Journal of Banking and Finance 98 (2019) 61–79

Contents lists available at ScienceDirect

Journal of Banking and Finance

journal homepage: www.elsevier.com/locate/jbf

How persistent are the effects of experience sampling on investor behavior?

Meike A.S. Bradbury[a], Thorsten Hens[a,b,d], Stefan Zeisberger[a,c,*]

[a] Department of Banking and Finance, University of Zurich, Switzerland
[b] Department of Finance, Norwegian School of Economics, Bergen, Norway
[c] Institute for Management Research, Radboud University, Nijmegen, The Netherlands
[d] Department of Economics, University of Lucerne, Switzerland

ARTICLE INFO

Article history:
Received 12 July 2016
Accepted 26 October 2018
Available online 3 November 2018

JEL classification:
D81
G11

Keywords:
Behavioral finance
Simulated experience
Experience sampling
Investment decision
Risk communication
Financial advice
Fintech

ABSTRACT

Investor behavior was shown to be considerably different when the risk-return tradeoff is presented by experience sampling as opposed to a descriptive communication. We analyze the persistency of this difference in a setting in which investors are faced with multiple decisions over time and are consequently able to adjust the risk level they initially chose. For this we use an experimental setting with repeated investment decisions over multiple trading days, and we also test a new form of risk simulation in which wealth paths over time are presented rather than just final outcomes. After investors' initial decisions, for which we confirm previous findings, we do not find persistent differences of simulation-based learning on investors' risk-taking behavior. With regards to trading volume, only a simulation in which investors see wealth paths and not only final outcomes leads to lower trading frequency soon after the initial asset allocation.

1. Introduction

The way investment risks and returns are presented to investors matters for investment decisions. One possible distinction is between experiencing likely returns in a risk simulation and a description thereof, related to research in judgement and decision making on the experience-description gap (Hertwig et al., 2004). With risk simulations investors can experience return distributions by random sampling of possible outcomes, rather than getting risk and potential returns described in verbal or graphical ways. Risk simulations were found to lead to higher risk-taking (Kaufmann et al., 2013; Ehm et al., 2014; Bradbury et al., 2015). These studies further show that investors who are informed via risk simulations understand the underlying return distribution more accurately and feel better informed after using risk simulations. Compared to alternatives such as "nudging" investors to save more, risk simulations aim at providing information in a more accessible way and do not take a normative perspective.[1]

Importantly, studies on experience sampling in finance focus on the point of time at which the investor makes an initial investment decision. What has not yet been analyzed is the persistency of these effects, i.e. if risk simulations can change investor behavior *after* the initial decision. This includes investors' reaction to intermediate short-term gains and losses, or longer-term risk-taking behavior. This is important to explore since findings in behavioral finance provide evidence that investors follow their emotional instincts during their investment journey and question their decisions along the way (Shefrin and Statman 1994; Odean 1999; Barber and Odean 2001). They might change their risk preferences in cases of market up- or downturns (Cohn et al., 2015), they react to previous gains and losses in systematic ways

* Corresponding author at: Department of Banking and Finance, University of Zurich, Zurich, Switzerland.
E-mail addresses: s.zeisberger@fm.ru.nl, stefan.zeisberger@uzh.ch (S. Zeisberger).

[1] Various studies have suggested different methods to increase saving rates, including, for example, pre-commitment (Thaler and Benartzi, 2004), elaborating on the value of future rewards (Weber et al., 2007) or presenting investors with their future selves (Hershfield et al., 2011). Generally, this fits into the literature of how the presentation form of information effects investors' decisions, see, for example, Benartzi and Thaler (1999), Hirshleifer and Teoh (2003), Beshears et al. (2017) and Shaton (2017).

https://doi.org/10.1016/j.jbankfin.2018.10.014
0378-4266/

(Shefrin and Statement 1985; Imas 2016) and frequent feedback combined with short periods of commitment might reduce risk-taking, as demonstrated in lab settings ("myopic loss aversion"; Benartzi and Thaler 1995; Gneezy and Potters 1997; Thaler et al., 1997).[2] Actual experience might be more salient or could be perceived as more important by investors than initially gained simulated experience. Hence, it is not clear to what extent effects of risk simulations such as higher risk-taking persist over time. Against this background, this study aims at analyzing the persistency of the effects risk simulations have on investor behavior once we leave the point of time the investor makes her initial investment decision and can deviate from initial allocations for the rest of the investment horizon.

Given our focus on a *series* of investment decisions, we additionally introduce an alternative way of communicating investment risk. In this alternative investors are not only presented with final investment outcomes as in currently proposed risk simulation tools but they also experience wealth paths over time, which lead to final outcomes. The underlying idea of this "wealth path simulation" is to increase the salience of the fact that intermediate losses can occur but are usually offset over longer time horizons. We contrast this wealth path simulation with a risk simulation of only final outcomes and with a descriptive risk communication.

Our research aims at answering two main questions: First, what is the effect of risk simulations on investors' risk-taking behavior over an extended period of time, on their reactions to intermediate gains and losses, and on their trading activity? Secondly, what effect does a wealth path simulation have on investment decisions compared to simulating just final outcomes?

Our findings can be summarized as follows: Generally, we confirm previously reported results of risk simulations on the *initial* investment decision. We observe that investors who receive return information via a simulation feel better informed about the investment decision and show a higher objective understanding of return distributions. As a consequence, investors show a higher willingness to take risks. With regard to our first research question, we find that investors who are not informed via simulations tend to increase their risk-taking gradually after receiving feedback on actual investment outcomes, converging over time to the risk levels taken by investors who were informed via experience-based learning. The differences in reactions to previous gains and losses are small and limited to early trading decisions soon after the initial decision. We do not find that risk simulations change investors' overall trading volume in our setting.

With regard to our second research question, comparing the different forms of risk simulations, we find that presenting investors with final outcomes seems to have the strongest effect on initial risk-taking, while the effect appears to be slightly weaker for the wealth path simulation. However, wealth path simulations seem to decrease short-term reactions to intermediate outcomes, at least in early periods. They also have the strongest effects on the initial understanding of the return distribution. We conclude that the "wealth path simulation" leads to similar effects as only simulating final outcomes despite presenting more information. The former might materialize in a slightly smaller increase in risk-taking but seems to reduce reactions to gains and losses in the short-term. In an extended time horizon, however, the effects of risk simulations are generally quite limited in our setting. From our results it cannot be expected that a one-time risk simulation education will have persistent effects on investor behavior in the field, at least not in the tested variants.

[2] However, Beshears et al. (2017) question the robustness of the lab results and suggest that they might be not easily transferable to a real-world setting. Shaton (2017), however, empirically (using a natural experiment) finds that displaying longer time horizons (a minimum of 12 months versus 1 month) on a regular basis increases risk-taking.

2. Experimental design

2.1. General setup and experimental task

We programmed a proprietary online experiment tailored to this study. Participants had to make 14 investment decisions, with a maximum of one decision per calendar day. Our intention for this longer time delay compared to a short-term lab setting is to better mimic delays of real-world investment decisions. While our investment delays do not correspond to the real-world, our choice is motivated by previous findings documenting a difference between short- and long-term settings. Most related, Beshears et al. (2017), who use a real one-year investment horizon with intermediate decisions, find that observations of strong differences in risk-taking in a short-term setting can turn into null results in a setting with weekly (or longer) decisions, in their case in a test on myopic loss aversion (similar to the Gneezy and Potters, 1997 experimental design). Similarly, Zeisberger et al. (2014) also find that adaptations of asset allocations depend on whether decisions are made directly after each other or with two weeks between them. Gold and Hester (2008) show that gambler's fallacy is reduced if outcome realizations are delayed. Similarly, risk simulations might have weaker effects after some time compared to a short-term lab setting. Generally, portfolio changes in the real world will not occur immediately after the initial investment decision, but much later, and these are the decisions we are interested in, which is why we opted for an experimental setting over multiple trading days mimicking a longer time horizon. Despite this design element which contrasts short-term lab settings, we acknowledge that our time horizon does not fully correspond with a 7-year time horizon we are simulating.

To facilitate participation, our participants received an e-mail reminder every day. We did not request participants to make decisions on weekends (but they could do so if they wanted). Still requesting 14 decisions, we allowed every participant to leave out a maximum of two working days in case they forgot to participate or did not have Internet access. However, in this case the experiment only "paused" for that particular participant and day, so that effectively every participant had to make 14 investment decisions to finish the experiment. If a participant missed more than two working days, this person could no longer participate. To ease access to the software and to facilitate the smooth functioning of the experiment, we made sure that our software worked on desktop computers and also on mobile devices, such as smartphones or tablet computers (e.g., iPad), guaranteeing the same layout on all devices, equally for different web browsers and versions thereof.

The experiment followed a general design and was split into three treatments (see Fig. 1).

This figure illustrates the overall sequence of the between-subject experimental design.

On the first day, which took on average 16 minutes to complete, participants had to answer some demographic questions, a financial literacy quiz (adapted from Lusardi and Mitchell, 2011) and typical risk profiling questions (for an overview and the exact wording see Appendix A). Following the questionnaire, participants were asked to choose an investment endowment (€5000, €25,000 or €100,000) that most closely represented their real financial circumstances (as used by Ehm et al., 2014). As we do not observe any significant differences in investment behavior with regard to the chosen investment level we pool the data for all further analyses.

Participants then received information on a risk-free and a risky asset. The *risk-free* asset was introduced as a fixed-term deposit

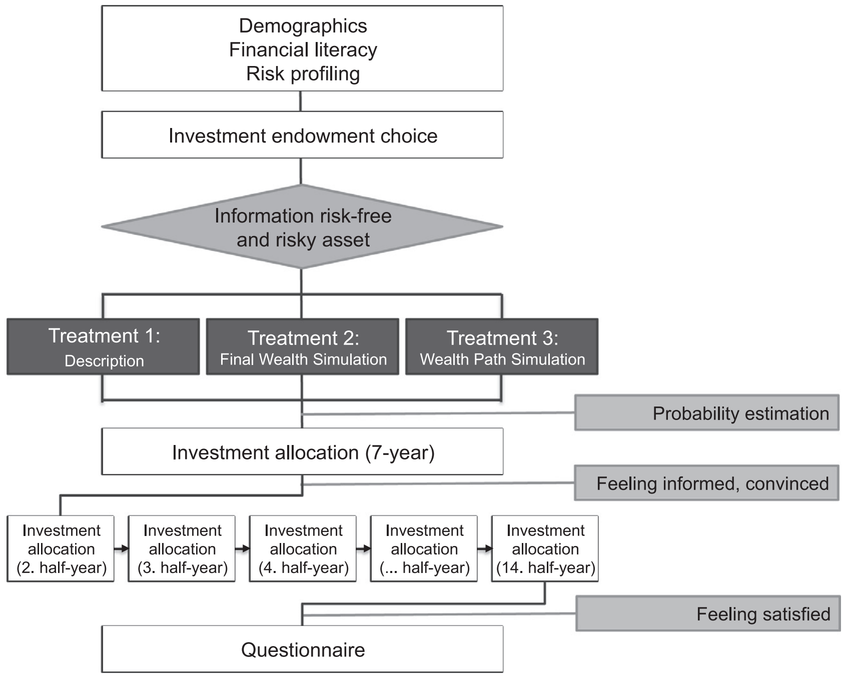


Fig. 1. General experimental setup.

account with a 1.8% rate of return p.a., which corresponds to the average rate of one-year U.S. Treasury Bills for the 10-year period before the experiment (12/31/2002–12/31/2012). To construct the return distribution of the *risky* asset, we calculated empirical half-year returns of the S&P 500 index over 49 years until 12/31/2012 using a one-day rolling time window, which resulted in approximately 12,500 overlapping half-year returns. We hence incorporated six-month autocorrelation structures. We then conducted a bootstrapping technique to construct seven-year returns. Participants were clearly informed that half-year returns are independent from each other. In one of our robustness checks, we constructed the return distribution using seven-year empirical returns (one-day rolling window resulting in 10,571 returns), hence capturing the autocorrelation structure for the whole seven-year investment horizon (see Section 4).

