HEIDELBERG UNIVERSITY
DEPARTMENT OF ECONOMICS

UNIVERSITÄT
HEIDELBERG
ZUKUNFT
SEIT 1386

Do women shy away from risky skill games?

Marco Lambrecht

Joerg Oechssler

AWI DISCUSSION PAPER SERIES NO. 717
June 2022

Do women shy away from risky skill games?*

Marco Lambrecht†
Hanken School of Economics & Helsinki GSE

Joerg Oechssler‡
University of Heidelberg

June 1, 2022

Abstract

A risky skill game is a game in which skill plays an important role but outcomes are also strongly influenced by random factors. Examples are poker or blackjack but also many economic activities like trading on financial markets. In an online experiment we let subjects choose how often they want to play a risky skill game. We find that women play only half as many rounds in risky skill games if the influence of chance is large. There is no gender difference if the influence of chance is small or if outcomes depend exclusively on chance.

JEL codes: C90, D82, D91, J16.
Keywords: gender, risk, competitiveness

*We thank Dirk Engelmann, Aldo Rustichini, Simeon Schudy and seminar participants at the University of Heidelberg and the Hanken School of Economics & Helsinki GSE for helpful comments. The University of Heidelberg provided the funding for this research.

†Arkadiankatu 7, PO Box 479, Fi-00101 Helsinki, email: marco.lambrecht@hanken.fi

‡Bergheimer Str. 58, 69115 Heidelberg, Germany, email: oechssler@uni-hd.de

1 Introduction

There are a number of professions where the underrepresentation of women is particularly glaring. In venture capital firms only 4.9% of partners are female (Women in VC, 2020). The picture is similar for senior roles in private equity and investment banking with the share of women being about 10%.[1] The pattern seems to hold even for some self-employed professions like day traders where possible labor market discrimination should play a smaller role (Zippia, 2021). Newell (2018) reports that the share of women among the more than 120,000 poker players entering the World Series of Poker every year is only about 5%.

The currently most accepted explanation for these findings is that such professions have the payoff structure of tournaments, introducing a strong element of competition among participants. Starting with the seminal papers of Gneezy, Niederle, and Rustichini (2003) and Niederle and Vesterlund (2007) a large literature has shown that women shy away from competition associated with such tournaments even if they are not worse in the underlying task than men.[2]

While we find the competition story fully convincing, there are some professions (e.g. day trading) where the competition aspect is less pronounced as often the participants obtain their results in private and payoffs depend on market outcomes rather than tournament outcomes. Tournaments of course have two consequences. On the one hand, they introduce competition. On the other hand they make outcomes much more risky than payment according to a piece rate, say.

It is this second aspect of riskiness on which we shall focus in this paper. We call a game a "risky skill game" if outcomes are partially determined by skill but chance plays an important role as well. Poker seems to be a good example. As shown by Duersch, Lambrecht, and Oechssler (2020) outcomes in poker are to more than 50% determined by chance. In poker it is possible that a completely unknown amateur wins the unofficial word championship, the World Series of Poker, something that would be utterly impossible in a skill game like chess.[3] We shall understand the term "game" loosely here such that it includes games against computers or games against nature. In this sense investment banking and day trading can be seen as risky skill games. Arguably, given the randomness

[1] See Prequin (undated) for private equity and Clarke (2020) for investment banking.

[2] See Niederle and Vesterlund (2011) for a survey.

[3] One famous example for this happening in poker is Chris Moneymaker's win in the WSOP in 2003 (Moneymaker and Paisner, 2005).

in the peer review process, publishing in academia can also be seen as a risky skill game.

While for many of the examples we gave, the two aspects of competition and riskiness are difficult to disentangle, in the experiments reported in this paper we are able to isolate riskiness by design. In particular, we let subjects play slot machines (one-armed bandits) with two reels and a stop button for the left reel. Subjects' payoff depends exclusively on their own performance and chance and they never learn the performance of other subjects. In the online experiment we let subjects choose how often they want to play the slot machine, where the incentives to play decrease over time. In our chance treatments, the stop button has no influence at all and outcomes depend purely on chance. In our skill treatments, subjects have (partial) control over the left reel. We have low variance treatments, in which the outcome depends only on the left reel and we have high variance treatments in which the outcomes depend on both reels and payoffs are scaled up accordingly.

Our main result is that in the High variance skill treatment, women play only half as many rounds as men even though they are just as good as men in stopping the left reel at the right moment. Yet we find no difference for the pure chance treatments or for the Low variance skill treatment. This indicates that the important factor is the interaction of skill and a significant amount of chance. The higher risk aversion of women cannot explain our finding since the payoff variance in the High variance chance treatment is even higher than in the High variance skill treatment and there we see no gender difference.

For robustness, we self-replicated our main result in a pre-registered study with high power. All of the main results replicate although the effect size is somewhat smaller in the replication. In regressions, we also control for risk aversion and several demographics. While risk aversion has the expected effect, it does not eliminate or even strongly diminish our main result.

So what could explain our finding? Further research may be needed but one possibility would be that women are more easily discouraged by a loss if they did everything correctly (stopped the left reel on time) but were unlucky with the right reel. This would be in line with earlier finding e.g. the interesting findings of Buser and Yuan (2019) that women give up more easily after losing in a math olympiad. It would also be troubling news for academia as it is certainly not rare in the publication process that one did everything correctly (wrote a good paper) but is nevertheless rejected by journals.

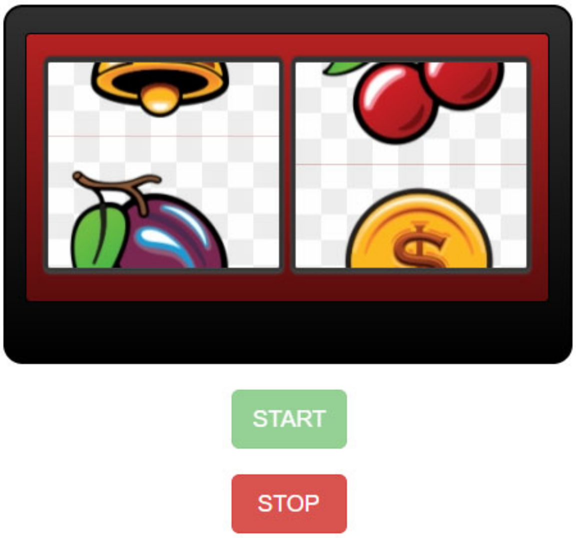


Figure 1: Snapshot of user interface with the two reels

2 Experimental Design

In the main experiment, subjects play repeatedly an online version of a slot machine or one-armed bandit. In each round they receive an endowment e and have to decide what amount $b \in \mathbb{N}$ they want to bet, with $1 \leq b \leq e$. The slot machine has two reels with 6 symbols each (see Figure 1). The right reel always stops randomly some time after the STOP-button is pressed. The behavior of the left reel depends on the treatment. In our *chance* treatments, the left reel also stops randomly. In our *skill* treatments, subjects are able to stop the left reel by pressing the stop button with a random delay, which is uniformly distributed on $[0ms, 500ms]$ to make the stopping time slightly less predictable. Nevertheless, given that the symbols always appear in the same order within a round, subjects with a good reaction time can become pretty good at stopping the left reel at the desired symbol.[4] Given the range of successful stopping one can meaningfully speak of a skill component in this gamble.

