universität innsbruck

Research Platform
Empirical and Experimental Economics

You can't always get what you want–An experiment on finance professionals' decisions for others

Matthias Stefan, Martin Holmén, Felix Holzmeister, Michael Kirchler, Erik Wengström

Working Papers in Economics and Statistics

2022-02

University of Innsbruck
http://uibk.ac.at/eeecon/

University of Innsbruck
Working Papers in Economics and Statistics

The series is jointly edited and published by

- Department of Banking and Finance
- Department of Economics
- Department of Public Finance
- Department of Statistics

Contact address of the editor:
research platform “Empirical and Experimental Economics”
University of Innsbruck
Universitaetsstrasse 15
A-6020 Innsbruck
Austria
Tel: + 43 512 507 71022
Fax: + 43 512 507 2970
E-mail: eeecon@uibk.ac.at

The most recent version of all working papers can be downloaded at
https://www.uibk.ac.at/eeecon/wopec/

For a list of recent papers see the backpages of this paper.

You can't always get what you want—An experiment on finance professionals' decisions for others

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

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

[*] *Corresponding author:* *matthias.stefan@uibk.ac.at*

Abstract

To study whether clients benefit from delegating financial investment decisions to an agent, we run an investment allocation experiment with 408 finance professionals (agents) and 550 participants from the general population (clients). In several between-subjects treatments, we vary the mode of decision-making (investment on one's own account vs. investments on behalf of clients) and the agents' incentives (aligned vs. fixed). We find that finance professionals show higher decision-making quality than participants from the general population when investing on their own account. However, when deciding on behalf of clients, professionals' decision-making quality does not significantly differ from their clients', neither when compensated with a fixed payment nor when facing aligned incentives. Our results further identify a considerable challenge in risk communication between agents and clients: While finance professionals tend to take into account principals' desired risk levels, the constructed portfolios by professionals show considerable overlaps in portfolio risk across different risk levels requested by principals. We argue that this result is due to differences in risk perception.

Keywords: Experimental finance, finance professionals, delegated decision-making, risk communication. *JEL Classification:* C93, G11, G41.

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 at the University of Innsbruck, the NOeG Conference (Austrian Economic Association) in Graz, the SFB Workshop in Innsbruck, the Experimental Finance Conference in Copenhagen, the Decision-Making for Others Workshop in Nijmegen, and the Workshop on Trust and Cooperation in Markets and Organizations in Stavanger 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).

"You can't always get what you want
But if you try sometime you'll find
You get what you need"

—Keith Richards & Mick Jagger (1969)

1. Introduction

In light of the great variety and increasing complexity of financial products, demand for financial advisory services and delegation of investment decisions has become more and more important. The large and ever-growing market for financial advice and delegated investments highlights the economic relevance of these services: As of 2021, the financial planning and advise industry in the United States has been estimated to $56.9 billion, and more than 230,000 personal financial advisers have been employed in the sector.[1] When finance professionals act on behalf of clients, the outcomes of their decisions can have substantial consequences for clients' financial well-being. Previous contributions to the literature, taking a critical look at financial advise, however, disclose various potential pitfalls for clients. For instance, preceding evidence suggests that investors who most need financial advice are least likely to obtain it (Bhattacharya et al., 2012; Bachmann and Hens, 2015), that advisory services do not help to reduce behavioral biases (Mullainathan et al., 2012), and that advised accounts—as compared to unadvised accounts—tend to perform worse in terms of risk-return tradeoffs (Hackethal et al., 2012). Moreover, Egan et al. (2019) document that misconduct is a prevalent feature in the industry for financial advice, and theoretical contributions by Inderst and Ottaviani (2012a,b) show that potential conflicts of interest and the advisor's incentive scheme might be a breeding ground for biased advisory services.

The critical stance on financial advice and delegation documented in the literature leads us to examine two focal preconditions for efficient delegation of financial decision-making in our paper, leading to the following research questions: *(a)* Do finance professionals (agents) make better decisions than their clients (principals) and does professionals' decision-making quality vary with the monetary incentives the face? This aspect is motivated by one of the main reasons for delegation, namely professionals' attributed superior ability in financial decision-making quality (and the related assumption that these abilities are exerted with more effort given the "right" incentives among professionals). *(b)* Can clients' investment preferences be purposefully communicated and are finance professionals able and willing to act upon them? We consider successful communication of risk attitudes an important precondition for the efficiency of delegated investments, as it constitutes a constraint for clients eventually ending up with what they ask for.

We run a controlled lab-in-the field online experiment with 408 finance professionals (advisers) and 550 participants from the general population (clients) in Sweden. In the main task of our experiment, all participants—i.e., both clients and advisers—were required to make 25 investment allocation decisions. In three between-subjects treatments, we varied whether finance professionals decide on their own behalf or on the account of clients, and, in the latter case, whether they face aliened incentives or receive a flat compensation for investing on behalf of others. When investing on a client's account, agents were

[1] https://perma.cc/9JXP-N83N

instructed to implement portfolios for one of four risk levels that could be requested by their clients. After having made their investment decisions, participants from the general population had the opportunity to delegate their investment choices to a financial professional. In case a client opted for delegation, they could indicate the risk level they want the agent to take on their behalf, knowing that they would be matched with a finance professional who constructed portfolios for the respective riskk profile.

Our main findings are as follows: As compared to participants from the general population, finance professionals show higher decision-making quality when deciding on their own account—though only for moderate and high levels of risk-taking—but not when investing on behalf of their clients. The latter turns out to hold irrespective of whether professionals are compensated with a flat payment or whether they face an aligned incentive scheme. We further show that finance professionals, on average, take into consideration their clients' desired risk levels, but construct portfolios with considerable overlaps in portfolio risk across clients' requested risk levels. This gives rise to the observation that clients who request different risk levels when delegating their investments to the agent eventually might end up with portfolios that are identical in terms of the risk-return tradeoff they involve. We provide evidence that this result emerges from systematic differences in the perception of risk between finance professionals and laypeople. The result suggests that risk communication constitutes a major challenge for the efficiency of delegating financial investments.

This paper contributes to the following strands of research: First, our study relates to the growing body of experimental literature investigating investment decisions on behalf of others, which has put forth a set of mixed results. Several studies report a "risky shift" in investment choices, 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). By contrast, some studies 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. We contribute by extending the pool of participants to those that take investment decisions in real-life, i.e., financial analysts, investment advisers, traders, fund managers, and brokers. Thus, we study the behavior of those agents that take these decisions professionally, arguably also because they are considered superior in financial decision-making. Second, our study relates to the small but growing literature on financial advise. The evidence on improved quality of advised portfolio decisions is mixed: Kramer (2012) finds increased diversification for advised portfolios but no differences in risk-adjusted returns. Similarly, Bluethgen et al. (2008) finds more diversification among advised portfolios, while costs are elevated and portfolios mostly predefined. Moreover, Hackethal et al. (2012) show that advised financial accounts exhibit lower returns, inferior Sharpe-ratios, and higher turn-over rates. Based on an audit study, Mullainathan et al. (2012) report that financial advise often does not reduce, but even increase existing behavioral biases in the interest of the advisers. However, evidence on delegated financial decision-making is indirect, since advise only turns into delegation, if it is adopted by clients. Indeed, existing evidence shows that advise is often not followed (Bhattacharya et al., 2012; Stolper, 2018). Thus, causal evidence on delegated financial decisions is still missing, which is our contribution to this strand of literature. Our findings, might, to some extent, contribute to the explanation of the mixed results outlined above: We find that professionals indeed try to customize portfolios to their clients (see also Kling et al., 2019; Rose, 2021), but do not show superior decision-making quality in our experiment when investing on their clients' behalf.

Third, our study relates to the literature on risk communication between agents and their clients. For instance, Gennaioli et al. (2015) argue that clients ask money managers to take risky investment decisions they are too anxious taking themselves (see Bucciol et al., 2019; Holzmeister et al., 2021, for related empirical and experimental evidence). However, the communication of risk between money managers and clients can be difficult, if the perception of risk systematically differs between the parties involved. Indeed, several studies show that clients' portfolio risk associated with delegated investments hinges on the agents' risk attitudes. For example, in an experimental study with finance professionals, Kirchler et al. (2020) show that professionals' self-assessed risk attitudes in financial matters explains risk-taking on behalf of clients. Building on household portfolio data, Foerster et al. (2017) report that adviser fixed effects explain considerably more variation in portfolio risk than a broad set of investor attributes. Similarly, Linnainmaa et al. (2019) show that most advisers invest their personal portfolios just as they advise their clients. These studies give rise to the question whether finance professionals are not willing (see Rose, 2021, for a discussion) or unable to customize their decisions. Again, causal evidence to provide an answer to this question is widely missing. We provide further insights into this question based on a controlled experimental setup and show that there is miscommunication of risk between clients and advisers that can be attributed to differences in risk perception. This can lead to sub-optimal portfolio allocations, as clients requesting different risk levels might end up with very similar portfolio risks. This finding is in line with the results documented in a study closely related to ours: Kling et al. (2019) show in their experiment that, even though agents show a high willingness to implement their clients' preferred investment profiles, the perception of risk profiles is very heterogeneous, which results in substantial miscommunication between clients and agents.

Finally, we add to the literature on financial agents' incentives. Often, agents face incentives not to correct investors' biased beliefs and inferior financial decisions (see, e.g., Mullainathan et al., 2012; Inderst and Ottaviani, 2012a,b) and clients seem to anticipate differences in decision-making quality depending on the agents' incentives. For instance, Holzmeister et al. (2021) show that clients delegate more frequently to professionals facing aligned incentives as compared to agents that are compensated with a flat payment. While such expectations appear to be substantiated in the study by Mullainathan et al. (2012), where advisers do not de-bias clients when incentives are misaligned, Kling et al. (2019) provide contrasting evidence and show that, even facing incentives to disregard clients' preferences, agents try to comply with the requested risk profiles. With our study, we provide further insights on the role of agents' incentives and show that decision-making quality does not systematically increase with professionals' aligned incentives.

2. Experimental Design

In what follows, we describe the allocation decision task, the treatment variations, the focal variables derived from the experiment, as well as the data collection and recruitment procedure. Note that there is a companion paper (Holzmeister et al., 2021) that is based on the same experiment. An earlier working paper version presented the results on various aspects of delegated investment decisions in a single manuscript (Holzmeister et al., 2019). During the revision process, we split our contribution into two separate papers to discuss our objectives, our results, and their implications in a more targeted manner. In particular, the companion paper focuses on a different set of variables to address the demand side of delegated invest-

ments; by contrast, the paper at hand zeros in on the supply side. Since both papers are based on the same dataset, substantial parts—such as the description of the experimental design, the recruitment, and relevant variables as well as their construction—are virtually identical in both papers.

Allocation decision task. The main task in our experiment is the allocation decision task introduced by Banks et al. (2018). The task consists of two blocks with 10 decisions with two binary assets and 15 decisions with five binary assets, respectively. Participants first read the instructions for the first block, and then had to correctly answer three comprehension questions to be able to proceed with the investment allocation task.[2] After the first ten decisions of the first block, participants read the instructions for the second block and continued with the next 15 decisions with the five assets. The order of the two blocks was fixed for all participants, but the order of decisions was randomized in each of the two blocks. Figure A1 in Appendix A shows screenshots of the allocation decision task with two and five assets, respectively.

Table 1: Return distributions of the available assets in the 25 opportunity sets. This table shows the 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 (sets #1–10) and five assets (sets #11–25, 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

[2] Participants 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.

For each of the 25 decisions, participants had to allocate an endowment of 100 SEK on the two or five assets.[3] The assets' returns depended on a coin toss (heads or tails) and were shown per 1 SEK invested.[4] 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.

