

Breuer, Wolfgang; Renerken, Thomas; Salzmann, Astrid Juliane

Article — Published Version

## Measuring risk-taking and patience in financial decision making

Review of Financial Economics

**Provided in Cooperation with:**

John Wiley & Sons

*Suggested Citation:* Breuer, Wolfgang; Renerken, Thomas; Salzmann, Astrid Juliane (2021) : Measuring risk-taking and patience in financial decision making, Review of Financial Economics, ISSN 1873-5924, Wiley, Hoboken, NJ, Vol. 40, Iss. 1, pp. 97-114, <https://doi.org/10.1002/rfe.1142>

This Version is available at:

<https://hdl.handle.net/10419/284740>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<http://creativecommons.org/licenses/by-nc-nd/4.0/>

# Measuring risk-taking and patience in financial decision making

Wolfgang Breuer<sup>1</sup>  | Thomas Rennerken<sup>1</sup>  | Astrid Juliane Salzmann<sup>2</sup> 

<sup>1</sup>RWTH Aachen University, Aachen, Germany

<sup>2</sup>LEM-CNRS (UMR 9221), IESEG School of Management, Lille, France

## Correspondence

Thomas Rennerken, RWTH Aachen University, Aachen, Germany.

Email: thomas.rennerken@bfw.rwth-aachen.de

## Abstract

We empirically compared the consistency among different kinds of measures for risk-taking and patience in a survey and an experiment. We evaluated how these variables relate to financial decisions, using a novel set of easy-to-apply survey questions as proxy. The main finding is that our results based on the novel survey questions describing financial behavior and on lifelike financial behavior in the laboratory are very similar. Therefore, relatively low-cost elicitation measures may be used to forecast decision making, which, in turn, may be described by our novel set of survey questions.

## KEYWORDS

decision making, preferences, risk attitude, time preferences

## JEL CLASSIFICATION

C81; D12; D81; G11

## 1 | INTRODUCTION

The literature is rich in evidence of substantial heterogeneity in individual investment decisions, such as the choice of portfolio structure and savings behavior. The way individuals build their financial portfolios has been an increasingly relevant issue for economists and policy-makers alike. Economic theory suggests that household portfolio choice depends, among other things, on an individual's level of patience and willingness to take risk. These factors are also understood as the root of many economic phenomena (Donkers & van Soest, 1999).

The study on financial decision making is challenging because individual preferences are difficult to measure. The literature has proposed distinct frameworks, such as the sophisticated models of the well-established prospect theory or the quasi-hyperbolic discounting model, to elicit relevant parameters. Collecting data on portfolio selection may also be complicated. Real-life data on portfolio choices can be difficult and expensive to obtain. Survey data are usually easier and cheaper to collect, but findings based on it cannot always be generalized. Incentivized laboratory experiments offer an interesting trade-off between generalizability and cost efficiency.

There is little consensus in the literature on how the level of patience and willingness to take risk should be measured to ensure consistent and reliable results. Financial advisors and private bankers still rely on rather rough estimates based on non-sophisticated measuring methods. Previous studies investigating different elicitation methods have revealed puzzling outcomes and large inconsistencies. A low correlation among different measures is a recurrent finding (see, e.g., Crosetto & Filippin, 2016; Szrek et al. 2012; Pedroni et al., 2017). Although self-reported measures for risk-taking show slightly higher consistency

in the relative ordering, the empirical rank order induced by lottery-based measures is almost identical to that expected from random ranking (Frey et al., 2017).

Previous studies have compared the suitability of different measures to elicit an individual's willingness to take risk and evaluated their usefulness for predicting financial behavior. Wärneryd (1996) asked questions on risk-taking, including self-descriptive measures and hypothetical lottery choices, but found that only the former are highly significant for explaining the riskiness of investment portfolios. Similarly, Hyll and Irrek (2015) concluded that simple measures of willingness to take risk appear to be more powerful predictors of portfolio allocation than complex lottery questions. Erner et al. (2013) examined whether the parameters of cumulative prospect theory are successful in predicting an individual's preference for different structured financial products; they found low predictive power of the elicited parameters. Other studies have examined the validity of different measures for predicting risk-taking behavior in the laboratory. Lönnqvist et al. (2015) compared incentivized lottery choice tasks with non-incentivized measures and found that only the self-assessment measures relate to risk-taking. Vieider et al. (2015) showed that there are correlations between different measures of risk attitudes, pointing to the existence of one underlying actual risk preference. Galizzi et al. (2016) revealed similar findings, indicating correlations between different measures, but also mixed evidence concerning external validity in the context of financial behavior. Bachmann et al. (2017) used experimental data and demonstrated that self-assessed measures of risk tolerance are not suitable predictors of risk-taking.

The comparison between real-life, experimental, and survey data for financial decision making, such as investment behavior, has not been sufficiently addressed in the previous literature. There is evidence that data collection method affects data quality (Duffy et al., 2005; Jäckle et al., 2010; de Leeuw, 2005; Newman et al., 2021). The literature on the effect of incentivization on data quality is not coherent (Davern et al., 2003; van den Brakel et al., 2006; Stecklov et al., 2018). When comparing survey and experimental data, individuals answering survey questions seemed to unknowingly misreport or knowingly misrepresent data to a large extent (Goldman et al., 1989; Cozby and Bates, 2019). The literature concerning the comparison of experimental and real-life data is rather scarce. In financial decision making, Li (2020) found that choices in the laboratory can predict, to a certain extent, real-life decisions.

Except for Bachmann et al. (2017), simple self-assessed measures seemed reliable when predicting actual risk-taking and patience. Given these conditions, easy-to-understand survey questions related to hypothetical investment behavior can be assumed to reflect behavior as, for example, in the laboratory. This is a potential finding that would make economic data collection rather easier than before.

First, this study contributes to the existing literature by proposing a novel and easy-to-apply survey measurement method to proxy actual risk-taking and patience. To this end, we investigated survey data of more than 700 individuals from nine different countries, thus ensuring sufficient variation in the dataset. Then, we used an experimental setting to create a more lifelike environment with incentivized investment choices. This approach aimed to check whether our novel questions for describing investment behavior in our survey can, in fact, proxy conclusions drawn from actual decision making in an experimental context. Second, this study expands the current literature by comparing the predictive power of different approaches on a large-scale basis while addressing both the complex problem of determining willingness to take risk and the easier issue of measuring the level of patience. This may validate the prevailing notion that elaborated decision-theoretical approaches only deliver a poor and inconsistent description of individuals' preferences. Along these lines, this study also investigated whether these poor descriptions exist because of an error-in-variables problem caused by mismeasurement.

The rest of our paper is organized as follows: Section 2 describes different measures to elicit willingness to take risk and the level of patience. Section 3 provides an overview of the dataset. Section 4 contains the empirical analysis of our survey data. Section 5 discusses the laboratory experiments. Section 6 addresses the error-in-variables problem. Section 7 is the conclusion.

## 2 | METHODS TO ASSESS WILLINGNESS TO TAKE RISK AND LEVEL OF PATIENCE

When comparing methods to predict risk-taking and patience, one can distinguish between sophisticated and simple measures. Sophisticated measures are widely used in decision making theory and economic literature, whereas simple methods, such as self-assessments, are used by practitioners in the financial services sector. An example of the latter can be seen in the recent rise in algorithm-based investment advice, so-called "robo-advice," which offers automated investment recommendations based on less sophisticated investor exploration questions. For the purpose of this study, it is necessary to understand the difference between these two approaches to assess willingness to take risk and level of patience.

## 2.1 | Simple self-assessment

Haliassos and Bertaut (1995) used a straightforward behavioral measure of risk tolerance. Respondents evaluated the following questions:

Which of the following statements comes closest to the amount of financial risk you are willing to take when you save or make investments?

- I would take substantial financial risks expecting to earn substantial returns.
- I would take above average financial risks expecting to earn above average financial returns.
- I would take average financial risks expecting to earn average financial returns.
- I am not willing to take any financial risks.

Haliassos and Bertaut (1995) used a similar formulation for self-descriptive patience:

Which of the following statements most applies to you if you needed to invest your money for a long period of time?

- I would tie up money for a long period of time to earn substantial returns.
- I would tie up money for an intermediate period of time to earn above average returns.
- I would tie up money for a short period of time to earn average returns.
- I am not willing to tie up money at all.

A range of studies has employed similar measures and demonstrated their usefulness in predicting actual risk-taking behavior in financial matters (Barasinska et al., 2012; Dohmen et al., 2011; Donkers & van Soest, 1999). However, with respect to an individual's level of patience, we are not aware of similar works examining the predictive power of these self-assessments.

## 2.2 | More sophisticated approaches

A more refined measurement of individual risk attitude is based on prospect theory. Its value function was introduced by Kahneman and Tversky (1979) and later extended to the cumulative prospect theory of Tversky and Kahneman (1992). Cumulative prospect theory is widely considered the most successful descriptive theory for decision making under risk. According to Kahneman and Tversky (1979), individuals are risk-seeking for gains with low probabilities and for losses with high probabilities, and risk averse for gains with high probabilities and for losses with low probabilities. In addition, Tversky and Kahneman (1992) established the concept of rank-dependent probability weighting.

