Incentives and Motivation in Dynamic Contests*

Arnd Heinrich Klein and Armin Schmutzler†

April 9, 2021

Abstract

This paper uses a contest setting to analyze the provision of intertemporal incentives in organizations. Should a principal repeatedly award small prizes or give a large prize that takes past performance into account? A simple theoretical model predicts higher efforts in the latter case. An experiment confirms this prediction, but the size of the effect is smaller than expected. This result reflects two observations of independent interest. First, there is a revenge effect for laggards in repeated contests: Laggards exert higher efforts than leaders with the same first-period effort level. Second, there is an intimidation effect for laggards in the single-prize case: Laggards exert lower efforts than leaders with the same first-period effort level. Moreover, we observe polarization in laggard behavior.

Keywords: Dynamic contests, rank-order tournaments, incentive systems, intimidation effect, revenge effect, experiments

JEL Codes: D20, M51

*We are grateful to Andreas Beerli, Jean-Michel Benkert, Dan Burghart, Alain Cohn, Vincent Crawford, Donja Darai, Somnath Datta, Florian Ederer, Ronald Indergand, Hedda Herdis Klein, Patrick Kline, Igor Letina, John Morgan, Yagiz Özdemir, Andrea Portmann, Eva Ranehill, Frédéric Schneider, Carmit Segal, Roberto Weber, Tony Williams, and seminar participants in Konstanz, Norwich (Contests: Theory and Evidence), Nuremberg, Schaffhausen (Zurich Workshop on Economics) and Zurich for helpful discussions and suggestions. Furthermore, we gratefully acknowledge financial support of the Swiss National Science Foundation (grant numbers 131854 and 151688).

†Arnd Heinrich Klein: University of Zurich; Armin Schmutzler: Department of Economics, University of Zurich and CEPR; armin.schmutzler@uzh.ch.

1 Introduction

Employees at General Electric in the 1980s and 1990s had a tough life: Every year, supervisors ranked their personnel according to relative performance into the best 20%, the middle 70%, and the worst 10% of their team. This ranking provided a basis for bonuses, promotions, and even firing decisions (Welch, 2005). While GE has replaced this rigid approach by more flexible mechanisms in the meantime,[1] the use of relative incentive schemes – and, in particular, contests – for bonuses and promotions within companies is still widespread (see, for example, Chlosta et al., 2014).[2]

As in the example of General Electric, contests are often dynamic. Firms usually make their employees compete against each other repeatedly, generating a series of relative performance signals. The dynamic nature of contests poses two interrelated questions: (1) Should firms consider the entire performance history when distributing bonuses and making promotion decisions, or should they focus only on the most recent performance signals? (2) How should firms distribute rewards across periods?

The answer to these questions is not obvious. On the one hand, basing agents' rewards on past performance softens competition later on, as future effort has smaller effects on winning chances. On the other hand, the expectation that effort will affect future winning chances intensifies competition in early periods. The problem becomes more complex if agents care about more than pure monetary preferences – risk aversion, loss aversion, joy of winning and social preferences are known to be potentially important effort determinants in contests. For instance, when organizations take past performance into account, some agents may be handicapped in current competition by large performance lags. As a result, they might experience frustration and give up competition. Conversely, when organizations frequently restart the competition, opponents have a more realistic chance to make up for past losses by exerting effort, and may even be willing to exert additional efforts, reflecting revenge motives. Finally, agents may display heterogeneity with respect to these motives. The effects of contest design on efforts are therefore hard to predict.

In view of the potentially complex agent responses to different incentive structures, this paper deals experimentally with dynamic contest design by comparing two-period contests with identical prize sum. The first contest consists of a sequence of two independent noisy rank-order tournaments with identical prizes in each period. Thus, in each period the firm only rewards the most recent performance signal. We compare this contest with one in which the principal gives only one prize after the second period, but considers the entire history and weighs the (noisy) performance observations from both periods equally. In this latter contest, relative performance is revealed to the agents before the second period.

[1] See McGregor (2013) for an overview on recent developments.

[2] Contests also play a crucial role beyond the human resources context (Konrad and Kovenock (2009)).

Even in the benchmark case that agents only care about expected monetary payoffs, the equilibrium analysis is non-trivial. Nevertheless, it is possible to show that both contests have an equilibrium with symmetric efforts in both periods. When there is only one prize, the equilibrium effort in the second period is decreasing in first-period performance difference. Intuitively, with increasing first-period performance asymmetry the second-period effort is less likely to determine the winner, which reduces the incentives of both players. Moreover, we find that the expected effort in the treatment with one prize is much higher than with two prizes (41% for our parameters).[3]

Some important aspects of the observed behavior are consistent with the equilibrium analysis of our benchmark model. First, in the single-prize treatment, second-period efforts depend negatively on performance differences. Second, the experiment shows that efforts in the single-prize treatment are higher in the treatment with one large prize than in the case with two small prizes. However, the difference in the two treatments is only about 10%, rather than the predicted 41%. This observation reflects the fact that there is substantial over-expenditure in the two-prize treatment in both periods, whereas with one prize it is much smaller in the first period and absent in the second.

To understand the observed behavior, we focus on the second period. Unlike in the one-prize case, with independent tournaments laggards (players whose first-period performance was worse than their competitor's) have the same winning chances in the second period as leaders (whose performance was better than their competitor's). Even though both agents choose the same effort in equilibrium, we observe that laggards exert more effort in the second period of the two-prize treatment than leaders. Next, we find a striking effect of relative first-period performance on second-period behavior in the single-prize treatment: Laggards exert lower effort in the second period than leaders with the same first-period effort, even though there is no difference in equilibrium. Together, these two observations help to explain our main results: The behavior of laggards contributes to the relatively low effort in the single-prize treatment as well as the high effort in the two-prize treatment. However, an important qualification is necessary: The observed behavior refers to the *average* laggard. A more careful look at the one-prize treatment reveals considerable heterogeneity. Whereas about 25% of the laggards choose the minimal allowable effort in the second period (most of them are lagging far behind), almost as many players choose the maximal allowable effort (most of them are close to the leader).