At the end of the experiment, one of a participant's own or, if participants chose to delegate their decisions to the agent, one of the agent's decisions was randomly chosen, and then a simulated coin toss determined the participant's payoff. 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. While we provide an overview of the four measures below, further details are provided in Appendix C.

First, violations of first order stochastic dominance (*FOSD*, Hadar and Russell, 1969) are measured by the difference between the expected return of a portfolio chosen by the participant and the maximum expected return of a feasible portfolio that has the same same minimum payoff as the chosen one. This idea can be illustrated using the example of opportunity set #1 as described in Table 1 above: Asset A_1 (B_1) yields a payoff of 1.20 SEK (3.60 SEK) if the coin shows up tails (heads) and 0.00 SEK otherwise. Suppose a participant i chooses portfolio $x_{i,1} = (80, 20)$, i.e., she allocates 80 SEK of the endowment to asset A_1 and 20 SEK to asset B_1. If the coin shows up heads, participant i receives a payoff of $20 \cdot 3.60 = 72$ SEK and if the coin shows up tails, she receives $80 \cdot 1.20 = 96$ SEK, thus yielding an expected portfolio return of 84 SEK. Now, this portfolio $x_{i,1}$ is first-order stochastically dominated: Choosing the allocation $x'_{i,1} = (60, 40)$ would guarantee the same minimum payoff ($60 \cdot 1.20 = 72$ SEK) but a higher maximum payoff ($40 \cdot 3.60 = 144$ SEK). The expected return of this alternative portfolio $x'_{i,1}$ is 108 SEK, thus exceeding the expected return of portfolio $x_{i,1}$ by 24 SEK. This difference in expected returns constitutes our measure of the *FOSD* violation. Note that choosing portfolio $x_{i,1}$ rather than $x'_{i,1}$ implies that a sizeable fraction of the prospective reward— conditional on the fact that both portfolios result in the same minimum payoff— would be left on the table, namely 24 SEK $\div$ 108 SEK $= 22.2\%$.

Second, violations of the *Generalized Axiom of Revealed Preferences* (*GARP*) are measured using the *Money Pump Index* (Echenique et al., 2011). The intuition behind our measure is that a decision-maker violating *GARP* could be exploited as a "money pump" by an arbitrageur who replicates the chosen portfolios at lower cost and sells them to the decision-maker at higher prices. Again, this measure can be illustrated using the same example of the participant i, making the investment in opportunity set #1 as described above. Now consider opportunity set #10, which is a mirror versions of set #1: Asset A (B) yields a payoff of 1.20 SEK (3.60 SEK) if the coin shows up heads (tails) and 0.00 SEK otherwise. Suppose participant i chooses portfolio $x_{i,10} = (90, 10)$. The portfolio choices $x_{i,1} = (80, 20)$ and $x_{i,10} = (90, 10)$ are not only first-order stochastically dominated (as per the argument sketched above), but also violate the *Generalized Axiom of Revealed Preferences.* The portfolio $x_{i,10}$ yields a payoff of $90 \cdot 1.20 = 108$ SEK if the coin shows up heads and a payoff of $10 \cdot 3.60 = 36$ SEK if the coin shows up tails. Note that the same portfolio is feasible

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

[4] We used the same returns as Banks et al. (2018), multiplied them by a factor of 1.5, and rounded them to one decimal place.

in opportunity set #1 without spending the entire endowment. Allocating 30 SEK of the endowment into asset A_1 and 30 SEK into asset B_1 would replicate the very same portfolio (with payoffs $30 \cdot 3.60 = 108$ SEK in case the coin shows up heads and $30 \cdot 1.20 = 36$ SEK in case it shows up tails). An arbitrageur could thus buy the replicated portfolio in opportunity set #1 yielding an expected return of 72 SEK at a "price" of only 60% of the endowment in opportunity set #1 and sell it to participant i at a "price" of 100% in opportunity set #10, involving a risk-free profit $0.40 \cdot 72\text{SEK} = 29\text{SEK}$. Following the same line of thought, $x_{i,1}$ can be constructed in opportunity set #10. To obtain a payoff of 72 SEK (heads) and 96 SEK (tails), respectively, in opportunity set #10, one would need to invest 60 SEK of the endowment in asset A_{10} and $26.\dot{6}$ SEK in asset B_{10}. Again, a hostile arbitrageur could buy the replicated portfolio (yielding an expected return of 84 SEK) at a "price" of $86.\dot{6}$ SEK in opportunity set #10 and sell it to participant i at a "price" of 100 SEK in opportunity set #1, involving a risk-free reward of 14 SEK. Since participant i's allocation decisions $x_{i,1}$ and $x_{i,10}$ can be replicated at an expense of less than 100% of the endowment in the respective other opportunity set each, she leaves a total of $29 + 14 = 43$ SEK—or 24% of the expected returns of the two portfolio decisions—in prospective rewards to a fictive arbitrageur.

Third, financial competence (*FC*) is a measure of a participant's ability to understand the task at hand: In our experiment, four opportunity 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. Consequently, two opportunity sets were presented four times each (#1 = #10 = #11 = #15 and #2 = #8 = #18 = #25). Assuming that a financially competent investor would understand the opportunity sets and invest consistently, we measure a participant's financial competence for identical opportunity sets as the absolute differences between the expected returns of the chosen portfolios.

Fourth, failure to minimize risk (*FMR*) is based on the assumption of a risk averse investor: In our experiment for two opportunity sets (#5 and #19 in Table 1), the expected return per 1 SEK invested was the same for all assets k. Consequently, all feasible portfolios in these opportunity sets share the same expected return. Choosing a fully-hedged portfolio (second-order) dominates all other feasible portfolios in these two opportunity sets. Thus, a participant's failure to minimize risk in these opportunity sets is measured as the standard deviation $SD_{i,j}$ of the particular portfolio allocation.

Each measure is averaged across all (relevant) opportunity sets. To construct a composite measure of decision-making quality, we conduct a principal component analysis of the four measures described above (see Table C1 in Appendix C). In particular, the first principle component constitutes our decision-making quality index (*DMQI*). Table C2 in Appendix C shows the pairwise correlations between the decision-making quality measures for both the general population sample and the finance professionals sample. Violations of *FOSD* and *GARP* and, to a lesser extent, the *FC* measures are positively correlated for both groups. While the *FMR* measure is also correlated with the other measures for the general population sample, the correlations for the finance professionals turn out to be insignificant.[5]

[5] It could be argued that this might be driven by the fact that the *FMR* measure does not necessarily capture decision mistakes, as failures to minimize risk could also be driven by preferences for risk-seeking. However, we follow Banks et al. (2018) in their choice of components for decision-making quality. Moreover, we use the first principal component of the four measures (in contrast to Banks et al. (2018); see SectionC in the Appendix) as our *DMQI* variable, with the consequence that uncorrelated influences—such as risk preference—should not pose a systematic issue for our measure.

Experimental treatments. Conditional on the subject pool, participants were randomly assigned to one of the treatments listed in Table 2. Irrespective of the treatment assigned, participants completed the 25-item allocation decision task described above first. Subsequently, participants from the general population could delegate their decisions to an agent. If they opted for delegation, their experimental payoff depended on the assigned agent's rather than their own decisions (details are provided below). This means that participants made the investment decision first and only then were informed about the opportunity to delegate the investment decisions. By this means, our experimental design allows studying whether delegation pays off for those who delegate *and* whether it would have been beneficial for those who choose to stick to their own decisions, since investment choices are observed for all participants. Moreover, we can study risk communication by comparing clients' and professionals' investment decisions conditional on the risk they take in the investment task.

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)

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 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. This treatment design with various types of agents is relevant for the companion paper (Holzmeister et al., 2021), where we study delegation decisions to human and algorithmic advisers.

If principals chose to delegate, they were asked to specify the risk (on a scale from 1 [no risk] to 4 [maximum return]) to be taken on their behalf by the agent, 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, 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 participants 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 her decisions, 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 decision task on a scale from 1 to 4, i.e., on the same scale clients face 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 financial 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 advise-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. 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.

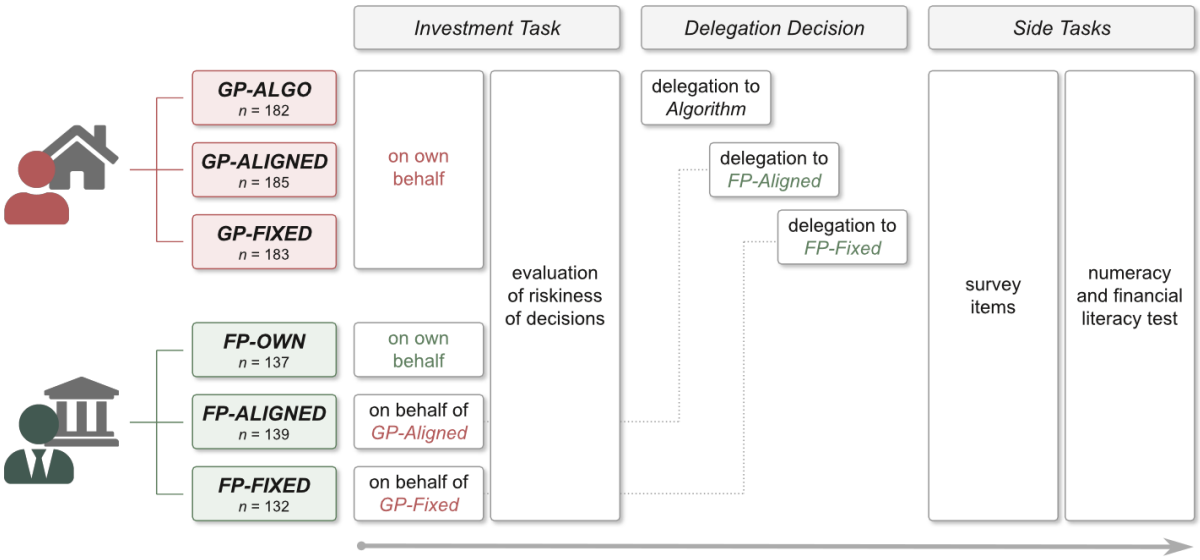

Figure 1: Flow chart of the experiment. This figure illustrates the sequence of tasks for participants in our experiment. First, participants were randomly assigned to one treatment and completed 25 investment decisions. Then, participants from the general population could delegate their investment decision to an agent in a delegation decision stage. Finally, all participants 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 all 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.

Recruitment and data collection. The experiment was conducted in Sweden in cooperation with *Statistiska centralbyrån* (*SCB*; Statistics Sweden), who invited participants for the experiment.[6] *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 participants from Sweden's working general population, excluding finance professionals. The sample of finance professionals consists of 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. Moreover, participants were informed that the study has been approved by the ethical review boards in Gothenburg and at *SCB*. 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.

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.[7] The experimental data was collected between January 4 and February 10, 2019.[8]

Sample descriptives. , The average participant in our experiment is 42 years old. 75% of participants from the finance professional sample and 55% of the general population sample are male. The relatively low fraction of females in the finance industry is typical for the job functions under consideration and has also been reported in previous studies (see, e.g., Kirchler et al., 2018; Weitzel et al., 2019). Both finance professionals and participants from the general population are highly educated: about 80% of finance practitioners and about 50% of the general population sample hold a university degree. The average gross

[6] *SCB* also provided a set of predefined variables of the participants' register data for those who completed the experiment. The participants' register data are used as control variables in Holzmeister et al. (2021) and Holmen et al. (n.d.), but do not enter any of the analyses presented in this paper. See Appendix A for further details.

[7] Thus, the average hourly salary for finance professionals and general population participants amounts to approximately $40. This is comparable to other studies with general population participants (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 (see, e.g., Haigh and List, 2005; Alevy et al., 2007; Weitzel et al., 2019).

