Economics Letters 209 (2021) 110137

Contents lists available at ScienceDirect

Economics Letters

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

Do MTurkers exhibit myopic loss aversion?[☆]

Rene Schwaiger [*], Laura Hueber

University of Innsbruck, Department of Banking and Finance, Universitätsstrasse 15, 6020 Innsbruck, Austria

ARTICLE INFO

Article history:
Received 27 April 2021
Received in revised form 4 October 2021
Accepted 22 October 2021
Available online 1 November 2021

JEL classification:
G10
G11
G41

Keywords:
Online experiment
Myopic loss aversion
Risk
Mturk

ABSTRACT

We present results from a highly powered online experiment with 937 participants on Amazon Mechanical Turk (MTurk) that examined whether MTurkers exhibit myopic loss aversion (MLA). The experiment consisted of measuring MLA-compliant behavior in two between-subjects treatments that differed only regarding the risk profile of the risky asset employed. We found no statistically significant evidence of MLA-compliant behavior for any of the two risk profiles among MTurkers in the full samples. However, we found evidence of MLA for one of the two risk profiles in some sub-samples where we screened-out participants based on processing times in the experiment.

1. Introduction

Myopic loss aversion (MLA) describes the behavior of individuals to frame decisions narrowly, i.e., to evaluate investments frequently or to segregate them, which is based on mental accounting (Kahneman and Tversky, 1984; Thaler, 1985; Kahneman and Lovallo, 1993; Thaler et al., 1997; Lee and Veld-Merkoulova, 2016), making them more prone to existing loss aversion (Kahneman and Tversky, 1979). Such behavior has been associated with a negative impact on individuals' financial decision making (Looney and Hardin, 2009). There exists a substantial body of empirical evidence reporting MLA-compliant behavior among university students in individual decisions (Keren and Wagenaar, 1987; Gneezy and Potters, 1997; Thaler et al., 1997; Bellemare et al., 2005; Langer and Weber, 2005; Fellner and Sutter, 2009), team decisions (Sutter, 2007), and experimental market situations (Gneezy et al., 2003). It has further been demonstrated that also individuals from the general population (Van der Heijden et al., 2012), financial experts (Haigh and List, 2005; Eriksen and Kvaloy, 2010; Larson et al., 2012), and private investors (Wendy and Asri, 2012) behave in accordance with MLA theory. Moreover, there exists evidence for MLA-compliant behavior in the contexts of retirement savings and insurances (Benartzi and Thaler, 1999; Papon, 2008).

We conducted an online experiment to investigate whether the concept of MLA can be generalized to the behavior of crowd workers on Amazon Mechanical Turk (MTurk), a subject pool frequently recruited for online social science experiments (Chandler and Shapiro, 2016). In doing so, we implemented the lottery investment framework established by Gneezy and Potters (1997) – the foundation for the most commonly used measurement of MLA (see e.g., Bellemare et al., 2005; Haigh and List, 2005; Fellner and Sutter, 2009) – on MTurk.

Given that the predictions of MLA theory do not explicitly differ between various types of mixed gambles as long as they are characterized by a positive expected value (Haisley et al., 2008), as an exploratory extension we also tested whether design changes regarding the risk profile of the lottery affect participant behavior. Therefore, in addition to the lottery by Gneezy and Potters (1997), we applied a second lottery based on Charness and Gneezy (2010), also characterized by a positive expected value, but more attractive in terms of both expected value and probabilities of winning and losing.[1]

☆ We thank one anonymous referee, Christoph Huber, Michael Kirchler, and Christian König for very valuable comments on previous versions of the paper. The financial support of the Austrian Science Fund (SFB F63), the Netzwerk Banking, Accounting, Auditing, Finance & IT (BAFIT), the Bank Austria Foundation and the University of Innsbruck is gratefully acknowledged.

* Corresponding author.
E-mail address: rene.schwaiger@uibk.ac.at (R. Schwaiger).

1 This choice was motivated by the fact that a pretest of a previous study by the authors (Hueber and Schwaiger, 2021), in which this lottery was used in a MLA framework, revealed behavior in student participants that was exactly opposed to the behavior predicted by MLA. Although the number of participants in this pre-test was very small, such reverse behavior did not occur in a similarly powered pre-test applying the lottery by Gneezy and Potters (1997).

https://doi.org/10.1016/j.econlet.2021.110137
0165-1765/

We found no MLA-compliant behavior in either treatment in the full samples, but we found evidence of MLA for the lottery by Gneezy and Potters (1997) in some sub-samples where we screened-out participants based on processing times in the experiment. In addition, we found no difference-in-difference effect between the two treatments, suggesting that the differences in risk profiles do not statistically significantly affect the impact of varying feedback and decision frequency on participants' risk-taking.

We contribute mainly to two strands in the literature. Specifically, we add to the general literature on MLA already discussed by applying the Gneezy and Potters (1997) MLA framework to MTurkers, i.e., a participant pool that, to the best of our knowledge, has not yet been investigated in this respect. In doing so, we address the generalizability of the concept across groups of people. Secondly, by varying the risk profiles of the lotteries, we contribute to a smaller part of the literature that examines the robustness and universality of MLA regarding differences in the characteristics of the underlying risky asset. It has been shown that the predictions of MLA theory are not robust to more realistic parameters of risky assets, such as actual financial market data (Beshears et al., 2017) as well as to mixed gambles with negative expected value, where the opposite of the predicted behavior can occur (Haisley et al., 2008). Similar results to the latter, but considering lotteries with positive expected value, have been provided by Langer and Weber (2001, 2005). The authors compellingly argued to extend the concept of MLA to the concept of myopic prospect theory (MPT) to explain non-unidirectional effects of varying feedback and decision frequency on decisions under risk with positive expected value. As Zeisberger et al. (2012, p. 46) have aptly put it: "*What can be learned [...] of research on myopia and investment is that [...] minor design issues that had not been considered to be relevant beforehand might have a major impact on the results.*" In addition, studies that have looked at the causes of MLA-conforming behavior, i.e., feedback and/or decision frequency, have provided mixed results (Bellemare et al., 2005; Langer and Weber, 2008; Fellner and Sutter, 2009).

2. Experimental design and procedure

Following the procedure by Gneezy and Potters (1997), the participants had to make a betting decision for each of nine rounds. Specifically, each participant i had to decide on a value $x_i \in [0, 200]$ of an initial endowment per round of 200 tokens to bet in a risky lottery. Participants were randomly assigned to one of two groups, i.e., sub-treatment H or sub-treatment L, which differed only in terms of feedback and decision frequency. In the H sub-treatment, participants chose the amount to bet in the risky lottery in each of the nine rounds and were informed after each round about the outcome of the lottery and their earnings from that round. In contrast, in the L sub-treatment, participants were asked to decide on the amount to bet in the lottery in rounds 1, 4, and 7 for three consecutive rounds. Decisions were binding for three rounds, so the amount bet in sub-treatment L remained unchanged for three consecutive rounds. Participants were informed about the outcomes of the lotteries and aggregated earnings only after every third round.[2] According to MLA theory, our prediction was that participants in the L sub-treatment will bet higher amounts than participants in the H sub-treatment, which is explained by a more advantageous perception of the lotteries when their results are presented in a more aggregated way. Both sub-treatments were implemented

Table 1
Treatment overview. The table provides a treatment overview varying the lottery properties across treatments GP and CG and the decision/feedback frequency across sub-treatments H and L.

Lottery	Treatment	Sub-treatment	
Gneezy and Potters (1997)	GP: $E(x_i) = 0.17$ for $x_i = 1$, $pr_{loss} = 2/3$	H	L
Charness and Gneezy (2010)	CG: $E(x_i) = 0.75$ for $x_i = 1$, $pr_{loss} = 1/2$	H	L

in two between-subjects treatments, i.e., GP and CG, to which participants were randomly assigned. In treatment GP, we applied the original lottery by Gneezy and Potters (1997) that reads as follows:[3]

> *You have a chance of 2/3 (67%) to lose the amount you bet and a chance of 1/3 (33%) to win two and a half times the amount you bet.*

In addition, in another treatment CG, we introduced the lottery established by Charness and Gneezy (2010) with the following risk profile:

> *You have a chance of 1/2 (50%) to lose the amount you bet and a chance of 1/2 (50%) to win two and a half times the amount you bet.*

Therefore, two treatments were obtained that differed only in terms of the risk profiles of the lotteries employed. Specifically, the lottery in treatment GP was characterized by an expected value $E(x_i) = 0.17$ for $x_i = 1$ with a loss probability $pr_{loss} = 67\%$. The lottery in CG was characterized by an expected value $E(x_i) = 0.75$ for $x_i = 1$ with a loss probability $pr_{loss} = 50\%$. According to the theory of MLA, we hypothesized that treatments would not differ regarding MLA-conform behavior.[4] For a given round t, participant i's earnings $\pi_{i,t}$ were given as follows:

$$\pi_{i,t} = \left\{ \begin{array}{lll} 200 + 2.5x_{i,t} & pr_{win}:\ 1/3\,(\text{GP}); & pr_{win}:\ 1/2\,(\text{CG}) \\ 200 - x_{i,t} & pr_{loss}:\ 2/3\,(\text{GP}); & pr_{loss}:\ 1/2\,(\text{CG}) \end{array} \right\} \tag{1}$$

For each treatment, we implemented the two sub-treatments, i.e., H and L, varying the decision and feedback frequency, as summarized in Table 1.

