On the Role of Benchmarks for Retail Investors

Vincent Eulenberg and Jonas Frey

Discussion Paper #2025/01

DFG Research Unit: “Consumer Preferences, Consumer Mistakes, and Firms’ Response” (FOR 5392)

On the Role of Benchmarks for Retail Investors*

Vincent Eulenberg† Jonas Frey‡

August 21, 2025

Abstract

Benchmarks are commonly included in charts depicting an asset's past returns, and this is mandatory in certain regulated documents in the US and the EU. However, there is little evidence of the impact of benchmarks on retail investors. We hypothesize that the provision of uninformative benchmarks may bias investors through a contrast effect. This effect predicts that investors will have lower (higher) return expectations and invest less (more) in a fund if its returns are shown together with a benchmark that outperformed (underperformed) the fund relative to a baseline with no benchmark. We test this in an experiment and find that benchmarks affect investors' return expectations and propensity to invest. The results show that benchmarks that underperform the fund increase return expectations and investment, consistent with a contrast effect. Conversely, benchmarks that outperform the fund increase return expectations even more strongly, contradicting the contrast effect.

JEL-Classification: G11, G41, D14, D91.
Keywords: Benchmarks, Household Finance, Contrast Effect, Expectations.

*We thank the German Research Foundation for financial support (project 462020252).
†University of Münster. Email: Vincent.Eulenberg@wiwi.uni-muenster.de
‡University of Gothenburg & University of Münster. Email:jonas.frey@economics.gu.se

1 Introduction

When retail investors consider investing in a mutual fund, they often receive not only information about the fund's past performance but also about the performance of a benchmark index. Benchmarks commonly appear in regulated investor information documents, fund prospectuses, and websites with fund information.[1] Despite this, there is little evidence about how such benchmarks affect retail investors' beliefs about a fund's future returns and their willingness to invest.

There are various ways in which benchmarks could influence investors' beliefs and decisions. On the one hand, they might provide valuable information that helps investors interpret a fund's past returns and make better-informed decisions. For instance, whether a fund outperformed its benchmark may be informative about the fund managers' skills. However, even if such outperformance reflects skill, it is unclear how valuable this information is since fund manager skill is not necessarily linked to higher future performance according to theoretical models (Berk and Green, 2004), and there is little empirical evidence that high past returns of mutual funds are predictive of their future returns (Choi and Zhao, 2020).

On the other hand, benchmarks might interfere with rational information processing. Recent evidence suggests that the format and context in which information is presented affect investors. For example, return expectations for an asset are lower if historical performance is displayed in return charts as opposed to price charts (Glaser et al., 2019), and investors are more likely to sell losing stocks if past returns are displayed less prominently (Frydman and Wang, 2020). In a similar vein, the presence of a benchmark could influence the evaluation of a mutual fund, even when it offers no new information.

To study the role of benchmarks on investing decisions, we ran a preregistered online experiment with 500 participants on Prolific. Our subjects came from a wide range of age groups, and 40% had investment experience, making the sample well-suited for examining retail investor behavior. In our experiment, participants saw a bar chart showing three years of annual historical returns of a mutual fund. In some cases, participants viewed the fund's returns in isolation; in others, they were also presented with the returns of a benchmark index in the same time period. Participants then had to forecast the return of the fund (and the benchmark where applicable) in the subsequent year. To ensure that the benchmarks do not provide relevant information for the evaluation of the mutual funds, we inform participants that the funds are U.S. mutual funds while the benchmarks consist of international stocks. Arguably, this makes the benchmarks sufficiently uninformative to make it impossible to infer a fund manager's skill from comparing the fund's performance to that of the benchmarks. Instead, such differences could simply reflect the distinct return characteristics inherent to the asset types in the benchmark compared to the fund. We chose uninformative benchmarks because of the evidence suggesting that funds strategically choose misspecified benchmarks that are easier to beat (Chen et al., 2025; Cremers et al., 2022; Elton et al., 2013; Mullally and Rossi, 2024; Sensoy, 2009).

[1] See Appendix A for examples and a discussion of the regulatory environment.

Following this forecasting task, participants had to decide how much of an endowment of 1000 experimental currency units (ECUs) they wanted to invest in the mutual fund for which they had just provided a return forecast. Each subject completed both tasks nine times, each time with a different mutual fund and benchmark. Importantly, we used a between-subjects design, presenting the same funds with different benchmarks across different subjects. This approach allowed us to causally measure the effect of a benchmark by comparing the decisions of subjects who evaluated the same fund but with a different (or no) benchmark.

Our experiment allows us to test the predictions that (i) investors' return expectations of a fund and (ii) their willingness to invest in the fund are influenced by the fund's performance relative to an uninformative benchmark. We hypothesize that investors become more optimistic when the fund outperforms the benchmark and more pessimistic when the benchmark outperforms the fund. The behavioral mechanism that yields this prediction is the well-known "contrast effect", which describes how the magnitude of a variable shown in the context of a smaller (larger) realization of a similar variable is overestimated (underestimated).[2] The contrast effect could influence retail investors through two channels: First, contrasts can visually distort perceptions, making one object appear larger or smaller depending on the size of nearby objects. Thus, the contrast effect could cause subjects to misjudge the size of the returns, which we present in a bar chart. Second, even if subjects correctly extract the information from the chart, the contrast effect could still affect their interpretation of the information, as it has been documented in many contexts where a visual channel is implausible (Hartzmark and Shue, 2018; Pepitone and DiNubile, 1976; Radbruch and Schiprowski, 2024). In any case, when a benchmark affects investors via behavioral channels, it could be detrimental to investors' abilities to estimate a fund's future performance correctly.

We find that, in line with our prediction, a benchmark that underperforms the fund increases the subjects' expected return of the fund by 0.21 percentage points and increases their investment in the fund by around 22 ECUs. In contrast to our prediction, the benchmark that outperforms the fund also increases the return expectation for the fund by 0.52 percentage points. However, it has no effect on the amount subjects invest in the fund. Overall, the results for the low benchmark are in line with the idea of a contrast effect, whereas the results for the high benchmark are not. This suggests that another mechanism drives the results, either independently of or in conjunction with the contrast effect. For instance, "wishful thinking" (Caplin and Leahy, n.d.) could explain why both benchmarks lead to more positive expectations, and the focus on funds that beat their benchmark in real-world investment contexts may explain why only the low-performing benchmark leads to more investment. However, our experiment is not designed to test for this explanation and distinguish it from other potential explanations.

Taken together, our results highlight that benchmarks affect our experimental subjects' re-

[2]The contrast effect has been supported in various domains. For instance, contrast effects influence scholarship candidate evaluations (Radbruch and Schiprowski, 2024), judgments of crime severity (Pepitone and DiNubile, 1976), dating decisions (Bhargava and Fisman, 2014), perception of beauty (Cash et al., 1983; Kenrick and Gutierres, 1980) and life satisfaction (Schkade and Kahneman, 1998). The contrast effect has also been found to influence the market response to earnings announcements (Hartzmark and Shue, 2018) and stock investment decisions (Antoniou et al., 2021; Kim and Hoffman, 2020).

turn expectations and investment decisions. Therefore, understanding the role of benchmarks in investment decisions is crucial and relevant to policy considerations.

In recognition of the importance of benchmarks, the Securities and Exchange Commission (SEC) mandated that mutual funds include a broad-based index in graphical comparisons in 1993. This decision was intended to standardize fund reporting and enhance investor understanding (Securities and Exchange Commission, 1993). In the example provided by the SEC, the S&P 500 was used as a benchmark (see Figure A.5). This may explain its widespread adoption, with 38.7% of fund assets benchmarked against the S&P 500 in 2020 (Chen et al., 2025). However, evidence suggests that over 40% of funds benchmarked against the S&P 500 were mismatched in terms of investment style (Sensoy, 2009). Even more concerningly, Hartzmark and Solomon (2022) show that many funds use a version of the index that does not account for dividends, which is an "especially uninformative measure of performance."

In addition to showing that funds choose misspecified benchmarks, both Sensoy (2009) and Hartzmark and Solomon (2022) demonstrate that outperforming a misspecified benchmark leads to increased fund flows and suggest that fund managers strategically choose mismatched benchmarks. Several other papers come to similar conclusions. For example, Elton et al. (2013) find that managers of separately managed accounts (SMAs), a type of investment vehicle for wealthy individuals and institutional investors, select benchmarks that significantly overstate the performance of the SMA, and investors respond to these inflated performance measures when allocating capital. Cremers et al. (2022) find that funds with such mismatched benchmarks tend to be riskier than their prospectus benchmarks, leading to average outperformance of the prospectus benchmarks but underperformance relative to benchmarks that better match their investment strategies. In a recent paper, Chen et al. (2025) show that mutual funds often change their benchmarks over time, with specialized funds choosing benchmarks that are easier to beat, typically featuring lower risk exposure. Likewise, Mullally and Rossi (2024) find that high-fee funds, broker-sold funds, and funds facing poor performance and outflows tend to change their benchmarks, often dropping those with higher past returns and adopting benchmarks with lower past returns. After adopting the new benchmark, such funds attract additional inflows despite continuing to underperform their peers.

These findings raise questions about the adequacy of current regulations regarding both the type of benchmarks fund providers are permitted to use and whether they should be required to provide a benchmark for comparison at all.

Our contribution to this literature is threefold. First, we provide causal evidence on the effect of a misspecified benchmark on the investment decisions of retail investors. We show that benchmarks that underperform the fund lead to more investment, which adds to the correlational evidence showing that funds that outperform their benchmarks tend to attract more fund flows. Second, we analyze whether biased return expectations are the channel through which benchmarks affect investments. We find evidence in line with this for benchmarks that underperform the fund but not for benchmarks that outperform the fund. Finally, we test whether the impact of benchmarks on expectations and investment decisions can be explained by contrast effects and find that contrast

effects alone cannot explain how benchmarks affect investors. This result is surprising, given that contrast effects have been shown to affect the market response to earnings announcements (Hartzmark and Shue, 2018) and the stock trading decisions of individual investors (Antoniou et al., 2021; Kim and Hoffman, 2020).

Overall, we find that benchmarks that underperform the fund lead to higher return expectations and increased investment, in line with our hypotheses. In contrast, benchmarks that outperform the fund have an even stronger effect on expectations, but no effect on investment, which contradicts our hypotheses.

Importantly, the inconsistent results for the benchmark that outperformed the fund are of lower practical relevance than the consistent results for the benchmark that underperformed the fund, given that there is a plethora of evidence that funds strategically choose benchmarks that they can beat, but no evidence that they choose benchmarks that outperform the fund. Our results provide causal evidence that choosing an underperforming benchmark can raise return expectations and investment.

2 Design and Implementation

We conducted an incentivized experiment with 500 subjects on Prolific. During the experiment, participants saw three years of annual historical returns for nine different mutual funds. An overview of the funds is provided in Table A.1. In some rounds of the experiment, the returns were presented in isolation, while in others they were shown alongside benchmark returns from the same period. There were three pairs of benchmarks, each pair consisting of a low and a high benchmark for the same time period: (i) NASDAQ (low) and Russell 2000 (high) for 2004–2006, (ii) S&P/TSX Composite (low) and BEL 20 (high) for 2012–2014, and (iii) Nikkei 225 (low) and NASDAQ (high) for 2014–2016. An overview of the benchmarks is given in Table A.2. For each benchmark pair, we selected three distinct mutual funds, each having a three-year average return exactly midway between the corresponding low and high benchmarks. This design allowed us to make causal inferences by employing a between-subjects approach, where each of the mutual funds was shown in isolation to one group of subjects, paired with a high benchmark for another group, and paired with a low benchmark for a third group.

We focused exclusively on positive returns for all funds and benchmarks to avoid confounding effects on the contrast effect, as negative returns could trigger loss aversion——a well-documented phenomenon where losses are weighted more heavily than gains (Kahneman and Tversky, 1979). To further mitigate any potential confounding effects, we designed the experimental setting to be as simple as possible, withholding additional information such as fund style, fund name, and time periods that would typically be available in real-world settings. This approach allowed us to isolate the effect of benchmark provision, avoiding potential interference from additional information. Finally, we intentionally framed the funds as "U.S. mutual funds" and the benchmarks as "reflecting large international stocks", which creates a mismatch between the two due to their differing risk

and return profiles. In line with the previous empirical literature, which showed that asset managers strategically select less informative benchmarks to shape investors' perceptions (Chen et al., 2025; Elton et al., 2013; Sensoy, 2009), we argue that mismatched benchmarks are uninformative about a fund manager's skill. Therefore, subjects should not change their return expectations and investment decisions depending on how the fund performed relative to the benchmark. However, we avoided explicitly stating that the benchmarks were irrelevant, as doing so might have led participants to perceive them as mere decoys and feel compelled to ignore them, potentially preventing any observable effect. This design is relevant from a practical perspective, as fund managers have been shown to strategically select less informative benchmarks to shape investors' perceptions (Sensoy, 2009).

After reviewing the returns for at least 10 seconds, participants had to forecast the return of the fund (and the benchmark, where applicable) in the subsequent year. A screenshot of the forecasting decision screen is given in Figure 1. Following this forecasting task, participants made an investment decision in which they had to decide how much of an endowment of 1000 experimental currency units they wanted to invest in the mutual fund they had just forecasted (see Figure A.10). Across the experiment, each participant completed a total of nine forecasting decisions and nine corresponding investment decisions (one for each fund). The decisions were divided into three categories: (i) decisions for mutual funds presented without benchmarks, (ii) decisions for mutual funds presented alongside benchmarks with higher average historical returns, and (iii) decisions for mutual funds presented with benchmarks exhibiting lower average returns. Subjects made three decisions for each category, and the sequence of the decisions was randomized at the participant level to mitigate ordering effects. We used different colors for the mutual fund and the benchmark in each round. The instructions are shown in the Appendix in Figures A.6, A.7, A.8 and A.9.