[8] In total, only a relatively small fraction of participants—especially for an online experiment—dropped out during the experiment. Overall, 68.9% of all participants 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. For comprehensive response rate analyses, refer to Appendix E.

income (from major employment) among finance professionals and participants from the general population amounts to SEK 722,046 ($sd = 547{,}815$) and SEK 393,706 ($sd = 259{,}726$), respectively. Further details on the socio-economic background of both samples are presented in Table E1 in Appendix F.

3. Results

A prerequisite for the efficiency of delegated investments is that agents score higher than principals in skills relevant for financial investments and decision-making quality. As a first step towards addressing our research questions, we thus start with testing whether this requirement is fulfilled by comparing the decision-making quality between finance professionals and laypeople when deciding on their own account. In a second step, we focus on the investment decisions of financial professionals made on behalf of clients. In particular, we examine whether agents' decision-making quality systematically differs when investing on a client's account as compared to when deciding on one's own behalf, and whether different incentive schemes affect the agents' susceptibility to poor investment decisions. Finally, we focus our attention on the question whether risky investments can be properly delegated to an agent in terms of communicating the desired risk level. In particular, we investigate whether clients and agents share the same understanding of what is risky, whether agents are capable (and willing) to implement the risk profile requested by clients, and whether clients eventually get what they ask for.

3.1. Decision-Making Quality

Assuming that individual-level risk preferences are perfectly mapped by the agents' decision, clients in our experiment can only benefit from delegating their investment decisions if agents, on average, show superior decision-making quality. Intuitively, one would hypothesize that finance professionals—as compared to laypeople—are equipped with higher levels of skills that are relevant to financial decision-making. Figure 2 supports this intuitive expectation: Finance professionals in our sample show significantly higher levels of both numeracy (*FP-**: $m = 5.3$, $sd = 1.6$; *GP-**: $m = 4.4$, $sd = 1.6$; two-sample t-test: $t(956) = 8.194, p < 0.001$) and financial literacy (*FP-**: $m = 5.4$, $sd = 0.9$; *GP-**: $m = 4.3$, $sd = 1.2$; two-sample t-test: $t(956) = 15.383, p < 0.001$) than participants from the general population. We deem these two measures compelling to proxy for financial sophistication. Indeed, decision-making quality in the investment task (*DMQI*) turns out being significantly correlated to both numeracy ($\rho = 0.193$, $p < 0.001$; $n = 958$) and financial literacy ($\rho = 0.170$, $p < 0.001$; $n = 958$). The sizeable differences in the two measures between subject pools and their correlation to *DMQI*, thus, are a first indication that clients could potentially benefit from delegating their investments.

To directly examine the proposition that finance professionals are less prone to subpar decisions in the experimental investment task than laypeople, we test whether *DMQI* differs systematically between the two subject pools. Since *DMQI* might be influenced by whether decisions are made on one's own account or on behalf of others and, in the latter case, by the incentive scheme faced by agents, we first restrict out attention to the baseline sample of finance professionals making investments for themselves but not for

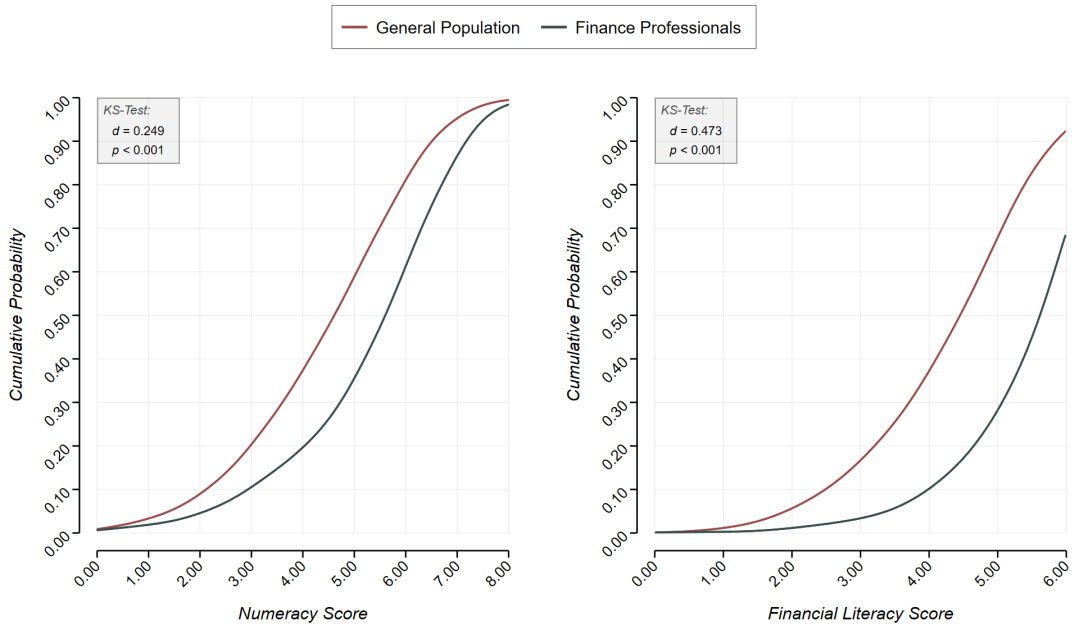

Figure 2: Numeracy and financial literacy scores. This figure shows empirical cumulative density distributions (based on Gaussian kernels with a bandwidth of 1) for participants' numeracy and financial literacy scores, respectively. Kolmogorov-Smirnov-tests are reported in the light gray boxes (n = 958).

clients (*FP-OWN*; $n = 137$).[9] A comparison of participants' *DMQI* in treatments *FP-OWN* and *GP-** suggests that finance professionals, on average, are indeed capable of making better investment decisions (model (1) in Table 3; $b = 0.320$, $p < 0.001$; $n = 687$). As to the economic magnitude of the difference in *DMQI*, finance professionals in *FP-OWN* score about 0.3 standard deviations higher in terms of decision-making quality as compared to participants from the general population, which corresponds to a small to moderate effect.[10] Since our measure of decision-making quality (*DMQI*) quantifies how much is left on the table in terms of returns, higher *DMQI* scores are—by construction of the experimental task—correlated with (risk-adjusted) expected returns. An ordinary least squares regression of the average expected return on *DMQI*, controlling for the portfolio's mean standard deviation, indicates that expected returns significantly increase with higher levels of decision-making quality ($b = 2.220$, $t(954) = 16.611$, $p < 0.001$; $n = 958$). This suggests that decision-making quality systematically increases expected investment returns above and beyond the inherent relationship between risk and returns. Thus, under the premise that agents are capable—and willing—to implement the principal's requested level of risk, delegating investments could indeed turn out beneficial for clients.

[9] When analyzing decision-making on one's own account, we include participants from all three general population treatments (*GP-**; $n = 550$). Taking into account all data from the general population sample (*GP-**) is feasible, since participants from the general population made their own decision first and only then decided whether or not to delegate. Clients were not aware of the opportunity to delegate their decisions to an agent when completing the investment task. This is also the reason why participants' *DMQI* in treatment *GP-ALGO* can plausibly enter the comparison between subject pools.

[10] The difference in *DMQI* between *GP-** and *FP-OWN* is reduced by about 55% ($b = 0.144$, $p = 0.082$; $n = 687$) when controlling for participants financial literacy and numeracy scores. This suggests that finance professionals' superior decision-making quality in the experimental task is—to a substantial part—due to better knowledge in finance and higher numeracy. Note, however, that we are primarily interested in whether finance professionals perform better in the task such that participants from the general population could benefit from delegating their investments. The question which factors explain the difference in *DMQI* is of secondary importance with respect to our research questions.

However, by construction of the experimental investment task, errors in decision-making are relatively less likely if the decision-maker is relatively more risk tolerant. For the sake of illustration, consider a risk neutral decision-maker: Choosing an allocation is straightforward as she will simply invest the entire endowment in the asset yielding the highest expected return. Instead, consider a highly risk averse decision-maker: In order to hedge against risk, the decision-maker has to choose a well-balanced portfolio of two or more assets using proper weights. Apparently, the likelihood of violating the principle of first order stochastic dominance (*FOSD*) and the generalized axiom of revealed preferences (*GARP*) is considerably larger for allocations lying in the interior of an opportunity set, as compared to allocations close to or on the boundary of the space. Thus, the superior decision-making quality of finance professionals could be driven—at least partly—by systematic differences in risk attitudes between pools rather than being the result of higher financial sophistication.

Table 3: Decision-making quality by subject pool. This table reports estimates from ordinary least squares regressions of the decision-making quality index (*DMQI*) on an indicator variable for finance professionals (*FP-OWN*), the mean standard deviation across the 25 decisions in the allocation decision task (*SD*; normalized to 1), and the interaction of *FP-OWN* and *SD*. Robust standard errors are reported in parentheses. * $p < 0.05$, ** $p < 0.005$.

	(1)	(2)	(3)
Subject Pool Indicator:			
FP-OWN	0.320**	0.297**	−0.051
	(0.082)	(0.091)	(0.176)
Controls:			
SD		0.219	−0.048
		(0.220)	(0.318)
Interaction Effects:			
FP-OWN × *SD*			0.766
			(0.390)
Constant:			
*GP-**	−0.077	−0.161*	−0.058
	(0.052)	(0.076)	(0.107)
F	15.303	10.789	11.294
$p > F$	0.000	0.000	0.000
Adj. R^2	0.011	0.011	0.014
Observations	687	687	687

Notes: The sample in the regressions includes all participants deciding on their own behalf, i.e., all participants from all general population treatments (*GP-**; $n = 550$) and finance professionals in treatment *FP-OWN* ($n = 137$).

Finance professionals in our sample, indeed, turn out being significantly less risk averse than participants from the general population: As illustrated in Figure 3, finance professionals' self-reported willingness to take risk exceeds the general population sample's willingness to take risk both in general (*FP-**: $m = 5.804$, $sd = 1.940$, *GP-**: $m = 4.8$, $sd = 2.1$; two-sample t-test: $t(956) = 7.557$, $p < 0.001$) as well as with respect to financial matters (*FP-**: $m = 6.1$, $sd = 2.1$, *GP-**: $m = 4.3$, $sd = 2.3$; two-sample

t-test: $t(956) = 12.259$, $p < 0.001$).[11] Participants' self-reported risk tolerance proves well-founded in the context of the experimental investment decisions. We find that finance professionals deciding on their own account (*FP-OWN*), on average, construct portfolios featuring significantly higher mean portfolio risk (*SD*) as compared to the general population sample (*FP-OWN*: $m = 0.5$, $sd = 0.3$; *GP-**: $m = 0.4$, $sd = 0.2$; two-sample t-test: $t(685) = 5.302$, $p < 0.001$). Consequently, the difference in *DMQI* between finance professionals and participants from the general population should be controlled for the systematic heterogeneity in risk-taking across the subject pools.

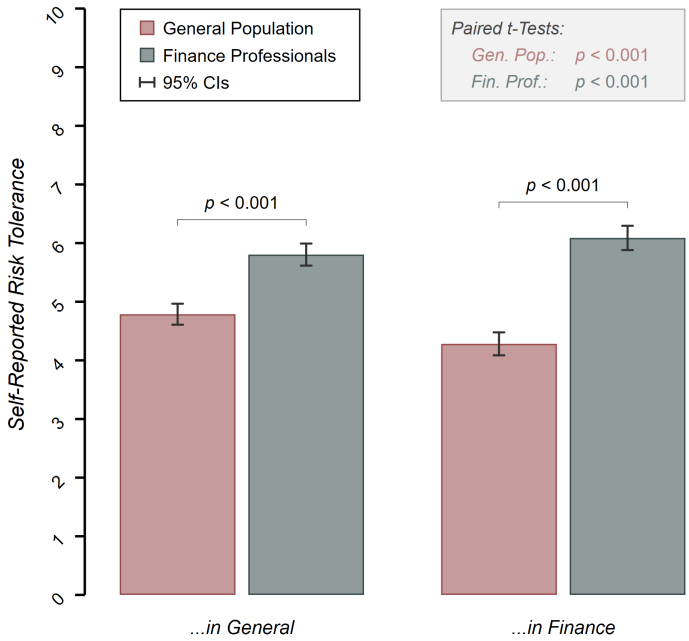

Figure 3: Self-reported risk tolerance. This figure depicts the mean levels of risk tolerance (self-reported on scales from 1 to 10) in general and in financial matters, respectively, separated for the general population sample and the finance professionals sample. p-values reported above the bars indicate comparisons between the subject pools and are based on two-sample t-tests (n = 958). p-values reported in the box are based on paired-sample t-tests with sample sizes of 550 (general population) and 408 (finance professionals), respectively.