We conducted a highly powered trial. Ex-post power analyses showed that our sample size of $N = 473$ in treatment GP, and $N = 464$ in treatment CG allowed us to obtain 80% power to reliably detect a small effect of Cohen's $d = 0.20$ with respect to average differences in risk-taking between H and L in both treatments.[5]

2 In round 3, the aggregated earnings from rounds 1–3; in round 6, the aggregated earnings from rounds 4–6; and in round 9, the aggregated earnings from rounds 7–9 were shown.

3 To ensure a valid comparison to Gneezy and Potters (1997), the instructions in our study were virtually equivalent to those in the original paper. Our instructions differed only with respect to the implementation of the lottery draw, which in our study was performed by the computer.

4 Following Langer and Weber (2001, 2005), we calculated whether MPT could explain possible reversed behavioral patterns when participants are confronted with the lottery in CG. Assuming estimated probability weights of γ^+: 0.61 and γ^-: 0.69 and weighting and value functions by Kahneman and Tversky (1992), for no values of α, β and λ, it follows that $S_1(x) > 0$ when $S_3(x) < 0$ holds simultaneously, i.e., the myopic value $S_1(x)$ of this lottery can never be positive if the non-myopic (aggregated) value $S_3(x)$ of this lottery is not positive at the same time. Thus, MPT would not predict that participants are willing to invest in this lottery in the myopic case (H) while not being willing to invest in the non-myopic case (L).

5 This effect was related to the measurement of MLA over all nine rounds, which we focused on in this paper. Specifically, in GP, we achieved a statistical power of approximately 99% to detect 71% of the original standardized effect size of Cohen's $d = 0.63$ in Gneezy and Potters (1997). Recent evidence on the replicability of social science experiments has provided an estimate of the average relative effect size of true positives that is approximately 71% (Camerer et al., 2016).

The experiment was conducted online with 937 US participants on Amazon MTurk.[6] The average age of the participants was 37 years, with 34% of participants being female and 66% of participants being male (see Table A.1 for details and further demographic and socioeconomic information).[7] Experimental sessions were held in August and September 2020 and January 2021. The average time participants spent on the experiment was 7.10 min (SD: 5.75 min). Participants received a flat fee of $0.75 plus an average bonus incentive of $1.45 (SD: 0.48) based on their decisions and lottery outcomes. The experiment was programmed using oTree (Chen et al., 2016).[8]

3. Results

Fig. 1 shows the average round bet over nine rounds as a percentage of the initial endowment of 200 tokens for both treatments, i.e., GP and CG, and sub-treatments H and L.

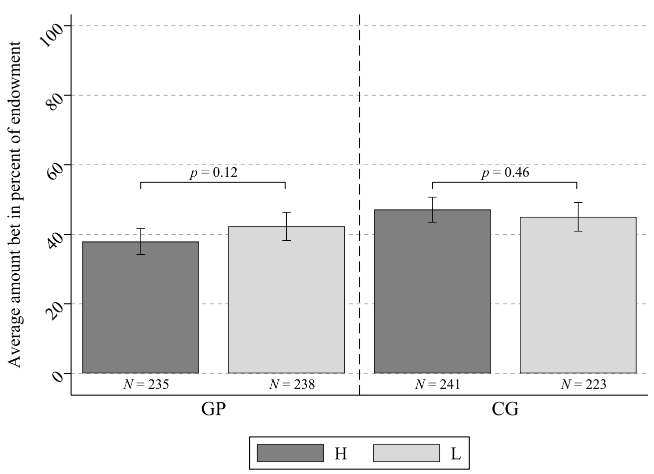


Fig. 1. Average amount bet in percent for both treatments and sub-treatments. The figure shows the average amount bet over nine rounds as a percentage of the initial endowment of 200 tokens for both treatments and sub-treatments (dark gray bars represent sub-treatment H and light gray bars represent sub-treatment L). Whiskers denote 95% confidence intervals. p indicates p-values of two-sided unpaired sample t-tests between sub-treatments H and L. GP: $N = 473$; CG: $N = 464$.

To begin our analyses, we first consider treatment GP, which was an online replication by Gneezy and Potters (1997).[9]

> **Result 1:** *MTurkers in treatment* GP *did not exhibit behavior consistent with MLA.*

Although displaying the sign predicted by MLA theory, a two-sided, unpaired sample t-test indicated that the small difference between H and L in terms of the average percentage bet in the lottery (Cohen's $d = 0.15$) is not statistically significant, as can be seen at the top of the corresponding first pair of bars in Fig. 1 (H: 0.379 - L: 0.423 $= -0.044$; $p = 0.12$; $N = 473$, see Table A.3 for details). Thus, in contrast to our hypothesis, we did not find evidence that MTurkers exhibit MLA-compliant behavior. The results contradict the findings of previous studies that have used this experimental design and have found statistically significant evidence of MLA-conforming behavior among different groups, e.g., university students or financial professionals (Gneezy and Potters, 1997; Gneezy et al., 2003; Bellemare et al., 2005; Fellner and Sutter, 2009).

> **Result 2:** *We did not find behavior consistent with MLA in treatment* CG *and did not find evidence of a variation in the difference in risk taking between* L *and* H *across treatments.*

As shown in Fig. 1, we found no statistically significant difference in MTurkers' risk-taking between H and L in treatment CG, as indicated by the corresponding p-value above the bars obtained from two-sided unpaired sample t-tests (H: 0.471 - L: 0.450 $= 0.021$; $p = 0.46$; $N = 464$, see Table A.3 for details).[10,11,12]

Finally, we ran multivariate Tobit regressions with the average lottery bets over nine rounds as dependent variable to examine the robustness of the results and to test for a difference-in-differences effect (see Table A.4 for details). All previous results were confirmed, but we found no variation in differences in amounts bet between participants in the H group and the L group across treatments, i.e., no statistically significant difference-in-difference effect, as indicated by the coefficient CG × L in models I and II in Table A.4 (model I: $p = 0.11$; model II: $p = 0.21$). The results were robust to permutation tests[13] and to the inclusion of participants' general and financial risk preferences, demographic, and socioeconomic characteristics.

6 In an exit questionnaire, we asked participants about their general and financial risk preferences, demographic and socioeconomic characteristics, such as information on age, gender, education, annual gross income, as well as their financial education and investment experience. The self-reported general and financial risk preferences were based on the German SOEP questionnaire (Dohmen et al., 2011).

7 We performed extensive randomization checks to test whether the distributions of demographic, socioeconomic, and risk-taking characteristics of participants differed between treatments and sub-treatments. We found no statistically significant differences in participant characteristics between treatments and sub-treatments, indicating a successful randomization procedure (see Table A.2 for details).

8 We refer to the Appendix for screenshots of the software. The experimental software can be accessed using the following link.

9 We applied significance levels of 5% and 0.5% for all statistical tests in this paper (Benjamin et al., 2017) and took a conservative approach by conducting two-sided tests, which was further justified by the empirically confirmed possibility of reverse effects (Langer and Weber, 2001, 2005).

10 Interestingly, although not statistically significant, an inverse pattern compared with that predicted by MLA seemed to occur, i.e., participants in sub-treatment H bet more compared to participants in sub-treatment L.

11 As robustness checks, we also performed the analyses in both treatments using the non-parametric Mann–Whitney U test, which confirmed the results.

12 The absence of a statistical support for the treatment effects is no sufficient evidence for null effects. As we were highly powered, we performed equivalence tests (TOST) to also test for equivalence with the null hypothesis in both treatments. We used the user-written program "tostt" in Stata (Dinno, 2017). We followed the approach by Juzek and Kizach (2019) to obtain more objective values for the parameter delta (δ) – the minimum worthwhile effect size – based on our data (GP: $\delta = \pm\, 0.09$; CG: $\delta = \pm\, 0.09$). For these values of δ, equivalence with the null hypothesis regarding the difference in risk-taking between H and L could be statistically supported (Tryon and Lewis, 2008) in both treatments (GP: $p(T > t_1) < 0.005$, $p(T > t_2) = 0.049$; CG: $p(T > t_1) = 0.007$, $p(T > t_2) < 0.005$). Conducting further equivalence tests in treatment GP, we were able to statistically rule out a difference in risk taking between H and L larger than about 9 percentage points (Cohen's $d = 0.30$). In the CG treatment, we were able to statistically rule out a difference larger than about 6.70 percentage points (Cohen's d = 0.22).

13 We used the user-written program "ritest" in Stata (Heß, 2017).

Result 3: *The null effect in Treatment* CG *remains stable even after applying various robustness checks based on the processing times of the participants in the experiment. However, in Treatment* GP, *some sub-samples show behavior consistent with MLA in these checks.*

A potential limitation of our study could be the data quality prevailing on the MTurk platform, which is discussed in detail in Appendix A.1. In addition, we present in this section in the Appendix the results of several robustness checks related to participants' time spent in the experiment and on the task instruction screen. Based on these analyses, we were able to confirm a null effect in Treatment CG, but found evidence of MLA for the lottery by Gneezy and Potters (1997) in some sub-samples in which we filtered out outliers based on total processing times in the experiment. Therefore, we remain cautious in interpreting the results in Treatment GP and note that noisy data may be a reason for the non-significant effect in the full sample in this treatment. MLA-compliant behavior reported in Gneezy and Potters (1997), thus, is generalizable only to more attentive and focused sub-populations on MTurk.