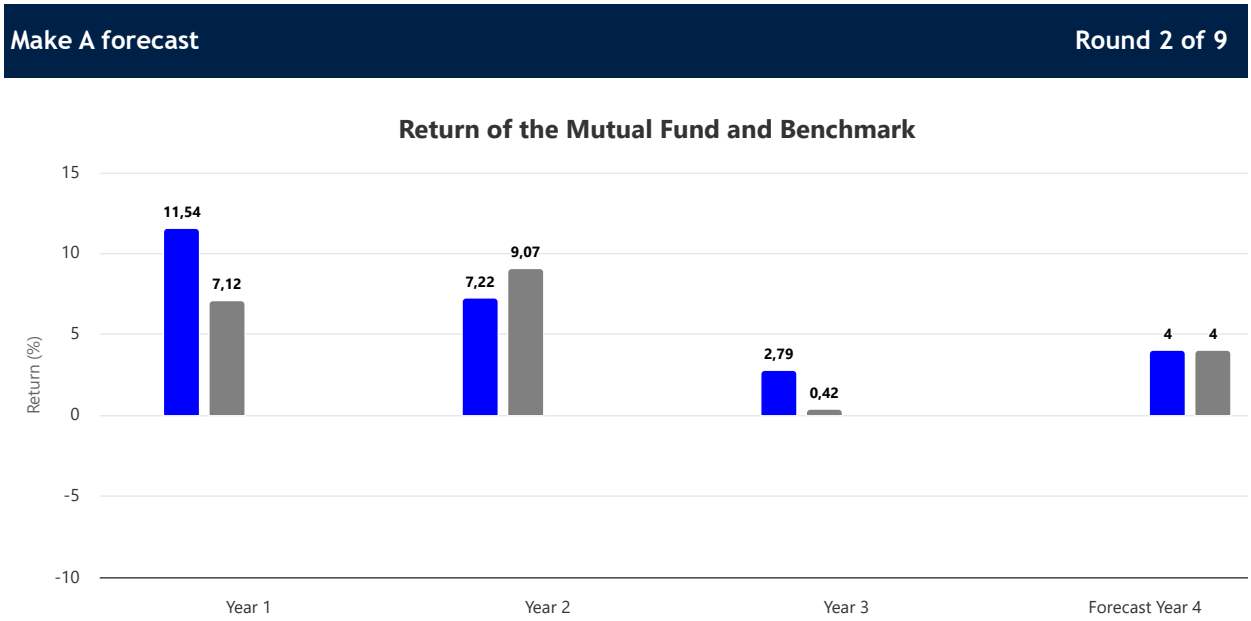
Remuneration: Each participant received a fixed remuneration of £2.50 for participating, supplemented by a variable bonus payment to incentivize responses. For the variable payment, either the forecasting or the investment decision from one of the nine rounds was randomly chosen to be implemented with real financial consequences. For the investment decision, the payment structure was designed to reflect the performance of the mutual fund. The payout of each participant was determined by the actual returns of the mutual fund in the following five years. The bonus payoff from the forecasting decision was higher the more accurate an individual's forecast was, i.e., the closer it was to the actual return of the mutual fund in the next year. Participants in the experiment were remunerated for the forecasting decision according to the following formula:[3]

$$\text{Bonus Payoff} = 2500 - 100 \cdot |\text{True Value (\%)} - \text{Forecast (\%)}| \tag{1}$$

The bonus payoff is expressed in experimental currency units, where 1000 ECUs are equivalent to £1. Average bonus payments for the forecasting and the investment decision were almost equivalent,

[3]For clarity, participants were not shown the formula. Instead, they saw an illustrative example showing how their forecasting accuracy affects their bonus payoff.

Figure 1: Forecasting Decision



Notes: The figure shows a screenshot of the forecasting page. Participants were presented with three years of annual returns data for a mutual fund and a corresponding benchmark (here: benchmark *low*). They were required to spend at least ten seconds reviewing the returns before predicting the fourth-year returns for both the fund and the benchmark.

at £1.15 and £1.16, respectively. Participants also received payments based on their scores in a cognitive reflection test comprising 10 questions, leading to total average bonus payments of £1.51 per participant. The median completion time for the experiment was 21.5 minutes.

Comprehension Check: After reading the instructions, participants had to undergo a simple attention check and answer two comprehension check questions. Participants who did not pass the comprehension check were given the opportunity to reread the instructions and attempt to answer the questions again. If they either failed the attention check or did not pass the comprehension check after two attempts, participants were excluded from the experiment. Out of 946 initial participants, 446 were unable to pass the comprehension check, resulting in a final dataset consisting of 500 participants.

Preregistration and Institutional Review: Our experimental design and main hypotheses have been preregistered before data collection with AsPredicted #159903.[4] The full experiment is available online at `https://forecasting-benchmarks.herokuapp.com`.

Our experiment has been reviewed and approved by the Institutional Review Board of the German Association for Experimental Economic Research (No. 2oXiyuBu).

3 Hypotheses

Existing evidence suggests that humans are susceptible to the contrast effect, whereby the perception of the absolute size of an entity—be it an object, the outcome of a variable, or a piece of information—is altered by the simultaneous perception of a comparable entity. The effect can manifest both through a visual channel and a more abstract information processing channel. An example of the visual channel can be seen in Figure 2, which illustrates that an object will appear larger if it is contrasted by smaller objects than if it is contrasted by larger objects. The figure's inner circles are the same size, but the left circle appears smaller than the right one. Likewise, in Figure 3, the mutual fund returns are the same, but the left bars might appear smaller than the right bars due to a contrast effect. An example of a contrast effect in a context unrelated to visual perception is that firms' post-earnings announcement returns tend to be more negative when other firms reported larger positive earnings surprises on the previous day (Hartzmark and Shue, 2018). A likely explanation for this is that the firm's own earnings look worse in comparison the better other firms did. In our experiment, benchmarks could influence subjects through both channels, and we do not try to disentangle the effect because the real-world decision situation that motivates our experiment also involves visual representations of benchmarks and, thus, both channels.

We hypothesize that both visual and informational contrast effects will influence subjects' beliefs about the funds' future returns. Specifically:

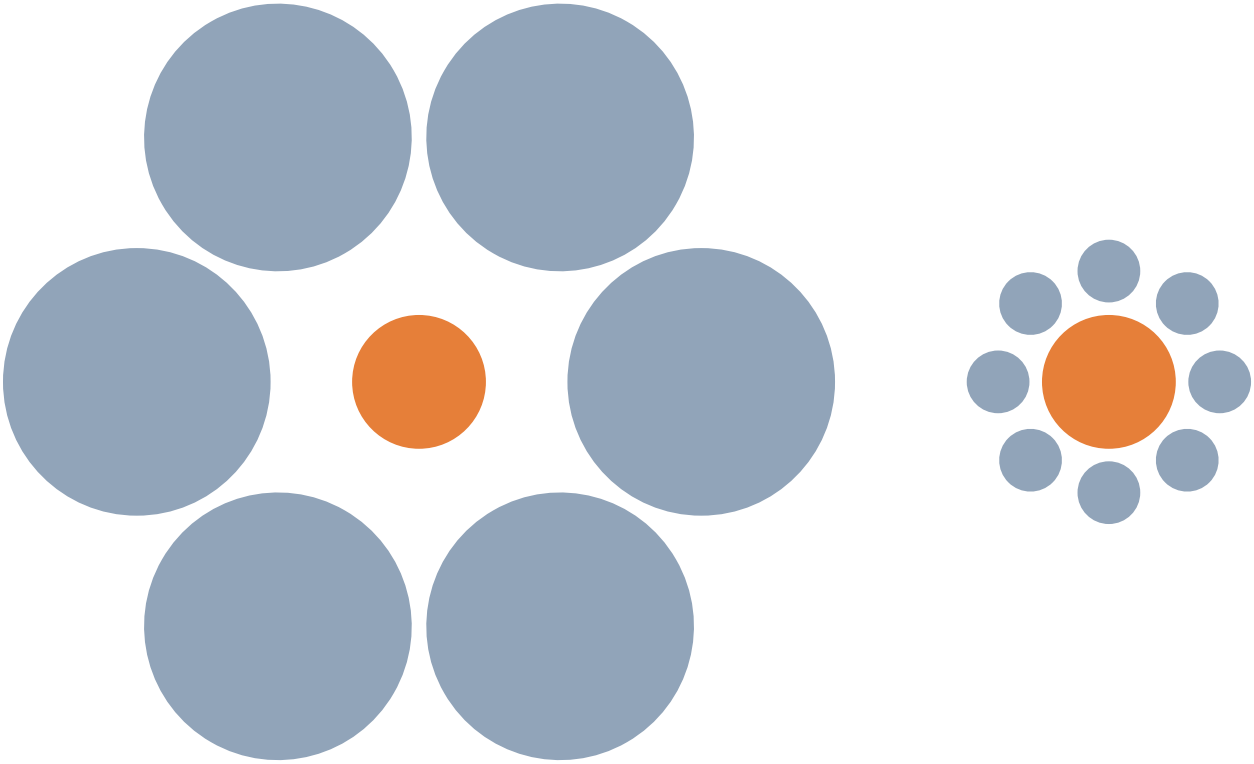
Hypothesis 1 *Subjects who see a mutual fund with a benchmark that has lower returns forecast higher future returns for the mutual fund than subjects who see the same mutual fund without a benchmark.*

Hypothesis 2 *Subjects who see a mutual fund with a benchmark that has higher returns forecast lower future returns for the mutual fund than subjects who see the same mutual fund without a benchmark.*

If the contrast effect influences return expectations, this should also influence participants' willingness to invest in the mutual fund because more positive return expectations should make investing more attractive. Moreover, even in the absence of a contrast effect on expectations, it

[4]The link to the preregistration is `https://aspredicted.org/8LL_DP7`. As mentioned in our preregistration, this is the second experimental run. We have conducted the same experiment with 500 participants before, but subjects were unable to recognize the difference between the low and the high benchmark, which is why we re-ran the experiment with more extreme benchmarks.

Figure 2: General Principle of the Visual Contrast Effect



Notes: This figure demonstrates the visual contrast effect: the two inner orange circles are identical in size, yet the left circle appears smaller due to the differing sizes of the surrounding elements. Source: Wikimedia Commons (public domain).

might directly affect subjects' investment decisions by influencing their subconscious assessment of the mutual fund. Therefore, we hypothesize:

Hypothesis 3 *Subjects who see a mutual fund with a benchmark that has lower returns invest more in the mutual fund than subjects who see the same mutual fund without a benchmark.*

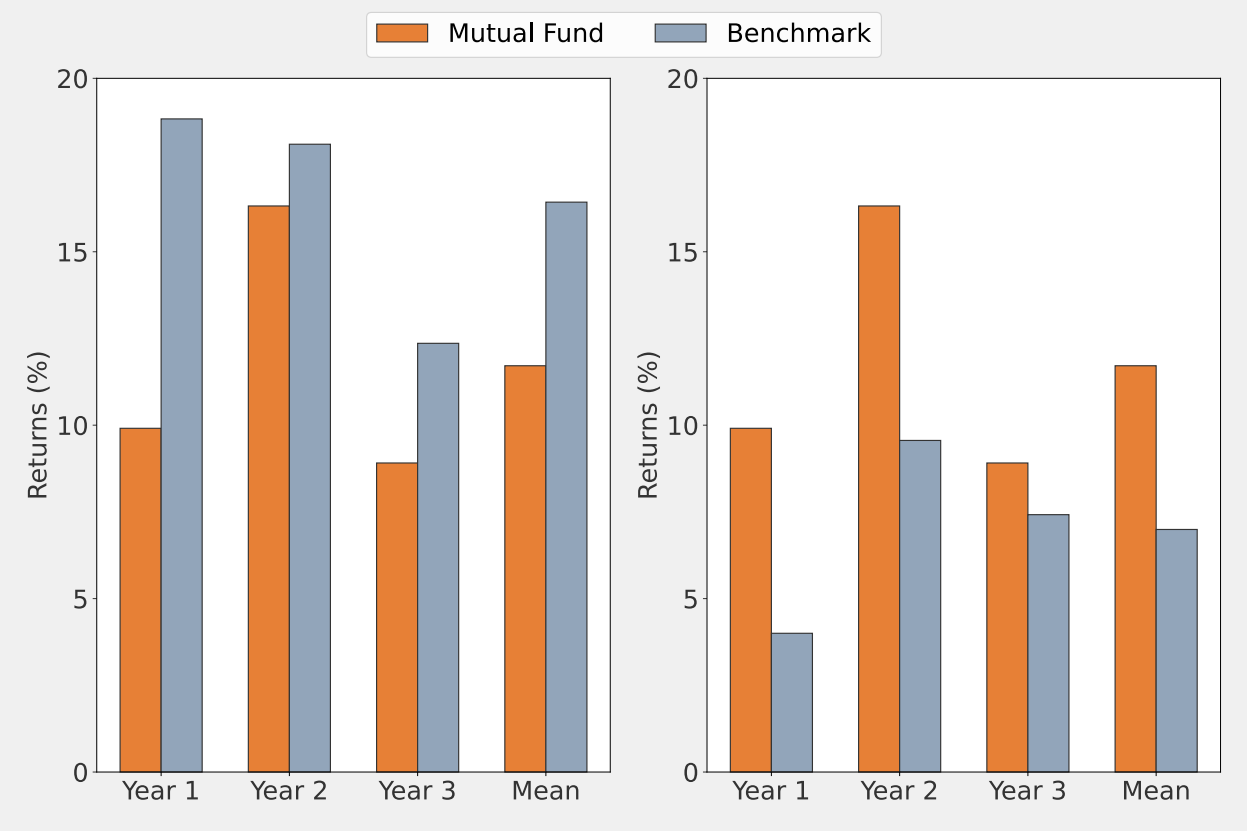
Hypothesis 4 *Subjects who see a mutual fund with a benchmark that has higher returns invest less in the mutual fund than subjects who see the same mutual fund without a benchmark.*

4 Results

Descriptive results are presented in Table 1. The forecasts for both the mutual fund and the benchmark are winsorized at the 1st and 99th percentile to deal with outliers.[5] In the condition *Benchmark High*, the forecast for the benchmark (10.97%) is larger than the forecast for the fund (9.38%). Analogously, in the condition *Benchmark Low*, the forecast for the benchmark (6.50%) is

[5]More than 99% of all forecasts range between -3% and 25%, but individual forecast values go as high as 900%.