Adjusting the difference in *DMQI* between subject pools for the variation in investment risk attenuates the effect only to a minor extent (less than 10%): Finance professionals in *FP-OWN*, on average, outperform the general population sample in terms of decision-making quality even when the potential heterogeneity in the vulnerability to poor decisions due to varying degrees of risk-taking is factored into the analysis (model (2) in Table 3; *FP-OWN*: $b = 0.320$, $p < 0.001$; $n = 687$). Investigating the difference-in-difference effect of risk-taking on *DMQI*, however, reveals a notable pattern. To illustrate risk-adjusted differences in *DMQI* between subject pools, we show the linear predictions of *DMQI* conditional on the mean portfolio risk—based on ordinary least squares regressions of *DMQI* on an indicator variable for finance professionals deciding on their own account (*FP-OWN*), the portfolio risk (*SD*), and the interaction of *FP-OWN* and *SD*—

[11] As an aside, we report that finance professionals self-report to be more risk tolerant in financial matters than in general matters (paired t-test: $t(407) = 3.995$, $p < 0.001$); participants in the general population sample indicate a reversed pattern, self-reporting to be less risk tolerant in financial matters (paired t-test: $t(549) = 8.1785$, $p < 0.001$).

in panel (A) of Figure 4. The corresponding regression model estimates are reported in column (3) in Table 3. As illustrated in the figure, *DMQI* increases with *SD* for finance professionals in *FP-OWN*, but does not vary with portfolio risk for the general population sample. This indicates that *GP-** participants fail to circumvent subpar decisions even if they do not shy away from risk. Notably, however, we find that finance professionals (investing on their own account) do not score better in terms of *DMQI* than participants from the general population when they are risk averse; they only outperform the *GP-** sample for moderate to high levels of portfolio risk. Put differently, if a decision-maker's risk attitude dictates her to choose a well-balanced (low-risk) portfolio, finance professionals are just as prone to poor decisions as participants from the general population; however, if a decision-maker's risk tolerance allows her to construct more risky portfolios, finance professionals are less susceptible to error.[12]

3.2. Investing on Behalf of Others

Given our findings on decision-making quality, two questions naturally arise in the context of delegated investment decisions: (i) Does finance professionals' decision-making quality systematically vary depending on whether they decide on behalf of clients or on their own account, and (ii) does the incentive scheme agents face affect the susceptibility to errors in their decision-making? To address these questions, we replicate the analysis reported above (focusing on *FP-OWN* vs. *GP-**) for the two treatments in which finance professionals invest on behalf of clients (*FP-ALIGNED* and *FP-FIXED*).

Panels (B) and (C) in Figure 4 show the linear predictions of finance professionals' *DMQI*, conditional on the mean portfolio risk (*SD*), in treatments *FP-ALIGNED* and *FP-FIXED*, respectively, as compared to the linear predictions of the general population sample's *DMQI*. As opposed to finance professionals deciding on their own account (*FP-OWN*; panel A), finance professionals do not significantly outperform participants from the general population when deciding on behalf of clients, irrespective of the mean portfolio risk. This suggests that—from the viewpoint of clients—there seems to be little scope to improve by means of delegating one's investment decisions to finance professionals.

At the first glance, these results appear to deliver a damning indictment of delegated investment decisions. Recall that finance professionals deciding on their own account (*FP-OWN*) only outperform participants from the general population in terms of decision-making quality for at least moderate levels of risk tolerance. Furthermore, recall that finance professionals in our sample, on average, are significantly more risk-tolerant than clients. The finding that finance professionals deciding on behalf of clients do not perform systematically better than participants from the general population, thus, might be the result of agents adapting to clients' requested risk profile. Evidence on the efficiency of risk communication in delegated investment decisions will be discussed in detail below. Yet, other potential mechanisms might be at play and could explain why finance professionals do not perform better than participants from the

[12] Note that the heterogeneous effect of portfolio risk (*SD*) on decision-making quality appears to be driven—at least partly—by correlational patterns between *SD* and our proxies of financial sophistication, i.e., numeracy and financial literacy. In particular, we report significantly positive correlations between *SD* and numeracy ($\rho = 0.196$, $p < 0.021$; $n = 137$) and financial literacy ($\rho = 0.336$, $p < 0.001$; $n = 137$) among finance professionals investing on their own account, whereas correlations turn out being considerably smaller and statistically insignificant for the *GP-** sample (*SD* vs. numeracy: $\rho = 0.082$, $p = 0.054$; *SD* vs. financial literacy: $\rho = 0.049$, $p = 0.251$; $n = 550$).

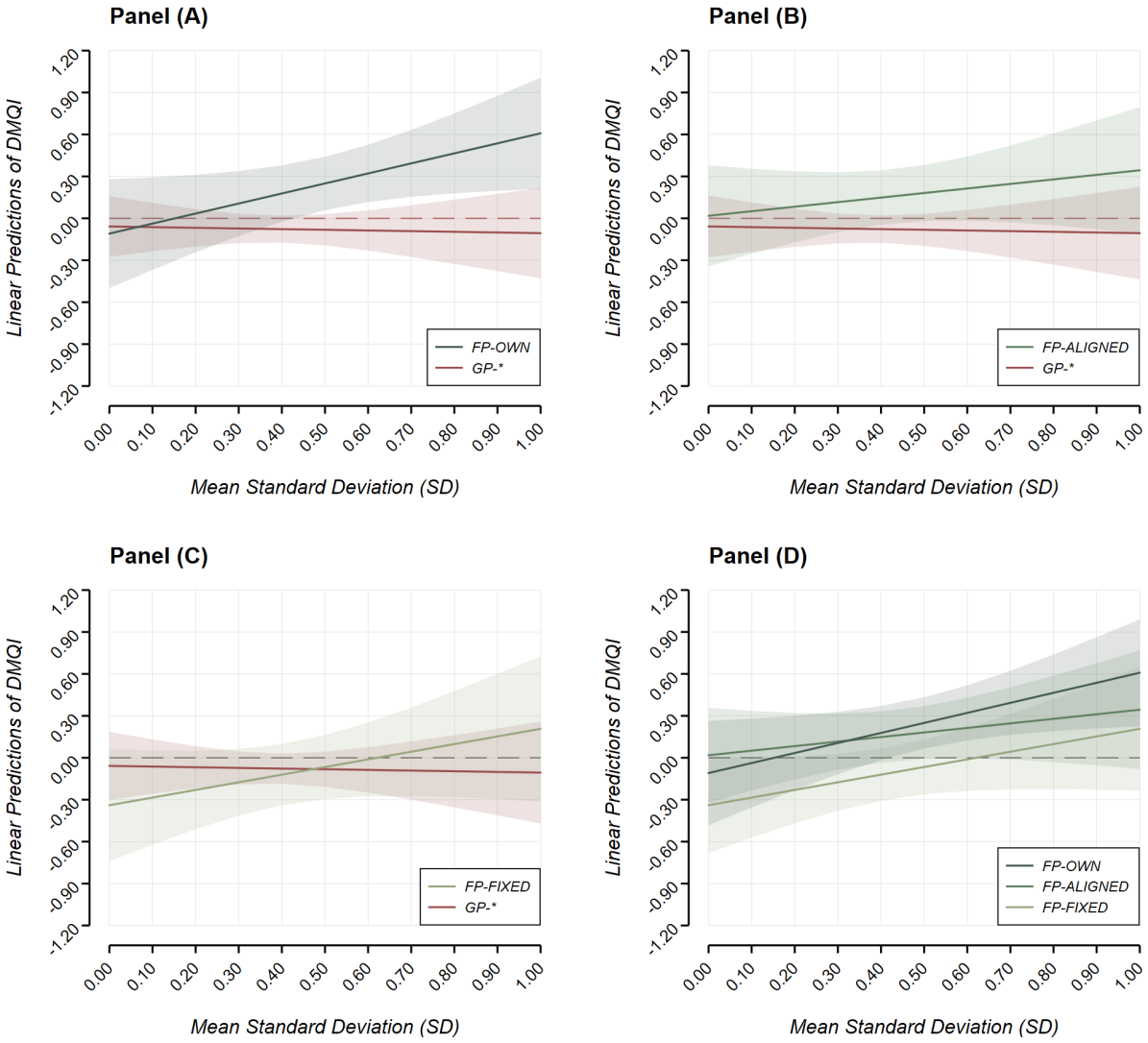

Figure 4: Decision-making quality (*DMQI*) conditional on portfolio risk. This table shows linear predictions of the decision-making quality index (*DMQI*) conditional on the mean standard deviation (*SD*; normalized to 1) across the 25 items in the allocation decision task (normalized to 1) after ordinary least squares regressions with robust standard errors (see Tables 3 and 4. **Panels (A)–(C).** Predictions of *DMQI* separated for the general population sample (pooled across treatments) and finance professionals deciding on their own behalf (*FP-OWN*), finance professionals deciding on behalf of clients receiving a flat payment (*FP-FIXED*), and finance professionals deciding on behalf of clients facing aligned incentives (*FP-ALIGNED*), respectively. **Panel (D)** Predictions of finance professionals' *DMQI* separated for the three treatments *FP-FIXED*, *FP-ALIGNED*, and *FP-OWN*.

general population when deciding on behalf of clients: (i) agents might put forth less of an effort when deciding on behalf of clients *per se*, and (ii) the incentive scheme an agent faces could further reduce agents' willingness to endeavor the best possible investment decisions.

To examine (i) whether finance professionals make less of an effort when deciding on behalf of others (as compared to when investing on their own account), we consider participants' response times to the 25 investment decisions in the experiment as a proxy. We report detailed analyses on decision times across subject pools, treatments, and tasks in Appendix G. With respect to the time spent on the investment decisions, we observe that finance professionals, on average, take *more* time when deciding on behalf of

clients as compared to when deciding on their own behalf, and as compared to clients' own decisions (see Tables G1 and G2 for details). These results suggest that finance professionals try hard to meet clients' expectations (e.g., their desired risk levels). Yet, we report that the time spent per decision does not significantly impact *DMQI*, neither among the general population, nor the finance professionals sample (see Appendix G for details).

Table 4: Finance professionals' decision-making quality by treatments. This table reports estimates from ordinary least squares regressions of the decision-making quality index (*DMQI*) on indicator variable for the treatments *FP-ALIGNED* and *FP-OWN*, the mean standard deviation across the 25 decisions in the allocation decision task (*SD*; normalized to 1), and the interaction of *FP-OWN* and *SD*. Robust standard errors are reported in parentheses. * $p < 0.05$, ** $p < 0.005$.