Following Weber et al. (2005), we did not reveal to the participants that a specific index and time period were used as this has been shown to influence risk perception and investment decisions. For the resulting distribution, the average return amounted to 7.0% p.a., after we included a 0.5% reduction to account for the management fee of a passive index fund. The return standard deviation was 16.9% p.a. All information was communicated to the participants in plain language to support their full understanding (see Appendices B and D for instructions).

Participants then had to allocate their chosen investment amount between the two assets for the seven-year period (see Section 2.3 for details). They were able to adjust the allocation to see how the risk-return profile cohered with the chosen allocation between risk-free and risky assets before choosing their final allocation. After making the investment decision, participants had to state how well informed they felt about the two assets and how convinced they were that they had chosen the appropriate allocation for themselves.

At the end of the first day, participants were informed about the details of the 14 decisions to be made, such that the following 14 “experiment days” represent a seven-year investment horizon with half-year intervals (participants were informed at the very beginning about the extended time period nature). Hence, one day represented half a year in the experiment aiming at mimicking better long-term decision-making, while not going for a 7-year experiment. Every day, participants were asked to make an investment decision, which related to the following half year, represented by one day in real time. Participants were also informed that if they changed their allocation, they would incur transaction costs. The costs amounted to 1% of the asset value shifted between the risk-free and the risky asset.

On all following days in the experiment, participants were informed about the past performance of their investment strategy (see Appendix B for a screen shot) before being asked if they want to adjust their asset allocation for the next experimental half year. Participants did not have to change their allocation, and as in reality, the default was to leave the asset allocation unchanged, but it was easy to change it by the use of a slider. Participants had to login every days, even if they wanted to keep their investment level.

2.2. Treatment design

The manner in which the risk-free and risky assets were presented to participants varied between three treatments in a between-subject design. Participants were randomly assigned to one of the three treatments. Appendices B–D provide illustrations of risk presentations in the treatments.

2.2.1. Description treatment (DESCRIP)

In treatment “Description” (DESCRIP) we informed participants in a descriptive and graphical manner about the relevant underlying return distribution of their investment (see Appendix B). We provided participants with the expected yearly return and the expected absolute amount after seven years. In the case of the risky asset, participants were also given the asset’s yearly return standard deviation. The expected yearly return, the expected absolute amount after seven years and the yearly standard deviation were provided in all three treatments in the same way. As most investors find standard deviations difficult to comprehend (Das et al., 2011) and because our empirical return distribution is not perfectly normal, we additionally provided a graphical illustration of the risky asset's return distribution using a detailed bar chart, aiming at enhancing participants' understanding of the index return distribution. Note that Beshears et al. (2017) found that showing any form of (in their case historical) graphical return illustration was the only effective intervention to increase initial risk-taking of individual investors (see Kaufmann et al., 2013 and Bradbury et al., 2015 for similar results). In this aspect, our DESCRIP treatment does provide more visually accessible information in the description and also the distribution treatment in Kaufmann et al. (2013) as we used a detailed histogram. Apart from that, we based our wording on Kaufmann et al. (2013) and information presentation.

2.2.2. Final wealth simulation treatment (FINAL)

The procedure in the treatment "Final Wealth Simulation" (FINAL) is based on the "risk tool" as proposed by Kaufmann et al. (2013). Rather than being presented with a return histogram, participants manually sampled 15 random returns from the return distribution of the risky asset by clicking with the computer mouse (as in experience sampling used for example in Hertwig et al., 2004). These returns were displayed on the screen consecutively. After the 15 draws, participants were free to continue with further manual draws if they wanted, up to 50 draws. No matter how many manual draws participants chose to make, a total of 50 random draws were plotted consecutively to ensure that the simulation did not result in major sampling errors (see also Bradbury et al., 2015).[3] Only after these 50 draws, participants were presented with the final return distribution, displayed graphically in the same diagram on the y-axis (see Appendix C).

2.2.3. Wealth path simulation treatment (PATH)

Both the DESCRIP and FINAL treatments focus on the wealth distribution at the *end* of the investment horizon. However, before reaching it, investors will also experience actual intermediate returns. These are only described in DESCRIP and FINAL in terms of yearly outcomes.[4] Against this background, we included an alternative means of investment risk communication in a third treatment. This treatment additionally visualizes the wealth paths towards achieving the final outcomes. More precisely, the "Wealth Path Simulation" (PATH) treatment adds to FINAL plotting also intermediate outcomes during the investment horizon, i.e., the wealth path to the final outcome, in half-year steps to achieve a smooth curve. The x-axis serves as a time horizon (see Appendix D), otherwise we used the same display as in FINAL.[5] All paths were displayed after each other with increasing speed. In addition, the highest and lowest interim wealth value and largest half-year loss were indicated for every path (see Appendix D). As in FINAL participants saw 50 paths, after which the same final return distribution was displayed.

A possible consequence of PATH is that investors might become less reactive when experiencing intermediate outcomes as the intermediate ups and downs of investors' wealth are more salient and visually accessible than in DESCRIP or FINAL. In other words, while in all treatments annual returns and risk is stated, intermediate returns are visualized in PATH (half-year return visualization to be precise). This can be seen as comparable to the experience-description gap for final wealth, but for PATH the information on intermediate returns is provided in a different way compared to FINAL and DESCRIP. Relatedly, Kaufmann et al. (2013) report that if participants can decide on their asset allocation anew those participants in the risk tool treatment showed a lower reaction to actual losses. However, in their case the second decision followed immediately the first one and was for a hypothetical and new and independent investment decision. So, while not entirely the same, potentially PATH could change the reactivity to previous intermediate outcomes or trading volume.

Another possible effect of PATH is, given insights from the literature on myopic loss aversion, that risk-taking is less influenced compared to FINAL. In lab settings, myopic loss aversion literature has demonstrated that showing aggregated return distributions can increase risk-taking (Benartzi and Thaler, 1999), so displaying short-term intermediate outcomes more saliently as in PATH might have the opposite effect. Shaton (2017) presents related empirical results about differences in the investment horizon presentation. She finds evidence that presenting investors with longer time horizons increases risk-taking. However, we also want to point out that the way we are displaying outcomes in PATH is different from other studies, so the mentioned previous findings can only give some indication and are not a strong predictor.

2.3. Participants and monetary incentives

We recruited a total of 952 participants (588 in main treatments and 364 for robustness checks) with the survey institute "Research Now" in Germany, which enabled us to use a broad and relatively representative sample of German citizens aged 18–65 years, motivated by Harrison and List (2004) and Belot et al. (2015). We prevented apprentices, school children and unemployed persons from taking part in the experiment to focus in potential real-world investors. We specified the gender distribution, two thirds being men as this more closely represents the actual gender distribution of financial decision makers (see, e.g., Barber and Odean, 2001). The mean age of participants was 40 years with a s.d. = 12.2 and 58.5% of the participants had completed secondary school at the minimum. Slightly more than half reported currently owning stocks or equity funds. Appendix F provides details.

We incentivized participants with real monetary rewards, also to reduce attrition in our study as it expands over multiple days. There was a fixed reward component and an incentive-compatible variable component. For the fixed component, each participant received €2.50 for the first day, €0.15 for every following day, and an additional €1.00 for finishing the experiment (making all 14 decisions in the required time). Hence, participants who finished the experiment were sure of receiving €5.60. For the additional, incentive-compatible component, each participant received a monetary endowment of €100.00 at the beginning of the experiment, which corresponded to their chosen experimental investment capital. 20 participants were randomly selected from those who finished and received the final outcome they had achieved over the seven-year investment horizon, based on the initial €100.00 endowment. We did not pay all participants since paying a fraction of participants with higher amounts was shown to have no significant effects on experimental outcomes and stated risk preferences, sometimes higher amounts increase the incentive effect, even for unknown winning probabilities (March et al., 2015). This "between-subjects random incentive system" has been frequently used (Dohmen et al., 2010; Haigh and List, 2005; Cohn et al., 2015; Kirchler et al., 2018). The average amount received was €132.40 with a substantial range from €87.80 to €278.90. All payment details were clearly communicated to all participants at the beginning of the study.

Of course, logging in every working day for a course of three weeks can create substantial non-monetary costs for the partic-

3 In line with previous studies, we did not find any significant effects of sampling errors or recency effects. We tested the average return sampled and the last five draws of the 50 and their possible influence on the initial asset allocation and total turnover for each participant.

4 Note that yearly returns and standard deviations are communicated while trading decisions are for simulated half-year returns in our setting.

5 The software animated each seven-year path by building it up time-wise in half-year steps from the day of the decision to the end of the investment horizon.

Table 1
Differences in allocation to the risky asset and participants' feeling informed and confident.
This table shows mean allocations to the risky asset for all three treatments. It also reports mean responses for feeling informed and confident, measured on six-point Likert scales. To compare allocation decisions two-tailed *t*-test statistics are reported. P-values in bold indicate significance at the 10% level.

Treatment	DESCRIP	FINAL	PATH
Allocation to risky asset	44.2%	51.3%	49.7%
p-values vs. *DESCRIP*		***p*** **= *.003***	***p*** **= *.018***
p-value FINAL vs. *PATH*			$p = .483$
Feeling informed	3.45	3.74	3.84
p-values vs. *DESCRIP*		***p*** **= *.022***	***p*** **= *.002***
p-value FINAL vs. *PATH*			$p = .493$
Feeling confident	3.88	3.99	4.05
p-values vs. *DESCRIP*		$p = .307$	$p = .179$
p-value FINAL vs. *PATH*			$p = .678$
N	211	188	189

Table 2
True and estimated probabilities of the risky asset's return distribution.
This table shows true values for return probabilities, the difference of participants' average estimate to the true values, and average absolute deviations of estimates from true values as well as standard deviations for all participants for three probability events. *P*-values are based on one-sided *t*-tests. *P*-values in bold indicate significance at 10% level. See Appendix G for the precise phrasing of the estimation prompts.