Our four treatments vary according to two dimensions (see Table 1 for an overview). As explained above, one dimension is whether the outcome depends purely on chance or

[4] Among subjects who played at least 10 rounds in treatment Low var skill, the average success rate was 57% and ranged from 20% to 86%.

at least partially on skill. The second dimension is low versus high variance. In the *Low var skill* treatment, subjects win if the left reel shows a shamrock. In this case they win 5 times their bet, $5b$. Otherwise they lose their bet. In the *High var skill* treatment, subjects win if both reels show shamrocks resulting in a payoff of $30b$, which keeps the expected payoff equal to the *Low var skill* treatment. Again they lose their bet if they do not meet the win condition.

The payoffs in our *chance* treatments are calibrated to match the average earnings of the skill treatments (assuming that subjects choose to bet their entire endowment $b = e$). In the *Low var chance* treatment, subjects are rewarded with 12 times their bet if they win ($12b$). Analogously, they win 72 times their bet, $72b$, in the *High var chance* treatment if the reels match the win condition.

Table 1: Treatments

Treatment	Left reel stops...	Win condition	Win	Obs
Low var skill	after pressing button	shamrock on left reel	$5b$	144
High var skill	after pressing button	shamrock on both reels	$30b$	151
Low var chance	randomly	shamrock on left reel	$12b$	151
High var chance	randomly	shamrock on both reels	$72b$	148

Note: In the skill treatments the left reel stops with a random delay in the range of 0 and 500ms after pressing the button.

After each round, subjects can choose to play another round or end the game. Their total payoff from this part of the experiment is the sum of the payoffs from all rounds. Our main interest in this experiment is how long subjects would play a game like this. Since we cannot force subjects to play with their own money, we have to pay them an endowment with which to play. However, in any online experiment, we can assume that subjects have opportunity cost for participating (at least the average hourly rate they could earn in another experiment that runs simultaneously). By reducing the endowment over time, we can assume that rational money maximizing subjects would stop playing at some point when the expected gain in our game is lower than their opportunity costs. We implement this by giving subjects an endowment of 50 in the first round and decrease e by 1 in each round. This yields a maximum of 50 rounds subjects could play. However, we would expect that money maximizing subjects reach their opportunity costs much earlier. Subjects who play all the way to the 50th round likely show some form of gambling disorder.

Before the main experiment subjects answer a questionnaire with a simple risk elicitation task (similar to Binswanger, 1981). Subjects have to choose one of five lotteries with two outcomes each (see Appendix). The lotteries are numbered such that smaller lottery numbers are indicative of more risk averse behavior. After the main experiment subjects are asked to fill in a questionnaire (see Appendix) with several standard questions to assess problem gambling and some demographics. To make sure that subjects stayed focused on the experiment, we include an incentivized attention check during this questionnaire.

The experiment was programmed in oTree (Chen et al. 2016) and implemented online via the platform Prolific. We were aiming for 150 gender-balanced subjects per treatment. In total, 594 subjects participated (see Table 1), of which 300 were female, 289 were male, and 5 did not provide (consistent) gender information within these two categories.[5] The sessions were run between November 2020 and June 2021. Subjects received a fixed payment of 600 points for finishing the experiment plus all points earned in the risk elicitation, all points earned in the main phase of the experiment, and all points earned in attention checks during the questionnaire. Points earned in the experiment were converted into GBP at a ratio of 800 points = 1 GBP. The average payment was approximately 2.60 GBP for less than 15 minutes. Some subjects spent more than 45 minutes on the experiment, and the highest payments exceeded 10 GBP. We include the instructions and screenshots of our experiment in the appendix.

3 Results

We start by analyzing the main variable if interest, namely how many rounds subjects decided to play. On average across all treatments, subjects played 11.8 rounds, with a minimum of 1 and a maximum of 50. While male subjects played 13.7 rounds (95% confidence interval $CI = [11.8, 15.5]$), females played only 10.1 rounds on average ($CI = [8.6, 11.6]$), which is significantly different according to a two-sided t-test with $p = .003$.[6] However, this difference is driven entirely by our *High var skill* treatment as can be seen in Figure 2, which shows the average number of rounds played separated by treatment and gender. There is no significant difference in the length of play for any of our other treatments across gender (t-tests, $p > .18$). The average rounds played by females in *High*

[5] For two subjects, the gender information provided by Prolific did not match their answers in our debriefing survey. Another two subjects responded to be of "other" gender, and one subject preferred not to answer. For all gender related questions we exclude those 5 subjects and only use the remaining 589 subjects.

[6] We use t-tests since the number of observations is sufficiently high. Using non-parametric MWU-tests gives qualitatively identical results.

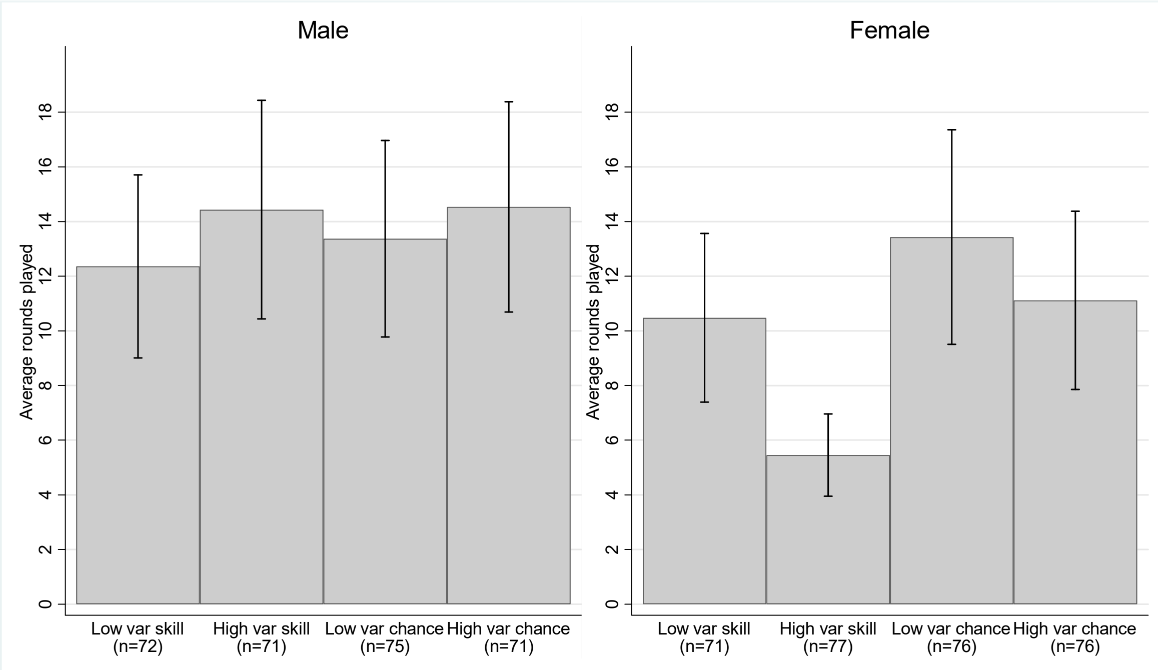


Figure 2: Average rounds played by subjects seperated by treatment and gender
Note: The error bars show 95%-confidence intervals around the mean.

var skill is significantly lower than all other bars (t-tests, $p < .01$). In particular, in *High var skill*, females play 5.45 rounds on average ($CI = [3.95, 6.96]$) while males play 14.44 rounds on average ($CI = [10.44, 18.44]$), t-test, $p < .001$). Thus, the main result of the paper is

Result 1 *In skill gambles with high variance, females play less than half as many rounds as males do.*

Further evidence for Result 1 comes from the simple OLS regression in Table 2, column (1) where we regress the number of rounds on dummies for all combinations of treatment and gender (with High var skill/male as baseline). The number of rounds played by females in High var skill is lower by almost 9 ($p = .004$). All other treatment/gender combinations are not significantly different from each other.