Based on these assumptions, a decision-maker chooses the outcome that maximizes his or her rank-dependent utility  $V = \sum \pi_i \cdot v(x_i)$  with  $v$  describing a value function and  $\pi$  specifying his or her decision weights. These are computed as

$$\pi_i = w\left(\sum_{j=i}^n p_j\right) - w\left(\sum_{j=i+1}^n p_j\right) \quad i \in (1, \dots, n) \quad (1)$$

where  $w$  is a probability weighting function. The value function

$$v(x) := \begin{cases} x^{\alpha^+}, & x \geq 0 \\ -\lambda \cdot (-x)^{\alpha^-}, & x < 0 \end{cases} \quad (2)$$

classifies the outcomes relative to a reference point. Therefore, positive values of  $x$  denote gains and negative ones losses relative to a reference point. The value function is concave for gains (outcomes above the reference point) and convex for losses (outcomes below the reference point). The parameter  $\lambda$  causes the value function to be steeper for losses than for gains, thereby modeling the loss aversion phenomenon. The  $\alpha^+$  and  $\alpha^-$  portray different degrees of (relative) risk aversion for gains and losses. Tversky and Kahneman's (1992) weighting function of probabilities is

$$w_\gamma(p) = \frac{p^\gamma}{(p^\gamma + (1-p)^\gamma)^{1/\gamma}} \quad (3)$$

and transforms the probability  $p$  of an outcome into a subjective probability weight using parameter  $\gamma$ . It models a pattern that underweights high probabilities and overweights small probabilities. Some researchers have distinguished between  $\gamma^+$  and  $\gamma^-$ ; however, we followed the reasoning of Tversky and Kahneman (1992) and assumed  $\gamma^+ = \gamma^- = \gamma$ .

We asked a series of questions in the form of a hypothetical binary lottery choice, similar to Barsky et al. (1997), to derive the relevant parameters. We applied the Gamble method by Currim and Sarin (1989) and the certainty equivalent method. These methods were designed for our purposes, as well as relatively easy to comprehend and intuitively appealing. Winning probabilities and stakes vary to cover a range of outcomes with a minimum number of questions, because the length of the survey was limited. For the binary prospects that we used, the rank dependence can be ignored.

Owing to space restrictions, we could not reproduce a detailed survey in this paper. The English version of the questionnaire is presented in Supplementary File Section VI. Our raw data are available on request. Generally speaking, the questions used to compute the relevant parameters in this case were stated in a formal, quantitative manner. For example:

Please state the amount of Z for which you are indifferent between both lotteries.

Lottery A

50 % chance to gain 20 \$

50 % chance to gain 200 \$

Lottery B

50 % chance to gain Z \$

50 % chance to gain nothing

Z should be \_\_\_\_\_ \$, such that lottery A is as attractive as lottery B.

These questions required more effort to understand the presented scenario compared with the self-assessment questions mentioned above. Because of the different nature of the question types and popularity of lottery-based questions, we decided to include them in our analysis. In this context, these questions could be considered a more sophisticated version of the simple self-assessment but with the same goal: to derive individual risk-taking parameters.

We were aware of the problems associated with calculating loss aversion parameters when using value functions as specified in (2). However, since it is the consistency with financial decision making among different measures and not the exact values of these measures we wanted to study and compare, we still used the equations presented above. Although the absolute value of  $\lambda$  might depend, for example, on scaling (see, e.g., Wakker, 2010, for further information), demonstrating a higher loss aversion in our survey leads to a higher value for  $\lambda$  and, therefore, does not (*ceteris paribus*) influence statistic significance in our regressions.

We used the theory of the quasi-hyperbolic discounting model to derive individual patience-related measures. Although many other elicitation mechanisms exist, this method is among the most commonly used and relatively easy to implement. Individuals tend to prefer smaller but earlier rewards instead of larger but later rewards. Moreover, economists generally agree that people discount the near future with a higher discount rate per period than the far future (Thaler, 1981). Such a behavior can be described by a utility function of the following form:

$$U(x_0, x_1, \dots, x_T) = u(x_0) + \sum_{t=1}^T \beta \cdot \delta^t \cdot u(x_t). \quad (4)$$

In (4), for  $\beta = 1$  and  $0 < \delta < 1$ , the discount factor would be  $\delta^t$  for any period  $t$ , and thus, a corresponding constant discount rate  $1/\delta - 1$  per period would result for all  $t$ . However, for  $0 < \beta < 1$ , the discount factor for the first period would only amount to  $\beta \cdot \delta$ , implying that the discount rate for the first period of  $1/(\beta \cdot \delta) - 1$  is higher than  $1/\delta - 1$ . Therefore, when  $0 < \beta < 1$  and  $0 < \delta < 1$ , people appear to be more patient in the long run and less patient in the immediate future. The parameter  $\beta$  is called the present bias, and  $\delta$  refers to the long-term discount factor. For  $\beta = 1$ , there is no present bias problem, and we simply arrive at the neoclassical intertemporal utility function, as introduced by Samuelson (1937). We again positioned questions for a hypothetical set of circumstances in which participants could take some money now or wait until later. We then asked how much it would require making waiting as attractive as receiving the money now. The amount of money was kept constant, but the questions varied with respect to the length of time to wait, for example:

Please consider the following alternatives:

A: a payment of 200 \$ now

B: a payment of X \$ in one month from now

X has to be at least \_\_\_\_\_ \$, such that B is as attractive as A.

- A: a payment of 200 \$ now  
 B: a payment of X \$ in one year from now  
 X has to be at least \_\_\_\_\_ \$, such that B is as attractive as A.

As can be seen, the questions were not as easy to understand as those in the simple self-assessment. These questions demanded some additional mental effort, making them a more sophisticated counterpart of the patience self-assessment.

We elicited the implicitly revealed parameters for both risk-taking and patience using a data-fitting approach, which is presented in detail in Supplementary File Section I.

An important concern arises from the fact that, as mentioned before, such parameter-related measurements are complicated and intellectually sophisticated. This particularly holds true for the determination of risk preferences based on (2) and (3). Respondents might misreport their preferences either knowingly or unknowingly. It is, therefore, probable that alternative survey instruments that are easier to understand can deliver similar (or even more) reliable information on individual preferences (Dohmen et al., 2011). Hence, we also refrained from employing more advanced versions of the models introduced above.

Previous studies have examined the consistency of the derived parameters with an individual's willingness to take financial risk but found no reliable evidence. Frijns et al. (2008) demonstrated that risk aversion measured through lottery choice is a determining factor of portfolio selection. However, Dimmock and Kouwenberg (2010) found that loss aversion and portfolio allocation decisions are not related. Erner et al. (2013) found a low predictive power of prospect theory for financial risk-taking. To the best of our knowledge, literature on the role of patience in portfolio allocation decisions is non-existent.

### 3 | DATASET

Our analysis was based on microeconomic data drawn from a specially designed survey to provide the most complete data. We elicited different measures of risk-taking and patience to obtain a rich set of information that is rarely available otherwise. In most cases, the survey was conducted at the beginning of a regular lecture at a university under the monitoring of a local field interviewer. The survey participants had mixed academic backgrounds; however, most of them were business students but at different stages of their studies. Participation was not mandatory. There was no explicit time limit for completing the survey. Individuals were able to provide information on the results if they contacted the interviewer after participation. We surveyed individuals in nine different countries, namely China, Germany, India, Iran, Italy, Singapore, South Korea, Spain, and Russia. These countries were not part of a specific selection process but chosen based on the authors' existing academic connections. This approach yielded a high response rate and, as the results of our analysis show, a broad cultural basis for our study, thus avoiding any kind of country-specific bias. We collected information on more than 700 university students across nine countries from 2010 to 2013. Because of the nature of the internationally conducted surveys, the procedure took some time. Nevertheless, the data collection per country did not take longer than three months each. We attempted to keep this time span as narrow as possible. As we did not compare absolute values of risk and time measures across countries, we did not think that the stability of preferences would be an issue in our research.

The large sample size allowed us to attain a comprehensive scope of individual variability. The survey was translated into local languages. Monetary payoffs were adjusted according to each country's purchasing power to assure comparability of results across countries. We collected information on the monthly income and expenses of the local students to verify whether the calculated amounts make economic sense or need to be adjusted accordingly. We removed the responses from non-native students from our sample. We also excluded participants whose answers showed indications of violation of first-order stochastic dominance, more exactly violations of internality (Gneezy et al. 2006), or monotonicity (as a manifestation of response errors, such as choosing to play a binary lottery at low probabilities for a gain but refraining from playing a lottery at high probabilities for the same gain). These participants were excluded to reduce the influence of outlier responses, as we were aware that our approach of direct matching is prone to numerous biases (see also Rieger et al., 2017). We eliminated an observation if there was at least one strong internality or monotonicity violation while allowing any number of weak ones. For the difference between strong and weak violations in this regard, see Rieger et al. (2015).

To measure individual decision making, we elicited answers to two novel and easy-to-apply hypothetical questions in our survey. These were related to the two facets of financial life in our analysis: risk and time.

Imagine you have an amount of 5,000 €. How much money would you invest in stocks/funds and fixed income assets?

Stocks/Funds \_\_\_\_\_

Fixed income assets \_\_\_\_\_

If you had the choice between a bond of one-year maturity with an interest rate of 3% p.a. and a bond of five-year maturity with an interest rate of 4.5% p.a., which option would you choose?