To explain the second-period deviations from equilibrium, we enrich our simple theoretical framework with two sources of behavioral variation. First, the payoff from winning includes a non-monetary component, interpreted as *joy of winning*, which varies across

[3] This is consistent with the more general analysis of two-period contests in Klein and Schmutzler (2017) according to which the contest with one big prize at the end is optimal in the sense that expected total efforts are higher than for any other combination of prize distribution and performance weights.

players and depends on the history of the game. In the two-prize treatment, this component is higher when a player has previously lost than when he has won. Joy of winning is a standard explanation for excess expenditure in contests. The assumption that joy of winning in two-prize treatments is higher for laggards than for leaders fits well with behavioral notions such as inequity aversion: A player who has previously lost will typically have lower first-period payoff than the opponent; inequity aversion would therefore increase the benefits from winning in the second period, suggesting that the observed high efforts stem from a revenge effect.

Second, we allow for the possibility that players' expected winning probabilities may depend on factors that do not affect the equilibrium in the baseline model and that players differ in the way in which they form these expectations. For instance, laggards might interpret the better relative performance of the opponent as a signal of high joy of winning and thus of high willingness to exert effort.

In this setup, we find that optimal behavior of laggards frequently involves exerting maximal effort (a *wake-up effect*) or giving up completely (an *intimidation effect*), depending on their joy of winning. This explains the polarization of laggard efforts in the one-prize treatment. The observed low efforts of laggards (which are decreasing in the size of the performance differential) then suggest that the intimidation effect prevails at the aggregate level. All told, the augmented model can explain why multiple prizes do not lead to much lower aggregate efforts than single-prize tournaments.

The rest of the paper proceeds as follows. In Section 2, we sketch the model and its results. Section 3 develops treatments and hypotheses and describes the experimental design. In Section 4, we present the results. Section 5 introduces the framework with which we interpret the experimental observations. Section 6 discusses the relation to the literature. Section 7 concludes.

2 The Benchmark Model

2.1 Assumptions

To capture the design problem of a principal that uses competitive incentives repeatedly, we use a particular parameterization of a two-stage contest analyzed in Klein and Schmutzler (2017).[4] Two risk-neutral agents, $i \in \{1, 2\}$, choose efforts $e_{it} \geq 0$ in periods $t \in \{1, 2\}$, with costs $K(e_{it}) = k(e_{it})^2/2$ for $k > 0$. At the end of each period t, the principal observes agent i's *performance* $s_{it} = e_{it} + \epsilon_{it}$, an imperfect effort measure. ϵ_{it} is a stochastic observation error, independently distributed across agents and periods. We

[4]We refer the reader to that paper for details and derivations.

assume that the difference of the observation errors $\Delta\epsilon_{it} = \epsilon_{it} - \epsilon_{jt}$ is normally distributed: $\Delta\epsilon_{it} \sim \mathcal{N}(0, \sigma^2)$. We denote this distribution and its density as $F()$ and $f()$, respectively.

The principal has a fixed budget W for prizes. She assigns the first-period prize $W_1 \in [0, W]$ to the agent with the highest performance in Period 1, so that agent i receives the prize if $s_{i1} > s_{j1}$. The interim performance is revealed to both agents. Furthermore, the principal assigns $W_2 = W - W_1$ as a second-period prize to the agent with the highest weighted sum of performance in Period 1 and Period 2: Agent i receives W_2 if $s_{i2} + \eta s_{i1} > s_{j2} + \eta s_{j1}$ for the performance weight $\eta \geq 0$. Hence, the contest is determined by the distribution of the prize money across periods (as implied by the first-period prize W_1) and the performance weight η.

2.2 Equilibrium Predictions

We first characterize the behavior of the agents in Period 2 given *relative first-period performance* $\Delta s_{i1} = s_{i1} - s_{j1}$. Second-period efforts in the unique equilibrium are

$$e_{i2}^*(\Delta s_{i1}) = f(\eta \Delta s_{i1})(W - W_1)/k. \tag{1}$$

Intuitively, agents weigh the marginal effort cost ke_{i2}^* against the marginal benefit from higher effort $(f(\eta\Delta s_{i1})(W - W_1))$. The resulting first-order condition (1) implies that greater asymmetry in first-period performance, i.e., higher $|\Delta s_{i1}|$ reduces the efforts of leaders and laggards by the same amount whenever the first-period weight η is positive. Moreover, an increase in η reduces second-period effort. First-period equilibrium efforts in the symmetric equilibrium are

$$e_1^*(\eta, W_1) = \frac{1}{k\sigma\sqrt{2\pi}}\left(W_1 + \frac{\eta}{\sqrt{1+\eta^2}}(W - W_1)\right). \tag{2}$$

Intuitively, when equalizing the marginal costs and benefits of higher first-period efforts, agents consider the immediate positive effects on winning W_1 as well as the improved chances of winning $W_2 = W - W_1$. For fixed total budget W, an increase in the second-period prize reduces the first-period prize by the same amount; the total effect always is a reduction of first-period efforts. Moreover, an increase in the first-period weight η increases first-period effort. Finally, the expected second-period efforts are

$$E(e_2^*(\eta, W_1)) = \frac{W - W_1}{k\sigma\sqrt{2\pi}\sqrt{1+\eta^2}}. \tag{3}$$

Note that expected second-period efforts are decreasing in the size of the weight η: The expected handicap for the laggard in the second period increases with the weight of past performance, which decreases (expected) second-period effort incentives for both players.