Figure 3: Visual Contrast Effect in Mutual Fund Evaluation



Notes: This figure demonstrates how the visual contrast effect could influence investor perception of mutual fund returns. The orange bars reflecting mutual fund returns on the left side and on the right side are identical in size. We hypothesize that participants will perceive them differently due to the different sizes of the adjacent benchmarks. The fund depicted here is "Fund 4", and the benchmarks are "Benchmark 3" (low) and "Benchmark 4" (high). For an overview of the funds and benchmarks, see Tables A.1 and A.2.

smaller than that for the fund (9.08%). Both differences are statistically significant at the 0.01% level, indicating that subjects correctly recognized that the high benchmark outperformed the fund and the low benchmark underperformed the fund. An overview of all fund and benchmark returns, as well as the respective average forecasts and investments in the funds for funds 1-3, 4-6, and 7-9, can be found in the Appendix in figures A.11, A.12, and A.13, respectively. Table A.1 presents an overview of the funds, and Table A.2 presents an overview of the benchmarks used in the experiment.

To test Hypothesis 1 and Hypothesis 2, which posit that participants who view the fund next to a benchmark with lower (higher) returns will overestimate (underestimate) the fund's return, we estimate the following regression equation:

$$ForecastFund_{if} = \beta_0 + \beta_1 BenchmarkHigh_{if} + \beta_2 BenchmarkLow_{if} + \lambda \mathbf{X}_i' + \delta_f + \varepsilon_{if}, \qquad (2)$$

Table 1: Summary of descriptive statistics

Demographics	**Mean**	**SD**
Age (in years)	29.2	8.8
Female (%)	51.3	50.0
Stock Market Participation (%)	40.0	49.0
$<$ £10,000 Personal Income (%)	58.6	49.3
Full-Time Job (%)	48.5	50.0
Student (%)	45.0	49.8
Comprehension + Completion Time	**Mean**	**SD**
First Try Comprehension Check (%)	41.2	49.3
Number Correct CRT Questions (Max 10)	7.2	2.27
Completion Time (in minutes)	24.3	11.71
Forecasting + Investment Decisions	**Mean**	**SD**
Forecast Fund (%)	9.10	4.56
Forecast Benchmark (%)	8.73	4.08
Investment in Fund	339.59	309.74
Fund Only		
Forecast Fund (%)	8.84	4.67
Investment in Fund	329.72	307.00
Benchmark High		
Forecast Fund (%)	9.38	4.36
Forecast Benchmark (%)	10.97	4.48
Investment in Fund	336.31	305.74
Benchmark Low		
Forecast Fund (%)	9.08	4.64
Forecast Benchmark (%)	6.50	3.68
Investment in Fund	352.75	314.49
Individuals	500	
Observations	4500	

Notes: The table shows means and standard deviations (SD) for demographics, task performance, and forecasting and investment decisions across conditions ("Fund Only," "Benchmark High," and "Benchmark Low").

where i and f index individuals and funds, respectively. The variable $ForecastFund$ represents subjects' forecasts of the funds' returns, and the binary dummy variables $BenchmarkHigh$ and $BenchmarkLow$ indicate whether or not an individual saw the fund alongside a benchmark with higher or lower returns, respectively. $\mathbf{X}$ is a vector of additional control variables, including age, gender, investment experience, personal income, and performance in the cognitive reflection test. Finally, δ reflects fund-fixed effects.

Hypothesis 3 and Hypothesis 4 state that observing fund returns and returns of the low (high) benchmark simultaneously leads to an increase (decrease) in the investment amount made by individuals. We test this hypothesis in a similar manner:

$$InvestmentFund_{if} = \beta_0 + \beta_1 BenchmarkHigh_{if} + \beta_2 BenchmarkLow_{if} + \lambda \mathbf{X}_i' + \delta_f + \varepsilon_{if} \quad (3)$$

The Effect of Benchmarks on Fund Return Expectations: The results of the OLS regression specifications are presented in Table 2.[6] Odd-numbered columns present regression specifications without control variables, while even-numbered columns include control variables.

Table 2: Baseline Regression

	(1) Forecast	(2) Forecast	(3) Investment	(4) Investment	(5) Investment	(6) Investment
High	0.522***	0.521***			5.646	5.662
	(5.09)	(5.06)			(0.85)	(0.85)
Low	0.209**	0.210**			22.14***	22.16***
	(2.11)	(2.11)			(3.25)	(3.24)
Forecast			19.57***	19.51***		
			(12.51)	(12.95)		
Mean of DV	9.1	9.1	339.6	339.6	339.6	339.6
Fund FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes	No	Yes
Observations	4500	4491	4500	4491	4500	4491
Adjusted R^2	0.473	0.478	0.121	0.153	0.078	0.110

Notes: Results from OLS regressions on benchmark dummies and fund forecasts. "High" and "Low" are dummy variables indicating whether subjects saw fund returns alongside a benchmark with higher or lower returns, respectively; "Fund Only" is the baseline category. "Forecast" represents the forecast of the fund, and "Investment" reflects the amount invested in the fund. Columns (2), (4) and (6) control for CRT performance, personal income, age and gender. Standard errors are clustered at the subject level. T-Statistics are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Consistent with Hypothesis 1, columns (1) and (2) show that subjects exposed to a mutual fund

[6]Note that the dependent variable "Investment" in columns (3) - (6) is truncated between 0 and 1000, suggesting a Tobit regression model could be more appropriate due to the bounded nature of the dependent variable (Amemiya, 1973; Tobin, 1958). We adhere to OLS here because of its conventional use in the literature but report a Tobit regression in Appendix Table A.3. The results are similar to those of the OLS regressions.

alongside a benchmark with a lower average return tend to forecast a higher return for the mutual fund in the subsequent year, compared to participants who observed the mutual fund without a benchmark ($\beta = 0.210, p < 0.05$). This is in line with the initially hypothesized contrast effect. However, in stark contrast to Hypothesis 2, the same holds true for subjects who view the mutual fund next to a benchmark with a higher average return, and the effect is even stronger in this case ($\beta = 0.521, p < 0.01$). Overall, these results provide mixed evidence, which is in line with our first but not with our second hypothesis. Therefore, while we find some evidence in line with contrast effects, alternative mechanisms are needed to fully explain our results. The fact that both the provision of a high benchmark and the provision of a low benchmark significantly influence expectations about the fund's prospects—despite the benchmarks being uninformative—indicates that investors react systematically and irrationally to such cues. While the effect of the low benchmark aligns with our hypothesis, the effect of the high benchmark clearly contradicts it, yet still demonstrates the behavioral relevance of benchmarks and underscores the need for further research to uncover the underlying mechanisms.

The Effect of Fund Return Expectations on Fund Investment Volume: Since funds with higher returns are, all else equal, a more attractive investment, subjects' return expectations for a fund should be positively linked to the amount they want to invest in the fund. We test this by regressing the amount subjects invested on their return expectations and report the results in columns (3) and (4) of Table 2.

The coefficient in our baseline regression without controls (column 3) indicates that a one percentage point increase in fund forecast is associated with a 22.50 points higher investment ($p < 0.01$), which corresponds to 6.6% of the sample mean. This effect is only slightly attenuated when including fund-fixed effects and controls in column (4) and remains statistically significant ($\beta = 19.40, p < 0.01$). These results suggest that return expectations for the fund do influence investment volume, raising the question of whether benchmarks affect fund investment volume via this mechanism.

The Effect of Benchmarks on Fund Investment Volume: Next, we directly test the influence of the benchmarks on the amount that subjects invest in the mutual fund. Hypothesis 3 suggests that seeing the fund alongside a benchmark with lower average returns increases the investment amount in the mutual fund compared to observing fund returns in the absence of a benchmark. As can be seen in columns (5) and (6) of Table 2, participants who observe the fund next to a benchmark with lower returns actually invest 22.16 points more of their initial 1000-point endowment ($p < 0.01$).

Moving to the effect of the high benchmark, we do not find a significant impact on the investment decision ($\beta = 5.662, p > 0.1$). That is, there is neither a negative effect as Hypothesis 4 would have predicted nor a positive effect that we might have expected to observe given the strong positive effect of the high benchmark on the forecast. The fact that the high benchmark has a stronger effect than the low benchmark on the forecast but no effect on the investment decision is not in line

with our previously defined hypotheses. One potential explanation is that participants erroneously believe that the fraction of endowment they do not invest in the mutual fund is automatically invested in the benchmark. In this case, the enhanced return expectations for the fund with the high benchmark would not necessarily translate into investment decisions as long as they do not exceed those of the benchmark. As shown in Figure A.10, we did not explicitly inform subjects again that any part of the endowment that was not invested would stay with them. If individuals were to assume incorrectly that the entire endowment must be invested, we would expect a distribution skewed towards the maximum (minimum) investment amount for *Benchmark Low* (*Benchmark High*). However, as illustrated in Figure A.16, histograms of investment amounts look relatively homogeneous across treatments. Moreover, since we conducted comprehension checks, we are confident that participants correctly and fully understood the instructions and thereby also know that they keep any uninvested funds.

In summary, we find that in line with the idea of a contrast effect, a benchmark that underperforms the mutual fund causes participants in our experiment to form more positive beliefs about its future returns and to allocate more money to the fund. However, our results for a higher benchmark are not in line with a contrast effect or any other theory we are aware of. One possible alternative explanation is that investors engage in "wishful thinking" as in the model of Caplin and Leahy (n.d.). They could successfully find a positive interpretation of both the benchmark that underperforms the fund ("the fund has beaten its benchmark") and of the benchmark that outperforms the fund ("the fund's returns will converge back to its benchmark"), explaining why both make them more optimistic. An explanation for why only the underperforming benchmark leads to higher investment could be that subjects are influenced by the emphasis put on beating a benchmark in real-world investment contexts. However, our experiment cannot distinguish this explanation from alternatives, and future research is needed to tease out the exact mechanism behind our results.

Heterogeneous effects: We examine whether real-world stock market experience influences responses to benchmark provision by splitting the sample into participants with and without stock market experience. We measured stock market experience based on participants' responses to the question, "Have you ever made investments (either personal or through your employment) in the common stock or shares of a company?" on Prolific. Participants were split into Stock Market Participants (SMP) and Non-Stock Market Participants (NSMP) based on their "yes" or "no" answers. Participants who selected "don't know" or "rather not say" were excluded from the analysis.

The results are presented in Table 3. The coefficients in columns (1) and (2) indicate that the effect of the low benchmark on the forecast for the subsample of SMP and NSMP is almost identical to the one in the full sample (column 2 of Table 2). However, the reduced sample size renders the effects insignificant. The effect of the high benchmark on fund forecasts is stronger in the SMP sample ($\beta = 0.764, p < 0.01$). In contrast, it has a lower economic and statistical significance in the NSMP sample ($\beta = 0.313, p < 0.05$), relative to the baseline regression. The impact of predicted fund returns on investment in columns (3) and (4) is qualitatively and quantitatively similar to that

Table 3: Stock Market Participation Sample Split

	(1) Forecast SMP	(2) Forecast NSMP	(3) Investment SMP	(4) Investment NSMP	(5) Investment SMP	(6) Investment NSMP
High	0.764***	0.313**			19.11*	-3.638
	(4.26)	(2.48)			(1.68)	(-0.44)
Low	0.210	0.204			38.25***	11.84
	(1.24)	(1.63)			(3.26)	(1.40)
Forecast			21.55***	17.66***		
			(9.12)	(9.10)		
Mean of DV	9.1	9.1	339.6	339.6	339.6	339.6
Fund FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1755	2637	1755	2637	1755	2637
Adjusted R^2	0.455	0.490	0.150	0.144	0.102	0.105

Notes: Results from OLS regressions on benchmark dummies and fund forecasts in the sample of Stock Market Participants (SMP) and Non-Stock Market Participants (NSMP). "High" and "Low" are dummy variables indicating whether subjects saw fund returns alongside a benchmark with higher or lower returns, respectively; "Fund Only" is the baseline category. "Forecast" represents the forecast of the fund, and "Investment" reflects the amount invested in the fund. Control variables include CRT performance, personal income, age and gender. Standard errors are clustered at the subject level. T-Statistics are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

in the baseline regression. However, SMP and NSMP differ in the magnitude of the effect of the benchmark on investment amounts. The results presented in column (5) suggest that the coefficient of the high benchmark is marginally statistically significant in the SMP sample ($\beta = 19.11, p < 0.1$). More interestingly, the effect of the low benchmark on investment volume has gained in economic significance ($\beta = 38.25, p < 0.01$), indicating that investment experience acquired through past stock market participation could enhance individuals' preference to invest in mutual funds that exceed market performance. In the sample of NSMP in column (6), neither the high nor the low benchmark influences investment decisions. In conclusion, we find that benchmarks that outperform the fund increase the return expectations for mutual funds more among participants with stock market experience than among those without. Moreover, both types of benchmarks have a larger effect on the investment decisions of stock market participants than on those of nonparticipants.

