Coordination under Loss Contracts*

Steffen Ahrens[1], Lea Bitter[2], and Ciril Bosch-Rosa[3]

[1]Chair of Macroeconomics, Freie Universität Berlin
[2]European Central Bank
[2]Chair of Macroeconomics, Technische Universität Berlin

November 10, 2022

Abstract

In this paper we study the effects that loss contracts—prepayments that can be clawed back later—have on group coordination when there is strategic uncertainty. To do so, we investigate the choices made by experimental subjects in a minimum effort game. In control sessions, incentives are formulated as a classic gain contract, while in treatment sessions, incentives are framed as an isomorphic loss contract. Contrary to most results in the loss contract literature, in our setup loss contracts backfire by reducing the minimum effort of groups and worsening the coordination between group members. Such results suggest that the success off loss contracts is context dependent and offer an explanation as to why loss contracts are not implemented more often in the wild.

Keywords strategic uncertainty, loss aversion, coordination, contract design, framing, experiment

JEL Classification C91 · D84 · G11 · G41

*Corresponding author: lea.bitter@ecb.europa.eu. We would like to thank Alex Rees-Jones, Dirk Engelmann, Frank Heinemann, and Muhammed Bulutay for their insightful comments. We also would like to thank Ailin Leng for sharing her z-Tree code with us and two anonymous referees and an advisory editor whose insightful comments have greatly improved the paper. All authors acknowledge the financial support from the Deutsche Forschungsgemeinschaft (DFG) through the CRC TRR 190 "Rationality and Competition". The views expressed herein do not necessarily reflect those of the European Central Bank.

1 Introduction

Coordination is an important aspect of many organizational settings. In those contexts, a single member's decision can impact the firm's business process. This is apparent in assembly lines or just-in-time inventory systems but also applies to other less obvious settings, such as the timely revision of a scientific paper or the initial delay of the EU to meet the Covid-19 vaccination schedule.[1]

In such settings, having all members of a team coordinate on a high effort can be complicated, as individuals face a risk-reward tradeoff: while exerting more effort might result in higher productivity, such effort might be wasted if someone else along the chain (the "weakest link") is not keeping up to speed. One way to mitigate this coordination failure is to increase the monetary benefits from coordination (Brandts and Cooper, 2006); yet, this measure is expensive. A recent cost-effective suggestion to increase effort is the use of so-called loss contracts (e.g., Hossain and List, 2012), where individuals receive a prepayment which is clawed back if they do not meet certain productivity targets.

The intuition for the application of loss contracts rests on the presence of loss aversion: since losses loom larger than gains, loss-averse individuals will work harder to avoid losing a dollar than to gain an additional dollar (e.g., Fryer Jr et al., 2012; Hossain and List, 2012; Imas et al., 2017). However, this intuition relies on a series of restrictive assumptions. For example, most of the previous literature on loss contracts assumes that higher levels of effort guarantee a higher payoff or unambiguously reduce the probability of a clawback. This assumption might hold for some individual decision-making situations but is not realistic in many setups that require groups to coordinate. In a situation with strategic uncertainty (e.g., when productivity relies on the weakest link), the interaction between loss aversion and risk aversion might backfire and induce individuals to exert lower effort.

As shown in Pierce et al. (2022), if the effort strategies of individuals involve a risk-reward tradeoff, then loss contracts might have perverse effects and drive individuals to reduce their efforts. The reason for this is that when losses loom larger than gains, the uncertainty of outcome-related losses might also loom larger than the potential gains, pushing individuals toward low effort "loss-minimizing" strategies. In a similar vein,

[1] The slow delivery of vaccines by AstraZeneca resulted in a substantial delay in the vaccination targets that the EU had set for 2021 (European Commission, 2021), turning the company into the weakest link of the vaccination chain.

Armantier and Boly (2015) show theoretically and experimentally that the introduction of loss contracts might have ambiguous effects on effort provision, a prediction in line with the experimental results of De Quidt (2018) and De Quidt et al. (2017).

Against this background, we study how loss contracts affect coordinated efforts within groups when strategic uncertainty is present. To do so, we design a between-subjects experiment in which subjects play multiple rounds of the "minimum effort game" (Van Huyck et al., 1990), also known as the "weakest link" game (e.g., Knez and Camerer, 1994; Riedl et al., 2016). As in a production chain, in this setup each subject's payoff depends on her own effort and the lowest effort of all group members. To study the effects of loss contracts, we set up two treatments: a control group with a "classic" gain contract and a treatment group with an isomorphic payoff function framed as a loss contract. Because the only difference between both treatments is how the payoffs are presented, any change in subjects' behavior can be attributed to the *framing* of the payoff function (Tversky and Kahneman, 1981).

We find that sessions with loss contracts result in lower group productivity compared to sessions with gain contracts. We also find that loss contracts worsen coordination among group members, which is reflected in a higher variance of the groups' effort choices. The combination of higher variance and lower effort levels under loss contracts results in substantially lower payoffs in these sessions. Our results show that risk aversion can explain part of this lower productivity. Additionally, those groups with a larger proportion of females exert higher minimum effort levels, coordinate better, and obtain higher payoffs.

Our study contributes to the literature on the effects of negatively framed incentives. While this literature is rich in the effect of loss contracts on individual worker effort (e.g., De Quidt et al., 2017; Imas et al., 2017; DellaVigna and Pope, 2018; Pierce et al., 2022), to the best of our knowledge, only a few papers have studied the effects that such contracts have on group coordination. In a field experiment, Hossain and List (2012) study the effects of loss contracts on group productivity and show that loss contracts increase group productivity. However, in their experiment, there is no strategic uncertainty.[2] Hong et al. (2015) study the effect of loss contracts with competing groups at the same factory as

[2] As explained on page five of the article, a subset of groups worked around belt lines with a speed that the group could alter or worked around guide rails with a fixed speed. It is unclear if strategic uncertainty existed in the two remaining groups (G3 and G4), but the results for these groups are mixed.

Hossain and List (2012) and find that groups incentivized with loss – instead of gain – contracts are more likely to win the contest.

In the laboratory, Iturbe-Ormaetxe et al. (2011) study the effects of framing using different thresholds in a public goods game. Cachon and Camerer (1996) study *loss avoidance* and forward induction (implicit communication about subjects' expectations) as an equilibrium selection refinement in median and minimum effort games. Hamman et al. (2007) study the effect of imposing a penalty or bonus conditional on specified outcomes, while Brandts and Cooper (2006) look at the effect that reducing previous bonus payments has on coordination. However, all of these laboratory experiments have different focuses and, with several behavioral aspects at play, they cannot isolate the effects of negatively framed incentives on group effort and coordination.

Overall, our contribution is to study the effects of loss contracts on group coordination and productivity when strategic uncertainty is present. To our knowledge, we are the first laboratory experiment to show that the effectiveness of loss contracts is context-dependent and may backfire in environments with strategic uncertainty. Given that such strategic uncertainty is common in organizational settings, our results might help explain the scarcity of loss contracts in the field and serve as a word of caution toward using loss contracts as a blanket solution to increase effort and productivity.

The paper is organized as follows. Section 2 presents our experimental design. Section 3 presents the experiment's results. Section 4 discusses the results. Finally, Section 5 concludes.

2 Experimental Design

2.1 Experimental Setup

We design a between-subjects experiment with two treatments: a gain contract and a loss contract. In both cases, subjects are divided into groups of six and simultaneously decide how much effort to exert in each given period. Subjects' payoffs decrease in their own effort and increase in the minimum effort chosen across all subjects in the group. Formally,

$$\Pi(e_i, e_{min}) = ae_{min} - be_i + C, \tag{1}$$

where e_i is the effort of subject i, e_{min} is the minimum effort across all subjects n in the group, a and b are parameters such that $a - b > 0$, and C is a constant to avoid negative payoffs. The parametrization follows Van Huyck et al. (1990), with $a = 20$ points, $b = 10$ points, and $C = 60$ points. At the end of the experiment, the exchange rate is of €1 for every 70 points, which is comparable to the rate in Engelmann and Normann (2010) and Leng et al. (2018). The game is played for ten consecutive periods, maintaining the same group composition. After each period, subjects receive feedback about the group's minimum effort and the resulting payoff. After the ten periods, we elicit several personality traits from our subjects. First, we measure subjects' cognitive ability using the CRT (Frederick, 2005), CRT2 (Thomson and Oppenheimer, 2016), and eCRT (Toplak et al., 2014) questions. Then we elicit their risk aversion, ambiguity aversion, and loss aversion through modifying the multiple price lists used in Rubin et al. (2018). Finally, subjects answer the short version of the Big Five personality traits suggested by Rammstedt and John (2007) and state their gender.

Our treatment is implemented through the framing of the payoffs. In the gain contract treatment, subjects are presented with the payoffs resulting from Equation (1), as depicted in the left panel of Table 1. The vertical axis of the payoff table denotes the effort choice of an individual subject i. The horizontal axis denotes the smallest effort level among subject i's group members. In the loss contract treatment, subjects are presented with isomorphic contracts that work through framing (Tversky and Kahneman, 1981). Subjects are endowed with 140 points before each period and are presented with the right panel of Table 1. Importantly, this second table does not represent a subject's final payoffs but the outcomes of all subjects' joint actions. To calculate the payoffs for each set of actions, subjects need to subtract the resulting outcome from their per-period endowment of 140 points. This is made clear in the instructions.

The experimental setup is designed to increase the strategic uncertainty of subjects. For example, in both panels of Table 1, the values presented are either all positive or all negative. By avoiding mixed gain-loss payoff tables as in Cachon and Camerer (1996) and Armantier and Boly (2015) or zeros as in Hamman et al. (2007), we prevent focal points.[3]

[3] Cachon and Camerer (1996) show that in minimum effort games with negative and non-negative outcomes, the latter act as focal points with subjects avoiding losses by ignoring all strategies that result in negative outcomes. By showing only positive or negative values, we can exclude such *loss avoidance* as an equilibrium selection principle.

	Gain Contract							Loss Contract						
	Minimum Choice within Group							*Minimum Choice within Group*						
	7	**6**	**5**	**4**	**3**	**2**	**1**	**7**	**6**	**5**	**4**	**3**	**2**	**1**
7	130	110	90	70	50	30	10	–10	–30	–50	–70	–90	–110	–130
6		120	100	80	60	40	20		–20	–40	–60	–80	–100	–120
5			110	90	70	50	30			–30	–50	–70	–90	–110
4				100	80	60	40				–40	–60	–80	–100
3					90	70	50					–50	–70	–90
2						80	60						–60	–80
1							70							–70

Table 1: Payoff tables presented to subjects. In both cases, rows represent own effort and columns represent the group's minimum effort. The left panel shows the gain contract treatment where subjects see their final payoff in points, and the right panel shows the loss contract table. Here, points are subtracted from subjects' initial endowment (140) and are not final payoffs.

Additionally, we do not allow peer-monitoring as is in Brandts and Cooper (2006), and unlike Cachon and Camerer (1996), we require *every* group member to increase their effort to reach a higher group payoff.

Given the high strategic uncertainty in our setup, we hypothesize that loss contracts will result in lower group minimum effort compared to gain contracts. The intuition for this hypothesis follows from Pierce et al. (2022), who show that in an environment where the results from increasing effort are non-deterministic, loss contracts can backfire. This is because, in such setups, workers face a risk-reward tradeoff where loss exposure can be offset through low effort levels. Since in our experiment higher effort levels increase the chance of a higher payoff but *also* the exposure to losses, we expect subjects under loss contracts to exert a lower level of effort to reduce their loss exposure. As Pierce et al. (2022) put it, "[introducing] loss framing incentivizes the [worker] to avoid losses."[4]

2.2 Experimental Procedure

The experiment was run at two different points in time (Spring 2019 and Spring 2021). The first series of sessions was run at the TU-WZB Experimental Economics Laboratory of the Technische Universität Berlin and it consisted of four sessions with gain contracts and four sessions with loss contracts. The second series of sessions was run online and, again,

[4]Note that such a "loss-minimizing" strategy is closely related to the "loss-avoidance" strategies as observed in Cachon and Camerer (1996). Furthermore, since the adoption of such a "loss-minimizing" strategy by a single group member is sufficient to pull down group productivity, loss exposure is most salient in such weakest-link environments.

it consisted of four sessions with gain contracts and four sessions with loss contracts.[5] In each session, we randomly divided subjects into 3 groups of 6 subjects for a total of 24 independent groups per treatment. The total number of subjects is 288, with the number of subjects equally divided across online and onsite sessions. The share of females is 35% in onsite sessions and 50% in online sessions.[6] Importantly, the subjects for both types of sessions came from the TU-WZB Experimental Economics Laboratory subject pool, were recruited using ORSEE (Greiner, 2015), and used the same software based on z-Tree (Fischbacher, 2007). For the online sessions, the software was distributed using z-Tree Unleashed (Duch et al., 2020). Sessions run onsite lasted less than one hour, with average earnings of €12.74 while those run online lasted around 90 minutes and the average payoff was of €20.07.[7]

3 Results

3.1 Minimum Effort

Figure 1 shows for each treatment the average minimum effort of all groups for each period (black dots), along with their 95% confidence intervals (vertical lines). The treatment effect is clear: under loss contracts the average minimum effort of groups is consistently lower than under gain contracts. A Mann-Whitney U test across treatments comparing the average effort of each group across all ten periods confirms such differences (p-value = 0.047).

