universität innsbruck

Research Platform
Empirical and Experimental Economics

Cognitive Skills and Economic Preferences in the Fund Industry

Adam Farago, Martin Holmén, Felix Holzmeister, Michael Kirchler, Michael Razen

Working Papers in Economics and Statistics

2019-16

University of Innsbruck
https://www.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.

Cognitive Skills and Economic Preferences in the Fund Industry

Adam Farago[†] Martin Holmén[†] Felix Holzmeister[‡]
Michael Kirchler[†,‡,*] Michael Razen[‡]

[†] University of Gothenburg, Department of Economics, Centre for Finance
[‡] University of Innsbruck, Department of Banking and Finance
[*] Corresponding author: michael.kirchler@uibk.ac.at

Abstract

By running a battery of experiments with fund managers, we investigate the impact of cognitive skills and economic preferences on their professional decisions. First, we find that fund managers' risk tolerance positively correlates with fund risk when accounting for fund benchmark, fund category, and other controls. Second, we show that fund managers' ambiguity tolerance positively correlates with the funds' tracking error from the benchmark. Finally, we report that cognitive skills do not explain fund performance in terms of excess returns. However, we do find that fund managers with high cognitive reflection abilities compose funds at lower risk.

JEL: C91, D91, G11, G41, J24
Keywords: Cognitive skills, economic preferences, fund managers, fund performance, experimental finance.

We thank Niclas Andren, Peter Bossaerts, Brice Corgnet, Matthias Heinz, Jürgen Huber, Matthias Stefan, Matthias Sutter, seminar participants of the C-SEB seminar at the University of Cologne, the Max Planck Institute for Research on Collective Goods in Bonn, Goethe University Frankfurt, Oslo Metropolitan University, University of Groningen, as well as conference participants of the NOEG 2019 (Austrian Economic Association) in Graz, Experimental Finance 2019 in Copenhagen, and the KWC-CFF 2019 workshop in Arild for valuable comments. We particularly thank all participating fund managers for the excellent collaboration and their enthusiasm. Financial support from the Austrian Science Fund FWF (START-grant Y617-G11 and SFB F63) and the Swedish Research Council (grant 2015-01713) is gratefully acknowledged. The software, including all instructions as used for the data collection, is available for download as a zipped *oTree* project at https://osf.io/dq3t8/ and as a live demo version via https://fea-2018-en.herokuapp.com.

1. Introduction

The mutual fund industry is at the heart of the finance sector and so are its main protagonists—fund managers. Two numerical examples of the fund industry's size document its relevance: in 2017, the value of total net assets of international open-end funds was $49.3 trillion and 114,000 open-end funds were registered, amounting to $2.7 trillion net sales.[1] Given the central role of fund managers in the finance industry, it is important to learn more about how their economic preferences and cognitive skills shape their professional decisions.

In this paper, we bridge two formerly loosely related, but important, strands of literature on fund management—namely, the one on fund managers' economic preferences and field behavior and the one on fund managers' cognitive skills and field behavior. We innovate by using well-established incentivized experiments to *directly measure* economic preferences and cognitive skills and combine these findings with *empirical fund data*. We find evidence that economic preferences and cognitive skills of fund managers correlate with fund dynamics and risk. We show that fund managers' risk tolerance positively correlates with fund risk and that fund managers' tolerance towards ambiguous outcomes is positively related to the funds' tracking error. Moreover, we show that cognitive abilities do not explain abnormal returns, but that fund managers with higher cognitive reflection skills construct funds with lower risk.

Our approach is similar to the idea of Fang and Wang (2015), who propose a three-tier framework to investigate how fund manager characteristics affect fund performance. They argue that a funds' *comprehensive performance* can be decomposed in two essential components, *excess return* and *risk*. In Table 1, we outline seminal studies on the impact of (i) economic preferences and/or (ii) cognitive skills on (a) performance and/or (b) risk among professionals and private investors. Most studies only focus on one particular aspect of economic preferences or cognitive skills and mainly rely on survey data or indirect measures.

As for non-professional market participants (e.g., private investors or students), there is a growing survey-based, experimental, and empirical literature investigating the role of *economic preferences* and its influence on investment decisions and portfolio returns (see Table 1). For instance, risk aversion has been shown to be negatively related to portfolio underdiversification (Dorn and Huberman, 2005), and ambiguity aversion tends to be positively correlated with portfolio risk (Bianchi and Tallon, 2018) and portfolio underdiversification (Dimmock et al., 2016b). *Cognitive skills* have also been shown to drive non-professional investors' decisions. For example, studies with private investors indicate that high-IQ investors earn higher Sharpe ratios (Grinblatt et al., 2011) and exhibit superior market timing and stock-picking skills as compared to low-IQ investors (Grinblatt et al., 2012). Moreover, various forms of cognitive skills (e.g., fluid intelligence, cognitive reflection abilities or a combination of various measures) predict traders' earnings and performance (Corgnet et al., 2018; Hefti et al., 2018; Noussair et al., 2016) in laboratory asset markets with students (see also Bruguier et al., 2010, for a related approach).

Turning to professional investors and the role of *economic preferences*, Bodnaruk and Simonov (2016) utilize non-incentivized questions on risk preferences and loss tolerance (see Table 1). They find that

[1] See the 2018 Investment Company Fact Book.

fund managers' self-reported risk preferences have no systematic effect on performance and portfolio risk, whereas managers with high levels of self-reported loss aversion construct mutual funds with lower downside risk and lower fund performance. In a natural field experiment with professional traders, Larson et al. (2016) detect myopic loss aversion also in traders' natural domain to negatively affect their performance. Thus, the question remains open whether these effects are robust to *direct* measures of economic preferences elicited in incentivized experiments, and whether there are correlations between risk-, ambiguity-, loss-tolerance and fund performance and risk, respectively.[2]

Turning to the effect of *cognitive skills* on decision-making of professional investors, existing studies have been resorting to various proxies. Fang and Wang (2015) report that fund managers with MBA or CFA degrees generate significantly higher excess returns and also achieve better comprehensive performance. Golec (1996) also documents a positive relation between fund performance and manager education. Similarly, Chevalier and Ellison (1999) report a positive relation between college SAT score and fund performance in the U.S. market.[3] Interestingly, there is lack of evidence regarding the relationship of cognitive skills and risk measures (e.g., Gastineau, 2004). Again, the question remains open whether *direct* measures of cognitive skills like fluid intelligence or cognitive reflection abilities correlate with mutual funds' performance and risk measures.

In this study, we innovate by using well-established tasks and incentivized experiments to *directly measure* cognitive skills and economic preferences and combine the resulting measures with *real-world fund data* from the same sample of fund managers. In particular, we address the following research questions:

RQ1: Do fund managers' cognitive skills like fluid intelligence, cognitive reflection, and theory of mind correlate with fund dynamics such as abnormal returns, fund risk, Sharpe Ratios, and tracking errors?

RQ2: Do fund managers' tolerance towards risk, losses, and ambiguity correlate with fund dynamics?

Our approach of combining experimentally elicited skills and preferences with real market data allows us to merge two important strands of literature among fund managers—i.e., (i) economic preferences and field behavior and (ii) cognitive skills and field behavior. Yet, we consider our approach exploratory,

[2] Additional and related evidence from experiments with finance professionals indicates that professionals often behave similarly as non-professionals. For instance, Sarin and Weber (1993) find that both students and experienced traders underprice an ambiguous asset in a small sample of experimental markets. They conjecture that ambiguous assets can cause psychological discomfort if the underlying stochastic process is unknown, combined with potential regret due to hindsight. Moreover, Cipriani and Guarino (2009) show that financial professionals are also prone to herding in a laboratory financial market environment. Also in a laboratory experiment, Haigh and List (2005) report that professional traders exhibit even more myopic loss aversion than students. Abdellaoui et al. (2013) show that finance professionals' behavior is better predicted by prospect theory than by expected utility and Razen et al. (2020) outline that professionals are equally prone to the domain effect (i.e., gain domain vs loss domain) in risk-taking than the general population. In contrast, Weitzel et al. (2019) and Cipriani et al. (2020) show that finance professionals and traders, respectively, are less prone to bubble formation in laboratory market experiments. Finally, Huber et al. (2021) show that finance professionals were more heavily impacted by the COVID-19 stock market crash than students, as the former show higher levels of risk aversion in investment experiments compared to experiments before March 2020. However, Angrisani et al. (2020) do not find differences in the impact of the COVID-19 stock market crash for professionals and students in non-finance related risk experiments.

[3] On the other hand, Fama and French (2010) conjecture that there is only superior performance—and hence indications of skill—in the extreme right tail of the distribution.

Table 1: Previous literature on the effects of economic preferences and cognitive skills on investment or trading performance (*Performance*) and risk taking (*Risk*). ↗ and ↘ indicate positive and negative effects, respectively.

A. Economic preferences

<table>
<tr><th>Characteristic</th><th>Performance</th><th>Risk</th><th>Sample</th><th>Data</th><th>Source</th><th>Relationship</th></tr>
<tr><td rowspan="2">Risk tolerance</td><td></td><td>Dorn and Huberman (2005)</td><td>Private investors</td><td>Field</td><td>Survey, retail bank data</td><td>Risk aversion ↘ underdiversification</td></tr>
<tr><td colspan="2">Bodnaruk and Simonov (2016)</td><td>Fund managers</td><td>Field</td><td>Survey, market</td><td>Risk aversion has no effect on downside risk and performance</td></tr>
<tr><td rowspan="2">Loss tolerance</td><td colspan="2">Bodnaruk and Simonov (2016)</td><td>Fund managers</td><td>Field</td><td>Survey, market</td><td>Loss aversion ↗ downside risk and ↘ performance</td></tr>
<tr><td>Larson et al. (2016)</td><td></td><td>Traders</td><td>Field</td><td>Natural experiment</td><td>Myopic loss aversion ↘ performance</td></tr>
<tr><td rowspan="2">Ambiguity tolerance</td><td colspan="2">Bianchi and Tallon (2018)</td><td>Private investors</td><td>Field</td><td>Survey, panel</td><td>Ambiguity aversion ↗ portfolio risk and performance</td></tr>
<tr><td></td><td>Dimmock et al. (2016a)</td><td>Private investors</td><td>Field</td><td>Survey</td><td>Ambiguity aversion ↗ portfolio underdiversification</td></tr>
</table>

B. Cognitive skills

Characteristic	*Performance*	*Risk*	*Sample*	*Data*	*Source*	*Relationship*
Cognitive skills or education	Grinblatt et al. (2011, 2012)		Private investors	Field	Panel, market	IQ ↗ performance
	Chevalier and Ellison (1999)		Fund managers	Field	Panel, market	SAT scores ↗ performance
	Golec (1996)		Fund managers	Field	Panel, market	Education ↗ performance
	Fang and Wang (2015)		Fund managers	Field	Panel, market	Education ↗ performance
		Gastineau (2004)	Fund managers	Field	Market	Superior trading policies / order management ↘ ETF tracking errors
	Hefti et al. (2018)		Students	Lab	Experiment	Mix of cognitive skills ↗ performance
	Hefti et al. (2018)		Students	Lab	Experiment	Combination of high cognitive and mentalizing skills ↗ performance
Cognitive reflection	Corgnet et al. (2018)		Students	Lab	Experiment	Cognitive reflection ↗ performance
	Noussair et al. (2016)		Students	Lab	Experiment	Cognitive reflection ↗ performance
Theory of mind	Bruguier et al. (2010)		Students	Lab	Experiment	Theory of mind ↗ performance in market prediction task
	Corgnet et al. (2018)		Students	Lab	Experiment	Theory of mind ↗ performance
Fluid intelligence	Corgnet et al. (2018)		Students	Lab	Experiment	Fluid intelligence ↗ performance

potentially stimulating the development of theoretical models combining various economic preferences or cognitive skills into a unified framework.

To address the above research questions, the following data were collected. In the first step, 92 fund managers from four large and mid-sized European Union (EU) member countries took part in our online experiment.[4] Then, we compiled daily and monthly time series data on the 412 funds that our participants managed between 2008 and 2019. The fund data includes, among others, returns, assets under management (AUM), and total expense ratios (ER) of the funds, as well as returns on their associated benchmarks. Finally, we matched the experimental data of each fund manager with their monthly funds' data.

With respect to the experimental part of our study, we followed the literature on lab experiments and ran three standard, non-incentivized tests to measure cognitive skills (Bruguier et al., 2010; Corgnet et al., 2018). To obtain scores in fluid intelligence, we administered a test similar to Corgnet et al. (2018). The test consisted of 18 Raven's advanced progressive matrices (Raven, 2000) in which fund managers had to solve diagrammatic puzzles. For cognitive reflection skills, we used the extended version of the cognitive reflection test (CRT, see Frederick, 2005) introduced by Primi et al. (2015) and Toplak et al. (2014). The questions are constructed such that they have an intuitive, but on reflection incorrect, response put forward by System 1; whereas the correct response requires the effortful activation of System 2 (Dual Process Theory, see Kahneman, 2011). For measuring theory of mind skills (TOM), which are important in detecting the informational content of investing by inferring others' intentions (Bruguier et al., 2010), we ran the "Reading-the-Mind-in-the-Eyes"-test proposed by Baron-Cohen et al. (2001).

Additionally, we ran incentivized experiments eliciting risk preferences and attitudes towards losses and ambiguity. Following Falk et al. (2018), we measured risk tolerance by means of a staircase procedure where fund managers had to make multiple decisions between a risky lottery and a (varying) safe payment. To measure loss tolerance, we ran a similar experiment with mixed lotteries, varying the potential losses while holding the alternative safe payment fixed. To assess ambiguity tolerance, we applied a modified version of the ambiguity experiment of Dimmock et al. (2016b), where subjects could choose between ambiguous and risky lotteries with varying probabilities of the latter.

Finally, alongside general demographic questions on age, experience in the industry, gender, and education, we added the five-item competition sub-scale of the Work and Family Orientation (WOFO) questionnaire introduced by Helmreich and Spence (1978), measuring preferences to compete.

First, we find a strong positive relationship between fund managers' risk tolerance and the volatility of the funds they manage. This indicates that fund managers with lower (higher) levels of risk tolerance compose funds with lower (higher) fund volatility. Importantly, this finding holds while controlling for fund managers' self-selection into fund categories (i.e., fixed income, international equity) and for additional variables like fund benchmark, industry experience, and fund size. In economic terms, our results predict that the benchmark-adjusted risk of funds run by managers with a risk tolerance close to risk neutrality is 14.0 percentage points higher compared to the funds run by managers with the lowest

[4] In the invitation and the welcome screen of the software we outlined not to disclose subjects' names and institutions. Moreover, we ensured not to disclose the countries of residence of the participating institutions in an upfront information letter.

risk tolerance in our experiment. In addition, we provide evidence that risk tolerance is negatively associated with a fund's Sharpe Ratio, implying that those fund managers with low risk tolerance, on average, earn higher risk-adjusted returns than their peers with high levels of risk tolerance.

Second, we observe that fund managers with lower levels of ambiguity tolerance manage their funds with lower tracking errors compared to their peers with higher levels of ambiguity tolerance. This finding indicates that tolerance towards ambiguous outcomes explains fund managers' propensity of deviations from the benchmark.[5]

Finally, we report that cognitive skills do not explain fund performance in terms of excess returns. However, we do find evidence that fund managers with high cognitive reflection abilities manage funds with lower risk when controlling for fund managers' self-selection into fund categories (i.e., fixed income, international equity) and for additional variables like fund benchmark, industry experience, and fund size. In economic terms, relative fund risk, on average, decreases by 4.5 percentage points per one-standard deviation increase in CRT.

With our paper, we, first, add to the literature investigating economic preferences and personality traits of business professionals in general (e.g., Adams et al., 2018; Graham et al., 2013; Kaplan et al., 2012; Malmendier and Tate, 2005) and, more specifically, to studies analyzing the interplay of economic preferences and portfolio dynamics among professional fund managers. As outlined, Bodnaruk and Simonov (2016) show that fund managers with high levels of self-reported loss aversion construct mutual funds with lower downside risk and show lower fund performance. Bodnaruk and Simonov (2016) run a short, non-incentivized survey to elicit fund managers' risk and loss aversion. We substantially extend their approach, as (i) we also account for the conjectures of Bruguier et al. (2010) and Corgnet et al. (2018) by studying the role of cognitive skills on behavior and performance and as (ii) we aim to establish a more comprehensive analysis of economic preferences by running a battery of incentivized experiments.

Second, we also add to the literature focusing on the impact of cognitive skills on investment behavior. Here, studies with private investors indicate that high-IQ investors show higher levels of stock market participation (Christelis et al., 2010), earn higher Sharpe ratios (Grinblatt et al., 2011), are less prone to the disposition effect, and exhibit superior market timing and stock-picking skills than low-IQ investors (Grinblatt et al., 2012). Regarding fund managers, studies document a positive relationship between fund performance and manager education (Golec, 1996) and between college SAT score and fund performance in the U.S. market (Chevalier and Ellison, 1999). Experimental finance literature contributes by analyzing the role of various cognitive skills for investment decisions: high cognitive reflection scores predict subjects' earnings in laboratory asset markets with students (Corgnet et al., 2015; Noussair et al., 2016), TOM correlates with subjects' skills in predicting price changes (Bruguier et al., 2010),[6] and all three concepts are joint predictors of students' performance (Corgnet et al., 2018). We

[5] Following literature on private investors, Boyle et al. (2012) and Dimmock et al. (2016b) show that ambiguity aversion is positively associated with home stock ownership, as foreign stocks are relatively more ambiguous (less familiar) than domestic stocks. Translated to our study, this would imply that deviating more strongly from the benchmark index (i.e., generating a higher tracking error) is more ambiguous and therefore less preferred by ambiguity-averse fund managers.

[6] However, DeMartino et al. (2013) show that TOM skills can also be detrimental when trading on financial markets. In a study using fMRI techniques, the authors report a mechanism by which social signals affect value computations in ventromedial prefrontal cortex, thereby increasing subject's proneness to ride financial bubbles.

extend this line of literature to the behavior of professional fund managers by showing that cognitive skills are negatively associated with fund risk. Generally speaking, our experimental approach allows us to study the link between fund managers' cognitive skills and their professional decisions directly without resorting to indirect proxies or artificial lab environments, thereby extending the approach of Corgnet et al. (2018) to the field.

Finally, we would like to note already upfront to not over-interpret the findings of our study. Strictly speaking, we cannot say much about the persistence of our findings and about the direction of causality. It could, for instance, be the case that past experiences in the financial market might have shaped fund managers' economic preferences as well (e.g., Guiso et al., 2018; Malmendier and Nagel, 2011). When turning to cognitive skills, however, evidence shows that at least these skills are stable over time (Stagnaro et al., 2018), indicating that the patterns in cognitive skills could be persistent. However, we leave the answer to the question of persistence and causality, especially for economic preferences, for future research. In five to ten years, we will analyze whether cognitive skills and economic preferences, elicited in our experiment, were able to predict (risk-adjusted) abnormal returns, fund risk, and tracking errors.

The remainder of the paper is organized as follows: In Section 2 we introduce the experimental design, the collection of the experimental and the empirical data, and our econometric approach. In Section 3 we present our results, and in Section 4 we discuss and conclude.

2. Experimental and Empirical Data

We exploit two sources of data: first, we collected data on cognitive skills and economic preferences by means of online experiments and, second, we matched the empirical fund time series with the experimental data of the fund managers who participated in the experiment. To ensure anonymity, fund managers' identities were replaced by randomly generated unique identifiers to match the experimental data with depersonalized, empirical fund data from various databases.

2.1. Experimental Fund Manager Data

We contacted approximately 900 fund managers via hard-copy letters and/or e-mails in which the study was outlined and which included personalized login credentials for participation in the online experiment. Ninety-four fund managers completed the experiment.[7] Fund managers were informed about the anonymous matching of the experimental data with the corresponding fund data. With the decision to participate, fund managers acknowledged to accept the informed consent of the experiment.