Table 2: OLS

	(1) rounds_played	(2) rounds_played	(3) rounds_played
low var skill male	−2.08	−2.18	−1.57
	(2.44)	(2.42)	(2.45)
low var skill female	−3.96	−3.43	−3.11
	(2.45)	(2.44)	(2.48)
high var skill female	−8.98***	−7.79***	−7.42***
	(2.40)	(2.42)	(2.46)
low var chance male	−1.06	0.24	0.44
	(2.42)	(2.48)	(2.50)
low var chance female	−1.00	0.66	0.13
	(2.41)	(2.45)	(2.49)
high var chance male	0.10	1.04	0.97
	(2.45)	(2.52)	(2.54)
high var chance female	−3.32	−1.29	−1.01
	(2.41)	(2.50)	(2.55)
risk tolerance		1.37***	1.25***
		(0.45)	(0.46)
avg. left reel correct		3.79	3.08
		(2.38)	(2.41)
gambling proclivity			−0.09
			(0.19)
demographic controls	no	no	yes
constant	14.44***	9.15***	11.06**
	(1.73)	(2.32)	(5.18)
N	589	589	589

Note: Standard errors in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

What could potentially explain this finding? It is well established that females have usually higher risk aversion than men. Since the variance of total payoff is, ceteris paribus, increasing in the rounds played, risk aversion may be one candidate for Result 1. To check for this explanation, we add the lottery choice in our risk elicitation task (column 2 in Table 2). While risk aversion has a significant effect with the expected sign (recall that higher lottery numbers correspond to less risk aversion), the coefficient of High var skill/female is only slightly smaller at -7.79. Questions 10 to 18 in the the questionnaire (see appendix) asked subjects 9 typical questions connected to gambling addiction (see e.g. Derevensky and Gupta, 2000, and Arthur et al., 2008). We built an index by summing the score of the 10 questions and call it gambling proclivity. The main result of the regressions is also robust to including the gambling proclivity index (which is not significant) and other demographics (age, income, education, profession, and religion) in column 3 of Table 2.

Table 3: Standard deviations and skewness of payoffs

	payoff per b		total game payoff	
	σ	skewness	σ	skewness
High var skill	7.13	3.72	928.3	2.93
High var chance	13.05	5.16	909.0	2.79

Note: Payoff per b is the average (across subjects) of the payoff subjects obtained by betting one point. Total game payoff is the average total payoff (across subjects) subjects obtained in the whole game (incl. endowments).

Another reason why risk aversion is unlikely to be an important driver for our main result is that the standard deviation of payoffs in High var chance is even higher than those in High var skill and yet females have no hesitation to play in the High var chance treatment. Table 3 shows the standard deviation and skewness of payoffs subjects obtained by betting one point in the two treatments. Standard deviation in High var chance is almost twice as high as in High var skill. Table 3 also shows the standard deviations (across subjects) of average total payoff subjects obtained in the whole game (incl. endowments), which are roughly the same in both treatments. However, total payoffs are of course endogenous as they depend on the bet size and the number of rounds, which are both chosen by subjects.

Most people seem to have a preference for skewness (see e.g. Golec and Tamarkin, 1998, and Mitton and Vorkink, 2007). When looking at payoff per bet b in Table 3 we see that skewness in High var chance is higher, which could potentially explain why playing

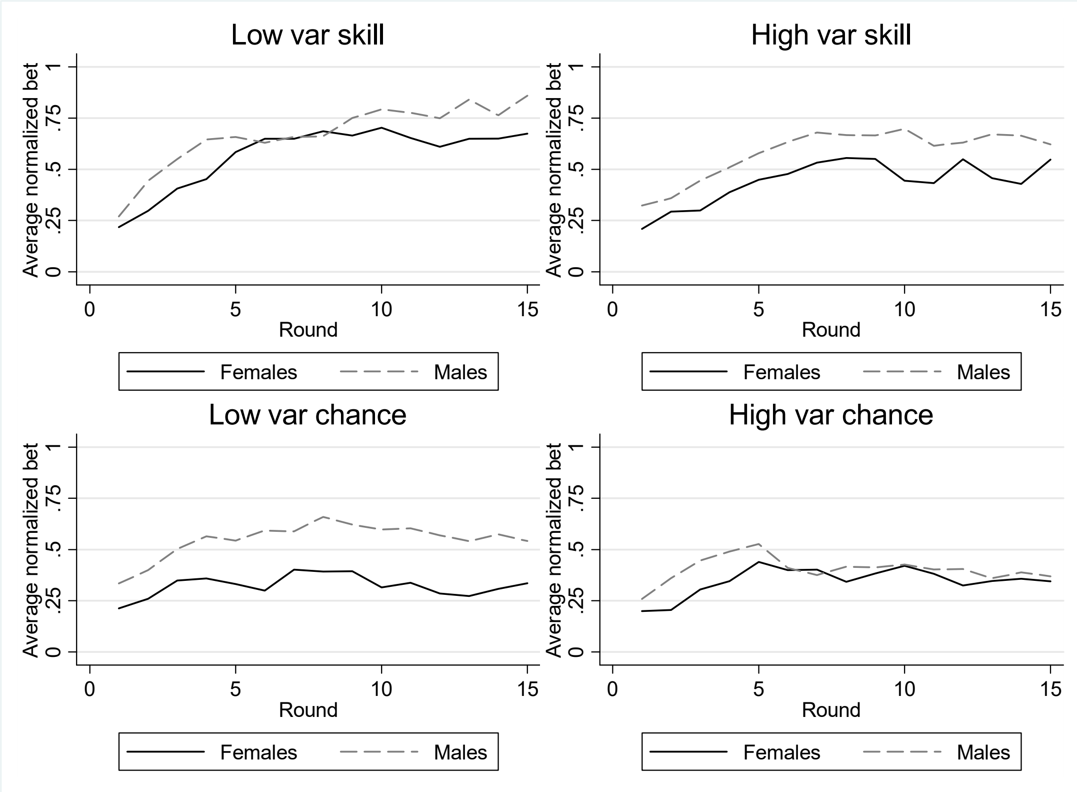


Figure 3: Average normalized bets by treatment and gender.
Note: Average normalized bets are calculated as percentage of the endowment in the current round.

in treatment High var chance seems more attractive.[7] However, in a repeated setting, one should look at (endogenous) total skewness (Ebert, 2020) and here the difference between the two treatments is small.

Could Result 1 potentially be explained by females compensating their smaller number of betting rounds by simply making larger bets in each round? Figure 3 shows average normalized bets which is the percentage of their endowment which subjects wager in each round and it shows that, to the contrary, males make larger bets in all treatments (t-tests, $p < .026$) except High var chance ($p = .066$), where bets are weakly larger.

Of course, one may wonder whether women are simply worse in stopping the left reel at a shamrock. However, Figure 4 shows that this is clearly not the case in treatment Low

[7]However not why this effect would only work for females.