	$P(r < 0\%)$	$P(r < r_f)$	$P(r > 100\%)$
True value [%]	**17.21**	**25.25**	**25.43**
DESCRIP vs. FINAL			
Average deviation of participants' estimates from true value [pp]			
DESCRIP	+15.1	+11.0	−0.5
FINAL	+13.9	+12.0	−2.1
Average absolute deviation of estimates from true value [pp]			
DESCRIP	22.1	22.2	21.6
FINAL	18.6	19.9	17.6
t-test	***p*** **= *.036***	$p = .102$	***p*** **= *.006***
Standard deviation of estimates			
DESCRIP	27.0	26.7	27.4
FINAL	21.3	22.7	22.5
DESCRIP vs. PATH			
Average deviation of participants' estimates from true value [pp]			
DESCRIP	+15.1	+11.0	−0.5
PATH	+10.8	+10.6	−2.0
Average absolute deviation of estimates from true value [pp]			
DESCRIP	22.1	22.2	21.6
PATH	15.2	18.7	17.2
t-test	***p*** **< *.001***	***p*** **= *.024***	***p*** **= *.002***
Standard deviation of estimates			
DESCRIP	27.0	26.7	27.4
PATH	19.2	22.0	21.5

ipants. Still 416 (71%) of the total 588 participants in the three main treatments logged in every day, as requested from them, and are thus considered in the analysis on investment behavior in repeated decisions (see Section 3.2). Participants also only were considered for the variable payment if they finished the whole study and logged in every day.

3. Results

3.1. The effect of information presentation on initial investment decision and understanding of the risk–return tradeoff

In this first section we include all $n = 588$ participants who initially took part in the study. Using only the 416 participants who finalized the study leads to qualitatively very similar results.[6] We find that the way in which risk is communicated influences considerably participants' initial allocations to the risky asset. We find a higher average allocation to the risky asset in the two simulation treatments compared to DESCRIP: FINAL 51.3%, PATH 49.7%, and DESCRIP 44.2%. Both differences to DESCRIP are significant (two-tailed *t*-test, FINAL vs. DESCRIP: $p = .003$, PATH vs. DESCRIP: $p = .018$). We find no significant difference between FINAL and PATH ($p = .483$). Risk simulations hence lead to increased risk-taking in our setting. The results are strongest for FINAL. The last finding is potentially explained by the fact that PATH shows also intermediate outcomes, possibly inducing some level of myopia.

We further asked participants how well informed they felt about the decision problem, and how confident they were of having made the right investment decision based on the information provided. Both questions were measured on a six-point Likert scale (1 = not at all, 6 = very) and were asked directly after participants' first investment decision. Based on previous findings, we expect to find higher values for feeling informed in the FINAL treatment (Kaufmann et al., 2013; Bradbury et al., 2015). The outcome for the PATH treatment, however, might not be as straightforward. This treatment could have been perceived as more complex and participants might have concentrated too much on the wealth paths, distracting them from the interpretation of the risk-return distribution *at the end* of the investment horizon.

We find evidence that investors feel significantly better informed if risk is communicated via FINAL compared to a DESCRIP, as anticipated (see Table 1). The same holds true for the communication via PATH. We do not find a significant difference in feeling confident between treatments. For FINAL and DESCRIP these findings are generally in line with those of Kaufmann et al. (2013) and Bradbury et al. (2015). Between PATH and FINAL we do not find any significant differences in feeling informed or being confident.

These measures might be subjective as they are based on self-reported answers. We also analyze an objective evaluation of individuals' understanding of possible financial consequences via specific probability judgments. Similar as in Kaufmann et al. (2013) participants had to estimate the probability of three possible events for the risky asset after seven years: a) achieving less than the initially invested amount (true probability = 17.2%), b) achieving less than the amount one would have realized by investing in the risk-free asset (true probability = 25.2%), and c) achieving more than twice as much as initially invested (true probability = 25.4%). These estimates provide an objective measure of how well investors understand the risk–return profile of the risky asset.

Generally, we find that participants on average overestimate the probability of a loss,[7] and they nearly correctly asses the upside potential over all treatments on average. Comparing the three probability estimates between treatments, we find the estimates in the FINAL and PATH treatments tend to be improved compared to DESCRIP (see Table 2) with regard to the average absolute deviations from the true values (*p*-values below 0.102 for FINAL compared to DESCRIP and below 0.024 for PATH compared to DESCRIP). It is interesting that PATH leads to similarly accurate results as FINAL, although wealth paths have the potential to distract investors with some irrelevant information with respect to the probability estimation after seven years. The average absolute deviations are even lowest in PATH (combined with lower *p*-values). The differences are significant for $P(r < 0\%)$ with $p = .049$, but not for the other two probability estimates.

[6] We also tested the robustness of our results by eliminating 6% of the participants who clicked fast through the instructions as they might have been inattentive. Our results are almost the same for all analyses. The same holds true for disregarding participants who took particularly long for the instructions.

[7] This is in line with Benartzi and Thaler (1999) and Weber et al. (2005), who also found their subjects to substantially overestimate the probability of a loss in a repeated play of simple gambles or more difficult distributions, which is comparable to an asset allocation decision context with a long-term investment horizon.

Investors are likely to spend more time on the decision in the simulation cases. We find this to be true in our experiment, too, with median completion times 7:20 min in DESCRIP, 10:16 in FINAL and 11:30 in PATH. When we restrict our analysis to participants who take between 6 and 12 minutes, a similar approach as in Kaufmann et al. 2013 ($n = 112$, 102 and 95 in DESCRIP, FINAL and PATH), we find qualitatively similar results with regard to the estimates, so the time spent does not seem to drive these results (two-sided t-test p-value for difference between FINAL and PATH with regard to $P(r < 0\%)$: 0.13).

Overall, these results are interesting as these effects of risk simulations are present even though there are sampling errors in the (active) sampling phase. Hence, even with these sampling errors, estimates tend to be more accurate compared to a description with an easy to read bar chart. We hypothesize that learning through risk simulations is more engaging.

Our effect sizes are smaller compared to some results of previous studies (Kaufmann et al., 2013; Bradbury et al., 2015). However, other research in this area did not include a graphical illustration in the form of an easy-to-read histogram in DESCRIP as in this study, except for the robustness check on "information asymmetry" in Bradbury et al. (2015). Similarly, Beshears et al. (2017) found that simply displaying a graph significantly increased risk-taking. We conclude that an appropriate presentation, such as in the form of a detailed and easy-to-read bar chart, seems to align investors' risk perception with actual risk, but not to the extent that a risk simulation does. Generally, we find evidence that risk simulations increase investors' risk-taking and increases understanding of risk-return tradeoffs. This is true not only for simulating final returns but also for the case in which investors are presented with wealth paths, where the latter also incorporates some irrelevant information with regard to final return estimates.

3.2. The impact of risk simulations on the persistency of investment decisions

With a closer alignment of perceived risk with actual risk, participants chose higher allocations to the risky asset in the FINAL and PATH treatments for their *initial* investment decision. We now turn to analyzing the investment behavior over the whole investment horizon, which means we focus on the $n = 416$ participants who finished the study. To finish the study, a participant needed to login on 14 different days of the experiment and this within the allowed total time (for details see Section 2.1).

3.2.1. Controlling for attrition bias

To ensure that our results are not driven by an attrition bias (early dropouts), we conducted a series of tests. First, we do not find any significant differences in dropout rates (participants who did not finish the experiment over the 14 trading days) between all three treatments: DESCRIP = 28.0%, FINAL = 31.9%, PATH = 28.0% in general. We also do not find any effect with regard to initial investments to the risk asset (for example higher dropouts for lower initial investments). Additionally, we compared the average half-year return of the dropout participants with those who completed the experiment. We do not find any significant difference between the average returns achieved, average returns until dropout amount to 1.83% in the case of the dropout participants and 1.89% for those who completed all 14 investment decisions (two-sided t-test $p = .776$). Furthermore, we do not find evidence that dropout behavior is driven by negative events: The last return that was experienced for the risky asset before a participant dropped out does not differ significantly from the overall average return realized, nor does it differ from the average return of those who remained in the experiment until the end. Overall, we conclude that participant dropouts do not seem to be linked in a systematic way to specific events or treatments. With regard to demographic variables we find that higher financial literacy, being male, and feeling more confident positively affected completing our study. Importantly, however, in none of our factorial regressions which we ran (non-reported, available upon request), we find any treatment dummy to be significant. We also tested, in the same regression, interaction effects of above mentioned variables with treatment dummies, and in none of these regressions we find any significant effects. We conclude that financial literacy, gender and feeling confidence affect dropouts overall, but this does not affect treatments differently, which is important for our analysis.

3.2.2. Risk-taking

Our results show that allocations in subsequent periods, i.e. after the initial allocation, remain significantly higher in the FINAL and PATH treatments up to a certain point in time (see Fig. 2). Participants in DESCRIP on average take seven periods of feedback, which corresponds to 3.5 years in the experimental framework, to reach a peak in their allocation to the risky asset, very close to the point in time at which the average allocation to the risky asset becomes similar between the three treatments. As a result, participants in DESCRIP show a significant increase in their allocation to the risky asset from the beginning of the investment horizon to $t = 3.5$ years by 7.5 percentage points (paired t-test, $p = .00$), whereas participants' investment risk in the FINAL and PATH treatments remain relatively stable on average over the first 3.5 years. This can suggest that participants initially presented with descriptive statements use actual feedback to find out how much risk they are willing to take, whereas participants presented with a form of risk communication that incorporates simulated experience start off with higher allocations and stay with them.

This figure shows the mean allocations with 95% confidence intervals to the risky asset over the full investment horizon for each treatment, per treatment.

What also becomes evident from Fig. 2 is that risk-taking is consistently highest in the FINAL treatment. It seems that presenting investors with final outcomes only vs. presenting them with wealth paths leads to higher allocations to the risky asset. However, the difference in allocation is only significant the first periods. Furthermore, allocations to the risky asset remain relatively stable in the FINAL and PATH treatments when comparing participants' initial investment decisions to all following decisions. This is different for treatment DESCRIP, where we find significant differences between the first allocation and allocations for periods 5 to 13.

Table 3 shows differences in the risk-taking behavior between treatments in more detail. We observe that in the FINAL treatment, allocations remain significantly higher in the first five periods and again towards the end of the investment period compared to DESCRIP. Similarly, but to a lesser extent, we observe risk-taking in the PATH treatment to be higher over the first four periods compared to DESCRIP, after which no further differences are observed. We conclude that risk simulations lead to higher risk-taking (particularly in FINAL), but we do not find risk-taking to be consistently higher for the entire investment period.

