

Hackethal, Andreas; Kirchler, Michael; Laudenbach, Christine; Razen, Michael;
Weber, Annika

Working Paper

On the role of monetary incentives in risk preference elicitation experiments

Working Papers in Economics and Statistics, No. 2020-29

Provided in Cooperation with:

Institute of Public Finance, University of Innsbruck

Suggested Citation: Hackethal, Andreas; Kirchler, Michael; Laudenbach, Christine; Razen, Michael; Weber, Annika (2020) : On the role of monetary incentives in risk preference elicitation experiments, Working Papers in Economics and Statistics, No. 2020-29, University of Innsbruck, Research Platform Empirical and Experimental Economics (eeecon), Innsbruck

This Version is available at:

<https://hdl.handle.net/10419/238253>

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.

On the role of monetary incentives in risk preference elicitation experiments

**Andreas Hackethal, Michael Kirchler, Christine Laudenbach,
Michael Razen, Annika Weber**

Working Papers in Economics and Statistics

2020-29



University of Innsbruck
Working Papers in Economics and Statistics

The series is jointly edited and published by

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

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

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

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

On the role of monetary incentives in risk preference elicitation experiments*

Andreas Hackethal[†] Michael Kirchler[‡] Christine Laudénbach[§]
Michael Razen[‡] Annika Weber[†]

August 30, 2021

Abstract

Incentivized experiments in which individuals receive monetary rewards according to the outcomes of their decisions are regarded as the gold standard for preference elicitation in experimental economics. These task-related real payments are considered necessary to reveal subjects’ “true preferences”. Using a systematic, large-sample approach with three subject pools of private investors, professional investors, and students, we test the effect of task-related monetary incentives on risk preferences in four standard experimental tasks. We find no systematic differences in behavior between and within subjects in the incentivized and non-incentivized regimes. We discuss implications for academic research and for applications in the field.

JEL: C91, D01, D81

Keywords: Risk Preferences, Incentives, Experimental Economics, Risk Aversion.

*We are grateful for helpful comments by Thomas Dohmen, Johannes Hermle, Felix Holzmeister, Alex Imas, Matthias Stefan, and Ruediger Weber on earlier versions of this paper. We thank Youpeng for excellent research assistance. Financial support from the Austrian Science Fund FWF (START-grant Y617-G11 Kirchler, and SFB F63) and the Austrian National Bank (grant 17788) is gratefully acknowledged. We gratefully acknowledge research support from the Leibniz Institute for Financial Research SAFE. There exist no conflicts of interest of any of the authors involved in this project. The study was pre-registered, and the data and replication materials will be made public upon publication of the paper. This project was ethically approved by the IRB at Goethe University Frankfurt. Declarations of interest: none.

[†]Goethe University Frankfurt and Leibniz Institute for Financial Research SAFE, Frankfurt

[‡]University of Innsbruck

[§]Goethe University Frankfurt

Risk is inherent to economic decision-making across many real-life domains, such as investments, health behaviors, or labor supply. As risk preferences are a fundamental determinant of decisions under risk, understanding how individuals' preferences feed into decisions is essential to the study of individual decision-making. As a result, assumptions about individuals' attitudes toward risk are central ingredients in many seminal models in economics and finance (e.g. Markowitz, 1952; Merton, 1969; Pratt, 1964; Barberis et al., 2001; Kahneman and Tversky, 1979). Researchers in the decision sciences, such as economics, finance, and neuroscience, commonly use controlled experiments to assess individuals' willingness to take risks. Experiments are often incentivized, which means that individuals receive *real monetary rewards dependent on the outcome of their decisions*. The rationale behind this procedure is the assumption that individuals reveal their true preferences only if the experimental tasks have salient monetary consequences (Smith, 1976; Harrison, 1994). This practice of using incentives contrasts with practices in other social sciences, most prominently psychology, where non-incentivized, hypothetical choices are common (Camerer and Hogarth, 1999; Hertwig and Ortmann, 2001).

In this paper, we are interested in the role of incentives in the context of risk preference-elicitation tasks in economics. In many experiments, these preferences are either the main variable of interest, or serve as a control variable or a model ingredient (e.g., Kamas and Preston (2012); Alan and Ertac (2018); Saccardo et al. (2018); Thunstroem and Ritten (2019)). Incentivizing these tasks does not only induce substantial (additional) monetary costs – effectively limiting sample sizes – but also increases administrative efforts (Dohmen et al., 2011). Moreover, complicated payoff formulas may unduly increase the complexity of the experimental design and arguably make choices less realistic to subjects (Read, 2005; Bardsley et al., 2020). The rise of online surveys has facilitated the recruitment of subjects on a large scale and opened the possibility of studying choices and preferences among subject pools other than students. Obtaining accurate measures of risk attitudes of non-standard subjects such as private and professional investors is essential to understand their financial behaviors and to gauge their impact on asset prices and the macroeconomy (Guiso and Sodini, 2013). However, using task-related incentives may not always be feasible in these settings.

Against this background, we use a systematic, large-scale approach to study the impact of task-related monetary incentives on experimentally elicited risk-preference measures. Our data set is unique by combining the following important dimensions: (i) we consider four standard experimental tasks instead of focusing on a single task and (ii) we study the choices of private and professional investors in addition to the choices of students. In total, we administer an online experiment of 1,480 participants, among them 821 private investors at a large German bank, 244 professional investors at various financial companies in the EU, and 415

students at the University of Innsbruck. The experimental tasks we consider are the *staircase procedure* by Falk et al. (2016, 2018), the *gamble-choice task* by Eckel and Grossman (2002), the *paired lottery choice task* by Holt and Laury (2002), and the *investment game* by Gneezy and Potters (1997). We randomly assign subjects to two incentive conditions. Respondents in the FLAT condition receive a fixed fee as a reward for participation. Respondents in the INCENTIVES condition – in addition to the fixed participation fee – receive a task-related payment according to the outcome of their decision in one randomly selected experimental task.

In 10 of the 12 in-sample mean comparisons in the main experiment, we find no systematic differences between the participants’ choices in the FLAT and the INCENTIVES condition. Only for the Holt and Laury (2002) elicitation task, we document a small increase in risk aversion for students and professional investors with incentives. Also, the standard deviations and distributions of individuals’ choices in each task by incentive condition do not differ significantly for all comparisons with the only exception being the distribution of switching points in the Holt and Laury (2002) task in the professional investor sample. The propensity to provide an extreme response does not significantly differ by incentive condition across all three subject pools. Moreover, we also find no systematic differences by incentive condition with respect to other aspects of decision quality, such as effort provided. Task-specific response times are similar and while drop-out rates in the experiment differ considerably across subject pools, results reveal no systematic differences in the propensity to drop out by incentive condition.

In two extensions with student samples, we (i) examine the role of incentives within subjects by running the respective other incentive condition six months later and (ii) add a third treatment where subjects receive a fixed participation fee equal to the average payout in the INCENTIVES condition. We are (i) able to confirm our results also in the within-subject analysis, where we again find no significant differences when comparing the average choices subjects made in the INCENTIVES to their choices in the FLAT condition in all four experimental tasks. We also do not find evidence for order effects, which may occur in a within-subject setting. Concerning the payoff, we (ii) find that the absolute level of the fixed participation fee does not alter our results. In particular, we show that student subjects behave virtually identically under two fixed participation fee conditions that vary in payout by a factor of three.

Our results complement prior research investigating hypothetical bias in decisions under risk. Studies of whether and how task-related incentives affect subjects’ risk-taking in economic experiments have produced mixed results. While some investigations find that subjects’ behavior is more risk averse when choices have real consequences (Holt and Laury, 2002, 2005; Harrison et al., 2005), other studies find no differences in subjects’ choices across in-

centive conditions (Beattie and Loomes, 1997; Kuehberger et al., 2002; Gneezy et al., 2015).¹ Furthermore, Smith and Walker (1993) and Camerer and Hogarth (1999) find that incentives lessen the variance of experimental measures. They argue that incentives might help to reduce instances of extreme outliers caused by otherwise inattentive or unmotivated subjects. Comparing real and hypothetical decisions, Camerer and Mobbs (2017) observe differences in brain activity, although not in all of the decision domains studied.² More recently, Etchart-Vincent and l’Haridon (2011) document that differences in the incentive scheme have no effect in the loss domain, whereas incentives matter for risk-taking in the gain domain. We contribute to this literature by providing a *comprehensive picture* of the role of incentives in the experimental elicitation of risk preferences. Rather than focusing on a single experimental task, we run a battery of commonly used risk-elicitation experiments to assess potential effects of incentivization on both location and dispersion. Importantly, with close to 1,500 respondents, our study is high-powered and allows for conclusive inferences even in the case of null results. While we cannot make statements about whether our findings hold with high stake sizes as well, we consider our deliberate focus on low and moderate stake sizes an advantage, as these are a common standard in the literature and increase the generalizability of our study for state-of-the-art procedures.³

Second, we add to the literature on the generalizability of findings obtained in laboratory experiments with standard student subjects (e.g., List, 2003; Haigh and List, 2005; Alevy et al., 2007). In general, existing studies find substantial variation in (risk) preferences across, but also within countries, which suggests that individual characteristics play an important role (e.g., Falk et al., 2018). Given that these characteristics matter for the preferences per se, they may also matter in the responsiveness to incentives in the elicitation process. Besides examining the role of incentives in experiments with students as subjects, we run our experiments with two large, non-standard subject pools of private and professional investors. Given the ramifications for asset prices and the macroeconomy, obtaining valid measures of these subjects’ risk preferences – and the interplay with task-related incentives – is of great interest to academics, regulators, and policy makers.

¹For extensive reviews, see Camerer (1995); Camerer and Hogarth (1999); Camerer and Mobbs (2017); Harrison (2006).

²Closely related, Dohmen et al. (2011) validate a non-incentivized survey item designed to measure individuals’ risk attitudes using a lottery-choice experiment and find that the general risk-attitude question strongly predicts behavior in the experiment. The survey item is also found to generate the best predictor of real-world risky behaviors.

³Encouragingly for our study, Enke et al. (2021) show that very high incentives, increasing the standard lab stakes to more than a monthly income, have hardly any effect on bias-proneness in four classical tasks (e.g., base-rate neglect or anchoring). The authors show that in none of the tasks these high stakes come even close to de-bias subjects.

1 Experimental Design

Experimental tasks

We consider four of the most widely used experimental tasks for eliciting risk preferences, which we introduce below. Table 1 provides details on the parameterization of the gambles involved in each of the tasks. The experimental instructions are provided in Online Appendix D. The euro amounts stated refer to the payoff parameters in the private and professional investor sample. To align stake sizes for all subject pools with the standards in the experimental economics literature, we divide these euro amounts by four to obtain corresponding payoff profiles for the student sample (see, e.g., Haigh and List, 2005; Kirchler et al., 2018).

- (i) The *staircase procedure* by Falk et al. (2016, 2018) (FA) aims to elicit subjects' certainty equivalent for a given lottery in a series of decisions. We ask subjects to choose between a lottery paying €45 or €0 with equal probability, and a safe payment of €24. Subjects who prefer the lottery in the first stage are offered a higher safe payment (€36) in the second decision, whereas subjects who prefer the safe payment are presented a lower safe payment (€12). The payoffs of the lottery remain constant across the decision rounds. In our specification, the payout of the safe alternative varies from €3 to €45. After four decision rounds, the staircase design allows the researcher to pin down a narrow interval for subjects' certainty equivalent as a measure of their risk preference, with certainty equivalents being higher for more risk-tolerant individuals. We provide an exposition of the entire sequence of decisions in the four decision rounds in Figure A1.
- (ii) The *gamble-choice task* by Eckel and Grossman (2002) (EG) asks subjects to choose their preferred lottery specification from a menu of six 50/50 gambles (see Table 1). The first lottery offers a secure payoff of €21 in both states of the world. Subsequently, the difference between the two possible payoffs widens, as the first payoff increases by €6 while the second payoff decreases by €3 in each subsequent lottery. Consequently, the rank of the lottery chosen, ranging from 1 to 6, serves as a measure of a subject's risk tolerance. Subjects with higher risk tolerance will choose lotteries farther down the list, as these offer higher expected returns at higher levels of risk.
- (iii) The *paired lottery choice task* by Holt and Laury (2002) (HL) presents subjects with 10 separate decisions between a lottery A that pays either €24.00 or €19.20 and a lottery B that pays either €46.20 or €1.20 (see Table 1). In the first decision, the probability for the high [low] state is 10% [90%] in both lotteries. In each subsequent decision, the probability of the high [low] state increases [decreases] by 10%. Hence, in each decision,

choosing lottery A is less risky than choosing lottery B. At the same time, however, the expected value of lottery A increases from €19.68 to €24.00, while the expected value of lottery B increases from €5.70 to €46.20. As a measure of individuals' willingness to incur risk, we mark the decision where a subject switches from lottery A to lottery B. Subjects with higher risk tolerance will switch earlier to lottery B.⁴ For ease of interpretation, we count the number of rows after the switching point, such that higher values imply higher risk tolerance.⁵

- (iv) The *investment game* by Gneezy and Potters (1997) (GP) stylizes an investment decision. In this task, subjects receive an initial endowment of €24 and are asked to decide which fraction to invest in a project that pays either 2.5 times the invested amount or €0, with equal probability (see Table 1).⁶ The amount not invested is kept in either state of the world. As is apparent from the parametrization, higher investments increase both expected value and variance of the payoff. We use the amount a subject invests in the risky project as a measure of risk tolerance such that higher values indicate higher levels of risk tolerance. Risk-neutral and risk-seeking subjects will invest their entire endowment.

Treatments and Payment of Subjects

We randomly assign subjects to one of the two treatments in a between-subjects design. In the FLAT condition, subjects receive a fixed fee as a reward for participating in the experiment. The fixed participation fee is €12 for subjects in the private and professional investor samples, and is €3 in the student sample. Respondents are explicitly informed that they will be asked to choose among several options with different *hypothetical* payoff profiles and that the payoffs resulting from their decisions will not actually be paid out. Payment of the fixed participation fee is independent of the choices made in the experiment. In addition to the participation fee, subjects in the INCENTIVES condition are paid the earnings resulting from their choice in one experimental task, which is randomly selected at the end of the experiment. In case the selected task involves a series of decisions, a second random draw determines the decision to

⁴Note that in the 10th decision, the higher payoff will be paid with certainty in both lotteries, so choosing the second lottery is a dominant strategy. This result provides validation of whether subjects have understood the task (Charness et al., 2020).

⁵Note that we allow participants to freely switch between options A and B as they move down the 10 decision rows. This lack of constraint allows for the possibility that individuals may behave inconsistently by choosing to switch between lotteries A and B more than once. We drop these observations from our main analyses but include a specific discussion on a potential association between incentives and the propensity to behave inconsistently in section 2.

⁶We use the same variant as Charness et al. (2020). Note that in the original version of Gneezy and Potters (1997), the probabilities for receiving 2.5 times the invested amount and zero are 1/3 and 2/3, respectively.

be paid out. Subjects are then paid according to their choice and the random outcome of the respective lottery. Subjects in both conditions are presented with the same experimental tasks and experimental instructions (except for minor differences necessary to explain the payment protocols).

Experimental Protocol

To determine the target number of subjects to be recruited, we performed a power analysis following Cohen (1988) for behavioral sciences. We aimed to maximize statistical power for each sample given the particular constraints with respect to recruitment possibilities. Applying the predefined target parameters to our realized sample sizes, our tests have 90%-*a priori* power to detect effect sizes as low as 0.23, 0.42, and 0.32 in mean differences between the INCENTIVES and FLAT condition for the sample of private investors, professional investors, and students, respectively. These numbers are in the range of small and small-to-medium effect sizes, as suggested by Cohen (1988). We provide details on our power analysis in Table C1 in Online Appendix C.