5 Conclusion

Although the returns of assets are commonly shown accompanied by the returns of a corresponding benchmark, little is known about how the benchmarks affect investors' beliefs about an asset's future returns and their willingness to invest. In this article, we experimentally investigate the

influence of benchmarks in the context of mutual funds. We demonstrate that the presence of a benchmark—whether it outperforms or underperforms the fund—leads to more positive return expectations. However, only a benchmark that underperforms the fund causes more investment. The results for the underperforming benchmark are in line with a contrast effect, which has been documented in many other decision contexts. However, the results for the outperforming benchmark are incompatible with a contrast effect. Our results indicate the need for future research in two areas: (1) the mechanism by which benchmarks influence investors and (2) the potential for strategic use of benchmarks to encourage investment in products that benefit providers more than investors.

References

Amemiya, Takeshi (1973). "Regression Analysis when the Dependent Variable Is Truncated Normal". In: *Econometrica* 41.6, pp. 997–1016. ISSN: 00129682, 14680262.

Antoniou, Constantinos, Junyang Guo, and Neil Stewart (2021). "It Depends Who You Ask: Context Effects in the Perception of Stock Returns". In: *SSRN Electronic Journal*. ISSN: 1556-5068. DOI: 10.2139/ssrn.3777778.

Berk, Jonathan B. and Richard C. Green (2004). "Mutual Fund Flows and Performance in Rational Markets". In: *Journal of Political Economy* 112.6, pp. 1269–1295. ISSN: 0022-3808. DOI: 10.1086/424739.

Bhargava, Saurabh and Ray Fisman (2014). "Contrast Effects in Sequential Decisions: Evidence from Speed Dating". In: *The Review of Economics and Statistics* 96.3, pp. 444–457. ISSN: 0034-6535. DOI: 10.1162/REST_a_00416.

Caplin, Andrew and John V Leahy (n.d.). "Wishful Thinking". In: ().

Cash, Thomas F., Diane Walker Cash, and Jonathan W. Butters (1983). ""Mirror, Mirror, on the Wall...?": Contrast Effects and Self-Evaluations of Physical Attractiveness". In: *Personality and Social Psychology Bulletin* 9.3, pp. 351–358. ISSN: 0146-1672. DOI: 10.1177/0146167283093004.

Chen, Huaizhi, Richard Evans, and Yang Sun (2025). "Self-Declared benchmarks and fund manager intent: "Cheating" or competing?" In: *Journal of Financial Economics* 165, p. 103975. ISSN: 0304-405X. DOI: https://doi.org/10.1016/j.jfineco.2024.103975.

Choi, James J and Kevin Zhao (2020). *Did Mutual Fund Return Persistence Persist?* Working Paper 26707. National Bureau of Economic Research. DOI: 10.3386/w26707.

Cremers, K. J. Martijn, Jon A. Fulkerson, and Timothy B. Riley (2022). "Benchmark Discrepancies and Mutual Fund Performance Evaluation". In: *Journal of Financial and Quantitative Analysis* 57.2, pp. 543–571. DOI: 10.1017/S0022109021000119.

Elton, Edwin J., Martin J. Gruber, and Christopher R. Blake (2013). "The Performance of Separate Accounts and Collective Investment Trusts*". In: *Review of Finance* 18.5, pp. 1717–1742. ISSN: 1572-3097. DOI: 10.1093/rof/rft038. eprint: https://academic.oup.com/rof/article-pdf/18/5/1717/26315664/rft038.pdf.

European Commission (2021). *Commission Delegated Regulation (EU) of 7.9.2021 Amending the Regulatory Technical Standards Laid Down in Commission Delegated Regulation (EU) 2017/653*. Official Journal of the European Union. Accessed: June 21, 2024.

European Parliament and Council of the European Union (2009). *Directive 2009/65/EC of the European Parliament and of the Council of 13 July 2009 on the Coordination of Laws, Regulations and Administrative Provisions Relating to Undertakings for Collective Investment in Transferable Securities (UCITS) (Recast)*. Official Journal of the European Union. Accessed: June 21, 2024.

European Securities and Markets Authority (2024). *Consolidated Questions and Answers on the PRIIPs Key Information Document (KID)*. Available at https://www.esma.europa.eu/sites/default/files/2023-05/JC_2023_22_-_Consolidated_JC_PRIIPs_Q_As.pdf. Last accessed 17 March 2025.

Frydman, Cary and Baolian Wang (2020). "The Impact of Salience on Investor Behavior: Evidence from a Natural Experiment". In: *Journal of Finance* 75.1, pp. 229–276. DOI: 10.1111/jofi.12851.

Glaser, Markus, Zwetelina Iliewa, and Martin Weber (2019). "Thinking about Prices versus Thinking about Returns in Financial Markets". In: *The Journal of Finance* 74.6, pp. 2997–3039. ISSN: 1540-6261. DOI: 10.1111/jofi.12835.

Hartzmark, Samuel M and David H Solomon (2022). "Reconsidering Returns". In: *The Review of Financial Studies* 35.1, pp. 343–393. DOI: 10.1093/rfs/hhab015.

Hartzmark, Samuel M. and Kelly Shue (2018). "A Tough Act to Follow: Contrast Effects in Financial Markets". In: *The Journal of Finance* 73.4, pp. 1567–1613. ISSN: 1540-6261. DOI: 10.1111/jofi.12685.

Kahneman, Daniel and Amos Tversky (1979). "Prospect Theory: An Analysis of Decision under Risk". In: *Econometrica* 47.2, pp. 263–291. ISSN: 00129682, 14680262.

Kenrick, Douglas T. and Sara E. Gutierres (1980). "Contrast Effects and Judgments of Physical Attractiveness: When Beauty Becomes a Social Problem". In: *Journal of Personality and Social Psychology* 38.1, pp. 131–140. ISSN: 1939-1315(Electronic),0022-3514(Print). DOI: 10.1037/0022-3514.38.1.131.

Kim, Jae Hyoung and Elizabeth Hoffman (2020). "Contrast Effects in Investment and Financing Decisions". In: *SSRN Electronic Journal*. ISSN: 1556-5068. DOI: 10.2139/ssrn.3256087.

Mullally, Kevin and Andrea Rossi (2024). "Moving the Goalposts? Mutual Fund Benchmark Changes and Relative Performance Manipulation". In: *The Review of Financial Studies* 38.4, pp. 1067–1119.

Pepitone, Albert and Mark DiNubile (1976). "Contrast Effects in Judgments of Crime Severity and the Punishment of Criminal Violators". In: *Journal of Personality* 33.4, pp. 448–459. ISSN: 0022-3514.

Radbruch, Jonas and Amelie Schiprowski (2024). "Interview Sequences and the Formation of Subjective Assessments". In: *The Review of Economic Studies*, rdae039. ISSN: 0034-6527. DOI: 10.1093/restud/rdae039.

Schkade, David A. and Daniel Kahneman (1998). "Does Living in California Make People Happy? A Focusing Illusion in Judgments of Life Satisfaction". In: *Psychological Science* 9.5, pp. 340–346. ISSN: 0956-7976. DOI: 10.1111/1467-9280.00066.

Securities and Exchange Commission (1993). *Final Rule: Disclosure of Mutual Fund Performance and Portfolio Managers*. Available at https://www.sec.gov/files/rules/final/33-6988.pdf. Last accessed 17 March 2025. Release Nos. 33-6988; IC-19382.

— (2023). *Form N-1A: Registration Statement Under the Securities Act of 1933 and/or Registration Statement Under the Investment Company Act of 1940*. Available at https://www.sec.gov/files/form-n-1a.pdf. Last accessed 17 March 2025.

Sensoy, Berk A. (2009). "Performance Evaluation and Self-Designated Benchmark Indexes in the Mutual Fund Industry". In: *Journal of Financial Economics* 92.1, pp. 25–39. ISSN: 0304-405X. DOI: 10.1016/j.jfineco.2008.02.011.

Tobin, James (1958). "Estimation of Relationships for Limited Dependent Variables". In: *Econometrica* 26.1, pp. 24–36. ISSN: 00129682, 14680262.

Appendix

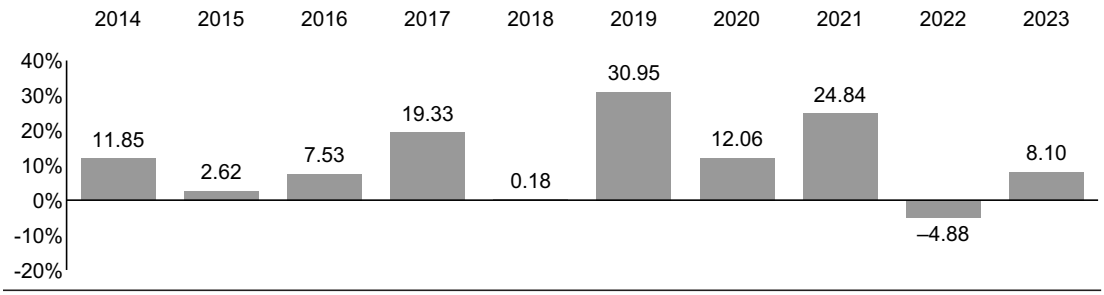
A Mutual Fund Reporting Standards in the US and Europe

In the United States, mutual funds must adhere to strict disclosure regulations, as specified in Form N-1A, enforced by the Securities and Exchange Commission (2023). These regulations require fund managers to report the 1-, 5-, and 10-year historical returns for both the fund and an appropriate benchmark (see Figure A.1). Similarly, it is common practice for mutual fund companies to display the historical returns of their funds alongside benchmarks in bar charts on their websites (see Figure A.2).

In contrast, the EU's approach, initiated with Directive 2009/65/EC in 2009, mandates fund managers to include historical performance data in Key Investor Information Documents (KIIDs) to enhance transparency and investor protection (European Parliament and Council of the European Union, 2009). In these regulated documents, fund managers were mandated to disclose the historical performance of the mutual fund alongside a relevant benchmark. With the introduction of the revised regulatory technical standards (RTS) in 2022, which mandated the replacement of the KIIDs with Key Information Documents (KIDs) (European Commission, 2021), it has become increasingly common to substitute historical returns with performance scenarios in such regulated documents, presenting potential future outcomes based on past performance data. However, when managers of a mutual fund that is managed in reference to a benchmark decide to display the historical returns of the fund, they are still obliged to show the past performance of the benchmark, too (European Securities and Markets Authority 2024). In the United Kingdom, KIIDs will remain in use until the end of 2026. Therefore, the past performance of both the mutual fund and a benchmark is still commonly shown together (see Figure A.3). Moreover, in fund prospectuses in the EU, the past performance of both the mutual fund and a benchmark is often displayed (see Figure A.4).