In Figure 2 we present the minimum choice of each group in each period (thin gray lines) for both the online (dashed) and onsite sessions (solid) as well as the mean minimum effort across all groups in each period (thick red line). Underlying such lines, a heat map displays darker colors if more groups have a minimum effort at that value. The figure suggests that the difference across treatments comes from a faster convergence to the

[5]The reason for running sessions online was the onset of the COVID-19 pandemic, which forced most laboratories to temporarily close down.

[6]Given the strong influence of gender on the results from the onsite sessions, in the online sessions we opted for an equal gender share.

[7]The difference in time needed for online and onsite sessions is due to the ID-checking protocol used in the TU-WZB Experimental Economics Laboratory for online sessions. Because identification of the subjects (ID check via webcam) required privacy, each subject was admitted sequentially into a private zoom room. To compensate for the extra time, all subjects are paid an extra "participation fee" of €7 on top of the regular €5 "show-up" fee, which explains the difference in payoffs.

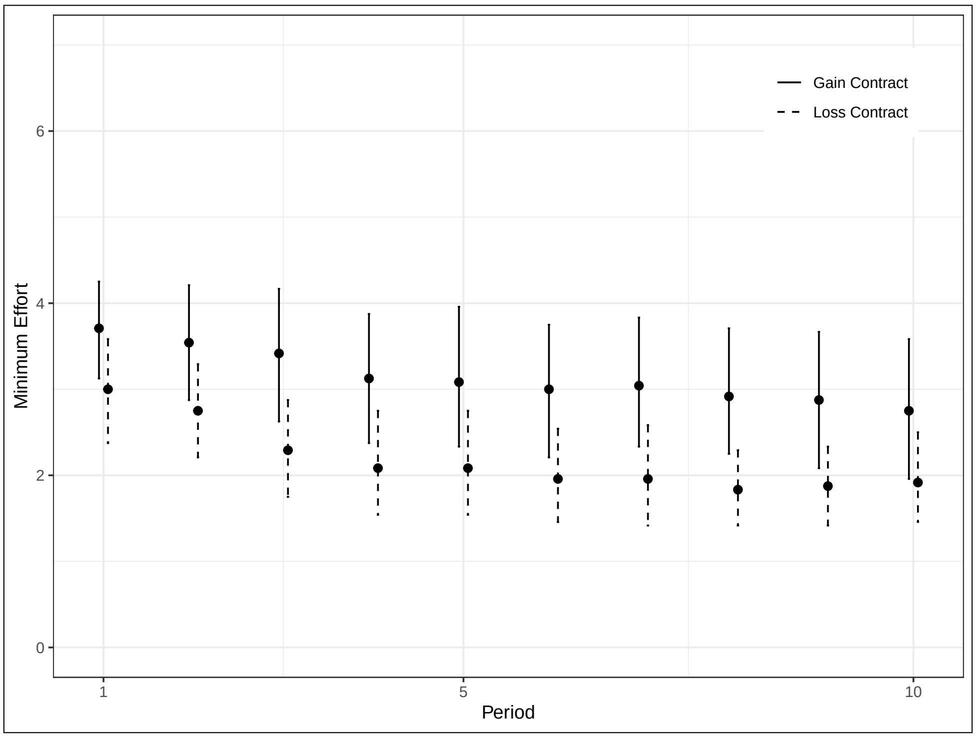


Figure 1: Summary of all minimum efforts per period across treatments. The figure presents the average minimum effort (black dot) and its bootstrapped 95% confidence interval (dashed line for loss contracts, solid line for gain contracts) per treatment.

lowest minimum effort under loss contracts than under gain contracts. A Mann-Whitney U test of the average effort made by each group for the first five and the last five periods confirms the asymmetric dynamics across treatments (p-value $= 0.067$ and p-value $= 0.014$, respectively).[8]

To study the data in a more disaggregated way, in Table 2 we present a random effects model with the per-group minimum effort (*periodmineffort*) for each period as the dependent variable. In column (1) we use two dummies. The first one takes a value of 1 if the session had loss contracts (*loss_contract*) and zero otherwise, while the second (*online*) takes a value of 1 if the session was run online and zero otherwise. In column (2) we control for the ratio of females per group (*female_ratio*) and the average value of different personality traits (e.g., *avg_risk_av* is the mean value of the risk aversion across all subjects of a group). Finally, in column (3) we include the average score of all subjects in the group for the Big Five personality traits. The controls used in Table

[8]In Table 16 of Appendix F we present a breakdown of the minimum effort levels of each group across periods. The table is complemented by a boxplot representing all the minimum effort choices across groups and treatments for the first and last five periods of the session in Figure 9.

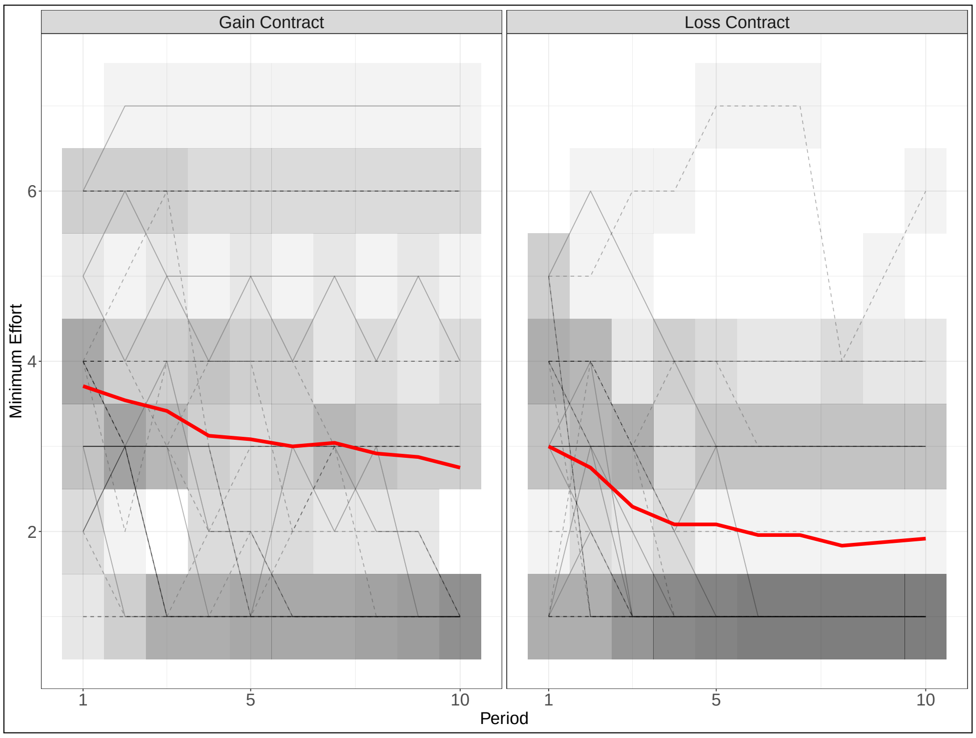


Figure 2: Summary of all minimum efforts per period across treatments. For both treatments, the thin gray lines represent the minimum effort played in each group, and the thick red line is the mean of this minimum effort. The dashed lines are for the online sessions and the solid lines for onsite sessions. Underlying we use a heat map which has a darker color when there are more groups with a minimum effort at that value.

	(1) *periodmineffort*	(2) *periodmineffort*	(3) *periodmineffort*
loss_contract	-0.971**	-1.119**	-1.130***
	(0.441)	(0.433)	(0.412)
online	0.088	-0.977*	-1.070**
	(0.441)	(0.517)	(0.501)
female_ratio		3.480***	4.409***
		(1.116)	(1.098)
avg_risk_av		-0.521**	-0.590***
		(0.220)	(0.213)
avg_loss_av		-0.056	-0.069
		(0.163)	(0.164)
avg_ambiguity_av		-0.012	-0.063
		(0.177)	(0.203)
avg_crt		0.305**	0.234
		(0.146)	(0.153)
constant	3.102***	4.719**	8.366**
	(0.450)	(2.292)	(4.140)
Big Five Averages	No	No	Yes
N	480	480	480

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2: Analysis of the minimum effort of groups across periods using a random effects GLS. In columns (1) to (3) we regress the minimum effort of each group for each period (*periodmineffort*) on the dummies for the loss contract (*loss_contract*) and online sessions (*online*), along with the average personality measures of the group (column (2)) and the average values of the Big Five (column (3)). All standard errors are clustered at the group level. For session level clustered errors, see Table 6 in Appendix B.

2 do not present any mulitcollinearity problems and the distribution of characteristics is balanced across treatments (see Appendix A for a further discussion on the different control measures). In all models, the errors are clustered at the group level. To address any concerns about session level effects, in Appendix B we reproduce all the tables in the main text with errors clustered at the session level. The results are qualitatively identical in all cases.

The results in Table 2 show that loss contracts have a negative effect on the minimum effort of groups, reducing the minimum effort by approximately 1 effort level on average. This effect is significant at the 1% level (p-value = 0.010 for column (2) and p-value = 0.006 for column (3)) and consistent with the treatment effects observed in Figure 1. Therefore, from all of the results presented in this section, we conclude that loss contracts result in a lower minimum effort of groups.

Result 1: A loss contract results in a lower minimum effort of groups.

Table 2 also shows that the gender composition of groups has a strong effect on their minimum effort. The more females are in a group, the higher is the minimum effort exerted. Furthermore, groups with higher average risk aversion show significantly lower groups minimum effort levels. By contrast, ambiguity aversion and loss aversion have no significant effects on the groups' minimum effort levels. Additionally, once we control for all other group characteristics, online sessions result in lower group minimum effort. The effect of online sessions might be attributed to a larger degree of "mistrust" and anonymity among subjects in comparison to onsite sessions (e.g., Chen and Li, 2009).[9] To understand better the differences between online and onsite sessions, in Appendix C we analyze our results by splitting the sample into onsite and online data. The results show that the treatment effects and session dynamics are similar across both types of sessions (see Figure 6) and that the interaction between treatment and online dummy is insignificant (Table 9). However, once we split the data, the effect of loss contracts is significant at the 5% level in onsite sessions, but not for online. We attribute this disparity in statistical significance to the decrease in power of our split samples and to

[9] Section 4 has a more detailed discussion on the effect of online experiments on mistrust and anonymity.

the differences in the gender ratios of online and onsite sessions.[10]

Finally, since the minimum effort game is extremely sensible to a single subject's decision, one could conjecture that the subject with the most extreme personality traits might drive the results. To test this, in Table 17 of Appendix F we replicate Table 2 using the most extreme personality values within each group as independent variables. The results are similar across both tables. Additionally, if we look at the *individual* effort levels of subjects and their *individual* personality measures, only the treatment dummy is statistically significant (see Table 11 in Appendix D).

3.2 Coordination

In this section, we analyze how loss contracts affect subjects' coordination within groups. Within-group coordination is important as a larger within-group dispersion of efforts result in more "wasted efforts" and, therefore, in lower efficiency.

To study such effects, in columns (1) to (3) of Table 3, we regress the variance of all effort choices, across all periods, for each group (*var*) on a dummy for loss contracts (*loss_contract*), a dummy for online sessions (*online*), the ratio of females in each group (*female_ratio*), and the average value of the different personality traits. In columns (4) to (6), we use a random effects model in which the dependent variable is the variance within each group for each period (*var_t*) using the same controls. In both cases, loss contracts result in higher within-group effort variance.

Result 2: A loss contract results in less coordination (larger variance) of effort choices within groups.

As in Table 2, group composition strongly influences how groups behave. The groups with a higher share of females and the groups with higher average CRT have less dispersed effort levels. In addition, it seems that groups with larger average loss aversion coordinate less. In Appendix E we investigate how personality traits affect the way subjects react to the observed minimum effort of their group. The results show that under loss contracts,

[10] While in online sessions gender was balanced, in onsite sessions it was not (see Table 5, Appendix A). Given the importance of gender for our results, the extremely gender-balanced groups of online sessions (see Figure 7, Appendix C) likely impacted the treatment effect and the weight of other personality measures in our results. For a longer discussion on the differences across online and onsite sessions, see Appendix C.