The experiment was administered online using Limesurvey. We recruited subjects for participation via e-mail. Through e-mail, we recruited student participants from the University of Innsbruck using Hroot (Bock et al., 2014). In addition, we invited private investors from a panel of 2,000 clients of a large German brick-and-mortar bank who regularly participate in short online surveys/experiments administered by Goethe University Frankfurt. Third, we recruited professional investors via two channels. Two-thirds of the professional investor sample were recruited from the proprietary subject pool of professional investors (www.before.world) at the University of Innsbruck, some of whom had participated in previous unrelated studies. The remaining professionals are fund managers from different European countries whom we identified via their fund affiliation using data from Morningstar.

To avoid potential selection bias into either of the incentive conditions, we use a standardized invitation letter for all subjects. Subjects learn about the payment protocol relevant for them only upon starting the experiment.

Data and sample characteristics

We collected data for our main analyses in April and May 2020.⁷ Overall, 1,512 subjects completed the experiment. Once the experiment started, we offered unlimited time to finish to avoid exerting time pressure on subjects who were potentially engaged with risk elicitation

⁷The second round of data for the within-subjects analyses among the student sample was collected in November 2020. Descriptive statistics for the robustness analyses (i.e., the student samples for treatment FLAT_HIGH and the within-sample, respectively), are provided in the online appendix.

experiments for the first time. However, to screen out participants who plausibly did not take the experiment seriously and to avoid potential noise due to outliers, we drop subjects in the top (99%) and bottom percentiles (1%) of the response time.⁸ The final sample consists of 1,480 subjects, comprising 821 subjects from the private investor sample, 244 from the financial professional investor sample, and 415 from the student sample. The median response time in the final sample is 13.22 minutes with a standard deviation of 9.88 minutes.

Panel A of Table 2 describes the private investor sample. Respondents in this sample are retail clients at a large German bank with a national branch network and are part of a regular online survey panel.⁹ Of the respondents, 26 percent are female, the average age is 53 years, and private investors' reported average net household income is €4,292. Of the private investors, 79 percent invest in stocks or stock mutual funds.

The private investor sample is well balanced along these characteristics across the two treatment arms, as indicated by the close averages and corresponding high p-values from a two-sided t-test of equal means in column 4 of Table 2 (Panel A). The only significant difference between subjects in the various incentive conditions arises in overall response time, which is significantly higher under the INCENTIVES condition. However, as we show below, this difference does not result from longer decision times in the single experimental tasks in the INCENTIVES condition, but is explained by subjects' need to spend more time reading longer texts outlining the payoff protocol in the INCENTIVES condition. This condition applies equally to all three subsamples.

Panel B of Table 2 describes the professional investors sample. Respondents in this sample are predominantly male (89 percent) and the average age is 42 years. The majority of professional investors are fund managers (35 percent), portfolio managers (18 percent), analysts (10 percent), and risk managers (7 percent). Again, the sample is well balanced across both treatment arms, as indicated by the close averages and corresponding high p-values from a two-sided t-test in column 4 of Table 2 (Panel B) in the Online Appendix.

Panel C of Table 2 characterizes the student sample. Most subjects in the student sample are female (57 percent), and the average age is 24 years. One in four student subjects invests in stocks or stock mutual funds. Both treatments in our student sample are well balanced along these characteristics.

⁸Our results are robust to varying this cutoff, as well as applying the cutoff to the three subsamples individually. The propensity toward excessive response times does not systematically vary with the treatment condition. We use the same cutoffs for students in the FLAT_HIGH.

⁹This panel consists of about 4,000 clients who signed up to regularly participate in online surveys for the purpose of academic research. For the present study, we selected 2,000 clients from the panel, over-sampling clients with an investment account (e.g. bonds, stocks, mutual funds) as opposed to clients holding a current account only. As private investor demographics have been elicited in an earlier survey wave, they are not re-elicited in the present study, a constraint that causes demographics to be missing for a small number of private investors who did not participate in this earlier survey.

2 Results

Risk-taking

Result 1: *Risk-taking of private investors, professional investors, and students does not differ across incentive conditions in three of the four experimental tasks. In the HL task, professional investors and students engage in slightly less risk-taking in the INCENTIVES than in the FLAT condition.*

Support: We start our analysis by comparing mean choices in the four experimental tasks between subjects in the FLAT and INCENTIVES condition. The upper panel in Figure 1 displays average choices by experimental task, incentive condition, and subject pool. Throughout, higher values (higher bars) indicate higher levels of risk tolerance. For FA, we display a value between 1 and 16 according to ordinal ranking of the resulting certainty equivalent. For EG, we display the rank (1 to 6) of the gamble chosen from the menu of six 50/50 gambles. In the HL-task, we present the number of decision rows left after switching to the higher-risk lottery, ranging from 0 to 10. For GP, we show the euro amount invested in the risky investment. To make choices comparable across subject pools, we divide the amount invested in the GP-task in the private and professional investor subsamples by four. GP hence takes on values between 1 and 6.¹⁰ The light (dark) shaded bars represent subjects' choices in the FLAT (INCENTIVES) condition. We report the p-values of two-sided t-tests for equality of mean choices in Table A1 in the Online Appendix.¹¹

We find no significant differences between subjects' choices in the various incentive conditions in three of the four experimental tasks (see Figure 1). In the case of HL, we find that subjects' behavior is slightly more risk-averse in the INCENTIVES condition than under the FLAT regime in the student sample (6.21 vs. 6.62, t-test, p-value=0.008, N=388) and the professional investor sample (6.05 vs. 5.57, t-test, p-value=0.032, N=226). In the lower panel of Figure 1, we display standardized treatment effects sizes, along with 95% confidence intervals. Within each subject pool, we standardize the choices in the different tasks by subtracting the mean and dividing by the standard deviation of the distribution of choices in the respective subject pool. We then regress standardized choices on an indicator variable for whether a subject has been assigned to the INCENTIVES condition. The figure confirms the above finding and illustrates the relative magnitude of the decrease in risk-taking in the HL task on the part of finance professionals and students in the INCENTIVES condition. In both cases, the difference amounts to 0.27 of a standard deviation.

¹⁰Investment amounts can be adjusted in steps of €0.50 (€2.00).

¹¹Importantly, we refrain from measuring risk preferences by calculating risk preference parameter. As our focus is on within-task comparisons, results of risk preference parameter estimates would be very similar to what we report in this paper.

In addition to comparing the means, we also compare variances and distributions of individuals' choices in each task by incentive condition, separately for each subject pool (see Table A2 and Figure A2 in the Online Appendix).¹² We find that F-Tests fail to reject the null of equal standard deviations under the INCENTIVES and the FLAT condition, respectively, for all comparisons. Further, Kolmogorov-Smirnov tests fail to reject the null hypothesis of equal distributions under both incentive regimes across tasks and subject pools, with the only exception being the distributions of the switching point in the HL task in the professional investor sample (KS-test, p-value=0.022, N=226; p-values of KS-tests are reported in Table A1 in the Online Appendix).

Incentives and effort

Result 2a: *In general, effort – as measured by decision times of subjects – does not differ across incentive conditions, tasks, and subject pools. Moreover, while drop-out rates in the experiment differ considerably across subject pools, results reveal no systematic differences in the propensity to drop out by incentive condition.*

Support: Another dimension along which task-related incentives may affect subjects' behavior in the experiments is the effort subjects apply in making decisions. We follow the literature in using decision time when making their choices as a measure of effort (Wilcox, 1993; Camerer and Hogarth, 1999) and we compare decision times in the four tasks by incentive condition. Subjects take on average 0.88 minutes to complete the FA, 1.13 minutes for the EG, 2.58 minutes for the HL, and 1.29 minutes for the GP task. As illustrated in Figure 2, we find no systematic differences in decision times by incentive condition across tasks and subject pools, except for professional investors in the INCENTIVES condition of the EG task (t-test, p-value = 0.01, N=244; see Table A3 in the Online Appendix). If we control for outliers, differences also get insignificant for that subsample. These results imply that the significantly higher *total* time spent in the experiment observed for subjects in the INCENTIVES condition across the three subject pools (see Table 2 in the Online Appendix) results from subjects taking more time to read the details of the task-related payoff protocol rather than from spending more time thinking about their choices.

As a second proxy for individuals' effort, we test whether task-related incentives act to increase subjects' perseverance, reducing the number of participants who exit the experiment prior to completion. Overall, 1,661 subjects started the experiment, of which 1,512 completed it (and 1,480 constitute the final data set after the winsorizing procedure outlined above),

¹²Panel A displays choices of private investors, and Panels B and C refer to professional investors and students, respectively. Light (dark) colored bars represent the choices of subjects in the FLAT (INCENTIVES) condition.

translating into a drop-out rate of 9 percent. We find that drop-out rates differ considerably across subject pools and are substantially lower in the student sample (below 3 percent) than in the private investor and professional investor samples (about 11 percent in both samples). However, in all subject pools, we do not find systematic differences in the propensity to drop out from the experiment across incentive conditions (see Panel A of Figure 3 and Table A3 in the Online Appendix).

Finally, we test whether task-related incentives reduce instances of inconsistent behavior in the arguably complex HL task. In the HL task, we allow participants to switch between options A and B as they move down the 10 decision rows. This opportunity may result in inconsistent behavior as subjects could switch between options more than once.¹³ Across all subject pools, 15 percent of participants make inconsistent choices, which compares rather favorably to other studies (Crosetto and Filippin, 2016; Charness et al., 2013). The prevalence of inconsistent choices differs greatly across subject pools (see Panel B of Figure 3 and Table A3 in the Online Appendix). While the choices of about 7 percent of subjects in the student and professional investor sample produce multiple switching points, 23 percent of respondents in the private investor sample show this kind of behavior. One obvious reason lies in different education levels among the three subject pools. We find the propensity to give inconsistent answers in the private investor sample to be negatively correlated with educational achievement and financial literacy. Among private investors who have completed a college degree (N=401), the share of respondents giving inconsistent answers amounts to 13 percent. Accordingly, Dave et al. (2010) show that less sophisticated subjects have trouble understanding the HL protocol. Importantly, we find no evidence that the incentive regime does affect the propensity to behave inconsistently in the HL task across the three subject pools.¹⁴

Result 2b: *In general, seriousness of subjects – as measured by variation in decisions – does not differ across incentive conditions.*

Support: Smith and Walker (1993) and Camerer and Hogarth (1999) argue that incentives might help to reduce instances of extreme outliers caused by otherwise inattentive or unmotivated subjects and hence lead to lower variance. From a methodological point of view, incentives could thus contribute to higher-quality data and improve statistical power (Camerer and Hogarth, 1999). We test this conjecture by pairwise comparisons of standard deviations across treatments (Table A2). Our results reveal no significant differences between

¹³Remember, a risk-averse agent should start out by choosing option A in the first decision, and then switch to option B at some point before the last decision.

¹⁴Holt and Laury (2002) observe the number of multiple switchers to be lower in real payoff treatments, compared to hypothetical choices. They also observe less inconsistent choices in higher-payoff treatments compared to their baseline real payoff treatment, with the difference being small, however.

incentivized and non-incentivized tasks ($p > 0.05$ for all comparisons). In fact, standard deviations vary much more across samples than across treatments. Using dispersion of results as an indicator for the effort subjects put in answering the risk elicitation tasks, we thus cannot confirm that incentives increase subject commitment.

Finally, we compare the probability of extreme choices for each task by incentive condition. To do so, we define an indicator equal to one if a respondent selects into the lowest or highest risk tolerance category, according to the respective measure. In Figure A3 in the Online Appendix, we illustrate that the propensity to provide an extreme response does not significantly differ by incentive condition at conventional significance levels of 5 percent or higher across all three subject pools, according to two-sided t-tests (see Table A4 in the Online Appendix).

Incentives and consistency in risk-taking across tasks

Result 3: *As measured by the standard deviation of an individual’s standardized choices, intra-subject consistency of risk-taking across the four experimental tasks is unaffected by the incentive condition.*

Support: We investigate whether the incentive regime affects the individual’s consistency of risk-taking (relative to their peers’ decisions) across the four experimental tasks. To calculate each subject’s standard deviation, we first standardize choices in the four experimental tasks by subtracting the mean and dividing by the standard deviation of the distribution of choices in the respective task and in the relevant subject pool. For each subject, we then compute the standard deviation over the subject’s standardized choices in the four tasks.¹⁵ We do not find any evidence that task-related incentives affect the intra-subject standard deviation of choices across the three subject pools (see Figure 4 and Table A3 in the Online Appendix). Interestingly, the intra-subject standard deviation decreases strongly when excluding the HL task from the consistency measure in general (see Figure A4 in the Online Appendix). Singular exclusion of any of the other three choices does not produce a similar effect.

3 Extensions

3.1 Within-subject Results

Until now, our experimental method was based on a between-subject design, randomly allocating subjects to either the INCENTIVES or the FLAT condition. Conditional on a random

¹⁵More precisely, note that this measure combines between and within elements. It shows how strongly a subject’s relative (compared to the peer sample) risk-taking varies across tasks.

assignment between groups, the between-subject approach is perceived to be the more conservative method. Within-subject designs may, however, be favorable in environments where subjects are likely to face repeated decisions (Charness et al. (2012)). Eliciting risk preferences with and without incentives in the same experimental session may decrease incentive effects due to a consistency preferences or experimenter demand effect based on the first decision made.

To overcome this concern, we conduct a second experiment (round) with students who participated in our first experiment (round) six months later in November 2020 (see, for example, Cavallo et al., 2017). We invited all students who participated in the first experiment and gave their consent to participate in a subsequent experiment (the second round, however, was not framed as a follow-up experiment in the narrow sense, as no references to the first round were made and also payments were administered entirely independently). Students who were previously assigned to the INCENTIVES condition were now assigned to the FLAT condition and vice versa. Overall, 213 students followed our invitation (i.e., response rate of 51.3%). For these 213 subjects, we are able to analyze the role of incentives in a within-subject design. To do so, we first replicate results from our between-subject analysis in Figure 1 within subjects. Panel A of Figure 5 shows that we are able to confirm our results also in the within-subject analysis. We find no significant differences if we compare the average choices subjects made in the INCENTIVES to their choices in the FLAT condition in all four experimental tasks. To better understand the heterogeneity across subjects, Panel B of Figure 5 depicts a scatter plot of the incentivized (y-axis) versus non-incentivized choices (x-axis) weighted by the frequency of occurrence. Points on the 45 degree line represent subjects who took the same decision under both treatment conditions. Systematic effects of incentives would be reflected by an overbalance of choices above (below) the 45 degree line, reflecting increased (decreased) risk taking in the presence of incentives. The scatter plots corroborate that there are no systematic incentive effects and that subjects tend towards taking the same or similar decisions under the two treatment conditions.

Finally, we also test for order effects to address potential issues with task recognition among the within sample. We do so by comparing the answers of those subjects being assigned to the INCENTIVES condition in the first wave to those who were assigned to this condition in the second wave (analogously, we analyze order effects for the FLAT condition). As average choices do not significantly differ across waves ($p > 0.05$ for all pairwise comparisons), we find no evidence for order effects.