Figure A.1: Performance section of the Vanguard Dividend Growth Fund

Annual Total Returns — Vanguard Dividend Growth Fund Investor Shares[1]



1 The year-to-date return as of the most recent calendar quarter, which ended on March 31, 2024, was 6.01%.

During the periods shown in the bar chart, the highest and lowest returns for a calendar quarter were:

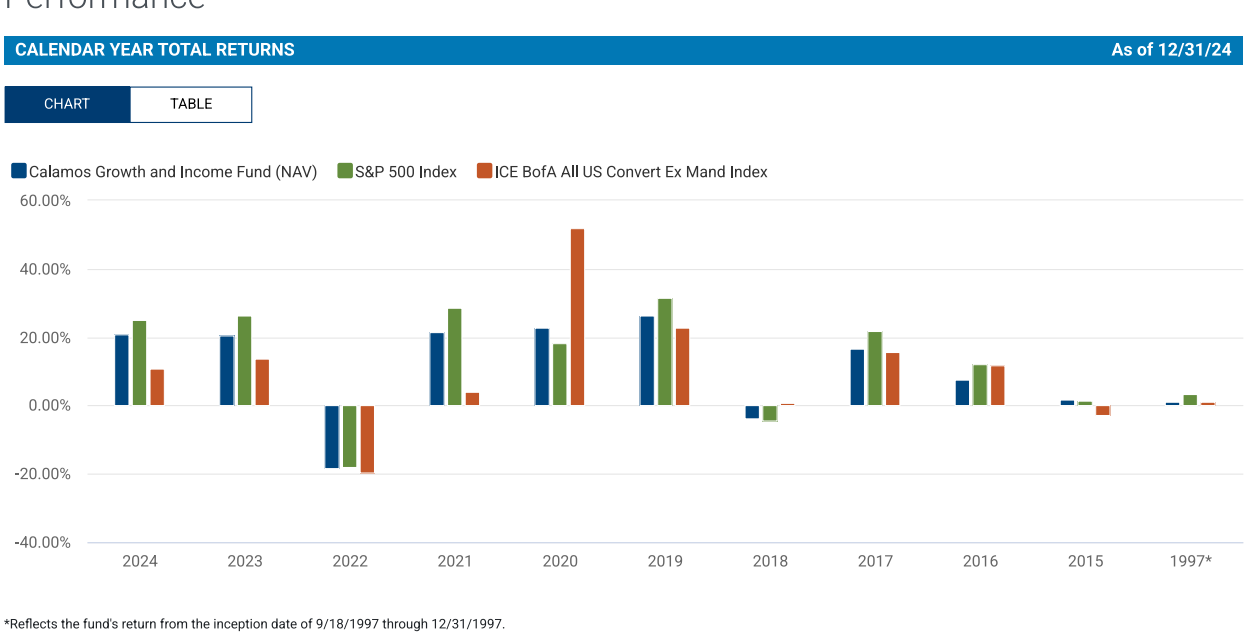
	Total Return	Quarter
Highest	13.84%	March 31, 2019
Lowest	-17.29%	March 31, 2020

Average Annual Total Returns for Periods Ended December 31, 2023

	1 Year	5 Years	10 Years
Vanguard Dividend Growth Fund Investor Shares			
Return Before Taxes	8.10%	13.50%	10.76%
Return After Taxes on Distributions	7.51	12.34	9.51
Return After Taxes on Distributions and Sale of Fund Shares	5.20	10.68	8.53
Dividend Growth Spliced Index (reflects no deduction for fees, expenses, or taxes)	14.52%	13.92%	10.77%
S&P U.S. Dividend Growers Index (reflects no deduction for fees, expenses, or taxes)	14.52	—	—
Dow Jones U.S. Total Stock Market Float Adjusted Index (reflects no deduction for fees, expenses, or taxes)	26.06	15.05	11.40

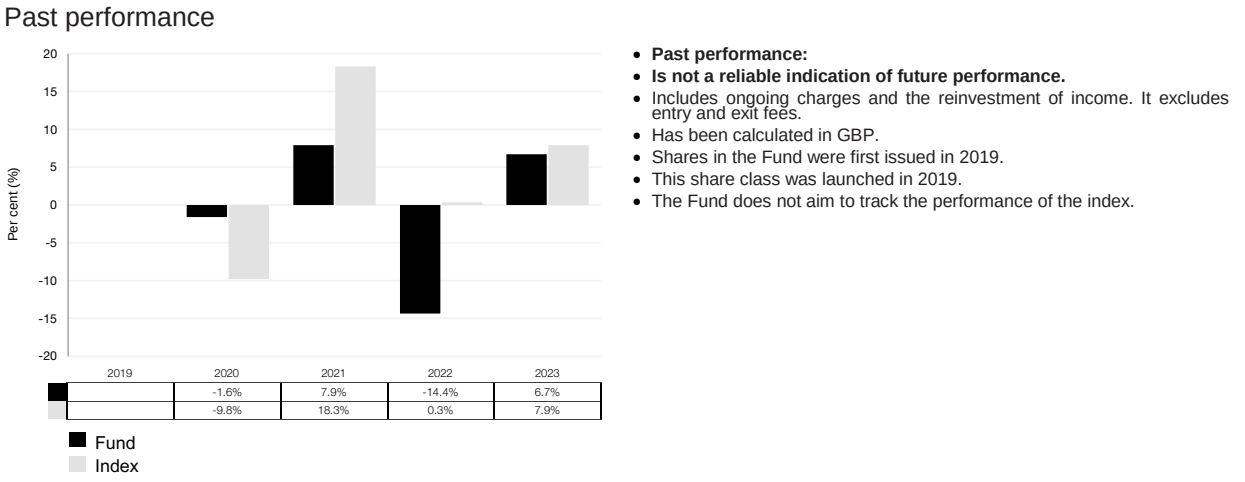
Notes: Summary prospectus as of 05/24/2024

Figure A.2: Bar chart showing the historical returns for the Calamos Growth and Income Fund



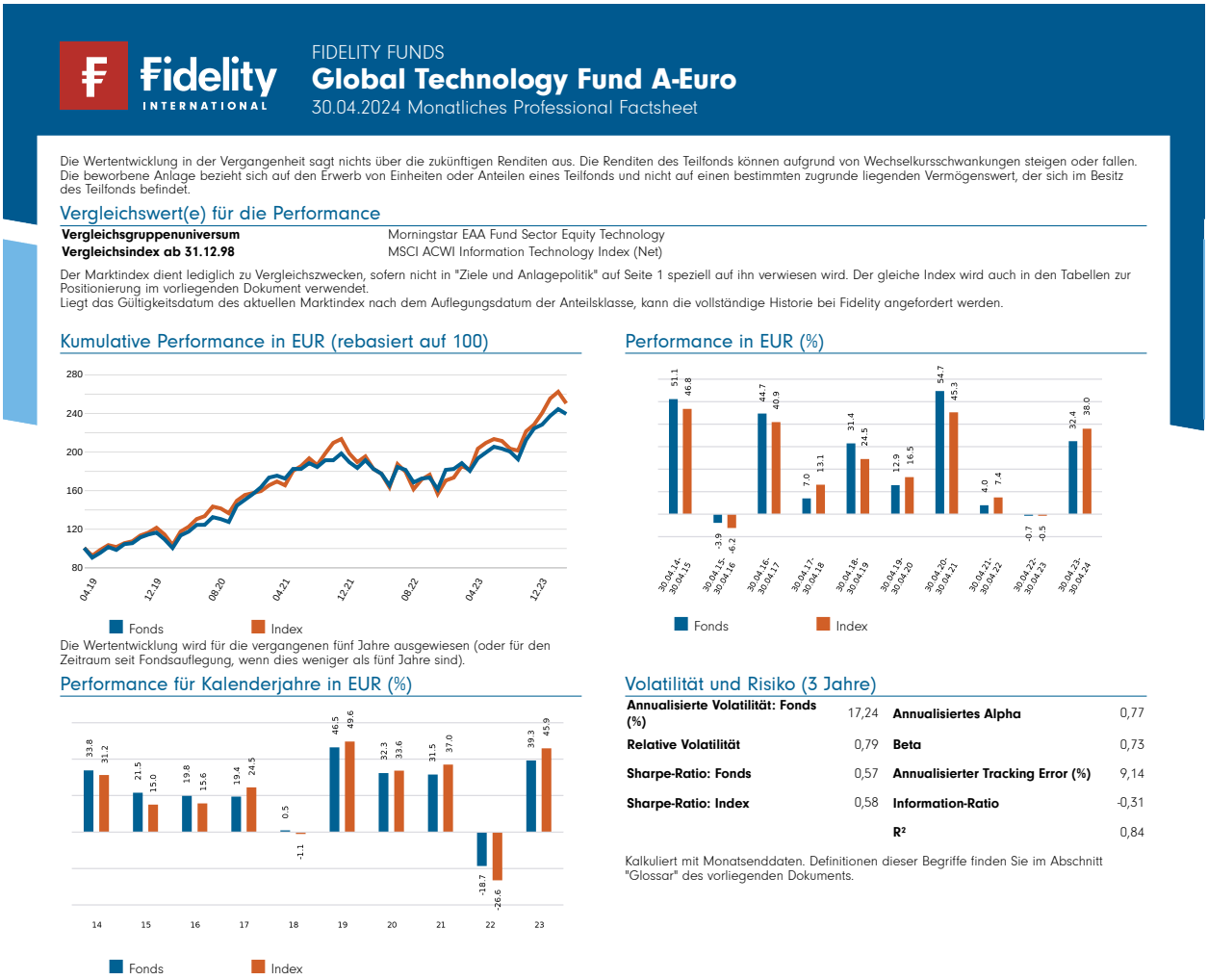
Notes: The fund is compared to relevant indices, as of 12/31/2024. Investors can view this data in either chart or table format, with the bar chart representation being the default option.

Figure A.3: Bar chart showing past performance section of the Vanguard Active U.K. Equity Fund



Notes: The chart is taken from the Key Investor Information Document (KIID) as of 05/03/2024. The benchmark shown is the FTSE All-Share Index.

Figure A.4: Performance section of the Fidelity Funds Global Technology Fund A-Euro



FIDELITY FUNDS

Global Technology Fund A-Euro

30.04.2024 Monatliches Professional Factsheet

Die Wertentwicklung in der Vergangenheit sagt nichts über die zukünftigen Renditen aus. Die Renditen des Teilfonds können aufgrund von Wechselkursschwankungen steigen oder fallen. Die beworbene Anlage bezieht sich auf den Erwerb von Einheiten oder Anteilen eines Teilfonds und nicht auf einen bestimmten zugrunde liegenden Vermögenswert, der sich im Besitz des Teilfonds befindet.

Vergleichswert(e) für die Performance

Vergleichsgruppenuniversum	Morningstar EAA Fund Sector Equity Technology
Vergleichsindex ab 31.12.98	MSCI ACWI Information Technology Index (Net)

Der Marktindex dient lediglich zu Vergleichszwecken, sofern nicht in "Ziele und Anlagepolitik" auf Seite 1 speziell auf ihn verwiesen wird. Der gleiche Index wird auch in den Tabellen zur Positionierung im vorliegenden Dokument verwendet.
Liegt das Gültigkeitsdatum des aktuellen Marktindex nach dem Auflegungsdatum der Anteilsklasse, kann die vollständige Historie bei Fidelity angefordert werden.

Kumulative Performance in EUR (rebasiert auf 100)

Die Wertentwicklung wird für die vergangenen fünf Jahre ausgewiesen (oder für den Zeitraum seit Fondsauflegung, wenn dies weniger als fünf Jahre sind).

Performance in EUR (%)

Performance für Kalenderjahre in EUR (%)

Volatilität und Risiko (3 Jahre)

Annualisierte Volatilität: Fonds (%)	17,24	**Annualisiertes Alpha**	0,77
Relative Volatilität	0,79	**Beta**	0,73
Sharpe-Ratio: Fonds	0,57	**Annualisierter Tracking Error (%)**	9,14
Sharpe-Ratio: Index	0,58	**Information-Ratio**	-0,31
		R²	0,84

Kalkuliert mit Monatsenddaten. Definitionen dieser Begriffe finden Sie im Abschnitt "Glossar" des vorliegenden Dokuments.

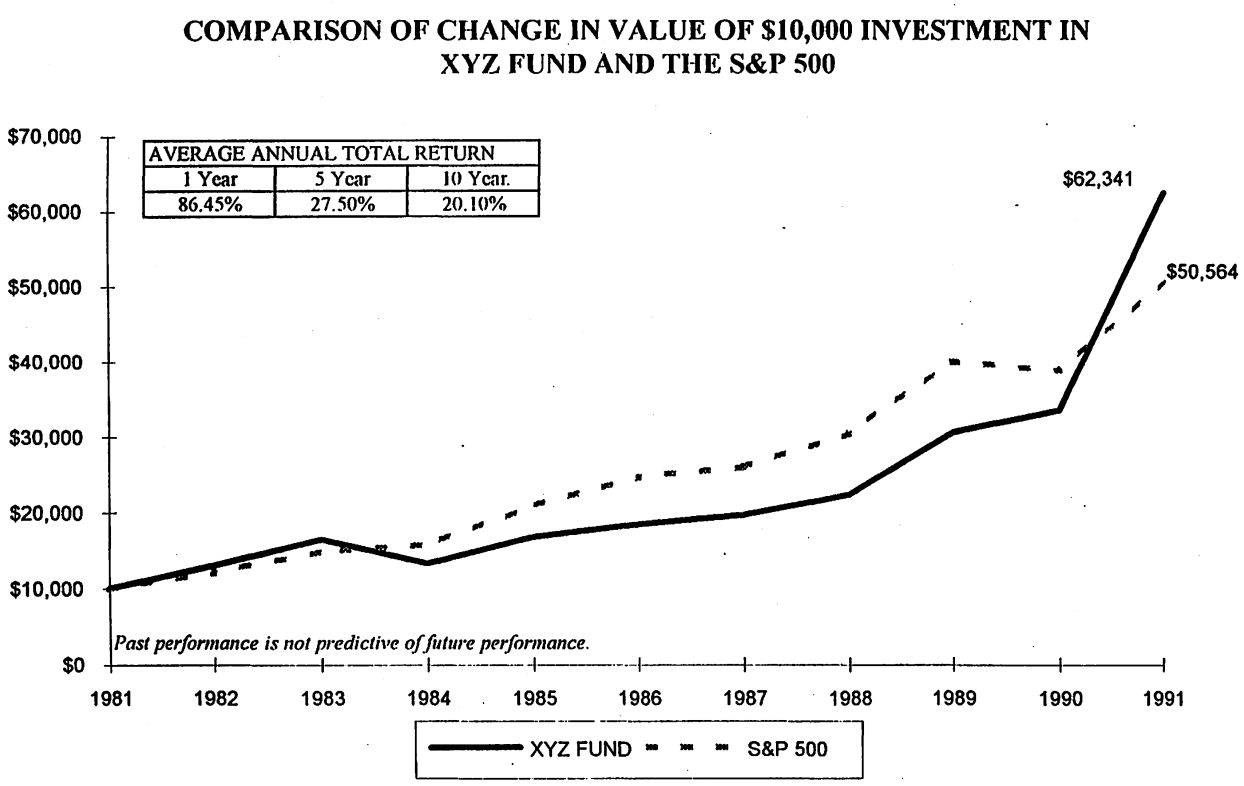
Performance bis 30.04.24 in EUR (%)