For a better understanding of how much investors vary their risk appetite we also measure *individual* absolute deviations from the initial allocation. After half of the investment horizon ($t = 7$ periods), we find some indication that allocation adaptations are smaller in the simulation treatments. Mean absolute adaptations are 17.1% in treatment DESCRIP vs. 15.5% in treatment FINAL ($p = .18$ one-sided t-test) and vs. 14.7% in treatment PATH ($p = .10$). We observe slightly larger differences if we focus on the first rounds, i.e. for $t = 5$ and $t = 6$, i.e. after investors have received feedback for 4 or 5 times. Hence, at least at the beginning, individual adaptations of asset allocation seem to be slightly lower when people are presented with wealth paths, while the effect size is

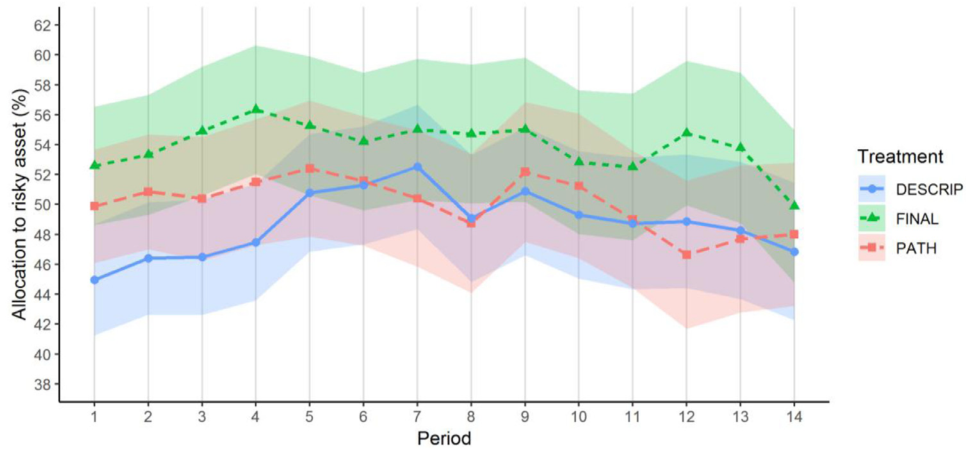


Fig. 2. Mean allocation to the risky asset with 95% confidence intervals, per treatment.

Table 3
Differences in allocation to the risky asset.
This table shows mean allocations to the risky asset between treatments. To compare allocation decisions, this table presents approximative two-sample Fisher-Pitman permutation tests. *P*-values in bold indicate significance.

Period	1	2	3	4	5	6	7	8	9	10	11	12	13	14
DESCRIP	45.0	46.4	46.5	47.5	50.8	51.3	52.5	49.1	50.9	49.3	48.7	48.9	48.3	46.9
FINAL	52.6	53.3	54.9	56.3	55.3	54.2	55.0	54.7	55.0	52.8	52.5	54.8	53.8	49.9
PATH	49.9	50.8	50.4	51.5	52.4	51.6	50.4	48.7	52.2	51.2	49.0	46.6	47.7	48.0
Description (DESCRIP) vs. Final Wealth Simulation (FINAL)														
$p =$	**.007**	**.015**	**.005**	**.003**	.148	.343	.437	**.080**	.211	.282	.264	**.081**	.112	.390
Description (DESCRIP) vs. Wealth Path Simulation (PATH)														
$p =$	**.071**	.106	.178	.171	.595	.922	.501	.913	.689	.554	.936	.512	.869	.734

small. This indicates that seems to be a small effect on the stability of individual risk appetite in early periods if simulations are used (further analyses below).

Regarding the second half of the investment horizon we find strong similarities between all treatments. Generally, risk-taking decreases towards the end of the investment horizon, and this to a similar extent between all the treatments. There are different possible explanations for this. One is that participants potentially aim to lock in gains (see, for example, Benzoni et al., 2007; Cocco et al., 2005; Gourinchas and Parker 2002). Similarly, a shortened investment horizon might cause investors to be more risk averse, to some extent consistent with myopic loss aversion (Benartzi and Thaler 1999). As participants gain wealth over time, this kind of behavior is also consistent with increasing relative risk aversion (IRRA), which holds for example with mean-variance preferences. This decrease in risk-taking in the second half of the investment period is interesting as it is in contrast to most short-term laboratory experimental results with university students.[8] In these, investments often go up as participants become more familiar with the environment and at the very end might reduce their investments.

3.2.3. Satisfaction

Does higher risk-taking in the simulation treatments reduce investor satisfaction due to a higher variance in returns—and hence possibly higher losses? Our results indicate that the answer is "no". Despite differences in risk-taking, we do not find differences in participants' satisfaction with their investment strategy. We measured satisfaction right after the end of the experiment by asking participants how satisfied they were with their choice of strategy, i.e. their allocation between the risk-free and the risky asset, using a six-point Likert scale (1 = not at all, 6 = very). The results of our regression analysis are reported in Table 4. We find that satisfaction is primarily driven by the final wealth and the return in

Table 4
Self-reported Satisfaction with Investment Result after 14 Periods.
This table shows regression results for self-reported satisfaction (Likert scale 1–6) explained by treatment and other variables. For income a median split was used (income_dummy).

Dependent variable: satisfaction with investment strategy (Likert scale 1–6)	
Treatment FINAL	−0.171
	(0.154)
Treatment PATH	−0.044
	(0.153)
Final Wealth (Initial = 100)	**0.010*****
	(0.002)
Return in period 1	**1.249****
	(0.590)
Return in period 14	0.640
	(0.622)
Male dummy	0.137
	(0.146)
Age	0.002
	(0.006)
Risk Attitude	0.040
	(0.054)
Income dummy	−0.061
	(0.130)
Constant	**2.705*****
	(0.359)
Observations	416
R^2	0.114
Adjusted R^2	0.094
Residual Std. Error	1.257 (df = 406)
F Statistic	5.597*** (df = 9; 406)

[8] See, for example, Gneezy and Potters (1997), Thaler et al. (1997), Langer and Weber (2008), Fellner and Sutter (2009) and Charness and Gneezy (2011). Mostly, a reduction is limited to the last period, as in Langer and Weber (2008) in the form of a "final-round effect".

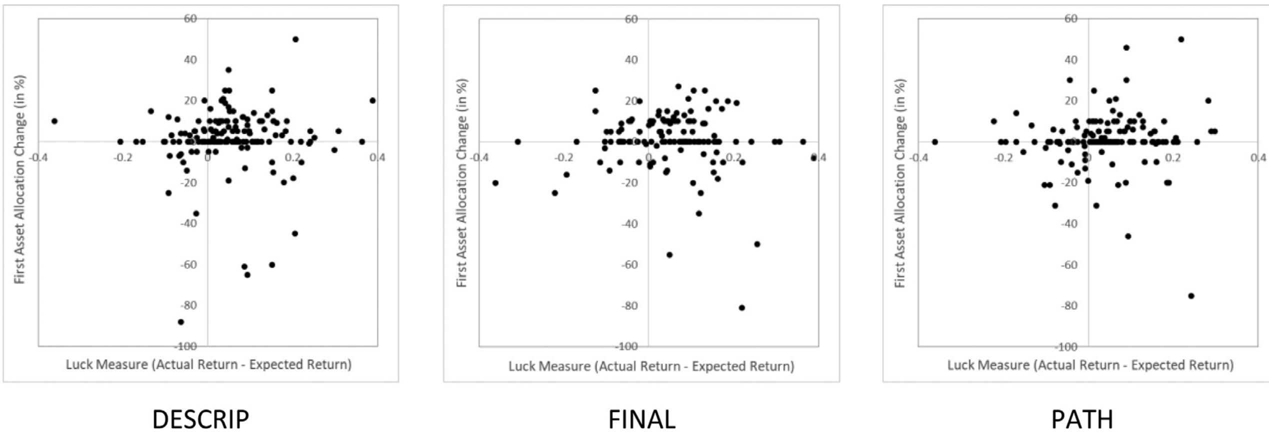


Fig. 4. Asset allocation changes in the first period.

period 1. While coefficients of our treatment dummies are slightly negative, differences are not significant. Hence, we do not find evidence that investors regret increased risk-taking behavior in hindsight when previously presented with a simulation.

3.2.4. Overall trading activity

Another element of investment behavior is trading activity. Kaufmann et al. (2013) found that the reactivity to losses is smaller in a risk simulation. In their case a hypothetical new decision had to be made and with information via a simulation participants did not reduce their investment in stocks as much compared to description after losses. The idea behind our PATH simulation was to make the intermediate returns more salient by visual display compared to DESCRIP and FINAL, so that PATH could potentially lead to lower trading activity. We find the lowest average and median volumes for PATH (median: 90% total turnover in relation to initial endowment). This is followed by DESCRIP (97%). The highest turnover is observed for FINAL (107.5%). However, none of the differences are statistically significant (PATH vs. DESCRIP: one-tailed t-test $p = .281$). We conclude that in our setting overall trading activity seems to be independent of the way risks are presented at the beginning of the investment horizon.

3.2.5. Reactivity to previous outcomes and intermediate trading activity

A possible reason for not finding any differences in trading activity between different treatments could stem from investors reacting more to actual outcomes than to simulated ones, particularly in later periods when the relative amount of actual experience is higher. Hence, consequences might only be present shortly after conducting the risk simulation. In a further analysis, we therefore compare only the first allocation change in each treatment, i.e. the change from the initial allocation. If risk simulations have an effect on how investors react to previous gains and losses, then this effect is likely to be highest for the first possible allocation change. While PATH has the lowest adaptation, we do not find significant differences between the three treatments. The average absolute change in allocation: DESCRIP: 9.1pp, FINAL: 8.8pp; PATH: 7.4pp. The difference between PATH and DESCRIP is not statistically significant (one-sided t-test p-value: 0.18). This finding also holds true if we split the analysis and analyze reactions to gains and losses separately.

This figure depicts the asset allocation changes in the first period for all participants in each of the three treatments separately in relation to a luck measure, i.e. actual return – expected return, similar to the analysis in Kaufmann et al. (2013). Positive numbers for asset allocation changes indicate increases from the initial allocation in the risky asset after the first period.

For more insights, Fig. 4 shows how investors react to the gain or loss (relative to the expected return which depends on their asset allocation) in the three treatments. We find that in each treatment, a majority of participants slightly increases risk-taking after the first period. However, there are some differences with respect to how many participants leave their allocation *un*changed after the first period, and it is highest in PATH and lowest in DESCRIP. DESRIP: 32.9%, FINAL: 33.6%, and highest in PATH with 41.2%.

As a more formal test, we analyze the likelihood that participants changed their initial allocation using Fisher's exact test in the early periods. We present results in Table 5. Here, we find evidence that changes in allocation from the initial allocation to the risky asset are less likely in treatment PATH compared to DESCRIP.

Generally speaking, we find some evidence that a PATH risk simulation lowers reactions to previous gains and losses in our setting if regarding the fraction of participants (not) changing their allocation. Our findings for FINAL seem less strong than the ones presented by Kaufmann et al. (2013). In their study, investor reactions to previous outcomes seem lower for FINAL compared to DESCRIP. We hypothesize that this difference could stem from a number of reasons: Lower reactions could be a result of the short-term lab setting Kaufmann et al. (2013) used. Alternatively, they only tested hypothetical allocation changes (non-incentivized) as opposed to our setting with real decisions. Third, in our setting we deal with intermediate rather than final outcomes. As an alternative reason, any of the results is due to statistical errors.

Given the insights from Beshears et al. (2017) about the importance of experimental time lags in investment decisions, we believe that risk simulations might only influence risk perception in the short-run (i.e. probability estimates of the return distribution) but not risk preferences in the longer-run, at least with regard to reactions to previous gains and losses. A possible explanation for not finding significant differences between FINAL and PATH could be that showing wealth paths might present "too much" information and investors have difficulty focusing on the important parts (however, at least they do not feel less informed as evidenced by results of see Table 1). Further tests with different experimental designs such as testing our treatments in a short-term lab setting, or testing investors' understanding of the return distribution at later stages could add more insights.