4. Conclusion

We conducted a highly powered online experiment with 937 participants on Amazon MTurk following the lottery framework by Gneezy and Potters (1997) to test whether MTurkers exhibit MLA in two treatments that differed in terms of lottery risk profiles. With the results of the full sample, we are unable to confirm MLA-compliant behavior for MTurkers in either treatment as we found small, statistically non-significant differences in risk taking between H and L and could statistically rule out standardized differences greater than Cohen's $d = 0.30$ (0.22) in GP (CG). In addition, we found no difference-in-difference effect between treatments, suggesting that the different risk profiles did not affect the differences in risk taking between H and L. Nonetheless, applying several robustness checks with respect to the processing times of the participants in the experiment, we found evidence of behavior consistent with MLA in Treatment GP. We conclude that the non-significant result with the full sample may be due to noisy data on MTurk. However, for Treatment CG, we could establish a robust null effect. Thus, the results of previous MLA studies, or at least the magnitude of the results, are generalizable only to a specific attentive sub-population on MTurk. Furthermore, with these findings, we join a growing body of scientific literature suggesting that the relationship between variations in decision and feedback frequency and risk-taking behavior is complex as it seems to depend on the underlying risk profile.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix

A.1. Robustness checks (data quality)

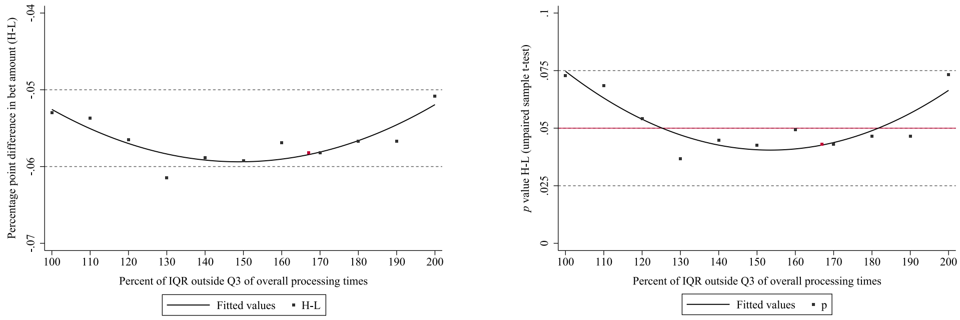
There are different findings regarding data quality on MTurk. On the one hand, crowd workers on MTurk have been shown to produce results similar to those in laboratory experiments or studies with representative population-based samples (see, e.g., Paolacci et al., 2010; Berinsky et al., 2012; Crump et al., 2013; Mullinix et al., 2015), to be similarly attentive or even more attentive than participants on other platforms (Thomas and Clifford, 2017), and to reliably and consistently report characteristics such as demographics and risk preferences (Johnson and Ryan, 2020). However, MTurk sampling has also been associated with specific problems that can lead to lower data quality (see e.g., Chmielewski and Kucker, 2020; Aguinis et al., 2021). In particular, fraudulent behavior by non-U.S. residents or participants using a virtual private server (VPS) appears to be an issue. There is evidence that such participants contribute to a lower signal-to-noise ratio, resulting in an attenuation of treatment effects by about 10–30 percent (see, e.g., Ahler et al., 2020; Kennedy et al., 2020).[14] We cannot rule out the possibility that such participants may have influenced our results, at least in part. Therefore, to check the robustness of our results, we examined data quality with respect to inattentive and otherwise problematic participants. Specifically, we used the amount of time participants spent throughout the overall experiment and, additionally, on the instruction screen for the task as proxies for data quality and tested whether several different ways of trimming the sample influence our results qualitatively. Although we did not implement specific attention or comprehension checks ex-ante and completion time in the experiment is not a perfect indicator of data quality, it does appear to be significantly, albeit weakly, correlated with other validity indicators (Chmielewski and Kucker, 2020). Furthermore, analyzing and trimming the sample with respect to completion times has the potential to significantly mitigate the degree of a present negative effect on data quality due to problematic participants. In particular, Ahler et al. (2020) have shown that potential trolls and participants with questionable IP addresses on Mturk over-proportionally appear to be (slow) outliers in terms of overall processing times. Therefore, we followed the procedure by Ahler et al. (2020) and excluded participants from the analyses who finished 167% of the interquartile range (IQR) outside the third quartile (Q3) of overall processing times (there were no fast outliers in our data using this method). We repeated the main analyses with these adjusted samples excluding (slow) outliers. The main results on differences in the average bet amount (H - L) over nine rounds in percent of the endowment and the respective p-values of two-sided unpaired sample t-tests [H - L] for this specific cut-off point and several different cut-off points between 100% and 200% of IQR are shown in Fig. A.1 for Treatment GP and Fig. A.2 for Treatment CG.

It is observable from Fig. A.1 that excluding slower participants who finished between 130% and 190% of IQR outside the third quartile of overall processing times in the experiment led to statistically significant evidence for MLA in the remaining sample in Treatment GP. However, visible from Fig. A.2, we found no cut-off points that led to qualitatively different results from the full sample in Treatment CG.[15]

In this analyses we only trimmed slow participants as the method did not detect any fast outliers. On average, it should have taken participants 2.06 min to meaningfully read the instructions (2032 characters) for the task (Trauzettel-Klosinski and Dietz, 2012). Strikingly, around 50% of participants spent less than 40 seconds on this screen. Although this seems problematic, there is evidence that speedsters on Mturk do not necessarily provide bad data quality (see, e.g., Ahler et al., 2020). As an additional robustness check, we symmetrically trimmed the sample based

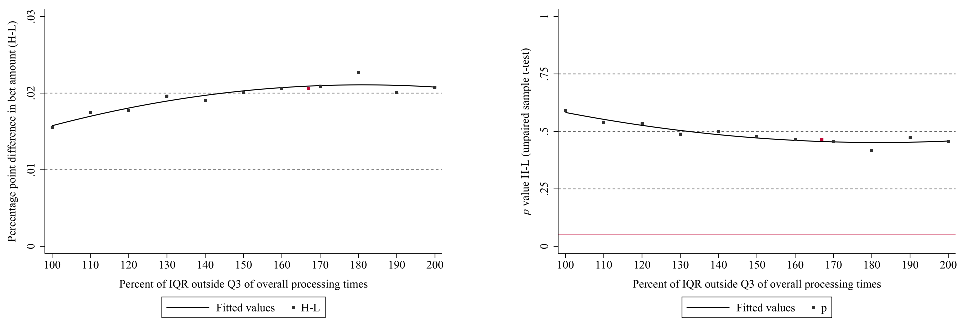
[14] On the other hand, bots on Mturk do not seem to pose as big a problem for data quality (anymore) as previously thought (see, e.g., Dreyfuss et al., 2018; Ahler et al., 2020; Kennedy et al., 2020).

[15] We repeated these analyses for processing times on the task-specific instruction screen, but found no qualitative difference in results in either treatment compared to the full sample (not shown).



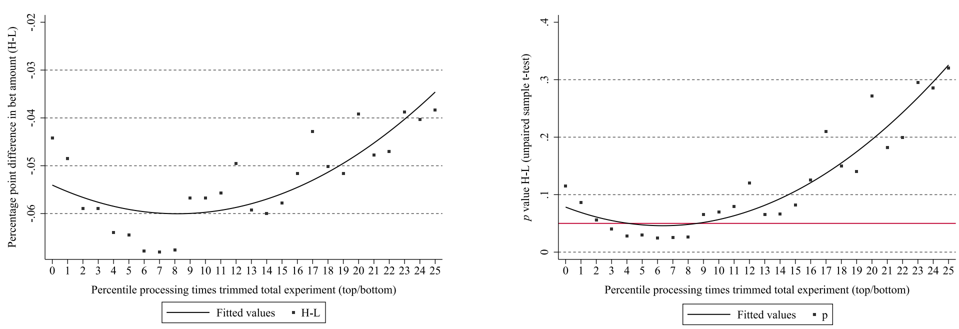
(a) H - L **(b)** p-values

Fig. A.1. Treatment GP: Difference in average risk taking (H - L) in sub-plot (a) and p-values (unpaired sample t-test [mean percent bet in H - mean percent bet in L], two-sided, alpha = 0.05) in sub-plot (b) for different percentages of IQR outside the third quartile (Q3) of overall processing times for the experiment. For example, at 150 percent, we excluded participants whose overall processing times were slower than (1.5 × IQR) + Q3 (there were no fast outliers according to this method). The horizontal red line in sub-plot (b) indicates a p-value of 0.05. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



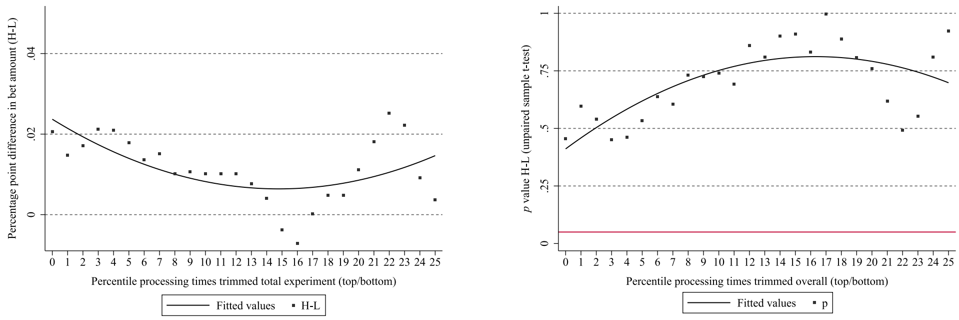
(a) H - L **(b)** p-values

Fig. A.2. Treatment CG: Difference in average risk taking (H - L) in sub-plot (a) and p-values (unpaired sample t-test [mean percent bet in H - mean percent bet in L], two-sided, alpha = 0.05) in sub-plot (b) for different percentages of IQR outside the third quartile (Q3) of overall processing times for the experiment. For example, at 150 percent, we excluded participants whose overall processing times were slower than (1.5 × IQR) + Q3 (there were no fast outliers according to this method). The horizontal red line in sub-plot (b) indicates a p-value of 0.05. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