	(1)	(2)	(3)
Treatment Indicators:			
FP-ALIGNED	0.269	0.263	0.358
	(0.152)	(0.153)	(0.255)
FP-OWN	0.356*	0.315*	0.230
	(0.146)	(0.149)	(0.235)
Controls:			
SD		0.531*	0.547
		(0.192)	(0.449)
Interaction Effects:			
FP-ALIGNED × *SD*			−0.221
			(0.542)
FP-OWN × *SD*			0.171
			(0.503)
Constant:			
FP-FIXED	−0.112	−0.333*	−0.340
	(0.132)	(0.135)	(0.189)
F	2.976	6.247	4.735
$p > F$	0.052	0.000	0.000
Adj. R^2	0.014	0.029	0.026
Observations	408	408	408

Finally, we test (ii) whether finance professionals' *DMQI* differs systematically between treatments. Column (1) in Table 4 reports the coefficient estimates of an ordinary least squares regression of *DMQI* on indicator variables for treatments *FP-ALIGNED* and *FP-FIXED*. As compared to investing on a client's account and receiving a flat payment (*FP-FIXED*), finance professionals tend to perform better when deciding on their own account (*FP-OWN*; $b = 0.356$, $p = 0.015$). The difference in *DMQI* between treatments *FP-FIXED* and *FP-ALIGNED*, however, turns out not to be statistically different from zero (*FP-ALIGNED*: $b = 0.0269$, $p = 0.078$). We further report that the decision-making quality of finance professionals does not significantly differ between the treatments *FP-ALIGNED* and *FP-OWN* (Wald test: $F(1, 405) = 0.750$, $p = 0.386$). Adjusting the treatment differences for portfolio risk (*SD*) leaves the coefficient estimates effectively unchanged (see model (2) in Table 4). The significant difference in *DMQI* between *FP-FIXED* and *FP-OWN* and

the lack of evidence for systematic differences in the comparisons *FP-FIXED* vs. *FP-ALIGNED* and *FP-ALIGNED* vs. *FP-OWN* leave us with indecisive evidence as to whether deciding on behalf of others negatively impacts professionals' decision-making quality. Yet, it appears appealing to embark on the interpretation that finance professionals tend to perform better when deciding on their own account as compared to when investing on behalf of clients, but that aligned incentives tend to alleviate the difference (as compared to rewarding agents with a fixed payment).

Overall, these results seem to be well compatible with the construal that finance professionals strive to customize their portfolio allocation to the risk profile requested by clients (which is, on average, lower than the risk level professionals would implement when investing for themselves), resulting in inferior investment decisions (as compared to the *DMQI* levels professionals would be capable to achieve when deciding based on their own preferences). Anyhow, the question whether finance professionals actually succeed in meeting their clients' preferences in terms of portfolio risk is yet to be answered.

3.3. Risk Communication in Delegated Decisions

Notwithstanding that delegation appears to offer little scope for improving investment decisions (evaluated in terms of *DMQI*), delegation could still be effective for the purpose of providing clients an opportunity to adjust the risk profile of their investment. Gennaioli et al. (2015) argue that principals prefer to hire money managers over investing on their own account because they are too anxious to decide on risky positions themselves, and hypothesize that the intention to increase risk-taking is one of clients' motives to delegate their investments to "money doctors." In our companion paper (Holzmeister et al., 2021), which is based on the same data set but focuses on the demand side of delegated investments, we show that clients, on average, request the agent to take higher levels of risk as compared to the perception of risk they took in their own investment decisions. Two questions that arise naturally are (i) whether clients and agents succeed in communicating risk profiles and (ii) whether clients eventually get what they ask for.

To address these questions, we start by plotting the frequencies of finance professionals' risk perception of the portfolios constructed in treatments *FP-ALIGNED* and *FP-FIXED* ($m = 2.5$, $sd = 1.0$) over the risk profiles they are asked to implement on the client's account ($m = 2.5$, $sd = 1.1$).[13] As illustrated in Figure 5, finance professionals, on average, strive to customize their investment decisions in such a way that the clients' requested risk level is met. In particular, about two thirds of the finance professionals indicate that they consider their investments to exactly match the risk profile requested by potential clients; 21% (13%) of the agents deviate from the requested risk level and indicate that they have taken more (less) risk. Although, on average, agents' risk perception turns out being marginally higher than the principals' desired risk level (paired t-test: $t(270) = 1.994$, $p = 0.0472$), finance professionals seem to comply with what they are asked for in their role as "money doctors." This analysis, however, is based on agents' risk *perception*. Since the perception of actual investment risk may vary between professionals and laypeople, these results do not allow inferring whether clients' indeed get what they ask for.

[13] Recall that finance professionals in the treatments *FP-ALIGNED* and *FP-FIXED* were randomly assigned one of four risk profiles they are asked to comply with when investing on the account of clients. Principals and agents were matched based on the treatment assignment *and* the risk profile only after the experiment has been completed. Thus, the desired risk profiles are exogenous to the decisions of finance professionals.

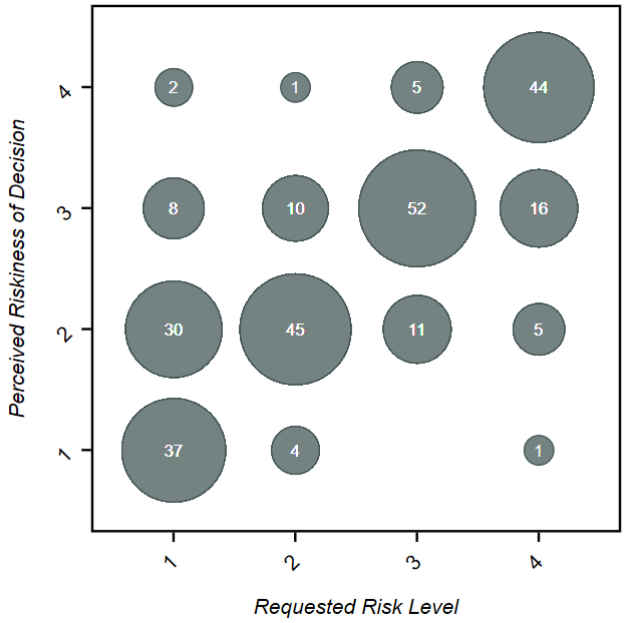

Figure 5: Desired risk conditional on risk perception. Risk level agents are asked to comply with when deciding on behalf of clients vs. agents' perception of the riskiness of their actual decisions ($n = 271$).

Recall that increasing the exposure to financial risk—and, thereby, increasing the expected return—has been argued to be one potential driver of delegating investments to money managers (Gennaioli et al., 2015). As discussed in detail in Holzmeister et al. (2021), we find empirical support for this claim by showing that clients that choose to delegate their investments, on average, ask the agent to take more risk as compared to how risky they perceive their investments to be when deciding on their own account. In particular, we find that 67.7% and 70.8% of clients who choose to delegate their investments to a finance professional in the treatments *GP-FIXED* and *GP-ALIGNED*, respectively, ask the agent to take more risk than they perceive to have implemented themselves when investing on their own account. Restricting our attention to those clients who seek to increase their risk-taking allows us to examine whether clients indeed get what they want. For those participants requesting a higher level of risk, delegation does indeed significantly increase portfolio risk: Regressions of the portfolio standard deviation in each of the 25 investments on an indicator variable for agents reveal that clients, on average, end up with significantly riskier positions when requesting a higher risk level to be implemented by a finance professional (as compared to the portfolio risk implied by their own decisions; *GP-FIXED*: $b = 0.133$, $t(141) = 3.299$, $p = 0.001$; *GP-ALIGNED*: $b = 0.137$, $t(152) = 3.311$, $p = 0.001$).[14] Figure H1 in Appendix H illustrates the results of a supplemen-

[14] The analyses are based on restricted samples: in particular, we include all observations of clients who choose to delegate *and* request a higher risk level as compared to the self-rated risk level of their own investments. When comparing portfolio risk of clients' own decisions to the risk of portfolios constructed by agents, we restrict the sample of finance professionals based on the incentives they face (i.e., *FP-FIXED* or *FP-ALIGNED* when focusing on clients in *GP-FIXED* or *GP-ALIGNED*, respectively). The samples, thus, comprise $142 \times 25 = 3{,}550$ and $153 \times 25 = 3{,}825$ observations in the analyses for the fixed payment and aligned incentive scheme, respectively. The analyses are based on ordinary least squares regressions of the portfolio standard deviation on an indicator for finance professionals, controlling for indicators variables for risk levels 2 to 4; standard errors are clustered on the individual level. As such, the coefficient estimates can be interpreted as the average difference in portfolio risk between clients' own decisions and the mean portfolio risk implemented by agents they could potentially be matched to.

tary analysis on the participant level, comparing the portfolio risk of clients' own investment decisions to the average risk of investments by agents deciding on behalf of clients (conditional on the requested risk profile) for those participants who request a higher risk level when delegating their decision to the agent. We observe that clients who seek to increase their risk exposure when delegating can expect to end up with significantly riskier positions by means of delegation in about 90% of the cases; the difference in portfolio risk between clients' own decisions and the average portfolio risk of potential matching partners is statistically significantly higher for about 60% of the observations. Taken together, these results indicate that finance professionals indeed try to adhere to the clients' requests in terms of risk exposure such that delegation eventually can be an effective means for clients to adjust their risk profile.

Revisiting risk perceptions, however, reveals a formidable challenge for delegated decision-making—risk communication. Figure 6 shows the cumulative distributions and the corresponding boxplots of portfolio risk (i.e., the mean standard deviation of the 25 allocation decisions in the investment task) conditional on the self-ascribed riskiness of participants' investments (risk profiles 1 to 4); we separate portfolios constructed by the general population sample (pooled across all treatments) from the portfolios constructed by agents deciding on behalf of clients (i.e., finance professionals in treatments *FP-FIXED* and *FP-ALIGNED*). Kolmogorov-Smirnov tests indicate that the distributions of portfolio risk associated with the four risk profiles differ significantly between the general population and the finance professional sample (see the associated p-values of Kolmogorov-Smirnov tests in Figure 6). This finding suggests that finance professionals and laypeople differ sharply in their perception of risk.[15] For risk profiles 1 and 2, the mean standard deviation (*SD*) of portfolios constructed by participants from the general population, on average, significantly exceeds the portfolio risk implemented by finance professionals deciding on behalf of clients. The difference in risk perceptions between the two subject pools, however, reverses for risk levels 3 and 4: the average standard deviation of portfolios constructed by finance professionals is significantly higher than the risk of clients' investment allocations associated with the same risk profiles. This result lends itself to the conclusion that clients compose more similar portfolios across the risk spectrum than professionals do, or, alternatively, that finance professionals tend to differentiate more explicitly between the various risk levels than laypeople. However, with respect to the question of whether a client's desired exposure to risk can be purposefully communicated when delegating investments to intermediaries, the discrepancy in the perception of risk between clients and finance professionals might pose a substantial hurdle. Whenever principals and agents do not sing from the same hymn sheet in terms of what is perceived risky and how to classify risk, clients will have a hard time getting what they actually seek—at least in terms of the risk associated with their investments.

Finally, we investigate the "outcome" of risk communication by examining the risk of portfolios constructed by agents on behalf of clients conditional on requested risk level of clients. Figure 7 illustrates the average standard deviation of agents' investment decisions on a client's account, separated for the risk profiles finance professionals are asked to implement as per the client's request. Apparently, the mean portfolio risk increases monotonically across the four risk profiles such that a client requesting a higher

[15] This finding contrasts the results by Holzmeister et al. (2020), who do not find evidence for systematic differences in risk perceptions between financial professionals and laypeople. Note, however, that the experimental setups used and the research questions addressed are not directly comparable between the two studies. In particular, Holzmeister et al., 2020 focus on potential differences in the *determinants* of risk perception based on a non-incentivized survey asking respondents to judge the risk of static return distributions, whereas participants in our experiment are required to self-assess their investment decisions in terms of riskiness.