- Bond of one-year maturity
- Bond of five-year maturity

These questions have not been used in previous literature. All other questions used in our survey have already been introduced in the literature. Although one might assume a high persistency across the different questions our survey comprised of, empirical evidence with regard to stability is neither consistent nor conclusive (Eckel et al., 2005; Deck et al., 2013; Bradford et al. 2014; Gürdal et al., 2017).

We calculated our indicator for risk-taking as the ratio of money invested in stocks or investment funds, as answered in our novel question. We coded our self-assessed risk-taking variable correspondingly, such that a higher score indicates a higher propensity to take risk. Our measure for time-related behavior was a dummy variable that takes the value of 1 if the individual's answer to our novel question is to choose the bond with longer maturity; otherwise, it is 0. We again encoded the self-assessed patience measure in the same direction, so that a higher value relates to a more patient behavior. Although the measures of riskiness and time orientation of portfolios appear crude, they are easy to interpret and understand and apparently have some relevance to household finance.

The survey covered information on demographic, economic, and cultural characteristics that may influence investment and saving decisions. Overall, the age of respondents in our sample varied from 17 to 32 years, with a mean age of 23 years. Of the total respondents, 46.4% were men. The median monthly income after adjustments according to each country's purchasing power at the time the survey was conducted was 458.33 €. The median wealth after adjustments amounted to 2,000 €. Respondents were relatively evenly distributed among the countries. Although there were obvious discrepancies, the differences in the income and wealth of students from different countries were relatively small. Nevertheless, we controlled for such differences in the regressions. Table 1 provides an overview of some basic facts about the data.

We examined the extent to which measured risk-taking and patience predict the answers to our novel survey questions concerning hypothetical financial behavior. The econometric baseline specification was a Tobit model for risk-taking and a logit model for patience with robust standard errors, estimated on a cross section of individuals and countries. The dependent variable indicated financial decision making. As regressors, we used a broad set of parameters for risk-taking and patience, as outlined in Section 2. All our specifications controlled for a set of socioeconomic characteristics that existing theory and empirical studies suggest as relevant for household portfolio choice. More specifically, we included respondents' age, gender, wealth, and income. We also included country dummies and individually measured cultural dimensions to disentangle the observed effects from the impacts of the societal and cultural background. These cultural dimensions, according to Hofstede (2001), play a more important role in our analyses as there is evidence that they heavily influence individual decision making (Breuer et al., 2014; Ferris et al., 2013; Frijns et al., 2013; Lievenbrück & Schmid, 2014; Rieger et al., 2015; Wang et al., 2016) despite a number of shortcomings (for an overview, see Bearden et al., 2006). We reported standardized coefficients so that the effects of different variables are easily comparable.

## 4 | SURVEY RESULTS

### 4.1 | Correlations

To obtain an overview, we analyzed how the survey measures relate to each other and other demographic factors. We first used a Shapiro–Wilk test to determine the correct method to measure our correlations. Table S.1 in Supplementary File Section II shows that none of our variables were normally distributed ( $p < .1$ ), which is why we ran a Spearman correlation test. Since gender represents a binary variable, it was excluded from the analyses in this section. The Spearman's correlations (see Table S.2 in Supplementary File Section II) indicated that our more sophisticated measures of patience correlated significantly with the demographic variables of age, income, and wealth. People of older age, with higher income, or with higher wealth demonstrated less patience. Interestingly, the self-assessed patience indicator was not significantly correlated with more sophisticated measures of patience. The corresponding statement about the self-assessed risk-taking indicator and other measures of risk-taking holds true as well. Risk aversion in gains was strongly positively correlated with risk aversion in losses, what one would expect. However, we found risk aversion in losses to be significantly negatively correlated with loss aversion. Our measure of probability weighting was significantly correlated with risk aversion in gains. We do not have explanations for the latter two

TABLE 1 Descriptive statistics, survey evidence (average values)

	All	China	Germany	India	Iran	Italy	Singapore	South Korea	Spain	Russia
Male	0.48	0.34	0.46	0.56	0.59	0.53	0.52	0.47	0.52	0.36
Age	23.13	18.61	24.81	26.62	23.11	22.96	22.48	22.85	23.34	23.50
Income	681.32	665.57	693.05	1,139.92	922.27	443.33	1,027.54	487.12	662.95	329.70
Wealth	8,357.66	1,805.18	9,004.55	12,872.50	7,315.21	6,619.68	9,872.11	8,101.59	6,197.28	9,505.47
Ratio of risky assets in individual portfolios	0.40	0.38	0.38	0.39	0.39	0.46	0.54	0.41	0.30	0.42
Ratio of individuals choosing the higher maturity bond	0.45	0.48	0.56	0.38	0.10	0.57	0.47	0.44	0.58	0.43
Self-assessed risk-taking	2.43	2.25	1.85	3.64	2.35	2.06	2.72	2.33	1.75	2.40
Self-assessed patience	3.01	2.80	3.00	3.13	2.75	3.14	3.18	3.10	2.87	2.65
Risk aversion (gains)	0.51	0.55	0.59	0.66	0.16	0.48	0.60	0.48	0.62	0.37
Risk aversion (losses)	0.60	0.63	0.79	0.67	0.17	0.60	0.72	0.55	0.75	0.45
Loss aversion	1.73	1.60	1.63	2.12	1.41	2.10	2.17	2.03	1.73	1.66
Probability weighting	0.65	0.62	0.53	0.57	0.84	0.49	0.70	0.68	0.70	0.66
Present bias parameter	0.40	0.26	0.58	0.52	0.50	0.13	0.73	0.41	0.25	0.39
Long-term discount factor	0.82	0.82	0.85	0.84	0.81	0.79	0.85	0.84	0.80	0.82
Uncertainty avoidance	65.46	104.65	53.91	55.16	56.33	73.52	33.43	47.61	82.66	41.07
Long-term orientation	46.91	49.14	44.90	29.74	55.32	46.68	46.27	57.57	59.00	44.23
Power distance	20.52	-1.87	27.04	26.09	34.95	22.05	26.54	12.33	20.33	29.91
Individualism	72.67	62.96	73.42	71.18	83.39	73.18	90.96	60.27	62.19	96.68
Masculinity	58.58	48.78	18.21	56.41	80.77	104.18	0.76	17.21	53.58	141.11
No. of observations in Table 3	420	89	66	51	36	39	30	43	20	46
No. of observations in Table 5	700	91	97	118	65	100	59	54	42	74



observations at this point. In summary, this is first tentative evidence that our different measures for willingness to take risk and the level of patience were almost uncorrelated and thus described different measurement approaches. Therefore, we can assume large differences in the predictive power for the answers to our savings and investment decisions.<sup>1</sup>

## 4.2 | Predicting risk-taking behavior

We analyzed the distribution of the answers to our risk self-assessment question and found sufficient variation across countries. Therefore, we could conclude that the participants understood this question well and did not compare themselves with other people in the same country, which would have led to very similar distributions instead (for a closer look, see Figure S.1 in Supplementary File Section III). Patterns of known country differences were also found in our data, but there were exceptions; for example, participants from Russia assessed themselves as quite risk-tolerant relative to the other countries, which is an unusual finding for their cultural background. However, the number of observations for some countries, including Russia, was relatively small and might not be completely representative of the respective country. As we were mainly interested in individual behavior and the possibilities of its prediction in the realm of saving and investment decisions, this finding did not pose too great a problem for our study. Moreover, we could not rule out that participants might have had different understandings of what “fixed income” means, but it was clear at all times that this option was the least risky one.

The next step of our empirical analysis regressed our newly introduced survey measure for risk-taking (named “*Risky Share*”) on a set of socio-demographic variables in model (1), Table 2. Subsequently, we added two different sets of variables that relate to risk-taking in models (2) and (3). Moreover, in model (4), we also added Hofstede's dimensions in the same way, as there is evidence that they exhibit an influence on risk-taking as well, as briefly described in Section 3. Finally, we ran a horseshoe regression in model (5) that included the two sets of risk-taking variables plus Hofstede's dimensions at a time. Therefore, model (5) was based on the following regression:

$$\text{Risky Share} = b_0 + b_1 \cdot \text{SART} + b_2 \cdot \alpha^+ + b_3 \cdot \alpha^- + b_4 \cdot \lambda + b_5 \cdot \gamma + b_6 \cdot \text{UAI} + b_7 \cdot \text{LTO} + b_8 \cdot \text{PDI} + b_9 \cdot \text{IDV} + b_{10} \cdot \text{MAS} + \mathbf{b} \cdot \mathbf{C} + \varepsilon \quad (5)$$