2.3 Optimal Contest

From (2) and (3), we can easily derive the optimal contest. We define *total efforts* as $e_{11}+e_{12}+e_{21}+e_{22}$ and *average efforts* (of agent i per period) as $(e_{i1}+e_{i2})/2$. Because efforts are symmetric in both periods, maximization of expected total and expected average efforts is equivalent, and we obtain the following result:

Proposition 1. *(i) Whenever $\eta > 0$, the optimal first-period prize is $W_1 = 0$.*
(ii) Whenever $W_2 > 0$ and thus $W_1 < W$, the optimal weight of past performance is $\eta = 1$.

Hence, the optimal contest has $W_1^* = 0$ and $\eta^* = 1$: There is one prize at the end of Period 2, with equal weight on both efforts. Result (i) captures a simple intuition: A positive first-period prize not only reduces the funds for inducing second-period efforts, it also weakens incentives for exerting first-period efforts with the goal of winning W_2. These two adverse effects dominate the positive effect of a higher W_1 on first-period efforts. Result (ii) is similarly intuitive. A higher η increases the marginal benefits of first-period efforts and decreases those of second-period efforts. The first effect dominates for low η; the second one for high η; the optimum is when these incentives are balanced.

3 Experimental Design

3.1 Treatments and Hypotheses

In our experiment, we set parameter values $k = 0.066$, $\sigma = 40$, and $W = 300$.[5] We compare the theoretically optimal single-prize contest (ONE) and a setting with two independent tournaments (TWO). Thus, in ONE, $W_1 = 0$, $W_2 = W$ and $\eta = 1$, whereas in TWO, $W_1 = W_2 = \frac{W}{2}$ and $\eta = 0$. Table 1 contains the predicted expected average effort (in total and in each period). Our central hypothesis follows directly from this table (and is implied by the general discussion in the previous section).

Hypothesis 1. *Efforts for each player and period, and thus average efforts, are higher under ONE than under TWO.*

Table 1 shows that the effects are substantial: Switching from two independent prizes to the optimal single-prize contest increases expected efforts in each period by 41%.

[5] $\Delta\epsilon_{it} \sim \mathcal{N}\left(0, 40^2\right)$ corresponds to $\epsilon_{it} \sim \mathcal{N}\left(0, 28.28^2\right)$

Table 1: Point predictions of efforts

	TWO	INT	ONE
Average effort	22.7	27.4	32.1
Effort in period 1	22.7	38.7	32.1
Effort in period 2	22.7	16.0	32.1

To improve our understanding of agent behavior, we add an intermediate treatment INT with two identical prizes (which is not optimal), but optimal first-period performance weight $\eta = 1$. This allows us to separate the effect of having one prize rather than two from the effect of putting weight on past performance.[6] By construction, predicted average efforts in INT must be higher than in TWO (approximately 20% by Table 1), because the weight has been optimally adjusted. Predicted average efforts in ONE must be higher than in INT (17% by Table 1), because the prize distribution in INT is not optimal. Finally, in INT predicted efforts in Period 1 are higher than in the two main treatments, whereas they are lower in Period 2: INT gives incentives for first-period efforts by a first-period prize and a first-period weight rather than only by a first-period weight.

3.2 Laboratory Experiment

We conducted three sessions with 32 participants each, each consisting of 30 rounds.[7] Within each session, every treatment was repeated for 10 rounds. The order of the treatments varied across sessions (see Table 2).[8] At the beginning of the sessions, the participants were randomly assigned to matching groups of size 8. In every round, pairs were randomly formed within matching groups.[9] Before the first treatment of a session, we distributed instructions about the general structure of the experiment and instructions specific to the first treatment, and the participants had to answer general control questions. Immediately before each treatment, the subjects received treatment-specific instructions and control questions.[10]

[6] Moreover, the INT treatment is interesting in its own right. It seems consistent with anecdotal evidence that firms who have one prize to distribute in each period pay attention to past performance.

[7] The sessions took place at the computer lab of the University of Zurich, Switzerland, in December 2013 and lasted for about 120 minutes, using z-Tree (Fischbacher, 2007). The subjects were recruited from university students of all fields except economics and psychology.

[8] The design is not perfectly balanced, as TWO precedes ONE twice, whereas ONE precedes TWO only once. However, as we show below, the main results are quite similar in the perfectly balanced subsample with only observations from Sessions 1 and 3.

[9] The participants were aware that in every round, the other participant in their pair was randomly chosen. The instructions did not mention the existence of matching groups.

[10] Section 3 of the Web Appendix contains the instructions. The instructions were read out aloud whenever they had been distributed. Section 4 of the Web Appendix contains the control questions.

Table 2: Order of treatments

	Session 1	Session 2	Session 3
Rounds 1 – 10	TWO	INT	ONE
Rounds 11 – 20	INT	TWO	INT
Rounds 21 – 30	ONE	ONE	TWO

Each round consisted of two periods. In Period 1, the subjects simultaneously chose efforts from the set $\{0, 0.5, ..., 54.5, 55\}$. Thereafter, random numbers were drawn from the given normal distribution.[11] Next, the computer determined the participants' performance as the sum of their effort level and the random number. We treated Period 2 in the same way as Period 1.[12] After Period 2, the computer calculated prizes. The participants' payoff from a particular round was equal to the sum of an endowment of 200 points and the prizes received in that round, net of effort costs. The endowment ensured that a participant's payoff from a round would never be below 0. At the end of each round, the participants learned their prizes, costs and net payoff. At the end of the session, the computer randomly chose one of the 30 rounds to determine payoffs. The participants were informed about the chosen round and were allowed to review all information received during the 30 rounds. Participants then received, individually and in private, their payoffs from the laboratory experiment and a pre-experimental questionnaire (see below) plus a participation fee of 10 Swiss Francs (CHF). The average total payoff was CHF 50.35, consisting of an average of CHF 36.28 from the laboratory experiment and an average of CHF 14.07 from the pre-experimental questionnaire.[13]

Discussion of Design Decisions Some of our design decisions involved trade-offs. First, even though the underlying model features continuous actions, we opted for a discrete design because the continuous model is easier to solve, whereas the discrete model is easier to present to subjects.[14] Second, we limited the maximal per-period effort to 55. As a budget limit is necessary, we could not completely avoid deviating from the exact set-up of the model. A per-period constraint seemed more appropriate than a global budget with potential spillover effects between rounds. In any event, our effort ceiling

[11] The instructions used a graph to illustrate the symmetry of the distribution and the greater probability of small random numbers (see page 15 of the Web Appendix).