Table 5
Fraction of subjects who left their allocation unchanged in comparison to first period (dummy variable).
This table shows the fraction of subjects for each treatment that did not change the allocation in comparison to the first period of investing (initial allocation) and one-sided Fisher's exact tests to test for differences.

	Allocation change in Period 2	Allocation change in Period 2 or 3	Allocation change in Period 2, 3 or 4	Allocation change in Period 2, 3, 4 or 5
DESCRIP	0.329	0.191	0.132	0.092
FINAL	0.336	0.250	0.188	0.148
PATH	0.412	0.265	0.213	0.184
p-value Fisher's exact test DESCRIP vs. FINAL	0.501	0.147	0.132	0.102
p-value Fisher's exact test DESCRIP vs. PATH	**.091**	**.087**	**.046**	**.018**
Observations	416	416	416	416

Table 6
Mixed effects regression.
This table reports mixed effects regression with the allocation to the risky asset as the dependent variable. Six-point Likert scale responses are treated as ordinal independent variables (risk attitude, feeling informed, feeling confident). Financial literacy is the score of financial literacy test. Standard errors are in brackets. * indicates significance at the 10% level, ** significance at the 5% level, and *** significance at the 1% level.

	Dependent variable: allocation to risky asset				
	(1)	(2)	(3)	(4)	(5)
Period	0.139	−0.235	−0.235	−0.235	0.139
	(0.154)	(0.172)	(0.172)	(0.172)	(0.154)
Treatment FINAL	**7.391*** **	**5.757* **	4.807	**5.486* **	**7.119*** **
	(2.822)	**(3.058)**	(3.036)	**(2.947)**	**(2.702)**
Treatment PATH	4.280	2.456	1.055	2.262	4.086
	(2.777)	(3.009)	(3.003)	(2.931)	(2.692)
Early periods (1–3) dummy		**−5.155*** **	**−5.155*** **	**−5.155*** **	
		(1.050)	**(1.050)**	**(1.050)**	
Feeling informed			**2.569** **	1.068	1.068
			(1.018)	(1.004)	(1.004)
Feeling confident			1.221	0.419	0.419
			(1.028)	(1.000)	(1.000)
Financial Literacy				**1.354* **	**1.354* **
				(0.794)	**(0.794)**
Gender (1 = male)				2.650	2.650
				(2.334)	(2.334)
Age				0.036	0.036
				(0.087)	(0.087)
University dummy				0.473	0.473
				(2.087)	(2.087)
High income dummy				0.488	0.488
				(2.107)	(2.107)
Risk aversion				**4.541*** **	**4.541*** **
				(0.879)	**(0.879)**
Period x Treatment FINAL	−0.289	−0.133	−0.133	−0.133	−0.289
	(0.227)	(0.254)	(0.254)	(0.254)	(0.227)
Period x Treatment PATH	**−0.394* **	−0.219	−0.219	−0.219	**−0.394* **
	(0.224)	(0.250)	(0.250)	(0.250)	**(0.224)**
Treatment FINAL x Early periods		2.154	2.154	2.154	
		(1.553)	(1.553)	(1.553)	
Treatment PATH x Early periods		2.406	2.406	2.406	
		(1.528)	(1.528)	(1.528)	
Constant	47.663***	51.571***	37.849***	23.903***	19.995***
	(1.908)	(2.068)	(3.927)	(5.660)	(5.603)
Observations	5,824	5,824	5,824	5,824	5,824
Log Likelihood	−24,650.780	−24,629.250	−24,618.450	−24,593.220	−24,614.750
Akaike Inf. Crit.	49,321.560	49,284.490	49,266.900	49,228.440	49,265.510
Bayesian Inf. Crit.	49,388.260	49,371.200	49,366.950	49,368.500	49,385.560

4. Robustness of results

4.1. Control variables

Our results remain stable when controlling for self-reported financial risk attitude, financial literacy, experience, being invested in real life and further demographic variables (see Appendix A for detailed information on the control variables) in a mixed effects regression as reported in Table 6.

We find self-reported financial risk attitude to be a strong predictor for the allocation to the risky asset, confirming the findings of a broad spectrum of literature in the financial decision-making context (see, e.g., Schooley and Worden 1996; Nosić and Weber 2010; Dohmen et al., 2011; Halko et al., 2012; Ehm et al., 2014). We also find financial literacy, measured by a quiz, to be a marginally significant indicator of risk-taking behavior, which supports evidence from existing literature that there is a link between financial literacy or sophistication and financial decision making (Campbell 2006; Agarwal et al., 2018). On the other hand, having a university degree, age, gender and income have hardly any predictive power in relation to participants' actual risk-taking. Models 2, 3 and 4 separate also early periods (1 to 3) and show the robustness of the results if these are analyzed individually.

Table 7
Allocation to the risky asset and predictability per treatment.
This table reports Tobit regression coefficient estimates with the allocation to the risky asset as the dependent variable. Standard errors are in brackets. Financial risk attitude treated as interval variable from Likert answers, financial literacy as score from objective test (0 to 6). Controls at the bottom of the table account for experience (Likert answer as interval), being invested (dummy), having a University degree (dummy), age, male and income (median dummy). They are indicated to be included in the regression analysis by yes but are hardly ever significant.
* indicates significance at the 10% level, ** 5% level, and *** 1% level (except constant).

Dependent variable: Allocation to the risky asset DESCRIP					
Investment period	1	4	7	10	14
Financial risk attitude	**7.86*** **	**6.47*** **	**9.02*** **	**8.66*** **	**9.65*** **
	(2.88)	(2.00)	(2.23)	(2.29)	(2.57)
Financial literacy	**2.12***	1.83	2.18	2.50	0.51
	(1.28)	(1.36)	(1.51)	(1.56)	(1.75)
Controls	Yes	Yes	Yes	Yes	Yes
Constant	14.05	15.54	26.88	17.86	0.73
	(10.41)	(11.09)	(12.30)	(12.68)	(14.23)
Observations	136	136	136	136	136
Log-likelihood	−598.30	−604.12	−604.04	−679.71	−669.56
Dependent variable: Allocation to the risky asset **FINAL**					
Investment period	1	4	7	10	14
Financial risk attitude	**3.72** **	**5.29** **	1.65	0.26	3.39
	(1.79)	(2.11)	(2.37)	(2.46)	(2.72)
Financial literacy	2.22	2.90	−1.80	0.65	2.40
	(1.56)	(1.83)	(2.06)	(2.14)	(2.36)
Controls	Yes	Yes	Yes	Yes	Yes
Constant	22.48	24.47	29.08	6.55	32.98
	(10.90)	(12.80)	(14.35)	(14.98)	(16.55)
Observations	128	128	128	128	128
Log-likelihood	−552.06	−542.63	−559.49	−567.62	−556.68
Dependent variable: Allocation to the risky asset **PATH**					
Investment period	1	4	7	10	14
Financial risk attitude	**3.91***	**5.30** **	2.97	1.94	2.65
	(2.11)	(2.34)	(2.69)	(2.92)	(2.75)
Financial literacy	2.72	1.60	3.40	1.15	−0.74
	(1.85)	(2.05)	(2.36)	(2.56)	(2.41)
Controls	Yes	Yes	Yes	Yes	Yes
Constant	34.82	26.80	22.03	24.62	37.72
	(11.91)	(13.17)	(15.20)	(16.46)	(15.49)
Observations	162	162	162	162	162
Log-likelihood	−598.30	−604.12	−604.04	−598.78	−609.82

We ran Tobit regressions for each treatment separately to see whether risk attitude and financial literacy have varying strengths in predictive power depending on the way in which risk is communicated (see Table 7). We find self-reported financial risk attitude to have considerable predictive power in the DESCRIP treatment for the allocations to the risky asset over all periods. This lnk is weaker for the FINAL and PATH treatments, in which the predictive power of risk attitude can be observed in the first few asset allocation decisions only. This is an interesting result since it could mean that classically used bank risk questionnaires might not be predictive for revealed investor risk preferences by actual investments. At least in our study, once investors have gained some simulated experience in a risk simulation, self-reported risk preferences have lower predictive power for the actual investment strategy.

4.2. Characteristics of return distribution

To ensure that our general conclusions are not limited to the specific return distribution of the risky asset, we repeated our experiment in the PATH treatment using a different return distribution. To construct the underlying return distribution, we took the actual empirical return distribution of the S&P 500 without smoothing it via a bootstrapping approach (see Fig. 5 for a comparison of the baseline and the control return distributions). We thus capture seven-year autocorrelation structures. Both distributions have the same expected yearly return and yearly standard deviation.

This figure compares the two different return distributions. The black bars represent the empirical distribution of the S&P 500 index since inception. The grey bars represent the return distribution of the S&P 500 index smoothed through a bootstrapping technique to resemble more closely a log-normal distribution.

We recruited 180 new participants with the same underlying socio-demographic characteristics, of whom a total of 125 completed all the investment decisions, which implies the same completion rate as in our three baseline treatments. We do not find any systematic difference in the risk-taking behavior to our PATH baseline results. Detailed analyses are provided in Appendix H. For the first investment decision, participants on average allocate 51.4% (s.d. = 23.9%) to the risky asset compared to 49.7% in the PATH treatment for the baseline return distribution, which is not significantly different from each other (two-tailed t-test, $p = .47$). We also find the same trading patterns as in our baseline study. In particular, participants in this alternative PATH treatment did not show significant changes in allocations over time on average. These results provide evidence that our findings are robust with regard to the return distribution, at least in the tested PATH treatment. Regarding trading behavior, we can confirm our previous results and find similar trading volumes for average and median levels. The median total trading turnover is 83% of the initial investment amount and does not differ from the original PATH treatment (two-tailed t-test $p = .52$).

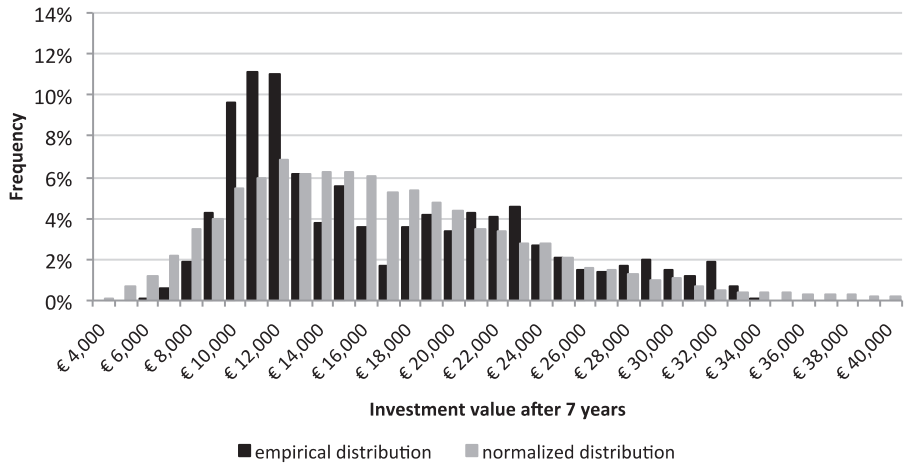


Fig. 5. Comparison of the empirical and normalized return distributions.