The experimental tasks were divided into three parts: (i) cognitive skills, (ii) incentivized economic preferences, and (iii) personality traits. Importantly, fund managers did not receive immediate feedback after each task, but were told in advance that they can select whether they want to receive feedback and background information on the experimental tasks after data collection has been completed. This was

[7] The total number of 900 invitations includes undelivered and returned mails, bounce-back e-mails, outdated or invalid (e-mail) addresses, etc. Thus, the response rate of roughly 10% should be considered being a conservative lower bound.

done to provide additional incentives to the fund managers to participate and to provide full disclosure of the background of the experiment. Upon completion of the experiment, one of the incentivized tasks was randomly chosen to determine the subjects' payout. In addition to their earnings from the corresponding task, subjects received a fixed participation fee of €25. Details on the experimental procedure and the feedback can be found in Appendix A.[8]

First, to measure fund managers' cognitive skills, we administered three different tasks. For cognitive reflection skills, we compiled a set of five questions taken from the extended cognitive reflection tests (CRT) proposed by Toplak et al. (2014) and Primi et al. (2015). The concept of cognitive reflection rests upon the dual-process theory framework (Kahneman, 2011). The questions in these tests are constructed in a way to have an intuitive, but on reflection incorrect, response put forward by System 1; the correct response requires the elaborate activation of System 2.[9] To obtain a score for fluid intelligence, we conducted a task similar to Corgnet et al. (2018), presenting 18 items from the Raven's Advanced Progressive Matrices (APM; Raven, 2000). For each item, subjects have to recognize the geometric pattern in a sequence and identify the missing element. The main objective of this test is to measure subjects' ability to solve novel problems, which is why it is also used to measure IQ. One additional advantage is that it can discriminate well even among high-IQ subjects. To measure theory of mind skills (TOM), we used 18 items of the "Reading-the-Eyes-in-the-Mind"-test proposed by Baron-Cohen et al. (2001). In this test, subjects are shown photographs of the eye region of different people and choose one of four feelings that best describe the mental state of the person whose eyes are shown. This test measures one's capacity to infer others' intentions, which, for instance, is important in detecting the information disseminated by the behavior of other market participants (Bruguier et al., 2010).[10]

Second, to measure economic preferences of the fund managers, we administered four incentivized experiments. Subjects were informed that, at the end of the experiment, one of these tasks would be randomly chosen and their decision in the respective task would determine their payout. Risk attitudes and inter-temporal preferences were elicited as in Falk et al. (2018). The task for loss tolerance was adapted from the procedure of Gaechter et al. (2010), while ambiguity tolerance was measured following the design introduced by Dimmock et al. (2016b). We increased consistency and comparability of the experiments by presenting all tasks in a staircase framework (see Figure S2 in Appendix A.2 for one example following Falk et al., 2018). In this setting, subjects face a set of path dependent decisions, offering two choices each. Along these decisions, one option stays the same, while the second option depends on the previous choice. Compared to single and multiple price list formats, this procedure offers the advantage to be concise without forfeiting precision in eliciting points of indifference. Moreover, as subjects are not informed about the staircase properties of the task, it is incentive-compatible.

In the risk preferences task, subjects first had to choose between a lottery paying €60 or €0 with

[8] The software, including all instructions as used for the data collection, is available for download as a zipped *oTree* project at https://osf.io/dq3t8/ and as a live demo version via https://fea-2018-en.herokuapp.com.

[9] For illustrative purposes, this is one of the questions: "Jerry received both the 15th highest and the 15th lowest mark in the class. How many students are in the class?" (Toplak et al., 2014). The (incorrect) intuitive answer (30 students) can be "overruled" upon reflection (29 students), which requires effortful System 2 processes.

[10] For the Raven's Advanced Progressive Matrices (APM) and the "Reading-the-Eyes-in-the-Mind"-test (TOM), we used shortened versions. The original tasks comprise 36 questions each, out of which we took every second question, starting with the first one of the original task. This was done to keep the overall time needed to complete the survey as short as possible without losing explanatory power. See also Bilker et al. (2012) and Olderbak et al. (2015) for data and a discussion on the usefulness of shortened versions of the APM- and the TOM-tests, respectively.

equal probability and a safe payment of €32. Subjects who preferred the lottery in the first stage were presented a higher safe payment in the second stage, while subjects who preferred the safe payment were presented a lower safe payment in the second stage. After four stages, this design allows to pin down a narrow interval for the subjects' certainty equivalents and hence an estimate of their risk preferences. Clearly, those subjects with high certainty equivalents are considered to show high levels of risk-tolerance. The payout of the safe alternative varied from €4 to €60.

In the time preferences task, the first decision problem asked subjects whether they preferred a payment of €20 today or a payment of €31 in 6 months. Those who selected the payment today were presented a higher future payment in the second stage while those who went with the future payment were presented a lower future payment in the second stage. Iterating this procedure reveals the implicit time discounting rate of the subjects. Note that we drop the variable on inter-temporal preferences (Patience) from the main analysis. The reason is that we lost part of the observations on Patience due to a runtime error issue with this task (around 10% of the sample), and therefore including this variable would result in a smaller sample. We show, in the robustness section, that all the significant results from the main analysis remain robust when Patience is added to analysis.

The loss aversion task started with the question whether subjects preferred to participate in a lottery that pays €22 or €−12 with equal probability. The positive payoff of €22 stayed the same in all questions. Subjects who rejected the lottery were presented with a lower negative payoff in the second stage while subjects who accepted the lottery were presented with a higher potential loss in the second stage. Iterating this procedure reveals the maximum loss subjects were willing to accept in order to obtain the chance of winning €22. According to this logic, subjects with a high tolerable maximum loss are the ones with high levels of loss tolerance. The range of varying negative payouts in the lottery varied from €−22.50 to €−1.50.[11]

In the ambiguity preferences task, the first decision problem asked subjects to choose between two lotteries. Each lottery offered the chance to win €60 or €0. While the probability of winning €60 was known to be 50% in one of the lotteries (risk), it was unknown in the other lottery (ambiguity). The ambiguous lottery remained unchanged throughout the task. Subjects who chose the risky lottery were presented with a new risky lottery offering a lower known probability of winning in the second stage, while subjects who chose the ambiguous lottery were presented a new risky lottery offering a higher known probability of winning in the second stage. Iterating this procedure reveals the matching probability at which subjects are indifferent between the ambiguous and the risky lottery. Thus, subjects who predominantly select the ambiguous lottery are the ones with high levels of ambiguity tolerance. The probabilities for winning €60 in the risky lottery ranged from 7% to 93%.

Third, we ran a test on measuring fund managers' attitudes towards competition. We used the 5-item subscale of the Work and Family Orientation (wofo) questionnaire proposed by Helmreich and Spence (1978), which is a widely used psychometric measure of individuals' self-assessed competitiveness, which was previously used in experiments with financial professionals (Kirchler et al., 2020).[12]

[11] Note that this parametrization together with the participation fee of €25 ensured that subjects could not incur losses in the experiment.

[12] Subjects answered the following five questions: "I enjoy working in situations involving competition with others"; "It is important to me to perform better than others on a task"; "I feel that winning is important in both work and games"; "It annoys me when other people perform better than I do"; "I try harder when I'm in competition with other people". The

Table 2: Summary statistics of scores in the experimental tasks ($n = 92$).

Task (Variable)	Mean	SD	Min	Max
CRT	3.97	1.09	1	5
APM	9.81	2.86	1	16
TOM	11.48	2.45	4	16
Risk Tolerance	27.52	6.94	10	50
Loss Tolerance	18.08	5.50	2.25	23.25
Ambiguity Tolerance	39.96	14.04	4	78
Competitiveness	26.59	4.53	14	35

Note: The cognitive reflection test consisted of 5 questions (CRT). The task measuring fluid intelligence comprised of 18 questions (APM). The "Reading-the-Eyes-in-the-Mind"-test measuring theory of mind comprised of 18 questions (TOM). The score for risk tolerance reflects the elicited certainty equivalent for a lottery paying €60 or €0 with equal probability (Risk Tolerance). The score for loss tolerance reflects the maximum potential loss subjects were willing to accept in order to have the chance of winning €22 (Loss Tolerance). The score for ambiguity tolerance represents the matching probability (in %) that left subjects indifferent between a risky lottery with the respective probability of winning and an ambiguous lottery with an unknown probability of winning (Ambiguity Tolerance; both lotteries paid €60 in the case of winning and €0 else). Competitiveness was measured as the sum of all five questions of the WOFO survey with Likert-scales ranging from 1 ("I do not agree at all") to 7 ("I fully agree") each (Competitiveness).

Questions on demographics concluded the experiment. In total, 94 fund managers completed the experiment, which, from an empirical perspective, might sound relatively low. However, in experimental studies with professional subjects, these numbers are in the upper range of comparable research.[13] We lose one fund manager, because we are not able to match at least one fund to him/her, and we lose one additional manager when excluding certain funds from the sample (see section 2.2 for details). Consequently, our base sample consists of 92 fund managers. The average age of the 92 fund managers in our final sample was 44 years, with an average tenure in the finance industry of 18 years. 95% of the fund managers were male.

The experiment was programmed in *oTree* (Chen et al., 2016), utilizing the ready-made applications introduced by Holzmeister (2017). The experimental sessions were conducted in December 2017 and January 2018. Completing the online experiment took fund managers on average 32 minutes (*SD* of 9 minutes). Payout to the subjects was administered via a third party specialized on micro-payments or via bank transfer.

Table 2 provides a descriptive overview of the experimental results. For the econometric analyses, we z-standardize the cognitive skill and competitiveness measures by subtracting means and dividing by standard deviations. This does not affect t-statistics, but it makes the economic interpretation of the corresponding coefficients more meaningful.[14] For the economic preference tasks, we conduct our

answers were provided on a Likert scale ranging from 1 (I do not agree at all) to 5 (I fully agree). The sum over all five questions finally enters our data analyses.

[13] For instance, Haigh and List (2005) run their study with 54 finance professionals, Cohn et al. (2014) tested 128 finance professionals—split across two experimental treatments, and Bodnaruk and Simonov (2016) conducted their fund manager study with 68 subjects and only use 52 managers in their main analysis.

[14] The motivation for standardizing cognitive skill and competitiveness measures is the absence of interpretable economic units. Moreover, accounting for potential differences in scaling among the WOFO questions, we also standardized each question separately before computing aggregated competitiveness scores.

analysis using the original metrics of the elicitation procedures.

2.2. Empirical Fund Data

We match the participants to empirical data on the funds they managed between January 2008 and December 2019. Matching the experimentally elicited cognitive skills and economic preferences with fund time series mainly from the past (i.e., before the experiment) warrants some more discussion. One alternative would be to run the experimental tasks in a first step and use them as predictors of fund performance and risk for subsequent years only. We leave this issue for future research, when we will analyze whether cognitive skills and economic preferences actually predict future (risk-adjusted) abnormal returns, fund risk, and tracking error in five to ten years. With respect to the chronology of experiment and fund data, Bodnaruk and Simonov (2016) used their survey questions also for a matched sample following both approaches separately (i.e., the empirical time series were selected once from the time span before and once after the survey was run) and they do report similar results. We account for this discussion by including data both before and after our experiment, albeit, of course, the larger part of our fund dataset covers a period before the experiment was conducted. Finally, cognitive skills appear to be relatively stable over time (Stagnaro et al., 2018, e.g.,), making the issue of experiments lagging or leading the empirical data less important. For instance, Böhm et al. (2018) use cognitive scores from the military enrollment tests at the age of 18 as a predictor for successful careers in finance. Moreover, a substantial body of literature uses the same approach of running experiments or surveys on economic preferences and relate its findings to portfolio choice of the preceding years (e.g., Bianchi, 2018; Bianchi and Tallon, 2018; Dimmock et al., 2016b; Riedl and Smeets, 2017). Hence, these studies implicitly assume relative stable preferences which is also partly backed up by literature (Meier and Sprenger, 2015).[15]

We use *Morningstar Direct* to match the participants to empirical data on the funds they manage. Table 3 provides a cross-sectional snapshot of the funds domiciled in our four sample countries at the end of 2017. Panel A describes all the funds that appear in Morningstar.[16] There are almost two thousand funds that constitute the underlying population, with an average assets under management of €363 million. The average gross return across the funds during 2017 was 9.5% and the funds have an average Morningstar rating of 3.2. Panel B of Table 3 shows the same summary statistics as Panel A for the funds where one of our participants is listed as a manager at the end of 2017. There are 240 such funds (the same person can be a manager of multiple funds at the same time). Our sample captures 12% of all funds both in terms of number of funds and in terms of total assets under management. The summary statistics are remarkably similar across Panels A and B, indicating that we have a fairly representative sample of funds.

[15] However, Guiso et al. (2018) find that measures of risk aversion among private investors—measured in 2007 and 2009—increased substantially after the financial crisis of 2007, indicating that risk preferences can vary following extreme events. However, the authors claim that it is unclear how persistent and long-lasting such fear-induced change in risk aversion are.

[16] Note that funds with no information about the manager are excluded (i.e., where the entry for the "Manager Name" variable in Morningstar is either blank or contains "Not Disclosed"). We make the exclusion, because even in the ideal case of a 100% response rate by the invited managers in the experiment, we would not be able to match these funds to our participants due to the lack of manager information in Morningstar. The excluded funds represent a little over 20% of the total assets under management in the four sample countries.

Table 3: Fund population at the end of 2017: The table provides summary statistics for the cross-section of funds domiciled in our four sample countries at the end of 2017. Panel A contains all funds with information on manager identity from Morningstar. The rest of the panels describe various subsets. Panel B contains funds where one of the participants from the experiment is listed as a manager at the end of 2017. Panel C contains all the funds run by a single manager at the end of 2017. Panel D contains funds where one of the participants from the experiment is listed as the single manager at the end of 2017. Panel E provides the distribution of the funds across the major investment categories for these four subsets.

	N	$Mean$	SD	p_{10}	p_{25}	p_{50}	p_{75}	p_{90}
Panel A: All funds								
AUM (mio. EUR)	1959	363.08	1228.47	9.13	26.24	89.54	307.87	847.07
Return in 2017 (%)	1749	9.50	7.72	1.36	3.85	8.53	13.56	19.31
MS rating (1 to 5)	1547	3.19	1.03	2	3	3	4	5
Panel B: Sample funds								
AUM (mio. EUR)	240	341.62	817.14	11.08	31.45	88.97	275.38	818.90
Return in 2017 (%)	208	9.68	7.11	1.22	4.24	9.06	13.39	18.31
MS rating (1 to 5)	199	3.30	1.01	2	3	3	4	5
Panel C: All single-managed funds								
AUM (mio. EUR)	1238	299.80	1175.36	8.44	24.47	73.00	255.23	745.19
Return in 2017 (%)	1161	9.37	7.71	1.46	3.84	8.36	13.54	18.96
MS rating (1 to 5)	970	3.10	1.00	2	2	3	4	4
Tenure (years)	1238	6.74	5.61	1.25	2.42	5.25	9.67	14.33
Panel D: Sample single-managed funds								
AUM (mio. EUR)	113	228.73	479.90	9.18	28.19	53.83	202.09	430.08
Return in 2017 (%)	105	9.13	6.66	1.53	4.24	8.49	12.84	16.81
MS rating (1 to 5)	98	3.17	0.98	2	2	3	4	4
Tenure (years)	113	6.99	4.77	1.92	3.42	5.92	9.17	14.67

Panel E: Category distribution (%)

	All	*Sample*	*All single*	*Sample single*
Fixed income	23.9	19.8	25.2	22.3
Equity	40.9	46.6	39.7	43.7
Allocation	27.0	28.6	26.7	29.5
Other	8.2	5.0	8.4	4.5

Note: *Return in 2017* is the gross return on the fund during 2017. The following three variables are taken at the end of 2017: AUM is the fund's assets under management, *MS rating* is the fund's Morningstar rating, and *Tenure* is the manager's tenure at the fund. The following statistics are shown: N is the number of non-missing observations, $Mean$ is the sample average, SD is the sample standard deviation, and p_j is the j-th percentile.

There are funds in the sample run by the single manager who participated in our experiments, and there are funds run by a group of managers, where our participating manager is one member of the group. The sample of single-managed funds provides a cleaner setting for our analysis, since in the case of the team-managed funds, we do not have information about the cognitive skills and economic preferences of the other members from the management team, or about the hierarchy within the group. Therefore, our main results are obtained using the sample of single-managed funds. Panel C of Table 3 provides a description of all the funds run by a single manager at the end of 2017. More than half of the total assets under management belongs to single-managed funds. Overall, single-managed funds are similar to the general population in terms of assets under management, gross return, and Morningstar rating . The average tenure of the single manager at the end of 2017 is close to 7 years. Panel D of Table 3 describes our sample of single-managed funds. The typical fund size is somewhat smaller in Panel D compared to Panel C, but the two panels are very close to each other along the other dimensions (performance and tenure). Finally, Panel E provides the distribution of the funds across the major investment categories, which turns out to be very similar across all subsets of the fund population discussed above. Overall, we consider our fund sample to be representative in the case of single-managed funds as well.

While Table 3 describes the sample as of December 2017, we do not rely only on these funds in our econometric analysis. Our participants may have also managed different funds before or after 2017, and leaving those out from the analysis could introduce survivorship bias. Therefore, our sample for the analysis constitutes *all* funds where our managers were active for at least one month between 2008 and 2019.[17] Using the procedure detailed in Appendix B.1, we are able to match at least one fund to 93 out of the 94 participants. Following the literature (see, e.g., Ibert et al., 2018), we eliminate money market funds and index funds, and we also exclude a few funds for which we are not able to obtain any fund-level time-series data. The resulting sample contains 92 managers and 412 funds. As mentioned earlier, the main analysis is carried out using the subset of single-managed funds. This subset consists of 65 managers and 209 funds.

After identifying the funds, we collect time series data on their net returns both on the daily (denoted by R^{net}_{id}) and monthly (denoted by R^{net}_{it}) frequencies. For each fund, we also obtain the monthly total expense ratios, ER_{it}, and the series of end-of-month assets under management, AUM_{it}. The main source of the data is the *Morningstar Direct* database, but we augment the ER and AUM series using the *Lipper* database and imputations. The final sample contains 28,369 unique fund-month observations. Details of the data collection and imputations are described in Appendix B.2.

Using the monthly data, we calculate the following variables related to fund performance:

$$R^{gross}_{it} = R^{net}_{it} + ER_{it} \tag{1}$$

$$R^{abn}_{it} = R^{gross}_{it} - R^{B}_{it} \tag{2}$$

$$V_{it} = AUM_{it-1} \cdot R^{abn}_{it} \tag{3}$$

The gross return, R^{gross}_{it}, represents the total return on the fund before expenses. R^{B}_{it} denotes the monthly return on the fund's benchmark. Hence the abnormal return, R^{abn}_{it}, is the total monthly return

[17] In unreported robustness checks we found that all results remain very similar (both qualitatively and quantitatively) if we (i) exclude funds where our managers have a tenure shorter than one year within the sample period, or (ii) we use 2008-2017 data on only those funds that our participants manage as of December 2017 (i.e., only the funds described in Table 3).

of the fund over its benchmark. We also calculate the value added, V_{it} (Berk and van Binsbergen, 2015), i.e., the product of assets under management and abnormal returns. Berk and van Binsbergen (2015) argue that the skill of a mutual fund manager equals the value her fund extracts from capital markets, which can be measured by V_{it}.

To measure the risk-adjusted performance and the amount of risk (relative to the benchmark) taken by the fund managers, we calculate the following monthly variables using daily data:

$$SR_{it} = \frac{E\left[R^{gross}_{id[t]}\right] - R_{ft}}{Std\left(R^{gross}_{id[t]}\right)} \tag{4}$$

$$RV_{it} = \frac{Std\left(R^{net}_{id[t]}\right)}{Std\left(R^{B}_{id[t]}\right)} \tag{5}$$

$$TE_{it} = Std\left(R^{net}_{id[t]} - R^{B}_{id[t]}\right), \tag{6}$$

where $R_{id[t]}$ denotes daily returns on fund i within month t, and R_{ft} is the one-month EURIBOR rate in month t (expressed in daily units).[18] $E[\cdot]$ denotes the expected value, while $Std(\cdot)$ is the standard deviation in the formulas above. SR_{it} is the Sharpe ratio of fund i in month t, which measures the risk-adjusted performance of the fund. The relative volatility, RV_{it}, measures the overall riskiness of the fund relative to the riskiness of the benchmark. A value of one indicates that, during month t, the fund's return volatility was the same as that of the benchmark, while values above (below) one indicate that the return volatility of the fund was higher (lower) than that of the benchmark. The tracking error, TE_{it}, measures risk in the fund's return that is due to active management decisions made by the portfolio manager. Importantly, these two variables measure risk-taking from two different points of view: RV_{it} compares the overall level of risk taken by the fund to its benchmark, while TE_{it} indicates how closely the fund mimics its benchmark index.