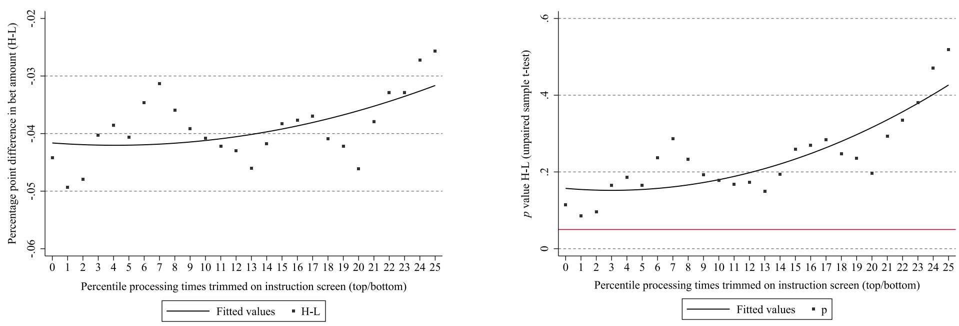
(a) H - L values **(b)** p-values

Fig. A.3. Treatment GP: Difference in average risk taking (H - L) in sub-plot (a) and p-values (unpaired sample t test [mean percent bet in H – mean percent bet in L], two-sided, alpha = 0.05) in sub-plot (b) for different cut-off points of overall processing times in the experiment. For example, for the 20th percentile, we excluded the participants with the 20% slowest and 20% fastest processing times in the experiment. The horizontal red line in sub-plot (b) indicates a p-value of 0.05. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



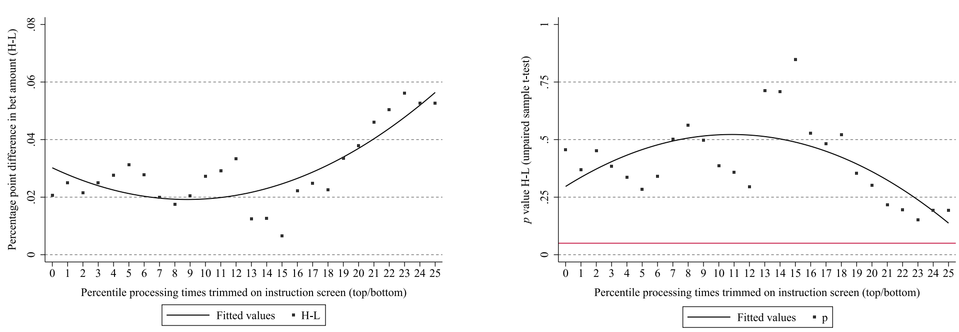
(a) H - L **(b)** p-values

Fig. A.4. Treatment CG: Difference in average risk taking (H - L) in sub-plot (a) and *p*-values (unpaired sample *t* test [mean percent bet in H – mean percent bet in L], two-sided, alpha = 0.05) in sub-plot (b) for different cut-off points of overall processing times in the experiment. For example, for the 20th percentile, we excluded the participants with the 20% slowest and 20% fastest processing times in the experiment. The horizontal red line in sub-plot (b) indicates a *p*-value of 0.05. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



(a) H - L **(b)** p-values

Fig. A.5. Treatment GP: Difference in average risk taking (H - L) in sub-plot (a) and *p*-values (unpaired sample *t* test [mean percent bet in H – mean percent bet in L], two-sided, alpha = 0.05) in sub-plot (b) for different cut-off points of processing times on the task relevant instruction screen. For example, for the 20th percentile, we excluded the participants with the 20% slowest and 20% fastest processing times on the instruction screen. The horizontal red line in sub-plot (b) indicates a *p*-value of 0.05. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



(a) H - L **(b)** p-values

Fig. A.6. Treatment CG: Difference in average risk taking (H - L) in sub-plot (a) and *p*-values (unpaired sample *t* tests [mean percent invested in H – mean percent invested in L], two-sided, alpha = 0.05) in sub-plot (b) for different cut-off points of processing times on the task relevant instruction screen. For example, for the 20th percentile, we excluded the participants with the 20% slowest and 20% fastest processing times on the instruction screen. The horizontal red line in sub-plot (b) indicates a *p*-value of 0.05. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

on overall processing times in the experiment to also account for fast participants, and present the results in Fig. A.3 for Treatment GP and Fig. A.4 for Treatment CG. We found a qualitative impact of this procedure on the results when we focused on the total processing times in the experiment in treatment GP.

As shown in Fig. A.3, excluding the top and bottom 3 to 8 percent of total processing times in the experiment led to statistically significant evidence of behavior consistent with MLA. Visible from Fig. A.4, there was no cut-off point that altered the results with the full sample qualitatively in Treatment CG. In addition, we asymmetrically trimmed the samples in the same systematic way. Here, compared to the full sample, the results remained qualitatively unchanged for all cut-off points when we only excluded participants with the fast processing times. However, when we focused only on participants with the slow processing times, we found qualitatively similar patterns as for the symmetric trimming (not shown), which is also consistent with the IQR procedure regarding outliers. Thus, particularly slow participants in terms of total time in the experiment appear to have a significant negative impact on data quality, consistent with Ahler et al. (2020).

As a final robustness check, we repeated the analyses, but this time based on the processing time on the instruction screen for the task. Fig. A.5 shows the results of the symmetric trimming procedure on the instruction screen for treatment GP and Fig. A.6 for Figure CG. As can be seen from both figures, symmetric trimming of the samples based on the times on the instruction screens did not lead to qualitative differences in either treatment compared to the full sample, thus, no statistically significant evidence for MLA. This also held true when we repeated the asymmetric procedure already applied for the overall time participants spent on the experiment (not shown).

A.2. Additional figures and tables

See Tables A.1–A.4.

A.3. Screenshots of the experiment

See Figs. A.7–A.18.

Table A.1
Descriptive statistics of demographic and socioeconomic variables including general and financial risk preferences. The table provides a descriptive overview regarding participants' age (AGE), gender (MALE), education (EDUCATION), annual gross income in USD (INCOME), professional or educational financial experience (FINANCIAL_SECTOR), investment experience regarding financial products for the past 5 years (INVEST_EXPERIENCE) and general (RISK_GENERAL) and financial risk-taking (RISK_FINANCIAL). RISK_GENERAL and RISK_FINANCIAL are ordinal variables ranging from 0 to 10, where 0 indicates participants' not being willing to take risks and 10 participants' being very willing to take risks. For the variables AGE, RISK_GENERAL and RISK_FINANCIAL the respective mean and standard deviation while for all other variables the relative distribution across categories is outlined.

Variable	Relative distribution across categories/means and standard deviations
AGE	Mean: 37.37; SD: 10.64
GENDER	Female: 0.34; Male: 0.66; Other: <0.01;
EDUCATION	No schooling: 0; Nursery school: <0.01; High school: 0.09; Associate degree: 0.04; Bachelor's degree: 0.61; Master's degree: 0.24; Doctoral degree: 0.01;
INCOME	0$–13.000$: 0.07; 13.000$–27.000$: 0.18; 27.000$–47.000$: 0.34; 47.000$–81.000$: 0.33; > 81.000$: 0.08;
FINANCIAL_SECTOR	Having worked in the finance sector/financial education: 0.47; Not having worked in the finance sector/no financial education: 0.53;
INVEST_EXPERIENCE	Having invested in financial products: 0.34; Not having invested in financial products: 0.66;
RISK_GENERAL	Mean: 8.02; SD: 2.70; (0: not at all willing to take risks, 10: very willing to take risks);
RISK_FINANCE	Mean: 7.90; SD: 2.77; (0: not at all willing to take risks, 10: very willing to take risks);

Table A.2
Randomization checks across treatments and sub-treatments. The variable AGE indicates the participants' age in years, MALE is a binary dummy taking the value of 0 for female subjects and 1 for male participants. STUDY_ECONOMICS is a binary variable, which equals 1 for participants enrolled in economics, business, or business law and 0 for all other study programs. FINANCIAL_SECTOR is a dummy taking the value of 1 for decision makers who have already worked in the financial sector or who have specific financial education and 0 for participants who have not. INVEST_EXPERIENCE represents a binary dummy taking the value of 1 for participants who have invested in financial products in the last five years. INCOME is an ordinal variable comprised of the total annual gross income quintiles in the US. EDUCATION is a 6-item ordinal variable taking the value of 0 for participants with nursery school completed up to a value of 6 for participants with a PhD. RISK_FINANCIAL is an ordinal variable representing self-reported risk preferences on a 10-point Likert scale in the financial domain. RISK_GENERAL is an ordinal variable representing self-reported risk preferences on a 10-point Likert scale in the general domain.