4.3. Transaction costs

It could be that transaction costs prevented our participants from making adaptations to their asset allocations. In a further robustness check with another 184 subjects (of whom 136 completed the study with 14 investment decisions), we re-ran the PATH treatment with the original return distribution, but we refrained from imposing transaction costs, i.e. shifting between both assets was free of charge. In the instructions, we therefore left out any information about transaction costs. We find that the initial allocation to the risky asset (which could not have been affected by transaction costs) equals 51.2% (s.d. = 22.5%) in the new PATH treatment, and it is not significantly different from the baseline treatment of 49.7% (two-tailed t-test, $p = .51$). Total trading turnover over all periods in relation to the initial endowment equals 76% (median value), which is not significantly different from the original PATH treatment (new PATH vs. original PATH: two-tailed t-test $p = .51$). Hence, we do not find that transaction costs significantly reduced allocation shifts. More detailed results are provided in Appendix H. Taking these results together with our original PATH results for the initial decision generally confirms the robustness of these (e.g. risk-taking level, feeling informed and confident, return distribution estimates).

5. Conclusion

Risk simulations based on experience sampling were shown to influence investment behavior for one-off investment decisions (Kaufmann et al., 2013; Bradbury et al., 2015). When possible outcomes are presented via a risk simulation, investors' risk perception is closer aligned with actual risk compared to less experiential ways of risk communication. Related findings in similar domains were documented, for example, in Cason and Samek (2015) who show that experimental subjects produce lower bubbles in laboratory asset markets when they have experienced market prices before, or Lusardi et al. (2014) who demonstrate effects on financial literacy and confidence in financial decisions when information is provided in a more experiential way via videos or tools. What is less clear is whether risk simulations have a *persistent* effect on investor behavior or if the effects are of short-term nature and are superimposed by actual investment experience. These questions are not only of academic interest but have also practical implications as they can provide new insights on the usefulness of such simulation tools in financial advice and planning.

Our research aims at answering two main questions. First, how persistent are the effects of experience sampling risk simulations on trading behavior? Second, what are the effects of a wealth path simulation ("PATH") compared to a simulation of final outcomes ("FINAL")? Generally, our results confirm previous findings for the initial investment decision. Participants in our study state feeling better informed about the investment decision, and they have a less biased understanding of return distribution if informed via a risk simulation tool. As a result, they invest more in risky assets initially.

As for our first research question, the higher risk-taking behavior when using risk simulations is only persistent to some degree. Investors who are initially informed in a descriptive manner (as opposed to a risk simulation) increase on average their investment levels over the first subsequent investment periods. One possible interpretation is that they might require actual investment experience to decide how much risk they are willing to take. However, the effect is not strong, and we do not find that risk simulations lead to persistently increased risk-taking behavior over the entire investment horizon. So, after having received some actual feedback, investment behavior seems to become independent of the way investors learned about the investment problem initially. In other words, risk simulations seem to influence mainly the initial asset allocation of investors at the time they are informed via the simulation. This is an interesting finding as Kaufmann et al. (2013) suggest that risk simulations could reduce investors' reactivity to losses. Our results do not necessarily contradict theirs as they tested it for an entirely new and hypothetical decision and not for intermediate ones as we did. However, our results could also be interpreted as being in line with Beshears et al. (2017) who find that robust results in a short-term setting might not hold when longer experimental time differences are used.

As for the differences in the type of risk simulation with regards to risk-taking, our second research question, our results show almost the same risk-taking between FINAL and PATH (slightly higher for FINAL). Also feeling informed, confident and return distribution estimates do not fundamentally differ between PATH and FINAL. That means that although PATH displays additional information, (potential) investors do not seem overwhelmed by this amount of additional information. With regard to trading behavior, we observe some evidence of reduced reactions to previous outcomes in the PATH treatment, but we do not find such results for FINAL, each comparing to DESCRIP. Hence, a PATH simulation could potentially slightly reduce reactions to intermediate gains and losses which a simulation with only final returns might not be able to do (at least when measured in adaptations from the initial allocation). However, we do not find significant differences in overall trading volume. It might be necessary to use risk simulations on a more permanent basis, not just for the initial investment decision, to observe differences here. Further treatments that we did not test such as a short-term lab setting can provide further insights. Alternatively, it can also be useful to ask investors at later stages about their understanding of the return distribution to better understand where the (non-)differences in investment behavior over time come from. Another interesting aspect to explore is to gain more knowledge on investors' short-term (half-year) return expectations. This could bring further insights into the underlying mechanisms of trading behavior, particularly in PATH.

As an additional result, we find that presenting investors with easy-to-read return histograms as a descriptive form of risk communication seems to partially close the experience-description gap. The relatively large differences in risk-taking in Kaufmann et al. (2013) or Bradbury et al. (2015) can to be partially explained by the fact that a purely verbal presentation of risk seems non-intuitive and difficult to comprehend for investors. Well-explained return histograms already make a difference to purely verbal investment communication, similar as in Beshears et al. (2017). However, importantly, we still find significant differences in initial risk-taking, risk perception and feeling informed between using well-described histograms and risk simulations. Lastly, we find evidence that risk simulations close the gap between investment levels of financially more and less literate investors. In our study, risk simulations were particularly influential with regard to risk-taking for less financially literate individuals.

Our results can also be interesting for business practice, in particular for financial advice and planning, e.g. in the case of robo-advisors. With robo-advisors there is no interaction with a human financial advisor, and risk communication is even more essential to provide a competitive service. Our study should allow insights on how to design risk profilers and learning tools that could be useful for financial institutions. Our findings can also be relevant to human advisors who are supported by interactive computer-based risk-profiling methods.

Acknowledgements

We are thankful to Charlotte Borsboom, James Choi, Helga Fehr-Duda, Eric Johnson, Christine Laudenbach, David Laibson, Miles Kimball, Markus Strucks, Pascal St-Amour, Martin Weber and participants at the following conferences and seminars: American Economic Association Chicago 2017, Behavioural Finance Working Group Conference in London 2016, Consumer Financial Protection Bureau (CFPB) research conference in Washington DC in 2015, Behavioral Decision Research in Management (BDRM) Conference in London 2014 and University of Innsbruck Finance Seminar 2014 for their valuable comments and suggestions. We also thank Achiel Fenneman for supporting us with many analyses, Marius Wolfensberger for his programming support and Behavioural Finance Solutions GmbH for financial support. This research was gratefully supported by the Swiss National Science Foundation (100018-149934) and by the University of Zurich via URPP Finreg. Previous titles of this paper have been: "How Risk Simulations Improve Long-Term Investment Decisions" and "Do Risk Simulations Lead to Persistently Better Investment Decisions?"

Appendix A. Demographics, financial literacy and risk profiling questionnaire

Variable	Response options/description
Financial risk attitude	How would you classify your willingness to take financial risks? Strongly avoid risks ○ ○ ○ ○ ○ ○ ○ Willing to take risks
Financial literacy*	Sum of six financial literacy questions (highest score = 6, lowest score = 0)
Experience	Self-reported investment experience with risky assets (such as stocks, derivatives, alternative investments, etc.) above 3 years equals 1, otherwise 0
Invested	Self-reported portion of current financial wealth invested in stocks or stock funds greater than 0% equals 1, otherwise 0
University	University degree equals 1, otherwise 0
Age	Age of the participant
Male	If gender male, then equals 1, otherwise 0
Income	Self-reported income based on predefined bandwidths (1 ≤ €20,000, 2 = €20,000–30,000, 3 = €30,000–40,000, 4 = €40,000–50,000, 5 = €50.000–60.000, 6 = €60.000–70.000, 7 ≥ €70.000)
* Questions:	Suppose you had €1,000 in a savings account and the interest rate was 2% per year. After 5 years, how much do you think you would have in the account if you left the money to grow? Imagine that the interest rate on your savings account was 2% per year and inflation was 3% per year. After 1 year, how much would you be able to buy with the money in this account? Buying a single company's stock usually provides a safer return than a stock mutual fund. True or false? With which asset class would you have achieved the highest return over the last 50 years? Which asset class shows the highest value fluctuation over the past 50 years? Imagine that you are invested in an asset that achieves a 10% return per year, how many years would it take to double your investment?

Appendix B. Screen shot of description (DESCRIP) treatment

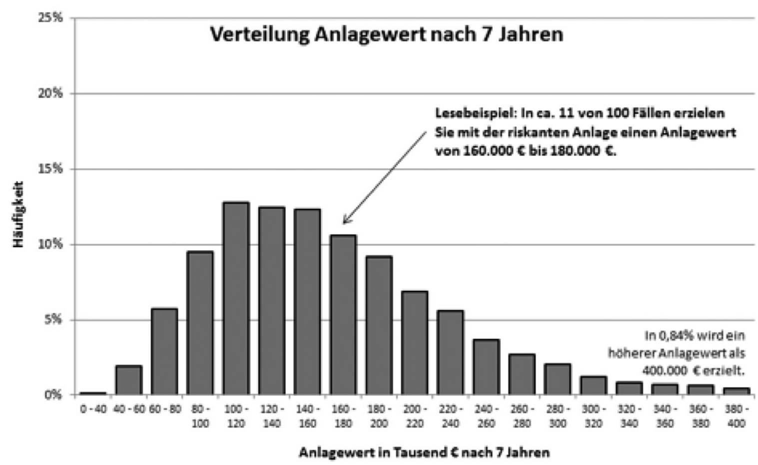
Fig. B1.

Riskante Anlage

Bei der riskanten Anlage handelt es sich um einen Aktienindex. Dieser bildet die Wertentwicklung von Aktien ab. Die **erwartete Rendite beträgt 7,0% pro Jahr**, d.h. **im Durchschnitt** können Sie eine Rendite von 7,0% erwarten. Die tatsächlich erzielte Rendite ist jedoch **nicht sicher**, sie kann deutlich höher aber auch deutlich tiefer ausfallen. Die jährliche **Standardabweichung**, d.h. die als „normal" betrachtete Abweichung von diesen 7,0%, beträgt **16,9%**.

Falls Sie Ihren zuvor gewählten Investitionsbetrag von 100.000 € ausschließlich in die riskante Anlage investieren, können Sie nach 7 Jahren einen Anlagewert von 164.100 € erwarten. Wie erwähnt ist dieser Wert jedoch nicht sicher, er könnte deutlich höher aber auch deutlich tiefer ausfallen.

Folgende Grafik zeigt Ihnen die Verteilung Ihres Anlagewertes am Ende der 7 Jahre, wenn Sie ausschließlich in die riskante Anlage investieren. Beispielsweise würden Sie in ca. 11 von 100 Fällen einen Anlageendwert zwischen 160.000 € und 180.000 € erzielen.



Beispielsweise ist es wahrscheinlich (**in 50 von 100 Fällen**), dass nach 7 Jahren daraus:

Zwischen **113.000 €** und **201.100 €** werden.

Mit hoher Wahrscheinlichkeit (**in 90 von 100 Fällen**) werden daraus:

Zwischen **72.200 €** und **296.500 €**.

Mit sehr hoher Wahrscheinlichkeit (**in 98 von 100 Fällen**) werden daraus:

Zwischen **52.000 €** und **394.200 €**.