Most of the above variables related to performance and risk-taking require return data on the fund's benchmark, R^B_{it}. We use the prospectus benchmark, which is self-declared by the fund company itself, because this benchmark is likely to have the highest influence on the manager's decisions. By comparing each fund to its prospectus benchmark, we follow the approach of, for instance, Ibert et al. (2018). However, our mutual fund databases do not report a prospectus benchmark for all funds in our sample. We are able to identify and obtain return data on the prospectus benchmark for 206 funds (50% of the 412 funds in our base sample). Since we would like to avoid losing half of our sample, we assign benchmarks to the remaining funds, if possible, via two further steps. First, similar to Ibert et al. (2018), we use the following benchmark assignment rule for equity funds: for each equity fund category defined by the *Morningstar* variable "Category," we find the most common benchmark among all open-ended mutual funds that have one of our four countries registered as "Domicile." This most common benchmark is assigned to all the funds in the given category. We are able to obtain benchmark return data for 43 additional funds this way. Second, we use *Lipper* to identify further benchmarks. *Lipper* independently assigns the "Lipper Technical Indicator Benchmark" to most of the funds in the database according to

[18] We require at least 18 daily observations to calculate the monthly expected returns and standard deviations. The $R^{gross}_{id[t]}$ values are obtained by adding the expense ratio (expressed in daily units) to the $R^{net}_{id[t]}$ observations. Note $R^{gross}_{id[t]}$ and $R^{net}_{id[t]}$ can be used interchangeably in the standard deviation calculations, since the expense ratio, ER, is constant within the month.

its assessment of the fund's investment strategy. This technical indicator benchmark is used for the remaining funds, if available. Benchmark return data is obtained for 55 additional funds via this step. Further details on the benchmark assignment process are described in Appendix B.3. Altogether, there are 304 funds with benchmark data. Note that there are 6 managers for whom we are not able to obtain benchmark data for any of the funds they manage. Consequently, these managers are left out from the analysis when the dependent variable requires return data on the benchmark.

Another possibility would be to replace the prospectus benchmark with an alternative benchmark obtained by fitting a multi-factor model to the fund's returns (see, e.g., Carhart, 1997; Fama and French, 1993, 2015). However, given the large heterogeneity of the funds in our sample with respect to country of domicile, allocation across asset classes, and geographical focus, it is ex-ante unclear what factors should be used. While we consider the prospectus benchmark to be the one that governs behavior and decisions of fund managers the most, we report – for the sake of completeness – robustness checks based on different factor model-based benchmarks in Appendix D. It is reassuring that our main results are qualitatively robust to various definitions of the funds' benchmarks.

Fund-level summary statistics are provided in Table 4. Panel A corresponds to the sample of 304 funds with benchmark return data, while Panel B presents the base sample of 412 funds. Comparison of the two panels reveals minor differences that are due to the availability of the benchmark return data. The following discussion focuses on Panel A, while the minor differences compared to Panel B are highlighted at the end of the section. Note that in our regression, we are going to use the variables from Table 4 expressed in the same units as indicated in the table.

The average fund (out of the 304) appears in the sample through 63 months, which is roughly half of the full sample period of 12 years. The average fund has AUM of €347 million, but the cross-sectional fund size distribution is wide and left skewed: the median AUM is €90 million, ten percent of the funds are smaller than €7 million, whereas ten percent of the funds manage more than €995 million. The average annual expense ratio is 1.31%, but investors can pay as little as 0.41% (10$^{\text{th}}$ percentile) and as much as 2.09% (90$^{\text{th}}$ percentile). The average annual gross return is 6.69% (the median is 7.17%), and there is a large variation across funds with the interdecile range taking values from −1.27% to 15.42%. The funds earn a slightly higher gross return than their benchmark on average: the mean (median) annual abnormal return is 0.29% (0.18%). However, the abnormal return after expenses ($R_i^{abn} - ER_i$) is negative both for the average and for the median fund (not reported in the table). The average annual value added is €3.8 million, but the median value added is close to zero, indicating that the distribution of V_i is considerably right skewed. The average (annualized) Sharpe ratio is 1.1, but there is a considerable variation across funds due to the large differences in terms of what asset classes and markets they focus on. Both the average and median RV_i are close to one, indicating that the return volatility of a typical fund is close to the volatility of its benchmark. However, RV_i can be as low as 0.76 (10$^{\text{th}}$ percentile) indicating that the fund's volatility is 24% lower than that of the benchmark, or can be as high as 1.20 (90$^{\text{th}}$ percentile), indicating that the fund is 20% more volatile than its benchmark. For half of the funds, we observe at least one month where our participant is the single manager. The funds' distribution across major investment categories is 65% equity funds, 21% fixed income funds, 10% allocation funds, and 4% others.

The comparison of the two panels in Table 4 shows that the tenure, expense ratio, Sharpe ratio, and

Table 4: Summary statistics of the sample at the fund level: The table provides summary statistics about the funds in our sample. For each variable (except the number of months in the sample), we calculate the time-series average of that variable for each fund for the period that the fund appears in our sample (i.e., when one of our participants is a manager). The table reports the cross-sectional summary statistics of the fund level time-series averages. Panel B corresponds to all funds in the sample, while Panel A corresponds to the subset of funds where return data on the benchmark is available.

Panel A: Funds with benchmark return data

	N	$Mean$	SD	p_{10}	p_{25}	p_{50}	p_{75}	p_{90}
in sample (months)	304	63.14	42.45	12.00	26.00	57.00	95.00	130.00
AUM_i (mio. EUR)	304	347.21	697.24	6.99	28.37	89.76	309.45	994.86
ER_i (annual %)	304	1.31	1.28	0.41	0.69	1.30	1.69	2.09
R_i^{gross} (annual %)	304	6.69	10.96	−1.27	2.92	7.17	11.79	15.42
R_i^{net} (annual %)	304	5.38	10.94	−1.86	1.50	5.78	10.29	14.54
R_i^{abn} (annual %)	304	0.29	4.58	−4.14	−1.32	0.18	1.81	4.57
V_i (annual mio. EUR)	304	3.82	19.49	−3.09	−0.61	0.06	2.11	13.42
SR_i (annual)	303	1.11	1.02	0.02	0.61	1.12	1.56	2.05
RV_i	299	1.00	0.19	0.76	0.94	1.01	1.08	1.20
TE_i (annual %)	302	10.90	8.62	2.11	4.71	8.47	16.09	21.13
Team-managed	304	0.49						
Fixed income	304	0.21						
Equity	304	0.65						
Allocation	304	0.10						
Rest	304	0.04						

Panel B: All funds

	N	$Mean$	SD	p_{10}	p_{25}	p_{50}	p_{75}	p_{90}
in sample (months)	412	64.41	43.05	12.00	27.00	57.00	96.00	132.00
AUM_i (mio. EUR)	401	304.38	634.63	6.82	22.07	75.26	246.35	869.26
ER_i (annual %)	405	1.27	1.14	0.42	0.74	1.27	1.62	2.03
R_i^{gross} (annual %)	399	6.43	9.82	−0.57	2.97	6.41	10.98	14.82
R_i^{net} (annual %)	406	5.17	9.74	−1.18	1.58	5.24	9.81	13.40
SR_i (annual)	406	1.17	1.01	0.02	0.62	1.16	1.65	2.27
Team-managed	412	0.49						
Fixed income	412	0.17						
Equity	412	0.50						
Allocation	412	0.28						
Rest	412	0.05						

Note: The gross return, R_i^{gross}, represents the total return on the fund before expenses and R_i^{net} indicates the total return after accounting for the expense ratio ER_i. The abnormal return, R_i^{abn}, is the total annual return of the fund over its benchmark. The value added Berk and van Binsbergen (2015), V_i, is the product of assets under management, AUM_i, and the abnormal return, R_i^{abn}. The Sharpe ratio, SR_i, is a measure of risk-adjusted performance. The relative volatility, RV_i, measures the overall riskiness of the fund relative to the riskiness of the benchmark. The tracking error, TE_i, measures risk in the fund's return that is due to active management decisions made by the portfolio manager. The following statistics are shown: N is the number of non-missing observations, $Mean$ is the sample average, SD is the sample standard deviation, and p_j is the j-th percentile.

return statistics are very similar in the full sample and the sub-sample with benchmark data. There are two differences between the two samples: (i) larger funds are somewhat over-represented in the sub-sample with benchmark data, as the average and median AUM is slightly higher in Panel A, and (ii) allocation funds are under-represented, while equity- and fixed income funds are over-represented in the sub-sample, relative to the full sample. This is due to the fact that smaller and allocation funds less frequently report a benchmark and that it is more difficult to come up with a suitable one.

2.3. Econometric Model

To answer our research questions on the impact of cognitive skills and economic preferences, we set up the following regression model. As dependent variables indicating fund performance, we use monthly gross returns, R_{it}^{gross}, monthly abnormal returns, R_{it}^{abn}, and monthly value added, V_{it}. As proxies for fund risk, we use relative volatility, RV_{it}, and tracking error, TE_{it}, serving as dependent variables as well. In addition, we proxy risk-adjusted returns by means of the funds' Sharpe ratios, SR_{it}.

As independent variables, we include all experimentally elicited variables measuring cognitive skills and we add fixed-effects for fund category and time. Moreover, we add a dichotomous variable indicating whether a fund was team-managed or not (Team) in the given month, Experience measuring years in the industry, and the log of assets under management at the end of the previous period, $\log(AUM)_{t-1}$. In the regressions, examining the effects on risk-taking, we also control for past fund performance. As it will be evident from the results, it is important to allow for a non-linear effect of past performance. Therefore, we use two variables that are non-linear transformations of the past fund return: $\min(R_{t-1}^{gross}, 0)$ is zero if the gross return in the previous month was positive, and equal to the gross return in the previous month if it was negative. Similarly, $\max(R_{t-1}^{gross}, 0)$ is zero if the gross return in the previous month was negative, and equal to the gross return in the previous month if it was positive. Including these two variables allows for the possibility to estimate a V-shaped effect of past return.

We use time fixed effects (i.e., year-month) and investment category fixed effects (i.e., the four categories shown in the last four rows of Table 4) in all our regressions. Standard errors are clustered at the manager level. Importantly, we drop our measure of fluid intelligence (apm) from our main analyses due to its high correlation with crt (see Table S1 in Appendix A.2), which leads to potential multicollinearity issues.[19] Finally, we would like to emphasize that we lower the p-value thresholds for statistical significance to 5.0% and 0.5% in all econometric specifications to reduce the likelihood of false positives.

[19] In a comprehensive robustness analysis, we re-estimate our main results using apm instead of crt. Coefficient estimates on all other experimental variables remain robust both in terms of statistical significance and economic magnitude. That is, the decision whether to use crt or apm in our regressions does not affect the results for the remainder of the experimental variables. The results of the robustness check are discussed in detail at the end of the *Results* section.

3. Results

Result 1: *Neither cognitive skills nor economic preferences and attitudes towards competition do contribute to abnormal returns or value added.*

Support: As outlined in Table 5, we observe that all coefficients of all experimental measures are insignificant for all performance variables (i.e., R^{gross}, R^{abn}, and V). This suggests that neither fund managers' attitudes towards risk, losses, and ambiguity, nor their cognitive skills or their attitudes towards competition have a systematic effect on the managed funds' abnormal returns or value added. Moreover, we find that the absence of significant results holds both for single-managed funds (columns 1, 3, and 5 in Table 5) and the sample of all funds. The positive coefficients for the Team dummy ($p < 0.05$), however, suggest that, on average, team managed funds outperform single-managed funds by approximately 80 basis points per year in abnormal returns (column R^{abn}). The latter result adds to a so far inconclusive literature: While some studies investigating the (risk-adjusted) performance of equity funds do not find differences between team-managed and single-managed funds (see, e.g., Bliss et al., 2008; Massa et al., 2010), others report that single-managed funds even outperform team-managed funds (see, e.g., Bär et al., 2011; Chen et al., 2004). On the contrary, Patel and Sarkissian (2017) provide evidence that team-managed funds add on up to 30–40 basis points per year to gross performance as compared to single-managed funds, and argue that the lack of evidence on performance benefits of team-managed funds in previously published studies is due to discrepancies in reported managerial structures in various data sources.[20]

Importantly, we focus on and mainly discuss the sample of single-managed funds as primary data set in all tables of the paper. We do this because we cannot control for the economic preferences and cognitive skills of the co-managers of the team-managed funds, making the effects less clear among those funds (see Bär et al., 2011, for evidence showing that team-managed funds can behave differently than single-managed funds).

Result 2: *Fund managers' risk tolerance is negatively related to a fund's Sharpe Ratio and positively correlated with fund volatility. At the same time, fund managers with high cognitive reflection abilities take fewer risks.*

Support: As indicated in column 1 of Table 6, we find a significantly negative relationship between the Sharpe Ratio and the fund manager's Risk Tolerance for single-managed funds. This finding indicates that individual's risk preferences explain the risk-adjusted performance of a managed fund, even beyond controls like the funds' category, assets under management, time fixed-effects, years in industry, and the other experimental measures. Thus, fund managers with lower levels of risk tolerance manage to perform better than fund managers with higher risk tolerance on a risk-adjusted basis. In economic terms, the annual Sharpe ratio of funds run by risk neutral managers (i.e., certainty equivalent

[20] Note that we obtained data on managerial structures from *Morningstar Direct* (MD); the study by Patel and Sarkissian (2017) uses the same database and shows that MD is considerably more accurate (relative to the *Securities and Exchange Commission*'s (SEC) records) than *Center for Research in Security Prices* (CRSP) and *Morningstar Principia* (MP)—the data sources of studies failing to find support for outperformance of team-managed funds as compared to single-managed funds. The authors estimate that the discrepancies in reported managerial structures in CRSP and MP data relative to SEC filings result in an underestimation of the team impact on fund performance of up to 50 basis points per year.

Table 5: Performance – gross returns, abnormal returns and value added: The table shows the results of ordinary least squares regressions of funds' abnormal returns and value added on cognitive skills and economic preferences/attitudes. The gross returns (R^{gross}, expressed in *annual percentage* units) indicate the gross returns of the funds without benchmark correction. The abnormal returns (R^{abn}, expressed in *annual percentage* units) are the funds' monthly gross returns over their benchmark. The value added (V, expressed in *annual mio. EUR* units) is the product of assets under management and abnormal returns. Standard errors are clustered at the manager level. Corresponding p-values are reported in parentheses. ** $p < 0.005$, * $p < 0.05$.

	Gross returns		*Abnormal returns*		*Value added*	
Dependent variable	R^{gross}	R^{gross}	R^{abn}	R^{abn}	V	V
Sample of funds	*Single*	*All*	*Single*	*All*	*Single*	*All*
CRT	−0.349	0.298	−0.181	−0.178	−0.380	1.017
	(0.239)	(0.358)	(0.520)	(0.465)	(0.764)	(0.460)
TOM	−0.067	−0.332	−0.280	−0.173	−0.850	0.168
	(0.814)	(0.259)	(0.260)	(0.345)	(0.267)	(0.844)
COMPETITIVENESS	−0.066	0.382	0.209	0.234	1.344	0.902
	(0.855)	(0.198)	(0.500)	(0.315)	(0.176)	(0.500)
RISK TOLERANCE	−0.088	0.003	−0.104	−0.020	0.001	0.186
	(0.052)	(0.928)	(0.060)	(0.521)	(0.993)	(0.451)
LOSS TOLERANCE	0.098	−0.007	0.060	0.032	−0.029	0.033
	(0.096)	(0.897)	(0.230)	(0.395)	(0.840)	(0.861)
AMBIGUITY TOLERANCE	0.013	−0.001	0.037	0.019	0.050	0.018
	(0.490)	(0.955)	(0.053)	(0.115)	(0.412)	(0.782)
$\log(AUM)_{t-1}$	0.070	−0.043	0.114	0.198*		
	(0.757)	(0.721)	(0.500)	(0.044)		
EXPERIENCE	−0.044	0.043	0.015	0.012	0.057	0.044
	(0.292)	(0.217)	(0.705)	(0.633)	(0.630)	(0.731)
TEAM		0.847		0.828*		2.702
		(0.088)		(0.042)		(0.187)
Constant	yes	yes	yes	yes	yes	yes
Category FE	yes	yes	yes	yes	yes	yes
Time FE	yes	yes	yes	yes	yes	yes
Number of observations	12,322	25,857	9,393	19,636	9,393	19,636
Number of managers	65	92	59	86	59	86
Adjusted R^2	0.439	0.504	0.086	0.101	0.031	0.040

Independent variables: CRT stands for the cognitive reflection score, comprised of 5 questions, measuring deliberate thinking. TOM stands for the "Reading-the-Mind-in-the-Eyes"-test, measuring theory of mind skills, i.e., the ability to infer the intention of others. The score for risk preferences (RISK TOLERANCE) reflects the elicited certainty equivalent for a lottery paying €60 or €0 with equal probability, with higher values indicating higher levels of risk tolerance. The measure for attitudes towards losses (LOSS TOLERANCE) reflects the maximum potential loss subjects were willing to accept in order to have the chance of winning €22. Again, the higher the number, the more tolerant towards losses a fund manager is (LOSS TOLERANCE). The score for ambiguity preferences (AMBIGUITY TOLERANCE) represents the matching probability (in %) that leaves subjects indifferent between a risky lottery with a certain probability of winning and an ambiguous lottery with an unknown probability of winning (both lotteries paid €60 in the case of winning and €0 else). COMPETITIVENESS is measured as the sum of the five standardized responses to the subscale of the WOFO, answered on scales ranging from 1 to 7 each. $\log(AUM)$ stands for the log of assets under management, EXPERIENCE indicates years in industry, and TEAM is a dichotomous indicator for team-managed funds.

of 30) is predicted to be 0.34 lower than the Sharpe ratio of funds run by managers with the lowest risk tolerance (see also Table 2 for the mean, standard deviation, minimum, and maximum of variable Risk Tolerance). This result, however, washes out for the full data set when team-managed funds are included (column 2 of Table 6).

The coefficient of Risk Tolerance is significantly positive in the regressions where the dependent variable is relative volatility. This result not only holds for single-managed funds, but also when team-managed funds are included, though with a lower magnitude in the latter case (see columns 3 and 4 of Table 6). This finding indicates that individual-level risk preferences, measured in our lottery experiment, explain part of the variation in the riskiness of a managed fund, even beyond controls like the funds' benchmark and category, assets under management, time fixed-effects, years in industry, and the other experimental measures. With this procedure we also control for potential effects of self-selection of fund managers into certain fund categories (e.g., fixed income vs. equity funds) and find the effect of risk preferences on top of the controls. We also consider the economic significance to be important. For single-managed funds (column 3 of Table 6), our results predict that those fund managers that are close to risk neutrality (i.e., certainty equivalent of around 30) run funds with a benchmark-adjusted risk that is 14.0% higher compared to those fund managers with the lowest risk tolerance.[21]

In addition, our analysis provides evidence that CRT scores are negatively related to relative volatility. As outlined in the third (and also in the fourth) column of Table 6, we report that the coefficients of CRT are negative ($p < 0.005$) for single-managed funds and the full sample alike. Again, we argue that this effect is also economically relevant. In economic terms, relative fund risk, on average, decreases by 4.5 percentage points per one-standard deviation increase in CRT (see column 3). Thus, in comparison to fund managers with a score of three out of five correctly answered questions in the CRT task, the relative volatility of the funds run by managers with the highest possible CRT score (five correct answers) is 8.3 percentage points lower in the case of single-managed funds (i.e., the estimated effect size of $0.045 \times 2 = 0.090$ divided by the standard deviation of 1.09).