Comparison	Variable	Test	Statistic	N
Treatment: GP vs. CG	MALE	Pearsons χ^2 Test	$chi^2 = 1.215$	937
Sub-treatment: H vs. L (GP)	MALE	Pearsons χ^2 Test	$chi^2 = 0.000$	473
Sub-treatment: H vs. L (CG)	MALE	Pearsons χ^2 Test	$chi^2 = 2.276$	464
Treatment: GP vs. CG	AGE	Kruskal–Wallis Test	$chi^2 = 0.476$	937
Sub-treatment: H vs. L (GP)	AGE	Kruskal–Wallis Test	$chi^2 = 0.000$	473
Sub-treatment: H vs. L (CG)	AGE	Kruskal–Wallis Test	$chi^2 = 0.790$	464
Treatment: GP vs. CG	EDUCATION	Pearsons χ^2 Test	$chi^2 = 2.306$	937
Sub-treatment: H vs. L (GP)	EDUCATION	Pearsons χ^2 Test	$chi^2 = 2.589$	473
Sub-treatment: H vs. L (CG)	EDUCATION	Pearsons χ^2 Test	$chi^2 = 7.417$	464
Treatment: GP vs. CG	RISK_FINANCIAL	Kruskal–Wallis Test	$chi^2 = 1.879$	937
Sub-treatment: H vs. L (GP)	RISK_FINANCIAL	Kruskal–Wallis Test	$chi^2 = 3.476$	473
Sub-treatment: H vs. L (CG)	RISK_FINANCIAL	Kruskal–Wallis Test	$chi^2 = 0.110$	464
Treatment: GP vs. CG	RISK_GENERAL	Kruskal–Wallis Test	$chi^2 = 0.580$	473
Sub-treatment: H vs. L (GP)	RISK_GENERAL	Kruskal–Wallis Test	$chi^2 = 1.472$	473
Sub-treatment: H vs. L (CG)	RISK_GENERAL	Kruskal–Wallis Test	$chi^2 = 0.000$	464
Treatment: GP vs. CG	INCOME	Pearsons χ^2 Test	$chi^2 = 1.752$	937
Sub-treatment: H vs. L (GP)	INCOME	Pearsons χ^2 Test	$chi^2 = 7.103$	473
Sub-treatment: H vs. L (CG)	INCOME	Pearsons χ^2 Test	$chi^2 = 5.278$	464
Treatment: GP vs. CG	INVEST_EXPERIENCE	Pearsons χ^2 Test	$chi^2 = 3.198$	937
Sub-treatment: H vs. L (GP)	INVEST_EXPERIENCE	Pearsons χ^2 Test	$chi^2 = 1.561$	473
Sub-treatment: H vs. L (CG)	INVEST_EXPERIENCE	Pearsons χ^2 Test	$chi^2 = 2.375$	464
Treatment: GP vs. CG	FINANCIAL_SECTOR	Pearsons χ^2 Test	$chi^2 = 2.313$	937
Sub-treatment: H vs. L (GP)	FINANCIAL_SECTOR	Pearsons χ^2 Test	$chi^2 = 0.115$	473
Sub-treatment: H vs. L (CG)	FINANCIAL_SECTOR	Pearsons χ^2 Test	$chi^2 = 1.923$	464

Note: * $p < 0.05$, ** $p < 0.005$.

Table A.3
Differences between treatments and sub-treatments. The table shows pairwise differences in the average bet amount over nine rounds in percent of the endowment between sub-treatments H and L in treatments GP and CG using two-sided unpaired sample t-tests. The table also shows pairwise differences in the total average bet amount (H + L) over nine rounds in percent of the endowment between treatments.

Treatments	obs	Sub-Treatment Difference: H-L	std. err.	comb. std. dev	pr(|T| > |t|)
GP	473	−0.044 (0.379–0.423)	0.028	0.305	0.115
CG	464	0.021 (0.471–0.450)	0.028	0.298	0.456
Pairwise comp.	**obs**	**Treatment Difference:**	**std. err.**	**comb. std. dev**	**pr(|T| > |t|)**
GP-CG	937	−0.060** (0.401–0.461)	0.020	0.303	0.002

Note: * $p < 0.05$, ** $p < 0.005$.

Table A.4
Multivariate Tobit regressions on differences between treatments and sub-treatments. The table shows multivariate Tobit regressions with the average percentage amount bet by participants as dependent variable. The variable CG is a binary dummy taking on the value 1 for participants in treatment CG and 0 for participants in treatment GP. L represents a binary dummy variable taking the value 1 for decision makers in the low-frequency feedback sub-treatment and 0 for their peers in the high-frequency feedback sub-treatment, i.e., H. CG × L represents an interaction term between CG and L. The variable AGE indicates the participants' age in years, MALE is a binary dummy taking the value of 0 for female subjects and 1 for male participants. FINANCIAL_SECTOR is a dummy taking the value of 1 for decision makers who have already worked in the financial sector or who have specific financial education and 0 for participants who have not. INVEST_EXPERIENCE represents a binary dummy taking the value of 1 for participants who have invested in financial products in the last five years. INCOME is an ordinal variable comprised of the total annual gross income quintiles in the US. EDUCATION is a 6-item ordinal variable taking the value of 0 for participants with nursery school completed up to a value of 6 for participants with a PhD. RISK_FINANCIAL is an ordinal variable representing self-reported risk preferences on a 10-point Likert scale in the financial domain. RISK_GENERAL is an ordinal variable representing self-reported risk preferences on a 10-point Likert scale in the general domain. "Permute p" reports the *p*-values of the corresponding coefficient, obtained from permutation tests with 1000 random draws.

	Model (I)	Model (II)
CG	0.032**	0.030**
	(0.010)	(0.010)
L	0.074	0.058
	(0.041)	(0.041)
CG × L	−0.023	−0.017
	(0.014)	(0.014)
AGE		0.001
		(0.001)
MALE		0.070**
		(0.022)
EDUCATION		−0.018
		(0.011)
FINANCIAL_SECTOR		−0.069**
		(0.024)
INCOME		0.010
		(0.011)
INVEST_EXPERIENCE		0.045
		(0.025)
RISK_FINANCIAL		0.005
		(0.006)
RISK_GENERAL		0.010
		(0.006)
Constant	0.348**	0.136
	(0.029)	(0.077)
Permute p CG × L	0.125	0.225
Observations	937	937
Prob > Chi^2	0.008	0.000
Pseudo R^2	0.017	0.059

$^{*}p < 0.05$, $^{**}p < 0.005$. Dependent variable: Average amount bet in percent of endowment ($\frac{x_i}{200}$). Standard errors in parentheses.

General Instructions

Dear participant,

Thank you for participating in this online experiment!

Please read the instructions for the experiment carefully. All statements in the instructions are true. Your payoff for this experiment also depends on how well you understood the instructions. The experiment and the analyses of the data are anonymous. Your answers will only be analysed for the purpose of scientific research.

The participation takes about 15 minutes. Please note that you will only receive a payoff if you have carried out the experiment to the end.

The experiment consists of one main part and an exit questionnaire. The instructions for the main part are displayed at the beginning of the experiment and can be accessed at any time during the main experiment by clicking on the button "*Instructions*".

All personal descriptions in the experiment apply to all genders. By clicking on "*Continue*" you accept the above conditions.

Continue

Fig. A.7. General instruction.

Instructions for the main part of the experiment

The main part of the experiment consists of 9 successive rounds. In each round you will start with an amount of 200 experimental currency units (ECU). You must decide which part "*x*" of this amount (between 0 ECU and 200 ECU) you wish to bet in the following lottery.

You have a chance of 2/3 (67%) to lose the amount you bet and a chance of 1/3 (33%) to win 2.5 times the amount you bet.

Hence, your earnings in the lottery are determined as follows. If you have decided to put an amount "*x*" in the lottery, then your earnings for the round are equal to $-x$ if you lose in the lottery and equal to $+2.5 \cdot x$ if you win in the lottery. Your **total earnings** for the round are equal to 200 ECU (your starting amount) plus your earnings in the lottery.

In the following table you see a general example of the calculation of your total earnings for the first round.

Amount bet	Realization of the lottery: *Win*	Realization of the lottery: *Loss*
x ($0 \le x \le 200$)	$200 + 2.5 \cdot x$	$200 - x$

After that, you will be informed about your total earnings for the first round. Then you are requested to indicate your choice for the next round. At the beginning of the second round you again start with an amount of 200 ECU, a part "*x*" of which you can bet in the lottery. The same procedure as described above determines your earnings for the second round (same calculation as in the table above). Again, at the end of the second round you will be informed about your total earnings for the second round. All subsequent rounds will also proceed in the same manner.

Please note that your total earnings for all rounds are collected, which means that in later rounds you cannot bet money you already earned. After the last round has been completed, your total earnings for all rounds will be summed and exchanged in Dollars at a rate of 1:1500 (divided by 1500). This amount determines your bonus payoff for the experiment.

Next

Fig. A.8. Specific instruction for treatment GP and sub-treatment H.

Instructions for the main part of the experiment

The main part of the experiment consists of 9 successive rounds. In each round you will start with an amount of 200 experimental currency units (ECU). You must decide which part "*x*" of this amount (between 0 ECU and 200 ECU) you wish to bet in the following lottery.

You have a chance of 2/3 (67%) to lose the amount you bet and a chance of 1/3 (33%) to win 2.5 times the amount you bet.

Please note that you fix your choice for the next three rounds. Thus, if you decide to bet an amount "*x*" in the lottery for round 1, then you also bet the same amount "*x*" in the lottery for rounds 2 and 3. Therefore, your decision applies for 3 consecutive rounds.

Hence, your earnings in the lottery for the three rounds are determined as follows. If you have decided to put an amount "*x*" in the lottery, then your earnings are equal to $-x$ for each loss in the lottery and equal to $+2.5 \cdot x$ for each win in the lottery. Your **total earnings** for the three rounds are equal to 600 ECU (three times your starting amount of 200 ECU) plus your total earnings for the 3 lottery rounds.

In the following table you see a general example of the calculation of your total earnings for the first 3 rounds.