**TABLE 2** Tobit analysis of risk behavior and risk-taking indicators, first step, survey evidence

	model (1)		model (2)		model (3)		model (4)		model (5)	
Male	0.074***	<i>0.023</i>	0.048**	<i>0.022</i>	0.046	<i>0.043</i>	0.071***	<i>0.023</i>	0.068	<i>0.042</i>
Age	0.000	<i>0.004</i>	0.003	<i>0.004</i>	0.003	<i>0.008</i>	−0.001	<i>0.004</i>	0.004	<i>0.008</i>
Income	0.000	<i>0.000</i>	0.000	<i>0.000</i>	0.000**	<i>0.000</i>	0.000	<i>0.000</i>	0.000**	<i>0.000</i>
Wealth	0.000	<i>0.000</i>	0.000***	<i>0.000</i>	0.000	<i>0.000</i>	0.000	<i>0.000</i>	0.000	<i>0.000</i>
Self-assessed risk-taking			0.113***	<i>0.015</i>					0.072**	<i>0.034</i>
Risk aversion (gains)					0.156	<i>0.112</i>			0.066	<i>0.110</i>
Risk aversion (losses)					0.004	<i>0.076</i>			0.040	<i>0.080</i>
Loss aversion					0.015	<i>0.026</i>			0.012	<i>0.021</i>
Probability weighting					0.181**	<i>0.073</i>			0.181**	<i>0.071</i>
Uncertainty avoidance							−0.001***	<i>0.000</i>	0.000	<i>0.000</i>
Long-term orientation							−0.001	<i>0.001</i>	−0.003**	<i>0.001</i>
Power distance							0.000	<i>0.000</i>	0.000	<i>0.000</i>
Individualism							0.000**	<i>0.000</i>	0.000	<i>0.000</i>
Masculinity							0.000	<i>0.000</i>	0.000	<i>0.000</i>
Country dummies	yes		yes		yes		yes		yes	
Observations	642		637		155		609		153	
Pseudo $R^2$	0.121		0.302		0.280		0.200		0.504	
$F$ statistic	7.313		12.380		2.390		4.730		2.949	
$p$ -Value	0.000		0.000		0.003		0.000		0.000	

Note: The dependent variable is the ratio of risky assets in individual portfolios. The independent variables consist of a group of socio-demographic variables and three different sets of variables for risk-taking as outlined in Section 3.

\*\*\*, \*\*, and \* denote significance at 1%, 5%, and 10%, respectively. Robust standard errors are reported in italics.

with *SART* referring to self-assessed risk-taking, and vector  $C$  describing all other control variables. Our analysis demonstrated that each set of measures for risk-taking and Hofstede's cultural dimensions had considerable predictive power with respect to the hypothetical risk-bearing choice of individuals. The explanatory power of the regression models increased significantly when we included indicators for risk propensity. The self-assessed risk-taking indicator thereby induced the highest increase in  $R^2$ . Of the four parameters that covered more sophisticated risk-taking measures, solely the probability bias appeared to have a significant impact on decision making.

The horserace regression reconfirmed that the self-descriptive risk-taking indicator has a dominant influence on decision making, although the impact of the probability bias in this special regression is significant as well. We noted that the sample size of the horserace regression ranked rather low, owing to a comparably high level of invalid responses for the implicitly estimated measures. Individuals appeared to incur difficulties in answering the underlying lottery questions consistently. Such violations of revealed measure conditions persisted even across all countries and were not uncommon when eliciting answers to lottery questions (Rieger et al., 2017).

The second step of our regression analysis subsequently added the significant variables from the previous regressions to the socio-demographic controls and substantiated our results. In this process, we restricted our analysis to the cross section of data points for which we had data on all of these variables available. This procedure achieved a better comparability of the impact of different risk measures while using a large sample size. This also eliminated any possible sample composition bias. The sample size for the models was 420 individuals, of which the highest number of participants were from China ( $n = 89$ ), followed by Germany ( $n = 66$ ) and India ( $n = 51$ ). The lowest number of observations was reported from Spain ( $n = 20$ ), whereas the number of participants in all other countries considered ranged between 30 (Singapore) and 46 (Russia, see Table 1).

Results in Table 3 paint a consistent picture. The self-assessed risk-taking indicator could best explain hypothetical financial risk-taking, as measured by our novel question. The pseudo  $R^2$  in the simple specification of model (2) was 32.4% while the explanatory power of the other models hardly exceeded 20%. Besides self-assessment, other variables measuring risk-taking achieved a very meager increase in the pseudo  $R^2$  of 7% at best of the baseline model (1) that only included socio-demographic controls.

This confirmed our findings that self-assessed indicators are consistently most important for explaining hypothetical risk behavior, as measured by our novel question. Probability bias no longer had a significant influence. The explanatory power of this model yielded a solid 42.2%, and the control variables of the regression models were generally in line with what was expected based on the previous literature. Male participants appeared to be more risk-taking. Age, income, and wealth did not significantly influence risk behavior in the majority of our regression models. Hofstede dimensions only partly turned out to be significant, and the coefficients were relatively close to zero.

### 4.3 | Predicting the degree of patience

With respect to the time self-assessment question, we found sufficient variation across countries (see Figure S.2 in Supplementary File Section II). In this case, patterns of known country differences in our data could not be spotted as easily as for the self-assessment of risk-taking.

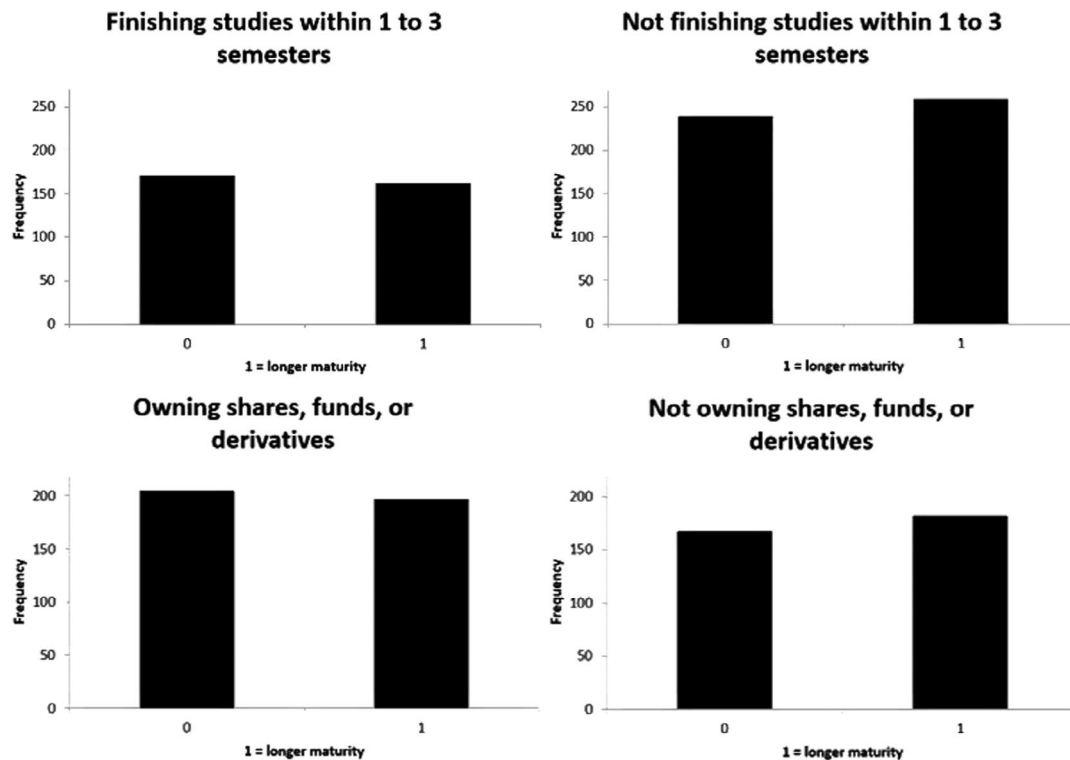
One might object that there are reasons other than preferences for our participants to choose the shorter maturity bond over the longer one. Students might expect interest rates to change over the next few years, or they might be at different points in their studies, leading to an anticipated need for money in a fixed period of time. We checked these issues by analyzing the participants' bond choices. We compared (a) the knowledge of financial markets, using investments in stocks, bonds, or funds against non-investors as proxies, and (b) the group of participants who are planning to finish their studies within the next one to three semesters against those who are going to study longer than three semesters (see Figure 1). As can be seen at first glance, the distributions did not differ greatly from each other. We performed a  $t$ -test and could not reject the null hypothesis of equal means in either case. Therefore, we supposed that students did not decide in different ways based on their market knowledge using financial market participation as proxy or based on their remaining study time. We could not rule out possible misunderstandings of when exactly the bond coupon is paid. For the five-year maturity option, participants might have understood the coupon payment to be made every year or alternatively to be paid as compound interest in one sum at the time of maturity. However, this did not cause too great a problem for our study since it was clear at all times that the initially invested amount of money, that is, the principal, is to be blocked for one or five years, respectively.