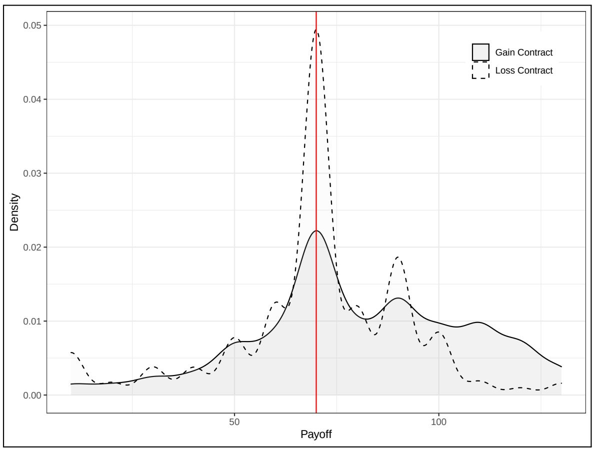
	OLS			Random Effects GLS		
	(1)	(2)	(3)	(4)	(5)	(6)
	var	*var*	*var*	*var_t*	*var_t*	*var_t*
loss_contract	1.063**	1.183**	1.248**	0.445	0.632**	0.590***
	(0.481)	(0.485)	(0.470)	(0.289)	(0.254)	(0.203)
online	-0.450	0.199	0.423	-0.121	0.149	0.152
	(0.481)	(0.615)	(0.680)	(0.289)	(0.312)	(0.311)
female_ratio		-2.999**	-4.702***		-1.274*	-1.861**
		(1.415)	(1.558)		(0.751)	(0.736)
avg_risk_av		0.027	0.198		-0.014	0.008
		(0.292)	(0.304)		(0.169)	(0.166)
avg_loss_av		0.162	0.109		0.173*	0.220***
		(0.168)	(0.188)		(0.094)	(0.079)
avg_ambiguity_av		-0.026	0.024		-0.116	-0.028
		(0.246)	(0.253)		(0.121)	(0.122)
avg_crt		-0.234	-0.101		-0.262***	-0.185**
		(0.167)	(0.172)		(0.081)	(0.087)
constant	2.518***	3.649	-1.798	1.575***	3.318**	-0.983
	(0.417)	(2.593)	(4.942)	(0.258)	(1.326)	(2.002)
Big Five Averages	No	No	Yes	No	No	Yes
N	48	48	48	480	480	480

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

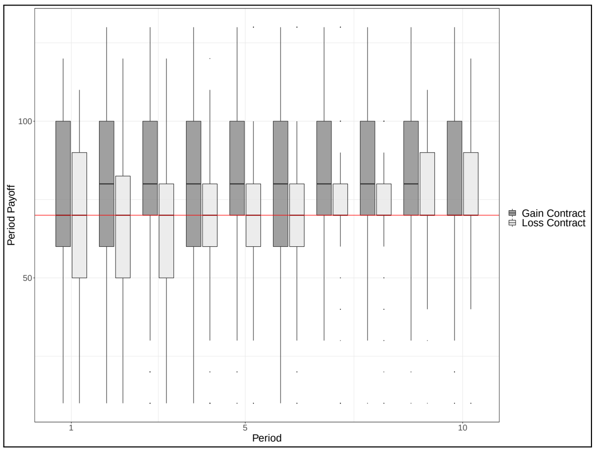
Table 3: Analysis of the aggregate variance in effort choices. In columns (1) to (3) we use OLS to regress the variance of all effort choices, across all periods, for each group (*var*) on the dummies for the loss contract (*loss_contract*) and online sessions (*online*), along with the average personality measures of the group (column (2)) and the average values of the Big Five (column (3)). In columns (4) to (6) we use a random effects model to regress the variance within each group for *each period* (*var_t*) on the dummies for the loss contract (*loss_contract*) and online sessions (*online*), along with the average personality measures of the group (column (5)) and the average values of the Big Five (column (6)). All standard errors are clustered at the group level. For session level clustered errors, see Table 7 in Appendix B.

individual risk aversion explains the behavior of those subjects who undercut the group's minimum effort.

3.3 Payoffs



(a) Payoffs across Treatments



(b) Payoffs per Period

Figure 3: Payoffs. The left panel presents the density plots of the per-period payoff of all subjects across treatments. The right panel disaggregates the data by period and treatment.

In this section, we analyze how loss contracts affect subjects' payoffs. To do so, in the left panel of Figure 3 we plot the density of the per-period payoff of all subjects across both treatments. The figure shows that the payoff distribution for gain contracts is skewed towards the higher payoffs (right of the graph), while the payoffs for loss contracts concentrate around the risk-dominant equilibrium payoff (i.e., 70 experimental units), which is marked with a red vertical line.[11] Moreover, in the right panel of Figure 3 we plot subjects' payoffs for each period. The panel shows that the median payoff is 70 experimental units *in every single period* under loss contracts, while it is greater than 70 in most periods under gain contracts. Another interesting feature of the right panel in Figure 3 is how the variance in payoffs seems to decrease faster under loss contracts as the experiment advances.

To quantify the effects of loss contracts on payoffs (*payoff*), in Table 4 we present a random effects models studying the effects of contract framing and group composition on

[11]There are other combinations by which a subject might get 70 experimental units. However, exerting the minimum effort is the only way a subject can guarantee these 70 experimental units.

	(1)	(2)	(3)
	payoff	*payoff*	*payoff*
loss_contract	-10.47**	-12.86**	-12.98***
	(5.243)	(5.125)	(4.796)
online	0.826	-10.51*	-10.96*
	(5.243)	(6.097)	(5.719)
female_ratio		37.81***	50.10***
		(13.39)	(13.83)
avg_risk_av		-4.421	-5.355*
		(2.841)	(2.927)
avg_loss_av		-1.574	-1.633
		(1.841)	(1.833)
avg_ambiguity_av		0.179	-0.357
		(2.103)	(2.471)
avg_crt		4.164**	3.100*
		(1.736)	(1.838)
constant	79.98***	88.05***	144.4***
	(5.438)	(26.48)	(48.30)
Big Five Averages	No	No	Yes
N	2880	2880	2880

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Analysis of the individual payoff of subjects across periods using a random effects GLS. In columns (1) to (3) we regress the individual payoffs for each subject (*payoff*) on the dummies for the loss contract (*loss_contract*) and online sessions (*online*), along with the average personality measures of the group (column (2)) and the average values of the Big Five (column (3)). All standard errors are clustered at the group level. For session level clustered errors, see Table 8 in Appendix B.

payoffs. The models confirm our above result; a loss contract is detrimental to subjects' payoffs.

Result 3: A loss contract results in lower payoffs.

Consistent with our previous results, we find significant gender effects for payoffs, so that groups with a higher share of females see significantly higher individual payoffs for all group members. On the other hand, neither online sessions nor the shares of risk aversion, ambiguity aversion, or loss aversion have explanatory power over payoffs.

4 Discussion

Our results show that contract framing has an effect. By manipulating how we present the payoff tables to subjects, sessions with loss contracts had lower minimum effort levels, less coordination within groups, and lower payoffs. This is in contrast with most of the previous literature where loss contracts had either a positive effect on individual effort (e.g., Hannan et al., 2005; Imas et al., 2017) or helped groups coordinate better (Hossain and List, 2012). However, unlike these previous cases, in our setup subjects face loss contracts under strategic uncertainty.

As shown in Pierce et al. (2022), implementing loss contracts in a stochastic context can backfire. This is because stochastic loss contracts present a risk-reward tradeoff that pushes loss-averse individuals to avoid any exposure to losses by taking a low-risk low-productivity strategy.[12] In our setup, such conservative strategies push subjects towards the risk-dominant equilibrium, which results in the race to the bottom that we observe in the data.

Using the individual level data collected from our subjects, we try to pin down the mechanism that drives loss contracts to backfire. Our results show that a higher average group risk aversion generally lowers the minimum effort of a group. Furthermore, while risk aversion does not affect individual effort choices across the board, it does explain the behavior of those subjects responsible for coordination failure under loss contracts, i.e., those subjects that undercut their group's minimum effort. Additionally, the higher

[12] Note that Cachon and Camerer (1996) also identify loss avoidance as an influential behavior selection principle.

the number of females, the higher the minimum effort and coordination levels of a group. Because of the strong and unexpected effects of gender, in Appendix G we study in depth the interaction between loss contracts and gender. The results show that the gender differences emerge mostly from the way in which females confront loss contracts, especially in the second half of the experiment.

One reason we cannot fully identify a behavioral pattern in the data is that subjects are making decisions in a *group*. This group interaction can introduce other complementary explanations to the low effort we observe, such as social preferences (Charness and Rabin, 2002) or the willingness to punish low effort contributors (Fehr et al., 2002). After all, it is likely that the negative externality of other subjects' actions (i.e., their low effort) is more salient when framed as a loss.[13] Such salience would explain why we see a much faster drop in effort levels under loss contracts than under gain contracts.

The influence of groups is also reflected in the negative impact of online sessions on the minimum effort of groups (see Table 2). This is consistent with the results from Chen and Li (2009) who show that low group identity leads to a deterioration in coordination. Based on social distance theory (Akerlof, 1997), the higher anonymity of online sessions translates into lower group identity. This is also supported by Charness et al. (2007) who find that reducing social distance leads to higher trust and reciprocity. Finally, the subjects' lack of familiarity with loss contracts might also be contributing to our results. Since loss contracts are not common, subjects might be unsure about how other subjects react to them, adding an extra layer of complexity that increases the strategic uncertainty of the setup. In a similar way that an increase in the size of groups increases the strategic uncertainty in experiments (Weber, 2006; Heinemann et al., 2009), the lack of subjects' familiarity with loss contracts might push them towards playing safer options.

5 Conclusion

From assembly line settings to the management of a global pandemic, the final outcome of many collaborations might depend on the weakest link. Yet, coordination on high effort is complicated (Devetag and Ortmann, 2007). It has been suggested that one way of

[13]If subjects are loss averse, low-effort behavior by other subjects will result in a bigger disutility under loss contracts than under gain contracts.

increasing the effort in the workplace is to present incentives as loss contracts (Hossain and List, 2012).

The literature studying loss contracts is large and points toward such contracts weakly increasing worker productivity in the field (e.g., Fryer Jr et al., 2012; Hossain and List, 2012; Hong et al., 2015) and in the laboratory (e.g., Hannan et al., 2005; Armantier and Boly, 2015; De Quidt, 2018). However, in all these cases, workers know that higher effort will result in a higher probability of keeping the salary (or bonus). Such a straightforward relationship between effort and final payoffs might not resemble many organizational settings which exhibit strategic uncertainty

In this paper, we create an environment of coordination under strategic uncertainty by implementing a minimum effort game (Van Huyck et al., 1990) in the laboratory. Laboratory experiments have proven to be a useful way to test the effects of different work environments and contracts schemes (Charness and Kuhn, 2011; Charness and Fehr, 2015; Herbst and Mas, 2015). They allow for full control of the environment and policies can be 'pre-tested' at a low cost (Plott, 1987), potentially avoiding multi-million losses (as, e.g., shown by Pierce et al. (2022))

Our results show that, unlike most of the previous literature, loss contracts reduce the effort of subjects. Such a drop in effort has strong welfare effects, leading to higher within-group effort variance and lower payoffs. The insight gained from the divergence between our results and those of previous studies is that loss contracts cannot be judged on their own but need to be evaluated in the context of their environment. Some settings might favor the introduction of loss contracts but, in other cases, implementing claw back policies might lead to destructive results.

Given the mostly positive outcomes reported in the literature, one would expect to see loss contracts implemented often. However, these are rare. This might be because many jobs require interactions with colleagues, contract partners, or service providers under strategic uncertainty, an environment in which loss contracts are likely to backfire. Thus, the results of our experiment can help explain why loss contracts are not commonly observed in the field.

Another interesting result of our experiment are the gender effects we observe. Groups with a larger proportion of females have higher minimum effort levels, are more coordinated, and thus have higher payoffs. These results seem to be driven mostly by gender differences in the loss contract sessions (see Appendix G). To our knowledge, we are the first to report gender effects under loss contracts. Furthermore, the previous literature on games with strategic complements reports that gender has no effect on the degree of coordination among experimental subjects (e.g., Dufwenberg and Gneezy, 2005; Heinemann et al., 2009; Engelmann and Normann, 2010; Di Girolamo and Drouvelis, 2015). Therefore, we also contribute to this literature by presenting evidence that the previously reported gender neutrality in games with strategic complements may not generalize to loss contracts.

To conclude, we study the effects of loss contracts in an environment with strategic uncertainty. Contrary to much of the previous literature, we show that if we relax some common assumptions, loss contracts can be detrimental and result in lower effort and coordination. Such results appear to be driven by the reluctance of subjects to expose themselves to losses and can help explain why loss contracts are not seen more often in the field.

References

Akerlof, G. A. (1997): "Social Distance and Social Decisions," *Econometrica*, 65, 1005–1027. Cited on page 17.

Armantier, O. and A. Boly (2015): "Framing of incentives and effort provision," *International Economic Review*, 56, 917–938. Cited on pages 3, 5, and 18.

Borghans, L., J. J. Heckman, B. H. Golsteyn, and H. Meijers (2009): "Gender differences in risk aversion and ambiguity aversion," *Journal of the European Economic Association*, 7, 649–658. Cited on page 47.