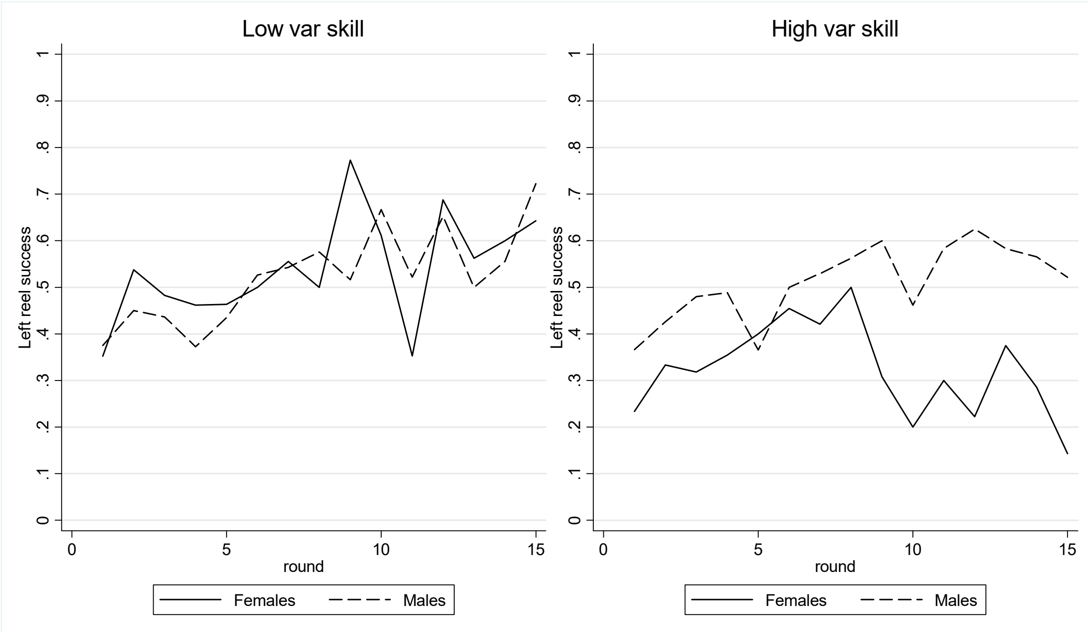


Figure 4: Share of successful stops of left reel

var skill. Both genders show an upward learning curve but neither in the first round nor on average in the first 15 rounds is there a significant difference in left reel success (t-tests, $p > .58$). Interestingly, this is different in High var skill for women. Men in high var skill show a very similar trend as in Low var skill. Women's success is already lower in the first round (t-test, $p = .079$) and the difference in the first 15 rounds is highly significant ($p = .0007$). Thus, it seems likely that women in High var skill are less motivated to perform well even though, motorically, they are just as good as men in stopping the reel.

Finally, we test whether women give up more easily after disappointing outcomes as suggested by Buser and Yuan (2019). We concentrate on treatment High var skill and run separate Cox proportional hazards models for men and women regarding the probability of stopping to play . In High var skill, there are two kind of potentially frustrating events: One can lose a bet due to the fact that one did not manage to stop the left reel correctly even though one was lucky with the right reel. Thus, the loss should be attributed to ability. Or, one can lose a round due to being unlucky with the right reel even though one stopped the left reel correctly. Thus, the loss should be attributed to bad luck. We enter dummies for these events as frust_ability and frust_luck, respectively.

The other explanatory variables in the Cox regressions are a dummy whether a subject has won in the current round (won), the current balance of accumulated payoffs (balance), and the number of rounds since the last win. Furthermore, we add some subject specific variables like the percentage of correctly stopping the left reel (as a measure of skill) and the choice from the risk elicitation task. Specifications (3) and (4) additionally contains demographics like age, income, as well as profession and religion categories.

The results in Table 4 show that frustration due to bad luck does not have a significant effect on the stopping probabilities in any regression. For women, however, we see a negative effect of frustration due to ability on stopping, contrary to what one could expect. Possibly women see such a missed opportunity rather as an encouraging sign that winning is possible. Having won in a current round increases the probability of stopping, but only for men. Balance is significant for women but the coefficients are very small. More skill (as measured by the percentage of stopping the left reel correctly) decreases the stopping probability, which seems plausible.

4 Replication of the main result

There could be some concern that our main Result 1 is an accidental finding as we have multiple possible hypotheses tests. To account for this we attempt a self-replication of the main result for treatment High var skill, which is preregistered and has sufficient power. We pre-registered the replication at the AEA RCT Registry (AEARCTR-0008858). The replication was run in February 2022, again on Prolific, with the same subject pool, the same software, and the same incentives. We recruited 152 new subjects, of which 75 were males, 75 were females and 2 did not provide (consistent) gender information.[8] Given the effect size we observed in High var skill of the original experiment, 150 subjects would result in a power of 99%.

Figure 5 shows the average rounds played by males and females in our replication of High var skill. While the difference is slightly smaller than in Figure 2, it is still quite large and significant (t-tests, $p = .011$). We also run the equivalent OLS regression for rounds_played

[8] As in the main experiment, we used only those subjects whose self-declared gender information agreed with that stored at Prolific.

Table 4: Cox proportional hazards model: probability of stopping

	(1) male	(2) female	(3) male	(4) female
frust_luck	0.08	0.06	0.07	0.03
	(0.30)	(0.30)	(0.30)	(0.29)
frust_ability	−0.88	−0.90*	−0.85	−1.00**
	(0.60)	(0.46)	(0.60)	(0.43)
won	1.47***	0.02	1.53***	0.04
	(0.42)	(0.45)	(0.40)	(0.43)
balance	0.00	0.00***	0.00	0.00***
	(0.00)	(0.00)	(0.00)	(0.00)
avg. left reel correct	−1.85***	−1.87***	−1.94**	−1.51*
	(0.71)	(0.70)	(0.86)	(0.81)
rounds since win	−0.03	−0.03	−0.03	−0.04
	(0.03)	(0.05)	(0.03)	(0.05)
risk tolerance	−0.11	−0.02	−0.11	−0.01
	(0.08)	(0.09)	(0.11)	(0.09)
demographic controls	no	no	yes	yes
N	1025	420	1025	420

Note: Clustered standard errors in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

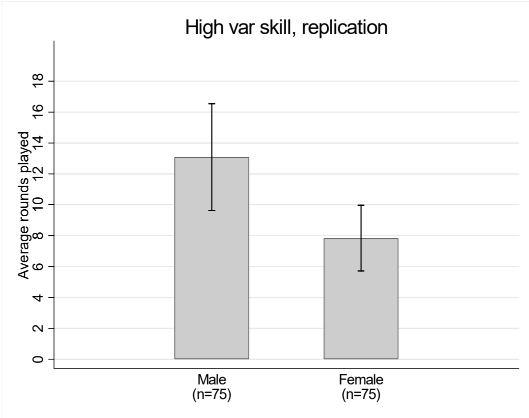


Figure 5: Average rounds played by subjects in replication treatment High var skill seperated by gender
Note: The error bars show 95%-confidence intervals around the mean.

and the Cox proportional hazards model for the probability of stopping. The results are shown in Appendix B in Tables 2 and 3. While some covariates are different and significant in the replication and not in the original study or vice versa, the main qualitative results are robust. In the replication, we also find no significant difference anymore in the ability of stopping the left reel at a shamrock between genders, neither for the first (t-tests, $p = .387$) nor for the first 15 rounds (t-tests, $p = .158$). This is consistent with our finding from Low var skill in the main experiment but makes it even more puzzling that women decide to stop so much earlier in High var skill.