Our analysis of patience proceeded in a manner similar to that presented in the previous section. We first regressed hypothetical time behavior (described by the variable "*Longer Maturity*") on the two different sets of measurements for patience plus the Hofstede scores in models (2) to (4), Table 4. The combined regression model (5) is as follows:

**TABLE 3** Tobit analysis of risk behavior and risk-taking indicators, second step, survey evidence

	model (1)	model (2)	model (3)	model (4)	model (5)	model (6)	model (7)						
Male	0.089***	0.028	0.062**	0.027	0.088***	0.028	0.084***	0.028	0.064**	0.027			
Age	-0.003	0.005	0.001	0.005	-0.003	0.005	-0.001	0.005	-0.004	0.005	0.002	0.005	
Income	0.000	0.000	0.000	0.000	0.000*	0.000	0.000	0.000	0.000	0.000	0.000	0.000	
Wealth	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	
Self-assessed risk-taking		0.117***		0.018								0.109***	0.018
Probability weighting			0.071*	0.038								0.035	0.037
Uncertainty avoidance					-0.001***	0.000						-0.001***	0.000
Long-term orientation									-0.002**	0.001		-0.001**	0.001
Individualism											0.000	0.000	0.000
Country dummies	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Observations	420	420	420	420	420	420	420	420	420	420	420	420	420
Pseudo $R^2$	0.149	0.324	0.162	0.214	0.179	0.159	0.159	0.159	0.179	0.159	0.159	0.159	0.422
$F$ statistic	3.709	7.174	3.955	4.558	3.794	3.695	3.695	3.695	3.794	3.695	3.695	3.695	7.475
$p$ -Value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Note: The dependent variable is the ratio of risky assets in individual portfolios. The independent variables consist of a group of socio-demographic variables and those variables for risk-taking that gained significance in Table 2. \*\*\*, \*\*, and \* denote significance at 1%, 5%, and 10%, respectively. Robust standard errors are reported in italics.



**FIGURE 1** Distributions of survey bond choices. The figure shows the distributions of bond choices (1 = longer maturity) by remaining study time and financial market participation

$$\text{Longer Maturity} = b_0 + b_1 \cdot \text{SAP} + b_2 \cdot \beta + b_3 \cdot \delta + b_4 \cdot \text{UAI} + b_5 \cdot \text{LTO} + b_6 \cdot \text{PDI} + b_7 \cdot \text{IDV} + b_8 \cdot \text{MAS} + \mathbf{b} \cdot \mathbf{C} + \varepsilon \quad (6)$$

with *SAP* referring to self-assessed patience, and vector *C* describing all other control variables.

Next, we established the relation between patience and survey-measured financial decision making in a cross section of exactly 700 data points for which we had data on all relevant variables available. As pointed out in Section 4.1, this procedure provided a more consistent view. The highest number of observations can be reported from India ( $n = 118$ ), followed by Italy ( $n = 100$ ), Germany ( $n = 97$ ), and China ( $n = 91$ ). Relatively few participants were disclosed from Russia ( $n = 74$ ), Iran ( $n = 65$ ), Singapore ( $n = 59$ ), South Korea ( $n = 54$ ), and Spain ( $n = 42$ ). In general, more individuals managed to answer our questions consistently compared with the risk-taking section, which might be traced back to an easier understanding.

Our findings were relatively consistent. The self-descriptive patience indicator consistently explained hypothetical individual decision making, as measured by our novel question, at the best possible rate. This indicator was able to increase the pseudo adjusted  $R^2$  from 3.4% in the baseline model to 4.7%. The horserace regression yielded 4.8% for the pseudo adjusted  $R^2$ . We did not observe a significant influence of gender or age in our regression models. An increase in income and wealth coincided with more patient time behavior.

The coefficient of the self-descriptive patience indicator was 1.504 in model (2) in Table 5. A move to the next higher value on the time preference scale thus increased the likelihood that an individual hypothetically chooses the bond with the higher maturity by 50%. The main conclusion we draw from our results is that self-assessed indicators affected hypothetical financial behavior as measured by our novel questions in an obvious and intuitive manner.

As our results are similar with respect to both risk-taking and patience and as parameter elicitation regarding time discounting is somewhat easier than parameter determination regarding prospect theory, the low predictive power of these two approaches may not be completely caused by the high level of complexity. Moreover, we concluded that our newly introduced hypothetical questions about risk and time behavior turned out to deliver results that were in line with the existing literature, thus presumably offering an easy-to-apply method to reliably reflect decision making in more realistic situations.

**TABLE 4** Logit analysis of time behavior and patience indicators, first step, survey evidence

	model (1)		model (2)		model (3)		model (4)		model (5)	
Male	1.041	<i>0.167</i>	1.061	<i>0.174</i>	1.069	<i>0.181</i>	0.970	<i>0.162</i>	1.036	<i>0.187</i>
Age	0.959	<i>0.027</i>	0.973	<i>0.028</i>	0.965	<i>0.029</i>	0.968	<i>0.028</i>	0.991	<i>0.032</i>
Income	1.000*	<i>0.000</i>	1.000**	<i>0.000</i>	1.000**	<i>0.000</i>	1.000**	<i>0.000</i>	1.000***	<i>0.000</i>
Wealth	1.000*	<i>0.000</i>	1.000**	<i>0.000</i>	1.000*	<i>0.000</i>	1.000	<i>0.000</i>	1.000	<i>0.000</i>
Self-assessed patience			1.523***	<i>0.163</i>					1.501***	<i>0.175</i>
Present bias parameter					1.337	<i>0.391</i>			1.299	<i>0.399</i>
Long-term discount factor					1.737	<i>1.539</i>			2.921	<i>2.712</i>
Uncertainty avoidance							1.001	<i>0.001</i>	1.002	<i>0.001</i>
Long-term orientation							1.007*	<i>0.004</i>	1.004	<i>0.004</i>
Power distance							1.000	<i>0.002</i>	1.000	<i>0.002</i>
Individualism							1.002	<i>0.002</i>	1.002	<i>0.002</i>
Masculinity							1.001	<i>0.001</i>	1.001	<i>0.001</i>
Country dummies	yes		yes		yes		yes		yes	
Observations	722		711		668		685		628	
$\chi^2$	50.290		63.980		49.170		47.270		57.420	
<i>p</i> -Value	0.000		0.000		0.000		0.000		0.000	
Pseudo $R^2$	0.064		0.081		0.067		0.070		0.088	
Pseudo $R^2$ adjusted	0.037		0.052		0.034		0.031		0.039	
Specificity	69.700		68.540		63.970		66.670		60.900	

Note: The dependent variable is the probability that an individual chooses the higher maturity bond. The independent variables consist of a group of socio-demographic variables and three different sets of variables for patience as outlined in Section 3.

\*\*\*, \*\*, and \* denote significance at 1%, 5%, and 10%, respectively. Robust standard errors are reported in italics.

**TABLE 5** Logit analysis of time behavior and patience indicators, second step, survey evidence

	model (1)		model (2)		model (3)		model (4)	
Male	0.976	<i>0.159</i>	1.015	<i>0.167</i>	0.945	<i>0.155</i>	0.985	<i>0.164</i>
Age	0.968	<i>0.028</i>	0.975	<i>0.028</i>	0.971	<i>0.028</i>	0.977	<i>0.028</i>
Income	1.000**	<i>0.000</i>	1.000**	<i>0.000</i>	1.000**	<i>0.000</i>	1.000**	<i>0.000</i>
Wealth	1.000**	<i>0.000</i>	1.000**	<i>0.000</i>	1.000**	<i>0.000</i>	1.000**	<i>0.000</i>
Self-assessed patience			1.504***	<i>0.162</i>			1.491***	<i>0.162</i>
Long-term orientation					1.007*	<i>0.004</i>	1.006*	<i>0.004</i>
Country dummies	yes		yes		yes		yes	
Observations	700		700		700		700	
$\chi^2$	47.460		60.910		48.550		61.010	
<i>p</i> -Value	0.000		0.000		0.000		0.000	
Pseudo $R^2$	0.061		0.076		0.065		0.080	
Pseudo $R^2$ adjusted	0.034		0.047		0.036		0.048	
Specificity	70.650		67.010		69.350		69.610	

Notes: The dependent variable is the probability that an individual chooses the higher maturity bond. The independent variables consist of a group of socio-demographic variables and those variables for patience that gained significance in Table 4.

\*\*\*, \*\*, and \* denote significance at 1%, 5%, and 10%, respectively. Robust standard errors are reported in italics.

## 5 | LABORATORY EXPERIMENT

To compare the outcomes of a more lifelike situation with our novel, simple, and easy-to-use risk and time behavior elicitation method, we complemented our survey procedure through an experimental study. As described before, the extent to which

preferences are misreported also depends on the data collection method. If our novel questions of hypothetical behavior are not subject to misreporting, we would expect the predictive power of the different risk and time measurements to be comparable for the survey and the experiment. Moreover, an experiment allowed us to address the following points that go beyond the scope of our survey:

- Offering performance-based monetary incentives
- Avoiding possible halo effects
- Determining risk and time decision making in a single setup

As we have elaborated before, previous studies have found that, all other things being equal, incentivization does not necessarily lead to different outcomes (Rieger et al., 2015; Vieider et al., 2015). However, the literature in this regard is still inconclusive. Although laboratory experiments typically suffer from smaller sample sizes and self-selection biases, offering real monetary incentives might ensure that participants are motivated to give thoughtful answers. By further disassociating the risk-taking/patience measurements and the actual decision making task, the experiment also helped avoid possible measurement errors that might solely arise because respondents tried to answer these questions consistently (halo effect). This might be particularly a problem with respect to patience because the simple investment problem presented in our novel question sounds, to some degree, similar to the self-assessment question. The experiment considered this in general and separated the corresponding questions sufficiently. Moreover, as described below, in the laboratory experiment, individuals now determine the risk and time structure of their portfolio simultaneously, loosening a potential mental connection to the previously answered self-assessment questions.