Fig. B.1. Description (DESCRIP) treatment.
English translation:
[Title:] Risky asset
[Text:] The risky asset is a stock index. It represents the growth of stocks. The expected return is 7% per year, which means you can expect a return of 7% on average per year. However, the actual return is unknown; it can turn out considerably higher but also considerably lower. The yearly standard deviation, i.e., the deviation from the 7% considered normal, is 16.9%.
If you invest the full investment amount of €100,000 previously chosen in the risky asset, you could expect a final amount of €164,100 after 7 years. However, as mentioned, this amount is not certain; it could turn out considerably higher but also considerably lower.
The following chart shows you the distribution of your initial wealth after 7 years if you had invested the full investment amount in the risky asset. For example, you would achieve a final wealth of between €160,000 and €180,000 in approximately 11 out of 100 cases.
[Chart:] Wealth distribution after 7 years
Reading example: In approximately 11 out of 100 cases you can achieve a final wealth of between €160,000 and €180,000 with the risky asset.
In 0.84% of cases, a final wealth greater than €400,000 is achieved.
[x-axis:] Final wealth in thousand € after 7 years
[y-axis:] Frequency
[Text:] It is likely, for example, that (**in 50 out of 100 cases**) after 7 years it will be:
between **€113,000** and **€201,000**.
It is very likely that (in 90 out of 100 cases) it will be:
between **€72,200** and **€296,500**.
It is highly likely that (in 98 out of 100 cases) it will be:
between **€52,000** and **€394,200**.

Appendix C. Screen shot of final wealth simulation (FINAL) treatment

Fig. C1.

Fig. C.1. Final Wealth Simulation (FINAL) treatment.
English translation of pop-up on screen (not shown in Fig. C.1):
|[Title:] Risky asset
[Text:] See description of the risky asset for the description treatment (DESCRIP)
You will be presented with possible, final values, for the case that you invest the full initial amount over 7 years in the risky asset. The final values are drawn randomly from the distribution described above. Click “draw” to view a further value. You need to draw at least 15 times. Afterwards, further values are drawn and automatically displayed one after another. In total, 50 random values will be displayed.

Appendix D. Screen shot of wealth path simulation (PATH) treatment

Fig. D1.

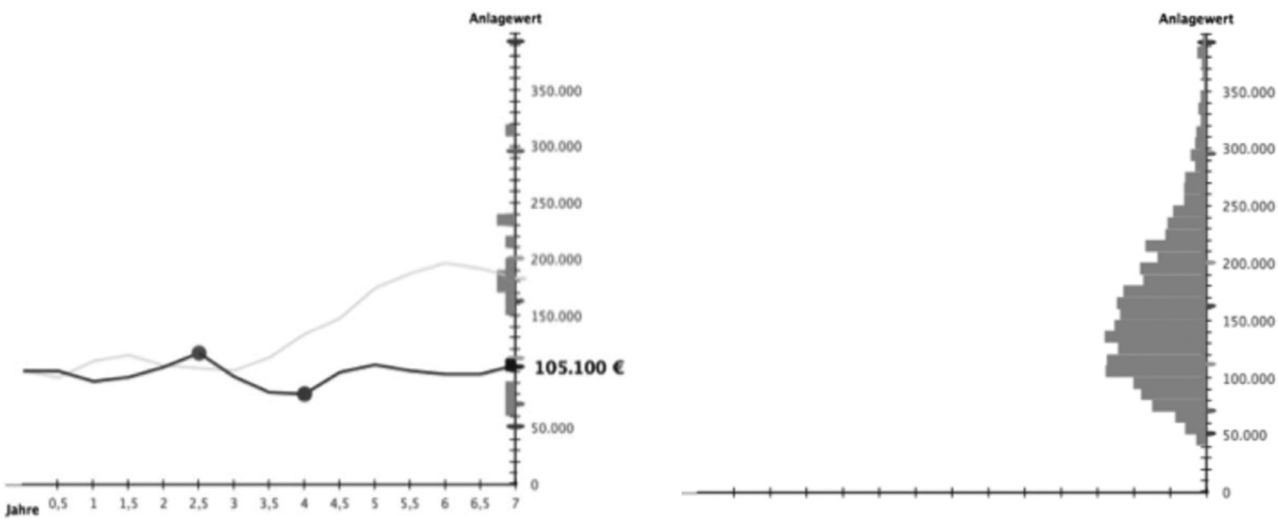


Fig. D.1. Wealth Path Simulation (PATH) treatment.
The left screen shot illustrates the build-up of the return distribution, as it was experienced by participants, and the right screen shot shows the final distribution.
English translation of pop-up on screen (not shown in Fig. D.1):
[Title:] Risky asset
[Text:] *Same as for the Final Wealth Simulation (FINAL) treatment.*

Appendix E. Screen performance overview

Fig. E1.

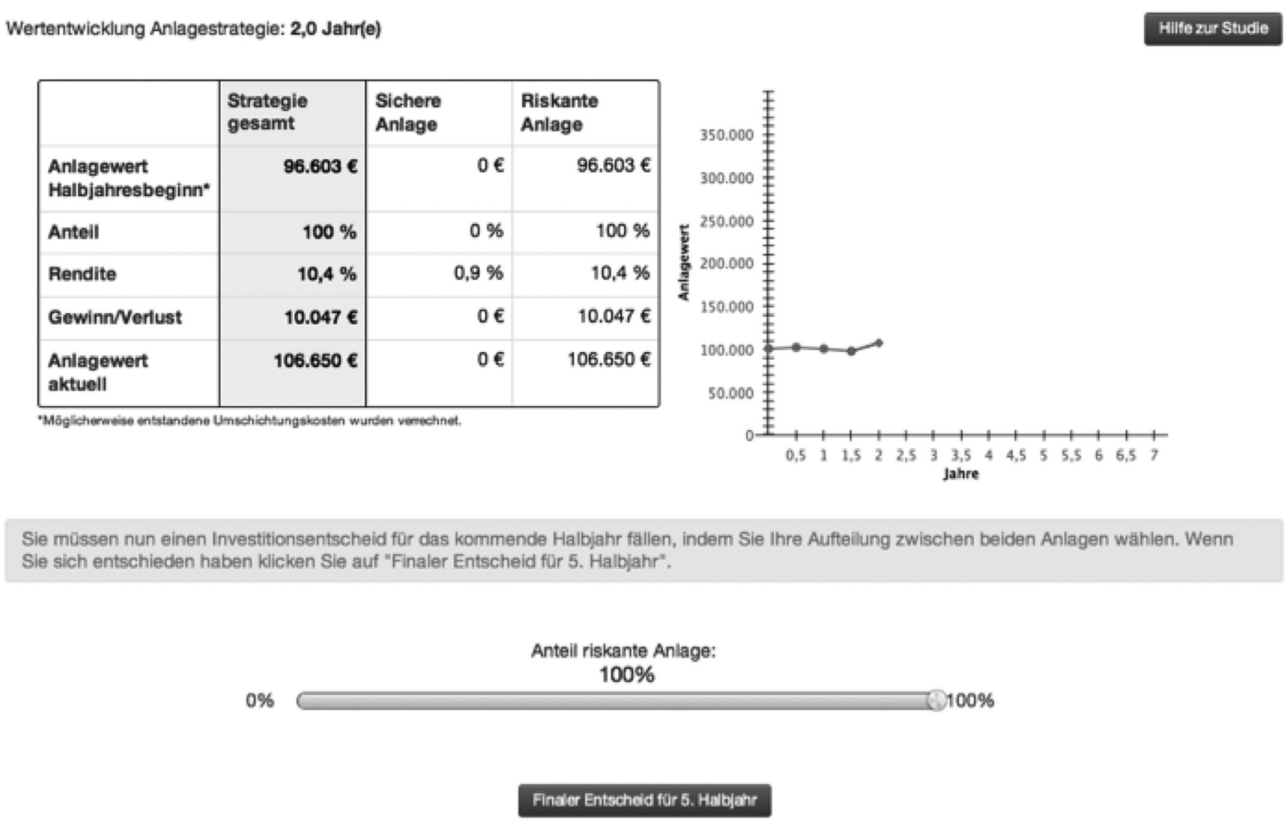


Fig. E.1. Performance overview of portfolio strategy.
English translation:
[Title:] Performance portfolio strategy
[Table:] First row: Overall strategy/risk-free asset/risky asset
First column: Investment value at beginning of half year/allocation/return/gain–loss/current investment value
[Slider:] Allocation to risky asset .

Appendix F. Summary statistics

Table F.1.

Table F.1
Overview of socio-economic characteristics of the participants.
This table shows summary statistics of socio-economic variables, the education level, and income distribution of participants within each treatment. Income ranges were derived based on the average gross income of German households (€3,871 per month in 2011 (Destatis, 2013), extending the range by two classes above the average).

Treatment	Description "DESCRIP"	Final Wealth Simulation "FINAL"	Wealth Path Simulation "PATH"
Age (average)	39.7	39.8	39.6
Min	18	18	18
Max	66	65	66
Gender (male)	65%	65%	71%
Stock/equity fund owners	57%	58%	52%
Education			
Still in school	0.5%	0.5%	0.4%
Secondary modern school	9.5%	9.0%	8.4%
Junior high school	30.8%	30.3%	32.7%
Secondary school	23.2%	22.3%	20.1%
University	35.1%	35.6%	37.0%
PhD	0.5%	1.6%	1.2%
N/A	0.5%	0.5%	0.3%
Income			
< €20,000	24.2%	22.3%	21.6%
€20,000–30,000	19.4%	16.5%	16.7%
€30,000–40,000	15.6%	15.4%	15.5%
€40,000–50,000	27.5%	29.8%	30.5%
€50,000–60,000	9.5%	9.0%	9.1%
€60,000–70,000	3.8%	6.9%	6.5%
> €70,000	0.0%	0.0%	0.0%
N	211	188	189

Appendix G. Probability estimations

Consider an investment in the aforementioned risky asset. Please estimate in how many out of 100 cases your investment amount will …

- Come out below your initial investment amount after 7 years.
- Come out below the amount you would have achieved by investing in the risk-free asset (*€113.300*) after 7 years.
- Come out above *€200.000* after 7 years.

Appendix H. Results for Robustness Checks (Different Return Distribution and No Transaction Costs)

Note: PATH (7Y) indicates the treatment for which the return distribution was calculated over 7 years (i.e. including 7y autocorrelation) and returns do not show zero serial correlation. PATH (FREE) indicates the treatment in which participants were not charged any transaction costs.

Fig. H1. and Table H.1. Table H.2.