Brañas-Garza, P., P. Kujal, and B. Lenkei (2019): "Cognitive reflection test: Whom, how, when," *Journal of Behavioral and Experimental Economics*, 82, 101455. Cited on page 47.

Brandts, J. and D. Cooper (2006): "A change would do you good.... An experimental study on how to overcome coordination failure in organizations," *American Economic Review*, 96, 669–693. Cited on pages 2, 4, and 6.

Cachon, G. P. and C. F. Camerer (1996): "Loss-avoidance and forward induction in experimental coordination games," *The Quarterly Journal of Economics*, 111, 165–194. Cited on pages 4, 5, 6, and 16.

Charness, G. and E. Fehr (2015): "From the lab to the real world," *Science*, 350, 512–513. Cited on page 18.

Charness, G., E. Haruvy, and D. Sonsino (2007): "Social distance and reciprocity: An Internet experiment," *Journal of Economic Behavior & Organization*, 63, 88–103. Cited on page 17.

Charness, G. and P. Kuhn (2011): "Chapter 3 - Lab Labor: What Can Labor Economists Learn from the Lab?" Elsevier, vol. 4 of *Handbook of Labor Economics*, 229–330. Cited on page 18.

Charness, G. and M. Rabin (2002): "Understanding social preferences with simple tests," *The Quarterly Journal of Economics*, 117, 817–869. Cited on page 17.

Chen, Y. and S. X. Li (2009): "Group identity and social preferences," *American Economic Review*, 99, 431–57. Cited on pages 11 and 17.

De Quidt, J. (2018): "Your loss is my gain: a recruitment experiment with framed incentives," *Journal of the European Economic Association*, 16, 522–559. Cited on pages 3 and 18.

De Quidt, J., F. Fallucchi, F. Kölle, D. Nosenzo, and S. Quercia (2017): "Bonus versus penalty: How robust are the effects of contract framing?" *Journal of the Economic Science Association*, 3, 174–182. Cited on page 3.

DellaVigna, S. and D. Pope (2018): "What motivates effort? Evidence and expert forecasts," *The Review of Economic Studies*, 85, 1029–1069. Cited on page 3.

Devetag, G. and A. Ortmann (2007): "When and why? A critical survey on coordination failure in the laboratory," *Experimental Economics*, 10, 331–344. Cited on page 17.

Di Girolamo, A. and M. Drouvelis (2015): "The role of gender composition and size of the group in a minimum effort game," *Economics Letters*, 137, 168 – 170. Cited on pages 19 and 48.

Duch, M. L., M. R. Grossmann, and T. Lauer (2020): "z-Tree unleashed: A novel client-integrating architecture for conducting z-Tree experiments over the Internet," *Journal of Behavioral and Experimental Finance*, 28, 100400. Cited on page 7.

Dufwenberg, M. and U. Gneezy (2005): "Gender & Coordination," in *Experimental Business Research*, Springer, vol. III, chap. 11, 253–262. Cited on pages 19 and 48.

Engelmann, D. and H.-T. Normann (2010): "Maximum effort in the minimum-effort game," *Experimental Economics*, 13, 249–259. Cited on pages 5, 19, and 48.

European Commission (2021): "Press statement by Commissioner Kyriakides on vaccine deliveries and on the vaccine export transparency scheme," *European Commission Press Release*. Cited on page 2.

Fehr, E., U. Fischbacher, and S. Gächter (2002): "Strong reciprocity, human cooperation, and the enforcement of social norms," *Human Nature*, 13, 1–25. Cited on page 17.

Fischbacher, U. (2007): "z-Tree: Zurich toolbox for ready-made economic experiments," *Experimental Economics*, 10, 171–178. Cited on page 7.

Frederick, S. (2005): "Cognitive Reflection and Decision Making," *Journal of Economic Perspectives*, 19, 25–42. Cited on page 5.

Fryer Jr, R. G., S. D. Levitt, J. List, and S. Sadoff (2012): "Enhancing the efficacy of teacher incentives through loss aversion: A field experiment," NBER Working Paper 18237, National Bureau of Economic Research. Cited on pages 2 and 18.

Greiner, B. (2015): "Subject pool recruitment procedures: organizing experiments with ORSEE," *Journal of the Economic Science Association*, 1, 114–125. Cited on page 7.

Hamman, J., S. Rick, and R. A. Weber (2007): "Solving coordination failure with "all-or-none' group-level incentives," *Experimental Economics*, 10, 285–303. Cited on pages 4 and 5.

Hannan, R. L., V. B. Hoffman, and D. V. Moser (2005): "Bonus versus penalty: does contract frame affect employee effort?" in *Experimental Business Research*, Springer, vol. II, 151–169. Cited on pages 16 and 18.

Heinemann, F., R. Nagel, and P. Ockenfels (2009): "Measuring Strategic Uncertainty in Coordination Games," *The Review of Economic Studies*, 76, 181–221. Cited on pages 17, 19, and 48.

Herbst, D. and A. Mas (2015): "Peer effects on worker output in the laboratory generalize to the field," *Science*, 350, 545–549. Cited on page 18.

Hong, F., T. Hossain, and J. A. List (2015): "Framing manipulations in contests: A natural field experiment," *Journal of Economic Behavior & Organization*, 118, 372 – 382. Cited on pages 3 and 18.

HOSSAIN, T. AND J. A. LIST (2012): "The behavioralist visits the factory: Increasing productivity using simple framing manipulations," *Management Science*, 58, 2151–2167. Cited on pages 2, 3, 4, 16, and 18.

IMAS, A., S. SADOFF, AND A. SAMEK (2017): "Do people anticipate loss aversion?" *Management Science*, 63, 1271–1284. Cited on pages 2, 3, and 16.

ITURBE-ORMAETXE, I., G. PONTI, J. TOMÁS, AND L. UBEDA (2011): "Framing effects in public goods: Prospect theory and experimental evidence," *Games and Economic Behavior*, 72, 439–447. Cited on page 4.

KNEZ, M. AND C. CAMERER (1994): "Creating Expectational Assets in the Laboratory: Coordination in 'Weakest-Link' Games," *Strategic Management Journal*, 15, 101–119. Cited on page 3.

LENG, A., L. FRIESEN, K. KALAYCI, AND P. MAN (2018): "A minimum effort coordination game experiment in continuous time," *Experimental Economics*, 21, 549–572. Cited on page 5.

PIERCE, L., A. REES-JONES, AND C. BLANK (2022): "The Negative Consequences of Loss-Framed Performance Incentives," NBER Working Paper 26619, National Bureau of Economic Research. Cited on pages 2, 3, 6, 16, 18, and 39.

PLOTT, C. R. (1987): "Dimensions of parallelism: Some policy applications of experimental methods," in *Laboratory Experimentation in Economics: Six Points of View*, New York, Cambridge University Press, chap. 7, 193–219. Cited on page 18.

RAMMSTEDT, B. AND O. P. JOHN (2007): "Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German," *Journal of Research in Personality*, 41, 203–212. Cited on page 5.

RIEDL, A., I. M. T. ROHDE, AND M. STROBEL (2016): "Efficient Coordination in Weakest-Link Games," *The Review of Economic Studies*, 83, 737–767. Cited on page 3.

RUBIN, J., A. SAMEK, AND R. M. SHEREMETA (2018): "Loss aversion and the quantity–quality tradeoff," *Experimental Economics*, 21, 292–315. Cited on page 5.

Thomson, K. S. and D. M. Oppenheimer (2016): "Investigating an alternate form of the cognitive reflection test," *Judgment and Decision Making*, 11, 99–113. Cited on page 5.

Toplak, M. E., R. F. West, and K. E. Stanovich (2014): "Assessing miserly information processing: An expansion of the Cognitive Reflection Test," *Thinking & Reasoning*, 20, 147–168. Cited on page 5.

Tversky, A. and D. Kahneman (1981): "The framing of decisions and the psychology of choice," *Science*, 211, 453–458. Cited on pages 3 and 5.

Van Huyck, J. B., R. C. Battalio, and R. O. Beil (1990): "Tacit coordination games, strategic uncertainty, and coordination failure," *The American Economic Review*, 80, 234–248. Cited on pages 3, 5, and 18.

Weber, R. A. (2006): "Managing growth to achieve efficient coordination in large groups," *American Economic Review*, 96, 114–126. Cited on page 17.

A Balanced Randomization

In Figure 4 we present the density plots for all the personality measures across treatments. The distributions look very similar, showing that the randomization of subjects into the different treatments resulted in a balanced distribution of characteristics. A series of Mann-Whitney U tests confirms that there are no statistical differences for personality across treatments. The only exception is the number of correctly answered CRT questions where those subjects in the loss contract sessions answered correctly a higher number of questions (Mann-Whitney U p-value = 0.017). It is also interesting to note that the correlation between the different personality measures is low (see Figure 5).

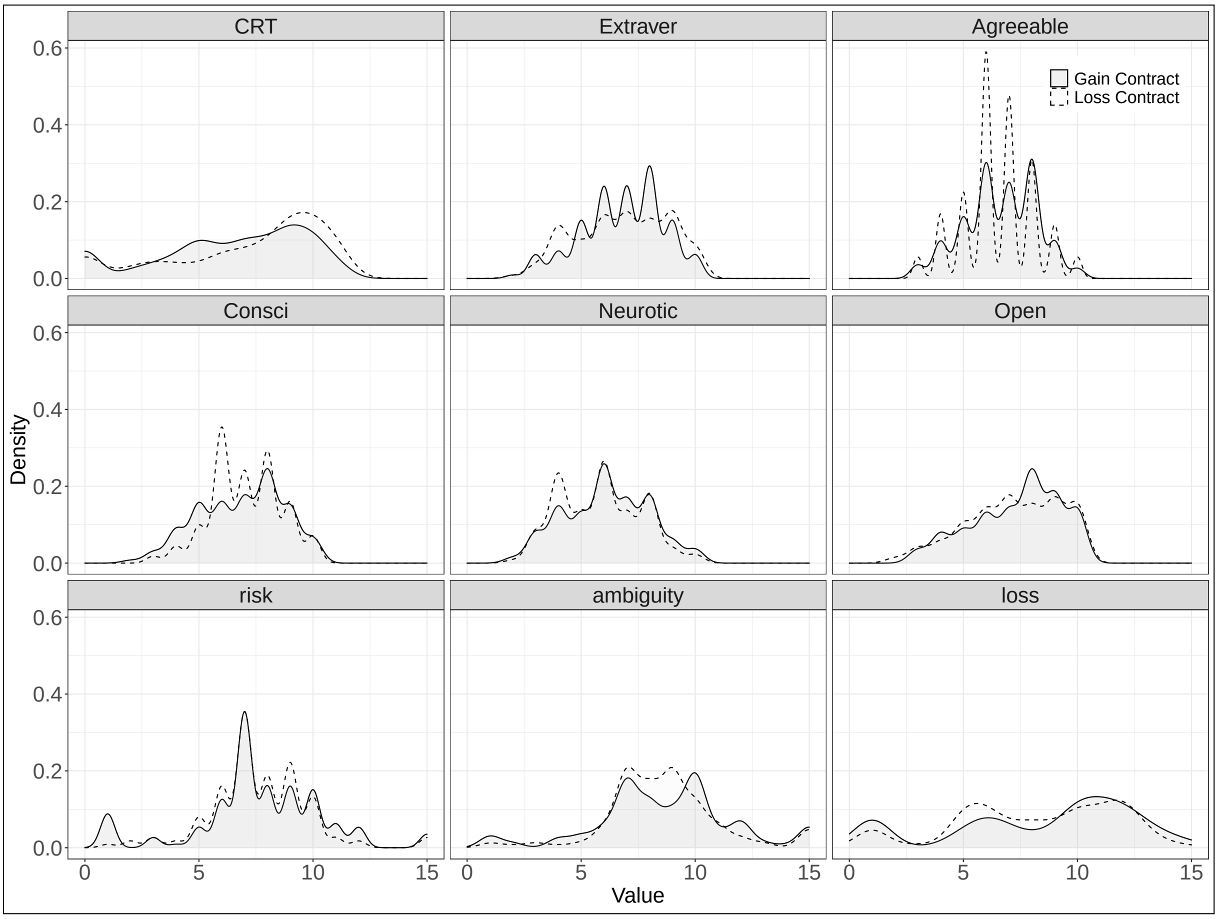


Figure 4: Density plots of each measure used as control for each treatment.