We conducted our laboratory experiment at a large German university using computers. Each participant completed the same questionnaire as described in Section 3. We removed our two hypothetical questions regarding financial behavior in terms of risk and time. Instead, participants were asked to make a portfolio allocation decision using the experimental task design of Kaufmann et al. (2013), given a hypothetical endowment of 5,000 €, for both one- and three-year maturity portfolios. All individuals knew that their final payments depended on their investment choices. The participants had to choose between the purchase of a riskless asset and a risky investment opportunity and state the share of money to be invested in each of those two options. The risk-free rates of return of the riskless asset were set at 1.5% p.a. for the one-year option and 2.0% p.a. for the three-year option. The risk-return profile of the risky investment opportunity was adapted from MSCI World Index data. First, a brief interactive introduction to probability distributions was provided, following the concept of Kaufmann et al. (2013). After making an allocation choice using a slider mechanism, a probability distribution of the possible outcomes of the chosen portfolio allocation based on a Monte Carlo simulation was presented. Participants could then decide to either change the allocation or accept the portfolio choice. This procedure was repeated for the allocation decision concerning a portfolio of longer maturity. Eventually, participants were asked to decide between either taking the one-year maturity portfolio and being paid one month after the event or taking the three-year maturity portfolio and being paid three months after the event. At this point, the allocation choices could no longer be altered. Moreover, participants could only see the probability distributions of their two portfolio choices involving the possible investment outcomes and the respective outcome probabilities at the time of maturity. The final payment, which was made via bank transfer, was then computed based on the probability distribution of the portfolio that the participant chose, while 500 € in the experiment corresponded to a real payment of 1 €. The final payments ranged between 8.07 € and 16.59 €, with an average payment of 12.03 € for an average of 45 min of work. Using this procedure, we could observe both, participants' risk decisions and time decisions in a single experimental setup. Two exemplary excerpts of the experiment are shown in Figures 2 and 3. Kaufmann et al. (2013) explained the experimental procedure in a more detailed way. This setup also allowed us to tackle the problem of different understandings of what "fixed income" means. The visualization component did not allow for any deviant interpretations, and the same reasoning could be adapted to possible misunderstandings of the coupon payment. The experimental validation might have been helpful in this context as well.

The final sample consisted of 163 observations. Again, participants whose answers indicated violation of first-order stochastic dominance because of response errors were excluded from the sample. This time, the share of excluded answers was considerably smaller because participants, in general, responded more consistently. This may be due to the fact that individuals participating in the laboratory experiment did so to earn money. Therefore, they might have felt obliged to read the questions very thoroughly and answer only after reflecting well. To this effect, incentivization may indeed be useful. However, as pointed out below, our main conclusions are not affected by incentivizing answers.

We defined the indicator of time-related behavior as a variable that takes the value of 1 if the individual chooses the investment with three-year maturity and otherwise 0, similar to our novel indicator of time-related behavior presented in Section 3. Since participants had to decide about the share to be invested in the respective investment options twice (once for one-year

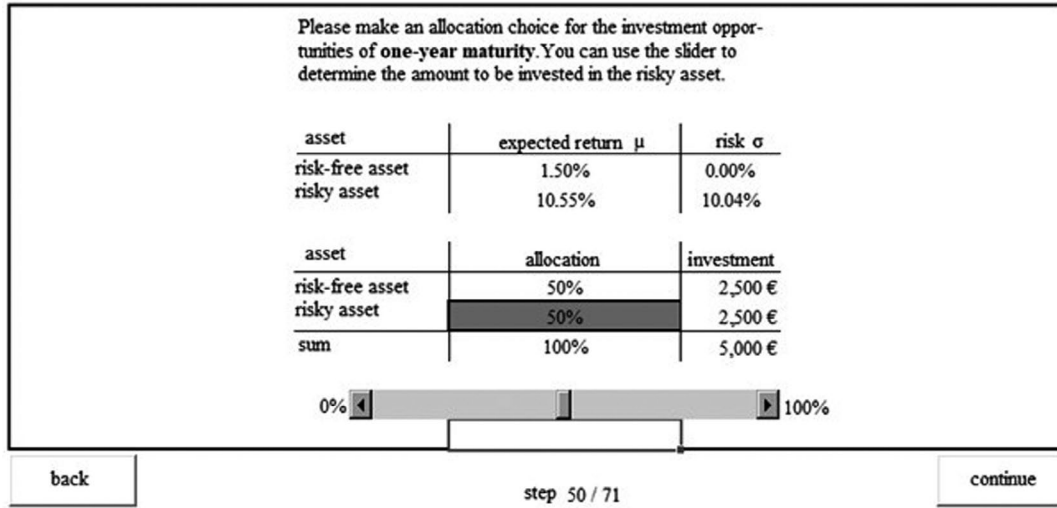


FIGURE 2 Experiment: Allocation mechanism. The figure shows the slider mechanism used by participants of the experiment to make an allocation decision

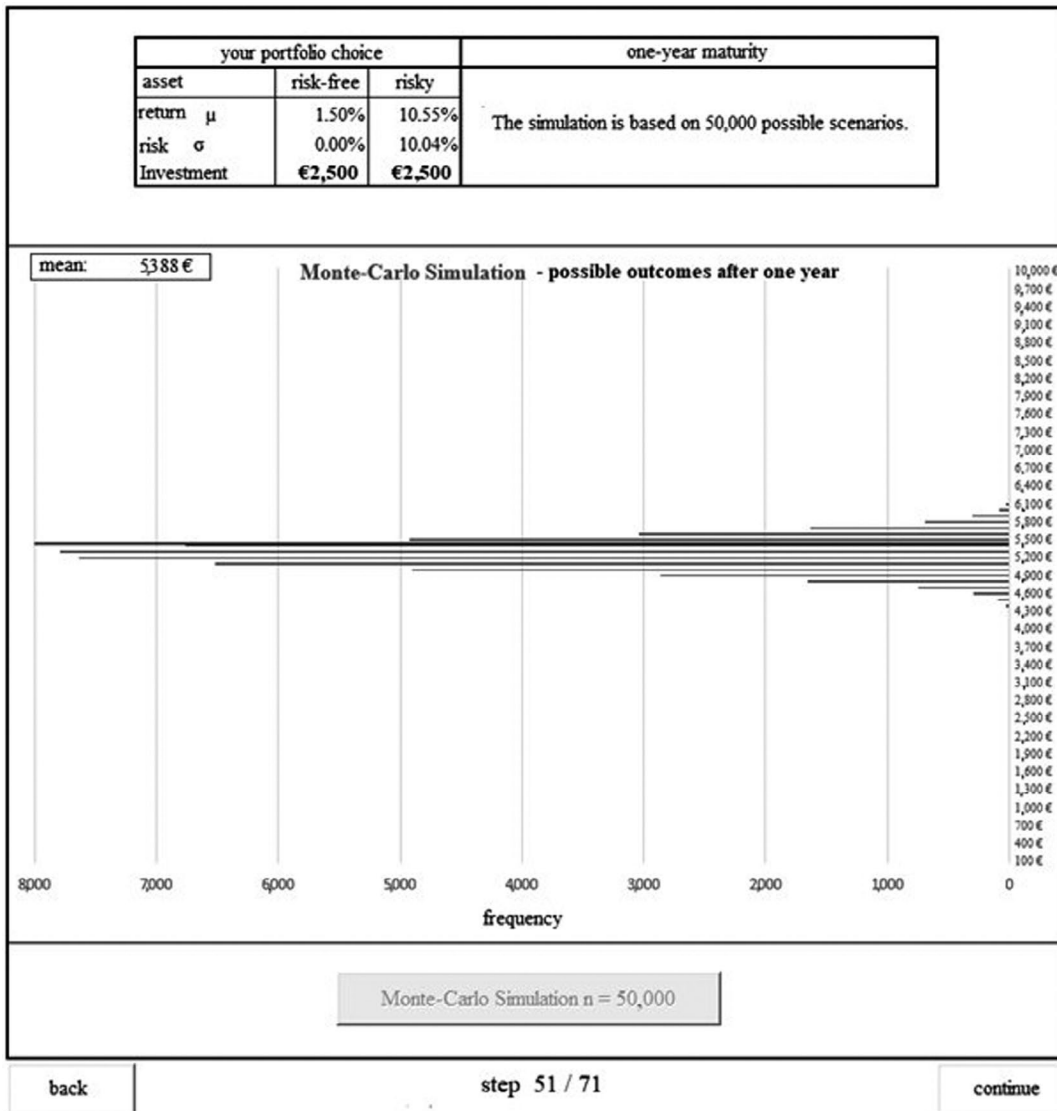


FIGURE 3 Experiment: Presentation of the probability distribution. The figure shows how the concept of a probability distribution has been presented to participants of the experiment

maturity and once for three-year maturity portfolios) in this analysis, the indicator of risk-related behavior was defined as each individual's mean share of money invested in the risky asset.

To allow for comparison between the two analyses presented in this study, we attempted to create a sample consisting of individuals of similar socioeconomic backgrounds as in our first study. The mean age of the participants in the sample was 24.28 years. The median monthly income was 650.00 €, and the median wealth was 3,000.00 €. Our ratio of male participants, however, was slightly higher than that in our survey setting. Although our experimental sample mainly consisted of students of German descent, we controlled for existing country-specific differences. Since our sample is rather restricted regarding country differences compared with our sample of the analyses described in the previous sections, we included a dummy variable that is 1 if the participant is not of German nationality and 0 otherwise. As in our survey, we also gathered information about Hofstede scores.