	1 Monat	3 Monate	Seit Jahresbeginn	1 Jahr	3 Jahre	5 Jahre	Seit 30.09.99*
Kumulatives Fondswachstum	-2,3	4,9	6,7	32,4	36,7	138,8	528,3
Kumulatives Indexwachstum	-4,6	4,2	9,4	38,0	47,5	149,6	551,8
Fonds pro Jahr im Durchschnitt	-	-	-	32,4	11,0	19,0	7,8
Index pro Jahr im Durchschnitt	-	-	-	38,0	13,8	20,1	7,9
Ranking in Vergleichsgruppenuniversum							
Y-Euro	38	92	107	99	13	11	
Gesamtzahl vergleichbarer Fonds	216	214	214	211	142	102	
Quartilsplatzierung**	1	2	2	2	1	1	

Angaben zur Fondsperformance, zur Volatilität und zu den Risikokennzahlen stammen von Fidelity. Bei der Berechnung der Wertentwicklung wird der Ausgabeaufschlag nicht berücksichtigt. Basis: Vergleich der Nettoinventarwerte, Wiederanlage der Erträge, auf EUR-Basis, nach Abzug von Gebühren. Zieht man von einem Investment einen Ausgabeaufschlag von 5,25 % ab, entspricht dies der Reduzierung einer jährlichen Wachstumsrate von 6 % über 5 Jahre auf 4,9 %. Es handelt sich hierbei um den höchsten anwendbaren Ausgabeaufschlag. Beträgt der von Ihnen zu zahlende Ausgabeaufschlag weniger als 5,25 %, ist die Auswirkung auf die Gesamtwertentwicklung geringer. Quelle der Marktindizes ist RIMES, andere Daten stammen von Fremdanbietern wie Morningstar.
* Performance-Anfangsdatum.
**Die Quartilsplatzierung gilt für die von Morningstar ermittelte primäre Anteilsklasse des Teilfonds, die von der in diesem Factsheet beschriebenen Anteilsklasse abweichen kann, und sie bezieht sich auf die Wertentwicklung im Zeitverlauf, die auf einer Skala von 1-4 bewertet wird. Eine Platzierung von 1 bedeutet, dass das platzierte Element zu den besten 25 % der Stichprobe gehört usw. Die Platzierungen basieren auf dem Wertentwicklungsverlauf aus dem Peer Group-Universum. Im Einklang mit der Methodik von „The Investment Association" können diese Aufzeichnungen eine Verlängerung der bisherigen Wertentwicklung einer früheren Anteilsklasse enthalten, und der Verlauf gilt möglicherweise nicht für die gleiche Klasse wie dieses Factsheet. Bei der Quartilsplatzierung handelt es sich um eine von Fidelity International durchgeführte interne Berechnung. Die Platzierung kann sich je nach Anteilsklasse unterscheiden.

Notes: A page from the factsheet as of 04/30/2024. The benchmark used is the MSCI ACWI Information Technology Index.

Figure A.5: Graphical Comparison of Mutual Fund vs. S&P 500



Notes: This graph, extracted from the "Final Rule: Disclosure of Mutual Fund Performance and Portfolio Managers" by the Securities and Exchange Commission (1993), shows how the SEC used the S&P 500 as an example of a broad-based index for required mutual fund comparisons at that time, contributing to the widespread use of the S&P 500 in fund performance evaluations.