Turning to the control variables of the analysis of relative volatility, we report that fund managers' risk-taking is strongly driven by the previous month's return realizations, and the effect is non-linear (see again columns 3 and 4). We observe significantly negative coefficients for $\min(R_{t-1}^{gross}, 0)$ (which can take on values ≤ 0) and significantly positive coefficients for $\max(R_{t-1}^{gross}, 0)$ (which can take on values ≥ 0). This means that both larger negative and larger positive returns in the previous month lead to higher risk-taking relative to the benchmark. This indicates a V-shaped reaction of risk-taking conditional on past performance. Let us highlight two further observations regarding past performance as a control. First, it is important to model the non-linearity of the effect; if past performance is included simply as a linear variable instead of the two non-linear variables above, these sign-dependent effects cancel each other out. Second, if we do not control for past performance (i.e., leave out $\min(R_{t-1}^{gross}, 0)$ and $\max(R_{t-1}^{gross}, 0)$ from the regressions) or control for past performance linearly, the coefficients on all remaining explanatory variables are effectively unchanged, hinting at robust patterns of the behavioral

[21] In an unreported analysis we confirm that the effect of Risk Tolerance on relative volatility is through the fund's volatility and not through the benchmark's volatility. That is, Risk Tolerance has a significant effect on the numerator in the definition of RV (see equation (5)), but not on the denominator. We prefer to use relative volatility, because it controls for the effect of managers' self-selection into certain funds.

Table 6: Sharpe Ratio and risk measures: The table shows the results of ordinary least squares regressions of funds' Sharpe Ratio (SR, expressed in *annual* units), relative volatility (RV, expressed as a *ratio*), and tracking error (TE, expressed in *annual percentage* units) on cognitive skills and economic preferences/attitudes. The Sharpe Ratio, SR, measures the abnormal return per unit of fund risk, the relative volatility, RV, accounts for the overall riskiness of the fund relative to the riskiness of the benchmark. The tracking error, TE, measures risk in the fund's return that is due to active management decisions, i.e., the standard deviation of the difference between the fund's net return and the benchmark return. Standard errors are clustered at the manager level. Corresponding p-values are reported in parentheses. ** $p < 0.005$, * $p < 0.05$.

	Sharpe Ratio		*Relative volatility*		*Tracking error*	
Dependent variable	SR	SR	RV	RV	TE	TE
Sample of funds	*Single*	*All*	*Single*	*All*	*Single*	*All*
CRT	0.087	0.127	−0.045**	−0.030**	0.484	0.469
	(0.104)	(0.110)	(0.002)	(0.004)	(0.541)	(0.433)
TOM	0.033	−0.065	−0.041*	−0.003	−0.355	−0.360
	(0.615)	(0.405)	(0.038)	(0.888)	(0.546)	(0.462)
Competitiveness	−0.032	0.052	0.018	0.003	0.067	1.445*
	(0.611)	(0.420)	(0.446)	(0.790)	(0.914)	(0.010)
Risk Tolerance	−0.017*	0.002	0.007*	0.003**	−0.010	0.060
	(0.019)	(0.875)	(0.016)	(0.004)	(0.882)	(0.304)
Loss Tolerance	−0.006	−0.004	0.003	0.000	0.114	0.061
	(0.598)	(0.766)	(0.558)	(0.973)	(0.203)	(0.454)
Ambiguity Tolerance	0.007	0.005	−0.001	−0.001	0.103*	0.088*
	(0.153)	(0.369)	(0.679)	(0.606)	(0.011)	(0.038)
$\min(R_{t-1}^{gross}, 0)$	0.006	−0.005	−0.013**	−0.014**	−0.845**	−0.822**
	(0.831)	(0.806)	(0.001)	(< 0.001)	(< 0.001)	(< 0.001)
$\max(R_{t-1}^{gross}, 0)$	−0.007	−0.016	0.013**	0.010**	0.226*	0.235**
	(0.680)	(0.416)	(< 0.001)	(< 0.001)	(0.006)	(0.002)
$\log(AUM)_{t-1}$	0.064	0.021	0.011	0.019*	−1.074**	−0.627**
	(0.116)	(0.470)	(0.314)	(0.015)	(0.001)	(0.004)
Experience	0.003	0.011	0.000	0.000	0.146	0.149
	(0.685)	(0.260)	(0.964)	(0.937)	(0.161)	(0.054)
Team		−0.047		0.041		−2.138*
		(0.726)		(0.152)		(0.023)
Constant	yes	yes	yes	yes	yes	yes
Category FE	yes	yes	yes	yes	yes	yes
Time FE	yes	yes	yes	yes	yes	yes
Number of observations	12,180	25,713	8,795	18,873	8,795	18,873
Number of managers	65	92	58	86	58	86
Adjusted R^2	0.426	0.432	0.156	0.117	0.582	0.556

Independent variables: CRT stands for the cognitive reflection score, comprised of 5 questions, measuring deliberate thinking. TOM stands for the "Reading-the-Mind-in-the-Eyes"-test, measuring theory of mind skills, i.e., the ability to infer the intention of others. The score for risk preferences (Risk Tolerance) reflects the elicited certainty equivalent for a lottery paying €60 or €0 with equal probability, with higher values indicating higher levels of risk tolerance. The measure for attitudes towards losses (Loss Tolerance) reflects the maximum potential loss subjects were willing to accept in order to have the chance of winning €22. Again, the higher the number, the more tolerant towards losses a fund manager is (Loss Tolerance). The score for ambiguity preferences (Ambiguity Tolerance) represents the matching probability (in %) that leaves subjects indifferent between a risky lottery with a certain probability of winning and an ambiguous lottery with an unknown probability of winning (both lotteries paid €60 in the case of winning and €0 else). Competitiveness is measured as the sum of the five standardized responses to the subscale of the WOFO, answered on scales ranging from 1 to 7 each. $\min(R_{t-1}^{gross}, 0)$ is zero if the gross return in the previous month was positive, and equal to previous month's gross return if it was negative; similarly, $\max(R_{t-1}^{gross}, 0)$ is zero if the gross return in the previous month was negative, and equal to previous month's gross return if it was positive. $\log(AUM)$ stands for the log of assets under management, Experience indicates years in industry, and Team is a dichotomous indicator for team-managed funds.

variables. We find no impact of all the other control variables on the relative riskiness of the fund to its benchmark.

Besides gaining insights on drivers of fund performance and risk, we also interpret our study as a test for the external validity of experimentally elicited economic preferences. The finding that our risk preference measure is significantly related to relevant measures of fund risk hints at the external validity of those experimentally elicited risk preferences. Our non-significant impact of loss aversion on fund risk (i.e., relative volatility), however, is in contrast to the result of Bodnaruk and Simonov (2016), showing that fund managers who exhibit high levels of loss aversion in a non-incentivized survey construct funds with lower downside risk and exhibit lower fund performance. As part of our robustness checks later, we also consider relative semi-volatility as a dependent variable, which is a measure of downside risk. We find that the results using relative semi-volatility are practically identical to those using relative volatility, i.e., there is a significant effect of Risk Tolerance but not of Loss Tolerance. The difference compared to the result of Bodnaruk and Simonov (2016) might be due to several differences in the experimental and empirical methodology between their study and ours. We interpret the findings as a sign that (generally defined) risk preferences have an impact on fund risk, and that it is not obvious how to distinguish experimentally elicited measures of risk aversion and loss aversion.

Result 3: *Fund managers' preferences for ambiguous outcomes are positively correlated with funds' tracking errors.*

Support: As shown in the fifth column of Table 6, we observe a pattern of ambiguity tolerance being positively related to the managed funds' tracking error. This pattern is significant both for the single-managed funds and for all funds (column 6), but the effect is weaker, as expected, when team-managed funds are included. This finding indicates that when fund managers are more tolerant towards ambiguous outcomes, they manage funds that deviate stronger from their benchmark compared to fund managers with lower levels of ambiguity tolerance. In economic terms our data predict that those fund managers that are close to ambiguity neutrality (i.e., matching probability of around 50% in the experiment) exhibit a tracking error that is 4.7 percentage points larger (in annual terms) compared to their peers with the lowest ambiguity tolerance.

Similar to patterns identified for the funds' relative volatility, we report that fund managers' tracking errors are dependent on previous months' return realizations (see column 5 in Table 6). We observe significantly negative coefficients for $\min(R_{t-1}^{gross}, 0)$, indicating that fund managers' tracking errors increase with increasing negative gross returns in the previous month. In addition, we find a similar, yet less pronounced effect on tracking errors of larger positive past gross returns. As for relative volatility, this hints at a V-shaped pattern of past return-based tracking errors with a stronger effect on the downside. Again, including a linear variable for past performance or removing both variables leaves the effects of the remaining explanatory variables unchanged. Taking the results on relative volatility and tracking error together, our findings suggest that fund managers facing negative gross returns in the previous month subsequently increase the risk of their portfolios, which in turn leads to higher tracking errors, i.e., higher deviations from the benchmark. Again, all findings hold for the full sample including team-managed funds as well (see column 6 in Table 6).

Moreover, we find a negative impact of fund size, $\log(AUM)_{t-1}$, on the tracking error, indicating that larger funds are managed with lower tracking errors. We also find that team-managed funds operate with lower tracking error, on average, than single-managed funds.

Robustness checks: We run several robustness checks *on the sample of single-managed funds.* The corresponding results are tabulated in Appendix D.

Our first robustness check is concerned with the number of observations per manager in a given month. It is typical for a manager to manage multiple funds at the same time. As a result, multiple fund observations can correspond to each manager-month combination in the main analysis. We adjust for this feature of the data by clustering the standard errors at the manager level in our main regressions. To offer an alternative solution, we follow Ibert et al. (2018) and create one observation for each manager-month by aggregating the variables across funds in this robustness check. To be in line with the single-managed sample, we start by keeping all manager-month combinations where at least half of the total assets managed by the individual is through single-managed funds.[22] Then, we create manager-month level observations by taking the weighted average of all the variables across the funds run by the manager in that month, where the weights correspond to the relative fund sizes (measured by AUM at the beginning of the month). To deal with team-managed funds, we divide the fund's AUM equally among all the managers in the fund.

Table S6 shows the results, which are very similar to the corresponding results from the main analysis (in Tables 5 and 6) both in terms of economic magnitude and statistical significance. The two notable changes are that Ambiguity Tolerance becomes significant at the 5% level in the abnormal return regression, while Risk Tolerance becomes insignificant at the 5% level in the Sharpe ratio regression. However, the changes in the coefficient estimates and p-values are not substantial in these two cases either. Overall, the results are robust to using a different approach for handling the fact that our participants manage multiple funds simultaneously.

The second robustness check provides alternative definitions for the fund's benchmark. In the main analysis we use the funds' prospectus benchmark to calculate all the performance measures that require return data on a benchmark. Alternatively, one can use a factor model to create "benchmark" returns for the fund (see, e.g., Carhart, 1997; Fama and French, 1993, 2015, for popular factor models). We strongly prefer our main approach of using the prospectus benchmark for at least two reasons. First, given the large heterogeneity of the funds in our sample with respect to country of domicile, allocation across asset classes, and geographical focus, it is ex-ante unclear what factors should be used, and the risk of using misspecified factor models is high. Second, the prospectus benchmark is likely the one that governs the behavior and decisions of fund managers the most, and therefore of interest to us. Nevertheless, in this section we present results using factor model-based benchmarks. The advantage of this approach is that we can also use funds where we were not able to identify a prospectus benchmark.

[22] The average (median) number of simultaneously managed funds is 2.6 (2). It is typical for single-managers to be "pure" single-managers. That is, in 85% of the remaining manager-month observations, our participant is the single manager in all his/her funds, and only 15% of the manager-month observations are such that our participant manages some of his/her funds alone and some as part of a team.

The first model we use is inspired by the Fama-French three-factor model and includes four factors: the European equity index return (MSCI) in excess of the one-month EURIBOR rate (the equity market factor), the European size and value factors from Kenneth French's data library, and the euro bond aggregate index return (Barclays) in excess of the one-month EURIBOR rate (the fixed-income factor).[23] We use European factors since all our funds in the sample are domiciled in Europe, and we include the fixed-income factor, since we cannot use only equity factors, as a lot of the funds are allocation and fixed-income funds. One might argue that using European factors is not adequate, since several funds in our sample have a non-European geographical focus. Therefore, the second model we consider is a global four-factor model that includes the global equity index return (MSCI) in excess of the one-month U.S. Treasury bill rate (global equity), the global bond aggregate index return (Barclays) in excess of the one-month U.S. T-bill rate (global fixed income), the European equity index return (MSCI) in excess of the one-month EURIBOR rate (European equity), and the euro bond aggregate index return (Barclays) in excess of the one-month EURIBOR rate (European fixed income). For each fund in our sample, the above two models are estimated using monthly return data if the fund has at least 24 monthly observations in our sample. This implies only those funds are included in this robustness check that have at least a two-year tenure in our sample.

After estimating the factor model, the model-based benchmark of a specific fund is considered to be the fitted values from the model without the intercept. Therefore, the abnormal return for fund i in month t, R_{it}^{abn}, will be the intercept (alpha) plus the residual from the regression. This also implies that the average abnormal return for the fund is the intercept from the regression. To have a measure corresponding to tracking error in this framework, we take daily fitted benchmark returns and calculate TE_{it} as in equation (6). Consequently, TE_{it} in this framework can also be interpreted as the idiosyncratic volatility of the fund relative to a specific factor-model.

Table S7 reports the results of the factor model-based benchmark approach. All the coefficients are insignificant for abnormal returns (columns 1 and 3), which is identical to Result 1 in the main analysis. The results for tracking error (columns 2 and 4) show that the coefficient on Ambiguity Tolerance remains significant, which is in line with Result 3 of the main analysis. Overall, the results on R_{it}^{abn} and TE_{it} seem generally robust to the alternative benchmark assignment approach that uses factor-models.

The third robustness check is concerned with the cross correlation between various measures of preferences and skills. As shown in Appendix A.2, the correlations between the experimental variables across the participants are generally low, but there are some significant ones, e.g., between Risk Tolerance and Loss Tolerance. Since all experimental variables are introduced simultaneously to the regressions in Tables 5 and 6, the possibility of spurious correlations might be a concern. Therefore, we also estimate the effect of the experimental variables from separate regressions. Table S8 reports the results; every reported coefficient comes from a different regression where the indicated variable is the single experimental measure, but control variables and fixed effects are the same as in Tables 5 and 6 for the corresponding dependent variable. Our findings remain robust both in terms of statistical significance and economic magnitude. The only notable difference is that the coefficient on Risk Tolerance

[23] We also considered a model, inspired by the Fama-French five-factor model, where we add the European investment and profitability factors to these. The (unreported) results are almost identical to the results shown in Appendix D.

in the Sharpe ratio regression loses significance at 5% level when estimated separately. However, its magnitude remains similar and it is still significant at the 10% level.

The fourth robustness exercise relates to the sample period. The financial crisis of 2008-2009 is included in the sample period of our main analysis, and one might be concerned that differences in our managers' reactions during the crisis might drive the overall results. Therefore, we re-estimate our main specifications with significant experimental variables excluding the financial crisis. In particular, we use the sample period 2010-2019. The results are reported in the first three columns of Table S9. Our findings remain robust both in terms of statistical significance and economic magnitude.

The fifth robustness check is connected to the experimental variable PATIENCE, measuring time preferences of our participants. As we pointed out previously, we lost a few observations of the experimental measure on inter-temporal preferences (10 participants) due to a runtime error issue in the experimental procedure. As is shown in the last three columns of Table S9, the significant experimental coefficients from the main analysis are robust to adding PATIENCE as an additional regressor. By omitting the variable PATIENCE from the main analysis, we can keep the 10 participants with the missing PATIENCE observations (around 10 percent of the sample).

The sixth robustness check is related to the experimental variable APM, measuring fluid intelligence. We drop APM from our main specifications due to its high correlation with CRT (see Table S1 in Appendix A.2), which leads to potential multicollinearity issues. In this robustness exercise, we re-estimate our main regressions with dropping CRT and adding APM instead. The results are reported in Table S10. Coefficient estimates on all other experimental variables remain robust both in terms of statistical significance and economic magnitude. That is, the decision to use CRT or APM in our regressions does not affect the results for the rest of the experimental variables. Regarding the coefficients on CRT and APM, there are two notable observations: First, CRT has a significant effect on relative volatility, while APM does not (compare column 3 of Table 6 with column 5 of Table S10). Nevertheless, both coefficient estimates are negative and point in the same direction. Second, APM has a significant effect on tracking error, while CRT does not (compare column 5 of Table 6 with column 6 of Table S10). However, the coefficient on APM is only marginally significant considering our benchmark significance level of 5%, and it does not remain significant in most of the robustness checks that we carry out in the paper.[24] Therefore, we refrain from further interpretation of this result.

The seventh robustness check is concerned with the period used for calculating the fund risk measures. In the main analysis we use daily returns over a month to calculate the Sharpe ratio (SR_{it}), relative volatility (RV_{it}), and tracking error (TE_{it}) measures. One might argue that one month of daily observations is not enough to accurately estimate these risk measures. In this robustness check we use daily data over 6-month periods to calculate the risk measures, e.g.,

$$RV_{ih} = \frac{Std\left(R^{net}_{id[h]}\right)}{Std\left(R^{B}_{id[h]}\right)}\,, \tag{7}$$

[24] In particular, the coefficient on APM is not significant at the 5% level (i) if one (aggregated) observation is used for each manager-month, (ii) if factor models are used to create benchmark returns for each fund, (iii) if PATIENCE is included as an additional independent variable, or (iv) if tracking error is calculated using daily returns over half-year periods (as in the next robustness exercise). The results of these robustness checks for APM are available upon request.

where $R_{id[h]}$ denotes daily returns within half-year h. The resulting unit of observation in these regressions is fund-half year (i.e., we have two observations within a year). Columns 1, 2, and 4 in Table S11 show that the results on SR, RV and TE are robust to calculating these measures on different frequencies.

Finally, we consider an alternative measure of fund riskiness, referred to as relative semi-volatility (RSV), which is based on the target semi-deviations of returns with the target return being zero:

$$RSV_{ih} = \frac{Std\left(R^{net}_{id[h]} \mid R^{net}_{id[h]} < 0\right)}{Std\left(R^{B}_{id[h]} \mid R^{B}_{id[h]} < 0\right)}. \tag{8}$$

Longer periods than a month are needed to reliably estimate RSV to ensure that there is a suitable number of negative return observations, and therefore we only include RSV in the analysis when the risk measures are calculated on a half-year frequency. Comparing the columns 2 and 3 of Table S11 to each other, we can see that the conclusions obtained when using relative semi-volatility as the dependent variable are very similar to those obtained when relative-volatility is used.

4. Discussion and Conclusion

In this paper, we address the question whether fund managers' cognitive skills, economic preferences, and attitudes toward competition can explain fund performance and dynamics. We study 92 fund managers, managing 412 mutual funds, from four large and mid-sized countries in the European Union. We matched the experimental data of the fund managers with the time series of the funds they manage and we controlled for various variables as, among others, the funds' benchmarks, fund categories, assets under management, and years of experience in all analyses.

First, we report that neither cognitive skills nor economic preferences turn out to have a systematic correlation with fund performance, neither in terms of abnormal returns nor in terms of value added. Second, we find a strong and positive relationship between fund managers' risk tolerance and fund volatility. This indicates that fund managers with low (high) levels of risk tolerance, on average, compose funds with lower (higher) fund volatility, relative to the benchmark. Importantly, this finding holds while controlling for fund managers' self-selection into different fund categories as well as for other economic preferences and cognitive skills, and for additional variables like fund benchmark, fund category, industry experience, and fund size. In addition, we provide evidence that risk tolerance is negatively related to a fund's Sharpe Ratio, implying that those fund managers with low risk tolerance earn higher risk-adjusted returns than their peers with high levels of risk tolerance. Moreover, we do find evidence that fund managers with high cognitive reflection abilities (CRT scores) manage funds with lower fund risk. Finally, we observe that fund managers with lower levels of ambiguity tolerance manage their funds with lower tracking errors compared to peers with higher levels of ambiguity tolerance. This indicates that ambiguity tolerance correlates with fund managers' propensity to "risk" deviating from the benchmark.

Our results have several implications for the fund industry. First, we show that measurable individual characteristics of fund managers indeed affect their professional decisions. In Appendix C, we provide further evidence that individual manager characteristics explain a considerable part of the variation in those fund specific variables that are used to obtain our main results, i.e., in fund volatility and tracking error. On the one hand, economic preferences correlate with relative fund risk and also risk-adjusted performance. On the other hand, cognitive abilities are related to relative fund risk. Our findings imply that abnormal fund performance does not depend on economic preferences and cognitive skills. However, our findings show that the relative riskiness of the fund seems to be the attribute fund managers influence with their risk preferences and cognitive skills.