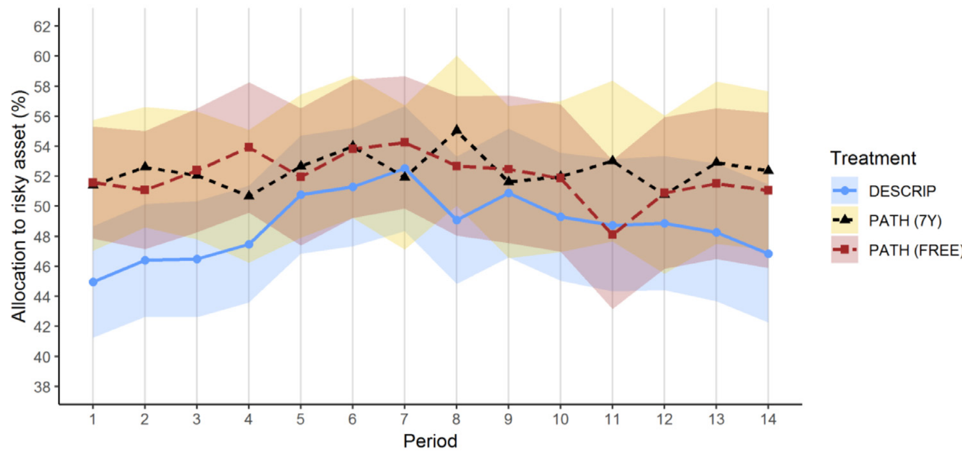


Fig. H.1. Mean allocation to the risky asset with 95% confidence intervals, per treatment.
This figure shows the mean allocations with 95% confidence intervals to the risky asset over the full investment horizon for each treatment.

Table H.1
Differences in allocation to the risky asset and participants' feeling informed and confident.
This table shows mean allocations to the risky asset for all three treatments. It also reports mean responses for feeling informed and confident, measured on six-point Likert scales. To compare allocation decisions two-tailed *t*-test statistics are reported. P-values in bold indicate significance at the 10% level.

Treatment	DESCRIP	PATH (7Y)	PATH (FREE)
Allocation to risky asset	44.2%	51.4%	51.2%
p-values vs. *DESCRIP*		***p*** **= .003**	***p*** **= .003**
Feeling informed	3.45	3.83	3.79
p-values vs. *DESCRIP*		***p*** **= .003**	***p*** **= .010**
Feeling confident	3.88	3.86	4.04
p-values vs. *DESCRIP*		*p* = .901	*p* = .194
N	211	180	184

Table H.2
True and estimated probabilities of the risky asset's return distribution.
This table shows true values for return probabilities, the difference of participants' average estimate to the true values, and average absolute deviations of estimates from true values as well as standard deviations for all participants for three probability events. *P*-values are reported for one-tailed *t*-tests. We use absolute deviations from estimated to true values to avoid opposite signs canceling each other out. *P*-values in bold indicate significance at 10% level. See Appendix G for the precise phrasing of the estimation prompts.

	$P(r < 0\%)$	$P(r < r_f)$	$P(r > 100\%)$
True value [%]	**17.21**	**25.25**	**25.43**
DESCRIP vs. PATH (7Y)			
Average deviation of participants' estimates from true value [pp]			
DESCRIP	+15.1	+11.0	−0.5
PATH (7Y)	+10.2	+6.8	−4.9
Average absolute deviation of estimates to true value			
DESCRIP	22.1	22.2	21.6
PATH (7Y)	16.4	19.4	16.6
t-test	***p*** **= .002**	***p*** **= .050**	***p*** **= .000**
Standard deviation of estimates			
DESCRIP	27.0	26.7	27.4
PATH (7Y)	21.5	23.3	19.8
DECRIP vs. PATH (FREE)			
Average deviation of participants' estimates from true value			
DESCRIP	+15.1	+11.0	−0.5
PATH (FREE)	+9.8	+10.8	−0.7
Average absolute deviation of estimates			
DESCRIP	22.1	22.2	21.6
PATH (FREE)	15.0	18.5	17.6
t- test	***P*** **= .000**	***p*** **= .020**	***p*** **= .007**
Standard deviation of estimates			
DESCRIP	27.0	26.7	27.4
PATH (FREE)	19.2	22.4	22.9

References

Agarwal, S., Pan, J., Qian, W., 2018. Age of Decision: Pension Savings Withdrawal and Consumption and Debt Response Working Paper http://ssrn.com/abstract=2487658.

Barber, B.M., Odean, T., 2001. Boys will be boys: gender, overconfidence, and common stock investment. Quart. J. Econ. 116, 261–292.

Benartzi, S., Thaler, R.H., 1995. Myopic loss aversion and the equity premium puzzle. Quart. J. Econ. 110, 73–92.

Benartzi, S., Thaler, R.H., 1999. Risk aversion or myopia? Choices in repeated gambles and retirement investments. Manag. Sci. 45, 364–381.

Benzoni, L., Collin-Dufresne, P., Goldstein, R.S., 2007. Portfolio choice over the life-cycle when the stock and labor markets are cointegrated. J. Finance 62, 2123–2167.

Beshears, J., Choi, J., Laibson, D., Madrian, B., 2017. Does aggregated returns disclosure increase portfolio risk-taking. Rev. Financ. Stud. 30, 1971–2005.

Belot, M., Duch, R., Miller, L., 2015. A comprehensive comparison of students and non-students in classic experimental games. J. Econ. Behav. Organ. 113, 26–33.

Bradbury, M.A.S., Hens, T., Zeisberger, S., 2015. Improving investment decisions with simulated experience. Rev. Finance 19, 1019–1052.

Campbell, J.Y., 2006. Household finance. J. Finance 61, 1553–1604.

Cason, T.N., Samek, A., 2015. Learning through passive participation in asset market bubbles. J. Econ. Sci. Assoc. 1 (2), 170–181.

Charness, G., Gneezy, U., 2011. Strong evidence for gender differences in risk taking. J. Econ. Behav. Organ. 83, 50–58.

Cohn, A., Engelmann, J., Fehr, E., Maréchal, M., 2015. Evidence for countercyclical risk aversion: an experiment with financial professionals. Am. Econ. Rev. 105, 860–885.

Cocco, J., Gomes, F.J., Pascal Maenhout, J., 2005. Consumption and portfolio choice over the life-cycle. Rev. Financ. Stud. 18, 491–533.

Das, S., Markowitz, H., Scheid, J., Statman, M., 2011. Portfolios for investors who want to reach their goals while staying on the mean-variance efficient frontier. J. Wealth Manag. 14, 25–31.

Destatis, 2013. Statistical Year Book: Germany and International 2013 Technical Report.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., 2010. Are risk aversion and impatience related to cognitive ability? Am. Econ. Rev. 100, 1238–1260.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., Wagner, G.G., 2011. Individual risk attitudes: measurement, determinants, and behavioral consequences. J. Eur. Econ. Assoc. 9, 522–550.

Ehm, C., Kaufmann, C., Weber, M., 2014. Volatility inadaptability: investors care about risk, but cannot cope with volatility. Rev. Finance 18, 1387–1423.

Fellner, G., Sutter, M., 2009. Causes, consequences, and cures of myopic loss aversion – an experimental investigation. Econ. J. 119, 900–916.

Gourinchas, P.-O., Parker, J., 2002. Consumption over the life cycle. Econometrica 70, 47–91.

Gneezy, U., Potters, J., 1997. An experiment on risk-taking and evaluation periods. Quart. J. Econ. 112, 631–645.

Gold, E., Hester, G., 2008. The gambler's fallacy and the coin's memory. In: Krueger, J.I. (Ed.), Rationality and Social Responsibility: Essays in Honor of Robyn Mason Dawes. Taylor & Francis Group, New Work, pp. 21–46.

Haigh, M., List, J., 2005. Do professional traders exhibit myopic loss aversion? An experimental analysis. J. Finance 60, 523–534.

Halko, M-L., Kaustia, M., Alanko, E., 2012. The gender effect in risky asset holdings. J. Econ. Behav. Organ. 83, 66–81.

Harrison, G., List, J., 2014. Field experiment. J. Econ. Lit. 42, 1009–1055.

Hershfield, H.E., Goldstein, D.G., Sharpe, W.F., Fox, J., Yeykelis, L., Carstensen, L.L., Bailenson, J.N., 2011. Increasing saving behavior through age-progressed renderings of the future self. J. Mark. Res. 48, 23–37.

Hertwig, R., Barron, G., Weber, E.U., Erev, I., 2004. Decisions from experience and the effect of rare events in risky choice. Psychol. Sci. 15, 534–539.

Hirshleifer, D., Teoh, S.H., 2003. Limited attention, information disclosure, and financial reporting. J. Account. Econ. 36, 337–386.

Imas, A., 2016. The realization effect: risk-taking after realized versus paper losses. Am. Econ. Rev. 106, 2086–2109.

Kaufmann, C., Weber, M., Haisley, E., 2013. The role of experience sampling and graphical displays on one's investment risk appetite. Manage. Sci. 59, 323–340.

Kirchler, M., Lindner, F., Weitzel, U., 2018. Rankings and risk-taking in the finance industry. J. Finance 73, 2271–2302.

Langer, T., Weber, M., 2008. Does commitment or feedback influence myopic loss aversion? An experimental analysis. J. Econ. Behav. Organ. 67, 810–819.

Lusardi, A., Mitchell, O.S., 2011. Financial literacy around the world: an overview. J. Pension Econ. Finance 10, 497–508.

Lusardi, A., Samek, A., Kapteyn, A., Glinert, L., Hung, A., Heinberg, A., 2014. Visual tools and narratives: new ways to improve financial literacy. J. Pension Econ. Finance 10, 1–27.

March, C., Ziegelmeyer, A., Greiner, B., Cyranek, R., 2015. Monetary Incentives in Large-Scale Experiments: A Case Study of Risk Aversion Working Paper.

Nosić, A., Weber, M., 2010. How risky do I invest: the role of risk attitudes, risk perceptions and overconfidence. Decis. Anal. 7, 282–301.

Odean, T., 1999. Do investors trade too much. Am. Econ. Rev. 89, 1279–1298.

Schooley, D.K., Worden, D.D., 1996. Risk aversion measure: comparing attitudes and asset allocation. Financ. Serv. Rev. 5, 87–99.

Shaton, M., 2017. The Display of Information and Household Investment Behavior. Federal Reserve Board of Governors Working Paper https://ssrn.com/abstract=2662694.

Shefrin, H., Statman, M., 1985. The disposition to sell winners too early and ride losers too long: the theory and evidence. J. Finance 40, 777–790.

Shefrin, H., Statman, M., 1994. Behavioral capital asset pricing theory. J. Financ. Quant. Anal. 29, 323–349.

Thaler, R.H., Benartzi, S., 2004. Save more tomorrow: using behavioral economics to increase employee saving. J. Polit. Econ. 112, 164–187.

Thaler, R., Tversky, A., Kahneman, D., Schwarz, A., 1997. The effect of myopia and loss aversion on risk-taking: an experimental test. Quart. J. Econ. 112, 647–661.

Weber, E.U., Johnson, E.J., Milch, K.F., Chang, H., Brodscholl, J.C., Goldstein, D.G., 2007. Asymmetric discounting in intertemporal choice: a query theory account. Psychol. Sci. 18, 516–523.

Weber, E.U., Siebenmorgen, N., Weber, M., 2005. Communicating asset risk: how name recognition and the format of historic volatility information affect risk perception and investment decision. Risk Anal. 25, 597–609.

Zeisberger, S., Langer, T., Weber, M., 2014. Do Changes in reporting Frequency Really Influence Investors' Risk-Taking Behavior? Myopic Loss Aversion Revisited Working Paper. http://ssrn.com/abstract=1786360.