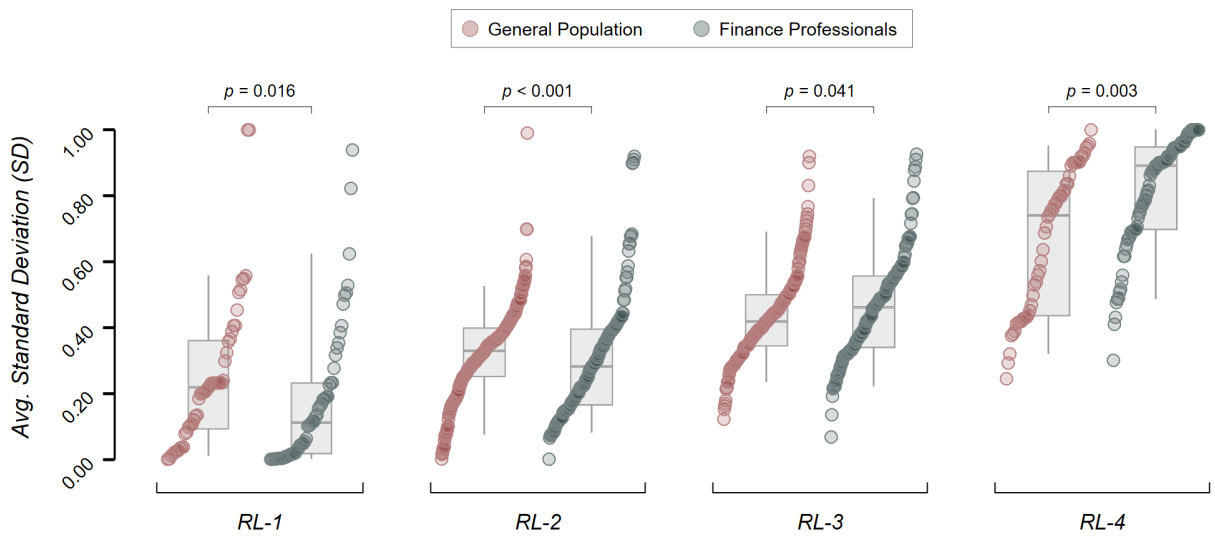

Figure 6: Portfolio risk conditional on self-ascribed risk profiles. This figure shows cumulative distributions and the corresponding boxplots of the mean standard deviation of the 25 allocations in the investment task (normalized to 1), conditional on the perceived riskiness of their choices (*RL-1* ... *RL-4*), separated for the general population 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 $n_4 = 136$.

(lower) risk level will, on average, end up with a more (less) risky investment position.[16] Yet, the heterogeneity in the risk of the investments issued by the agents turns out to be remarkable. As demonstrated in Figure 7, we find that each of the four risk profiles spans almost the full range of the potential portfolio risk spectrum, involving that the distributions of portfolio risk vastly overlap for the different risk levels. This result implies that clients, indicating different levels of risk bearing abilities when delegating their investment decisions, might eventually end up with similar levels of portfolio risk. For example, 25% of the portfolios constructed by finance professionals for principals requesting risk profile 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 imply higher risk than 25% of the allocations constructed for risk level 3 / 4.

To further illustrate the severity of the overlaps in portfolio risk designed for different risk levels, we present a supplementary analysis in Figure H2 in Appendix H. In particular, we construct ten equally sized categories of portfolio risk—$[0.0, 0.1), [0.1, 0.2), \ldots, [0.9, 1.0]$—and tabulate the frequency distribution for each category over the four risk profiles. Moreover, for each of the ten risk classes, we calculate the Herfindahl-Hirschman-Index, a diversity measure defined as $HHI = \sum_k s_k^2$ with s_k denoting the share of portfolios associated with risk profile $k = \{1,2,3,4\}$. By construction of the index, *HHI* takes a minimum value of 0.25 (if $s_1 = \cdots = s_4 = 0.25$) and a maximum value of 1 (if $s_k = 1$ for either $k \in \{1,2,3,4\}$). While the *HHI* is as high has 0.79 and 0.86 for the lowest and highest risk classes, respectively, it turns out being as low as 0.29 for intermediate risk categories. This suggests that portfolio allocations close to the boundaries of the risk spectrum are unambiguously mapped onto the highest and lowest risk profiles, whereas portfolios in the interior of the set of feasible investment risks are not distinctly ascribed to one of

[16] An ordinary least squares regression of the mean standard deviation (*SD*) of allocations compiled by finance professionals investing on behalf of clients (i.e., *FP-ALIGNED* and *FP-OWN*; $n = 271$) on indicator variables for risk profiles 2 to 4 shows that the differences in *SD* between the four risk levels is statistically significant ($p < 0.001$) for all pairwise comparisons.

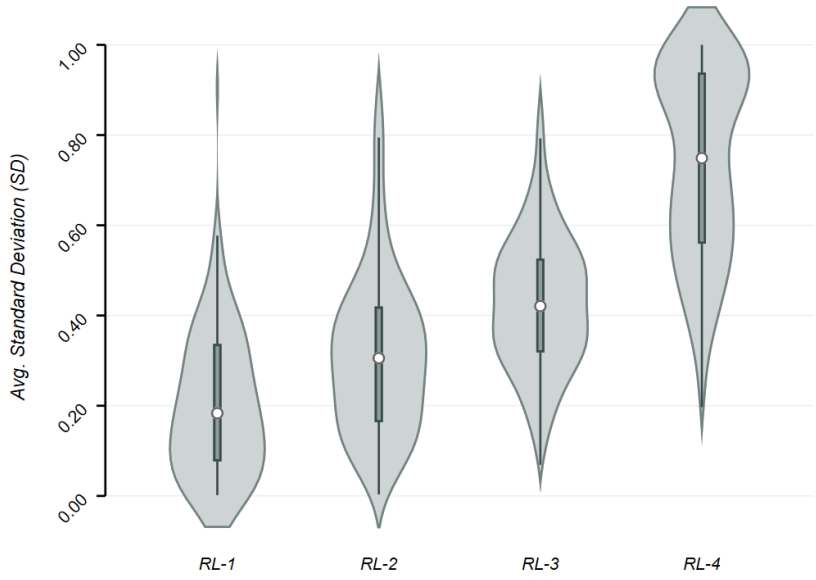

Figure 7: Portfolio risk conditional on clients' requested risk level. This figure illustrates the distribution (Gaussian kernel) of the average standard deviation of the 25 allocations in the investment task (normalized to 1), conditional on the risk level (*RL-1* ... *RL-4*) indicated by clients for treatments *FP-FIXED* and *FP-ALIGNED*. Sample sizes for *RL-1* through *RL-4* are $n_1 = 77$, $n_2 = 60$, $n_3 = 68$, and $n_4 = 66$.

the communicated risk levels. Overall, the figure makes clear that portfolios built for distinct risk profiles by finance professionals may render analogical in terms of the actual risk they involve. Put differently, two clients with sharply differing risk attitudes—thus requesting their agents to invest according to two distinct risk profiles—might well end up with portfolio allocations that are very similar in terms of the risk they involve. If so, at least one of the two clients is likely not to be pleased with the delegated decisions.

4. Conclusion

In this paper, we reported the results from a controlled online experiment with finance professionals (serving as agents) and participants from the general population in Sweden (acting as clients). All participants—i.e., both clients and advisers—completed a 25-item investment task, similar to the one introduced by Banks et al. (2018), which allows to quantify their decision-making quality. We varied whether finance professionals decide on their own behalf or on the account of clients, and, in the latter case, whether they face an aliened incentive scheme or receive a fixed compensation for investing on behalf of others. When investing on a client's account, agents were instructed to comply to one of four risk levels. Participants from the general population had the opportunity to delegate their investment to a finance professional. When delegating their decisions, clients indicated how much risk they want the agent to take on their behalf and were matched to a finance professional according to their requested risk profile.

We found that finance professionals, on average, show higher decision-making quality than participants from the general population only when deciding for themselves. Yet, the edge in decision-making quality

turned out being small to moderate in size since finance professionals only perform significantly better for moderate and high levels of portfolio risk. Notably, however, finance professionals did not outperform participants from the general population when investing on the account of clients, irrespective of the incentive scheme they face. Furthermore, we showed that finance professionals tried to customize their investments on behalf of clients in such a way that the resulting portfolio allocations comply with risk profiles requested by clients. However, differences in risk perception turned out to result in a risk communication problem: Portfolios constructed by finance professionals exhibited considerable overlaps in risk across the various risk profiles they were asked to comply with, such that clients eventually might have ended up with considerable more or less risk than what they asked for. Evidence suggested that this result has emerged from systematic differences in the perception of risk between finance professionals and participants from the general population.

Our study involves several potential limitations, as already indicated in our companion paper (Holzmeister et al., 2021). First, the investment decisions in our experiment are an abstraction of real-world investment choices, as are all economics and finance experiments and models. While the simplification and abstraction implied by our experimental design might limit the generalizability of our results, the upside of our approach is that we are able to identify empirically unobservable variables, of which the following are the most relevant ones for our research questions: in observational data it is difficult (i) to separate advised and delegated investments, (ii) to measure decision-making quality, (iii) to attain direct evidence on the extent to which investments are customized to the client's needs, and (iv) to investigate the causal effect of agents' incentive schemes on investment decisions on behalf of clients. Another potential limitation is that the experimental incentives, as compared to salaries of finance professionals, could be considered relatively low. However, in terms of the stake size of incentives, our study joins the ranks of previous research with finance professionals (e.g., Haigh and List, 2005; Alevy et al., 2007; Kirchler et al., 2018; Weitzel et al., 2019). Anyhow, the high quality of the experimental data (e.g., reasonable amount of time spent, high decision-making quality, and low number of outliers) suggests that the participants took the experiment seriously. Furthermore, our results should be considered in light of a potential sample selection bias. Some of the differences in socio-demographic characteristics between the participants in our experiment and those that were invited but did not take part turn out to be statistically significant (see Table E1 in Appendix F). For both subject pools, the number of participants with a University degree is clearly higher in our sample than among those that did not participate – even though numeracy and financial literacy is significantly higher among finance professionals. The high level of education in our sample could explain the high levels of decision-making quality among the general population sample (see Appendix C for details). However, this sample selection bias does not necessarily impair our results' policy relevance, since real-world markets for delegation and advice are affected by self-selection too. Private investors are usually a biased sample from the general population, with males, older, and wealthier people with higher financial sophistication being over-represented (see, e.g. Collins, 2012; Hackethal et al., 2012; Calcagno and Monticone, 2015). Eventually, given the limitations of our study, we call for caution in generalizing the results presented in this paper.

We deem our findings instructive for future research and informative for policy discussions regarding financial advise alike. First and foremost, we identify a substantial problem in risk communication, which is at the core of delegatory and advisory services in finance. Despite its apparent significance, to the best of our knowledge, the question whether risk can be properly communicated between agents and clients has

received almost no consideration in research. Noteworthy exceptions are the discussion of client-adviser matching in Rose (2021) and, in particular, the study by Kling et al. (2019), identifying miscommunication of risk between agents and clients. The hurdle imposed by miscommunication is particularly relevant for policy making and regulatory requirements, where, so far, the main focus has been on the processing and the disclosure of information as well as the elicitation of clients' investment preferences. However, given that the perception of risk can differ sharply between clients and agents, it remains unclear as for how these regulatory requirements can actually improve the efficiency of delegated and advised investments. For this reason, we are convinced that more research on the pitfalls of risk communication is needed and that risk perceptions should receive more attention (see, e.g. Holzmeister et al., 2022) to better understand which factors drive the efficiency of delegated and advised financial investments. Moreover, tackling potential miscommunication might also help to increase clients' trust in delegatory and advisory services and could strengthen the perception that agents indeed act in the client's interest, something—good news—finance professionals already do in our experiment.

References

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.

Bachmann, K. & Hens, T. (2015). Investment competence and advice seeking. *Journal of Behavioral and Experimental Finance*, *6*, 27–41.