Finally, when looking at gender, the proportion of females across treatments is very similar. In control we have 63/144 females (∼43%), while in treatment we have 59/144 (∼41%). A detailed breakdown of the number of males and females across sessions is

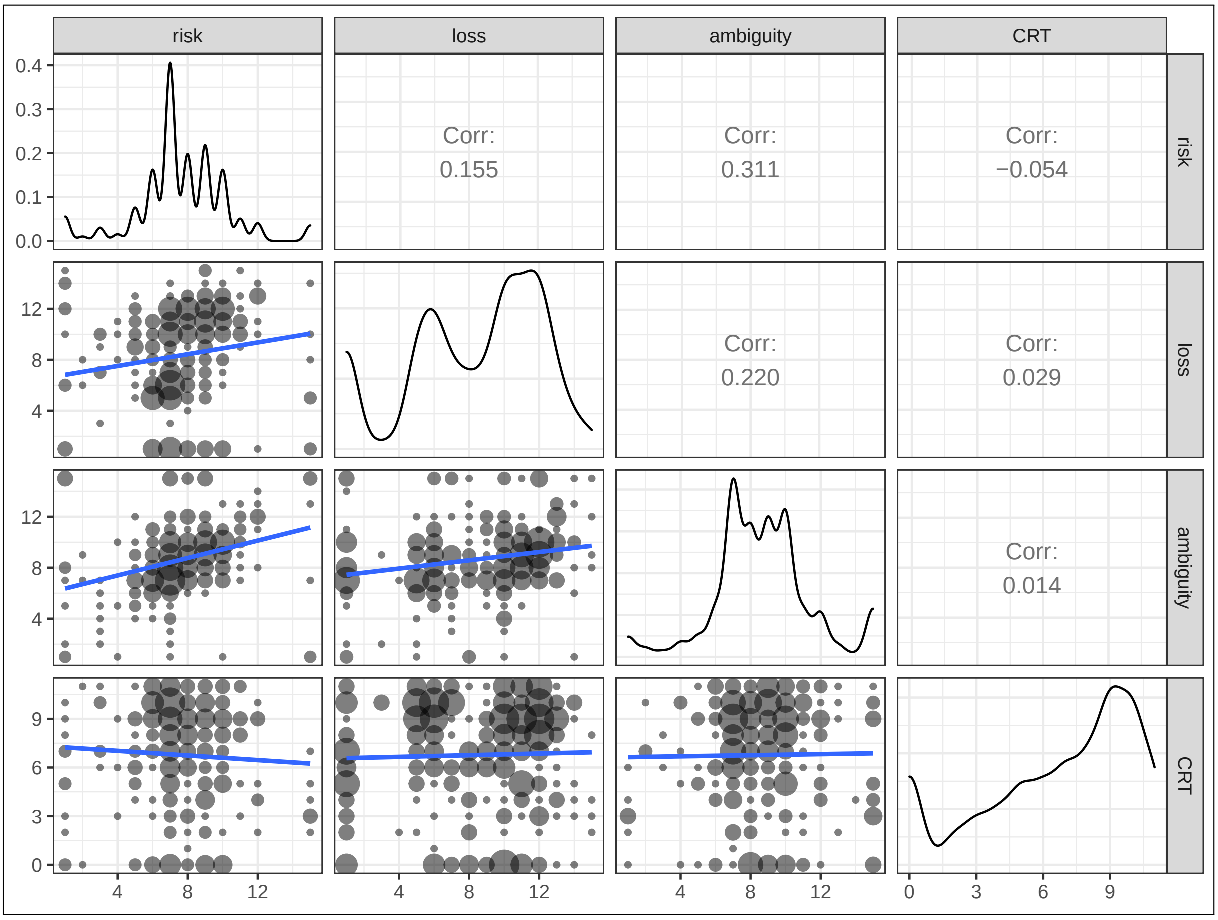


Figure 5: Correlations across different personality measures (risk aversion, loss aversion, ambiguity aversion, and CRT score). In the diagonal we plot the density estimates for each measure. The lower triangle shows a scatter plot between each pair of measures, while the upper triangle shows Pearson's correlation (ρ).

provided in Table 5. A test of proportions finds no statistical differences in the number of females across the two treatments (p-value = 0.634).

Gain Contract		
Session	Male	Female
Onsite		
1	10	8
2	12	6
3	11	7
4	12	6
Total	*45*	*27*
Online		
5	11	7
6	9	9
7	9	9
8	7	11
Total	*36*	*36*

Loss Contract		
Session	Male	Female
Onsite		
1	7	11
2	15	3
3	12	6
4	15	3
Total	*49*	*23*
Online		
5	11	7
6	9	9
7	8	10
8	8	10
Total	*36*	*36*

Table 5: Number of males and females per session. On the left we present the number of males and females per session in gain contract treatment. On the right we present the number of males and females per session in the loss contract treatment. In total 166 males and 122 females participated in our experiment.

B Wild-Block-Bootstrap Tables

In this section we replicate Table 2, Table 3, and Table 4 clustering the errors at the session level.

	(1)	(2)	(3)
	periodmineffort	*periodmineffort*	*periodmineffort*
loss_contract	-0.971**	-1.119**	-1.130***
	(0.467)	(0.493)	(0.429)
online	0.088	-0.977	-1.070
	(0.476)	(0.713)	(0.820)
female_ratio		3.480***	4.409***
		(1.207)	(1.248)
avg_risk_av		-0.521*	-0.590*
		(0.306)	(0.320)
avg_loss_av		-0.056	-0.069
		(0.206)	(0.216)
avg_ambiguity_av		-0.012	-0.063
		(0.212)	(0.213)
avg_crt		0.305*	0.234
		(0.185)	(0.184)
constant	3.102***	4.719*	8.366*
	(0.609)	(2.693)	(4.673)
N	480	480	480
Big Five Averages	No	No	Yes

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Analysis of the minimum effort of groups across periods using OLS. In columns (1) to (3) we regress the minimum effort of each group for each period (*periodmineffort*) on the dummies for the loss contract (*loss_contract*) and online sessions (*online*), along with the average personality measures of the group (column (2)) and the average values of the Big Five (column (3)). All wild block bootstrap standard errors are clustered at the session level. For group level clustered errors, see Table 2.

	OLS			Random Effects GLS		
	(1)	(2)	(3)	(4)	(5)	(6)
	var	*var*	*var*	*var_t*	*var_t*	*var_t*
loss_contract	1.063**	1.183***	1.248***	0.445	0.632**	0.590***
	(0.403)	(0.407)	(0.389)	(0.324)	(0.302)	(0.219)
online	-0.450	0.199	0.423	-0.121	0.149	0.152
	(0.312)	(0.459)	(0.518)	(0.336)	(0.366)	(0.432)
female_ratio		-2.999**	-4.702***		-1.274*	-1.861**
		(1.225)	(1.134)		(0.680)	(0.936)
avg_risk_av		0.027	0.198		-0.014	0.008
		(0.348)	(0.324)		(0.265)	(0.215)
avg_loss_av		0.162	0.109		0.173**	0.220***
		(0.176)	(0.172)		(0.082)	(0.063)
avg_ambiguity_av		-0.026	0.024		-0.116	-0.028
		(0.188)	(0.205)		(0.121)	(0.098)
avg_crt		-0.234	-0.101		-0.262***	-0.185**
		(0.163)	(0.171)		(0.096)	(0.090)
constant	2.518***	3.649**	-1.798	1.575***	3.318***	-0.983
	(0.388)	(1.606)	(3.152)	(0.404)	(1.141)	(1.781)
Big Five Averages	No	No	Yes	No	No	Yes
N	48	48	48	480	480	480

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Analysis of the aggregate variance in effort choices. In columns (1) to (3) we use OLS to regress the variance of all effort choices, across all periods, for each group (*var*) on the dummies for the loss contract (*loss_contract*) and online sessions (*online*), along with the average personality measures of the group (column (2)) and the average values of the Big Five (column (3)). In columns (4) to (6) we use a random effects model to regress the variance within each group for *each period* (*var_t*) on the dummies for the loss contract (*loss_contract*) and online sessions (*online*), along with the average personality measures of the group (column (5)) and the average values of the Big Five (column (6)). All wild block bootstrap standard errors are clustered at the session level. For group level clustered errors, see Table 3.

	(1)	(2)	(3)
	payoff	*payoff*	*payoff*
loss_contract	-10.47*	-12.86**	-12.98**
	(6.131)	(6.212)	(5.346)
online	0.826	-10.51	-10.96
	(6.266)	(9.259)	(10.449)
female_ratio		37.81**	50.10***
		(15.330)	(17.744)
avg_risk_av		-4.421	-5.355
		(4.290)	(4.465)
avg_loss_av		-1.574	-1.633
		(2.196)	(2.429)
avg_ambiguity_av		0.179	-0.357
		(2.706)	(2.698)
avg_crt		4.164*	3.100
		(2.410)	(2.401)
constant	79.98***	88.05***	144.4**
	(8.015)	(31.677)	(57.551)
N	2880	2880	2880
Big Five Averages	No	No	Yes

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Analysis of the individual payoff of subjects across periods using OLS. In columns (1) to (3) we regress the individual payoffs for each subject (*payoff*) on the dummies for the loss contract (*loss_contract*) and online sessions (*online*), along with the average personality measures of the group (column (2)) and the average values of the Big Five (column (3)). All wild block bootstrap standard errors are clustered at the session level. For group level clustered errors, see Table 4.

C Online and Onsite

In Figure 6 we reproduce Figure 1, but separating the data into the onsite and online sessions. The figure shows an almost identical pattern of behavior in both types of sessions, with the average effort in gain contracts always above that of loss contracts. Additionally, in both cases, the average minimum effort declines as the experiment progresses. The main difference across both figures is a steeper but delayed drop of minimum effort of onsite loss contract sessions compared to online sessions and slightly larger confidence intervals for loss contracts in online sessions.

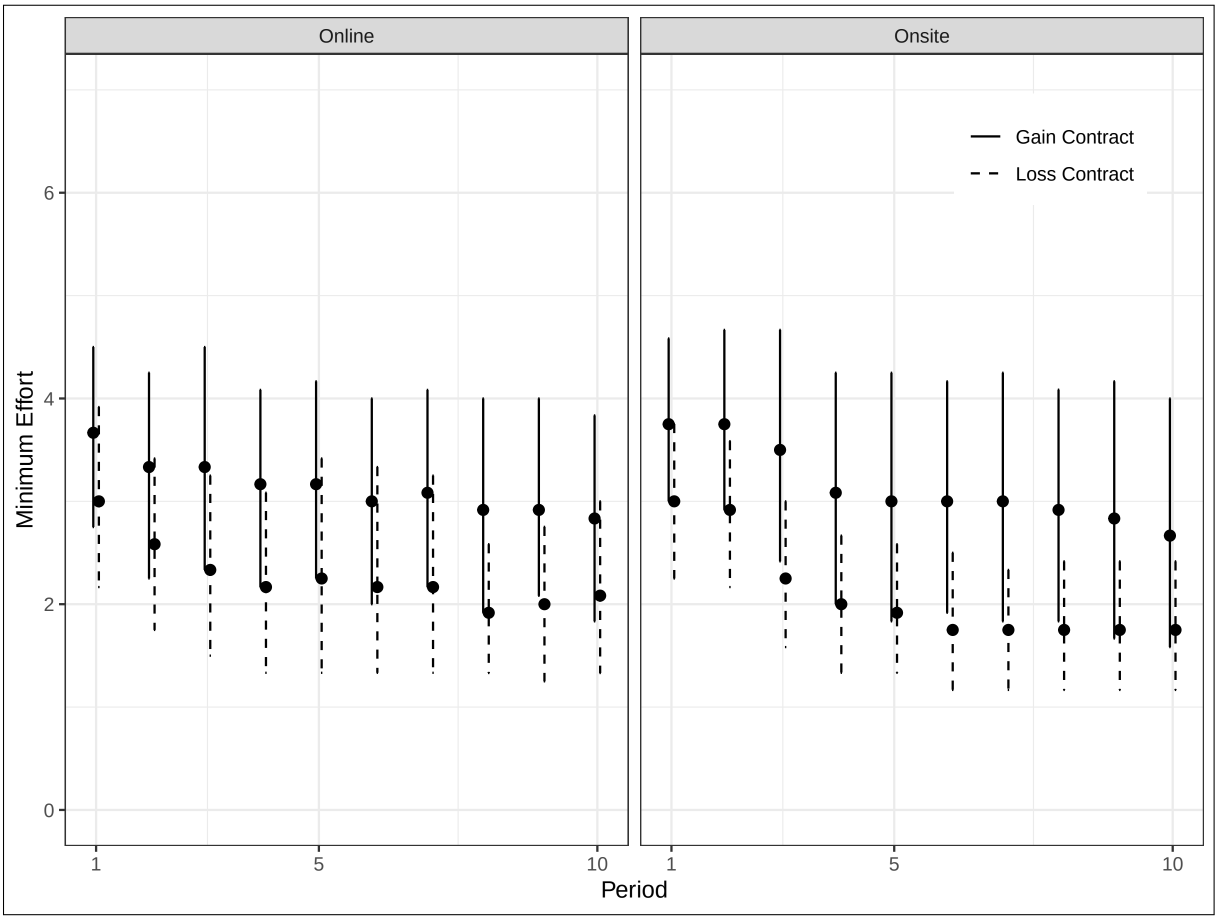


Figure 6: Summary of all minimum efforts per period across treatments by type of session (online or onsite). The figure presents the average minimum effort (black dot) and its bootstrapped 95% confidence interval (dashed line for loss contracts, solid line for gain contracts) per treatment.