3.2 Average Payment Amount

In our design, average payments are lower in the FLAT condition than in the INCENTIVES condition (see Table 2). To control for the sensitivity of our results to the *absolute amount* paid to subjects, we also include a third incentive condition in the first experimental round, FLAT_HIGH, for the student sample (i.e., students were randomly allocated to one of the three treatments). Under this regime, student subjects receive a fixed participation fee equal to the average payout of students in the INCENTIVES condition, amounting to €9. Summary statistics for this subsample are provided in Table A5. The table shows that the samples are balanced with regard to students’ personal characteristics.

Figure A5 in the Online Appendix reproduces Figure 1 comparing choices in the student sample by incentive condition, including the third condition FLAT_HIGH. We find that the absolute level of the fixed participation fee does not alter our results. Student subjects behave virtually identically under the FLAT_HIGH and FLAT condition, as illustrated by the rightmost set of bars. Consequently, student subjects in the INCENTIVES condition take risks similar to those in both flat fee conditions FLAT_HIGH or FLAT in the FA, EG, and GP tasks. Again, we cannot make statements about whether our findings hold with more extreme (much higher) stake sizes as well, given that the intention of our study is to test the role of incentives for state-of-the-art procedures in the sense of standard stake sizes.

4 Conclusion

We use a systematic, large-sample approach with three subject pools of private investors, professional investors, and students, to investigate the impact of task-related monetary incentives on risk preferences, elicited in four standard experimental tasks: the *staircase procedure* by Falk et al. (2016, 2018), the *gamble-choice task* by Eckel and Grossman (2002), the *investment game* by Gneezy and Potters (1997), and the *paired lottery choice task* by Holt and Laury (2002). We find no significant differences between the choices of subjects in the different incentive conditions in 10 of the 12 in-sample comparisons across subject pools. Only in the Holt and Laury (2002) task do professional investors and students behave in a slightly more risk averse manner under a task-related incentive regime than under a regime where subjects receive a flat fee for participation. These results do not change when absolute differences in payment amounts are accounted for in the student sample. We also find no systematic differences across incentive conditions with respect to task-specific response times, drop-out rates, inconsistent choice behavior, and within-subject consistency in risk-taking across tasks.

Our analyses so far remain silent as to why the HL task produces the only significant differences between monetarily incentivized and hypothetical choices in some of the tests.

Our results show that exclusion of the HL choice increases within-subject consistency in choices across the four experimental tasks, providing evidence that individuals tend to behave "differently" in this task. Future research could tackle questions as to whether incentives matter more in complex tasks or whether incentives interact with specific features of the tasks, such as the ability to capture risk-seeking behavior, which is particularly inherent to HL, but absent in some other tasks (e.g., EG).¹⁶

Importantly, our results do not necessarily extend to experimental tasks other than the risk-preference elicitation tasks covered. For example, real task-related incentives have been shown to matter in valuation tasks, where subjects regularly overstate their valuation of alternatives or objects if choices are only hypothetical (List and Gallet, 2001). Similarly, numerous studies show that purely hypothetical tasks overstate socially desirable behaviors in subjects, such as altruism, cooperativeness, and patience (see Camerer and Mobbs (2017) for a recent review). In these settings, incentives are an integral part of the experimental design and it is hard to doubt the necessity to incentivize these tasks (Bardsley et al., 2020). We also do not argue that experimental subjects need not be paid at all. While some people may be intrinsically motivated to participate and will respond truthfully to experimental tasks (Read, 2005), payment of a flat reward for participation plausibly increases the willingness to participate and may help reduce selection into participation. As we paid all subjects a fixed reward for participation, we can only hypothesize about the potential effects of fixed participation fees.

Given the importance of risk preferences to researchers, policymakers, and industry professionals, future research should strive to settle the current methodological issues in the elicitation of individuals' attitudes toward risk. This obligation holds for large-scale lab-in-the-field experiments, where preferences are often elicited only as a control variable and where time and money are scarce. It also holds for applied settings, most prominently the elicitation of risk preferences as part of the financial advisory process under MiFID II (see <https://www.esma.europa.eu/policy-rules/mifid-ii-and-mifir>; retrieved July 1, 2020).¹⁷ Thus, having lean experimental protocols that produce accurate measures of individuals' risk preferences is critical. In weighing complexity of the experimental design against the accuracy of

¹⁶See Crosetto and Filippin (2016) for a recent discussion of how characteristics of risk elicitation tasks such as complexity, availability of a safe option, or the range of risk attitudes may affect measured behaviors. To date, the literature contains no discussion of how these factors may interact with the presence or absence of task-related incentives.

¹⁷More generally, several studies have tested the power of preferences using standard experimental tasks to explain individuals' risky behaviors in the lab and in the field. In a recent study, Charness et al. (2020) show that incentivized standard measures of risk preferences are found to have very limited external validity (i.e., field behavior is measured with financial, health, and employment decisions in a representative Dutch population). In addition, those authors show that simpler measures perform better than the more complex measures.

preference measures, an important criterion is that of learning more about potential hypothetical bias in standard risk elicitation tasks. Our results imply that the degree of hypothetical bias is limited in experiments administered online to private and professional investors (with non-task-related incentives).

References

- Alan, Sule, Seda Ertac. 2018. Fostering patience in the classroom: Results from randomized educational intervention. *Journal of Political Economy* **126**(5) 1865 – 1911.
- Alevy, Jonathan E, Michael S Haigh, John A List. 2007. Information cascades: Evidence from a field experiment with financial market professionals. *Journal of Finance* **62**(1) 151–180.
- Barberis, Nicholas, Ming Huang, Tano Santos. 2001. Prospect theory and asset prices. *The Quarterly Journal of Economics* **116**(1) 1–53.
- Bardsley, Nicholas, Robin Cubitt, Graham Loomes, Peter Moffatt, Chris Starmer, Robert Sugden. 2020. *Experimental Economics: Rethinking the Rules*. Princeton University Press, Princeton and London.
- Beattie, Jane, Graham Loomes. 1997. The impact of incentives upon risky choice experiments. *Journal of Risk and Uncertainty* **14**(2) 155–168.
- Bock, Olaf, Ingmar Baetge, Andreas Nicklisch. 2014. hroot: Hamburg registration and organization online tool. *European Economic Review* **71** 117–120. doi: 10.1016/j.eurocorev.2014.07.003.
- Camerer, Colin. 1995. Individual decision making. *Handbook of Experimental Economics* .
- Camerer, Colin, Dean Mobbs. 2017. Differences in behavior and brain activity during hypothetical and real choices. *Trends in Cognitive Sciences* **21**(1) 46–56.
- Camerer, Colin F, Robin M Hogarth. 1999. The effects of financial incentives in experiments: A review and capital-labor-production framework. *Journal of Risk and Uncertainty* **19**(1-3) 7–42.
- Cavallo, Alberto, Guillermo Cruces, Ricardo Perez-Truglia. 2017. Inflation expectations, learning, and supermarket prices: Evidence from survey experiments. *American Economic Journal: Macroeconomics* **9**(3) 1–35.
- Charness, Gary, Thomas Garcia, Theo Offerman, Marie Claire Villeval. 2020. Do measures of risk attitude in the laboratory predict behavior under risk in and outside of the laboratory? *Journal of Risk and Uncertainty* **forthcoming**.
- Charness, Gary, Uri Gneezy, Alex Imas. 2013. Experimental methods: Eliciting risk preferences. *Journal of Economic Behavior & Organization* **87** 43 – 51.

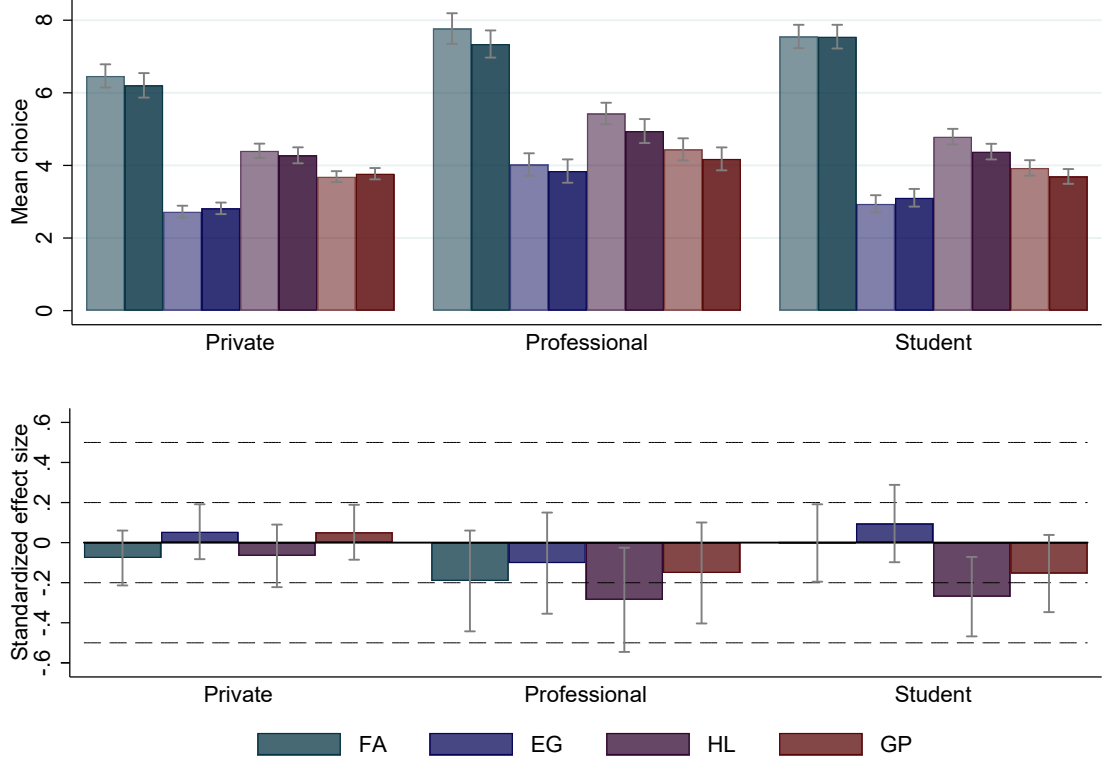
- Charness, Gary, Uri Gneezy, Michael A. Kuhn. 2012. Experimental methods: Between-subject and within-subject design. *Journal of Economic Behavior & Organization* **81** 1 – 8.
- Cohen, Jacob. 1988. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Lawrence Erlbaum Associates, Hillsdale, NJ.
- Crosetto, Paolo, Antonio Filippin. 2016. A theoretical and experimental appraisal of four risk elicitation methods. *Experimental Economics* **19**(3) 613–641.
- Dave, Chetan, Catherine C Eckel, Cathleen A Johnson, Christian Rojas. 2010. Eliciting risk rferences: When is simple better? *Journal of Risk and Uncertainty* **41**(3) 219–243.
- Dohmen, Thomas, Armin Falk, David Huffman, Uwe Sunde, Jürgen Schupp, Gert G. Wagner. 2011. Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association* **9**(3) 522–550.
- Eckel, Catherine C., Philip J. Grossman. 2002. Sex differences and statistical stereotyping in attitudes toward financial risk. *Evolution and Human Behavior* **23**(4) 281 – 295. doi: 10.1016/S1090-5138(02)00097-1.
- Enke, Benjamin, Uri Gneezy, Brian Hall, David C. Martin, Vadim Nelidov, Theo Offerman, Jeroen van de Ven. 2021. Cognitive biases: Mistakes or missing stakes? Working Paper.
- Etchart-Vincent, Nathalie, Olivier l’Haridon. 2011. Monetary incentives in the loss domain and behavior toward risk: An experimental comparison of three reward schemes including real losses. *Journal of Risk and Uncertainty* **42**(1) 61–83.
- Falk, Armin, Anke Becker, Thomas Dohmen, David Huffman David, Uwe Sunde. 2016. The preference survey module: A validated instrument for measuring risk, time, and social preferences Working Paper.
- Falk, Armin, Anke Becker, Thomas Dohmen, Benjamin Enke, David Huffman, Uwe Sunde. 2018. Global Evidence on Economic Preferences. *The Quarterly Journal of Economics* **133**(4) 1645–1692.
- Filippin, Antonio, Paolo Crosetto. 2016. A reconsideration of gender differences in risk attitudes. *Management Science* **62**(11) 3138–3160. doi:10.1287/mnsc.2015.2294.
- Gneezy, Uri, Alex Imas, John List. 2015. Estimating individual ambiguity aversion: A simple approach Working Paper.
- Gneezy, Uri, Jan Potters. 1997. An experiment on risk taking and evaluation periods. *The Quarterly Journal of Economics* **112**(2) 631–645. doi:10.1162/003355397555217.

- Guiso, Luigi, Paolo Sodini. 2013. Household finance: An emerging field. George Constantinides, Milton Harris, René Stulz, eds., *Handbook of the Economics of Finance*, vol. 2. Elsevier, 1397–1532.
- Haigh, Michael S, John A List. 2005. Do Professional Traders Exhibit Myopic Loss Aversion? An Experimental Analysis. *Journal of Finance* **60**(1) 523–534.
- Harrison, Glenn W. 1994. Expected utility theory and the experimentalists. John D Hey, ed., *Experimental Economics*. Springer, 43–73.
- Harrison, Glenn W. 2006. Hypothetical bias over uncertain outcomes. John A List, ed., *Using Experimental Methods in Environmental and Resource Economics*. The name of the publisher, Northampton, MA, 201–213.
- Harrison, Glenn W, Eric Johnson, Melayne M McInnes, E Elisabet Rutström. 2005. Risk aversion and incentive effects: Comment. *American Economic Review* **95**(3) 897–901.
- Hertwig, Ralph, Andreas Ortmann. 2001. Experimental practices in economics: A methodological challenge for psychologists? *Behavioral and Brain Sciences* **24**(3) 383–403.
- Holt, Charles A., Susan K. Laury. 2002. Risk aversion and incentive effects. *American Economic Review* **92**(5) 1644–1655. doi:10.1257/000282802762024700.
- Holt, Charles A, Susan K Laury. 2005. Risk aversion and incentive effects: New data without order effects. *American Economic Review* **95**(3) 902–912.
- Kahneman, Daniel, Amos Tversky. 1979. Prospect theory: An analysis of decision under risk. *Econometrica* **47**(2) 263–291.
- Kamas, Linda, Anne Preston. 2012. The importance of being confident; gender, career choice, and willingness to compete. *Journal of Economic Behavior & Organization* **83**(1) 82 – 97.
- Kirchler, Michael, Florian Lindner, Utz Weitzel. 2018. Rankings and risk-taking in the finance industry. *Journal of Finance* **73**(5) 2271–2302.
- Kuehberger, Anton, Michael Schulte-Mecklenbeck, Josef Perner. 2002. Framing decisions: Hypothetical and real. *Organizational Behavior and Human Decision Processes* **89**(2) 1162 – 1175.
- List, John, Craig Gallet. 2001. What experimental protocol influence disparities between actual and hypothetical stated values? *Environmental & Resource Economics* **20**(3) 241–254.