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

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.

Bluethgen, R., Gintschel, A., Hackethal, A., & Mueller, A. (2008). Financial advice and individual investors' portfolios. *Working Paper*.

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.

Bucciol, A., Cavasso, B., & Zarri, L. (2019). Can Risk-Averse Households Make Risky Investments? The Role of Trust in Others. *The Scandinavian Journal of Economics*, *121*(1), 326–352.

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.

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.

Collins, J. M. (2012). Financial advice: A substitute for financial literacy? *Financial Services Review*, *21*(4), 307–322.

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.

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*.

Egan, M., Matvos, G., & Seru, A. (2019). The market for financial adviser misconduct. *Journal of Political Economy*, *127*(1), 233–295.

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*.

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.

Hackethal, A., Haliassos, M., & Jappelli, T. (2012). Financial advisors: A case of babysitters? *Journal of Banking and Finance*, *36*(2), 509–524.

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.

Holmen, M., Holzmeister, F., Kirchler, M., Stefan, M., & Wengström, E. (n.d.). Economic preferences and personality traits among finance professionals and the general population. *SSRN Working Paper; available at https://ssrn.com/abstract=3779944*.

Holzmeister, F., Holmen, M., Kirchler, M., Stefan, M., & Wengström, E. (2019). Delegated Decision-Making in Finance. *SSRN Working Paper; available at https://ssrn.com/abstract=3498890*.

Holzmeister, F., Holmen, M., Kirchler, M., Stefan, M., & Wengström, E. (2021). Delegation decisions in finance. *SSRN Working Paper; available at https://papers.ssrn.com/abstract=4018656*.

Holzmeister, F., Huber, C., & Palan, S. (2022). A critical perspective on the conceptualization of risk in behavioral and experimental finance. In S. Füllbrunn & E. Haruvy (Eds.), *Handbook of Experimental Finance*. Edward Elgar Publishing.

Holzmeister, F., Huber, J., Kirchler, M., Lindner, F., Weitzel, U., & Zeisberger, S. (2020). What drives risk perception? A global survey with financial professionals and lay people. *Management Science*, *66*(9), 3977–4002.

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

Inderst, R. & Ottaviani, M. (2012b). 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.*

Kramer, M. (2012). Financial advice and individual investor portfolio performance. *Financial Management, 41*(2), 395–428.

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

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

Rose, J. (2021). Same Same, *not* different: Client-Advisor Matching in the Finance Industry. *Working Paper.*

Stolper, O. (2018). It takes two to tango: Households' response to financial advice and the role of financial literacy. *Journal of Banking and Finance, 92*, 295–310.

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.

Appendices

for Online Publication

You can't always get what you want—An experiment on finance professionals' decisions for others

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

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

[*] *Corresponding author: matthias.stefan@uibk.ac.at*

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/zhnj5/. 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/wej2k/. 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/bxhju/. 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.[17]

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 participant 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. 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 register data from *SCB* for each participant who completed all tasks in the experiment. Since they are relevant for Holzmeister et al., 2021 and Holmen et al., n.d., please refer to these two papers for a full list and discussion of register variables. In the paper at hand, we are interested in potential difference between financial professionals and the general population, but not where these differences originate from (for a more detailed discussion see Holmen et al., n.d.). This is why we disregard the participants' register data in this paper.

[17] Please note that the register data obtained from *SCB* may not be publicly shared.

Figure A1: Screenshots of the allocation decision task. The figure shows screenshots of the main experimental task as displayed to participants. 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.

B. Allocation Decision Task

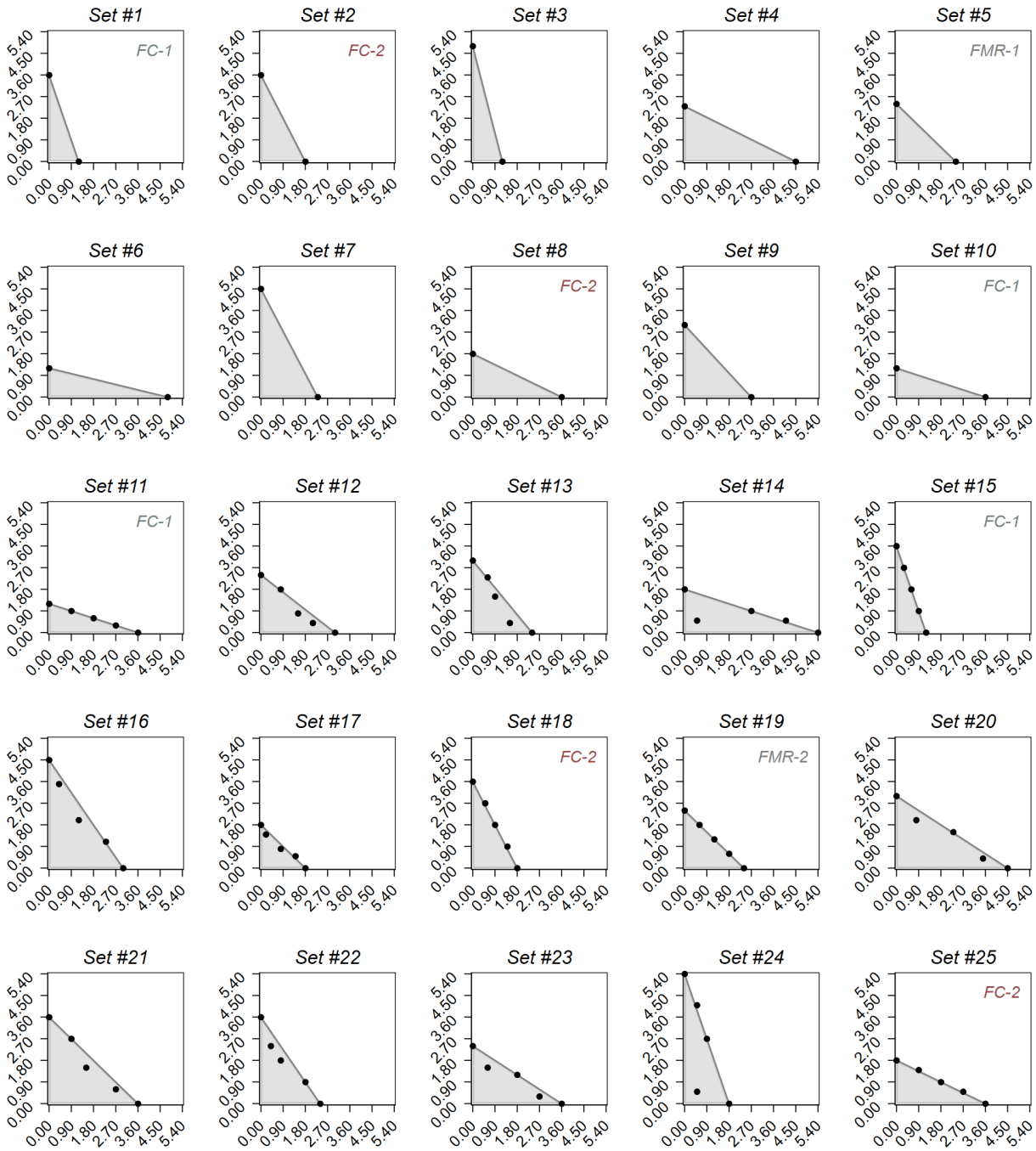

Figure B1: Opportunity sets in the allocation decision task. In each panel of this figure, the vertical (horizontal) axis indicates the return per 1 SEK invested if the coin shows up heads (tails). Each dot indicates a single asset. The labels *FC-1*, *FC-2*, *FMR-1*, and *FMR-2* denote particular opportunity 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

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

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., $SD_i = {}^1\!/_{25} \cdot \sum_{j=1}^{25} SD_{i,j}$. To take into account that the various opportunity sets allow for different levels of standard deviation, we normalize the standard deviation to 1: $SD_i = {}^{SD_i}\!/_{\max(SD)}$.

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:

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

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., $\textit{FOSD}_i = 1/23 \cdot \sum_{j=1}^{23} \textit{FOSD}_{i,j}$.

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.,

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

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., $\textit{MPI}_i = 1/300 \cdot \sum_{m=1}^{25} \sum_{n=1}^{25} \textit{MPI}_{i,m,n} \; \forall \; m > n$.

Financial competence (*FC*). Four opportunity 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). In addition, two of these particular four opportunity sets 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 FC_1 and 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 $J_2 = \{\#2, \#8, \#18, \#25\}$. Thus, J_1 and J_2 are sets containing identical opportunity sets presented as mirror images in the two-asset- or the five-asset-frame, respectively. 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 portfolio that is fully hedged and, thus, has zero risk (second-order) dominates all other feasible portfolios in these two opportunity sets. The failure to minimize risk for participant 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*). The four measures address different aspects of decision making quality. We use the first principal component of the four measures as a proxy of decision-making quality (*DMQI*) for each participant (please note that this approach differs from the one in Banks et al., 2018). The underlying principal component analysis is summarized in Table C1. By construction, *DMQI* has 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.

The four measures tend to be correlated, as can be seen in Table C2 reporting Pearson correlation coefficients between *FOSD*, *MPI*, *FC*, and *FMR*, separated for the general population and finance professional sample. The study by Banks et al. (2018), conducted with participants from the general population, reports correlation coefficients of comparable magnitude.

Table C1: Principal component analysis of the four decision-making quality measures. This table outlines the four principal components of the *DMQI*, i.e., (i) violations of first order stochastic dominance (*FOSD*), (ii) money pump index (*MPI*), (iii) financial competence, and (iv) failure to minimize risk (*FMR*); $n = 958$. **Panel (A)** reports the eigenvalue and the proportion of explained variance for each of the four components. **Panel (B)** reports unrotated components.

Panel (A)

Component	*Eigenvalue*	*Difference*	*Proportion*	*Cumulative*
Comp. #1	2.034	1.046	0.509	0.509
Comp. #2	0.989	0.279	0.247	0.756
Comp. #3	0.710	0.444	0.178	0.933
Comp. #4	0.267	.	0.067	1.000

Panel (B)

Component	*Comp. #1*	*Comp. #2*	*Comp. #3*	*Comp. #4*
FOSD	0.621	−0.099	−0.318	0.710
MPI	0.618	0.026	−0.360	−0.698
FC	0.457	−0.228	0.859	−0.048
FMR	0.154	0.968	0.179	0.081

Table C2: Correlations between the decision-making quality measures. This table reports Pearson correlations between the decision-making quality measures (i) violations of first order stochastic dominance (*FOSD*), (ii) money pump index (*MPI*), (iii) financial competence (*FC*), (iv) failure to minimize risk (*FMR*), as well as the unified measure of decision-making quality (*DMQI*), separated for the general population sample (lower triangular matrix; $n = 550$) and the finance professionals sample (upper triangular matrix; $n = 408$). p-values are reported in parentheses.

	FOSD	*MPI*	*FC*	*FMR*	*DMQI*
FOSD		0.776	0.334	−0.084	0.740
		(< 0.001)	(< 0.001)	(0.090)	(< 0.001)
MPI	0.703		0.327	0.075	0.871
	(< 0.001)		(< 0.001)	(0.131)	(< 0.001)
FC	0.418	0.379		−0.131	0.685
	(< 0.001)	(< 0.001)		(0.008)	(< 0.001)
FMR	0.281	0.261	0.240		0.222
	(< 0.001)	(< 0.001)	(< 0.001)		(< 0.001)
DMQI	0.766	0.861	0.762	0.451	
	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	

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.[18] 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 participants' 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 likely 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 participants' 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,

[18] 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.[19] 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 F1 in Appendix F.