In Table 9 we reproduce Table 2 with an interaction between the online and treatment dummy ($loss\_contract \times online$). The interaction does not detect a statistically significant impact of running our treatment sessions online, reinforcing the results from Figure 6.

	(1)	(2)	(3)
	periodmineffort	*periodmineffort*	*periodmineffort*
loss_contract	-1.067*	-1.287*	-1.458**
	(0.617)	(0.658)	(0.608)
online	-0.00833	-1.175	-1.505*
	(0.728)	(0.838)	(0.817)
loss_contract × *online*	0.192	0.338	0.650
	(0.883)	(0.888)	(0.802)
female_ratio		3.481***	4.533***
		(1.125)	(1.136)
avg_risk_av		-0.533**	-0.639***
		(0.235)	(0.231)
avg_loss_av		-0.0481	-0.0464
		(0.167)	(0.167)
avg_ambiguity_av		-0.0361	-0.0994
		(0.189)	(0.212)
avg_crt		0.304**	0.227
		(0.145)	(0.153)
constant	3.150***	5.064**	8.765**
	(0.554)	(2.522)	(4.392)
N	480	480	480
Big Five Averages	No	No	Yes

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 9: Analysis of the minimum effort of groups across periods using a random effects GLS and the interaction between the treatment and online sessions. In columns (1) to (3) we regress the minimum effort of each group for each period (*periodmineffort*) on the dummies for the loss contract (*loss_contract*) and online sessions (*online*), along with their interaction. In column (2) we add the average personality measures of the group and in column (3) the average values of the Big Five. All standard errors are clustered at the group level.

In Table 10 we reproduce Table 2, but separating the data into onsite and online sessions. Several things stand out in this table. First, for both online and onsite sessions, the sign of the treatment dummy (*loss_contract*) is in line with the results of Table 2 and with Figure 6, where effort levels are lower under loss contract. Second, while the treatment dummy is statistically significant at the 1% level in Table 2, once we split the data into online and onsite sessions, the dummy is significant (at the 5%) for the onsite sessions, but not for the online sessions. While lower power in the split samples might explain some of the differences in the significance of the treatment effect, variance in the gender distribution across the two types of sessions may also play a role. In onsite sessions we had a majority of males, whereas in online sessions gender was perfectly balanced. Because subjects were randomly assigned within each session, gender is very balanced in the online session groups but not in onsite sessions, where several groups are all male (see Figure 7). As we discuss in Appendix G, the ratio of females in loss contract sessions plays a big role in determining the effect size of our treatment (see Table 18 and Figure 10 in Appendix G). Therefore, we believe that the lack of variance in the gender composition of online groups can help explain the lower treatment effects in such sessions. Furthermore, the homogeneity in the ratio of females in online sessions might also explains why we cannot reproduce the gender effects of onsite sessions.

Overall, the results of the online and onsite sessions are very similar (see Figure 6). In both cases loss contracts have a detrimental effect on the minimum effort of groups and in both cases the group dynamics are similar. Moreover, if we interact the treatment and online dummy, the interaction is not significant (see Table 9). While the effects are stronger in onsite sessions, based on the evidence presented in this appendix, we conclude that the online data supports our onsite findings and reinforces the original results.

	Onsite			Online		
	(1)	(2)	(3)	(4)	(5)	(6)
	periodmineffort	*periodmineffort*	*periodmineffort*	*periodmineffort*	*periodmineffort*	*periodmineffort*
loss_contract	-1.067*	-1.151*	-1.272**	-0.875	-0.288	-0.610
	(0.623)	(0.691)	(0.640)	(0.637)	(0.646)	(0.632)
female_ratio		2.644**	3.857***		2.124	3.776
		(1.163)	(1.163)		(3.347)	(2.892)
avg_risk_av		-0.279	-0.595*		-1.158***	-1.115***
		(0.274)	(0.317)		(0.253)	(0.306)
avg_loss_av		-0.175	-0.342		0.287	0.375**
		(0.202)	(0.259)		(0.190)	(0.163)
avg_ambiguity_av		0.00392	-0.0640		-0.454	-0.785***
		(0.303)	(0.400)		(0.334)	(0.252)
avg_crt		0.255	0.290		0.372*	0.351*
		(0.182)	(0.206)		(0.218)	(0.185)
constant	3.150***	4.325	4.579	3.142***	9.045**	31.90***
	(0.560)	(3.143)	(5.593)	(0.477)	(4.169)	(5.872)
N	240	240	240	240	240	240
Big Five Averages	No	No	Yes	No	No	Yes

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 10: Analysis of the minimum effort of groups across online and onsite sessions using a random effects GLS. In columns (1) to (3) we use only the observations for onsite sessions and regress the minimum effort of each group for each period (*periodmineffort*) on the dummies for the loss contract (*loss_contract*) (column (1)), along with the average personality measures of the group (column (2)) and the average values of the Big Five (column (3)). In columns (4) to (6) we use only the observations for online sessions and regress the minimum effort of each group for each period (*periodmineffort*) on the dummies for the loss contract (*loss_contract*) (column (4)), along with the average personality measures of the group (column (5)) and the average values of the Big Five (column (6)). All standard errors are clustered at the group level.

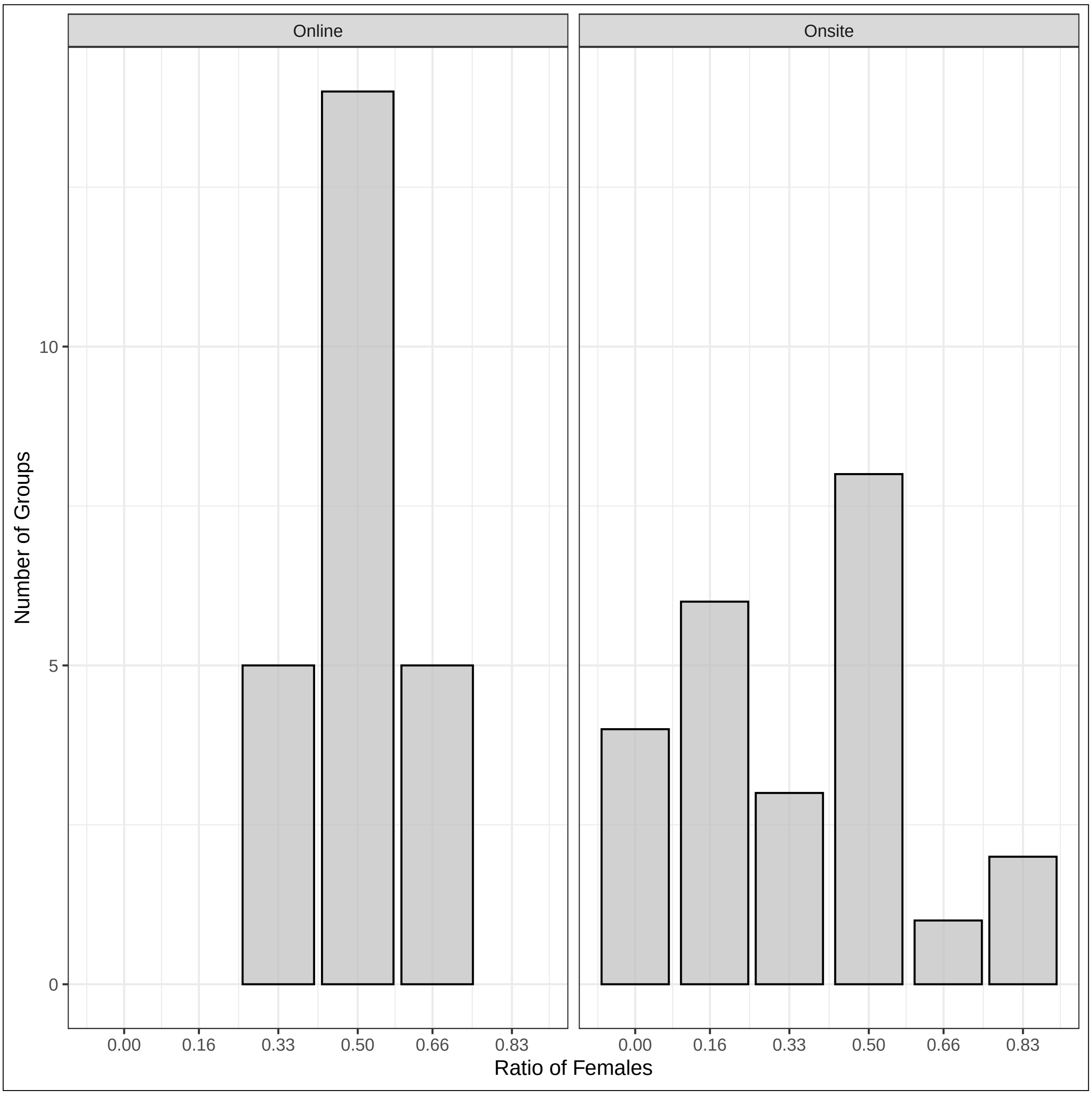


Figure 7: Distribution of the share of females per group and session type. The bar graph presents the number of groups (vertical axis) for each possible ratio of females (horizontal axis) separated by online or onsite sessions.

D Individual effort provision

Table 11 presents a random effects model of subjects' individual effort ($effort_{i,t}$) across all periods t. Unlike Table 2, in Table 11 the independent variables are individual personality measures and not group averages. The findings of Table 11 complement and confirm the results of Table 2; loss contracts do not only lower the *minimum* effort of groups, they also decrease subjects' individual effort levels. However, in contrast to Table 2 we do not detect a significant effect of online sessions, gender, or risk aversion on individual effort levels or on any other personality measures.

	(1) *effort*	(2) *effort*	(3) *effort*
loss_contract	-0.895**	-0.922**	-0.908**
	(0.390)	(0.389)	(0.388)
online	0.0924	-0.0367	0.00958
	(0.390)	(0.383)	(0.377)
female		0.229	0.290*
		(0.158)	(0.172)
risk_av		-0.0480*	-0.0506
		(0.0292)	(0.0327)
loss_av		-0.0287	-0.0286
		(0.0235)	(0.0251)
ambiguity_av		-0.0195	-0.0212
		(0.0304)	(0.0315)
crt		0.0349	0.0290
		(0.0283)	(0.0283)
constant	4.207***	4.726***	5.316***
	(0.370)	(0.522)	(1.067)
Big Five	No	No	Yes
N	2880	2880	2880

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 11: Analysis of the individual effort provision using a random effects GLS. In columns (1) to (3) we regress the individual effort levels for each subject (*effort*) on the dummies for loss contract (*loss_contract*) and online sessions (*online*) along with the individual personality measure (column (2)) and the individual values of Big Five (column (3)). All standard errors are clustered at the group level.

E Response to Observed Minimum Effort

To understand better the dynamics within groups, in this section we examine the way subjects reacted to the last observed minimum effort of the group. To do this, we define the variable $response_{i,t}$, which measures the difference between the effort exerted by subject i in period t ($effort_{i,t}$) and the minimum effort of the group g in period $t-1$ ($minimum_{g,t-1}$). Formally:

$$response_{i,t} = effort_{i,t} - minimum_{g,t-1}. \tag{2}$$

If the value of $response_{i,t}$ is above zero, then subject i played an effort level above last period's group minimum effort ($minimum_{g,t-1}$), trying to push the effort level of the group upwards. If $response_{i,t}$ takes a value below zero, then it means the subject undercut the effort level of her group, playing a lower effort than that observed in the previous period. Finally, if $response_{i,t}$ takes a value of zero, then subject i best-responded to the last observed minimum effort level of the group.

Table 12 shows the results of an OLS regression (column (1) to (3)) using only data for the second period of the experiment and a random effects GLS (column (4) to (6)) for periods 2 to 10.[14] The dependent variable in both cases is $response_{i,t}$. In line with the regression on individual effort provision (Table 11), the results indicate that none of the personality traits can significantly explain the behavior of subjects. Further, against our previous results, loss contracts do not seem to affect the response of subjects.

In Table 13 we use a logit model to study the likelihood that subjects *best respond* to the last minimum effort played by the group. The dependent variable is $best\_response_{i,t}$, which takes a value of 1 if the subject exerts the same level of effort as the minimum effort of its group in the previous period (i.e., if $effort_{i,t} = minimum_{g,t-1}$) and of 0 otherwise. In columns (1) to (3) we present a logit using only the data for period $t = 2$. In columns (4) to (6) we present a random effects model for all periods $t > 1$. We cannot detect any effect of personality traits on the likelihood to best respond to the last observed minimum effort of the group. Only the dummy for online sessions seems to have explanatory power, with online sessions increasing the likelihood to best respond to the group minimum effort

[14] We are especially interested in period $t = 2$ because it is the first period with feedback of the groups' behavior.