- List, John A. 2003. Does market experience eliminate market anomalies? *Quarterly Journal of Economics* **118**(1) 41–71.
- Markowitz, Harry. 1952. Portfolio selection. *The Journal of Finance* **7**(1) 77–91.
- Merton, Robert C. 1969. Lifetime portfolio selection under uncertainty: The continuous-time case. *Review of Economics and Statistics* 247–257.
- Pratt, John. 1964. Risk aversion in the small and in the large. *Econometrica* **32**(1–2) 122–136.
- Read, Daniel. 2005. Monetary incentives, what are they good for? *Journal of Economic Methodology* **12**(2) 265–276.
- Saccardo, Silvia, Aniela Pietrasz, Uri Gneezy. 2018. On the size of the gender difference in competitiveness. *Management Science* **64**(4) 1541 – 1554.
- Smith, Vernon L. 1976. Experimental economics: Induced value theory. *American Economic Review* **66**(2) 274–279.
- Smith, Vernon L., James M. Walker. 1993. Monetary rewards and decision costs in experimental economics. *Economic Inquiry* **31**(2) 245–261.
- Thunstroem, Linda, Chian Jones Ritten. 2019. Endogenous attention to costs. *Journal of Risk and Uncertainty* **59**(1) 1 – 22.
- Wilcox, Nathaniel T. 1993. Lottery choice: Incentives, complexity and decision time. *Economic Journal* **103**(421) 1397–1417.

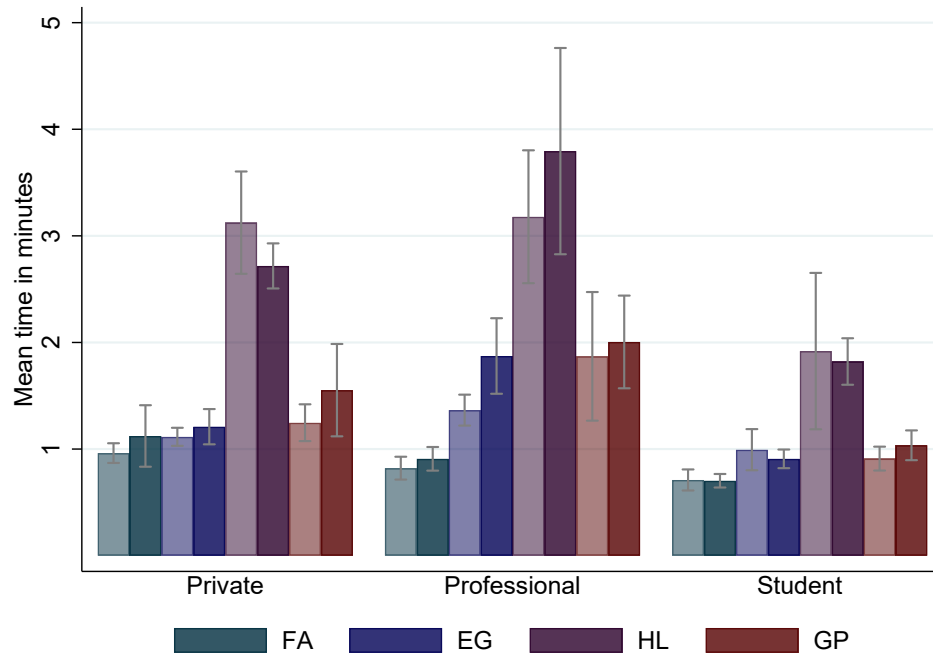
Main figures

Figure 1: Effect of task-related incentives on risk-taking



Notes: Upper panel: Mean choices by subject pool (3 blocks), task (4 colors), and incentive condition (2 shades). The light (dark) shaded bars represent the choices of subjects in the FLAT (INCENTIVES) condition. For all tasks, higher levels indicate greater risk-taking. Owing to the different nature of the underlying tasks, the absolute height of the bars cannot be compared across tasks. FA takes a value between 1 and 16, according to the ordinal rank of the certainty equivalent resulting from the last of the four choices in the staircase risk task. EG is the rank (1-6) of the gamble chosen from a menu of six 50/50 gambles, increasing in risk. HL is the number of decision rows left after switching to the higher-risk lottery, ranging from 0 to 10. GP is the euro amount invested in the risky project and takes values between 0 and 24 for private and professional investors, and values between 0 and 6 for students. For illustration purposes in the above graph, we align these values across samples by dividing the invested amount in the private and professional investor sample by 4. Lower panel: Standardized treatment effects of incentivization on risk-taking, by subject pool and task. Bars represent the coefficient of an indicator variable for whether a subject has been assigned to the INCENTIVES condition in regressions with standardized choices. We standardize choices by deducting the mean and dividing by the standard deviation in the respective subject pool. Error bars indicate 95%-confidence intervals.

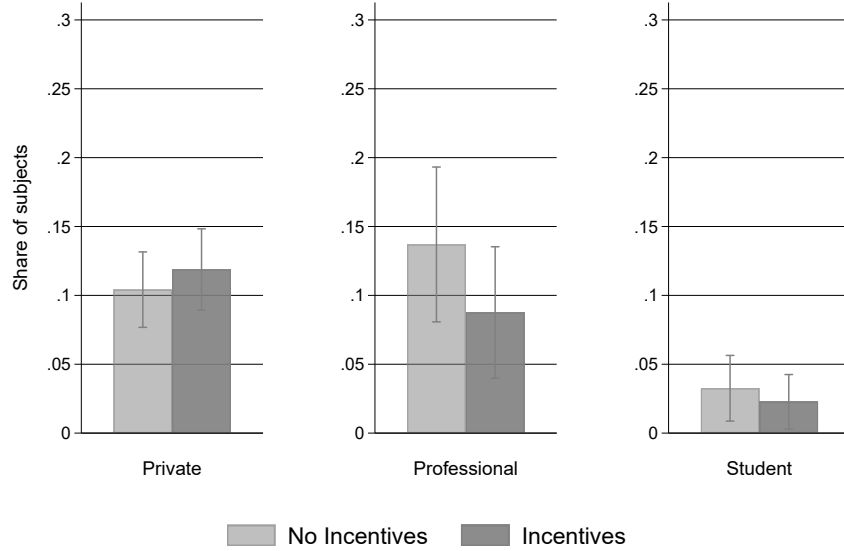
Figure 2: Effect of task-related incentives on decision times



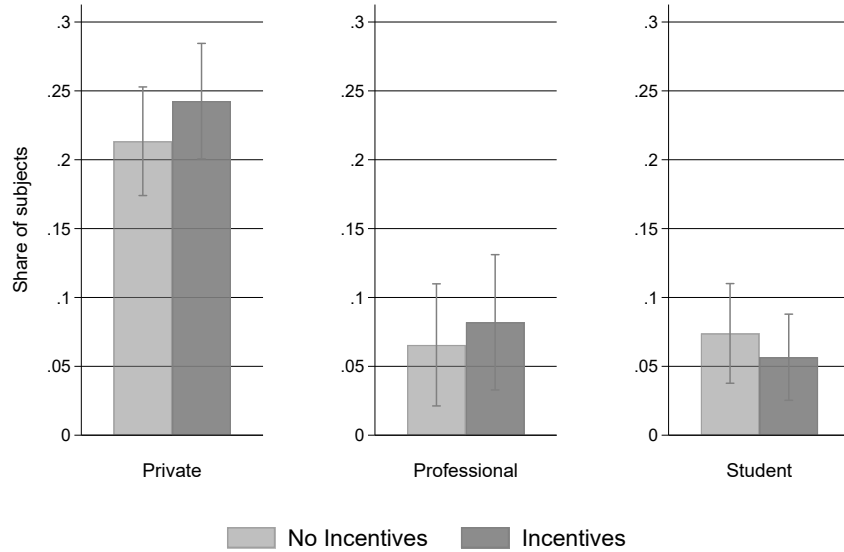
Notes: The figure compares the average decision times in minutes for the four experimental tasks (4 colors) by incentive condition (2 shades) separately for the three subject pools (3 blocks). The light (dark) shaded bars represent the choices of subjects in the FLAT (INCENTIVES) condition. Error bars indicate 95% confidence intervals.

Figure 3: Drop-out rates and inconsistent choices

Panel A: Drop-out rates

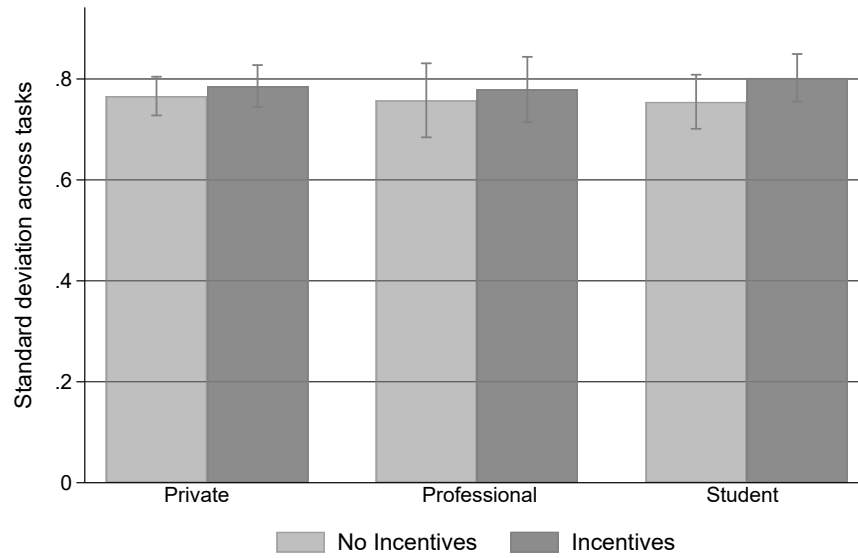


Panel B: Multiple switchers (HL)



Notes: The figure shows the propensity of respondents to drop out from the experiment (Panel A) and the share of subjects with inconsistent answers (multiple switching points) in the HL task (Panel B) by incentive condition, separately for the three subject pools. Drop out rates are calculated based on the overall number of respondents who started the experiment ($N = 1,882$), of which 1,727 completed it. Error bars indicate 95%-confidence intervals.

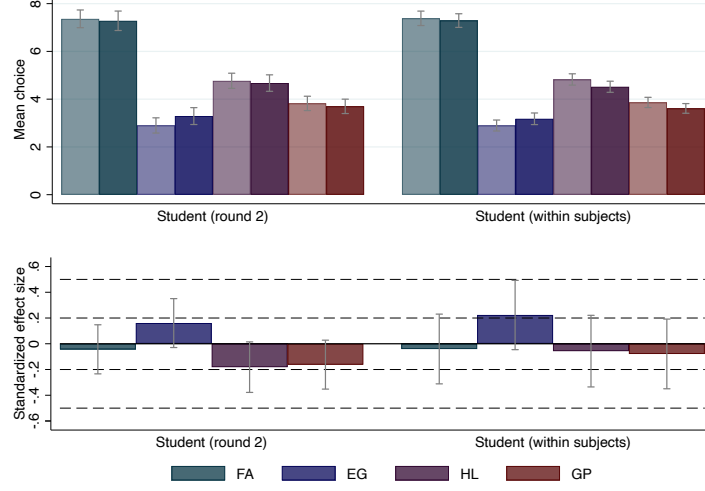
Figure 4: Within-subject consistency in risk-taking across tasks



Notes: The figure compares the mean within-subject standard deviation for the four experimental tasks by incentive condition separately for the three subject pools. We standardize choices in the single tasks by deducting the mean and dividing by the standard deviation of choices made in the given task in the relevant subject pool. We then calculate the within-subject standard deviation over a subject's four standardized choices. The light (dark) shaded bars refer to subjects in the FLAT (INCENTIVES) condition. Error bars indicate 95%-confidence intervals.

Figure 5: Task-related incentives and risk-taking within-subjects

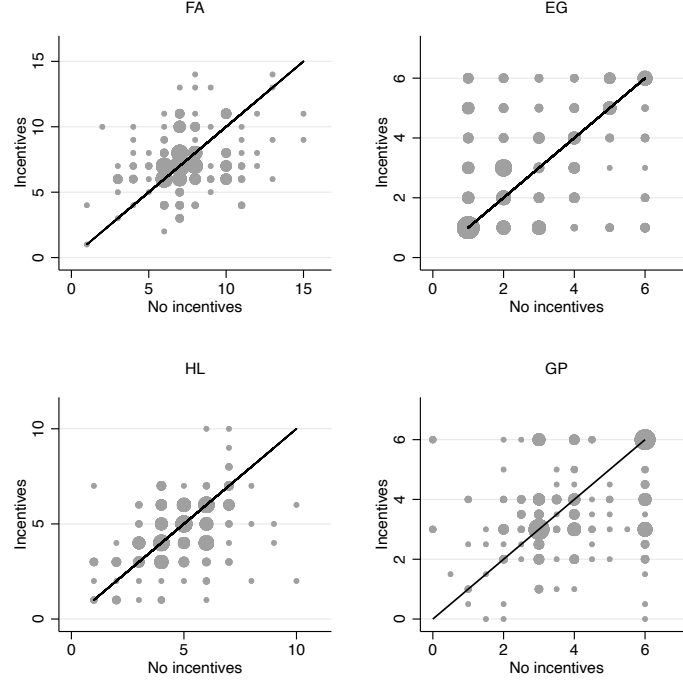
Panel A: Mean Choices and standardized treatment effects



Notes: Upper panel: Mean choices by task (4 colors) and incentive condition (2 shades) for the student sample. The light (dark) shaded bars represent the choices of subjects in the FLAT as well as in the INCENTIVES condition. The order of conditions was counterbalanced. For all tasks, higher levels indicate greater risk-taking. Owing to the different nature of the underlying tasks, the absolute height of the bars cannot be compared across tasks. FA takes a value between 1 and 16, according to the ordinal rank of the certainty equivalent resulting from the last of the four choices in the staircase risk task. EG is the rank (1-6) of the gamble chosen from a menu of six 50/50 gambles, increasing in risk. HL is the number of decision rows left after switching to the higher-risk lottery, ranging from 0 to 10. GP is the euro amount invested in the risky project and takes values between 0 and 6. Lower panel: Standardized treatment effects of incentivization on risk-taking, by task. Bars represent the coefficient of an indicator variable for the INCENTIVES condition. We standardize choices by deducting the mean and dividing by the standard deviation. Error bars indicate 95%-confidence intervals.

Figure 5 (continued): Task-related incentives and risk-taking within-subjects

Panel B: Scatter Plot of incentivized versus non-incentivized choices



Notes: The figure presents scatter plots of the incentivized (y-axis) versus non-incentivized choices (x-axis) for the four risk tasks in the student sample weighted by the frequency of occurrence. Larger circles represent higher frequencies. The black line represents the 45 degree line. FA takes a value between 1 and 16, according to the ordinal rank of the certainty equivalent resulting from the last of the four choices in the staircase risk task. EG is the rank (1-6) of the gamble chosen from a menu of six 50/50 gambles, increasing in risk. HL is the number of decision rows left after switching to the higher-risk lottery, ranging from 0 to 10. GP is the euro amount invested in the risky project and takes values between 0 and 6.