Amount bet	Realization of the lottery Round 1 - Round 2 - Round 3	Total earnings after 3 rounds
x ($0 \le x \le 200$)	*Win-Win-Win*	$600 + 2{,}5 \cdot 3 \cdot x$
x ($0 \le x \le 200$)	*Win-Win-Loss*	$600 - x + 2.5 \cdot 2 \cdot x$
x ($0 \le x \le 200$)	*Win-Loss-Win*	$600 - x + 2.5 \cdot 2 \cdot x$
x ($0 \le x \le 200$)	*Win-Loss-Loss*	$600 - 2 \cdot x + 2.5 \cdot x$
x ($0 \le x \le 200$)	*Loss-Win-Win*	$600 - x + 2.5 \cdot 2 \cdot x$
x ($0 \le x \le 200$)	*Loss-Win-Loss*	$600 - 2 \cdot x + 2.5 \cdot x$
x ($0 \le x \le 200$)	*Loss-Loss-Win*	$600 - 2 \cdot x + 2.5 \cdot x$
x ($0 \le x \le 200$)	*Loss-Loss-Loss*	$600 - 3 \cdot x$

After that, you will be informed about your total earnings for the first 3 rounds (1 to 3). Then you are requested to indicate your choice for the next three rounds (4 to 6). For each of the three rounds you again start with an amount of 200 ECU, a part "*x*" of which you can bet in the lottery. The same procedure as described above determines your earnings for the next three rounds (4 to 6). The subsequent three rounds (7 to 9) will also proceed in the same manner.

Please note that your total earnings for all rounds are collected, which means that in later rounds you cannot bet money you already earned. After the last round has been completed, your total earnings for all rounds will be summed and exchanged in Dollars at a rate of 1:1500 (divided by 1500). This amount determines your bonus payoff for the experiment.

Next

Fig. A.9. Specific instruction for treatment GP and sub-treatment L.

Main Part - Lottery bet

The experiment now starts with the main part.

Next

Fig. A.10. Introductory screen to the main part.

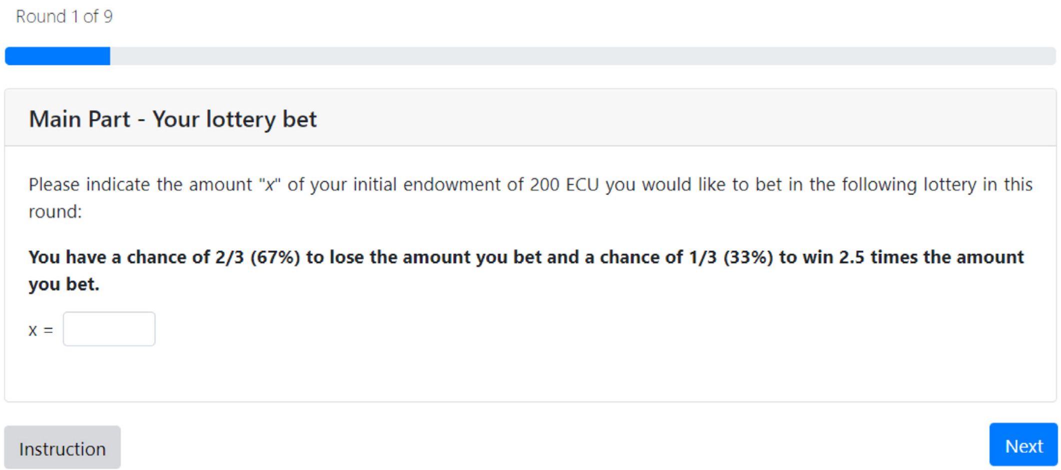


Fig. A.11. Decision screen in treatment GP and sub-treatment H.

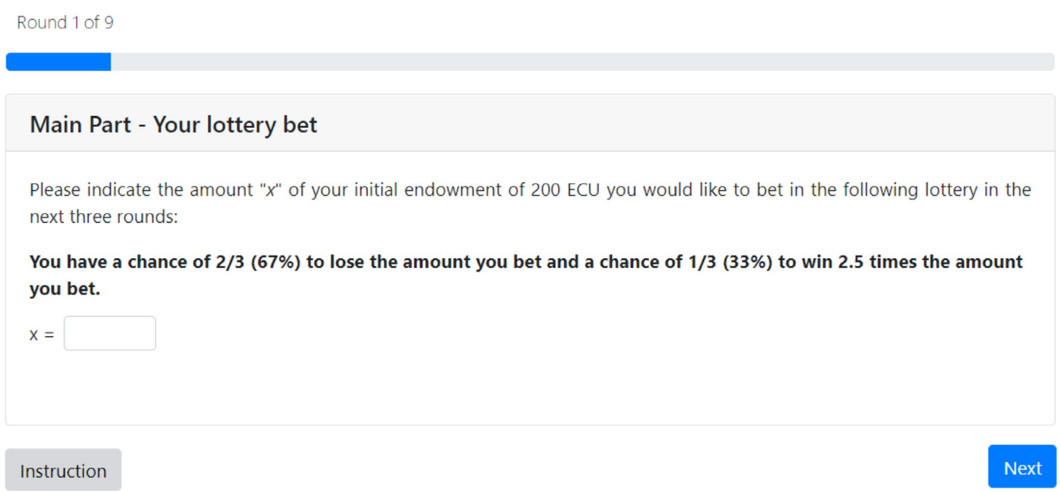


Fig. A.12. Decision screen in treatment GP and sub-treatment L.

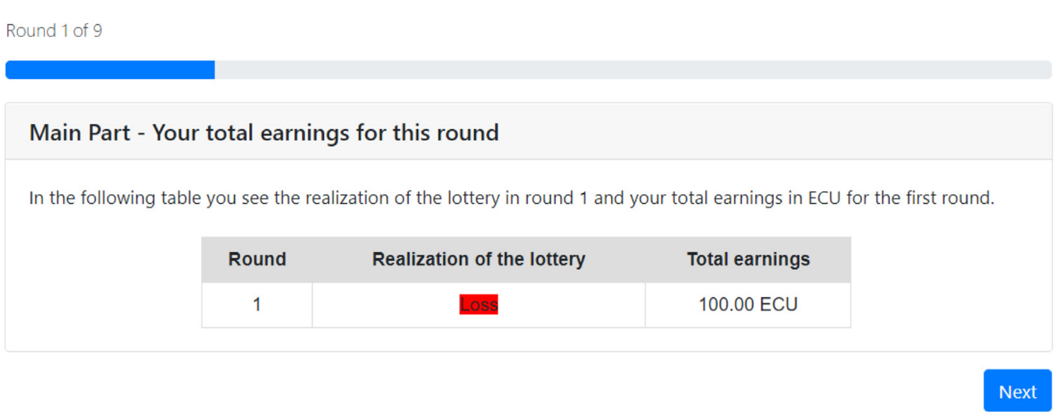


Fig. A.13. History screen in sub-treatment H.

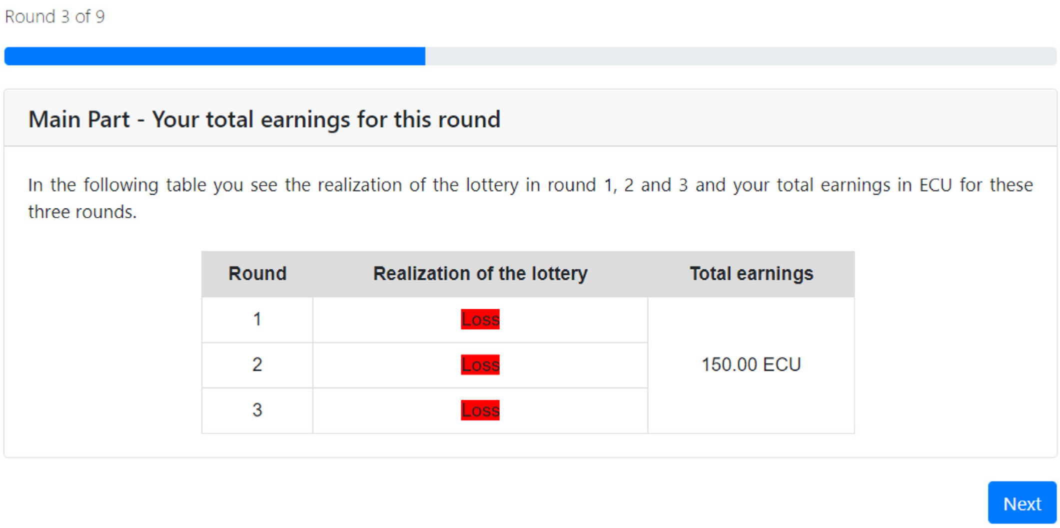


Fig. A.14. History screen in sub-treatment L.

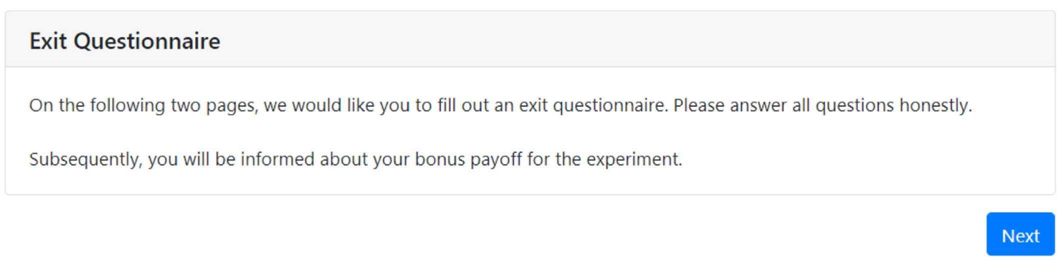


Fig. A.15. Introductory screen for the final questionnaire.