Second, from a company perspective, a stronger focus on eliciting cognitive skills and particularly economic preferences among fund managers could achieve a better match of fund risk (fund strategy) and manager's economic preferences. However, at this stage we should be cautious in our interpretation as we cannot say much about persistence of our findings with our data set. This particularly applies for economic preferences, as studies by, for instance, Malmendier and Nagel (2011) and Guiso et al. (2018) show that extreme events have the power to shape risk preferences. This channel seems to be less relevant for cognitive skills, as evidence is showing that at least cognitive reflection skills are stable over time (Stagnaro et al., 2018). However, given these identification problems regarding causality in our data set, we leave the answer to the question of persistence and causality for future research. As indicated in the introductory section, we will analyze whether cognitive skills and economic preferences serve as predictors for (risk-adjusted) abnormal returns, fund risk, and tracking errors in a couple of years from now.

Third, our results indicate that fund risk and risk-adjusted performance depend to a certain degree on the fund manager's preferences and abilities. From a customer's perspective, this has important consequences, as it is unclear whether these patterns fit the investors' preferences (who is not informed about the economic preferences and cognitive abilities of the fund manager). This observation is in line with the experimental finding of Kirchler et al. (2020), showing that financial professionals' self-assessed risk attitude predominantly explains risk-taking on behalf of third parties (customers). Our finding also relates to the empirical observations of Foerster et al. (2017) and Linnainmaa et al. (2020). Both studies show that financial advisors invest their personal portfolios just like they advise their clients—i.e., they trade too much, chase returns, prefer expensive, actively managed funds, and hold underdiversified portfolios. Given this potential mismatch between customers' and fund managers' risk preferences and managers' impact on fund risk, one could think of efficiency losses on behalf of the customers. However, it is unclear from a practical perspective, whether regulation could circumvent this problem by achieving a better match of fund managers and customers with respect to economic preferences. This appears to be much easier implemented in the financial advisory industry, even though the practical implementation might be difficult there as well.

Fourth, from a scientific perspective, we interpret our study as a test for the correlation between the behavior observed in laboratory experiments and real-world professional decision-making. We add to the scientific debate by means of reporting an effect of experimentally elicited economic preferences and cognitive skills on fund risk. Moreover, we also consider our approach exploratory to a certain extent, potentially stimulating the development of theoretical models combining various economic preferences or cognitive skills into a unified framework. For instance, theorists could start with the objective function of fund managers and try to model the impact of various preferences as well as cognitive skills in a unified framework. In our experimental and empirical study, we have shown that some preferences seem to matter for fund risk and we therefore leave this theoretical advancement for future research. We consider this to be a fruitful avenue to take.

Finally, we emphasize that our findings represent correlational rather than causal evidence, as we cannot rule out that fund managers' preferences might have been shaped by performance and risk of their funds (cognitive reflection skills, on the other hand, seem a rather stable personality trait). Another limitation is potential survivorship bias in our data, since we do not have experimental data on the fund managers who dropped out of the industry during our observation window, potentially due to poor performance. We address this issue in part as our time frame includes fund data in close proximity to the experiment, both before and after it was conducted. However, our dataset might allow us to provide some evidence on this issue a few years from now, when we can model economic preferences and cognitive skills as predictors of *future* fund performance and fund risk only, while also controlling for the relationship between these individual traits and the likelihood of exiting of the industry. However, this can only be tentative evidence, because the sample of those leaving the industry might be too small. Thus, further research with large samples of fund managers is needed.

References

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

Adams, R., Keloharju, M., & Knüpfer, S. (2018). Are CEOs born leaders? lessons from traits of a million individuals. *Journal of Financial Economics*, *130*, 392–408.

Angrisani, M., Cipriani, M., Guarino, A., Kendall, R., & Ortiz de Zarate Pina, J. (2020). Risk preferences at the time of covid-19: An experiment with professional traders and students. *Federal Reserve Bank of New York Staff Report, No. 927.*

Bär, M., Kempf, A., & Ruenzi, S. (2011). Is a Team Different from the Sum of its Parts? Evidence from Mutual Fund Managers*. *Review of Finance*, *15*(2), 359–396.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" Test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, *42*(2), 241–251.

Berk, J. B. & van Binsbergen, J. H. (2015). Measuring skill in the mutual fund industry. *Journal of Financial Economics*, *118*(1), 1–20.

Bianchi, M. (2018). Financial literacy and portfolio dynamics. *The Journal of Finance*, *73*(2), 831–859.

Bianchi, M. & Tallon, J.-M. (2018). Ambiguity preferences and portfolio choices: Evidence from the field. *Management Science*, *65*(84), 1455–1947.

Bilker, W. B., Hansen, J. A., Brensinger, C. M., Richard, J., Gur, R. E., & Gur, R. C. (2012). Development of abbreviated nine-item forms of the raven's standard progressive matrices test. *Assessment*, *19*(3), 354–369.

Bliss, R. T., Potter, M. E., & Schwarz, C. (2008). Performance characteristics of individually-managed versus team-managed mutual funds. *The Journal of Portfolio Management*, *34*(3), 110–119.

Bodnaruk, A. & Simonov, A. (2016). Loss-averse preferences, performance, and career success of institutional investors. *The Review of Financial Studies*, *29*(11), 13–21.

Böhm, M. J., Metzger, D., & Strömberg, P. (2018). *Since you're so rich, you must be really smart: Talent and the finance wage premium.*

Boyle, P., Garlappi, L., Uppal, R., & Wang, T. (2012). Keynes meets markowitz: The trade-off between familiarity and diversification. *Management Science*, *58*, 253–272.

Bruguier, A. J., Quartz, S. R., & Bossaerts, P. (2010). Exploring the nature of "trader intuition". *The Journal of Finance*, *65*(5), 1703–1723.

Carhart, M. M. (1997). On persistence in mutual fund performance. *Journal of Finance*, *52*(1), 57–82.

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.

Chen, J., Hong, H., Huang, M., & Kubik, J. D. (2004). Does fund size erode mutual fund performance? the role of liquidity and organization. *American Economic Review*, *94*, 1276–1302.

Chevalier, J. & Ellison, G. (1999). Are some mutual fund managers better than others? cross-sectional patterns in behavior and performance. *The Journal of Finance*, *54*(3), 875–899.

Christelis, D., Jappelli, T., & Padula, M. (2010). Cognitive abilities and portfolio choice. *European Economic Review*, *54*(1), 18–38.

Cipriani, M., Filippis, R. D., Guarino, A., & Kendall, R. (2020). Trading by Professional Traders: An Experiment. *Federal Reserve Bank of New York Staff Report, No. 927.*

Cipriani, M. & Guarino, A. (2009). Herd behavior in financial markets: An experiment with financial market professionals. *Journal of the European Economic Association*, *7*(1), 206–233.

Cohn, A., Fehr, E., & Maréchal, M. A. (2014). Business culture and dishonesty in the banking industry. *Nature*, *516*, 86–89.

Corgnet, B., DeSantis, M., & Porter, D. (2018). What makes a good trader? on the role of intuition and reflection on trader performance. *The Journal of Finance*, *73*(3), 1113–1137.

Corgnet, B., Hernán-González, R., Kujal, P., & Porter, D. (2015). The effect of earned versus house money on price bubble formation in experimental asset markets. *Review of Finance*, *19*(4), 1455–1488.

DeMartino, B., O'Doherty, J., Ray, D., Bossaerts, P., & Camerer, C. (2013). In the mind of the market: Theory of mind biases value computation during financial bubbles. *Neuron*, *79*, 1222–1231.

Dimmock, S. G., Kouwenberg, R., Mitchell, O. S., & Peijnenburg, K. (2016a). Ambiguity aversion and household portfolio choice puzzles: Empirical evidence. *Journal of Financial Economics*, *119*(3), 559–577. doi:https://doi.org/10.1016/j.jfineco.2016.01.003

Dimmock, S. G., Kouwenberg, R., Mitchell, O. S., & Peijnenburg, K. (2016b). Prices in experimental asset markets under uncertainty. *Journal of Financial Economics*, *119*, 559–577.

Dorn, D. & Huberman, G. (2005). Talk and action: What individual investors say and what they do. *Review of Finance*, *9*, 437–481.

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.

Fama, E. F. & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, *33*(1), 3–56.

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

Fama, E. F. & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, *116*(1), 1–22.

Fang, Y. & Wang, H. (2015). Fund manager characteristics and performance. *Investment Analysts Journal*, *44*(1), 102–116.

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.

Gaechter, S., Johnson, E. J., & Herrmann, A. (2010). *Individual-level loss aversion in riskless and risky choices* (Discussion Papers No. 2010-20). The Centre for Decision Research and Experimental Economics, School of Economics, University of Nottingham. Retrieved from https://ideas.repec.org/p/cdx/dpaper/2010-20.html

Gastineau, G. (2004). The benchmark index etf performance problem. *Journal of Portfolio Management*, *30*, 96–103.

Golec, J. H. (1996). The effects of mutual fund managers' characteristics on their portfolio performance, risk and fees. *Financial Services Review*, *5*(2), 133–147.

Graham, J. R., Harvey, C. R., & Puri, M. (2013). Managerial attitudes and corporate actions. *Journal of Financial Economics*, *109*, 103–121.

Grinblatt, M., Keloharju, M., & Linnainmaa, J. T. (2011). IQ and stock market participation. *The Journal of Finance*, *66*(6), 2121–2164.

Grinblatt, M., Keloharju, M., & Linnainmaa, J. T. (2012). IQ, trading behavior, and performance. *Journal of Financial Economics*, *104*(2), 339–362.

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

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

Hefti, A., Heinke, S., & Schneider, F. (2018). Mental capabilities, heterogeneous trading patterns and performance in an experimental asset market. *SSRN Working Paper No. 234*.

Helmreich, R. L. & Spence, J. T. (1978). The work and family orientation questionnaire: An objective instrument to assess components of achievement motivation and attitudes toward family and career. *JSAS Catalog of Selected Documents in Psychology*, *8*(35).

Holzmeister, F. (2017). oTree: Ready-made apps for risk preference elicitation methods. *Journal of Behavioral and Experimental Finance*, *16*, 33–38.

Huber, C., Huber, J., & Kirchler, M. (2021). Market shocks and professionals' investment behavior - evidence from the covid-19 crash. *Journal of Banking and Finance*, *forthcoming*.

Ibert, M., Kaniel, R., Van Nieuwerburgh, S., & Vestman, R. (2018). Are mutual fund managers paid for investment skill? *Review of Financial Studies*, *31*(2), 715–772.

Kahneman, D. (2011). *Thinking, fast and slow*. Ney York: Farrar, Straus and Giroux.

Kaplan, S. N., Klebanov, M. M., & Sorensen, M. (2012). Which ceo characteristics and abilities matter? *The Journal of Finance*, *67*(3), 973–1007.

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

Larson, F., List, J. A., & Metcalfe, R. D. (2016). *Can myopic loss aversion explain the equity premium puzzle? evidence from a natural field experiment with professional traders* (Working Paper No. 22605). National Bureau of Economic Research.

Linnainmaa, J. T., Melzer, B., & Previtero, A. (2020). The misguided beliefs of financial advisors. *The Journal of Finance*, *forthcoming*.

Malmendier, U. & Nagel, S. (2011). Depression babies: Do macroeconomic experiences affect risk taking? *Quarterly Journal of Economics*, *126*(1), 373–416.

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

Massa, M., Reuter, J., & Zitzewitz, E. (2010). When should firms share credit with employees? evidence from anonymously managed mutual funds. *Journal of Financial Economics*, *95*(3), 400–424.

Meier, S. & Sprenger, C. D. (2015). Temporal stability of preferences. *The Review of Economics and Statistics*, *97*(2), 273–286.

Noussair, C. N., Tucker, S., & Xu, Y. (2016). Futures markets, cognitive ability, and mispricing in experimental asset markets. *Journal of Economic Behavior & Organization*, *130*, 166–179.

Olderbak, S., Wilhelm, O., Olaru, G., Geiger, M., Brenneman, M. W., & Roberts, R. D. (2015). A psychometric analysis of the reading the mind in the eyes test: Toward a brief form for research and applied settings. *Frontiers in Psychology*, *6*, 1503.

Patel, S. & Sarkissian, S. (2017). To group or not to group? evidence from mutual fund databases. *Journal of Financial and Quantitative Analysis*, *52*(5), 1989–2021.

Primi, C., Morsanyi, K., Donati, M. A., & Hamilton, J. (2015). The development and testing of a new version of the cognitive reflection test applying item response theory (irt). *Journal of Behavioral Decision Making*, *29*(5), 453–469.

Raven, J. (2000). The Raven's progressive matrices: Change and stability over culture and time. *Cognitive Psychology*, *41*(1), 1–48.

Razen, M., Kirchler, M., & Weitzel, U. (2020). Domain-specific risk-taking among finance professionals. *Journal of Behavioral and Experimental Finance*, *forthcoming*.

Riedl, A. & Smeets, P. (2017). Why do investors hold socially responsible mutual funds? *The Journal of Finance*, *72*(6), 2505–2550.

Sarin, R. K. & Weber, M. (1993). Effects of ambiguity in market experiments. *Management Science*, *39*(5), 602–615.

Stagnaro, M., Pennycook, G., & Rand, D. (2018). Performance on the cognitive reflection test is stable across time. *Judgment and Decision Making*, *13*, 260–267.

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

Weitzel, U., Huber, C., Huber, J., Kirchler, M., Lindner, F., & Rose, J. (2019). Bubbles and Financial Professionals. *The Review of Financial Studies*, *33*(6), 2659–2696.

Online Appendix

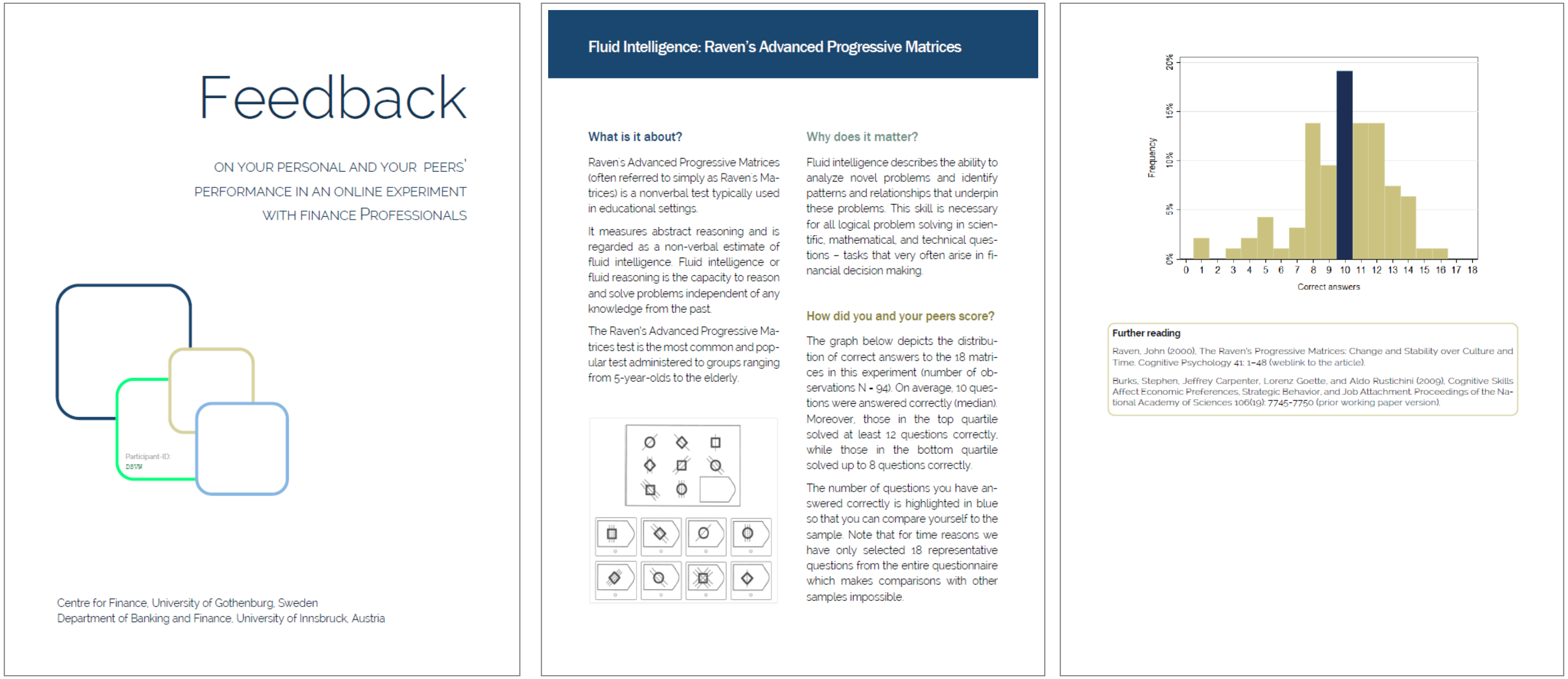
A. Details on the Experiment

In this section, the experimental tasks, the feedback map for the fund managers, and results of the various experimental tasks are described in more detail. The experiment has been conducted online using *oTree* (Chen et al., 2016). The software, including all instructions as used for the data collection, is available for download as a zipped *oTree* project at https://osf.io/dq3t8/ and as a live demo version via https://fea-2018-en.herokuapp.com. Participants have been invited via hard-copy letters and/ or e-mail, based on contact information available via funds' fact sheets, the webpages of institutions, and *Morningstar*.

A.1. Feedback Map

At the end of the experiment, fund managers could indicate whether they wished to receive personalized feedback (as a multi-page *.pdf-file distributed via e-mail) once the data collection has been completed. The feedback maps contained general information about each task and why the measured skill may potentially matter for financial decision-making. Moreover, subjects received their own scores as well as summary statistics about the performance of their peers participating in our experiment. Figure S1show the title page and, as one example, the feedback pages for Raven's Advanced Progressive Matrices (APM).[25]

[25] Note that the sample feedback shown in Figure S1 includes the full sample of 94 participants who completed the experiment. Since no fund data could be obtained for two participants, the remainder of this section refers to the sample of $n = 92$ fund managers.



Feedback

ON YOUR PERSONAL AND YOUR PEERS' PERFORMANCE IN AN ONLINE EXPERIMENT WITH FINANCE PROFESSIONALS

Participant-ID:

Centre for Finance, University of Gothenburg, Sweden
Department of Banking and Finance, University of Innsbruck, Austria

Fluid Intelligence: Raven's Advanced Progressive Matrices

What is it about?

Raven's Advanced Progressive Matrices (often referred to simply as Raven's Matrices) is a nonverbal test typically used in educational settings.

It measures abstract reasoning and is regarded as a non-verbal estimate of fluid intelligence. Fluid intelligence or fluid reasoning is the capacity to reason and solve problems independent of any knowledge from the past.

The Raven's Advanced Progressive Matrices test is the most common and popular test administered to groups ranging from 5-year-olds to the elderly.

Why does it matter?

Fluid intelligence describes the ability to analyze novel problems and identify patterns and relationships that underpin these problems. This skill is necessary for all logical problem solving in scientific, mathematical, and technical questions – tasks that very often arise in financial decision making.

How did you and your peers score?

The graph below depicts the distribution of correct answers to the 18 matrices in this experiment (number of observations N = 94). On average, 10 questions were answered correctly (median). Moreover, those in the top quartile solved at least 12 questions correctly, while those in the bottom quartile solved up to 8 questions correctly.

The number of questions you have answered correctly is highlighted in blue so that you can compare yourself to the sample. Note that for time reasons we have only selected 18 representative questions from the entire questionnaire which makes comparisons with other samples impossible.

Further reading

Raven, John (2000). The Raven's Progressive Matrices: Change and Stability over Culture and Time. Cognitive Psychology 41: 1–48 (weblink to the article).

Burks, Stephen, Jeffrey Carpenter, Lorenz Goette, and Aldo Rustichini (2009). Cognitive Skills Affect Economic Preferences, Strategic Behavior, and Job Attachment. Proceedings of the National Academy of Sciences 106(19): 7745–7750 (prior working paper version).

Figure S1: Feedback map: The figure shows the title page (left) and, as one example, the information and feedback provided for the Advanced Progressive Matrices task (APM; middle and right). Background information and feedback comparing the individual performance to all participating peers has been presented in a similar way for all other measures elicited in the experiment.