B Instructions of the Experiment

Figure A.6: General Instructions

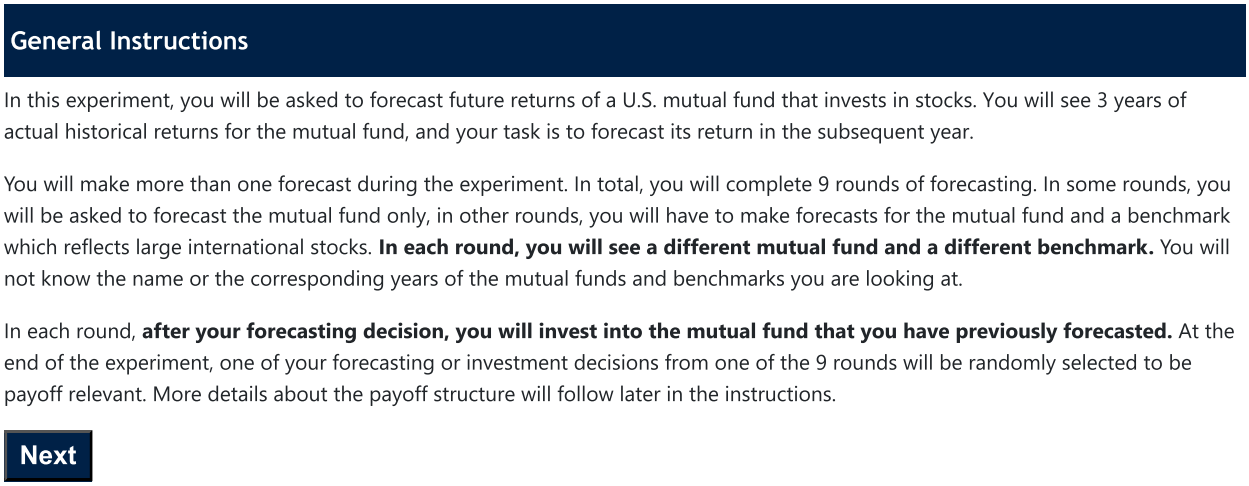


Figure A.7: Instructions: Mutual Funds and Benchmarks

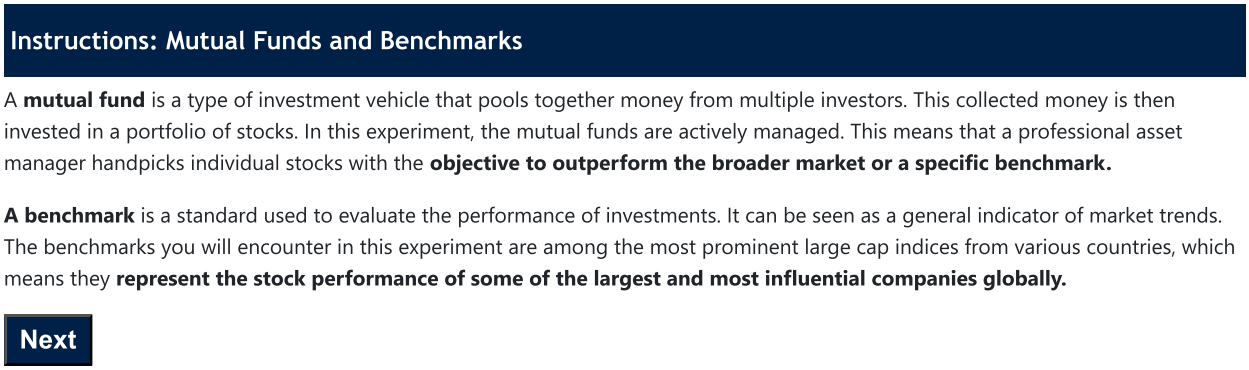


Figure A.8: Payment Instructions

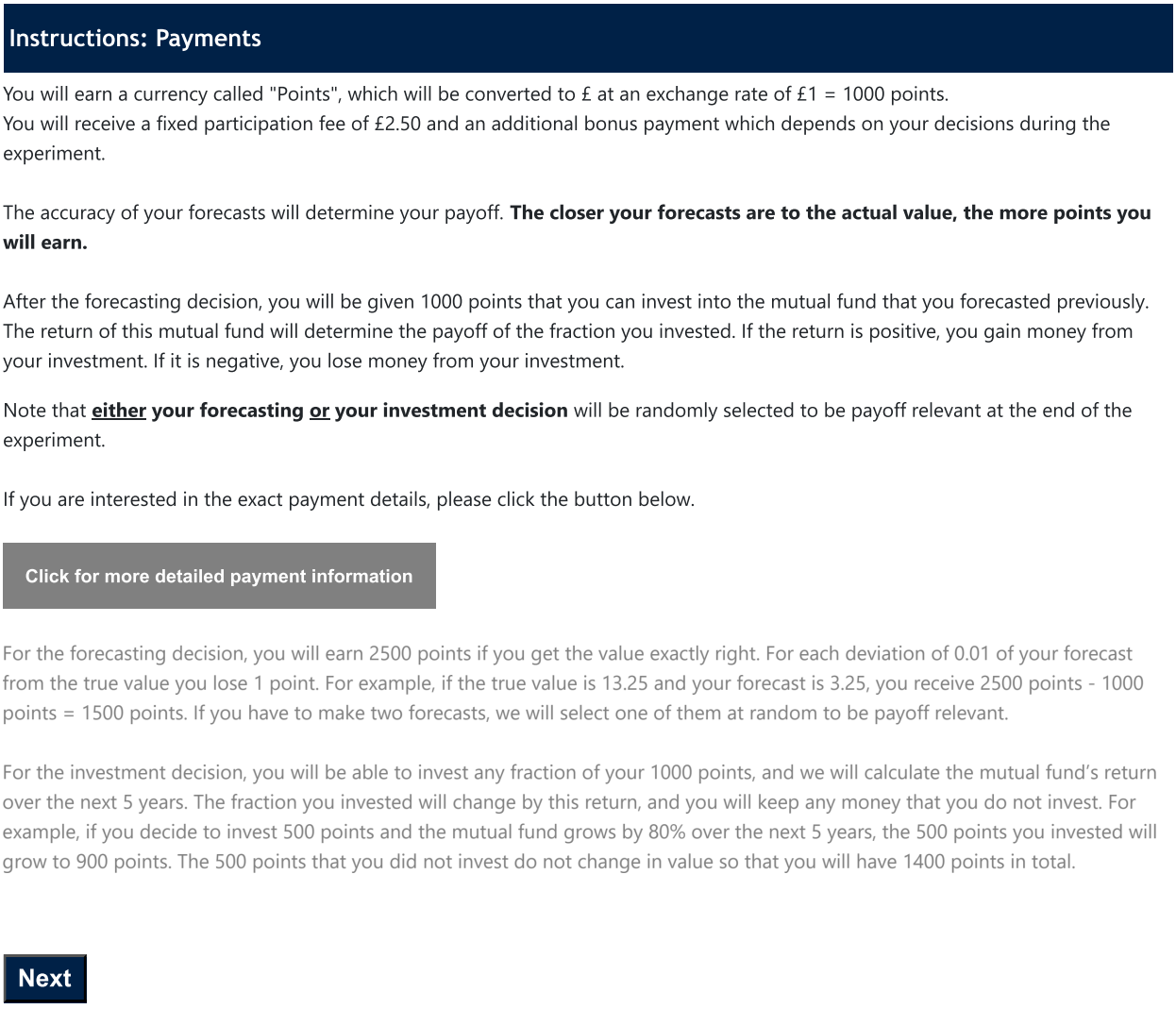


Figure A.9: Comprehension and Attention Check

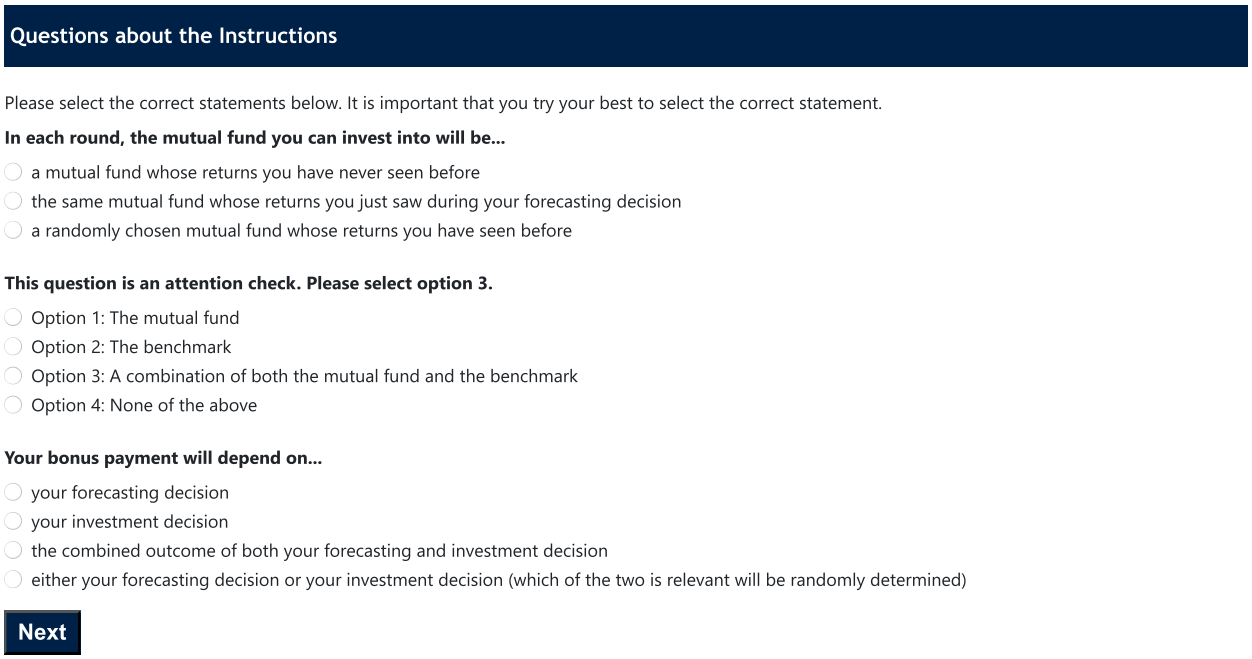
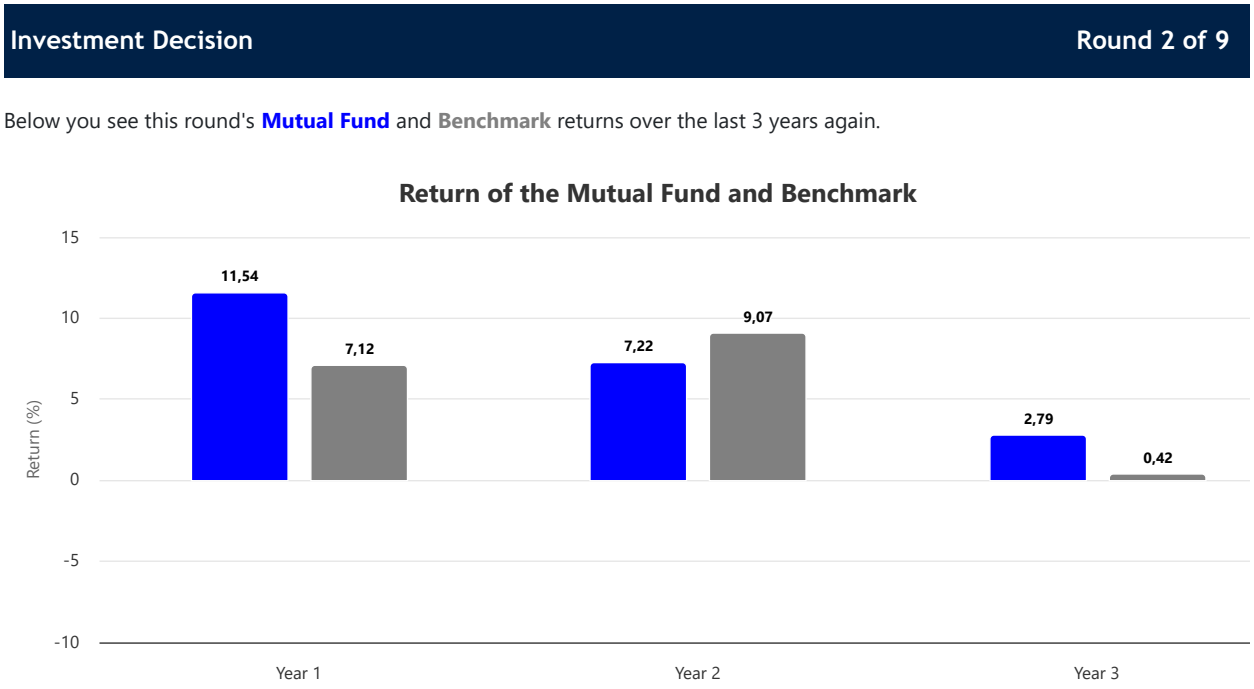


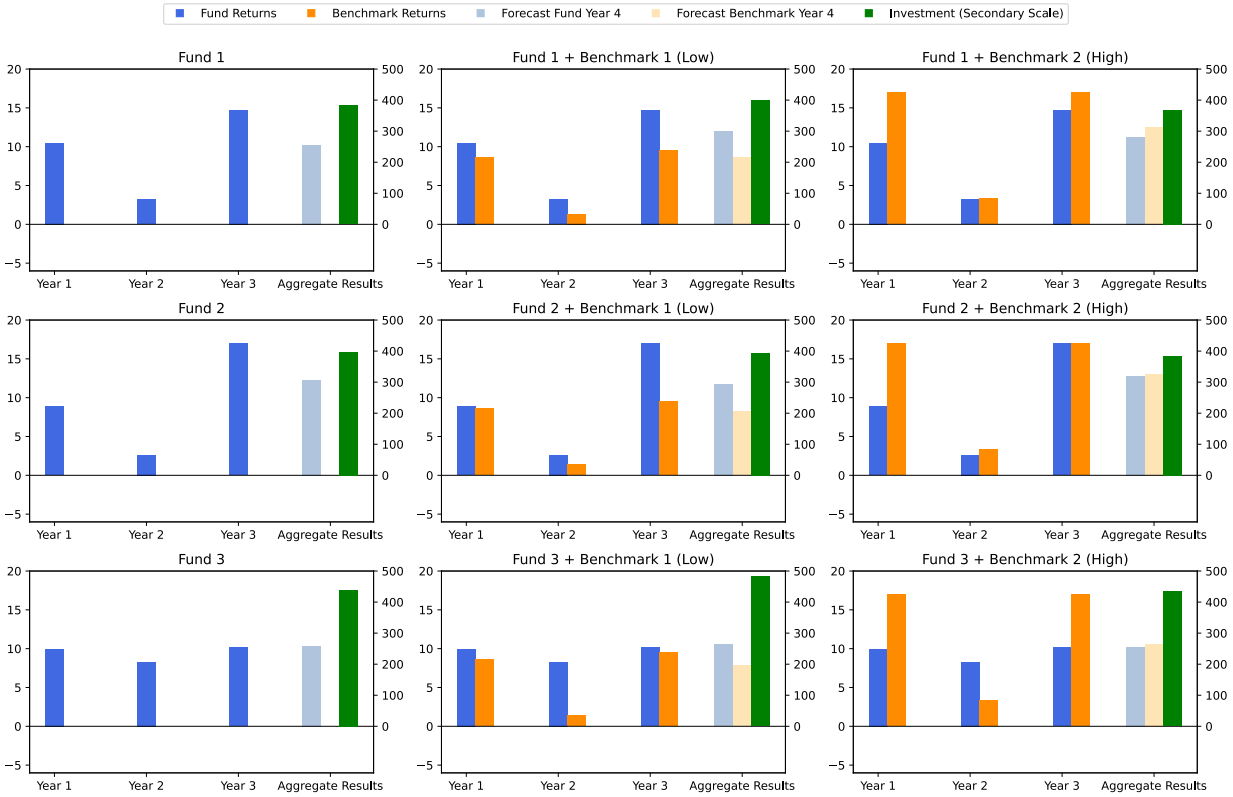
Figure A.10: Investment Decision Screen



Notes: After forecasting the fourth-year returns of both the mutual fund and its benchmark on the previous screen, participants were asked to indicate how much of their initial 1,000-point investment they were willing to allocate to the fund. The compensation for participants' investments was based on the actual performance of the mutual fund over the subsequent five years.

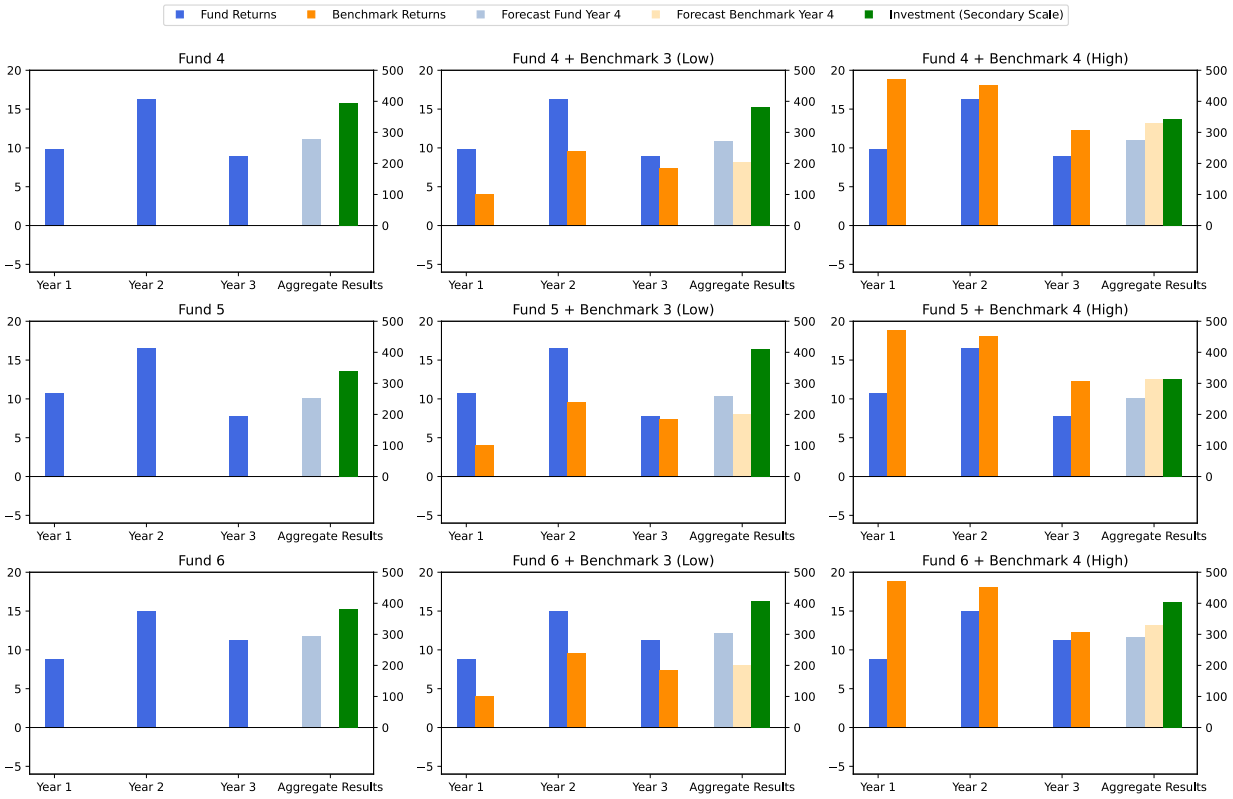
C Additional Experimental Results

Figure A.11: Returns, Forecasts and Investment Amounts for Funds 1-3



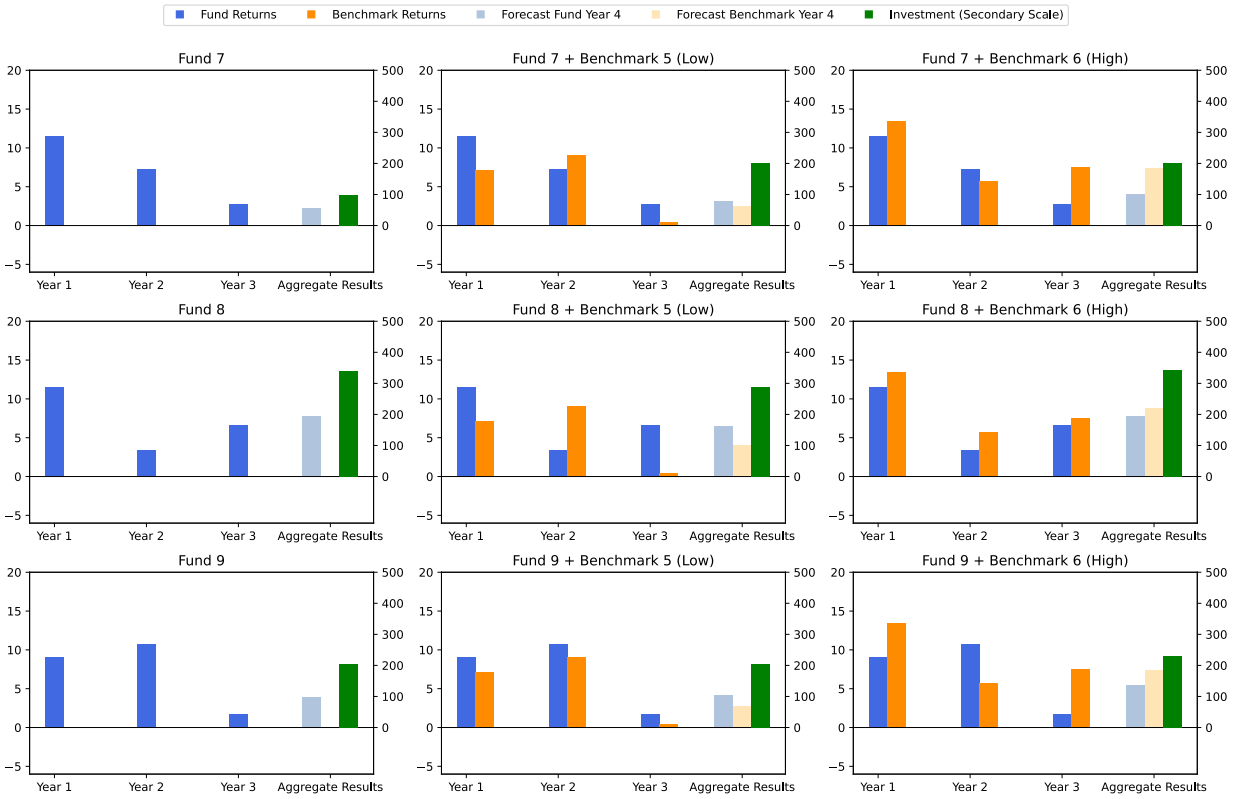
Notes: For each fund, participants were divided into three equal groups for the experiment. The first group observed the mutual fund in isolation (left section), the second group viewed the fund alongside a benchmark with lower average returns (middle section), and the third group was shown the fund with a benchmark exhibiting higher returns (right section). Fully opaque bars indicate the actual returns of the fund (blue) and the benchmark (orange) as presented to participants during the study. Bars with reduced opacity depict the average forecasted fourth-year returns for the fund (blue) and the benchmark (orange). The green bars show the mean investment volume that participants were willing to invest. For an overview of the funds and benchmarks, see Tables A.1 and A.2.