5 Conclusion

To explain the remarkable underrepresentation of women in some professions where skill and chance interact strongly, we study risky skill games in an online experiment. We find no difference in the duration of play between men and women when outcomes depend exclusively on chance. We also find no difference when outcome depend almost exclusively on skill. However, when a win requires both skill and luck, we find that women play significantly fewer rounds than men do even though they are equally good at performing the task.

Thus, besides the well-established fact that women shy away from competition, we have identified another and complementary reason why women are underrepresented in some professions. In our design, there is no competition as subjects never learn the performance of other subjects and their payoff is exclusively based on their own actions and/or chance. In this sense our results are in line with some recent evidence that a gender gap can occur even in the absence of competition (van Veldhuizen, 2021).

One possible explanation for our results is that women are more likely than men to attribute a loss to lacking ability rather than just bad luck. In fact, Alnamlah and Gravert (2020) show that women are more likely to drop out when they are being told that lacking ability was the cause of failure and less likely to drop out when being told that luck was to blame. However, this explanation seems less applicable for our experiment since subjects could easily see whether their loss was due to skill (getting the wrong symbol on the left reel) versus due to chance (when the right reel showed the wrong symbol).

References

[1] Alnamlah, M. and Gravert, C. (2020), She Could Not Agree More: The Role of Failure Attribution in Shaping the Gender Gap in Competition Persistence, mimeo, University of Copenhagen.

[2] Arthur, D., Tong, W. L., Chen, C. P., Hingπ, A. Y., Sagara-Rosemeyer, M., Kua, E. H., & Ignacio, J. (2008). The validity and reliability of four measures of gambling behaviour in a sample of Singapore university students. *Journal of Gambling Studies*, 24(4), 451-462.

[3] Chen, D. L., Schonger, M., & Wickens, C. (2016). oTree—An open-source platform for laboratory, online, and field experiments. J*ournal of Behavioral and Experimental Finance*, 9, 88-97.

[4] Clarke, P. (2020), Under 10% of top City dealmakers are women: 'It's still very testosterone-fuelled', *Financial News*, August 10, available at `https://www.fnlondon.com/articles/under-10-of-top-city-dealmakers-are-women-its-still-very-testosterone-fuelled-20200810`

[5] Derevensky, J. L., & Gupta, R. (2000). Prevalence estimates of adolescent gambling: A comparison of the SOGS-RA, DSM-IV-J, and the GA 20 questions. *Journal of Gambling Studies*, 16(2), 227-251.

[6] Ebert, S. (2020). On Taking a Skewed Risk More Than Once. Available at SSRN 3731565.

[7] Gneezy, U., Niederle, M., & Rustichini, A. (2003). Performance in competitive environments: Gender differences. *Quarterly Journal of Economics*, 118(3), 1049-1074.

[8] Golec, J., & Tamarkin, M. (1998). Bettors love skewness, not risk, at the horse track. *Journal of Political Economy*, 106(1), 205-225.

[9] Mitton, T., & Vorkink, K. (2007). Equilibrium underdiversification and the preference for skewness. *The Review of Financial Studies*, 20(4), 1255-1288.

[10] Moneymaker, C., & Paisner, D. (2005). Moneymaker: How an amateur poker player turned $40 into $2.5 million at the World Series of Poker. Harper Collins.

[11] Newell, J. (2018), WSOP Female Participation Lowest in Six Years, `https://www.legaluspokersites.com/news/wsop-female-participation-lowest-in-six-years/14918/`

[12] Niederle, M., & Vesterlund, L. (2007). Do women shy away from competition? Do men compete too much?. *Quarterly Journal of Economics*, 122(3), 1067-1101.

[13] Niederle, M., & Vesterlund, L. (2011). Gender and competition. *Annu. Rev. Econ.*, 3(1), 601-630.

[14] Prequin (undated), Women in private equity, available at `https://docs.preqin.com/reports/Preqin-Women-in-Private-Equity-February-2019.pdf`

[15] van Veldhuizen, R. (2021), Gender Differences in Tournament Choices: Risk Preferences, Overconfidence or Competitiveness? forthcoming, *Journal of the European Economic Association.*

[16] Women in VC (2020), The Untapped Potential of Women-led Funds, available at `https://assets.ctfassets.net/jh572x5wd4r0/7qRourAWPj0U9R7MN5nWgy/711a6d8344bcd4fbe0f1a6dcf766a3c0/WVC_Report_-_The_Untapped_Potential_of_Women-Led_Funds.pdf`

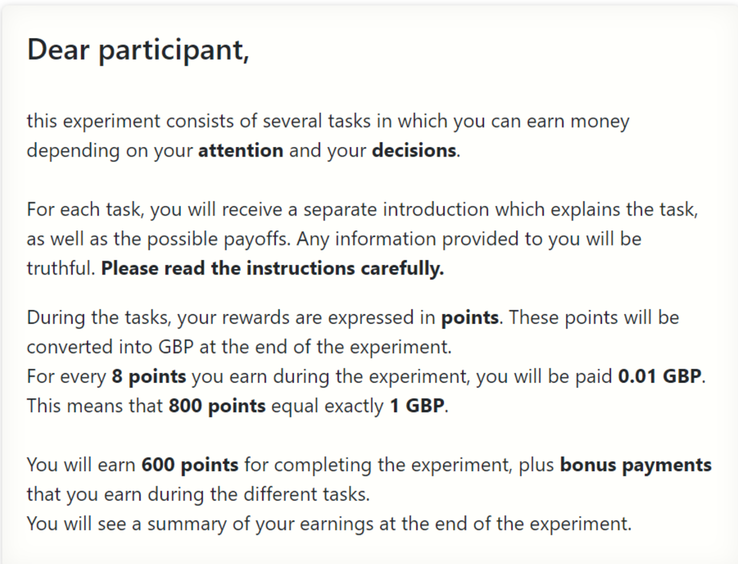
[17] Zippia (2021), Day trader statistics by gender, available at `https://www.zippia.com/day-trader-jobs/demographics/`

A Experimental Instructions & Details

Table 1 provides the order of stages in our experiment. Subsection A.1 shows our risk elicitation task. Depending on our treatment, we provided our participants with instructions describing the game stage. We include the instructions for treatment *Low var skill* in subsection A.2. Note that the instructions for our other treatments are identical, except for the behavior of the first wheel, the win condition and the reward parameter for winning. In subsection A.3, we include the comprehension quiz questions which participants had to pass before proceeding to the practice round of the game stage. Subsection A.4 shows fictional screens similar to those that participants could see during the game stage. Finally, we provide the instructions and questions of our debriefing survey in subsection A.5.

Table 1: Stages in our experiment

	Stage
1	Welcome screen
2	Risk elicitation task
3	One armed bandit instructions
4	One armed bandit quiz
5	One armed bandit trial round
6	One armed bandit game
7	Debriefing survey
8	Payoff summary


Dear participant,

this experiment consists of several tasks in which you can earn money depending on your **attention** and your **decisions**.

For each task, you will receive a separate introduction which explains the task, as well as the possible payoffs. Any information provided to you will be truthful. **Please read the instructions carefully.**

During the tasks, your rewards are expressed in **points**. These points will be converted into GBP at the end of the experiment.
For every **8 points** you earn during the experiment, you will be paid **0.01 GBP**.
This means that **800 points** equal exactly **1 GBP**.