	OLS (t=2)			**Random Effects GLS** (t>1)		
	(1)	(2)	(3)	(4)	(5)	(6)
	response	*response*	*response*	*response*	*response*	*response*
loss_contract	0.264	0.256	0.263	0.0324	0.0379	0.0504
	(0.297)	(0.297)	(0.292)	(0.126)	(0.124)	(0.119)
online	-0.153	-0.143	-0.0556	0.0262	-0.00312	-0.0204
	(0.297)	(0.324)	(0.330)	(0.126)	(0.128)	(0.131)
female		-0.297	-0.213		-0.0747	-0.128
		(0.220)	(0.252)		(0.0710)	(0.0886)
risk_av		-0.0130	-0.0183		-0.00940	-0.00797
		(0.0388)	(0.0389)		(0.0200)	(0.0200)
loss_av		0.00215	0.00992		-0.0161	-0.0205*
		(0.0346)	(0.0349)		(0.0109)	(0.0117)
ambiguity_av		-0.0252	-0.0226		-0.0177	-0.0160
		(0.0391)	(0.0395)		(0.0156)	(0.0155)
crt		-0.00218	-0.00989		-0.0156	-0.0131
		(0.0309)	(0.0321)		(0.0118)	(0.0119)
constant	1.750***	2.188***	3.459***	0.880***	1.388***	0.842*
	(0.227)	(0.619)	(1.001)	(0.102)	(0.220)	(0.454)
Big Five	No	No	Yes	No	No	Yes
N	288	288	288	2592	2592	2592

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 12: Analysis of the response of subjects to the last observed minimum effort of the group (*response*). In columns (1) to (3) we use an OLS regression to analyze the response of subjects to the minimum effort of the group observed in period one. In columns (4) to (6), we regress the response of subjects across all periods using random effects GLS. We regress *response* on the dummies for loss contract (*loss_contract*) and online sessions (*online*) along with the individual personality measure (column (2) & (5)) and the individual values of Big Five (column (3) & (6)). All standard errors are clustered at the group level.

of period $t = 1$. As in Table 12, the treatment dummy (*loss_contract*) is not statistically significant.

To study the undercutting behavior of subjects, Table 14 displays the frequencies of the variable $response_{i,t}$ per period. The table shows a clear convergence of $response_{i,t}$ towards zero. This convergence comes both from the positive and negative side of the table. Interestingly, Table 14 also shows that only a minority of subjects ever pick an effort level below the last observed minimum effort of the group. Out of the 2,880 decisions made across all periods and sessions, only 73 responses were below the observed $minimum_{g,t-1}$. Since some subjects repeatedly went below the observed $minimum_{g,t-1}$, in total, only 53 out of the total 288 subjects (18%) ever undercut their last observed group minimum.[15] On the other hand, a substantial number of subjects pick effort values above the last observed minimum.

The descriptive behavior observed in Table 14 could explain the variance observed in Table 3. Because subjects are seldom willing to undercut their groups' past minimum behavior ($minimum_{g,t-1}$), this value acts as a *de facto* lower bound for their choices. This is confirmed in Figure 8, where we present the boxplots for each level of $minimum_{t-1}$ across both treatments. The level of dispersion of responses is greater the lower the group's minimum. In other words, groups that reach high minimum effort values present little variance in the response of its components. Because the minimum effort of groups drops faster under loss contracts, this "wider" space for subjects to pick effort levels translates into higher variance, more wasted efforts, and contributes to the lower payoffs reported in Section 3.3.

Finally, an interesting insight comes from analyzing those subjects who are driving the results: i.e., those who undercut the last observed group effort. In Table 15 we interact our treatment variable with the individual level of risk aversion for those subjects that undercut the group at least once in the whole session. The results show that the interaction between risk aversion and loss contracts is negative and significant at the 1%. This result is in line with the prediction of Pierce et al. (2022) who argue that loss contracts under uncertainty might affect the risk tolerance of agents and push them towards less risky strategies.

[15] A probit model using a dummy variable for undercutting the effort of the group (not presented) reveals no statistically significant personality traits of these subjects.

	Logit (t=2)			**Random Effects Logit** (t>1)		
	(1)	(2)	(3)	(4)	(5)	(6)
	best response	*best response*	*best response*	*best response*	*best response*	*best response*
loss_contract	0.181	0.201	0.212	0.250	0.257	0.253
	(0.297)	(0.301)	(0.307)	(0.276)	(0.273)	(0.266)
online	0.727**	0.957***	0.919***	-0.0896	0.0345	0.0792
	(0.297)	(0.326)	(0.344)	(0.274)	(0.278)	(0.283)
female		-0.187	-0.251		-0.148	-0.113
		(0.327)	(0.355)		(0.174)	(0.181)
risk_av		0.104	0.105		0.0302	0.0287
		(0.0693)	(0.0691)		(0.0302)	(0.0329)
loss_av		0.0110	0.00552		0.0240	0.0306
		(0.0460)	(0.0489)		(0.0256)	(0.0279)
ambiguity_av		0.0736	0.0724		0.0511	0.0537
		(0.0603)	(0.0603)		(0.0335)	(0.0343)
crt		-0.0405	-0.0381		-0.000228	-0.00462
		(0.0398)	(0.0416)		(0.0290)	(0.0285)
constant	-1.918***	-3.260***	-4.017**	0.00575	-0.868	-0.136
	(0.275)	(1.037)	(1.578)	(0.259)	(0.548)	(0.920)
Big Five	No	No	Yes	No	No	Yes
N	288	288	288	2592	2592	2592

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 13: Analysis of the response of subjects to the last observed minimum effort of the group (*response*). In columns (1) to (3) we use a logit regression to analyze whether a subject played a best response to the minimum effort of the group observed in period one. In columns (4) to (6), we conduct a random effects logit regression across all periods to investigate whether subjects best responded to the observed minimum of the previous period. We regress the dummy variable *best response* on the dummies for loss contract (*loss_contract*) and online sessions (*online*) along with the individual personality measure (column (2) & (5)) and the individual values of Big Five (column (3) & (6)). All standard errors are clustered at the group.

Gain Contract										
	Period									
Response	2	3	4	5	6	7	8	9	10	Total
-3	0	0	1	0	0	0	0	0	0	1
-2	3	3	2	2	1	0	1	1	0	13
-1	5	3	3	1	4	2	4	1	6	29
0	**26**	**47**	**49**	**73**	**76**	**85**	**89**	**96**	**103**	**644**
1	38	45	50	37	31	30	33	29	18	311
2	29	16	20	13	16	14	11	5	8	132
3	27	12	9	10	8	6	3	6	2	83
4	7	11	4	6	3	3	2	3	1	40
5	7	5	2	1	1	4	0	2	4	26
6	2	2	4	1	4	0	1	1	2	17
Total	144	144	144	144	144	144	144	144	144	1296

Loss Contract										
	Period									
Response	2	3	4	5	6	7	8	9	10	Total
-4	2	0	0	0	0	0	0	0	0	2
-3	1	1	0	0	0	0	1	0	0	3
-2	0	2	1	0	2	0	0	0	0	5
-1	3	8	5	2	2	0	0	0	0	20
0	**30**	**38**	**58**	**69**	**86**	**97**	**106**	**108**	**113**	**705**
1	31	43	38	40	32	32	23	23	17	279
2	26	11	13	13	8	4	3	5	7	90
3	25	18	11	10	5	4	5	6	3	87
4	9	13	7	6	2	2	1	1	1	42
5	3	6	4	1	3	0	2	0	0	19
6	14	4	7	3	4	5	3	1	3	44
Total	144	144	144	144	144	144	144	144	144	1296

Table 14: Frequency of responses to the group minimum effort of the previous period for period 2 to 10. The upper (lower) panel presents the number of times that each frequency of responses under gain contracts (loss contracts), whereas a response of zero (in bold) corresponds to the best response.

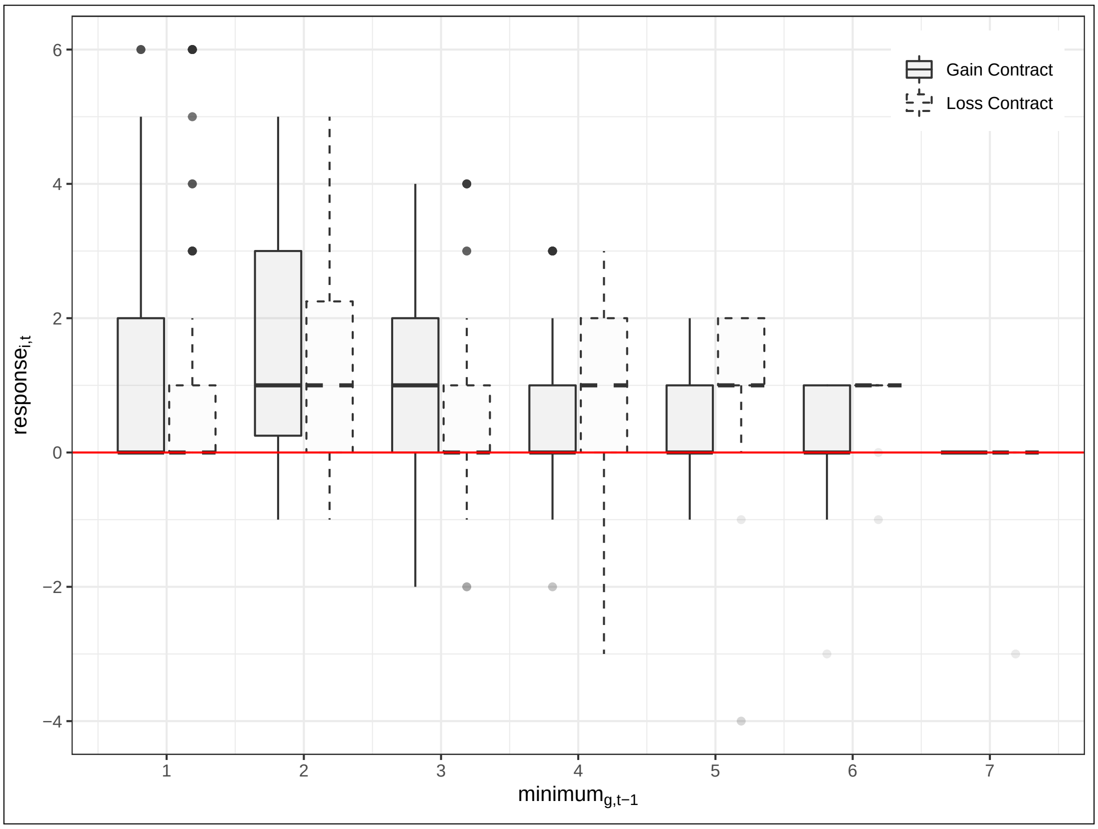


Figure 8: Distribution of responses. Boxplots representing the distribution of responses ($response_{i,t}$) to the observed group minimum effort of the previous period ($minimum_{g,t-1}$) across treatments.

	(1)	(2)	(3)
	response	*response*	*response*
loss_contract	1.118**	1.319**	1.596**
	(0.436)	(0.567)	(0.678)
risk_av	0.104***	0.0997**	0.114**
	(0.0229)	(0.0434)	(0.0500)
loss_contract × *risk_av*	-0.186***	-0.215***	-0.253***
	(0.0647)	(0.0770)	(0.0964)
online	-0.121	-0.205	-0.272
	(0.234)	(0.225)	(0.258)
female		0.299	0.291
		(0.188)	(0.215)
loss_av		0.0461*	0.0483
		(0.0258)	(0.0305)
Ambiguity_av		-0.0360	-0.0489
		(0.0309)	(0.0300)
crt		-0.0282	-0.0310
		(0.0245)	(0.0302)
constant	-2.105***	-2.047***	-1.449
	(0.227)	(0.341)	(0.967)
N	73	73	73
Big Five	No	No	Yes

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 15: Analysis of the response of subjects to the last observed minimum effort of the group (*response*) and the interaction between the treatment and subjects' risk aversion using a random effects GLS. In column (1) we regress the response of subjects who undercut their groups on the dummy for loss contract (*loss_contract*), risk aversion (*risk_av*), their interaction, as well as a dummy for online sessions (*online*). In column (2) we add extra personality controls, and in column (3) we add the Big Five measures. All standard errors are clustered at the group level.