Main tables

Table 1: Overview of risky choices in the single tasks

Task	Choice(s)		
FA	Sequence of four choices between a fixed amount A and a 50/50 gamble B		
	Choice	Option A	Option B
	(1)	€24 with p=1	€45 with p=0.5; €0 with p=0.5
	(2) if (1) = A	€12 with p=1	€45 with p=0.5; €0 with p=0.5
	(2) if (1) = B	€36 with p=1	€45 with p=0.5; €0 with p=0.5

	See Figure A1 in Online Appendix A for all conditional sequences of choices.		
EG	Choice of preferred lottery from a menu of six 50/50 gambles		
	Option 1	€21 with p=0.5; €21 with p=0.5	
	Option 2	€27 with p=0.5; €18 with p=0.5	
	Option 3	€33 with p=0.5; €15 with p=0.5	
	Option 4	€39 with p=0.5; €12 with p=0.5	
	Option 5	€45 with p=0.5; €9 with p=0.5	
	Option 6	€52 with p=0.5; €2 with p=0.5	
HL	Ten separate choices between two lotteries A and B		
	Choice	Option A	Option B
	(1)	€24 with p=0.1; €19.20 with p=0.9	€46.40 with p=0.1; €1.20 with p=0.9
	(2)	€24 with p=0.2; €19.20 with p=0.8	€46.40 with p=0.2; €1.20 with p=0.8
	(3)	€24 with p=0.3; €19.20 with p=0.7	€46.40 with p=0.3; €1.20 with p=0.7
	(4)	€24 with p=0.4; €19.20 with p=0.6	€46.40 with p=0.4; €1.20 with p=0.6
	(5)	€24 with p=0.5; €19.20 with p=0.5	€46.40 with p=0.5; €1.20 with p=0.5
	(6)	€24 with p=0.6; €19.20 with p=0.4	€46.40 with p=0.6; €1.20 with p=0.4
	(7)	€24 with p=0.7; €19.20 with p=0.3	€46.40 with p=0.7; €1.20 with p=0.3
	(8)	€24 with p=0.8; €19.20 with p=0.2	€46.40 with p=0.8; €1.20 with p=0.2
	(9)	€24 with p=0.9; €19.20 with p=0.1	€46.40 with p=0.9; €1.20 with p=0.1
(10)	€24 with p=1	€46.40 with p=1	
GP	Decision what fraction of €24 to invest in a project that pays 2.5 times the amount invested or 0 with equal probability.		
	€24 - € invested + 2.5 × € invested, with p=0.5		
	€24 - € invested, with p=0.5		

Notes: The table presents the gambles involved in the four risk preference elicitation tasks. Euro values stated refer to the parametrization in the private and professional investor sample. For subjects in the student sample, all values are divided by 4. FA takes a value between 1 and 16, according to the certainty equivalent resulting from the last of the four choices in the staircase risk task. EG is the rank (1-6) of the gamble chosen from a menu of six 50/50 gambles, increasing in risk. HL is the number of decision rows left after switching to the higher-risk lottery, ranging from 0 to 10. GP is the EUR amount invested in the risky project and takes values between 0 and 6. Investment amounts can be adjusted in steps of €2. Higher values imply higher risk tolerance across all four tasks.

Table 2: Descriptive statistics

Panel A. Private investors					
	ALL	INCENTIVES	FLAT		
	Mean (SD)	Mean	Mean	P-value	Obs.
	(1)	(2)	(3)	(4)	(5)
Female	0.25 (0.44)	0.28	0.23	0.102	809
Age	52.98 (15.67)	53.53	52.44	0.323	809
HH net income	4,250 (2,402)	4,151	4,347	0.288	676
Stock investor	0.79 (0.41)	0.79	0.79	0.949	821
Smartphone	0.19 (0.39)	0.18	0.19	0.750	821
Total time	16.33 (10.07)	17.33	15.37	0.005	821
Payoff	21.71 (16.38)	31.72	12.00	0.000	821
Panel B. Professional investors					
	ALL	INCENTIVES	FLAT		
	Mean (SD)	Mean	Mean	P-value	Obs.
	(1)	(2)	(3)	(4)	(5)
Female	0.11 (0.31)	0.11	0.11	1.000	244
Age	41.89 (9.50)	42.19	41.59	0.624	244
<i>Job position</i>					
- Fund manager	0.35	0.37	0.34	0.594	244
- Portfolio manager	0.18	0.15	0.20	0.241	244
- Analyst	0.10	0.09	0.11	0.528	244
- Risk manager	0.07	0.08	0.06	0.453	244
Smartphone	0.05 (0.22)	0.03	0.07	0.238	244
Total time	20.46 (12.41)	22.77	18.15	0.003	244
Payoff	21.72 (17.17)	31.43	12.00	0.000	244

Table 2: cont.

Panel C. Students

	ALL	INCENTIVES	FLAT		
	Mean (SD)	Mean	Mean	P-value	Obs.
	(1)	(2)	(3)	(4)	(5)
Female	0.57 (0.50)	0.58	0.54	0.159	415
Age	23.94 (3.22)	24.16	23.71	0.220	415
Net income	748.88 (370.51)	727.14	772.54	0.167	403
Stock investor	0.25 (0.43)	0.23	0.26	0.480	415
Smartphone	0.04 (0.19)	0.03	0.04	0.551	415
Total time	13.45 (8.34)	14.35	12.52	0.026	415
Payoff	5.38 (4.02)	7.67	3.00	0.000	415

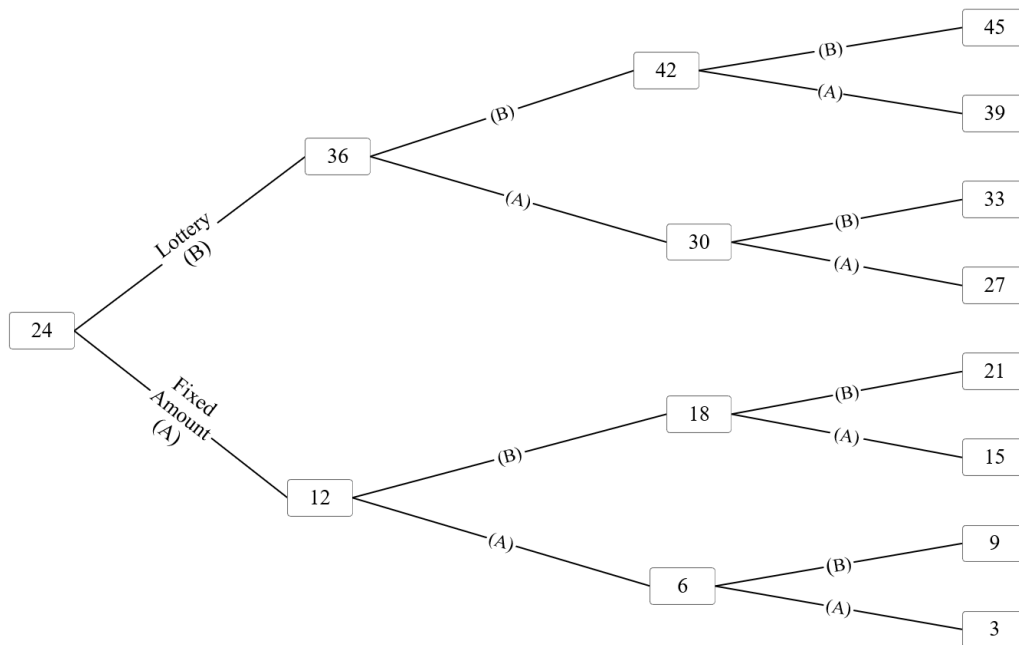
Notes: Panel A shows summary statistics for the 821 respondents in the private investor sample. Panel B shows summary statistics for the 244 respondents in the finance professionals sample. Panel C shows summary statistics for the 628 participants in the student sample. Information of respondents' household net income is only available for 676 of the 821 respondents in the private investor sample, due to non-response. Similarly, 12 of the 415 students choose "prefer not to answer". Note that private investor demographics have been elicited in an earlier survey wave and were not re-elicited for the purpose of our study. This causes demographics to be missing for a number of private investors who did not participate in this earlier survey. Stock investor is an indicator equal to one for participants who invest in stocks or stock mutual funds. Smartphone is an indicator of whether the respondent has participated in the experiment using a smartphone. Total time is the time (in minutes) a subject has spend to complete the entire experiment. Payoff is the final payoff participants receive after completing the experiment. It is fixed in the FLAT conditions. For subjects in the INCENTIVES condition, it depends on the choice and resulting outcome in one randomly determined experimental task. Task-related payoffs in the INCENTIVES condition are in addition to the fixed participation fee paid to subjects in the FLAT condition. Potential payoffs in the student sample result from dividing payoff options presented to private investors and finance professionals by 4. Column 4 reports p-values from a two-sided t-test of equal means between subjects in the INCENTIVES and FLAT condition.

Online Appendix: On the role of monetary incentives in risk preference elicitation

Andreas Hackethal^{1,2}, Michael Kirchler³, Christine Laudenbach¹,
Michael Razen³, Annika Weber^{1,2}

A Additional Figures

Figure A1: Sequence of choices in the staircase procedure FA



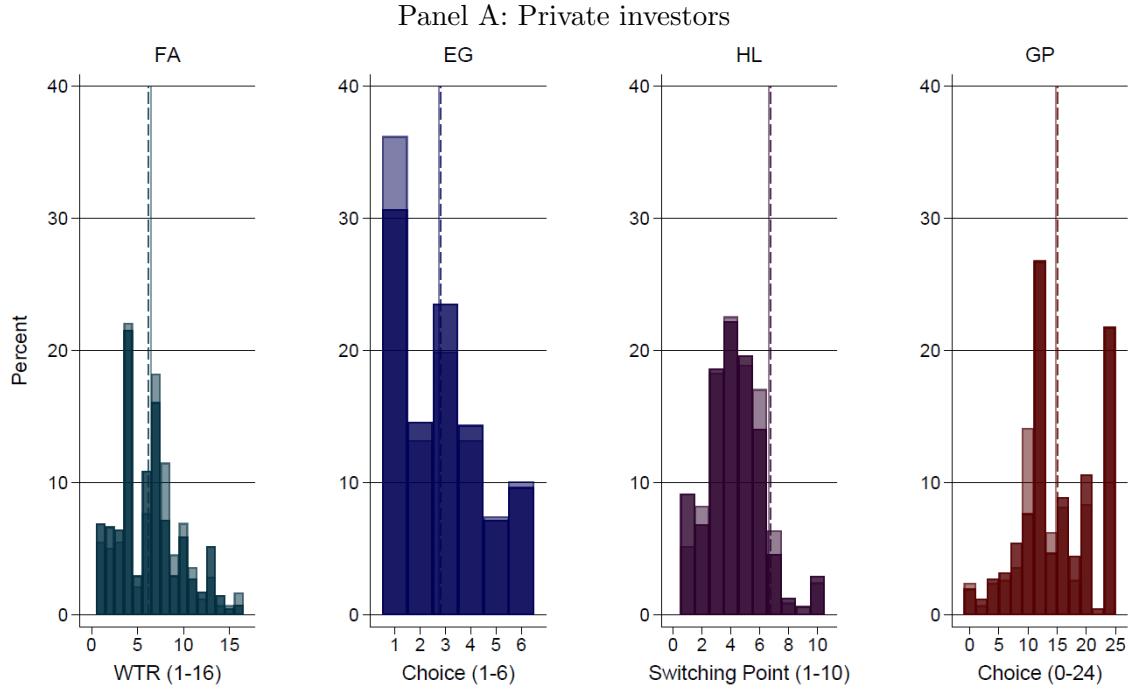
Notes: The staircase procedure following FA works as follows. First, each respondent was asked whether they would prefer to receive 24 EUR for sure or whether they preferred a 50:50 chance of receiving 45 EUR or nothing. In case the respondent opted for the fixed amount (A), the fixed amount being offered in the second question decreased to 12 EUR. If, on the other hand, the respondent opted for the lottery (B), the safe amount was increased to 36. Working further through the tree follows the same logic.

¹Goethe University Frankfurt

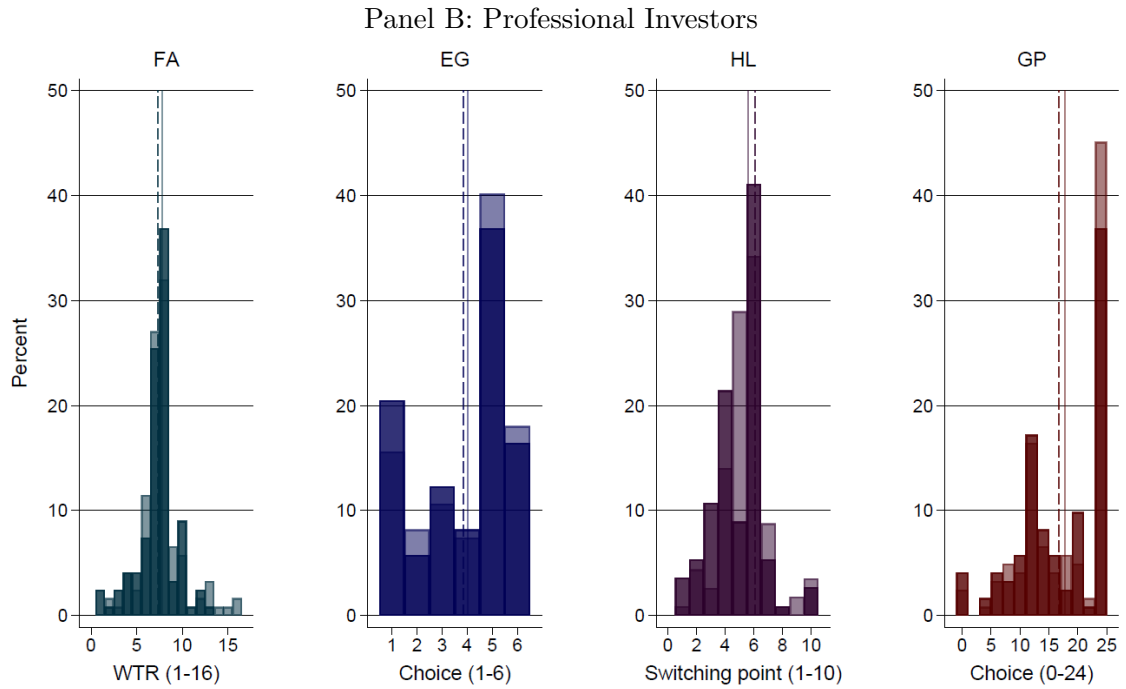
²Leibniz Institute for Financial Research SAFE, Frankfurt

³University of Innsbruck

Figure A2: Effect of task-related incentives on risk-taking - Distributions



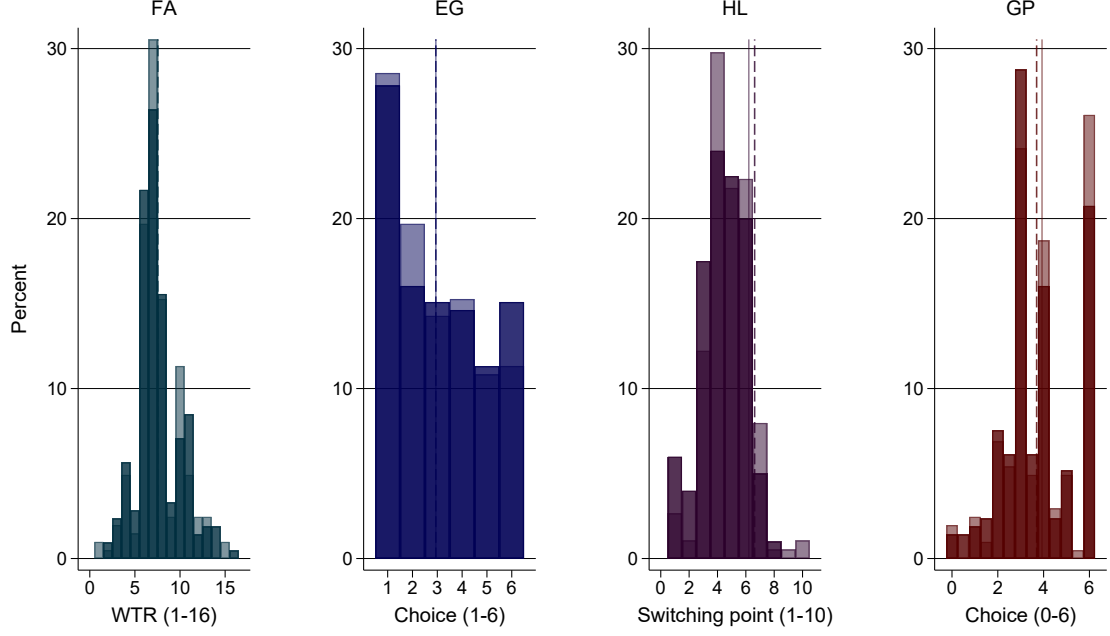
The light (dark) shaded bars represent the choices of subjects in the FLAT (INCENTIVES) condition.