A.2. Details on the Experimental Tasks

Below, we provide further details on the experimental protocols of the tasks used to elicit cognitive abilities and economic preferences in the online experiment. The distributions of scores in the experimental tasks are depicted in Figure S3; correlations between the measures are summarized in Table S1.

Fluid Intelligence. Raven's Advanced Progressive Matrices (APM; Raven, 2000) are designed to measure fluid intelligence. We presented subjects with 18 increasingly difficult items (instead of the 36 items in the original version) where they had to infer the missing element of a given diagrammatic puzzle. In particular, we used every second item, starting with the first puzzle. For further details, we refer to the demo version of the software (https://fea-2018-en.herokuapp.com).

Cognitive Reflection Test. Cognitive reflection tests are designed to measure subjects' ability to consciously reflect on their intuitive responses. These types of tests were first established by Frederick (2005) and have been used widely since. To avoid potential recognition effects by the subjects, we decided to use questions from newer versions of the test proposed by Toplak et al. (2014) and Primi et al. (2015). Each question was displayed on a separate screen; the order has been randomized to avoid order effects. In particular, we included the following five questions (correct answers in parentheses):

- 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)
- Jerry received both the 15th highest and the 15th lowest mark in the class. How many students are in the class? (29 students)
- A man buys a pig for $60, sells it for $70, buys it back for $80, and sells it finally for $90. How much has he made? ($20)
- If three elves can wrap three toys in one hour, how many elves are needed to wrap six toys in two hours? (3 elves)
- In an athletics team, tall members are three times more likely to win a medal than short members. This year the team has won 60 medals so far. How many of these have been won by short athletes? (15 medals)

Theory of Mind. Theory of mind is a social sensitivity-skill that refers to the capacity of "reading the minds" of other people. In the "Reading-the-Mind-in-the-Eyes"-test introduced by Baron-Cohen et al. (2001), subjects have to infer the emotional state of a person from a picture showing only their eye region. In each of 18 trials, subjects had to select the correct emotion from a list of four adjectives. For each of the four potential answers, we provided participants with synonyms and an example sentence using the adjective (describing an emotional state) in an easy-to-understand context. To make sure participants understand the task, we implemented one practice trial, providing them with feedback about whether their choice has been correct. We used a subset of 18 pictures from the original menu of 36 pictures, out of which we took every second question, starting with the first one of the original task. The number of correctly chosen emotions serves as our measure of "reading the mind" skills (TOM). For further details, we refer to the demo version of the software (https://fea-2018-en.herokuapp.com).

Risk Preferences. The staircase risk elicitation method by Falk et al. (2018) allows to infer a subject's certainty equivalent for a given lottery, ensuring consistent answers. In four iterative, path-dependent questions, subjects decided between a lottery that pays €60 or €0 with equal probability and a certain payment that varies from question to question (see Figure S2 for a graphical representation of the task). Precisely, due to the limited number of iterations, the staircase approach yields intervals for the certainty equivalents. The midpoints of the intervals constitute our measure for subjects risk attitudes (Risk Tolerance). The task was implemented using the icl app put forward by Holzmeister (2017).

Time Preferences. We used the same staircase approach proposed by Falk et al. (2018) to measure time preferences, implemented via a modified version of the icl app of Holzmeister (2017). In four path-dependent questions, subjects had to decide whether they preferred a payment of €20 today or a certain higher amount in 6 months. The future premium increased in the next question when a subject opted for the payment today and decreased when the subject chose the future payment. Similarly to the risk elicitation task, this approach yields intervals for the time premia required by subjects to wait 6 months. The intervals' midpoints serve as our measure for participants' patience (Patience). For a facilitated interpretation, we compute the future premium subjects were willing to give up in order to receive the payment today (i.e., we multiplied the time premia by -1). Thus, higher values represent higher patience.

Loss Tolerance. The task to elicit participants' attitudes toward losses is based on the exercise proposed by Gaechter et al. (2010). However, to align the task with the experiments to elicit risk and time preferences, we transformed the elicitation procedure into an interactive, path-dependent series of questions, utilizing the icl app put forward by Holzmeister (2017). In each question, subjects decided whether they wished to participate in a lottery paying either €22 or some negative amount with equal probability. In the end, the task reveals intervals for each subject's maximum accepted loss in order to have the chance to win €22 (Loss Tolerance).

Ambiguity Tolerance. We followed the setup of Dimmock et al. (2016b) to elicit ambiguity tolerance. As in the original task, subjects had to choose between two urns containing 100 balls of blue and orange color each. At the end of the experiment, one ball was drawn randomly from the chosen urn. If the ball was blue, the subject would win €60; if the ball was orange, the subject would win nothing. While the distribution of blue and orange balls (i.e., the probability of winning) was known in the first urn, the probability was unknown for the second urn. In the first decision, the known distribution offered a 50% chance of winning. Subjects who chose the known distribution (risk) were presented a lower known probability of winning in the second question, while subject who chose the unknown distribution (ambiguity) were presented a higher known probability of winning in the second question. This procedure reveals a matching probability that leaves subjects indifferent between the risky and the ambiguous alternative, constituting our measure (Ambiguity Tolerance) of ambiguity preferences.

Competitiveness. The Work and Family Orientation (wofo) questionnaire of Helmreich and Spence (1978) is a widely used psychometric measure of individuals' competitiveness. Subjects answered how strongly they agree with a certain statement about their attitudes towards competition on a scale from 1 to 7. The competitiveness score (Competitiveness) is then computed as the sum of the individual answers. In particular, participants answered the following five questions:

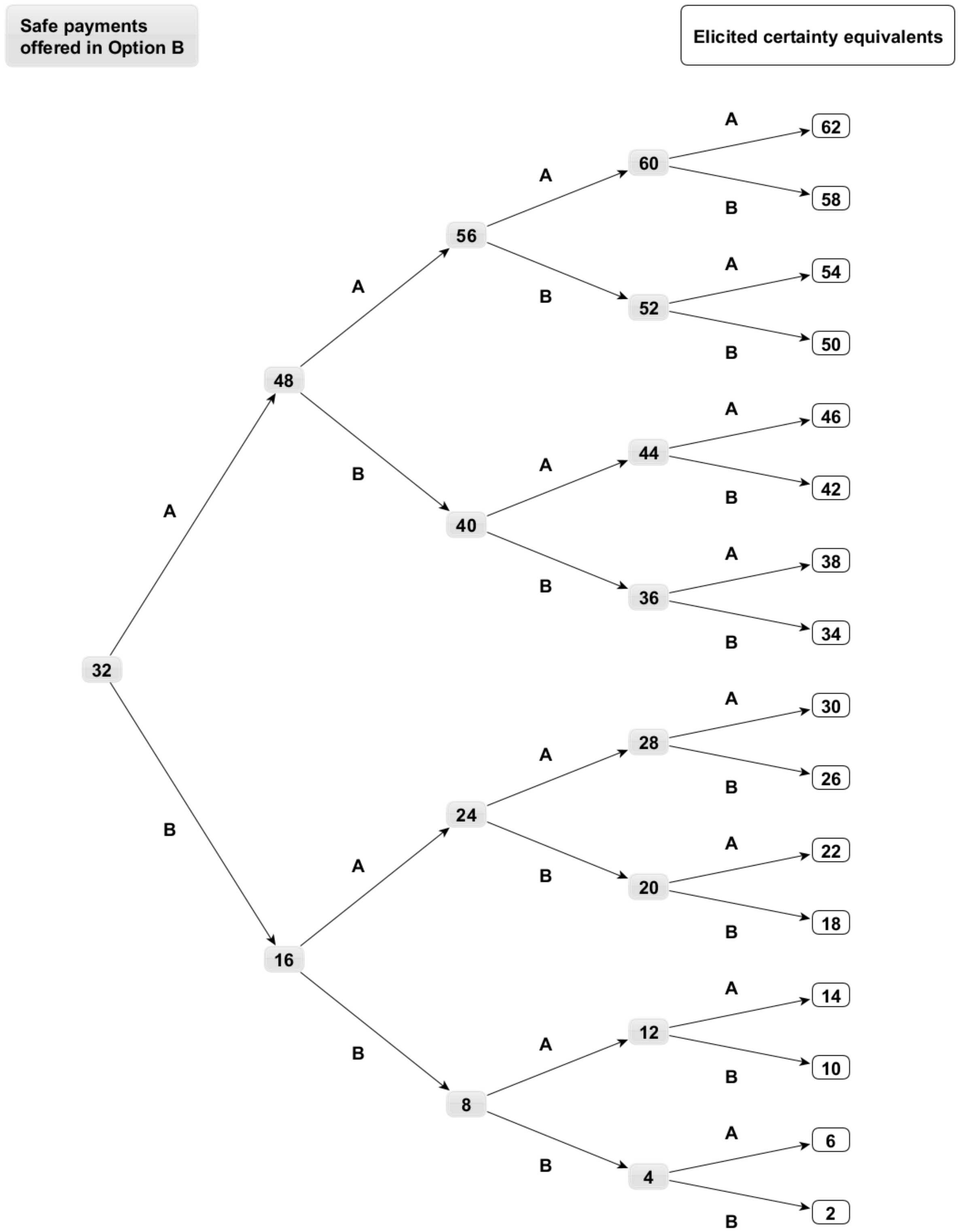


Figure S2: The staircase risk preference elicitation procedure is based on Falk et al. (2018). The final column shows the midpoints of the elicited intervals for the certainty equivalents (±2 Euros). The iterative methods for eliciting attitudes towards time discounting, losses, and ambiguous outcomes have been implemented based on the same structuring.

- I enjoy working in situations involving competition with others.
- It is important to me to perform better than others on a task.
- I feel that winning is important in both work and games.
- It annoys me when other people perform better than I do.
- I try harder when I'm in competition with other people.

A.2.1. Correlation Between the Experimental Variables

Table S1: Pearson correlation coefficients between the experimental variables. ** $p < 0.005$, * $p < 0.05$. $n = 92$.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) CRT	1.00						
(2) APM	0.49**	1.00					
(3) TOM	0.13	0.02	1.00				
(4) RISK TOLERANCE	0.07	0.14	−0.01	1.00			
(5) LOSS TOLERANCE	−0.03	0.00	−0.11	0.24*	1.00		
(6) AMBIGUITY TOLERANCE	0.07	0.22*	0.02	0.06	0.01	1.00	
(7) COMPETITIVENESS	0.15	0.13	0.06	−0.09	−0.11	0.05	1.00

Note: CRT stands for the cognitive reflection score, comprised of 5 questions, measuring deliberate thinking. TOM stands for the "Reading-the-Mind-in-the-Eyes"-test, measuring theory of mind skills, i.e., the ability to infer the intention of others. The score for risk preferences (RISK TOLERANCE) reflects the elicited certainty equivalent for a lottery paying €60 or €0 with equal probability, with higher values indicating higher levels of risk tolerance. The measure for attitudes towards losses (LOSS TOLERANCE) reflects the maximum potential loss subjects were willing to accept in order to have the chance of winning €22. Again, the higher the number, the more tolerant towards losses a fund manager is (LOSS TOLERANCE). The score for ambiguity preferences (AMBIGUITY TOLERANCE) represents the matching probability (in %) that leaves subjects indifferent between a risky lottery with a certain probability of winning and an ambiguous lottery with an unknown probability of winning (both lotteries paid €60 in the case of winning and €0 else). COMPETITIVENESS is measured as the sum of the five standardized responses to the subscale of the WOFO, answered on scales ranging from 1 to 7 each.

B. Details on the Fund Data

B.1. Matching Funds to Participants

We start with the 94 fund managers who participated in our experiments, and use the *Morningstar Direct* database to match the managers to their funds. We obtain time series data on all open-ended mutual funds (both active and inactive) that have one of our four countries registered as "Domicile" or "Region of Sale" in the database. *Morningstar* provides the manager history for each fund, which contains the full name of the fund's managers (current and past) together with the tenure for each manager (start date and end date).

In a first step, we searched for all the funds (i) where the full name of one of our participants appears as the fund's manager during the period from January 2008 to December 2019, and (ii) where the fund company matches the participant's workplace at the time of the experiment. We manage to match at least one fund to 93 of our 94 participants and identify 411 funds altogether.[26] In a second step, we

[26] For one participant, we are not able to match any of the funds. According to this participant's LinkedIn profile, she/he started to work for the given company as a portfolio manager in mid 2016 and finished early 2018. She/he might have been too junior to be listed as a manager for any particular fund during this time, hence she/he does not show up in our fund data.

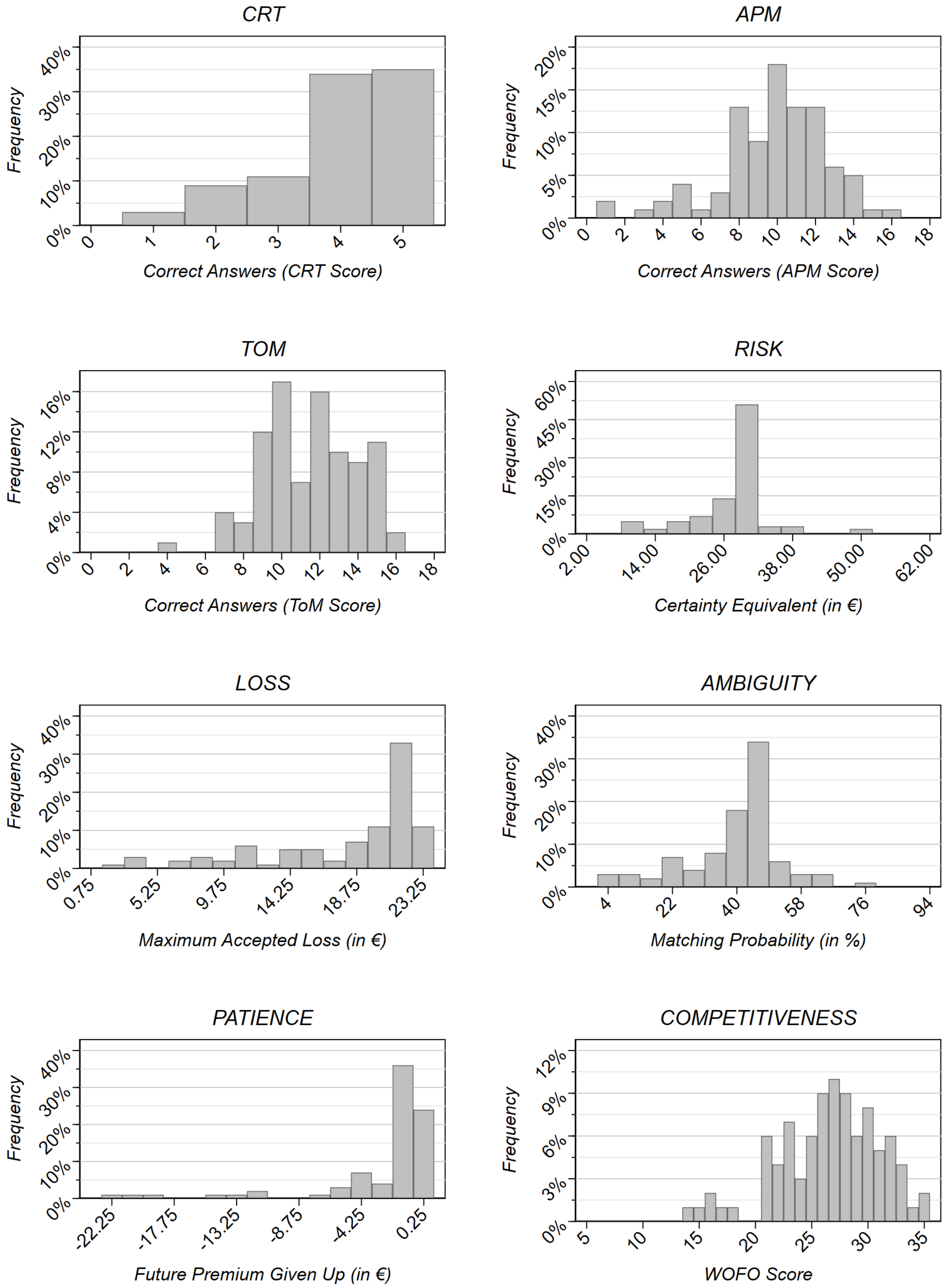


Figure S3: Histograms of the eight tasks on cognitive skills, economic preferences, and attitudes towards competitiveness. The values for economic preferences represent midpoints of the corresponding elicited intervals. The scaling of the x-axes ranges from the potential minimum to the potential maximum value of the task. $n = 92$.

Table S2: Number of (non-missing) fund-month observations for each variable.

	Fund-month observations	Number of Funds
Sample	28,369	412
Assets under management (AUM)	26,436	401
Net return (in €)	27,070	406
Total Expense Ratio (ER)	26,962	405

augment the compiled list with funds where the full name of one of our participants appears as manager during the sample period, but the fund company does not match the participant's workplace at the time of the experiment if two conditions are satisfied: (i) the manager's tenure at this fund ended before December 2019, and (ii) we are able to verify that it is likely to be the same person (e.g., through the manager's biography on the current employer's website or the manager's LinkedIn profile). This results in 22 additional funds. In a third step, following the literature (see, e.g., Ibert et al., 2018)), we eliminate money market mutual funds (5 funds identified as such by the variable "Broad Category Group" in *Morningstar*) and index funds (9 funds identified as such by *Morningstar* or by the word "index" in their name). Finally, we exclude 7 funds for which we are not able to obtain any fund-level time-series data (the details of collecting the fund-level data is described in subsection B.2). Our resulting base sample consists of 412 funds managed by 92 managers for a total of 28,369 unique fund-month observations.[27] Note that we lose one additional manager when excluding the money market funds, index funds, and funds with no time-series data.

To ensure anonymity, after matching the participants to their funds, the managers' names are replaced with randomly generated, unique identifiers to match the de-personalized fund data with the experimental data at a later stage.

B.2. Monthly and Daily Fund-Level Data

In this section we describe how the fund-level data on assets under management, returns, and expense ratios are obtained. Mutual funds in our sample can have multiple share classes. We start with data on the share classes (identified by ISIN code) and then aggregate to the fund level. Table S2 provides a brief description of the number of non-missing fund-month observations for each variable. In the following subsections, we provide details for the three variables of interest for our research questions.

B.2.1. Assets Under Management

We start by assembling share class level assets under management (AUM) data using the following steps:

- **Step 1**: We retrieve share class level data from *Morningstar*.
- **Step 2**: We also retrieve share class level data from *Lipper*. For share classes with no AUM data at all in *Morningstar* (i.e., the whole AUM series is missing), we use AUM values from *Lipper*, if available.

[27] For a given manager-fund observation, if the start date (end date) does not coincide with the first (last) day of the month, we include that month in the manager's tenure if she/he was the manager for at least 20 calendar days during that month.

Table S3: Summary of steps involved for creating fund level observations for AUM.

		Frequency	Percent
Step 1:	Only *Morningstar* data	25,061	94.80%
Step 2:	Only *Lipper* data	136	0.51%
Step 3:	*Morningstar* and *Lipper* data	1,137	4.30%
Step 4:	Imputations needed	102	0.39%

- **Step 3**: For share classes where we have AUM data from both *Morningstar* an *Lipper* we calculate

$$AUMratio_{it} = \frac{AUM_{it}^{Morningstar}}{AUM_{it}^{Lipper}} \tag{9}$$

  for months where both datasets report an observation. Then we replace the missing *Morningstar* values with *Lipper* values for the share classes where $0.99 < \overline{AUMratio_i} < 1.01$, i.e., where the AUM values from the two datasets are very close to each other in those months when both datasets report a value.

- **Step 4**: We impute the missing AUM values following Ibert et al., 2018. Only missing values in the middle of AUM series are imputed using past AUM values, the return on the share class, and a factor that adjusts for flow rates. Let $[t_0, t]$ and $[t+n, T]$ be periods where the share class has AUM data, i.e., $[t+1, t+n-1]$ is the period with missing values. The missing values are then filled according to

$$AUM_{ik} = F_i \cdot AUM_{ik-1}\,(1 + R_{ik}) \qquad \text{for } k \in [t+1, t+n-1] \tag{10}$$

$$F_i = \left(\frac{1}{\prod_{k=t+1}^{t+n}(1+R_{ik})} \cdot \frac{AUM_{it+n}}{AUM_{it}} \right)^{\frac{1}{n}}, \tag{11}$$

  where i denotes the share class, F_i is the factor adjusting for flow rate during the missing period, and R_{ik} is the net return on the share class in month k.