[12] First, the subjects simultaneously chose efforts. Then, the computer drew random numbers, calculated the performance levels, and displayed both performance levels as well as relative performance Δs_{i2}.

[13] In the laboratory experiment, 10 points were worth CHF 1. The exchange rate at the time of the experiment was CHF 1.23 per EUR and CHF 0.91 per USD.

[14] The instructions contained a table and a graph depicting the costs of each allowable input level (see page 14 of the Web Appendix).

is very generous: It is 2.4 (1.7) times as high as the predicted effort in each period in ONE (TWO), and around 20% above any second-period equilibrium effort. Thus, the equilibrium of the unconstrained game is also an equilibrium of the constrained game. In Section 4.3, we discuss how the effort constraint might have influenced results. Third, we chose a within-subject design to increase statistical power.[15] To mitigate the problem that this makes the experiment harder to understand, we interrupted each session when the treatment changed, read out the modified rules and asked control questions. Nevertheless, we cannot rule out that with more time to learn over-expenditure might have been smaller.

3.3 Pre-experimental Questionnaire

We used a pre-experimental questionnaire to elicit individual attributes that may explain variation in behavior.[16] Earlier literature as, for example, reviewed by Dechenaux et al. (2015) has shown that social preferences, risk aversion, loss aversion and non-monetary preferences for winning (*joy of winning*) can potentially explain individual heterogeneity in contests. To elicit social preferences, we used the SVO Slider Measure developed by Murphy et al. (2011).[17] The measure allows us to divide the participants into four classical social value orientation types, based on their choices in a sequence of six dictator games: *Altruists*, who maximize the payoff of the other, *prosocials*, who maximize joint payoff, *individualists*, who maximize their own payoff, and *competitors*, who maximize the difference between their own and the other's payoff. Elicitation of social value orientation identified no participant in the competitive category. Consistent with the findings of Murphy et al. (2011), the majority of the participants (60 out of 96) were prosocials. The remaining subjects were individualists.[18] We therefore only distinguish between prosocial and individualistic participants by using the dummy variable PROSOC to indicate prosociality. To assess risk aversion and loss aversion, we used lottery tasks similar to those used by Dohmen et al. (2011) and Gächter et al. (2010).[19] We attempted to capture heterogeneous preferences towards winning using the Revised Competitiveness Index (RCI) of Houston et al. (2002), which measures "a desire to win in interpersonal situations" (p. 31) based on questions about competition in daily life.[20]

[15] Charness et al. (2012) discusses the advantages and disadvantages of within-subject treatments.

[16] Subjects completed the questionnaire using the online tool Qualtrics at least eight days ahead of the laboratory experiment.

[17] Murphy et al. (2011) argue that the SVO Slider Measure is a reliable and valid method to elicit social value orientation. For further details, see Section 1.1 in the Web Appendix.

[18] There was also one altruist, which we added to the individualistic category.

[19] According to these measures 80% (90%) of the subjects had risk (loss) aversion.

[20] Harris and Houston (2010) argue that the index is reliable and correlates positively with other indices of competitiveness (Houston et al., 2002). Section 1.4 of the Web Appendix contains further details. We observe that the mean of RCI (45.7) is close to the mean reported by Houston et al. (2002) (48.5).

4 Experimental Results

We describe the main treatment effects in Section 4.1. Sections 4.2 and 4.3 compare the behavior of leaders and laggards in Period 2 and investigate the effect of performance differences. Section 4.4 shows how player characteristics influence behavior.

4.1 Main Treatment Effects

Figure 1 shows that mean efforts in Period 1(2) are 12% (7%) higher in ONE than in TWO. These differences are at least marginally significant.[21]

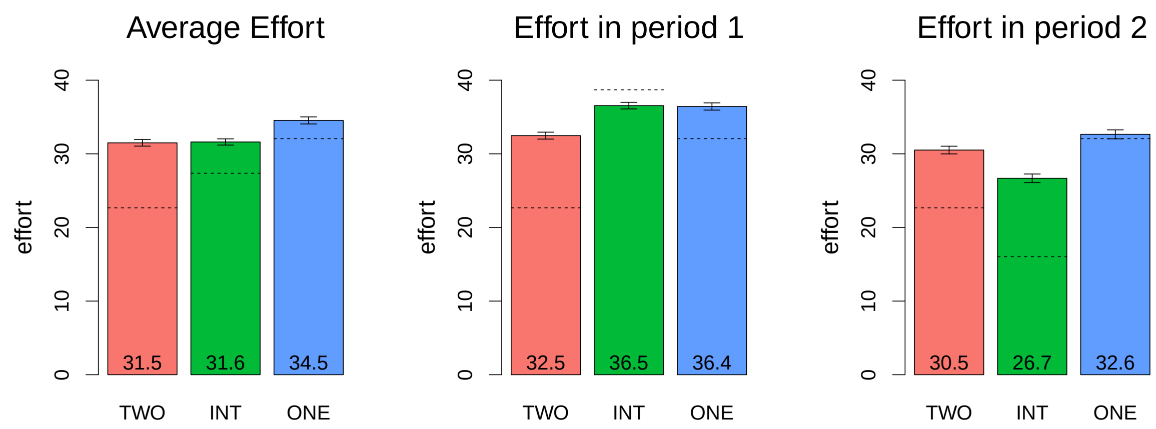


Figure 1: Means of efforts across participants. Heights of bars and values at bottom of bars correspond to means of efforts. Lengths of whiskers at top of bars are equal to standard errors of the means. Dotted lines depict Nash predictions. $N = 960$ per treatment. Sample: All participants.