You will earn **600 points** for completing the experiment, plus **bonus payments** that you earn during the different tasks.
You will see a summary of your earnings at the end of the experiment.

Figure 1: Welcome screen

A.1 Risk elicitation task

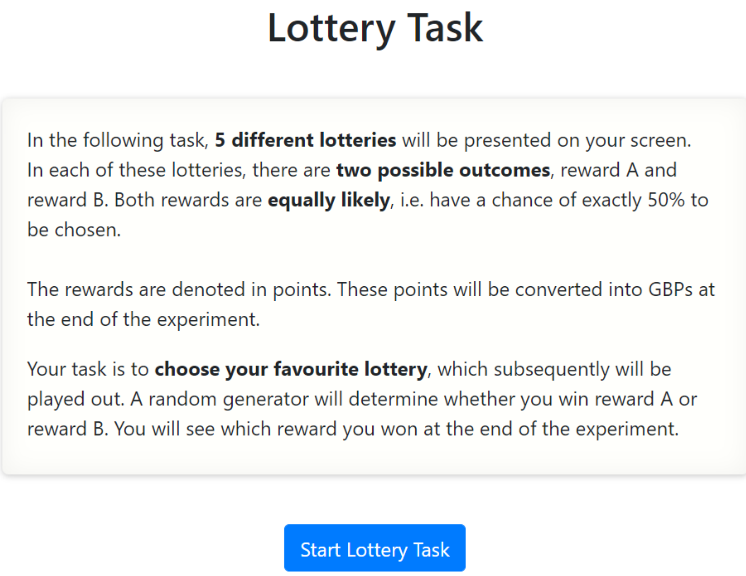
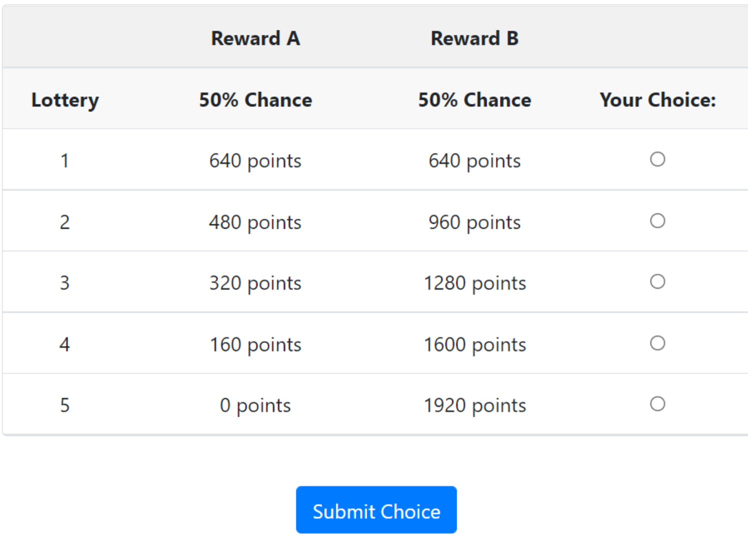


Figure 2: Lottery task instructions



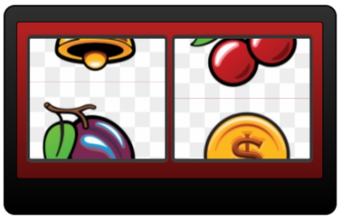
	Reward A	Reward B	
Lottery	**50% Chance**	**50% Chance**	**Your Choice:**
1	640 points	640 points	○
2	480 points	960 points	○
3	320 points	1280 points	○
4	160 points	1600 points	○
5	0 points	1920 points	○

Figure 3: Lottery task decision screen

A.2 Game stage instructions

The next stage of this experiment is a **reaction time game**. Before the start of the game, you will go through a **comprehension quiz**. You have to answer all questions correctly in order to proceed. If you fail to answer all questions correctly, you will come back to this page.

In the game, you will see 2 windows with wheels of 6 different pictures flashing. You can stop the left wheel by pressing a button labelled **"STOP"**. However, this button does not react instantly, so stopping could be **delayed** slightly. The second wheel will stop **randomly** and all pictures are equally likely.



You win the game if **the left wheel shows a shamrock**. If the left wheel does not show a shamrock, you lose the game.

You can **choose how many rounds you want to play**. After each round, you will be asked whether you want to play another round or not. If you decide not to play another round, the game ends.

How to earn money?

At the beginning of each round, you will receive some points. In the first round, you will receive **50 points**. The points you receive will **decrease by 1** after each round. In the second round, you will receive 49 points and so on. In case you continue playing until you reach 0 points, the game will end automatically.

Before the game starts, you must decide how many points you want to bet. The minimum bet is 1 point. The maximum bet is all points you received for this round. If you win the game, you get **5 times** your bet added to your account. If you lose the game, you lose the points you bet. The points that you do not bet will go directly into your account.

Here is an example. Suppose you receive 45 points and you bet 10 points. If you win the game, 85 points are added to your account since you earn 5 times your bet of 10 points, plus the 35 points that you did not bet. If you lose the game, only 35 points are added to your account.

After each round, you will be asked whether you want to play another round or leave the game and move on to the next stage. If you decide to leave the game, your account balance will be saved and converted into GBP at the end of the experiment.

Before the actual game starts, there will be a **trial round** which will **not affect your payoff**. This is to help you get familiarized with the interface and the game.

Attention:
For technical reasons, you will **lose your bet** if you **reload/refresh the game page**.

Please make sure to NOT reload/refresh the game page.

Figure 4: Game stage instructions

A.3 Game stage quiz

- How many points will you receive in your first round?
(40, 50, 60 - *Correct answer: 50*)
- **Low var treatments:** What is your win condition?
(Shamrock at the left, Coin at the right, Lemon at both - *Correct answer: Shamrock at the left*)
- **High var treatments:** What is your win condition?
(Lemon at the left, Coin at the right, Shamrock at both - *Correct answer: Shamrock at both*)
- **Skill treatments:** What determines the picture of the second wheel?
(The second wheel stops randomly, The STOP-button influences the second wheel, The picture of the first wheel influences the second wheel - *Correct answer: The second wheel stops randomly*)
- **Chance treatments:** What determines the pictures of the wheels?
(The wheels stop randomly, The STOP-button influences the wheels, The picture of the first wheel influences the second wheel - *Correct answer: The wheels stop randomly*)
- What is the minimum bet that you can place?
(0, 1, 10 - *Correct answer: 1*)