The light (dark) shaded bars represent the choices of subjects in the FLAT (INCENTIVES) condition.

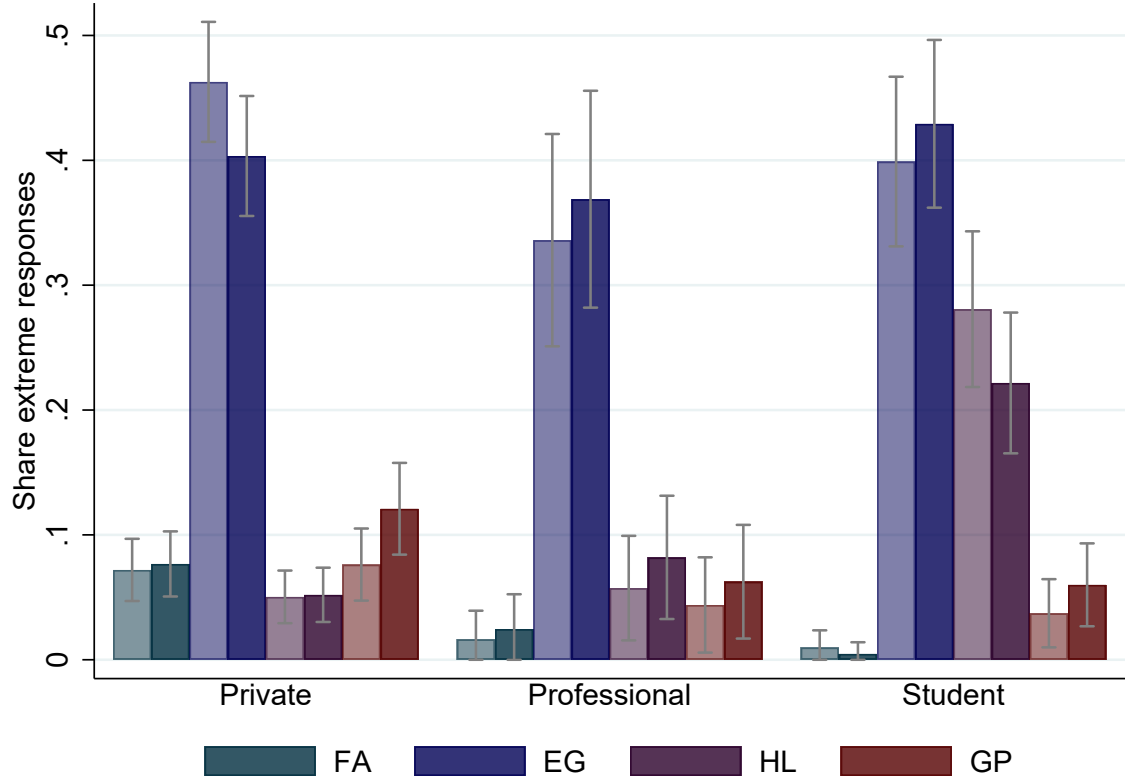
Figure A2 (continued): Effect of task-related incentives on risk-taking - Distributions

Panel C: Students



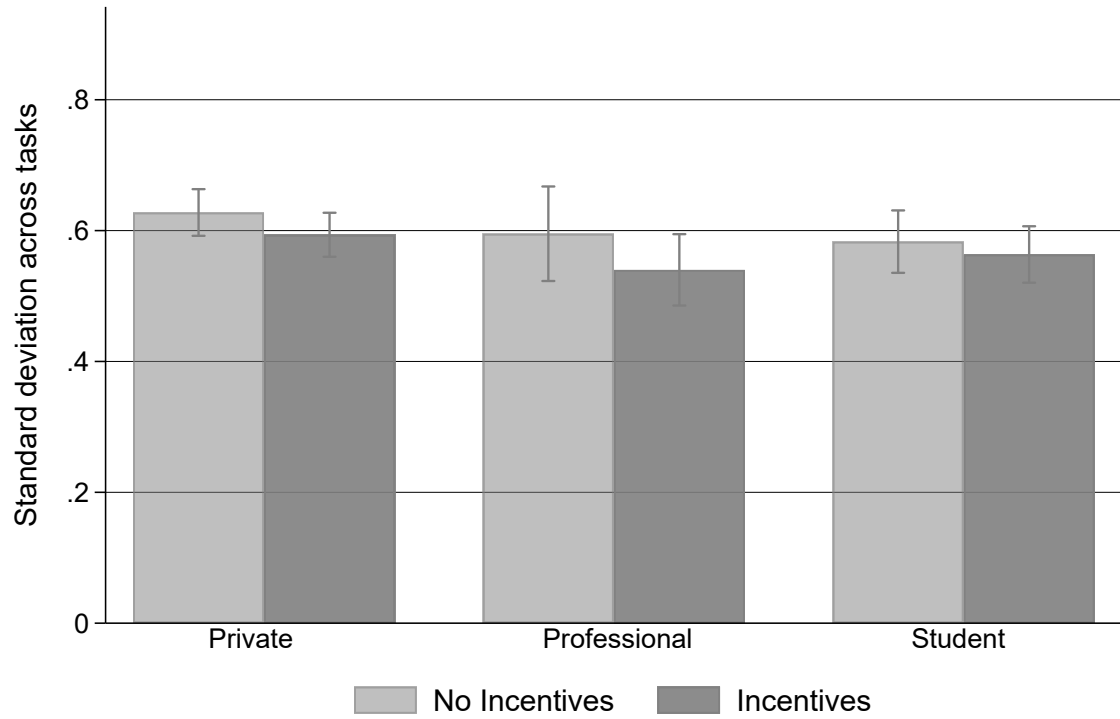
Notes: The figure shows the distributions of respondents' choices in the four experimental tasks (FA in emerald, EG in blue, HL in purple, and GP in maroon) by incentive condition, separately for each subject pool. The light (dark) shaded bars represent the choices of subjects in the FLAT (INCENTIVES) condition. The solid (dashed) line represents the mean in the FLAT (INCENTIVES) condition. In the FLAT condition, subjects receive a fixed participation reward only, amounting to €12 for private investors and professional investors, and to €3 for students. Subjects in the INCENTIVES condition *in addition* are paid the earnings resulting from their choice in one randomly determined experimental task. Panel A represents choices in the private investor sample. Panel B (C) represents choices in the professional investor (student) sample. FA takes a value between 1 and 16, according to the certainty equivalent resulting from the last of the four choices in the staircase risk task. EG is the rank (1-6) of the gamble chosen from a menu of six 50/50 gambles, increasing in risk. HL is the number of decision rows left after switching to the higher-risk lottery, ranging from 0 to 10. GP is the EUR amount invested in the risky project and takes values between 0 and 24 (0 and 6 for students). We report the p-values of Kolmogorov-Smirnov tests for equality of distributions in Table A1.

Figure A3: Task-related incentives and extreme choices



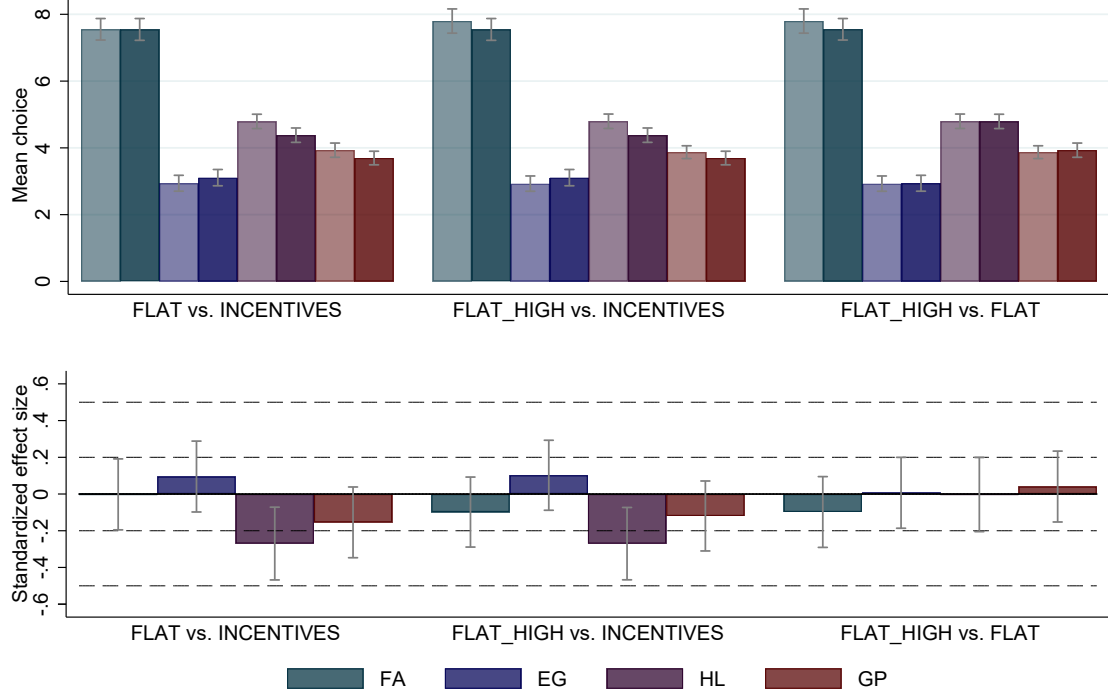
Notes: The figure shows the share of respondents who choose the least or the most risky option in the four experimental tasks (FA in emerald, EG in blue, HL in purple, and GP in maroon) by incentive condition, separately for the three subject pools. The light (dark) shaded bars represent the choices of subjects in the FLAT (INCENTIVES) condition. In the FLAT condition, subjects receive a fixed participation reward only, amounting to €12 for private investors and professional investors, and to €3 for students. Subjects in the INCENTIVES condition *in addition* are paid the earnings resulting from their choice in one randomly determined experimental task. The left block of bars represents probabilities of extreme choices in the private investor sample. The middle (right) block represents probabilities of extreme choices in the professional investor (student) sample. Error bars indicate 95%-confidence intervals. We report the p-values of two-sided t-tests for equality of mean choices in Table A4. Note that the higher share of extreme answers for EG compared to the other tasks is not surprising as there are only 6 possible choices. Hence, the 2 extreme answers account for 1/3 of the decision space.

Figure A4: Consistency in responses across tasks – excluding HL



Notes: The figure compares the mean within-subject standard deviation for three of the four experimental tasks - excluding HL - by incentive condition separately for the three subject pools. We standardize choices in the single tasks by deducting the mean and dividing by the standard deviation of choices made in the given task in the relevant subject pool. We then calculate the within-subject standard deviation over a subject's three standardized choices. The light (dark) shaded bars refer to subjects in the FLAT (INCENTIVES) condition. In the FLAT condition, subjects receive a fixed participation reward only, amounting to €12 for private investors and professional investors, and to €3 for students. Subjects in the INCENTIVES condition *in addition* are paid the earnings resulting from their choice in one randomly determined experimental task. Error bars indicate 95%-confidence intervals.

Figure A5: Effect of task-related incentives on risk-taking - Student sample



Notes: The upper panel shows pairwise comparisons of the mean choices in the four different tasks (FA in emerald, EG in blue, HL in purple, and GP in maroon) across subjects in the three different incentive conditions in the student subject pool. The left block compares subjects in the FLAT condition (light shaded bars) to subjects in the INCENTIVES condition (dark shaded bars). The middle block compares subjects in the FLAT_HIGH condition (light shaded bars) to subjects in the INCENTIVES condition (dark shaded bars). The right block compares subjects in the FLAT_HIGH condition (light shaded bars) to subjects in the FLAT condition (dark shaded bars). Subjects in the FLAT (FLAT_HIGH) condition receive a fixed participation fee of €3 (€9). Subjects in the INCENTIVES condition receive a fixed participation fee of €3 plus the earnings resulting from their choice in one randomly determined task. FA takes a value between 1 and 16, according to the certainty equivalent resulting from the last of the four choices in the staircase risk task. EG is the rank (1-6) of the gamble chosen from a menu of six 50/50 gambles, increasing in risk. HL is the number of decision rows left after switching to the higher-risk lottery, ranging from 0 to 10. GP is the EUR amount invested in the risky project and takes values between 0 and 6. Investment amounts can be adjusted in steps of €0.50. Higher values imply higher risk tolerance across all four tasks. The lower panel shows standardized treatment effects. We standardize the choices of student subjects in the four different task by subtracting the mean and dividing by the standard deviation of the distribution of choices in the subject pool. We then regress standardized choices on an indicator for whether a subject has been assigned to the comparison incentive condition (INCENTIVES for the left and middle blocks, FLAT for the right block). Error bars indicate 95%-confidence intervals.

B Additional Tables

Table A1: Effect of task-related incentives on risk-taking

Panel A. Private investors

	FA	EG	HL	GP
FLAT	6.47	2.73	4.40	14.76
INCENTIVES	6.21	2.82	4.27	15.09
Δ	-0.26	0.09	-0.12	0.33
t-test (p-value)	0.272	0.436	0.406	0.460
KS test (p-value)	0.177	0.525	0.956	0.583
Observations	821	821	634	821

Panel B. Professional investors

	FA	EG	HL	GP
FLAT	7.77	4.02	5.43	17.77
INCENTIVES	7.34	3.84	4.95	16.72
Δ	-0.43	-0.18	-0.48	-1.05
t-test (p-value)	0.135	0.425	0.032	0.237
KS test (p-value)	0.983	0.998	0.022	0.652
Observations	244	244	226	244

Panel C. Students

	FA	EG	HL	GP
FLAT	7.55	2.94	4.79	3.93
INCENTIVES	7.55	3.11	4.38	3.70
Δ	0.00	0.17	-0.41	-0.24
t-test (p-value)	0.984	0.333	0.008	0.116
KS test (p-value)	0.996	0.983	0.125	0.345
Observations	415	415	388	415

Notes: The table reports differences in risk-taking in the four different experimental task by incentive condition and subject pool. In the FLAT condition, subjects receive a fixed participation reward only, amounting to €12 for private investors and professional investors, and to €3 for students. Subjects in the INCENTIVES condition in addition are paid the earnings resulting from their choice in one randomly determined experimental task. Panels A, B and C show differences by incentive condition for the 821 subjects in the private investor sample, the 244 subjects in the professional investor sample, and the 638 respondents in the student sample, respectively. Mean differences that are significant at least at the 5 percent level are printed in bold. We report p-values of a two sided t-test of equal means and p-values of a Kolmogorov-Smirnov test of equal distributions.