The central design implication of the benchmark model is confirmed: Average efforts are higher with one large prize at the end (and equal weights) than with two prizes. However, the treatment effects are smaller than predicted. Observed average efforts under ONE are only 10% higher than under TWO rather than the predicted 41%, while the difference is significant.[22,23] We summarize our observations as follows:

[21] Observations are clustered within matching groups. We therefore use the signed-rank test proposed by Datta and Satten (2008). The unit of observation is the mean of a participant's (average or period-specific) effort across all 10 rounds within a treatment. The p-values are $p = 0.005$ and $p = 0.076$, respectively. We obtain similar results with Wilcoxon signed-rank tests based on matching-group means or Tobit estimations.

[22] The p-value is $p = 0.006$. When using only data from (S1) and (S3), second-period efforts no longer differ significantly between ONE and TWO. Nevertheless, the difference between mean efforts in the entire game becomes slightly higher (14%).

[23] When the efforts of different players are complementary, it may be more compelling to take the minimum of these efforts as an objective. The difference between the minimal efforts of the players is even smaller than the difference between average efforts (7%).

Result 1. *Efforts in each of the periods $t = 1, 2$, and thus average efforts, are higher in ONE than in TWO, but the difference between treatments is smaller than predicted.*

A careful look at Figure 1 suggests the following sharper statements; their significance is confirmed by Table A1 in the Appendix.

Result 2. *In ONE and TWO, average efforts are higher than predicted. The same is true for efforts in each period, except for the second-period efforts in ONE. In Period 1, there is less over-expenditure under ONE than under TWO.*

Hence, the positive effect of moving from TWO to ONE is so small because there is substantial over-expenditure in the former case and less over-expenditure in the latter.[24]

While we focus on the comparison between ONE and TWO, the comparison with INT is instructive. Average efforts in INT are similar to average efforts in TWO, even though they are almost 21% higher in equilibrium.[25] Thus, contrary to the prediction, the adjustment of weights in itself (TWO $\rightarrow$ INT) does not have a positive effect on efforts; only the combination of a weight and prize adjustment (TWO $\rightarrow$ ONE) increases average efforts. However, the adjustment of weights changes the distribution of efforts across periods.

4.2 Second-Period Efforts: Leaders vs. Laggards

We have identified two main deviations from equilibrium at the aggregate level. First, average efforts are higher than predicted. Second, the difference between observed and predicted efforts is higher for TWO than for ONE, in particular, in Period 2. The first observation is not surprising: Excessive effort is widespread in contests, which is usually attributed to joy of winning (see Dechenaux et al., 2015).[26] In the following, we therefore focus on explaining why there is more over-expenditure in TWO than in ONE.

Table 3 shows substantial excess spending in Period 1. In both treatments, the excess spending of leaders (players with positive relative performance Δs_{i1}) falls by about 40% from Period 1 to Period 2. The mean over-expenditure of laggards (subjects with $\Delta s_{i1} < 0$)) increases (by 1.7) in TWO, whereas it falls in ONE (by 2.0). Thus differences in the behavioral adjustments of leaders and laggards across treatments help to explain why over-expenditure is relatively high in TWO, but not in ONE.

[24] Similarly, when observations from (S2) are eliminated, there is substantial over-expenditure under TWO, but not under ONE. Also, note that the above results refer to all rounds. Figure A1 in the appendix shows that efforts under TWO are considerably lower in later rounds, in particular, in Period 1. This pattern is absent in ONE. Therefore, the difference in average efforts under ONE and TWO is closer to the equilibrium prediction in later rounds than when all rounds are considered. Nevertheless, Results 1 and 2 still apply in the final round.

[25] Mean efforts in Period 1 in INT are slightly below equilibrium, but above equilibrium in Period 2.

[26] However, over-expenditure is less common in rank-order tournaments (Dechenaux et al., 2015).

Table 3: Means of over-expenditure for leaders and laggards

Policy	TWO			ONE		
Over-expenditure	Period 1	Period 2	Change	Period 1	Period 2	Change
Leader	13.4	7.7	-5.7	8.1	4.9	-3.2
Laggard	6.2	7.9	1.7	0.6	-1.4	-2.0

Over-expenditure is calculated as observed effort minus predicted effort (conditional on first-period performance difference for second period in ONE).

Table 4 presents the results of OLS regressions of over-expenditure and the change in over-expenditure on dummy variables for leaders and laggards, controlling for first-period efforts (expressed as deviations from the mean). This procedure addresses the potential endogeneity of a subject's status as leader or laggard: With first-period effort as a control, the coefficients of the dummy variables capture only the variation caused by the random component in determining whether a player is leader or laggard (see the robustness discussion at the end of this section).[27]

Column (1) of Table 4 shows that in TWO a player with average first-period effort chooses 4.88 units of efforts more than predicted in Period 2 when ending up as a leader, whereas the over-expenditure is 10.91 when ending up as a laggard. Thus, even though the outcome of the first period has no effect on the expected marginal monetary value of effort in the second period and on the corresponding equilibrium, the chance component links the two periods by influencing who becomes leader or laggard. The over-expenditure of laggards is significantly higher than for leaders with the same first-period effort.

The massive over-expenditure of laggards in TWO disappears in ONE. Column (2) of Table 4 shows that the second-period behavior of average laggards in ONE is close to the equilibrium prediction. At the same time, leaders still engage in excess spending (by 2.72 units), although less than in TWO. Thus, the difference between leaders and laggards apparent in TWO is essentially reversed in ONE.