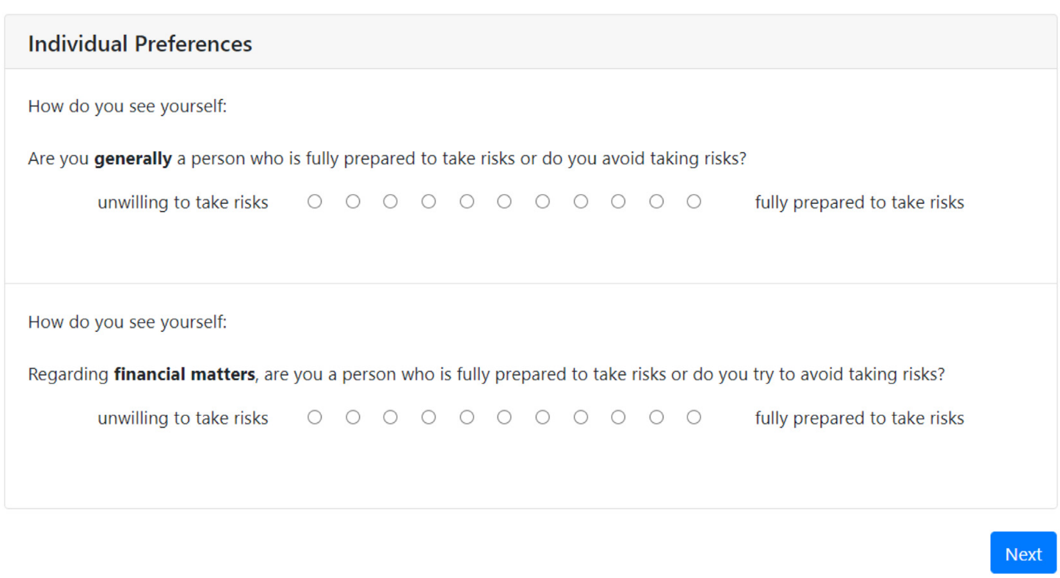


Fig. A.16. General and financial risk preferences.

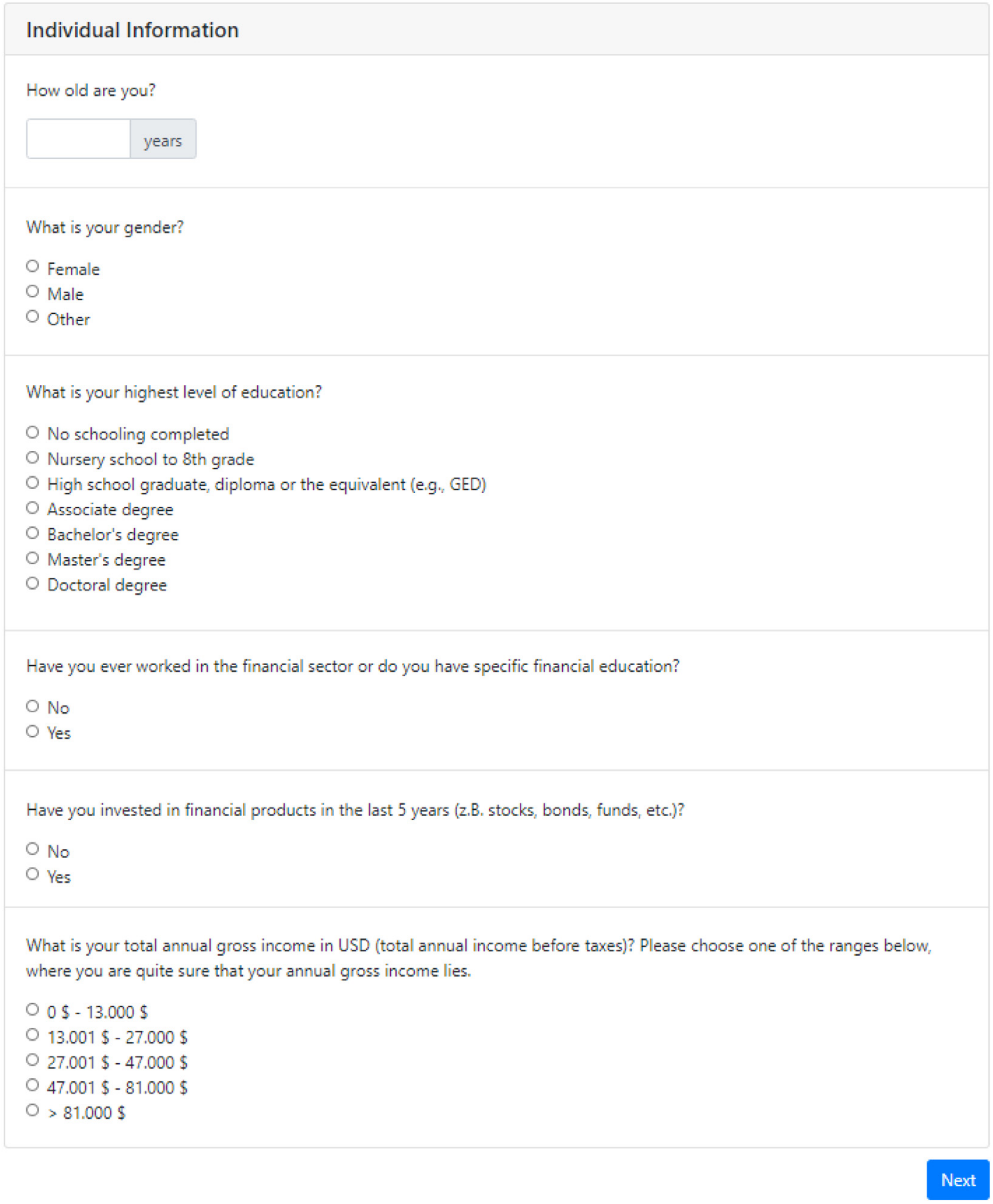


Fig. A.17. Personal questions.

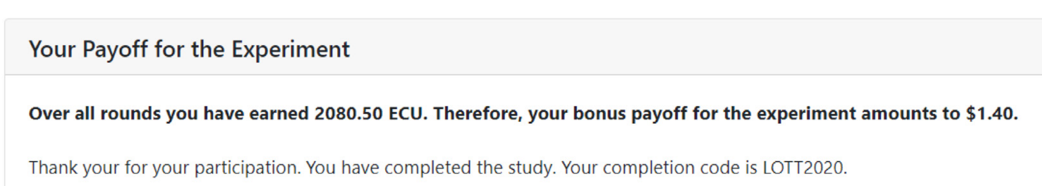


Fig. A.18. Payoff screen.

References

Aguinis, Herman, Villamor, Isabel, Ramani, Ravi S., 2021. MTurk research: Review and recommendations. J. Manag. 47 (4), 823–837. http://dx.doi.org/10.1177/0149206320969787.

Ahler, Douglas J., Roush, E., Sood, Gaurav, 2020. The Micro-Task Market for "Lemons": Data Quality on Amazon's Mechanical Turk. Working Paper, Florida State University, URL http://www.gsood.com/research/papers/turk.pdf.

Bellemare, Charles, Krause, Michaela, Kröger, Sabine, Zhang, Chendi, 2005. Myopic loss aversion: Information feedback vs. investment flexibility. Econom. Lett. 87 (3), 319–324. http://dx.doi.org/10.1016/j.econlet.2004.12.011.

Benartzi, Shlomo, Thaler, Richard H., 1999. Risk aversion or myopia? Choices in repeated gambles and retirement investments. Manage. Sci. 45 (3), 364–381. http://dx.doi.org/10.1287/mnsc.45.3.364.

Benjamin, Daniel J., Berger, James O., Johannesson, Magnus, Nosek, Brian A., Wagenmakers, Eric-Jan, Berk, Richard, Bollen, Kenneth A., Brembs, Björn, Brown, Lawrence, Camerer, Colin, et al., 2017. Redefine statistical significance. Nat. Hum. Behav. 2, 6–10. http://dx.doi.org/10.1038/s41562-017-0224-0.

Berinsky, Adam J., Huber, Gregory A., Lenz, Gabriel S., 2012. Evaluating online labor markets for experimental research: Amazon.com's mechanical turk. Political Anal. 20 (3), 351–368. http://dx.doi.org/10.1093/pan/mpr057.

Beshears, John, Choi, James J., Laibson, David, Madrian, Brigitte C., 2017. Does aggregated returns disclosure increase portfolio risk taking? Rev. Financ. Stud. 30 (6), 1971–2005. http://dx.doi.org/10.1093/rfs/hhw086.

Camerer, Colin F., Dreber, Anna, Forsell, Eskil, Ho, Teck-Hua, Huber, Jürgen, Johannesson, Magnus, Kirchler, Michael, Almenberg, Johan, Altmejd, Adam, Chan, Taizan, Heikensten, Emma, Holzmeister, Felix, Imai, Taisuke, Isaksson, Siri, Nave, Gideon, Pfeiffer, Thomas, Razen, Michael, Wu, Hang, 2016. Evaluating replicability of laboratory experiments in economics. Science 251 (6280), 1433–1436.

Chandler, Jesse, Shapiro, Danielle, 2016. Conducting clinical research using crowdsourced convenience samples. Annu. Rev. Clin. Psychol. 12, 53–81. http://dx.doi.org/10.1146/annurev-clinpsy-021815-093623.

Charness, Gary, Gneezy, Uri, 2010. Portfolio choice and risk attitudes: An experiment. Econ. Inq. 48 (1), 133–146. http://dx.doi.org/10.1111/j.1465-7295.2009.00219.x.