F Extra Figures & Tables

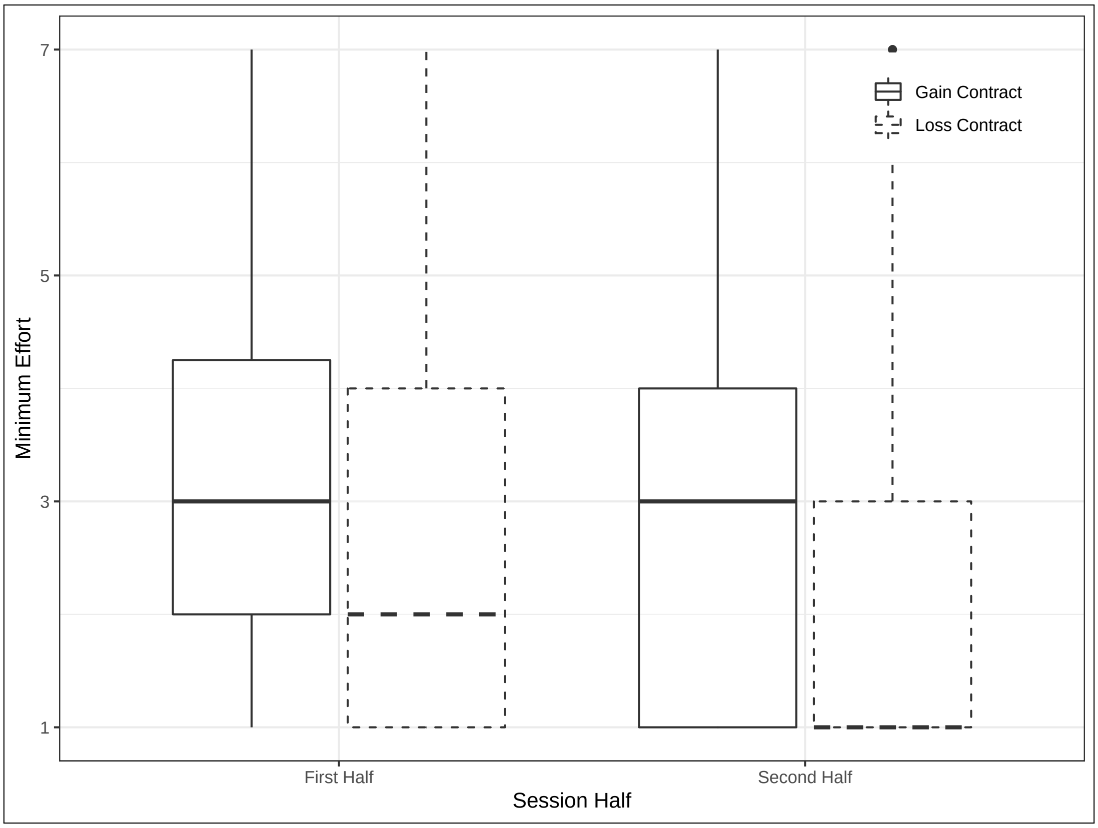


Figure 9: Comparison of the distribution of minimum efforts across treatments for the first and second half of the experiment.

Gain Contract											
	Period										
Minimum effort	1	2	3	4	5	6	7	8	9	10	Total
1	2	4	7	7	8	8	8	9	10	12	75
2	3	1	0	3	3	3	2	2	2	0	19
3	5	9	6	4	3	4	6	5	4	4	50
4	8	4	4	5	4	4	2	3	2	3	39
5	2	1	2	1	2	1	2	1	2	1	15
6	4	4	4	3	3	3	3	3	3	3	33
7	0	1	1	1	1	1	1	1	1	1	9
Total	24	24	24	24	24	24	24	24	24	24	240
Loss Contract											
	Period										
Minimum effort	1	2	3	4	5	6	7	8	9	10	Total
1	7	7	11	13	14	15	15	15	15	15	127
2	1	3	2	3	1	1	1	1	1	1	15
3	5	6	7	3	5	5	5	5	5	5	51
4	7	6	2	4	3	2	2	3	2	2	33
5	4	1	1	0	0	0	0	0	1	0	7
6	0	1	1	1	0	0	0	0	0	1	4
7	0	0	0	0	1	1	1	0	0	0	3
Total	24	24	24	24	24	24	24	24	24	24	240

Table 16: Frequency of effort levels played as minimum effort within a group in each period of each treatment. In the upper (lower) panel we present the number of times that each level of effort (first column) was played as minimum effort for each period (columns 2-11) under gain contracts (loss contracts).

	(1)	(2)	(3)
	periodmineffort	*periodmineffort*	*periodmineffort*
loss_contract	-0.971**	-0.920**	-1.000**
	(0.441)	(0.427)	(0.435)
online	0.0875	-0.0682	0.0944
	(0.441)	(0.501)	(0.505)
female_ratio		2.044*	2.543**
		(1.183)	(1.015)
maxg_risk_av		0.0642	0.0799
		(0.112)	(0.112)
maxg_loss_av		-0.117	-0.130
		(0.190)	(0.173)
maxg_ambiguity_av		0.0353	-0.0224
		(0.117)	(0.124)
minimumg_crt		0.118	0.0824
		(0.0798)	(0.0783)
constant	3.102***	2.328	7.763
	(0.450)	(2.791)	(4.891)
Big Five Averages	No	No	Yes
N	480	480	480

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 17: Analysis of the minimum effort of groups across periods using a random effects GLS. In columns (1) to (3) we regress the minimum effort of each group for each period (*periodmineffort*) on the dummies for the loss contract (*loss_contract*) and online sessions (*online*), along with the most extreme personality measures of the group (column (2)) and the most extreme values of the Big Five (column (3)). All standard errors are clustered at the group level.

G Gender

Given the strong gender effects detected in our analysis, in this section we investigate its effects on the minimum effort of groups. Our results show that groups with a higher share of females coordinate on higher effort levels (Table 2), coordinate better (Table 3), and consequently, have a higher average payoff (Table 4). When analyzing personality traits across genders, we find that females are more risk averse than males and score lower on the CRT test. These are well known facts in the literature (e.g., Borghans et al., 2009; Brañas-Garza et al., 2019, respectively). However, it cannot explain the treatment differences as gender is balanced across treatments (see Appendix A).

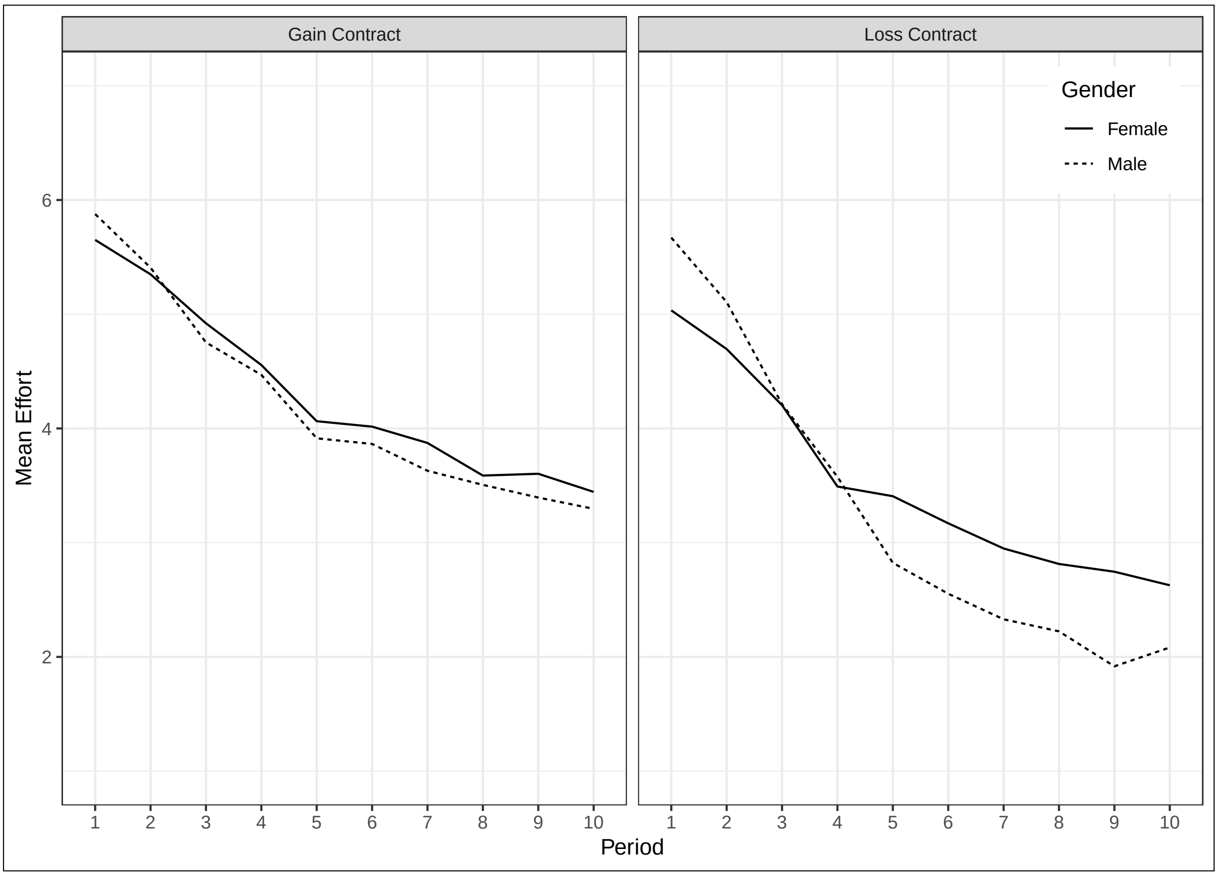


Figure 10: Mean effort decisions by gender and treatment. The left panel shows the mean effort decision of all females per period (solid line) and the mean effort decisions of all males per period (dashed line) for gain contracts. The right panel shows the same for loss contracts.

In Figure 10 we plot the average choices of females and males in each period for both treatments. The figure shows that most of the gender differences stem from the loss contract treatment. A series of Mann-Whitney U tests confirm this. While we cannot reject the null that males and females exert the same level of effort at the beginning of the session under gain contracts (p-value = 0.227), we can for loss contracts (p-value = 0.005). This difference is driven by females changing their behavior under loss contracts. A Mann-Whitney U test comparing the initial effort of males across treatments cannot reject the null of same effort across treatments (p-value = 0.442). In contrast, the same test finds a significant difference in initial female effort across treatments (p-value = 0.024).

Moreover, as seen in Figure 10, the differences across gender reverses as the experiment progresses only under loss contracts. While males start exerting significantly more effort in the first period, by the end of the session they are considerably below the effort of females. In Table 18 we present a random effects GLS regression in which the dependent variable is the group minimum effort using either the gain contracts (columns (1) to (3)) or loss contracts (columns (4) to (6)). In all models, we interact the ratio of females in each group with the period (*period*) of the session. Such interaction has no effect on the evolution of the group minimum effort when using only the data from the gain contract sessions. However, once we use the loss contract data, we see a strong positive effect of the interaction between the ratio of females in each group and the session period. In other words, the results of Table 18 show that groups with a larger proportion of females have a higher minimum effort as the sessions progress under loss contracts but not under gain contracts.

The results that gender has no effect under gain contracts is in line with the existing literature on gender effects in coordination games with strategic complements (e.g., Dufwenberg and Gneezy, 2005; Heinemann et al., 2009; Engelmann and Normann, 2010; Di Girolamo and Drouvelis, 2015). However, the differences that we observe under loss contracts indicate that this might not be a generalizable result and deserves further study.

	Gain Contract			Loss Contract		
	(1)	(2)	(3)	(4)	(5)	(6)
	periodmineffort	*periodmineffort*	*periodmineffort*	*periodmineffort*	*periodmineffort*	*periodmineffort*
female_ratio	0.195	2.686*	1.342	2.171**	2.771**	4.581***
	(1.645)	(1.619)	(1.335)	(0.983)	(1.257)	(1.525)
period	-0.145***	-0.145***	-0.145***	-0.203***	-0.203***	-0.203***
	(0.0487)	(0.0491)	(0.0497)	(0.0525)	(0.0529)	(0.0535)
female_ratio × *period*	0.107	0.107	0.107	0.219**	0.219**	0.219**
	(0.0936)	(0.0944)	(0.0954)	(0.0946)	(0.0954)	(0.0964)
online	-0.106	-1.635*	-1.294	-0.426	-0.454	-0.853
	(0.768)	(0.930)	(1.064)	(0.417)	(0.418)	(0.530)
avg_risk_av		-0.924***	-1.050***		-0.262	-0.717**
		(0.227)	(0.228)		(0.331)	(0.361)
avg_loss_av		-0.324	-0.379*		0.134	0.167
		(0.210)	(0.217)		(0.146)	(0.151)
avg_ambiguity_av		0.0509	0.110		0.0609	-0.0265
		(0.288)	(0.263)		(0.164)	(0.219)
avg_crt		0.664***	0.927***		0.0352	-0.0276
		(0.191)	(0.207)		(0.127)	(0.186)
constant	3.651***	8.440**	2.399	2.121***	2.006	5.191
	(0.773)	(3.663)	(4.827)	(0.491)	(2.893)	(8.217)
Big Five Averages	No	No	Yes	No	No	Yes
N	240	240	240	240	240	240

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 18: Analysis of the minimum effort of groups across periods and the interaction between the ratio of females in each group with the period by treatment using a random effects GLS. In columns (1) to (3), we use only data from gain contract sessions, in columns (4) to (6) we use data only from loss contract sessions. We regress the minimum effort of each group for each period (*periodmineffort*) on the ratio of females in each group (*female_ratio*), the period (*period*), their interaction, and the dummy for online sessions (*online*) along with the individual personality measure (column (2) & (5)) and the individual values of Big Five (column (3) & (6)). All standard errors are clustered at the group level.