Table A2: Standard deviation of choices

	All	Private investors			Professional investors			Students		
	INCENTIVES	FLAT	P-value	INCENTIVES	FLAT	P-value	INCENTIVES	FLAT	P-value	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Std. Dev.										
FA	3.02	3.45	3.32	-0.13 (0.45)	2.09	2.35	0.26 (0.20)	2.41	2.33	-0.08 (0.61)
EG	1.75	1.63	1.69	0.06 (0.46)	1.80	1.73	-0.07 (0.67)	1.80	1.72	-0.08 (0.52)
HL	1.77	1.95	1.82	-0.14 (0.19)	1.76	1.60	-0.16 (0.30)	1.54	1.49	-0.05 (0.63)
GP	1.60	1.59	1.57	-0.02 (0.79)	1.76	1.69	-0.07 (0.67)	1.51	1.54	0.03 (0.75)
Obs (FA/EG/GP)	1,480	417	404	821	122	122	244	203	212	415
Obs (HL)	1,248	328	306	634	114	112	226	188	200	388

Notes: The table reports standard deviations of the choices in the four different tasks, by incentive condition and subject pool. FA takes a value between 1 and 16, according to the ordinal rank of the certainty equivalent resulting from the last of the four choices in the staircase risk task. EG is the rank (1-6) of the gamble chosen from a menu of six 50/50 gambles, increasing in risk. HL is the number of decision rows left after switching to the higher-risk lottery, ranging from 0 to 10. GP is the euro amount invested in the risky project and takes values between 0 and 24 for private and professional investors, and values between 0 and 6 for students. For comparability, we align these values across samples by dividing the invested amount in the private and professional investor sample by 4. In the FLAT condition, subjects receive a fixed participation reward only, amounting to €12 for private investors and professional investors, and to €3 for students. Subjects in the INCENTIVES condition in addition are paid the earnings resulting from their choice in one randomly determined experimental task. We report p-values of an F-Test on the equality of standard deviations for each sample.

Table A3: Effort and consistency of choices across tasks

	All	Private investors			Professional investors			Students		
		INCENTIVES	FLAT	P-value	INCENTIVES	FLAT	P-value	INCENTIVES	FLAT	P-value
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Decision time										
FA	0.88	1.12	0.96	-0.16 (0.29)	0.91	0.82	-0.09 (0.26)	0.70	0.71	0.01 (0.90)
EG	1.13	1.21	1.11	-0.09 (0.31)	1.87	1.37	-0.51** (0.01)	0.91	0.99	0.09 (0.42)
HL	2.58	2.72	3.12	0.41 (0.13)	3.79	3.18	-0.62 (0.29)	1.82	1.92	0.10 (0.80)
GP	1.29	1.55	1.25	-0.31 (0.19)	2.00	1.87	-0.14 (0.72)	1.04	0.91	-0.12 (0.17)
HL: multiple switchers	0.15	0.24	0.21	-0.03 (0.32)	0.08	0.07	-0.02 (0.63)	0.06	0.07	0.02 (0.48)
Drop-out rate	0.08	0.12	0.10	-0.01 (0.48)	0.09	0.14	0.05 (0.19)	0.02	0.03	0.01 (0.53)
Indiv. SD	0.78	0.79	0.77	-0.02 (0.50)	0.78	0.76	-0.02 (0.67)	0.80	0.75	-0.05 (0.19)
Observations	1,480	404	417	821	122	122	244	212	203	415

Notes: The table reports differences in task-specific decision times in minutes, frequencies of multiple switching in the HL task, drop-out rates as well as the within-subject standard deviation of an individual's choices across the four tasks by incentive condition and subject pool. In the FLAT condition, subjects receive a fixed participation reward only, amounting to €12 for private investors and professional investors, and to €3 for students. Subjects in the INCENTIVES condition in addition are paid the earnings resulting from their choice in one randomly determined experimental task. Drop out rates are calculated based on the overall number of respondents who started the experiment ($N = 1,661$), of which 1,512 completed it. To calculate the within-subject standard deviation, we standardize subjects' choices in the four experimental tasks by subtracting the mean and dividing by the standard deviation of the distribution of choices in the respective task in the relevant subject pool. For each subject, we then calculate the standard deviation of the standardized choices in the four experimental tasks. We report p-values of two sided t-tests of equal means.

Table A4: Probability of extreme choices

	All	Private investors			Professional investors			Students		
	INCENTIVES	FLAT	P-value	INCENTIVES	FLAT	P-value	INCENTIVES	FLAT	P-value	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Extreme choice										
FA	0.04	0.08	0.07	-0.00 (0.79)	0.02	0.02	-0.01 (0.65)	0.00	0.01	0.01 (0.54)
EG	0.41	0.40	0.46	0.06 (0.09)	0.37	0.34	-0.03 (0.59)	0.43	0.40	-0.03 (0.53)
HL	0.07	0.12	0.08	-0.04 (0.06)	0.06	0.04	-0.02 (0.53)	0.06	0.04	-0.02 (0.30)
GP	0.12	0.05	0.05	-0.00 (0.92)	0.08	0.06	-0.02 (0.45)	0.22	0.28	0.06 (0.17)
Observations	1,480	404	417	821	122	122	244	212	203	415

Notes: The table reports differences in the propensity to choose the least or the most risky option in the four different tasks, by incentive condition and subject pool. In the FLAT condition, subjects receive a fixed participation reward only, amounting to €12 for private investors and professional investors, and to €3 for students. Subjects in the INCENTIVES condition in addition are paid the earnings resulting from their choice in one randomly determined experimental task. We report p-values of a two sided t-test of equal means for each sample.

Table A5: Descriptive statistics: Student sample with high flat fee

	ALL	INCENTIVES	FLAT_HIGH		
	Mean (SD)	Mean	Mean	P-value	Obs.
	(1)	(2)	(3)	(4)	(5)
Female	0.58 (0.49)	0.58	0.59	0.762	425
Age	24.11 (3.54)	24.16	24.06	0.546	425
Net income	738.13 (373.73)	727.14	749.28	0.8788	417
Stock investor	0.23 (0.42)	0.23	0.23	0.887	425
Smartphone	0.04 (0.18)	0.03	0.04	0.804	425
Total time	13.08 (6.87)	14.35	11.81	0.000	425
Payoff	8.34 (8.34) (3.30)	7.67	9.00	0.000	425

Notes: This table shows summary statistics for the 425 participants in the INCENTIVES versus FLAT_HIGH conditions in the student sample. Information of respondents' household net income is only available for 417 students. Stock investor is an indicator equal to one for participants who invest in stocks or stock mutual funds. Smartphone is an indicator of whether the respondent has participated in the experiment using a smartphone. Total time is the time (in minutes) a subject has spend to complete the entire experiment. Payoff is the final payoff participants receive after completing the experiment. It is fixed in the FLAT_HIGH condition. For subjects in the INCENTIVES condition, it depends on the choice and resulting outcome in one randomly determined experimental task. Task-related payoffs in the INCENTIVES condition are in addition to the fixed participation fee paid to subjects in the relevant FLAT_HIGH condition. Column 4 reports p-values from a two-sided t-test of equal means between subjects in the INCENTIVES and FLAT_HIGH condition.

Table A6: Order effects in student within-sample

	INCENTIVES			FLAT		
	Round 1 (1)	Round 2 (2)	P-value (3)	Round 1 (4)	Round 2 (5)	P-value (6)
Average choice						
FA	7.30	7.28	-0.02 (0.95)	7.41	7.36	-0.04 (0.89)
EG	3.07	3.29	0.22 (0.38)	2.89	2.90	0.01 (0.98)
HL	4.38	4.67	0.30 (0.21)	4.89	4.77	-0.12 (0.61)
GP	3.53	3.70	0.17 (0.42)	3.91	3.82	-0.09 (0.70)
Obs (FA/EG/GP)	110	103	213	103	110	213
Obs (HL)	106	98	204	91	104	195

Notes: The table reports average choices in the four different tasks, by incentive condition and wave. In the FLAT condition, subjects receive a fixed participation reward only, amounting to €12 for private investors and professional investors, and to €3 for students. Subjects in the INCENTIVES condition in addition are paid the earnings resulting from their choice in one randomly determined experimental task. We report p-values of a t-Test on the equality of means for each sample.

C Additional Analyses

Details on the power analysis

Our goal is to provide the basis for conclusive inference, including the case where we would not reject the null hypothesis of no effect of monetary incentivization. To this end, we seek to achieve a sufficiently high power for our statistical tests. In the absence of an indication of how large the effect of incentivization on average choices might be, we followed the reference points suggested by Cohen (1988) for behavioral sciences. As a lower bound for our analysis, we sought a probability of 90% to detect a ‘small’ effect size of less than 0.5 of a standard deviation.⁴ This result corresponds to a sample size of at least $N = 85$ per incentive condition. For the most difficult to recruit sample, that of professional investors, we targeted a sample size of $N = 100$. For the samples of students and private investors, which are easier to recruit, we aimed at sample sizes that would allow us to detect even smaller effect sizes of $d = 0.33$ and $d = 0.20$ (the latter being the lowest threshold for power analysis as suggested by Cohen (1988)), respectively, rounding to targeted sample sizes of $N = 200$ and $N = 500$.

Table C1: Power analysis

	INCENTIVES vs FLAT
Private investors ($N_1 = 404$, $N_2 = 417$)	0.226
Professional investors ($N_1 = 122$, $N_2 = 122$)	0.415
Students ($N_1 = 212$, $N_2 = 203$)	0.318

Notes: The table provides an overview of the number of subjects in each incentive condition and reports the corresponding effect sizes (Cohen’s d) we can detect with 90% power.

Our average realized sample sizes per group were 410, 122, and 208 for private investors,

⁴Cohen (1988) argues that a ‘medium’ effect size of $d = 0.5$ is ‘large enough to be visible to the naked eye’. To put effect sizes into the perspective of risk elicitation experiments, we refer to the extensive meta-study by Filippin and Crosetto (2016), who analyze the effect of gender on risk-taking. They find an average effect size of $d = 0.55$ for both the investment game and the gamble-choice task, and $d = 0.17$ for the multiple price list. However, while, some controversy persists on the effect of gender on risk-taking, the importance of incentivizing preference elicitation tasks seems to be almost universally accepted among experimental economists. We therefore believe that a threshold *below* the median effect of gender is a conservative benchmark for the presumed effect of incentivization.

professional investors, and students, respectively. These numbers are all well above the lower bound of $N = 85$ to detect effect sizes below $d = 0.5$ with 90% power. The corresponding *ex ante* effect sizes we are able to detect given our realized numbers of observations are summarized in Table C1.

D Instructions of the Online Experiment

Intro screen incentives treatment



Thank you for participating in this study!

The survey takes about **15-20 minutes**. Participation is voluntary and anonymous. Responses are analyzed for the purpose of academic research only.

We reward your participation with **real monetary benefits**. The **starting value of your reward is 12 EUR**. Depending on your answers in the survey, your payoff can increase or decrease. On average, you can expect a payoff of **36 EUR**. You can choose whether to receive your final payoff via bank transfer or as an online Amazon gift card.

In the first six tasks of this survey, we ask you to choose among several options with different payoff profiles. To determine your final payoff, at the end of the survey, **one of these six tasks** is selected at random. The outcome of one of your decisions in the respective task determines by how much your final payoff increases above or decreases below the initial 12 EUR. All tasks are equally likely to be relevant for your payoff.

Note that there are no "right" or "wrong" answers to the questions in this survey. This questionnaire is about your own preferences. Always choose the option that you prefer.

Note:

The survey includes content that is not displayed optimally on smartphones. We therefore ask you to answer the survey on a **desktop computer, laptop or tablet**, if possible.

Amazon gift cards are distributed via e-mail. If you choose to receive a gift card, please indicate your e-mail address at the end of the survey. If you prefer to receive your payoff via bank transfer, please indicate your bank details at the end of the survey. E-mail addresses and bank details are stored separately from your answers and are deleted immediately after sending out the rewards.

If you have any questions, you can contact us by e-mail (finanzforum@finance.uni-frankfurt.de).

Next

[Privacy statement](#)

Intro screen flat treatment

 universität
innsbruck

 GOETHE
UNIVERSITÄT
FRANKFURT AM MAIN

Load unfinished survey

Exit and clear survey

Thank you for participating in this study!

The survey takes about **15-20 minutes**. Participation is voluntary and anonymous. Responses are analyzed for the purpose of academic research only.

We reward your participation with **real monetary benefits**. The **value of your reward is 12 EUR**. You can choose whether to receive your final payoff via bank transfer or as an online Amazon gift card.

In the first six tasks of this survey, we ask you to choose among several options with different hypothetical payoff profiles. Note that the payoff amounts that result from your decisions will not actually be paid out.

Note that there are no "right" or "wrong" answers to the questions in this survey. This questionnaire is about your own preferences. Always choose the option that you prefer.

Note:

The survey includes content that is not displayed optimally on smartphones. We therefore ask you to answer the survey on a **desktop computer, laptop or tablet**, if possible.

Amazon gift cards are distributed via e-mail. If you choose to receive a gift card, please indicate your e-mail address at the end of the survey. If you prefer to receive your payoff via bank transfer, please indicate your bank details at the end of the survey. E-mail addresses and bank details are stored separately from your answers and are deleted immediately after sending out the rewards.

If you have any questions, you can contact us by e-mail (finanzforum@finance.uni-frankfurt.de).

Next

[Privacy statement](#)

The *staircase procedure* by Falk et al. (2016, 2018) (FA)

Task: Decision sequence

Instructions

In the following **four decisions**, we ask you to choose between receiving a **fixed amount** (Option A) or participating in a **lottery** (Option B) which either pays **45 EUR** or **0 EUR** with equal probability (50/50).

The payoffs of the lottery are the same in each of the four decisions. The **fixed amount** is **different in each decision**.

If one of these four decisions is drawn at random to determine your payoff, your final earnings result as follows. If you choose to participate in the lottery, a fair coin is tossed to determine the outcome of the lottery. If the coin comes up "heads", your payoff increases by the positive amount, in addition to the initial payment. If the coin comes up "tails", you receive no additional payment. If you choose to receive the fixed amount, your payoff is equal to the fixed amount, in addition to the initial payment.

Next

Task: Decision sequence

Decision 1 of 4

Which option do you prefer?

Option A: You receive a fixed amount of 24 EUR.

Option B: You receive either 0 EUR or 45 EUR with equal probability (50/50).

Option A: You receive a fixed amount of 24 EUR.

Option B: You receive either 0 EUR or 45 EUR with equal probability (50/50).

☐ Option A: Fixed amount

☐ Option B: Lottery

Next

Task: Decision sequence

Decision 2 of 4

Which option do you prefer?

Option A: You receive a fixed amount of 36 EUR.

Option B: You receive either 0 EUR or 45 EUR with equal probability (50/50).

☐ Option A: Fixed amount

☐ Option B: Lottery

Next

Task: Decision sequence

Decision 3 of 4

Which option do you prefer?

Option A: You receive a fixed amount of 42 EUR.

Option B: You receive either 0 EUR or 45 EUR with equal probability (50/50).

☐ Option A: Fixed amount

☐ Option B: Lottery

Next

Task: Decision sequence

Decision 4 of 4

Which option do you prefer?

Option A: You receive a fixed amount of 45 EUR.

Option B: You receive either 0 EUR or 45 EUR with equal probability (50/50).

☐ Option A: Fixed amount

☐ Option B: Lottery

Next

The *gamble-choice task* by Eckel and Grossman (2002) (EG)

Task: Coin toss

You are asked to choose your favorite option among the six options displayed in the table below.

Each option has **two possible payoffs**. A coin is tossed to determine which of the two amounts is paid out. Both payoffs hence result **with equal probability**.

<u>Option</u>	<u>Coin comes up "heads"</u> (payoff in EUR)	<u>Coin comes up "tails"</u> (payoff in EUR)
1	21 EUR	21 EUR
2	27 EUR	18 EUR
3	33 EUR	15 EUR
4	39 EUR	12 EUR
5	45 EUR	9 EUR
6	52 EUR	2 EUR

Please choose **your preferred option**.