Focusing instead on the change of behavior between periods, Columns (3) and (4) of Table 4 show that a leader in TWO (ONE) who exerted average efforts in Period 1 reduces efforts by 4.61 (1.65). The behavior of laggards differs sharply: Whereas there is even an (insignificant) increase of over-expenditure in TWO between periods (+0.69), the decline

[27] Whether a participant finishes the first period as a leader or a laggard depends on both first-period effort choices, and the random shock. Her first-period effort choice is correlated with her second-period effort choice if it reflects persistent traits. Therefore, LEADER and LAGGARD, the explanatory variables of interest in Table 4, are endogenous. By including first-period effort as an explanatory variable (EFFORT 1*), the coefficients are identified only based on variation between leaders and laggards with the same first-period effort level. As it is outside of the control of participants with the same first-period effort level whether they finish the first period as leaders or laggards, the remaining variation between leaders and laggards is exogenous. Similar arguments hold for the following regressions.

Table 4: Effect of first-period outcome on second period

Model	(1)	(2)	(3)	(4)
Dep. var.	Over-exp. in 2		Change of over-exp.	
Policy	TWO	ONE	TWO	ONE
EFFORT 1*	0.86***	0.57***	-0.29***	-0.43***
	(0.000)	(0.000)	(0.001)	(0.000)
LEADER	4.88***	2.72**	-4.61***	-1.65*
	(0.000)	(0.015)	(0.000)	(0.082)
LAGGARD	10.91***	0.70	0.69	-3.67***
	(0.000)	(0.253)	(0.359)	(0.000)
N	960	960	960	960
Number of clusters	12	12	12	12
Adj. R^2		0.24	0.17	0.15
Log-likelihood	-3408.58			
Bootstrap samples	9999	9999	9999	9999

Tobit/ordinary least squares regressions. Dependent variable is calculated as observed minus predicted second-period effort (conditional on first-period performance difference for ONE) minus observed plus predicted first-period effort, or observed second-period effort minus predicted second-period effort (conditional on first-period performance difference for ONE). EFFORT 1* is calculated as deviation from mean effort under the corresponding policy. Bootstrapped p-values given in parentheses, computed using pairs cluster bootstrap-t (see Section 2 in the Web Appendix) with standard errors clustered on matching group. Sample: All participants. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

in over-expenditure in ONE is stronger than for leaders (-3.67). This is consistent with the observations in Table 3: As laggards increase their efforts relative to leaders between the two periods in TWO, the initial difference between the two groups essentially disappears, so that absolute second-period efforts of leaders and laggards are similar.

We summarize the main differences in the behavior of leaders and laggards as follows.

Result 3. *(i) In TWO, efforts of laggards in Period 2 are higher than those of leaders. (ii) In ONE, efforts of laggards in Period 2 are lower than those of leaders.*

To conclude, the similarity of average efforts under ONE and TWO reflects the two main observations above: Due to the low laggard effort in the former case and the high laggard effort in the latter case, the treatment effect is smaller than predicted.

Robustness By controlling for first-period efforts, the above identification strategy accounts for any endogeneity of first-period outcomes that may result from persistently different propensities to exert effort. To address further endogeneity concerns, we perform two robustness tests: (a) including participant-level fixed effects in the regression, and (b) using Propensity Score Matching (PSM) instead of a regression.

(a) Participant-level fixed effects further limit the identification of the coefficients to within-participant variation of being a leader or a laggard. Because fixed effects demean the data, we can only estimate the difference in over-expenditure between leaders and laggards, not the absolute levels of over-expenditure. Our estimates imply that leaders exert 4.2 units less (1.4 units more) over-expenditure than laggards in TWO (ONE), which is consistent with the results reported in columns (1) and (2) of Table 4.[28]

(b) As an alternative identification strategy, we use PSM to compare participants with the same likelihood of becoming a leader conditional on their first-period effort choice. Again, the results are consistent with those reported in Table 4, as we find that leaders exert 5.4 units less (2.0 units more) over-expenditure than laggards in TWO (ONE).

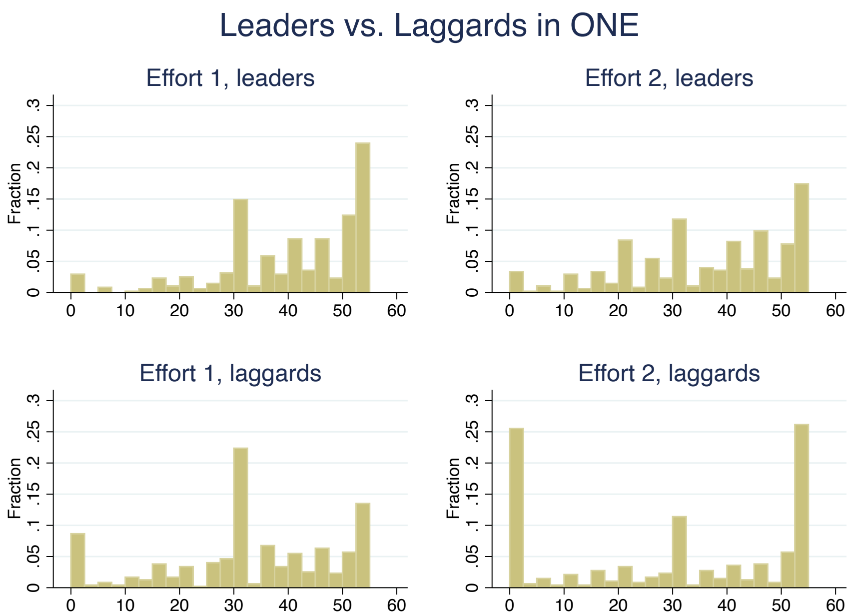
4.3 Performance Differences and Effort Heterogeneity

The aggregate behavior of laggards (and, to a lesser extent, leaders) in the second period of ONE conceals considerable heterogeneity. Figure 2 shows that many laggards (24%) choose zero efforts, whereas about the same fraction (23%) chooses the maximal effort. This result is remarkable as such polarization arises neither for leaders in Period 2 nor for either type of players in Period 1.[29] Moreover, note that, in ONE, more laggards choose maximal efforts than leaders (23% vs 16%). This suggests that, in a completely unconstrained setting, the intimidation effect might have been less pronounced, so that the average behavior of laggards would have been closer to the theoretical prediction of the benchmark model. In treatment TWO, a bias from the effort ceiling seems unlikely, as the fraction of leaders and laggards at the effort ceiling is about the same (around 10%).