Figure A.12: Returns, Forecasts and Investment Amounts for Funds 4-6



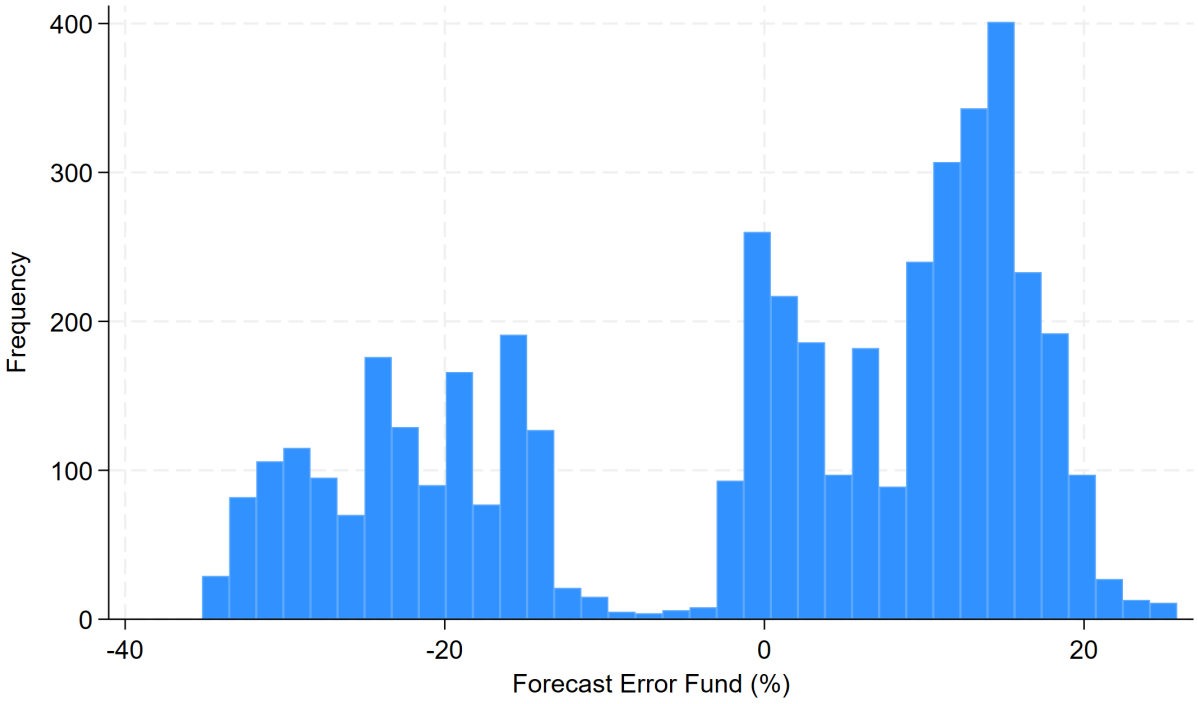
Notes: For each fund, participants were divided into three equal groups for the experiment. The first group observed the mutual fund in isolation (left section), the second group viewed the fund alongside a benchmark with lower average returns (middle section), and the third group was shown the fund with a benchmark exhibiting higher returns (right section). Fully opaque bars indicate the actual returns of the fund (blue) and the benchmark (orange) as presented to participants during the study. Bars with reduced opacity depict the average forecasted fourth-year returns for the fund (blue) and the benchmark (orange). The green bars show the mean investment volume that participants were willing to invest. For an overview of the funds and benchmarks, see Tables A.1 and A.2.

Figure A.13: Returns, Forecasts and Investment Amounts for Funds 7-9



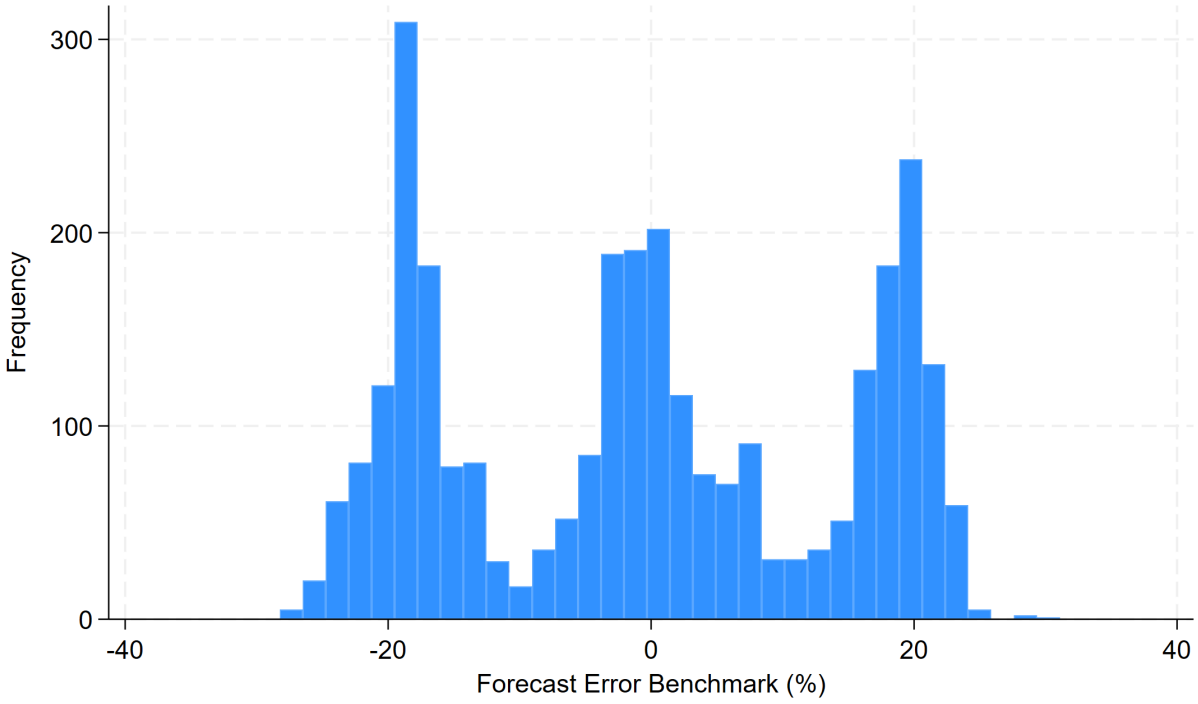
Notes: For each fund, participants were divided into three equal groups for the experiment. The first group observed the mutual fund in isolation (left section), the second group viewed the fund alongside a benchmark with lower average returns (middle section), and the third group was shown the fund with a benchmark exhibiting higher returns (right section). Fully opaque bars indicate the actual returns of the fund (blue) and the benchmark (orange) as presented to participants during the study. Bars with reduced opacity depict the average forecasted fourth-year returns for the fund (blue) and the benchmark (orange). The green bars show the mean investment volume that participants were willing to invest. For an overview of the funds and benchmarks, see Tables A.1 and A.2.

Figure A.14: Histogram of Forecast Errors of the Fund.



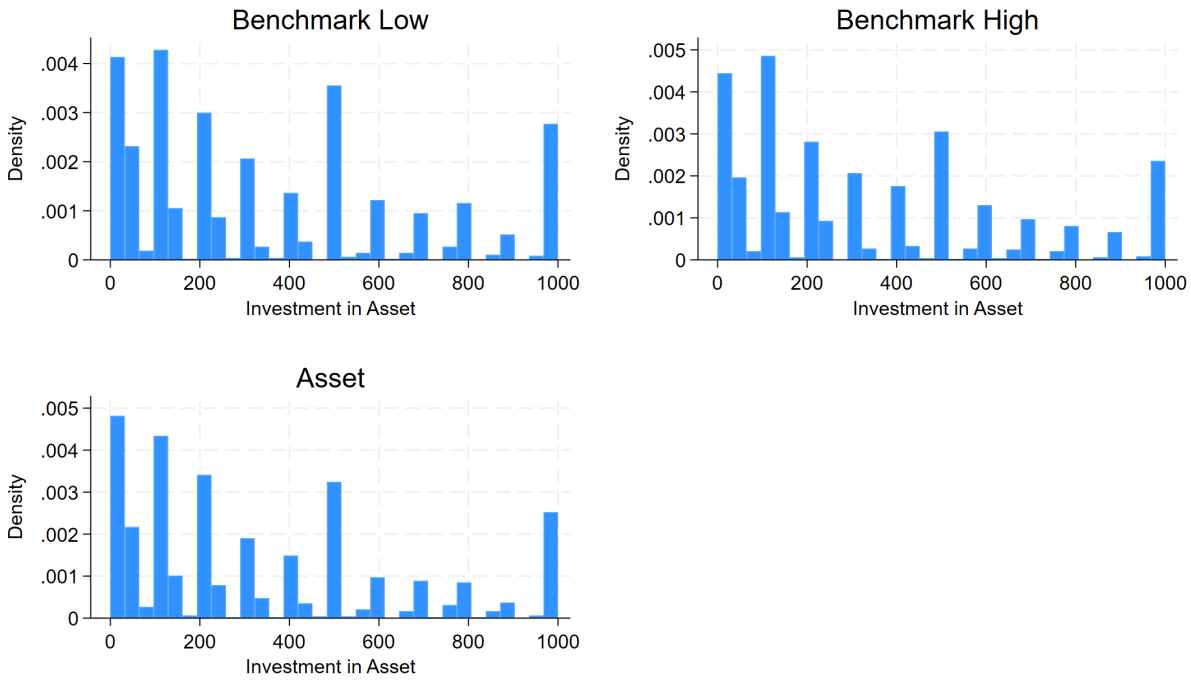
Notes: Forecast error is defined as the difference between participants' forecasted return for the fourth year (winsorized at the 1st and 99th percentiles) and the actual fourth-year return of the fund.

Figure A.15: Histogram of Forecast Errors of the Benchmark.



Notes: Forecast error is defined as the difference between participants' forecasted return for the fourth year (winsorized at the 1st and 99th percentiles) and the actual fourth-year return of the benchmark.

Figure A.16: Investment Amount Histograms.



Notes: The histograms reflect the distribution of investment amounts in the mutual fund when the fund is shown alongside a benchmark with lower returns (top left), a benchmark with higher returns (top right), and without a benchmark (bottom).

Table A.1: Overview of Funds

Fund	Fund Name	Time Frame	Mean Return	Standard Deviation	CRSP Fund Number	CRSP Style Code
1	PIMCO Funds: StocksPLUS Fund. Institutional Class Shares	2004-2006	9.47%	4.75%	23478	EDYB
2	Washington Mutual Investors Fund, Inc. Class 529-B Shares	2004-2006	9.47%	5.88%	31964	EDYB
3	Calamos Investment Trust: CALAMOS Growth & Income Fund. Class I Shares	2004-2006	9.47%	0.83%	6970	M
4	Calvert Social Investment Fund: Calvert Balanced Portfolio. Class C Shares	2012-2014	11.71%	3.28%	7028	M
5	Ultra Series Fund: Madison Target Retirement 2030 Fund. Class I Shares	2012-2014	11.70%	3.64%	41432	EDYB
6	Hennessy Funds Trust: Hennessy Equity and Income Fund. Investor Class Shares	2012-2014	11.70%	2.55%	2841	M
7	Deutsche Investment Trust: Deutsche Capital Growth Fund. Class C Shares	2014-2016	7.18%	3.57%	8593	EDYG
8	Thornburg Investment Trust: Thornburg Value Fund. Class R3 Shares	2014-2016	7.20%	3.31%	29855	EDYB
9	Metropolitan Series Fund: T Rowe Price Large Cap Growth Portfolio. Class A Shares	2014-2016	7.21%	3.92%	39086	EDYG

Notes: There are three sets of funds: Funds 1–3, 4–6, and 7–9. Within each set, the funds share the same average return, which differs across sets.

Table A.2: Overview of Benchmarks

Benchmark	**Benchmark Name**	**Time Frame**	**Mean Return**	**Standard Deviation**
1	NASDAQ	2004-2006	6.49%	3.64%
2	Russell 2000	2004-2006	12.44%	6.45%
3	S&P/TSX Composite	2012-2014	6.99%	2.29%
4	BEL 20	2012-2014	16.43%	2.89%
5	Nikkei 225	2014-2016	5.54%	3.70%
6	NASDAQ	2014-2016	8.88%	3.28%

Notes: Benchmarks 1 and 2 correspond to the low and high benchmarks presented alongside Funds 1-3, respectively. Benchmarks 3 and 4 serve as the low and high benchmarks for Funds 4-6, respectively, while Benchmarks 5 and 6 are the low and high benchmarks presented alongside Funds 7-9, respectively.

Table A.3: Tobit regression results for investment decisions.

	(1)	(2)	(3)	(4)
	Dependent Variable: Investment			
Forecast	27.92***	23.91***		
	(21.35)	(13.06)		
High			6.103	5.892
			(0.75)	(0.72)
Low			25.20***	25.12***
			(2.99)	(2.97)
Mean of DV	339.6	339.6	339.6	339.6
Fund FE	No	Yes	Yes	Yes
Controls	No	Yes	No	Yes
Observations	4500	4491	4500	4491
Pseudo R^2	0.010	0.014	0.007	0.009

Notes: "High" and "Low" are dummy variables indicating whether subjects saw fund returns alongside a benchmark with higher or lower returns, respectively; "Fund Only" is the baseline category. The dependent variable "Investment" reflects the amount invested in the fund and is censored between 0 and 1000. Columns (2) and (4) control for CRT performance, personal income, age and gender. Standard errors are clustered at the subject level. T-Statistics are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.