A.4 Screenshots of one armed bandit stage

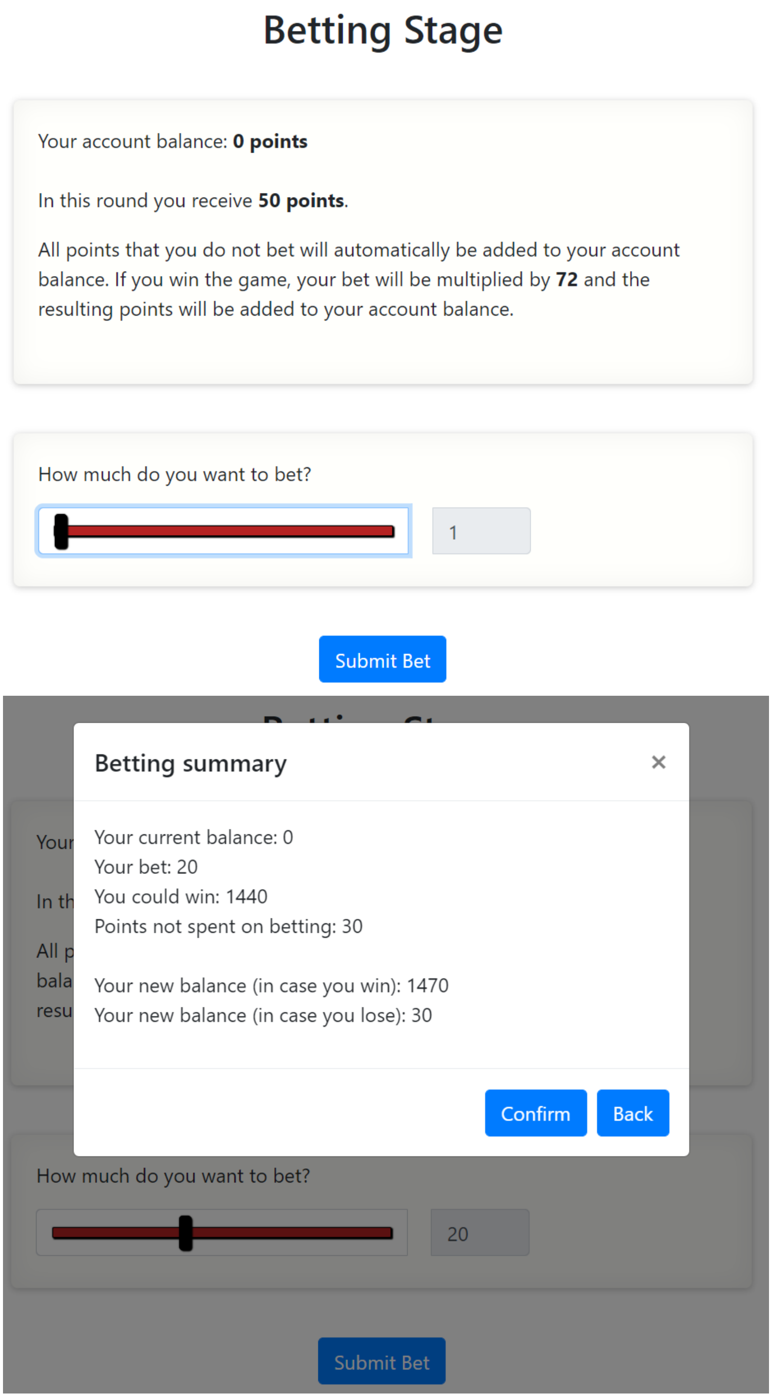


Figure 5: Game stage betting screens

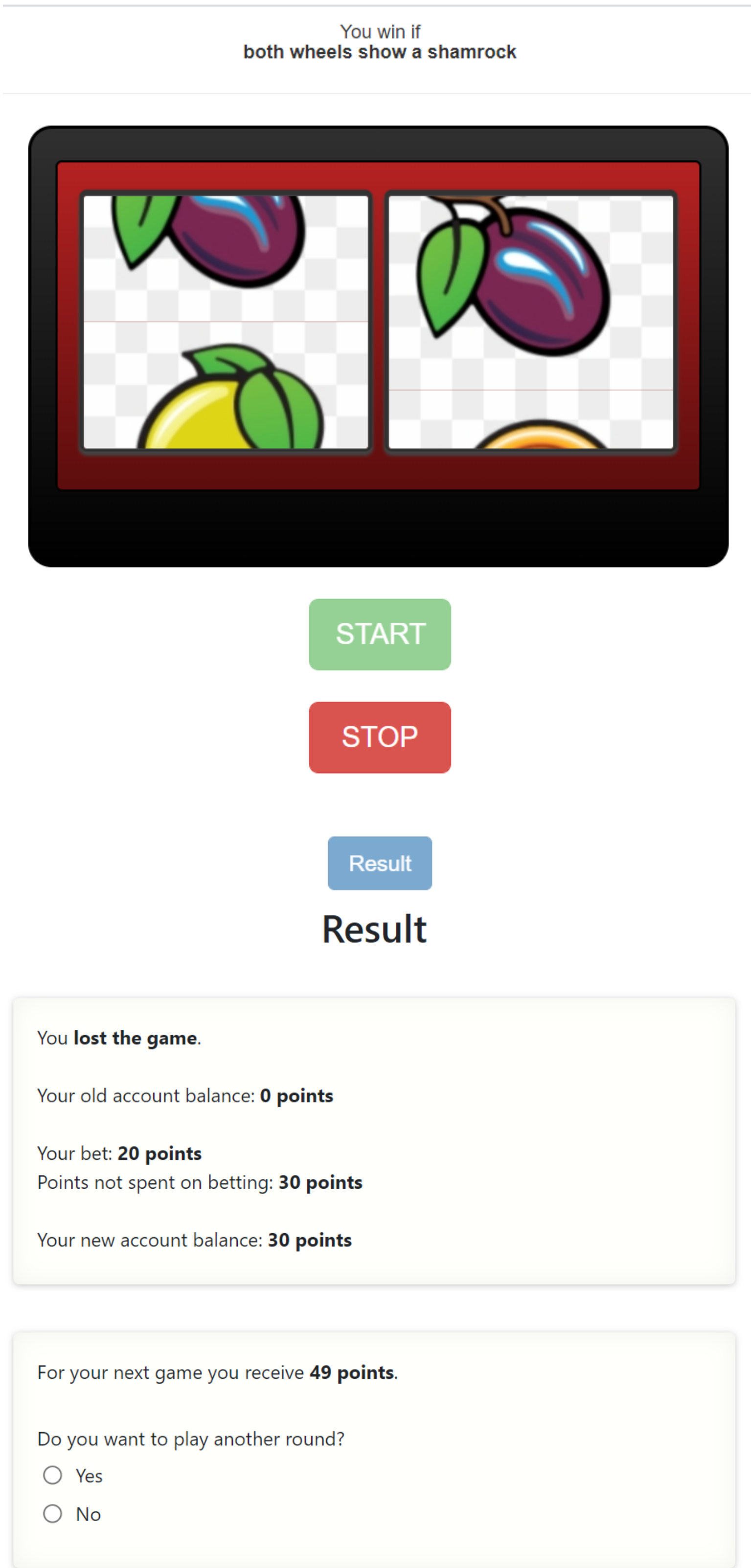


Figure 6: Game stage screens

A.5 Debriefing survey

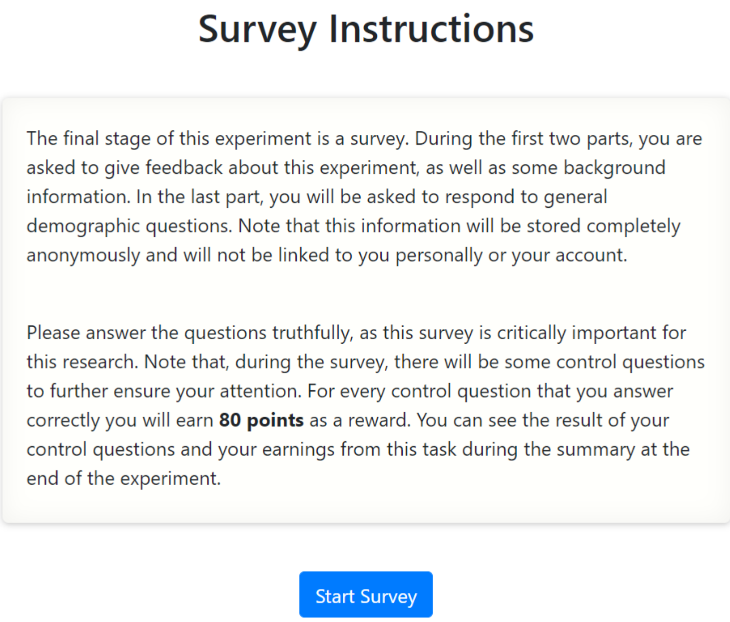


Figure 7: Debriefing survey instruction screen

1. I played the game longer than I wanted to.
   (Strongly agree, Agree, Disagree, Strongly disagree)
2. Sometimes I felt like stopping the game, but yet I played another round.
   (Strongly agree, Agree, Disagree, Strongly disagree)
3. It was stressful to play the game.
   (Strongly agree, Agree, Disagree, Strongly disagree)
4. While playing the game, I forgot worries I had on my mind.
   (Strongly agree, Agree, Disagree, Strongly disagree)
5. At times, I got angry during the game.
   (Strongly agree, Agree, Disagree, Strongly disagree)
6. When I lost a round, I wanted to win back my losses as soon as possible.
   (Strongly agree, Agree, Disagree, Strongly disagree)
7. When I won a round, I felt the urge to win even more.
   (Strongly agree, Agree, Disagree, Strongly disagree)
8. Praying helped me to win.
   (Strongly agree, Agree, Disagree, Strongly disagree)
9. I regret playing the game.
   (Strongly agree, Agree, Disagree, Strongly disagree)
10. How often do you play slot machines (for example, in casinos or pubs)?
    (Very frequently, Occasionally, Rarely, Never)
11. How often do you play games for money?
    (Very frequently, Occasionally, Rarely, Never)

12. How often do you bet money on sports competitions?
    (Very frequently, Occasionally, Rarely, Never)
13. How often did you get in trouble because of playing/betting?
    (Very frequently, Occasionally, Rarely, Never)
14. How often did you borrow money in order to play or bet?
    (Very frequently, Occasionally, Rarely, Never)
15. How often did you lie to others related to playing and betting?
    (Very frequently, Occasionally, Rarely, Never)
16. When chance is against you this round, it is more likely that chance will be in your favour in the next round.