The dispersion in first-period efforts suggests that the dispersion in laggard behavior revealed by Figure 2 partly reflects exogenous player heterogeneity. In addition, Figure 3 shows that heterogeneity reflects the size of performance differentials. In this figure, we split the sample in two halves - pairs with *weak* and *strong laggards*, depending on the difference to the leader. Around 40% of the weak laggards give up the race and exert zero effort, whereas slightly less than 20% exert maximal effort. Conversely, around 30% the strong laggards exert maximal efforts, whereas less than 10% give up completely. Thus,

[28]Including fixed effects only (without including first-period effort) does not address endogeneity of LEADER and LAGGARD, as the corresponding regressions would compare the same participants at different first-period effort levels. The effect of becoming a leader or a laggard is likely to differ even for the same participant depending on how much effort the participant has invested in the first period.

[29]Figure A2 in the Appendix contains histograms of first- and second-period efforts for the other two treatments. Extreme choices are very rare in TWO, but common in INT.



$N = 960$. Sample: All participants.

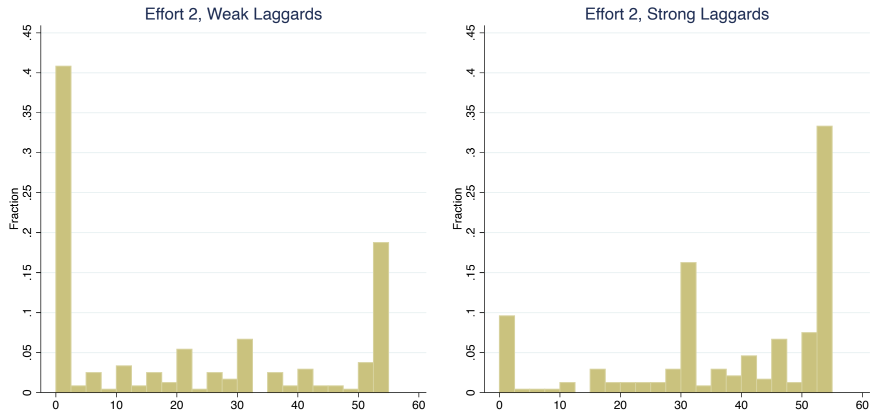
Figure 2: Distribution of efforts under ONE

polarization reflects differences between strong and weak laggards as well as heterogeneity within each type of laggard.

We therefore investigate the role of the first-period performance difference in more detail. In TWO, performance differences have no effect on the second-period equilibrium, as both periods are independent. Column (1) of Table 5 confirms this prediction. In ONE, the results are again very different. Column (2) shows that, as predicted, the *actual* effort in ONE falls for both players as the asymmetry increases. However, contrary to the prediction, this reduction is smaller for leaders than for laggards.[30] Digging deeper, Column (3) compares the behavior of marginal leaders and laggards in ONE for whom first-period performance was the same, so that the second period determines the outcome. For average first-period efforts and identical first-period performance ($\Delta s_{i1} = 0$), the second-period effort is 7.53 (6.59) units lower than in the Nash equilibrium for leaders (laggards). As the asymmetry increases, both leaders and laggards move towards over-expenditure, but the effect is stronger for leaders, as the respective interaction coefficients 0.30 and 0.22 show. This means that the adverse effect of increasing asymmetry on efforts is not as strong as predicted, but greater for laggards than for leaders. As a result, the observation for marginal leaders and laggards is reversed for *average leaders* and *average laggards* (for whom the performance difference takes average values): The average

[30]The coefficients are -0.27 and -0.43, respectively.

leader exerts 2.65 units more effort than in equilibrium; the corresponding (insignificant) coefficient for average laggards is 0.78 (Column (4)).



$N = 480$. Sample: Laggards with above (below) median first-period performance difference.

Figure 3: Distribution of second-period efforts under ONE, strong and weak laggards

4.4 Other Determinants of Behavior

We now briefly summarize our results about the effects of player characteristics.

4.4.1 Second-Period Efforts

Social Preferences Table A2 in the appendix investigates the role of social preferences on second-period behavior. The high over-expenditure of laggards in TWO is almost independent of whether they are prosocial or not (see Column (2)). Efforts of individualistic leaders are similar to those of individualistic laggards. Prosocial leaders, however, exert lower effort than individualistic leaders with the same first-period effort (see the weakly significant coefficient on the interaction term in Column (2)). In ONE, the interaction terms are not significant, casting doubt on a role of social preferences for the low efforts of laggards.[31,32]

Gender Effects Related to a substantial literature on gender effects in contests (see Section 6), Table A2 also shows regressions with a gender dummy. In treatment TWO, the result that efforts are higher for laggards than for leaders when controlling for first-period

[31]However, the coefficient of the interaction term for leaders in Column (5) is not far from significance, suggesting similarly the possibility that prosocial leaders exert less effort than individualistic leaders.