Chen, Daniel L., Schonger, Martin, Wickens, Chris, 2016. oTree—An open-source platform for laboratory, online, and field experiments. J. Behav. Exp. Finance 9, 88–97. http://dx.doi.org/10.1016/j.jbef.2015.12.001.

Chmielewski, Michael, Kucker, Sarah C., 2020. An mturk crisis? Shifts in data quality and the impact on study results. Soc. Psychol. Personal. Sci. 11 (4), 464–473. http://dx.doi.org/10.1177/1948550619875149.

Crump, Matthew J.C., McDonnell, John V., Gureckis, Todd M., 2013. Evaluating Amazon's mechanical turk as a tool for experimental behavioral research. PLoS One 8 (3), e57410. http://dx.doi.org/10.1371/journal.pone.0057410.

Dinno, Alexis, 2017. Tostregress: Linear regression tests for equivalence. Stata software package.

Dohmen, Thomas J., Falk, Armin, Huffman, David, Schupp, Juergen, Sunde, Uwe, Wagner, Gert, 2011. Individual risk attitudes: Measurement, determinants, and behavioral consequences. J. Eur. Econom. Assoc. 9 (3), 522–550. http://dx.doi.org/10.1111/j.1542-4774.2011.01015.x.

Dreyfuss, Emily, Barrett, Brian, Newman, Lily H., 2018. A bot panic hits amazon's mechanical turk. URL https://www.wired.com/story/amazon-mechanical-turk-bot-panic. (Last accessed: 2021-09-20).

Eriksen, Kristoffer, Kvaloy, Ola, 2010. Myopic investment management. Rev. Finance 14 (3), 521–542. http://dx.doi.org/10.1093/rof/rfp019.

Fellner, Gerlinde, Sutter, Matthias, 2009. Causes, consequences, and cures of myopic loss aversion–An experimental investigation. Econ. J. 119 (537), 900–916. http://dx.doi.org/10.1111/j.1468-0297.2009.02251.x.

Gneezy, Uri, Kapteyn, Arie, Potters, Jan, 2003. Evaluation periods and asset prices in a market experiment. J. Finance 58 (2), 821–837. http://dx.doi.org/10.1111/1540-6261.00547.

Gneezy, Uri, Potters, Jan, 1997. An experiment on risk taking and evaluation periods. Q. J. Econ. 112 (2), 631–645. http://dx.doi.org/10.1162/003355397555217.

Haigh, Michael S., List, John A., 2005. Do professional traders exhibit myopic loss aversion? An experimental analysis. J. Finance 60 (1), 523–534. http://dx.doi.org/10.1111/j.1540-6261.2005.00737.x.

Haisley, Emily, Mostafa, Romel, Loewenstein, George, 2008. Myopic risk-seeking: The impact of narrow decision bracketing on lottery play. J. Risk Uncertain. 37, 57–75. http://dx.doi.org/10.1007/s11166-008-9041-1.

Van der Heijden, Eline, Klein, Tobias J., Müller, Wieland, Potters, Jan, 2012. Framing effects and impatience: Evidence from a large scale experiment. J. Econ. Behav. Organ. 84, 701–711. http://dx.doi.org/10.1016/j.jebo.2012.09.017.

Heß, Simon, 2017. Randomization inference with stata: A guide and software. Stata J. 17 (3), 630–651.

Hueber, Laura, Schwaiger, Rene, 2021. Debiasing Through Experience Sampling: The Case of Myopic Loss Aversion. Working Papers in Economics and Statistics 2021-01, URL https://econpapers.repec.org/paper/innwpaper/2021-01.htm.

Johnson, David, Ryan, John Barry, 2020. Amazon mechanical turk workers can provide consistent and economically meaningful data. South. Econ. J. 87 (1), 369–385. http://dx.doi.org/10.1002/soej.12451.

Juzek, Tom S., Kizach, Johannes, 2019. How to Set Delta in the two-one-sided T-tests procedure (TOST). J. Res. Des. Stat. Linguist. Commun. Sci. 5 (1–2), 153–169. http://dx.doi.org/10.1558/jrds.39002.

Kahneman, Daniel, Lovallo, Dan, 1993. Timid choices and bold forecasts: A cognitive perspective on risk taking. Manage. Sci. 39 (1), 17–31. http://dx.doi.org/10.1287/mnsc.39.1.17.

Kahneman, Daniel, Tversky, Amos, 1979. Prospect theory: An analysis of decision under risk. Econometrica (ISSN: 00129682) 47 (2), 263–291. http://dx.doi.org/10.2307/1914185.

Kahneman, Daniel, Tversky, Amos, 1984. Choices, values, and frames. Am. Psychol. 39 (4), 341–350, https://psycnet.apa.org/doi/10.1037/0003-066X.39.4.341.

Kahneman, D., Tversky, A., 1992. Advances in prospect theory: Cumultative representation of uncertainty. J. Risk Uncertain. 5, 297–323.

Kennedy, Ryan, Clifford, Scott, Burleigh, Tyler, Waggoner, Philip D., Jewell, Ryan, Winter, Nicholas J.G., 2020. The shape of and solutions to the MTurk quality crisis. Political Sci. Res. Methods 8 (4), 614–629. http://dx.doi.org/10.1017/psrm.2020.6.

Keren, Gideon, Wagenaar, Willem A., 1987. Violation of utility theory in unique and repeated gambles. J. Exp. Psychol. 13 (3), 387–391. http://dx.doi.org/10.1037/0278-7393.13.3.387.

Langer, Thomas, Weber, Martin, 2001. Prospect theory, mental accounting, and differences in aggregated and segregated evaluation of lottery portfolios. Manage. Sci. 47 (5), 716–733. http://dx.doi.org/10.1287/mnsc.47.5.716.10483.

Langer, Thomas, Weber, Martin, 2005. Myopic prospect theory vs. myopic loss aversion: How general is the phenomenon? J. Econ. Behav. Organ. 56, 25–38. http://dx.doi.org/10.1016/j.jebo.2003.01.005.

Langer, Thomas, Weber, Martin, 2008. Does commitment or feedback influence myopic loss aversion? An experimental analysis. J. Econ. Behav. Organ. 67 (3–4), 810–819. http://dx.doi.org/10.1016/j.jebo.2006.05.019.

Larson, Francis, List, John A., Metcalfe, Robert D., 2012. Can Myopic Loss Aversion Explain the Equity Premium Puzzle? Evidence From a Natural Field Experiment with Professional Traders. NBER Working Paper Series 22605.

Lee, Boram, Veld-Merkoulova, Yulia, 2016. Myopic loss aversion and stock investments: An empirical study of private investors. J. Bank. Financ. 70, 235–246.

Looney, Clayton Arlen, Hardin, Andrew M., 2009. Decision support for retirement portfolio management: Overcoming myopic loss aversion via technology design. Manage. Sci. 55 (10), 1688–1703. http://dx.doi.org/10.1287/mnsc.1090.1052.

Mullinix, Kevin J., Leeper, Thomas J., Druckman, James N., Freese, Jeremy, 2015. The generalizability of survey experiments. J. Exp. Political Sci. 2 (2), 109–138. http://dx.doi.org/10.1017/XPS.2015.19.

Paolacci, Gabriele, Chandler, Jesse, Ipeirotis, Panagiotis G., 2010. Running experiments on Amazon Mechanical Turk. Judg. Decis. Mak. 5 (5), 411–419.

Papon, Thomas, 2008. The effect of pre-commitment and past-experience on insurance choices: An experimental study. Geneva Risk Insur. Rev. 33, 47–73. http://dx.doi.org/10.1057/grir.2008.8.

Sutter, Matthias, 2007. Are teams prone to myopic loss aversion? An experimental study on individual versus team investment behavior. Econom. Lett. 97 (2), 128–132. http://dx.doi.org/10.1016/j.econlet.2007.02.031.

Thaler, Richard, 1985. Mental accounting and consumer choice. Mark. Sci. 4 (3), 177–266. http://dx.doi.org/10.1287/mksc.4.3.199.

Thaler, Richard H., Tversky, Amos, Kahneman, Daniel, Schwartz, Alan, 1997. The effect of Myopia and loss aversion on risk taking: An experimental test. Q. J. Econ. 112 (2), 647–661. http://dx.doi.org/10.1162/003355397555226.

Thomas, Kyle A., Clifford, Scott, 2017. Validity and Mechanical Turk: An assessment of exclusion methods and interactive experiments. Comput. Hum. Behav. 77, 184–197. http://dx.doi.org/10.1016/j.chb.2017.08.038.

Trauzettel-Klosinski, Susanne, Dietz, Klaus, 2012. Standardized assessment of reading performance: The new international reading speed texts IReST. Invest. Ophthalmol. Vis. Sci. 53 (9).

Tryon, Warren W., Lewis, Charles, 2008. An inferential confidence interval method of establishing statistical equivalence that corrects Tryon's (2001) reduction factor. Psychol. Methods 13 (3), 272–277. http://dx.doi.org/10.1037/a0013158.

Wendy, Wendy, Asri, Marwan, 2012. Psychological biases in investment decisions: An experimental study of myopic behavior in developing capital markets. J. Indones. Econ. Bus. 27 (2), 143–158.

Zeisberger, Stefan, Langer, Thomas, Weber, Martin, 2012. Why does myopia decrease the willingness to invest? Is it myopic loss aversion or myopic loss probability aversion? Theory and Decision 72, 35–50. http://dx.doi.org/10.1007/s11238-010-9236-1.