Descriptive statistics and regression models are presented in Supplementary File Section IV. Compared with our survey findings, the results of our laboratory experiment mostly painted a consistent picture. In terms of risk-taking, the following results were found in all our models: the coefficient of the self-assessed measures had a significant impact; the findings for probability weighting were inconsistent; and the other measures were not significant at all. Concerning patience, only the self-assessed measure turned out to be highly significant, while no other relevant coefficient was of any significance, except for one single model in which the Hofstede cultural dimension of masculinity was significant at the 10% level. Therefore, our findings strongly confirm the survey results and point to the fact that our newly introduced questions were able to proxy decision making well enough to ensure delivering results that are comparable to lifelike situations. A second important observation is that self-assessed indicators worked best for predicting behavior.

## 6 | TACKLING THE ERROR-IN-VARIABLES PROBLEM

Because of the general nature of experiments and surveys, we could not rule out that there were measurement errors in the explanatory variables leading to biased estimators. We were aware of this problem. Regarding the complexity of lottery-choice-based questions and our results presented in the previous sections, and more precisely looking at the widely appearing non-significance of the variables derived from lottery-choice-based questions, we suspect that there might be measurement errors in the set of explanatory variables. In general, there are three possibilities to address this issue: one may use an instrumental variables approach, rely on the dynamic panel estimators of Arellano and Bond (1991), or apply the measurement error-consistent generalized method of moments/cumulants estimators (Erickson et al., 2017; Erickson & Whited, 2000), to our lottery-measured explanatory variables. The dynamic panel estimators from Arellano and Bond (1991) are based on lagged regressors and thus require panel data. Replacing the regressors we wanted to analyze with instrumental variables does not make sense. Therefore, for our purposes, only the third approach can be employed. Because this approach is not suited for binary dependent variables, it can only be applied to our risk-taking dataset.

The exact outcomes of our analysis are presented in the Supplementary File Section V. We found that our results were very sensitive to the choice of moments/cumulants and their highest order. Although there were indications that, when considering an error-in-variables model, lottery-based measures of risk-taking may gain reliability, the explanatory power of these results when applying the error-in-variables approach remains questionable. This is because there is no clear structure regarding the significance of other risk-taking measures besides self-assessment. The latter variable still had a consistent significant impact in all our models, except once. However, this may be an interesting starting point for future research.

## 7 | CONCLUSION

A recurrent theme in economics is the description of individual risk-taking and patience measures and their relation to choice behavior. The novel feature of our study lies in linking different risk-taking and patience measures to a new and easy-to-use survey approach utilized to proxy lifelike investment situations. Moreover, we replicated and extended the existing literature concerning the predictive power of more or less sophisticated risk-taking and patience elicitation methods.

Our evidence was broadly consistent with the hypothesis that individual indicators play an important role in investment decisions. Analyses based on our novel survey questions regarding investment behavior achieved high comparability with those relying on actual decision making in an experimental context. Another finding was that, in general, self-assessed indicators consistently generated the best predictions of behavior. It appears that the parameters based on more sophisticated elicitation methods have little importance in explaining the risk and time dimension of portfolio selection, and our results were inconsistent in



this regard. A plausible explanation is that ordinary respondents may incur problems in answering more sophisticated choice questions between different alternatives. The relatively low number of observations for these questions because of many invalid answers supports this conjecture and casts doubt on such methods. We tried to tackle this issue with mixed results using special regression approaches for such a measurement problem regarding explanatory variables. Although this might increase the empirical relevance of lottery-choice-induced parameter computations, the results strongly depended on the manner by which such regression models are calibrated, and there is no straightforward way to unambiguously determine the best model variant. Apparently, there is a need for a regression method that considers the error-in-variables problem but yields more stable results than the methods currently available. Methodological progress might help mitigate measurement problems caused by sophisticated indicator elicitation procedures in the future. Then, we believe, rather complex measurement methods may be more helpful than they are for the time being.

We tentatively conclude that economic data concerning financial decision making can be generated reliably using low-cost survey questions that proxy actual decision making in an experimental context. In addition, our results justify the current practices of financial advisors or private bankers, who use rather rough methods to measure their clients' risk-taking and patience attitudes. Our findings support prior evidence that demonstrates persistent discrepancies between the predictions of traditional, more sophisticated models and behavior (Andersen et al., 2006; Erner et al., 2013; Lönnqvist et al., 2015).

However, our data had some drawbacks. First, we only examined university students and not a representative sample of the total population. Although this caveat has to be kept in mind, we believe our study remains valuable in providing evidence on the explanatory power of different variables as this sample selection offers important advantages. On the one hand, because of their quantitative background, students of business administration and economics, as well as in the field of STEM (science, technology, engineering, and mathematics) can understand the numeric formulations of the questions without great difficulty. Asking a representative sample of the population the same numerical questions involving probabilities might have been less straightforward. Previous studies have demonstrated that student samples incur fewer violations of revealed measure conditions (von Gaudecker et al., 2011). On the other hand, they constitute a relatively homogeneous and comparable group across countries, helping to reduce the influence of country-level background variables, and they represent future investors. In addition, previous studies have documented that student samples reveal behavioral patterns that do not differ much from those of non-student samples (Druckman & Kam, 2011; Falk et al., 2013). Therefore, we are confident that our conclusions remain valid even for more heterogeneous samples. Second, we only compared survey and laboratory data; therefore, we cannot draw any direct conclusions about real-life portfolio selection decisions. There is a need to compare survey data based on our novel questions with genuine, real-life decision making. However, we leave this issue for future research.

## ACKNOWLEDGMENTS

We gratefully acknowledge the numerous comments of three anonymous reviewers and the helpful remarks of participants of IFABS 2019 Conference in Angers, France, which have been extraordinarily helpful in developing our study.

## ORCID

Wolfgang Breuer  <https://orcid.org/0000-0001-5470-4263>

Thomas Rennerken  <https://orcid.org/0000-0002-5830-6059>

Astrid Juliane Salzmann  <https://orcid.org/0000-0002-8291-5225>

## ENDNOTE

<sup>1</sup> We would like to thank an anonymous reviewer for highlighting these issues in this section. Lack of space does unfortunately not permit us to elaborate deeper on these relationships.

## REFERENCES

- Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2006). Elicitation using multiple price list formats. *Experimental Economics*, 9, 383–405. <https://doi.org/10.1007/s10683-006-7055-6>
- Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies*, 58, 277–297. <https://doi.org/10.2307/2297968>
- Bachmann, K., Hens, T., & Stössel, R. (2017). Which measures predict risk taking in a multi-stage controlled investment decision process? *Financial Services Review*, 26, 339–365.
- Barasinska, N., Schäfer, D., & Stephan, A. (2012). Individual risk attitudes and the composition of financial portfolios: Evidence from German household portfolios. *The Quarterly Review of Economic and Finance*, 42, 1–14. <https://doi.org/10.1016/j.qref.2011.10.001>
- Barsky, R. B., Juster, F. T., Kimball, M. S., & Shapiro, M. D. (1997). Preference parameters and behavioral heterogeneity: An experimental approach in the health and retirement study. *The Quarterly Journal of Economics*, 112, 537–579. <https://doi.org/10.1162/003355397555280>