After assembling the share class level AUM data, we aggregate it to the fund level. That is, for each fund-month observation, we aggregate AUM across all share classes of the fund to get the fund-level AUM. Altogether, we have 26,436 non-missing fund-month AUM observations. Table S3 summarizes which of the above steps are involved when creating the fund-level observations. The table reveals that the majority of the fund-month observations (94.8%) rely solely on the *Morningstar* database.

B.2.2. Net Return

Monthly frequency. We retrieve share class level monthly net returns from *Morningstar.* All returns are converted into Euro for comparability across different countries. *Morningstar* has a very thorough coverage of returns, so we do not augment the data with additional sources. We aggregate the share class level returns to the fund level. Altogether, we have 27,070 non-missing fund-month return observations. When aggregating to the fund level, we have the following options:

- If the fund has only one share class in the given month, we take the return on the single share class (52.3% of the observations).

- If the fund has multiple share classes in the given month but we observe the same return across all share classes with a return observation, we take this common return (10.2% of the observations).
- If the fund has multiple share classes in the given month and the return varies across the share classes, we take the AUM-weighted return across the share classes (37.5% of the observations).

Daily frequency. In order to measures the riskiness of mutual funds, we rely on daily fund returns. Therefore, we also retrieve share class level daily net returns from *Morningstar* for the relevant funds over the sample period. All returns are converted into Euro for comparability across different countries. Then we aggregate the returns to the fund-level, yielding a total of 623,364 non-missing fund-day return observations. When aggregating to the fund level, similar to the case of the monthly returns, we have the following options:

- If the fund has only one share class in the given month, we take the return on the single share class (54.1% of the observations).
- If the fund has multiple share classes in the given month but we observe the same return across all share classes with a return observation, we take this common return (5.7% of the observations).
- If the fund has multiple share classes in the given month and the return varies across the share classes, we take the AUM-weighted return across the share classes (40.1% of the observations).
- If the fund has multiple share classes in the given month and we have return observations on multiple share classes but no AUM observations on any of the share classes, we take the simple average of the return across the share classes (0.1% of the observations).

When AUM data is needed, we use the monthly AUM described in Section B.2.1.

B.2.3. Total Expense Ratio

Expense ratios (ER) are reported yearly in both *Morningstar* and *Lipper*, so we start by assembling yearly share class level ER data using the following steps:

- **Step 1**: We retrieve share class level data from *Morningstar*.
- **Step 2**: We also retrieve share class level data from *Lipper*. For share classes with no ER data at all in *Morningstar* (i.e., the whole ER series is missing), we use ER values from *Lipper*, if available.
- **Step 3**: For share classes where we have ER data from both *Morningstar* and *Lipper* we calculate

$$ERdiff_{it} = \left| ER_{it}^{Morningstar} - ER_{it}^{Lipper} \right|, \tag{12}$$

  for years where both datasets have an observation. Then we replace the missing *Morningstar* values with *Lipper* values for the share classes where $\overline{ERdiff}_i < 10bps$ and $\max\left(ERdiff_{it}\right) < 25bps$, i.e., where the mean difference is not larger than 10 basis points per year and the maximum difference is not larger than 25 basis points per year during those years when they both datasets report a value.
- **Step 4**: For share classes where the ER series from *Mornigstar* stops earlier than the ER series from *Lipper*, and the two databases had exactly the same value for the last available joint observation, we use the *Lipper* values for the remaining period.

After the steps outlined above, we use imputations similar to Ibert et al., 2018:

- **Step 5**: For share classes where the ER series is constant, i.e., the smallest ER is equal to the largest ER for all existing observations, the missing ER observations are filled with this constant value.
- **Step 6**: For share classes that have missing values in the middle of ER series, the missing values are imputed using past ER values and ER growth rates. Let $[t_0, t]$ and $[t+n, T]$ be periods where the share class has ER data, i.e., $[t+1, t+n-1]$ is the period with missing values. The missing values are then filled according to

$$ER_{ik} = \left(\frac{ER_{it+n}}{ER_{it}}\right)^{\frac{1}{n}} \cdot ER_{ik-1} \quad \text{for } k \in [t+1, t+n-1]\,, \tag{13}$$

  where i denotes the share class.
- **Step 7**: For share classes that have missing values at the tails of the ER series, we test if the ER series follow a linear time trend. If they do, we replace the missing ER values with the forecast values from the linear model. Let $[t_0, t]$ and $[t+n, T]$ be periods where the share class has missing ER observations, i.e, $[t+1, t+n-1]$ is the period with ER data. We estimate the model

$$\log\left(ER_{ik}\right) = a_i + b_i k + \varepsilon_{ik} \quad \text{for } k \in [t+1, t+n-1]\,, \tag{14}$$

  and fill the missing ER values

$$ER_{ik} = \exp\left(\hat{a}_i + \hat{b}_i k\right) \quad \text{for } k \in [t_0, t] \cup [t+n, T]\,. \tag{15}$$

  if the p-value of $\hat{b}_i$ is less than or equal to 5% and $n \geq 6$. If these conditions are violated, we fill the missing ER values at the left (right) tail of the series with the mean values of the first (last) three ER values.

After the above steps, we create monthly share class level data by assigning $\frac{ER_{it}}{12}$ to each month of year t for share class i (similar to Ibert et al., 2018). Then we aggregate the share class-level expense ratios to the fund level. Altogether, we have 26,962 non-missing monthly ER observations. When aggregating to the fund level, we have the following options:

- If the fund has only one share class in the given month, we take the ER on the single share class (43.4% of the observations).
- If the fund has multiple share classes in the given month but we observe the same ER across all share classes with a ER observation, we take this common ER (22.8% of the observations).
- If the fund has multiple share classes in the given month and ER varies across the share classes, we take the AUM-weighted ER across the share classes (33.4% of the observations).
- If the fund has multiple share classes in the given month and we have ER observations on multiple share classes but no AUM observations on any of the share classes, we take the simple average of ER across the share classes (0.4% of the observations).

Table S4 summarizes how much we rely on the different data sources and imputations when creating the fund level observations. The table reveals that we have to rely more heavily on *Lipper* and imputations to gather the ER data, than for returns and AUM.

Table S4: Summary of steps involved for creating fund level observations for ER.

		Frequency	Percent
Step 1:	Only *Morningstar* data	8,736	32.40%
Step 2:	Only *Lipper* data	1,419	5.26%
Step 3–4:	*Morningstar* and *Lipper* data	9,467	35.11%
Step 5–7:	Imputations needed	7,340	27.22%

B.3. Benchmark Assignment

Prospectus benchmark. A prospectus benchmark is reported in *Morningstar* for 255 funds in our sample (62% of all funds). We find monthly returns for the benchmark indices, expressed in Euro, on *Morningstar*, *Lipper*, or *Datastream*. Then the returns on the benchmarks are assigned to the respective funds. Note that some funds have linear combinations of indices as their benchmark. When the benchmark is a linear combination of different indices, the benchmark return is only calculated for those fund-months when return data on all benchmark constituents are available. Altogether, we are able to assign a prospectus benchmark return to 13,159 fund-month observations, covering 206 funds (81% of the 255 that have a reported prospectus benchmark and 50% of our full sample).

Lipper benchmark. *Lipper* independently assigns the "Lipper Technical Indicator Benchmark" to most of the funds in the database according to its assessment of the fund's investment strategy. The technical indicator benchmark is assigned to 374 funds in our sample (91% of all funds). We find monthly returns for the benchmark indices, expressed in Euro, on *Morningstar*, *Lipper*, or *Datastream*. Altogether, we are able to assign a *Lipper* benchmark return to 17,111 fund-month observations, covering 262 funds (70% of the 374 that have an assigned *Lipper* benchmark and 64% of our full sample).

Hand assigned benchmark. We also assign benchmarks "by hand" to equity funds in our sample (identified as such by the variable "Broad Category Group" in *Morningstar*). For each equity fund category defined by the *Morningstar* variable "Category", we find the most common benchmark among all open-ended mutual funds (both active and inactive) that have one of our four countries registered as "Domicile". This most common benchmark is assigned to all the funds in the given category. Altogether, we have a "manually" assigned benchmark return for 11,121 fund-month observations, covering 168 funds.

C. Variation Explained by Manager Fixed Effects

In this appendix, we study how much of the variation in our dependent variables is explained by time-invariant manager characteristics. In particular, we examine the increase in (adjusted) R^2-values when adding manager fixed effects to the regression models that otherwise only include time-varying covariates. This exercise helps us understand the importance of manager characteristics (which include, among many other attributes, cognitive skills and economic preferences elicited in our experiments) in explaining the variation that we see in various facets of fund performance. We use our main sample of single-managed funds.

Table S5: Adjusted R-squared values of regressions with time-, manager-, and fund fixed effects: The table shows adjusted R^2-values from ordinary least squares regressions of funds' performance and risk characteristics on various sets of fixed effects. R^{gross} is the gross returns of the fund, R^{abn} is the fund's gross returns over its benchmark, and V is value added. SR denotes the Sharpe Ratio, RV is the overall volatility of the fund relative to the volatility of its benchmark, and TE is the tracking error. The set(s) of fixed effects used in the regressions is showed in the first column.

	R^{gross}	R^{abn}	V	SR	RV	TE
Time FE	43.6%	8.5%	2.9%	42.0%	6.5%	41.7%
Time + Manager FE	44.0%	9.0%	4.8%	44.1%	36.6%	73.9%
Time + Fund FE	44.1%	9.7%	5.6%	44.8%	55.3%	79.8%

Independent variables: lagged $\log(AUM)$ in the regressions where the dependent variable is R^{gross} or R^{abn}. Lagged $\log(AUM)$ together with $\min(R^{gross}_{t-1}, 0)$ and $\max(R^{gross}_{t-1}, 0)$ in the regressions where the dependent variable is SR, RV, or TE.

We also compare the effect of manager fixed effects to that of fund fixed effects. Note that there are considerably more funds than managers in the sample, and in the vast majority of the cases the dichotomous variable corresponding to a specific manager is the sum of a collection of fund indicators, because all the funds belong to only that one manager in the data.[28] This structure of the data implies that the fund fixed effects are expected to explain a larger part of the cross-sectional variation in the dependent variables than the manager fixed effects. However, the comparison is still useful for assessing the relative importance of manager characteristics.

Table S5 presents adjusted R^2-values from three regressions for each dependent variable used in our paper. For the regressions in the first row of Table S5, time (year-month) fixed effects are used alongside time varying fund characteristics as control variables. Manager fixed effects are added in the second row, and fund fixed effects are added in the third row. In case of the performance-related dependent variables (R^{gross}, R^{abn}, and V), neither the added manager fixed effects, nor the added fund fixed effects increase the R^2-values considerably. Likewise, we do not observe a substantial increase in R^2 when focusing on the funds' Sharpe Ratio (SR). In light of this, it is not surprising that we do not find any significant effects of the experimental measures in the main analysis for these dependent variables. In case of the risk-related dependent variables (RV and TE), however, manager- and fund fixed effects give rise to a substantial increase in the R^2-values. This is reassuring, as our main results are based on these dependent variables. Given the above discussion on the expected explanatory power of the manager- versus fund fixed effects, the results in Table S5 show that manager characteristics are highly relevant. The increase in (adjusted) R^2 is considerably bigger going from the first row to the second, compared to the increase when going from the second to the third row, especially in the case of TE.

[28] There are 65 managers in the single-managed fund sample, who altogether have managed 209 funds for some time during the period 2008–2019. Among these, there are only 5 funds for which more than one of our participants were managers for some time during this time period.

D. Results of Robustness Checks

Table S6: Robustness check for handling multiple funds run by a manager: The table shows the results of ordinary least squares regressions of funds' performance and risk characteristics on cognitive skills and economic preferences/attitudes. R^{gross} is the gross returns of the fund without benchmark correction, R^{abn} is the fund's gross returns over its benchmark, and V is value added - the product of assets under management and abnormal returns. SR denotes the Sharpe Ratio, RV is the overall volatility of the fund relative to the volatility of the benchmark, and TE is the tracking error measured as the standard deviation of the difference between the fund's net return and the benchmark return. Manager-month level observations are used, which are created by taking the weighted average of each dependent variable across the funds run by a manager in a given month. The weights are the relative fund sizes measured by AUM at the beginning of the month. Standard errors are clustered at the manager level. Corresponding p-values are reported in parentheses. ** $p < 0.005$, * $p < 0.05$.

Dependent variable	R^{gross}	R^{abn}	V	SR	RV	TE
CRT	−0.093	−0.528	−1.870	0.025	−0.055*	−0.303
	(0.862)	(0.229)	(0.456)	(0.625)	(0.022)	(0.705)
TOM	−0.557	−0.106	−0.540	−0.054	−0.031	0.081
	(0.233)	(0.745)	(0.570)	(0.438)	(0.203)	(0.887)
Competitiveness	−0.348	−0.064	1.048	0.000	0.037	0.302
	(0.397)	(0.860)	(0.335)	(1.000)	(0.160)	(0.615)
Risk Tolerance	−0.094	−0.102	−0.139	−0.013	0.007*	0.108
	(0.064)	(0.137)	(0.266)	(0.072)	(0.006)	(0.114)
Loss Tolerance	0.056	0.061	0.016	−0.005	0.000	−0.010
	(0.518)	(0.419)	(0.935)	(0.659)	(0.994)	(0.924)
Ambiguity Tolerance	0.024	0.059*	0.092	0.002	0.001	0.105**
	(0.308)	(0.027)	(0.165)	(0.743)	(0.713)	(0.001)
$\log(AUM)_{t-1}$	0.792	0.412		0.109*	−0.003	−1.632**
	(0.087)	(0.236)		(0.033)	(0.857)	(0.001)
Experience	−0.054	0.066	0.155	0.007	0.003	0.017
	(0.360)	(0.234)	(0.411)	(0.448)	(0.354)	(0.860)
Constant	yes	yes	yes	yes	yes	yes
Category FE	yes	yes	yes	yes	yes	yes
Time FE	yes	yes	yes	yes	yes	yes
Number of observations	5,140	4,469	4,465	5,136	4,404	4,404
Number of managers	63	58	58	63	57	57
Adjusted R^2	0.435	0.119	0.045	0.457	0.153	0.582

Independent variables: CRT stands for the cognitive reflection score, comprised of 5 questions, measuring deliberate thinking. TOM stands for the "Reading-the-Mind-in-the-Eyes"-test, measuring theory of mind skills, i.e., the ability to infer the intention of others. The score for risk preferences (Risk Tolerance) reflects the elicited certainty equivalent for a lottery paying €60 or €0 with equal probability, with higher values indicating higher levels of risk tolerance. The measure for attitudes towards losses (Loss Tolerance) reflects the maximum potential loss subjects were willing to accept in order to have the chance of winning €22. Again, the higher the number, the more tolerant towards losses a fund manager is (Loss Tolerance). The score for ambiguity preferences (Ambiguity Tolerance) represents the matching probability (in %) that leaves subjects indifferent between a risky lottery with a certain probability of winning and an ambiguous lottery with an unknown probability of winning (both lotteries paid €60 in the case of winning and €0 else). Competitiveness is measured as the sum of the five standardized responses to the subscale of the WOFO, answered on scales ranging from 1 to 7 each. $\log(AUM)_{t-1}$ is the lagged log of total assets under management, and Experience is the years the manager spent in the industry.

Table S7: Robustness check for abnormal returns and tracking error using factor-model based benchmarks: The table shows the results of ordinary least squares regressions of funds' abnormal returns and tracking error on cognitive skills and economic preferences/attitudes. The abnormal returns, R^{abn}, are the funds' monthly gross returns over their benchmark. The tracking error, TE, measures standard deviation of the difference between the fund's net return and the benchmark return. For each fund, the benchmark return is calculated by estimating a factor model on the fund's returns and taking the fitted values (without the intercept) from the model. We consider two factor models: a global four-factor model (containing global equity, global fixed income, European equity, European fixed income factors) and a European Fama-French 3-factor type model (containing European equity market, size, value, and fixed income factors). Standard errors are clustered at the manager level. Corresponding p-values are reported in parentheses. ** $p < 0.005$, * $p < 0.05$.

	Global four-factor model		*Fama-French (Europe) model*	
Dependent variable	R^{abn}	TE	R^{abn}	TE
Sample of funds	*Single*	*Single*	*Single*	*Single*
CRT	0.113	−0.304	0.044	−0.237
	(0.691)	(0.561)	(0.885)	(0.623)
TOM	0.232	−0.233	0.171	−0.201
	(0.209)	(0.520)	(0.460)	(0.538)
Competitiveness	−0.079	0.252	−0.238	0.180
	(0.717)	(0.507)	(0.326)	(0.579)
Risk Tolerance	−0.055	0.037	−0.047	0.036
	(0.071)	(0.283)	(0.090)	(0.259)
Loss Tolerance	0.032	0.115	0.059	0.125*
	(0.449)	(0.066)	(0.169)	(0.031)
Ambiguity Tolerance	0.019	0.056*	0.023	0.052*
	(0.197)	(0.036)	(0.191)	(0.024)
Constant & Controls	yes	yes	yes	yes
Category FE	yes	yes	yes	yes
Time FE	yes	yes	yes	yes
Number of observations	11,920	11,766	11,920	11,766
Number of managers	62	62	62	62
Adjusted R^2	0.101	0.642	0.127	0.636

Independent variables: CRT stands for the cognitive reflection score, comprised of 5 questions, measuring deliberate thinking. TOM stands for the "Reading-the-Mind-in-the-Eyes"-test, measuring theory of mind skills, i.e., the ability to infer the intention of others. The score for risk preferences (Risk Tolerance) reflects the elicited certainty equivalent for a lottery paying €60 or €0 with equal probability, with higher values indicating higher levels of risk tolerance. The measure for attitudes towards losses (Loss Tolerance) reflects the maximum potential loss subjects were willing to accept in order to have the chance of winning €22. Again, the higher the number, the more tolerant towards losses a fund manager is (Loss Tolerance). The score for ambiguity preferences (Ambiguity Tolerance) represents the matching probability (in %) that leaves subjects indifferent between a risky lottery with a certain probability of winning and an ambiguous lottery with an unknown probability of winning (both lotteries paid €60 in the case of winning and €0 else). Competitiveness is measured as the sum of the five standardized responses to the subscale of the WOFO, answered on scales ranging from 1 to 7 each. Controls refer to the nonlinear transformations of the past gross return ($\min(R^{gross}_{t-1}, 0)$ and $\max(R^{gross}_{t-1}, 0)$), lagged log of assets under management ($\log(AUM)_{t-1}$), and years in industry (Experience). In order to save space, the coefficients on the control variables are not reported.

Table S8: Robustness check for including the experimental measures separately: The table shows coefficient estimates from ordinary least squares regressions of funds' performance and risk characteristics on cognitive skills and economic preferences/attitudes. Each coefficient estimate is obtained from a separate regression of the dependent variable (column heading) on the particular covariate (row heading), adjusted for the same control variables and fixed effects as in Table 5. The first three columns use the same regressions as the single-manager regression from Table 5 with the same dependent variable, respectively. The only difference is that all six experimental measures are jointly included in the regressions of Table 5, while only the experimental measure indicated in each row is included in the regressions of this table. Similarly, the last three columns use the same regressions as the corresponding single-manager regressions from Table 6, with the same modifications. Corresponding p-values are reported in parentheses. ** $p < 0.005$, * $p < 0.05$.

