0.400 |
| 0.20 – 0.30 | 0.382 | 0.382 | 0.206 | 0.029 | 0.336 |
| 0.30 – 0.40 | 0.250 | 0.271 | 0.396 | 0.083 | 0.299 |
| 0.40 – 0.50 | 0.162 | 0.270 | 0.405 | 0.162 | 0.290 |
| 0.50 – 0.60 | 0.107 | 0.071 | 0.500 | 0.321 | 0.370 |
| 0.60 – 0.70 | 0.000 | 0.118 | 0.294 | 0.588 | 0.446 |
| 0.70 – 0.80 | 0.100 | 0.200 | 0.200 | 0.500 | 0.340 |
| 0.80 – 0.90 | 0.000 | 0.111 | 0.111 | 0.778 | 0.630 |
| 0.90 – 1.00 | 0.095 | 0.048 | 0.000 | 0.857 | 0.746 |

**Figure G2: Number of portfolios with similar portfolio risk across risk levels.** This figure shows the fraction of finance professionals' portfolios (when deciding on behalf of principals, i.e., in treatments *FP-FIXED* and *FP-ALIGNED*) across equally-sized classes of portfolio risk (normalized to 1) over the four risk levels. The color coding increases with the cell's magnitude. The column HHI refers to the Herfindahl-Hirschman-Index, a diversity index defined as $HHI = \sum_k s_k^2$ with $s_k$ denoting the share in risk level $k = \{1,2,3,4\}$. HHI takes a minimum value of 0.25 (if $s_k = 0.25 \; \forall k$) and a maximum value of 1 (if $s_k = 1 \; \forall k$).