[19] 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 participants 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* ○ exactly 2,000 SEK ○ less than 2,000 SEK ○ 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 ○ exactly the same ○ *less than today* ○ 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 ○ *exactly the same* ○ less than today ○ do not know
Q4	When an investor spreads his money among different assets, does the risk of losing money in general:	○ increase ○ *decrease* ○ stay the same ○ do not know
Q5	If the interest rate falls, what should happen to bond prices?	○ *rise* ○ fall ○ stay the same ○ none of the above ○ 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* ○ he has lent money to firm B ○ he is liable for the firm B's debt ○ none of the above ○ do not know

E. Response Rate Analysis

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 χ^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.

For the finance professionals group, an analysis of response rates shows that men responded to a greater extent than women, and that finance professionals 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 extent as 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, 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 degree 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.

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. χ^2-tests (with k-1 degrees of freedom) and the corresponding p-values are reported.

	General Population			*Finance Professionals*		
	Resp.	*No Resp.*	χ^2 / p	*Resp.*	*No Resp.*	χ^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	

F. Descriptive Results

In the following, we present a set of descriptive results for all (self-reported) measures elicited in the experiment. With the exception of risk tolerance, none of these variables enter the analyses presented in this paper; rather these variables are considered as covariates in the analyses of the demand side of delegated investment decisions presented in our companion paper (Holzmeister et al., 2021). Yet, a comparison of the two subject pools—participants from the general population and professionals from the finance industry—along these measures is interesting in itself as it sheds further light on the differences between the two samples in our experiment.

Table F1: 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/Hypothetical 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 Financial Matters	5.60	6.85	−1.249**
	(3.67)	(3.33)	[0.231]
Use Expertise of Finance Professionals	3.58	2.21	1.376**
	(3.19)	(2.81)	[0.214]
Delegate to Finance Professionals	3.36	1.32	2.039**
	(3.11)	(2.16)	[0.192]
Delegate to Investment 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.

G. 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 G1. 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 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 G2.

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 participant 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 participants' proneness to poor investment decisions is not significantly driven by the (average) time they spend on each decision, neither in the general population sample ($b = 0.005$, $t(548) = 1.314$, $p = 0.189$, $n = 550$), nor in the finance professionals sample ($b = 0.003$, $t(406) = 1.261$, $p = 0.208$, $n = 408$).

Table G1: 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 G2: 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 G1. * $p < 0.05$, ** $p < 0.005$.

	*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

H. Supplementary Analyses

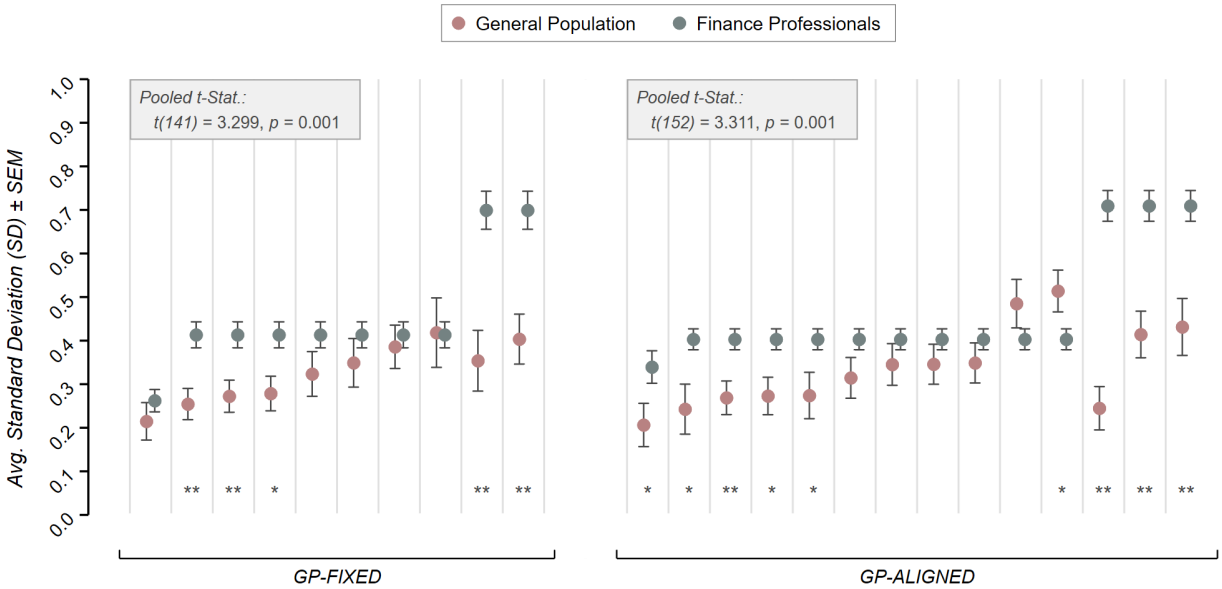

Figure H1: Portfolio risk of clients asking the agent to take more risk when delegating. This figure shows the average portfolio risk across the 25 investment decisions (*SD*; normalized to 1) of those clients that choose to delegate *and* ask the agent to take more risk than they believe they took in their own decisions, separated for the treatments *GP-FIXED* and *GP-ALIGNED* (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 on the 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 profile 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 H2: 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 $s_1 = \cdots = s_4 = 0.25$) and a maximum value of 1 (if $s_k = 1$ for either $k \in \{1,2,3,4\}$).

University of Innsbruck - Working Papers in Economics and Statistics
Recent Papers can be accessed on the following webpage:

https://www.uibk.ac.at/eeecon/wopec/

2022-02 **Matthias Stefan, Martin Holmén, Felix Holzmeister, Michael Kirchler, Erik Wengström:** You can't always get what you want–An experiment on finance professionals' decisions for others

2022-01 **Toman Barsbai, Andreas Steinmayr, Christoph Winter:** Immigrating into a Recession: Evidence from Family Migrants to the U.S.

2021-32 **Fanny Dellinger:** Housing Support Policies and Refugees' Labor Market Integration in Austria

2021-31 **Albert J. Menkveld, Anna Dreber, Felix Holzmeister, Jürgen Huber, Magnus Johannesson, Michael Kirchler, Sebastian Neusüss, Michael Razen, Utz Weitzel and et al:** Non-Standard Errors

2021-30 **Toman Barsbai, Victoria Licuanan, Andreas Steinmayr, Erwin Tiongson, Dean Yang:** Information and Immigrant Settlement

2021-29 **Natalie Struwe, Esther Blanco, James M. Walker:** Competition Among Public Good Providers for Donor Rewards

2021-28 **Stjepan Srhoj, Melko Dragojević:** Public procurement and supplier job creation: Insights from auctions

2021-27 **Rudolf Kerschbamer, Regine Oexl:** The effect of random shocks on reciprocal behavior in dynamic principal-agent settings

2021-26 **Glenn E. Dutcher, Regine Oexl, Dmitry Ryvkin, Tim Salmon:** Competitive versus cooperative incentives in team production with heterogeneous agents

2021-25 **Anita Gantner, Regine Oexl:** Respecting Entitlements in Legislative Bargaining - A Matter of Preference or Necessity?

2021-24 **Silvia Angerer, E. Glenn Dutcher, Daniela Glätzle-Rützler, Philipp Lergetporer, Matthias Sutter:** The formation of risk preferences throughsmall-scale events

2021-23 **Stjepan Srhoj, Dejan Kovač, Jacob N. Shapiro, Randall K. Filer:** The Impact of Delay: Evidence from Formal Out-of-Court Restructuring

2021-22 **Octavio Fernández-Amador, Joseph F. Francois, Doris A. Oberdabernig, Patrick Tomberger:** Energy footprints and the international trade network: A new dataset. Is the European Union doing it better?

2021-21 **Felix Holzmeister, Jürgen Huber, Michael Kirchler, Rene Schwaiger:** Nudging Debtors to Pay Their Debt: Two Randomized Controlled Trials

2021-20 **Daniel Müller, Elisabeth Gsottbauer:** Why Do People Demand Rent Control?

2021-19 **Alexandra Baier, Loukas Balafoutas, Tarek Jaber-Lopez:** Ostracism and Theft in Heterogeneous Groups

2021-18 **Zvonimir Bašić, Parampreet C. Bindra, Daniela Glätzle-Rützler, Angelo Romano, Matthias Sutter, Claudia Zoller:** The roots of cooperation

2021-17 **Silvia Angerer, Jana Bolvashenkova, Daniela Glätzle-Rützler, Philipp Lergetporer, Matthias Sutter:** Children's patience and school-track choices several years later: Linking experimental and field data

2021-16 **Daniel Gründler, Eric Mayer, Johann Scharler:** Monetary Policy Announcements, Information Schocks, and Exchange Rate Dynamics

2021-15 **Sebastian Bachler, Felix Holzmeister, Michael Razen, Matthias Stefan:** The Impact of Presentation Format and Choice Architecture on Portfolio Allocations: Experimental Evidence

2021-14 **Jeppe Christoffersen, Felix Holzmeister, Thomas Plenborg:** What is Risk to Managers?

2021-13 **Silvia Angerer, Daniela Glätzle-Rützler, Christian Waibel:** Trust in health care credence goods: Experimental evidence on framing andsubject pool effects

2021-12 **Rene Schwaiger, Laura Hueber:** Do MTurkers Exhibit Myopic Loss Aversion?

2021-11 **Felix Holzmeister, Christoph Huber, Stefan Palan:** A Critical Perspective on the Conceptualization of Risk in Behavioral and Experimental Finance

2021-10 **Michael Razen, Alexander Kupfer:** Can increased tax transparency curb corporate tax avoidance?

2021-09 **Changxia Ke, Florian Morath, Anthony Newell, Lionel Page:** Too big to prevail: The paradox of power in coalition formation

2021-08 **Marco Haan, Pim Heijnen, Martin Obradovits:** Competition with List Prices

2021-07 **Martin Dufwenberg, Olof Johansson-Stenman, Michael Kirchler, Florian Lindner, Rene Schwaiger:** Mean Markets or Kind Commerce?

2021-06 **Christoph Huber, Jürgen Huber, and Michael Kirchler:** Volatility Shocks and Investment Behavior

2021-05 **Max Breitenlechner, Georgios Georgiadis, Ben Schumann:** What goes around comes around: How large are spillbacks from US monetary policy?

2021-04 **Utz Weitzel, Michael Kirchler:** The Banker's Oath And Financial Advice

2021-03 **Martin Holmen, Felix Holzmeister, Michael Kirchler, Matthias Stefan, Erik Wengström:** Economic Preferences and Personality Traits Among Finance Professionals and the General Population

2021-02 **Christian König-Kersting:** On the Robustness of Social Norm Elicitation

2021-01 **Laura Hueber, Rene Schwaiger:** Debiasing Through Experience Sampling: The Case of Myopic Loss Aversion.

University of Innsbruck

Working Papers in Economics and Statistics

2022-02

Matthias Stefan, Martin Holmén, Felix Holzmeister, Michael Kirchler, Erik Wengström

You can't always get what you want–An experiment on finance professionals' decisions for others

Abstract
To study whether clients benefit from delegating financial investment decisions to an agent, we run an investment allocation experiment with 408 finance professionals (agents) and 550 participants from the general population (clients). In several between-subjects treatments, we vary the mode of decision-making (investment on one's own account vs. investments on behalf of clients) and the agents' incentives (aligned vs. fixed). We find that finance professionals show higher decision-making quality than participants from the general population when investing on their own account. However, when deciding on behalf of clients, professionals' decision-making quality does not significantly differ from their clients', neither when compensated with a fixed payment nor when facing aligned incentives. Our results further identify a considerable challenge in risk communication between agents and clients: While finance professionals tend to take into account principals' desired risk levels, the constructed portfolios by professionals show considerable overlaps in portfolio risk across different risk levels requested by principals. We argue that this result is due to differences in risk perception.

ISSN 1993-4378 (Print)
ISSN 1993-6885 (Online)