- Bearden, W. O., Money, R. B., & Nevins, J. L. (2006). Multidimensional versus unidimensional measures in assessing national culture values: The Hofstede VSM 94 example. *Journal of Business Research*, *59*, 195–203. <https://doi.org/10.1016/j.jbusres.2005.04.008>
- Bradford, D., Courtemanche, C., Heutel, G., McAlvanah, P., & Ruhm, C. (2014). Time preferences and consumer behavior. Working Paper 20320, NBER.
- Breuer, W., Riesener, M., & Salzman, A. J. (2014). Risk aversion vs. Individualism: What drives risk taking in household finance? *The European Journal of Finance*, *20*, 446–462.
- Cozby, P. C., & Bates, S. C. (2019). *Methods in behavioral research*. McGraw-Hill Education.
- Crosetto, P., & Filippin, A. (2016). A theoretical and experimental appraisal of four risk elicitation methods. *Experimental Economics*, *19*, 613–641.
- Currim, I. S., & Sarin, R. K. (1989). Prospect versus utility. *Management Science*, *35*, 22–41. <https://doi.org/10.1287/mnsc.35.1.22>
- Davern, M., Rockwood, T. H., Sherrod, R., & Campbell, S. (2003). Prepaid monetary incentives and data quality in face-to-face interviews: Data from the 1996 Survey of Income and program participation incentive experiment. *The Public Opinion Quarterly*, *67*, 139–147. <https://doi.org/10.1086/346012>
- De Leeuw, E. G. (2005). To mix or not to mix data collection modes in surveys. *Journal of Official Statistics*, *21*, 233–255.
- Deck, C., Lee, J., Reyes, J. A., & Rosen, C. C. (2013). A failed attempt to explain within subject variation in risk taking behavior using domain specific risk attitudes. *Journal of Economic Behavior & Organization*, *87*, 1–24. <https://doi.org/10.1016/j.jebo.2012.11.010>
- Dimmock, S. G., & Kouwenberg, R. (2010). Loss-aversion and household portfolio choice. *Journal of Empirical Finance*, *17*, 441–459. <https://doi.org/10.1016/j.jempfin.2009.11.005>
- Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., & Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, *9*, 522–550. <https://doi.org/10.1111/j.1542-4774.2011.01015.x>
- Donkers, B., & van Soest, A. (1999). Subjective measures of household preferences and financial decisions. *Journal of Economic Psychology*, *20*, 613–642. [https://doi.org/10.1016/S0167-4870\(99\)00027-6](https://doi.org/10.1016/S0167-4870(99)00027-6)
- Druckman, J. N., & Kam, C. (2011). Students as experimental participants: A defense of the 'narrow data base'. In D. P. Green, J. H. Kuklinski, & A. Lupia (Eds.), *Cambridge Handbook of Experimental Political Science* (pp. 41–57). Cambridge University Press.
- Duffy, B., Smith, K., Terhanian, G., & Bremer, J. (2005). Comparing data from online and face-to-face surveys. *International Journal of Market Research*, *47*, 615–639. <https://doi.org/10.1177/147078530504700602>
- Eckel, C., Johnson, C., & Montmarquette, C. (2005). Saving decisions of the working poor: Short- and long-term horizons. In G. W. Harrison, J. Carpenter, & J. A. List (Eds.), *Field experiments in economics* (pp. 219–260). Emerald Group Publishing Limited.
- Erickson, T., Parham, R., & Whited, T. M. (2017). Fitting the errors-in-variables model using high-order cumulants and moments. *Stata Journal*, *17*, 116–129. <https://doi.org/10.1177/1536867X1701700107>
- Erickson, T., & Whited, T. M. (2000). Measurement error and the relationship between investment and Q. *Journal of Political Economy*, *108*, 1027–1057.
- Erner, C., Klos, A., & Langer, T. (2013). Can prospect theory be used to predict an investor's willingness to pay? *Journal of Banking and Finance*, *37*, 1960–1973. <https://doi.org/10.1016/j.jbankfin.2012.12.008>
- Falk, A., Meier, S., & Zehnder, C. (2013). Do lab experiments misrepresent social preferences? The case of self-selected student samples. *Journal of the European Economic Association*, *11*, 839–852. <https://doi.org/10.1111/jeea.12019>
- Ferris, S. P., Jayaraman, N., & Sabherwal, S. (2013). CEO overconfidence and international merger and acquisition activity. *Journal of Financial and Quantitative Analysis*, *48*, 137–164. <https://doi.org/10.1017/S0022109013000069>
- Frey, R., Pedroni, A., Mata, R., Rieskamp, J., & Hertwig, R. (2017). Risk preference shares the psychometric structure of major psychological traits. *Science Advances*, *3*, 1–13. <https://doi.org/10.1126/sciadv.1701381>
- Frijns, B., Gilbert, A., Lehnert, T., & Tourani-Rad, A. (2013). Uncertainty avoidance, risk tolerance and corporate takeover decisions. *Journal of Banking & Finance*, *37*, 2457–2471. <https://doi.org/10.1016/j.jbankfin.2013.02.010>
- Frijns, B., Koellen, E., & Lehnert, T. (2008). On the determinants of portfolio choice. *Journal of Economic Behavior & Organization*, *66*, 373–386. <https://doi.org/10.1016/j.jebo.2006.04.004>
- Galizzi, M. M., Machado, S. R., & Miniaci, R. (2016). Temporal stability, cross-validity, and external validity of risk preferences measures: Experimental evidence from a UK representative sample. Working Paper, 67554, LSE. <https://doi.org/10.2139/ssrn.2822613>
- Gneezy, U., List, J.-A., & Wu, G. (2006). The uncertainty effect: When a risky prospect is valued less than its worst possible outcome. *The Quarterly Journal of Economics*, *121*, 1283–1309. <https://doi.org/10.1093/qje/121.4.1283>
- Goldman, N., Moreno, L., & Westoff, C. F. (1989). Collection of survey data on contraception: An evaluation of an experiment in Peru. *Studies in Family Planning*, *20*, 147–157. <https://doi.org/10.2307/1966569>
- Gürdal, M., Kuzubaş, T., & Saltoğlu, B. (2017). Measures of individual risk attitudes and portfolio choice: Evidence from pension participants. *Journal of Economic Psychology*, *62*, 186–203. <https://doi.org/10.1016/j.joep.2017.06.010>
- Haliassos, M., & Bertaut, C. C. (1995). Why do so few hold stocks? *The Economic Journal*, *105*, 1110–1129. <https://doi.org/10.2307/2235407>
- Hofstede, G. (2001). *Culture's consequences: Comparing values, behaviors, institutions, and organizations across nations*. Sage Publications.
- Hyll, W., & Irrek, M. (2015). *The impact of risk attitudes on financial investments*. Working Paper, IWH.
- Jäckle, A., Roberts, C., & Lynn, P. (2010). Assessing the effect of data collection mode on measurement. *International Statistic Review*, *78*, 3–20. <https://doi.org/10.1111/j.1751-5823.2010.00102.x>
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*, 263–291. <https://doi.org/10.2307/1914185>
- Kaufmann, C., Weber, M., & Haisley, E. (2013). The role of experience sampling and graphical displays on one's investment risk appetite. *Management Science*, *59*, 323–340. <https://doi.org/10.1287/mnsc.1120.1607>
- Li, K. K. (2020). *A horse race of preferences for predicting portfolio choice*. Working Paper, Shenzhen Audencia Business School.

- Lievenbrück, M., & Schmid, T. (2014). Why do firms (not) hedge? Novel evidence on cultural influence. *Journal of Corporate Finance*, 25, 92–106. <https://doi.org/10.1016/j.jcorpfin.2013.10.010>
- Lönnqvist, J.-E., Verkasalo, M., Walkowitz, G., & Wichardt, P. C. (2015). Measuring individual risk attitudes in the lab: Task or ask? An empirical investigation. *Journal of Economic Behavior and Organization*, 119, 254–266.
- Newman, A., Bavik, Y. L., Mount, M., & Shao, B. (2021). Data collection via online platforms: Challenges and recommendations for future research. *Applied Psychology*, 1–23.
- Pedroni, A., Frey, R., Bruhin, A., Dutilh, G., Hertwig, R., & Rieskamp, J. (2017). The risk elicitation puzzle. *Nature Human Behaviour*, 1, 803–809. <https://doi.org/10.1038/s41562-017-0219-x>
- Rieger, M. O., Wang, M., & Hens, T. (2015). Risk preferences around the world. *Management Science*, 61, 637–648. <https://doi.org/10.1287/mnsc.2013.1869>
- Rieger, M. O., Wang, M., & Hens, T. (2017). Estimating cumulative prospect theory parameters from an international survey. *Theory and Decision*, 82, 567–596. <https://doi.org/10.1007/s11238-016-9582-8>
- Samuelson, P. A. (1937). A note on measurement of utility. *Review of Economic Studies*, 4, 155–161. <https://doi.org/10.2307/2967612>
- Stecklov, G., Weinreb, A., & Carletto, C. (2018). Can incentives improve survey data quality in developing countries? Results from a field experiment in India. *Journal of the Royal Statistical Society*, 181, 1033–1056. <https://doi.org/10.1111/rssa.12333>
- Szrek, H., Chao, L.-W., Ramlagan, S., & Peltzer, K. (2012). Predicting (un)healthy behavior: A comparison of risk-taking propensity measures. *Judgment and Decision Making*, 7, 716–727.
- Thaler, R. (1981). Some empirical evidence on dynamic inconsistency. *Economics Letters*, 8, 201–207. [https://doi.org/10.1016/0165-1765\(81\)90067-7](https://doi.org/10.1016/0165-1765(81)90067-7)
- Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5, 297–323. <https://doi.org/10.1007/BF00122574>
- Van den Brakel, J. A., Vis-Visschers, R., & Schmeets, J. J. G. (2006). An experiment with data collection modes and incentives in the Dutch Family and Fertility Survey for young Moroccans and Turks. *Field Methods*, 18, 321–334. <https://doi.org/10.1177/1525822X06287533>
- Vieider, F. M., Chmura, T., Fisher, T., Kusakawa, T., Martinsson, P., Mattison Thompson, F., & Sunday, A. (2015). Within- versus between-country differences in risk attitudes: Implications for cultural comparisons. *Theory and Decision*, 78, 209–218. <https://doi.org/10.1007/s11238-014-9418-3>
- Von Gaudecker, H.-M., van Soest, A., & Wengstrom, E. (2011). Heterogeneity in risky choice behavior in a broad population. *American Economic Review*, 101, 664–694. <https://doi.org/10.1257/aer.101.2.664>
- Wakker, P. (2010). *Prospect theory for risk and ambiguity*. Cambridge University Press.
- Wang, M., Rieger, M. O., & Hens, T. (2016). The impact of culture on loss aversion. *Journal of Behavioral Decision Making*, 30, 270–281.
- Wärneryd, K.-E. (1996). Risk attitudes and risky behavior. *Journal of Economic Psychology*, 17, 749–770.

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

**How to cite this article:** Breuer, W., Renerken, T., & Salzman, A. J. (2022). Measuring risk-taking and patience in financial decision making. *Review of Financial Economics*, 40, 97–114. <https://doi.org/10.1002/rfe.1142>