Dependent variable $\longrightarrow$	R^{gross}	R^{abn}	V	SR	RV	TE
CRT	−0.431	−0.170	−0.084	0.086	−0.038*	0.975
	(0.166)	(0.593)	(0.946)	(0.179)	(0.009)	(0.268)
TOM	−0.247	−0.425	−0.871	0.009	−0.038	−0.520
	(0.419)	(0.087)	(0.247)	(0.905)	(0.084)	(0.445)
Competitiveness	0.193	0.542	1.438	0.007	0.001	0.468
	(0.623)	(0.134)	(0.075)	(0.931)	(0.977)	(0.553)
Risk Tolerance	−0.077	−0.082	−0.045	−0.013	0.005*	0.073
	(0.061)	(0.054)	(0.715)	(0.090)	(0.038)	(0.439)
Loss Tolerance	0.064	0.053	0.025	−0.012	0.005	0.140
	(0.215)	(0.317)	(0.867)	(0.446)	(0.267)	(0.195)
Ambiguity Tolerance	0.004	0.026	0.056	0.006	0.000	0.114**
	(0.877)	(0.160)	(0.254)	(0.251)	(0.835)	(0.002)

Independent variables: CRT stands for the cognitive reflection score, comprised of 5 questions, measuring deliberate thinking. TOM stands for the "Reading-the-Mind-in-the-Eyes"-test, measuring theory of mind skills, i.e., the ability to infer the intention of others. The score for risk preferences (Risk Tolerance) reflects the elicited certainty equivalent for a lottery paying €60 or €0 with equal probability, with higher values indicating higher levels of risk tolerance. The measure for attitudes towards losses (Loss Tolerance) reflects the maximum potential loss subjects were willing to accept in order to have the chance of winning €22. Again, the higher the number, the more tolerant towards losses a fund manager is (Loss Tolerance). The score for ambiguity preferences (Ambiguity Tolerance) represents the matching probability (in %) that leaves subjects indifferent between a risky lottery with a certain probability of winning and an ambiguous lottery with an unknown probability of winning (both lotteries paid €60 in the case of winning and €0 else). Competitiveness is measured as the sum of the five standardized responses to the subscale of the WOFO, answered on scales ranging from 1 to 7 each.

Table S9: Robustness check for excluding the 2008-2009 crisis period and including time preferences: The table shows the results of ordinary least squares regressions of funds' Sharpe ratio, relative volatility, and tracking error on cognitive skills and economic preferences/attitudes. SR denotes the Sharpe Ratio, RV is the overall volatility of the fund relative to the volatility of the benchmark, and TE is the tracking error measured as the standard deviation of the difference between the fund's net return and the benchmark return. In the first three columns, a shorter sample period is used compared to the main analysis, namely Jan 2010 – Dec 2019. In the last three columns, the original sample period is used (Jan 2008 – Dec 2019), and there is an additional regressor measuring the time preferences of the managers (Patience). Standard errors are clustered at the manager level. Corresponding p-values are shown in parentheses. ** $p < 0.005$, * $p < 0.05$.

	Shorter sample period: Jan 2010 - Dec 2019			*Variable Patience added*		
Dependent variable	SR	RV	TE	SR	RV	TE
Sample of funds	*Single*	*Single*	*Single*	*Single*	*Single*	*Single*
CRT	0.079	−0.047**	0.550	0.095	−0.040*	0.415
	(0.134)	(0.001)	(0.454)	(0.073)	(0.005)	(0.563)
TOM	0.039	−0.036	−0.245	0.049	−0.032	−0.453
	(0.548)	(0.068)	(0.659)	(0.506)	(0.129)	(0.368)
Competitiveness	−0.034	0.021	0.317	−0.027	0.019	−0.547
	(0.616)	(0.384)	(0.590)	(0.721)	(0.510)	(0.366)
Risk Tolerance	−0.018*	0.007*	−0.040	−0.017*	0.006*	0.003
	(0.019)	(0.024)	(0.536)	(0.032)	(0.023)	(0.965)
Loss Tolerance	−0.004	0.003	0.151	−0.005	0.004	0.159
	(0.766)	(0.495)	(0.065)	(0.743)	(0.378)	(0.082)
Ambiguity Tolerance	0.008	−0.001	0.077*	0.005	−0.001	0.080*
	(0.140)	(0.692)	(0.045)	(0.309)	(0.752)	(0.021)
Patience				−0.015	−0.006	−0.051
				(0.181)	(0.077)	(0.578)
$\min(R^{gross}_{t-1}, 0)$	0.041	−0.019**	−0.827**	0.013	−0.013**	−0.906**
	(0.171)	(< 0.001)	(< 0.001)	(0.651)	(0.001)	(< 0.001)
$\max(R^{gross}_{t-1}, 0)$	−0.047	0.015**	0.225**	−0.012	0.013**	0.236*
	(0.082)	(< 0.001)	(0.001)	(0.501)	(< 0.001)	(0.008)
$\log(AUM)_{t-1}$	0.065	0.009	−0.979**	0.070	0.023*	−1.359**
	(0.123)	(0.459)	(0.002)	(0.116)	(0.045)	(< 0.001)
Experience	0.003	0.001	0.158	0.005	0.002	−0.029
	(0.706)	(0.796)	(0.100)	(0.676)	(0.469)	(0.756)
Constant	yes	yes	yes	yes	yes	yes
Category FE	yes	yes	yes	yes	yes	yes
Time FE	yes	yes	yes	yes	yes	yes
Number of observations	11,006	7,892	7,892	10,959	7,699	7,699
Number of managers	65	58	58	58	51	51
Adjusted R^2	0.416	0.174	0.546	0.430	0.167	0.634

Independent variables: CRT stands for the cognitive reflection score, comprised of 5 questions, measuring deliberate thinking. TOM stands for the "Reading-the-Mind-in-the-Eyes"-test, measuring theory of mind skills, i.e., the ability to infer the intention of others. The score for risk preferences (Risk Tolerance) reflects the elicited certainty equivalent for a lottery paying €60 or €0 with equal probability, with higher values indicating higher levels of risk tolerance. The measure for attitudes towards losses (Loss Tolerance) reflects the maximum potential loss subjects were willing to accept in order to have the chance of winning €22. Again, the higher the number, the more tolerant towards losses a fund manager is (Loss Tolerance). The score for ambiguity preferences (Ambiguity Tolerance) represents the matching probability (in %) that leaves subjects indifferent between a risky lottery with a certain probability of winning and an ambiguous lottery with an unknown probability of winning (both lotteries paid €60 in the case of winning and €0 else). Competitiveness is measured as the sum of the five standardized responses to the subscale of the WOFO, answered on scales ranging from 1 to 7 each. $\min(R^{gross}_{t-1}, 0)$ is zero if the gross return in the previous month was positive, and equal to previous month's gross return if it was negative; similarly, $\max(R^{gross}_{t-1}, 0)$ is zero if the gross return in the previous month was negative, and equal to previous month's gross return if it was positive. $\log(AUM)$ stands for the log of assets under management and Experience indicates years in industry.

Table S10: Robustness check for using APM instead of CRT: The table shows the results of ordinary least squares regressions of funds' performance and risk characteristics on cognitive skills and economic preferences/attitudes. R^{gross} is the gross returns of the fund without benchmark correction, R^{abn} is the fund's gross returns over its benchmark, and V is value added - the product of assets under management and abnormal returns. SR denotes the Sharpe Ratio, RV is the overall volatility of the fund relative to the volatility of the benchmark, and TE is the tracking error measured as the standard deviation of the difference between the fund's net return and the benchmark return. Standard errors are clustered at the manager level. Corresponding p-values are reported in parentheses. ** $p < 0.005$, * $p < 0.05$.

Dependent variable	R^{gross}	R^{abn}	V	SR	RV	TE
APM	0.048	0.076	0.312	0.026	−0.029	1.264*
	(0.871)	(0.802)	(0.770)	(0.709)	(0.172)	(0.045)
TOM	−0.093	−0.292	−0.887	0.037	−0.040*	−0.431
	(0.745)	(0.241)	(0.252)	(0.589)	(0.050)	(0.430)
Competitiveness	−0.042	0.189	1.298	−0.040	0.013	0.153
	(0.906)	(0.552)	(0.193)	(0.528)	(0.608)	(0.815)
Risk Tolerance	−0.095	−0.109	−0.010	−0.016*	0.006*	−0.006
	(0.053)	(0.068)	(0.946)	(0.015)	(0.028)	(0.932)
Loss Tolerance	0.097	0.061	−0.030	−0.007	0.003	0.092
	(0.105)	(0.195)	(0.848)	(0.591)	(0.490)	(0.279)
Ambiguity Tolerance	0.008	0.033	0.038	0.008	−0.001	0.083*
	(0.723)	(0.081)	(0.548)	(0.115)	(0.536)	(0.016)
$\min(R^{gross}_{t-1}, 0)$				0.009	−0.015**	−0.840**
				(0.733)	(< 0.001)	(< 0.001)
$\max(R^{gross}_{t-1}, 0)$				−0.011	0.014**	0.228*
				(0.492)	(< 0.001)	(0.016)
$\log(AUM)_{t-1}$	0.058	0.111		0.066	0.011	−1.100**
	(0.798)	(0.521)		(0.102)	(0.308)	(< 0.001)
Experience	−0.025	0.024	0.078	−0.001	0.002	0.155
	(0.570)	(0.522)	(0.520)	(0.949)	(0.549)	(0.097)
Constant	yes	yes	yes	yes	yes	yes
Category FE	yes	yes	yes	yes	yes	yes
Time FE	yes	yes	yes	yes	yes	yes
Number of observations	12,322	9,393	9,393	12,180	8,795	8,795
Number of managers	65	59	59	65	58	58
Adjusted R^2	0.439	0.086	0.031	0.425	0.138	0.590

Independent variables: APM stands for the score on the test based on Raven's Advanced Progressive Matrices, measuring fluid intelligence. TOM stands for the "Reading-the-Mind-in-the-Eyes"-test, measuring theory of mind skills, i.e., the ability to infer the intention of others. The score for risk preferences (Risk Tolerance) reflects the elicited certainty equivalent for a lottery paying €60 or €0 with equal probability, with higher values indicating higher levels of risk tolerance. The measure for attitudes towards losses (Loss Tolerance) reflects the maximum potential loss subjects were willing to accept in order to have the chance of winning €22. Again, the higher the number, the more tolerant towards losses a fund manager is (Loss Tolerance). The score for ambiguity preferences (Ambiguity Tolerance) represents the matching probability (in %) that leaves subjects indifferent between a risky lottery with a certain probability of winning and an ambiguous lottery with an unknown probability of winning (both lotteries paid €60 in the case of winning and €0 else). Competitiveness is measured as the sum of the five standardized responses to the subscale of the WOFO, answered on scales ranging from 1 to 7 each. $\min(R^{gross}_{t-1}, 0)$ is zero if the gross return in the previous month was positive, and equal to previous month's gross return if it was negative; similarly, $\max(R^{gross}_{t-1}, 0)$ is zero if the gross return in the previous month was negative, and equal to previous month's gross return if it was positive. $\log(AUM)$ stands for the log of assets under management and Experience indicates years in industry.

Table S11: Robustness check for Sharpe Ratio, relative volatility, relative semi-volatility, and tracking error when calculating these measures based on daily data over half a year: The table shows the results of ordinary least squares regressions of funds' Sharpe Ratio, relative volatility, relative semi-volatility, and tracking error on cognitive skills and economic preferences/attitudes. The Sharpe Ratio, SR, measures the abnormal return per unit of fund risk. The relative volatility, RV, stands for the overall riskiness of the fund relative to the riskiness of the benchmark. Relative semi-volatility, RSV, measures the funds' down-side risk relative to the downside risk of the benchmark and the tracking error, TE, measures the standard deviation of the difference between the fund's net return and the benchmark return. Standard errors are clustered at the manager level. Corresponding p-values are shown in parentheses. ** $p < 0.005$, * $p < 0.05$.

Dependent variable *Sample of funds*	SR *Single*	RV *Single*	RSV *Single*	TE *Single*
CRT	0.068	-0.048**	-0.043*	0.447
	(0.108)	(0.003)	(0.008)	(0.582)
TOM	0.032	-0.042	-0.042	-0.264
	(0.518)	(0.054)	(0.063)	(0.647)
COMPETITIVENESS	-0.067	0.021	0.031	0.029
	(0.165)	(0.454)	(0.293)	(0.963)
RISK TOLERANCE	-0.019**	0.007*	0.007*	-0.014
	(< 0.001)	(0.014)	(0.010)	(0.846)
LOSS TOLERANCE	0.003	0.003	0.003	0.126
	(0.716)	(0.553)	(0.516)	(0.163)
AMBIGUITY TOLERANCE	0.005	0.000	-0.001	0.098*
	(0.168)	(0.823)	(0.704)	(0.016)
$\min(R_{t-1}^{gross}, 0)$	0.018*	-0.004	-0.004	-0.238**
	(0.041)	(0.110)	(0.082)	(0.002)
$\max(R_{t-1}^{gross}, 0)$	-0.007	0.003*	0.003	0.056
	(0.413)	(0.022)	(0.065)	(0.185)
$\log(AUM)_{t-1}$	0.047	0.006	0.005	-1.096**
	(0.120)	(0.669)	(0.704)	(0.001)
EXPERIENCE	-0.001	0.000	0.002	0.141
	(0.900)	(0.917)	(0.589)	(0.169)
Constant	yes	yes	yes	yes
Category FE	yes	yes	yes	yes
Time FE	yes	yes	yes	yes
Number of observations	1,826	1,316	1,316	1,316
Number of managers	63	57	57	57
Adjusted R^2	0.451	0.184	0.147	0.621

Independent variables: CRT stands for the cognitive reflection score, comprised of 5 questions, measuring deliberate thinking. TOM stands for the "Reading-the-Mind-in-the-Eyes"-test, measuring theory of mind skills, i.e., the ability to infer the intention of others. The score for risk preferences (RISK TOLERANCE) reflects the elicited certainty equivalent for a lottery paying €60 or €0 with equal probability, with higher values indicating higher levels of risk tolerance. The measure for attitudes towards losses (LOSS TOLERANCE) reflects the maximum potential loss subjects were willing to accept in order to have the chance of winning €22. Again, the higher the number, the more tolerant towards losses a fund manager is (LOSS TOLERANCE). The score for ambiguity preferences (AMBIGUITY TOLERANCE) represents the matching probability (in %) that leaves subjects indifferent between a risky lottery with a certain probability of winning and an ambiguous lottery with an unknown probability of winning (both lotteries paid €60 in the case of winning and €0 else). COMPETITIVENESS is measured as the sum of the five standardized responses to the subscale of the WOFO, answered on scales ranging from 1 to 7 each. $\min(R_{t-1}^{gross}, 0)$ is zero if the gross return in the previous month was positive, and equal to previous month's gross return if it was negative; similarly, $\max(R_{t-1}^{gross}, 0)$ is zero if the gross return in the previous month was negative, and equal to previous month's gross return if it was positive. $\log(AUM)$ stands for the log of assets under management, and EXPERIENCE indicates years in industry.

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/

2019-16 **Adam Farago, Martin Holmen, Felix Holzmeister, Michael Kirchler, Michael Razen:** Cognitive Skills and Economic Preferences in the Fund Industry

2019-15 **Christopher Kah, Daniel Neururer:** Generiert der stationäre Buchhandel positive Nachfrageeffekte und verhilft dadurch dem Kulturgut Buch bei seiner Verbreitung? - Ein natürliches Experiment

2019-14 **Stjepan Srhoj, Michael Lapinski, Janette Walde:** Size matters? Impact evaluation of business development grants on SME performance

2019-13 **Andrea M. Leiter, Engelbert Theurl:** DETERMINANTS OF PREPAID SYSTEMS OF HEALTHCARE FINANCING - A WORLDWIDE COUNTRY-LEVEL PERSPECTIVE

2019-12 **Michael Razen, Michael Kirchler, Utz Weitzel:** Domain-Specific Risk-Taking Among Finance Professionals

2019-11 **Jonathan Hall, Rudolf Kerschbamer, Daniel Neururer, Eric Skoog:** Uncovering sophisticated discrimination with the help of credence goods markups - evidence from a natural field experiment

2019-10 **Daniela Glätzle-Rützler, Philipp Lergetporer, Matthias Sutter:** Collective intertemporal decisions and heterogeneity in groups

2019-09 **Morten Hedegaard, Rudolf Kerschbamer, Daniel Müller, Jean-Robert Tyran:** Distributional Preferences Explain Individual Behavior Across Games and Time

2019-08 **Daniel Müller, Sander Renes:** Fairness Views and Political Preferences - Evidence from a representative sample

2019-07 **Florian Lindner, Michael Kirchler, Stephanie Rosenkranz, Utze Weitzel:** Social Status and Risk-Taking in Investment Decisions

2019-06 **Christoph Huber, Julia Rose:** Individual attitudes and market dynamics towards imprecision

2019-05 **Felix Holzmeister, Jürgen Huber, Michael Kirchler, Florian Lindner, Utz Weitzel, Stefan Zeisberger:** What Drives Risk Perception? A Global Survey with Financial Professionals and Lay People

2019-04 **David M. McEvoy, Tobias Haller, Esther Blanco:** The Role of Non-Binding Pledges in Social Dilemmas with Mitigation and Adaptation

2019-03 **Katharina Momsen, Markus Ohndorf:** When do people exploit moral wiggle room? An experimental analysis in a market setup

2019-02 **Rudolf Kerschbamer, Daniel Neururer, Matthias Sutter:** Credence goods markets and the informational value of new media: A natural field experiment

2019-01 **Martin Geiger, Eric Mayer, Johann Scharler:** Inequality and the Business Cycle: Evidence from U.S. survey data

2018-18 **Matthias Sutter, Jürgen Huber, Michael Kirchler, Matthias Stefan, Markus Walzl:** Where to look for the morals in markets?

2018-17 **Rene Schwaiger, Michael Kirchler, Florian Lindner, Utz Weitzel:** Determinants of investor expectations and satisfaction. A study with financial professionals

2018-16 **Andreas Groll, Julien Hambuckers, Thomas Kneib, Nikolaus Umlauf:** LASSO-Type Penalization in the Framework of Generalized Additive Models for Location, Scale and Shape

2018-15 **Christoph Huber, Jürgen Huber:** Scale matters: Risk perception, return expectations, and investment propensity under different scalings

2018-14 **Thorsten Simon, Georg J. Mayr, Nikolaus Umlauf, Achim Zeileis:** Lightning prediction using model output statistics

2018-13 **Martin Geiger, Johann Scharler:** How do consumers interpret the macroeconomic effects of oil price fluctuations? Evidence from U.S. survey data

2018-12 **Martin Geiger, Johann Scharler:** How do people interpret macroeconomic shocks? Evidence from U.S. survey data

2018-11 **Sebastian J. Dietz, Philipp Kneringer, Georg J. Mayr, Achim Zeileis:** Low visibility forecasts for different flight planning horizons using tree-based boosting models

2018-10 **Michael Pfaffermayr:** Trade creation and trade diversion of regional trade agreements revisited: A constrained panel pseudo-maximum likelihood approach

2018-09 **Achim Zeileis, Christoph Leitner, Kurt Hornik:** Probabilistic forecasts for the 2018 FIFA World Cup based on the bookmaker consensus model

2018-08 **Lisa Schlosser, Torsten Hothorn, Reto Stauffer, Achim Zeileis:** Distributional regression forests for probabilistic precipitation forecasting in complex terrain

2018-07 **Michael Kirchler, Florian Lindner, Utz Weitzel:** Delegated decision making and social competition in the finance industry

2018-06 **Manuel Gebetsberger, Reto Stauffer, Georg J. Mayr, Achim Zeileis:** Skewed logistic distribution for statistical temperature post-processing in mountainous areas

2018-05 **Reto Stauffer, Georg J. Mayr, Jakob W. Messner, Achim Zeileis:** Hourly probabilistic snow forecasts over complex terrain: A hybrid ensemble postprocessing approach

2018-04 **Utz Weitzel, Christoph Huber, Florian Lindner, Jürgen Huber, Julia Rose, Michael Kirchler:** Bubbles and financial professionals

2018-03 **Carolin Strobl, Julia Kopf, Raphael Hartmann, Achim Zeileis:** Anchor point selection: An approach for anchoring without anchor items

2018-02 **Michael Greinecker, Christopher Kah:** Pairwise stable matching in large economies

2018-01 **Max Breitenlechner, Johann Scharler:** How does monetary policy influence bank lending? Evidence from the market for banks' wholesale funding

University of Innsbruck

Working Papers in Economics and Statistics

2019-16

Adam Farago, Martin Holmen, Felix Holzmeister, Michael Kirchler, Michael Razen

Cognitive Skills and Economic Preferences in the Fund Industry

Abstract
By running a battery of experiments with fund managers, we investigate the impact of cognitive skills and economic preferences on their professional decisions. First, we find that fund managers' risk tolerance positively correlates with fund risk when accounting for fund benchmark, fund category, and other controls. Second, we show that fund managers' ambiguity tolerance positively correlates with the funds' tracking error from the benchmark. Finally, we report that cognitive skills do not explain fund performance in terms of excess returns. However, we do find that fund managers with high cognitive reflection abilities compose funds at lower risk.

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