If this decision is drawn at random to determine your payoff, your final earnings result as follows. To determine which of two amounts is paid out, a fair coin is tossed. If the coin comes up "heads", your payoff increases by the higher amount, in addition to the initial payment. If the coin comes up "tails", your payoff increases by the lower amount, in addition to the initial payment.

- ☐ **Option 1** [Heads: 21 EUR or Tails: 21 EUR]
- ☐ **Option 2** [Heads: 27 EUR or Tails: 18 EUR]
- ☐ **Option 3** [Heads: 33 EUR or Tails: 15 EUR]
- ☐ **Option 4** [Heads: 39 EUR or Tails: 12 EUR]
- ☐ **Option 5** [Heads: 45 EUR or Tails: 9 EUR]
- ☐ **Option 6** [Heads: 52 EUR or Tails: 2 EUR]

Next

The *paired lottery choice task* by Holt and Laury (2002) (HL)

Task: Probability scenarios

You are asked to choose between two options, Option A and Option B, in each of the 10 scenarios below.

- Option A is a lottery that pays either **24.00 EUR** or **19.20 EUR**.
- Option B is a lottery that pays either **46.20 EUR** or **1.20 EUR**.

Starting from scenario 1, the **chances of the higher payoff increase** in each subsequent scenario as you move down the table.

	Option A					Option B			
	or:					or:			
	Probability	Payoff	Probability	Payoff		Probability	Payoff	Probability	Payoff
(1)	10%	24.00 EUR	90%	19.20 EUR		10%	46.20 EUR	90%	1.20 EUR
(2)	20%	24.00 EUR	80%	19.20 EUR		20%	46.20 EUR	80%	1.20 EUR
(3)	30%	24.00 EUR	70%	19.20 EUR		30%	46.20 EUR	70%	1.20 EUR
(4)	40%	24.00 EUR	60%	19.20 EUR		40%	46.20 EUR	60%	1.20 EUR
(5)	50%	24.00 EUR	50%	19.20 EUR		50%	46.20 EUR	50%	1.20 EUR
(6)	60%	24.00 EUR	40%	19.20 EUR		60%	46.20 EUR	40%	1.20 EUR
(7)	70%	24.00 EUR	30%	19.20 EUR		70%	46.20 EUR	30%	1.20 EUR
(8)	80%	24.00 EUR	20%	19.20 EUR		80%	46.20 EUR	20%	1.20 EUR
(9)	90%	24.00 EUR	10%	19.20 EUR		90%	46.20 EUR	10%	1.20 EUR
(10)	100%	24.00 EUR	0%	19.20 EUR		100%	46.20 EUR	0%	1.20 EUR

Please indicate whether you prefer **Option A** or **Option B** in each of the 10 scenarios.

If this decision is drawn at random to determine your payoff, your final earnings result as follows. In a first step, one of the 10 scenarios is selected at random. To determine the outcome of the option you selected for this scenario, one ball is drawn at random from an urn containing 100 blue and red balls. The proportion of blue and red balls corresponds to the probabilities of respective scenario (e.g., 10 blue and 90 red balls for Scenario 1). If a blue (red) ball is drawn, your payoff increases by the higher (lower) amount, in addition to the initial payment.

	Option A	Option B	
24.00 EUR with 10% probability or 19.20 EUR with 90% probability	<input type="radio"/>	<input type="radio"/>	46.20 EUR with 10% probability or 1.20 EUR with 90% probability
24.00 EUR with 20% probability or 19.20 EUR with 80% probability	<input type="radio"/>	<input type="radio"/>	46.20 EUR with 20% probability or 1.20 EUR with 80% probability
24.00 EUR with 30% probability or 19.20 EUR with 70% probability	<input type="radio"/>	<input type="radio"/>	46.20 EUR with 30% probability or 1.20 EUR with 70% probability
24.00 EUR with 40% probability or 19.20 EUR with 60% probability	<input type="radio"/>	<input type="radio"/>	46.20 EUR with 40% probability or 1.20 EUR with 60% probability
24.00 EUR with 50% probability or 19.20 EUR with 50% probability	<input type="radio"/>	<input type="radio"/>	46.20 EUR with 50% probability or 1.20 EUR with 50% probability
24.00 EUR with 60% probability or 19.20 EUR with 40% probability	<input type="radio"/>	<input type="radio"/>	46.20 EUR with 60% probability or 1.20 EUR with 40% probability
24.00 EUR with 70% probability or 19.20 EUR with 30% probability	<input type="radio"/>	<input type="radio"/>	46.20 EUR with 70% probability or 1.20 EUR with 30% probability
24.00 EUR with 80% probability or 19.20 EUR with 20% probability	<input type="radio"/>	<input type="radio"/>	46.20 EUR with 80% probability or 1.20 EUR with 20% probability
24.00 EUR with 90% probability or 19.20 EUR with 10% probability	<input type="radio"/>	<input type="radio"/>	46.20 EUR with 90% probability or 1.20 EUR with 10% probability
24.00 EUR with 100% probability	<input type="radio"/>	<input type="radio"/>	46.20 EUR with 100% probability

Next

The *investment game* by Gneezy and Potters (1997) (GP)

Task: Project investment decision

You receive **24 EUR**. You are asked to decide how much of this amount you want to invest in a project.

With a **probability of 50%**, the project is successful and pays **2.50 EUR** for each Euro invested. With equal probability of 50%, the project is not successful and you lose the amount you invested.

You can adjust the amount you wish to invest in steps of 2 EUR.

Those EUR you do not invest are yours to keep.

Here are the payoff rules again:

Probability	Project outcome	Your payoff (EUR)
50%	Success	<i>amount kept</i> <i>+ 2.5 x amount invested</i>
50%	No success	<i>amount kept</i>

Please indicate **how many EUR** you wish to **invest** in the project.

If this decision is drawn at random to determine your payoff, your final earnings result as follows. To determine if your investment is a success, a fair coin is tossed. If the coin comes up "heads", the project is successful. If the coin comes up "tails", the project is not successful. The amount by which your final payoff increases in addition to the initial payment hence depends on how much you choose to invest in the project, and on whether or not the project is successful.

0 EUR 2 EUR 4 EUR 6 EUR 8 EUR 10 EUR 12 EUR 14 EUR 16 EUR 18 EUR 20 EUR 22 EUR 24 EUR

Next

University of Innsbruck - Working Papers in Economics and Statistics

Recent Papers can be accessed on the following webpage:

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

- 2020-29 **Andreas Hackethal, Michael Kirchler, Christine Laudenbach, Michael Razen, Anika Weber:** [On the role of monetary incentives in risk preference elicitation experiments](#)
- 2020-28 **Andrej Gill, Matthias Heinz, Heiner Schumacher, Matthias Sutter:** [Trustworthiness in the Financial Industry](#)
- 2020-27 **Matthias Sutter, Michael Weyland, Anna Untertrifaller, Manuel Froitzheim:** [Financial literacy, risk and time preferences - Results from a randomized educational intervention](#)
- 2020-26 **Rene Schwaiger, Jürgen Huber, Michael Kirchler, Daniel Kleinlercher, Utz Weitzel:** [Unequal Opportunities, Social Groups, and Redistribution](#)
- 2020-25 **Roman Inderst, Martin Obradovits:** [Competitive Strategies when Consumers are Relative Thinkers: Implications for Pricing, Promotions, and Product Choice](#)
- 2020-24 **Martin Obradovits, Philipp Plaickner:** [Price-Directed Search and Collusion](#)
- 2020-23 **Helena Fornwagner, Oliver P. Hauser:** [Climate action for \(my\) children](#)
- 2020-22 **Esther Blanco, Natalie Struwe, James M. Walker:** [Incentivizing public good provision through outsider transfers: experimental evidence on sharing rules and additional requirements](#)
- 2020-21 **Loukas Balafoutas, Helena Fornwagner, Rudolf Kerschbamer, Matthias Sutter, Maryna Tverdostup:** [Diagnostic Uncertainty and Insurance in Credence Goods Markets](#)
- 2020-20 **Anna Ulrichshofer, Markus Walzl:** [Customer Disputes, Misconduct, and Reputation Building in the Market for Financial Advice](#)
- 2020-19 **Anna Ulrichshofer, Markus Walzl:** [Social Comparison and Optimal Contracts in the Competition for Managerial Talent](#)
- 2020-18 **Martin Obradovits, Philipp Plaickner:** [Searching for Treatment](#)
- 2020-17 **Jun Honda:** [The Gender-Punishment Gap revisited](#)
- 2020-16 **Jun Honda:** [The Relation between Rankings and Risk-Taking in the Labor Market for Financial Advice](#)

- 2020-15 **Christina Bannier, Eberhard Feess, Natalie Packham, Markus Walzl:** [Differentiation and Risk-Aversion in Imperfectly Competitive Labor Markets](#)
- 2020-14 **Felix Holzmeister, Rudolf Kerschbamer:** [oTree: The Equality Equivalence Test](#)
- 2020-13 **Parampreet Christopher Bindra, Graeme Pearce:** [The effect of priming on fraud: Evidence from a natural field experiment](#)
- 2020-12 **Alessandro De Chiara, Marco A. Schwarz:** [A Dynamic Theory of Regulatory Capture](#)
- 2020-11 **Christoph Huber, Jürgen Huber, Michael Kirchler:** [Market shocks and professionals' investment behavior - Evidence from the COVID-19 crash](#)
- 2020-10 **Elisabeth Gsottbauer, Daniel Müller, Samuel Müller, Stefan T. Trautmann, Galina Zudenkova:** [Social class and \(un\)ethical behavior: Causal versus correlational evidence](#)
- 2020-09 **Parampreet Christopher Bindra, Rudolf Kerschbamer, Daniel Neururer, Matthias Sutter:** [Reveal it or conceal it: On the value of second opinions in a low-entry-barriers credence goods market](#)
- 2020-08 **Robert Steiger, Eva Posch, Gottfried Tappeiner, Janette Walde:** [Effects of climate change on tourism demand considering individual seasonal preferences](#)
- 2020-07 **Fang Liu, Alexander Rasch, Marco A. Schwarz, Christian Waibel:** [The role of diagnostic ability in markets for expert services](#)
- 2020-06 **Matthias Stefan, Jürgen Huber, Michael Kirchler, Matthias Sutter, Markus Walzl:** [Monetary and Social Incentives in Multi-Tasking: The Ranking Substitution Effect](#)
- 2020-05 **Michael Razen, Jürgen Huber, Laura Hueber, Michael Kirchler, Matthias Stefan:** [Financial Literacy, Economic Preferences, and Adolescents' Field Behavior](#)
- 2020-04 **Christian König-Kersting, Johannes Lohse, Anna Louisa Merkel:** [Active and Passive Risk-Taking](#)
- 2020-03 **Christoph Huber, Jürgen Huber:** [Bad bankers no more? Truth-telling and \(dis\)honesty in the finance industry](#)
- 2020-02 **Dietmar Fehr, Daniel Müller, Marcel Preuss:** [Social Mobility Perceptions and Inequality Acceptance](#)
- 2020-01 **Loukas Balafoutas, Rudolf Kerschbamer:** [Credence goods in the literature: What the past fifteen years have taught us about fraud, incentives, and the role of institutions](#)
- 2019-21 **Felix Holzmeister, Martin Holmen, Michael Kirchler, Matthias Stefan, Erik Wengström:** [Delegated Decision-Making in Finance](#)

- 2019-20 **Julia Rose, Michael Kirchler, Stefan Palan:** Status and Reputation Nudging
- 2019-19 **Felix Holzmeister, Matthias Stefan:** The risk elicitation puzzle revisited: Across-methods (in)consistency?
- 2019-18 **Katharina Momsen, Markus Ohndorf:** Information Avoidance, Selective Exposure, and Fake(?) News-A Green Market Experiment
- 2019-17 **Stjepan Srhoj, Bruno Skrinjaric, Sonja Radas, Janette Walde:** Closing the Finance Gap by Nudging: Impact Assessment of Public Grants for Women Entrepreneurs
- 2019-16 **Adam Farago, Martin Holmen, Felix Holzmeister, Michael Kirchler, Michael Razen:** Cognitive Skills and Economic Preferences in the Fund Industry
- 2019-15 **Christopher Kah, Daniel Neururer:** Generiert der stationäre Buchhandel positive Nachfrageeffekte und verhilft dadurch dem Kulturgut Buch bei seiner Verbreitung? - Ein natürliches Experiment
- 2019-14 **Stjepan Srhoj, Michael Lapinski, Janette Walde:** Size matters? Impact evaluation of business development grants on SME performance
- 2019-13 **Andrea M. Leiter, Engelbert Theurl:** Determinants of prepaid systems of healthcare financing - A worldwide country-level perspective
- 2019-12 **Michael Razen, Michael Kirchler, Utz Weitzel:** Domain-Specific Risk-Taking Among Finance Professionals
- 2019-11 **Jonathan Hall, Rudolf Kerschbamer, Daniel Neururer, Eric Skoog:** Uncovering sophisticated discrimination with the help of credence goods markups - evidence from a natural field experiment
- 2019-10 **Daniela Glätzle-Rützler, Philipp Lergetporer, Matthias Sutter:** Collective intertemporal decisions and heterogeneity in groups
- 2019-09 **Morten Hedegaard, Rudolf Kerschbamer, Daniel Müller, Jean-Robert Tyran:** Distributional Preferences Explain Individual Behavior Across Games and Time
- 2019-08 **Daniel Müller, Sander Renes:** Fairness Views and Political Preferences - Evidence from a representative sample
- 2019-07 **Florian Lindner, Michael Kirchler, Stephanie Rosenkranz, Utze Weitzel:** Social Status and Risk-Taking in Investment Decisions
- 2019-06 **Christoph Huber, Julia Rose:** Individual attitudes and market dynamics towards imprecision
- 2019-05 **Felix Holzmeister, Jürgen Huber, Michael Kirchler, Florian Lindner, Utz Weitzel, Stefan Zeisberger:** What Drives Risk Perception? A Global Survey with Financial Professionals and Lay People

- 2019-04 **David M. McEvoy, Tobias Haller, Esther Blanco:** [The Role of Non-Binding Pledges in Social Dilemmas with Mitigation and Adaptation](#)
- 2019-03 **Katharina Momsen, Markus Ohndorf:** [When do people exploit moral wiggle room? An experimental analysis in a market setup](#)
- 2019-02 **Rudolf Kerschbamer, Daniel Neururer, Matthias Sutter:** [Credence goods markets and the informational value of new media: A natural field experiment](#)
- 2019-01 **Martin Geiger, Eric Mayer, Johann Scharler:** [Inequality and the Business Cycle: Evidence from U.S. survey data](#)

University of Innsbruck

Working Papers in Economics and Statistics

2020-29

Andreas Hackethal, Michael Kirchler, Christine Laudenbach, Michael Razen, Annika Weber

On the role of monetary incentives in risk preference elicitation experiments

Abstract

Incentivized experiments in which individuals receive monetary rewards according to the outcomes of their decisions are regarded as the gold standard for preference elicitation in experimental economics. These task-related real payments are considered necessary to reveal subjects' "true preferences". Using a systematic, large-sample approach with three subject pools of private investors, professional investors, and students, we test the effect of task-related monetary incentives on risk preferences in four standard experimental tasks. We find no systematic differences in behavior between and within subjects in the incentivized and non-incentivized regimes. We discuss implications for academic research and for applications in the field.

ISSN 1993-4378 (Print)

ISSN 1993-6885 (Online)