    (Strongly agree, Agree, Disagree, Strongly disagree)
17. When chance is in your favor this round, it will most likely still be in your favour in the next round.
    (Strongly agree, Agree, Disagree, Strongly disagree)
18. You can influence the outcome of chance draws by applying systematic strategies.
    (Strongly agree, Agree, Disagree, Strongly disagree)

B Additional results of the replication

Table 2: OLS in replication

	(1) rounds_played	(2) rounds_played	(3) rounds_played
high var skill female	-5.24**	-4.46**	-3.77*
	(2.04)	(2.02)	(2.27)
risk tolerance		-0.13	-0.23
		(0.68)	(0.71)
avg. left reel correct		10.46***	9.52**
		(3.55)	(3.82)
gambling proclivity			-0.13
			(0.30)
demographic controls	no	no	yes
constant	13.08***	8.67***	23.01***
	(1.44)	(2.82)	(8.60)
N	150	150	150

Note that High var skill/male serves as baseline. Standard errors in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

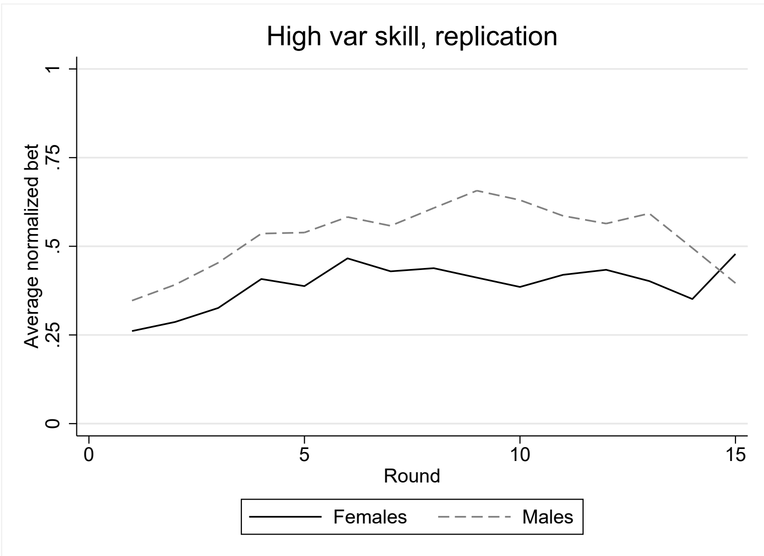


Figure 8: Average normalized bets by gender in replication.
Note that bets for females are significantly lower (t-test, $p < .001$).

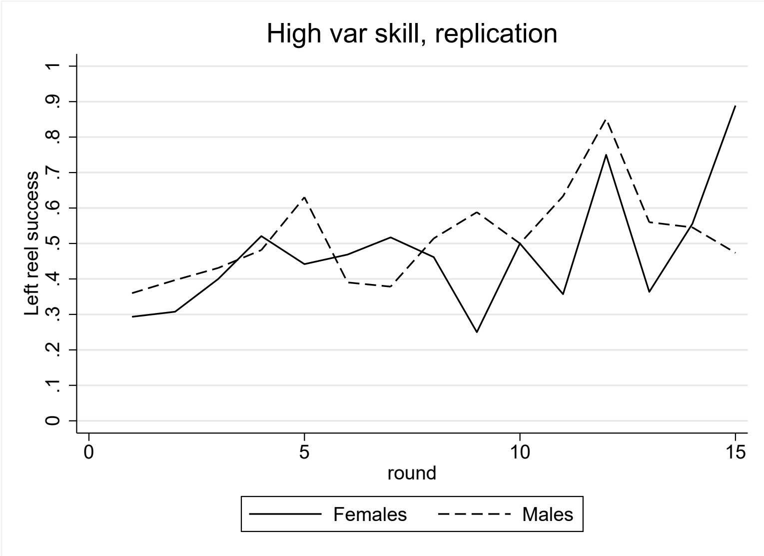


Figure 9: Share of successful stops of left reel in replication

Table 3: Cox proportional hazards model: probability of stopping in replication

	(1) male	(2) female	(3) male	(4) female
frust_luck	0.55*	0.16	0.57*	0.18
	(0.29)	(0.27)	(0.30)	(0.28)
frust_ability	0.47	-0.70	0.55	-0.70
	(0.42)	(0.52)	(0.43)	(0.53)
won	1.48***	1.46***	1.51***	1.56***
	(0.33)	(0.33)	(0.35)	(0.37)
balance	0.00***	-0.00	0.00***	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)
avg. left reel correct	-2.91***	-1.71**	-2.93***	-2.07**
	(0.86)	(0.69)	(1.11)	(0.82)
rounds since win	0.07**	0.02	0.06*	0.03
	(0.03)	(0.03)	(0.03)	(0.04)
risk tolerance	-0.02	0.05	-0.02	0.08
	(0.07)	(0.09)	(0.08)	(0.09)
demographic controls	no	no	yes	yes
N	981	588	981	588

Note: Clustered standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$