[32]Column (1) of Table A3 shows that, in INT, leaders and laggards behave similarly, whereas laggards exerted higher effort in TWO and less in ONE. By Column (2), this conclusion does not depend on prosociality. By Column (4), the effect of asymmetry on efforts lies between those in ONE and TWO. Finally, by Column (3), the gender effects are similar in INT and in TWO.

efforts is much stronger for females (see Column (3)). Not only do female laggards exert significantly more efforts than male laggards; in addition, female leaders exert less efforts than male leaders (though the difference is not significant). In ONE, however, there is no clear gender effect (see Column (6)).[33]

4.4.2 First-Period Efforts

Recall that in Period 1 there is more over-expenditure in TWO than in ONE, though the difference is smaller than in Period 2. We therefore now summarize how exogenous player characteristics and the previous experience of players affect first-period behavior, using explanatory variables such as player characteristics, the number of rounds and measures of previous exogenous performance shocks (see Table A4 in the appendix).[34]

Result 4. *Consider Period 1 of treatments ONE and TWO.*
(i) The mean effort of prosocial types is lower than for individualists; this effect is significant under TWO.
(ii) Loss aversion reduces efforts, but the effect is not significant.
(iii) There are no clear effects of competitiveness, risk aversion and past luck on efforts.

The results for prosociality and loss aversion are consistent with previous findings from lottery contests.[35] The low efforts of prosocial types are in line with their preference for joint profit maximization.[36,37] We note in passing that the results for INT are similar to the other treatments in that loss aversion and prosociality have negative signs – however, none of the coefficients is significant. While one might expect that previous luck would lower efforts, the only significant results we obtain are for INT.

[33] Moreover, we do not find any gender difference in the effects of the size of the performance asymmetry on behavior (regression output available upon request).

[34] These measures are (1) the sum of the first- and second-period effort difference (relative luck) in the previous round, (2) the share of occurrences of favorable (positive) observation error differences in the current treatments and (3) in the whole experiment.

[35] See Kong (2008), Shupp et al. (2013) and Hernandez-Lagos et al. (2017), respectively.

[36] The literature does not provide a clear hypothesis on the relationship between social value orientations and efforts. The SVO Slider Measure is difficult to compare to the social preference measures used in previous contest experiments. (see, e.g. Balafoutas et al. (2012))

[37] Previous research suggests that efforts should be decreasing in risk aversion (Millner and Pratt, 1991; Anderson and Freeborn, 2010; Sheremeta and Zhang, 2010; Price and Sheremeta, 2011, 2015; Sheremeta, 2011; Sheremeta et al., 2017) and loss aversion (Kong, 2008; Shupp et al., 2013) and increasing in the preference towards winning (Sheremeta, 2010; Price and Sheremeta, 2011, 2015; Sheremeta et al., 2017; Brookins and Ryvkin, 2014).

Table 5: Effect of first-period asymmetry on effort and over-expenditure in Period 2

Model	(1)	(2)	(3)	(4)
Dep. var.	Effort 2		Over-exp. eff. 2	
Policy	TWO	ONE	ONE	ONE
EFFORT 1*	0.86***	0.85***	0.59***	0.59***
	(0.000)	(0.000)	(0.000)	(0.000)
LEADER	28.44***	43.88***	-7.53***	2.65**
	(0.000)	(0.000)	(0.000)	(0.030)
LAGGARD	33.34***	47.29***	-6.59***	0.78
	(0.000)	(0.000)	(0.001)	(0.222)
$\|\Delta s_{i1}\| \cdot$ LEADER	-0.03	-0.27***	0.30***	
	(0.133)	(0.000)	(0.000)	
$\|\Delta s_{i1}\|^* \cdot$ LEADER				0.30***
				(0.000)
$\|\Delta s_{i1}\| \cdot$ LAGGARD	0.01	-0.43***	0.22***	
	(0.736)	(0.000)	(0.000)	
$\|\Delta s_{i1}\|^* \cdot$ LAGGARD				0.22***
				(0.000)
N	960	960	960	960
Number of clusters	12	12	12	12
Adj. R^2			0.37	0.37
Log-likelihood	-3408.06	-3144.02		
Bootstrap samples	9999	9999	9999	9999

Tobit/ordinary least squares regressions. Over-expenditure in second-period effort is calculated as observed second-period effort minus predicted second-period effort (conditional on first-period performance difference). EFFORT 1* is calculated as deviation from mean effort under the corresponding policy. $|\Delta s_{i1}|^*$ is calculated as deviation from mean $|\Delta s_{i1}|$ under the corresponding policy. Bootstrapped p-values given in parentheses, computed using pairs cluster bootstrap-t (see Section 2 in the Web Appendix) with standard errors clustered on matching group. Sample: All participants. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

5 Towards an Explanation of the Observations

The following observations made above deviate from the predictions of the benchmark model.

1. Except for laggards in Period 2 of ONE, effort is higher than predicted.
2. In TWO, average laggards exert higher effort than average leaders.
3. In ONE, laggards often choose extreme effort, with a greater share of weak laggards at the minimum and a greater share of strong laggards at the maximum.
4. In ONE, average laggards exert less effort than average leaders.

5.1 A Simple Framework

We now introduce a framework to explain these observations (and similar results in INT; we discuss the latter in subsection 8.2 in the appendix). The framework allows for non-monetary payoffs and more general assessments of winning probabilities than the benchmark model.

Assumption 1: *Player i's total payoff from winning W is* $\Pi_i = W + V_i$*, where*
(i) V_i *is positive and can vary across players;*
(ii) for fixed i and $|\Delta s_{i1}|$*,* V_i *is higher if i is a laggard than if i is a leader.*

The assumption allows for joy of winning while remaining agnostic about its exact source. The feature (ii) that, in an otherwise identical situation, joy of winning is higher for a laggard than for a leader could reflect social preferences such as inequity aversion (Fehr and Schmidt, 1999) or reciprocity (Dufwenberg and Kirchsteiger, 2004; Falk and Fischbacher, 2006).[38] However, even without social preferences, higher joy of winning for a laggard than for a leader could result from decreasing marginal benefits from obtaining a prize (the second prize won is less valuable than the first one).[39]

Next, in view of the complexity of the game, we relax the assumption made in Section 2 that the players assess winning probabilities based on equilibrium considerations and distributional assumptions. Not only may players have difficulties dealing with normal distributions, but they may also be unsure about the behavior of their opponent.

[38] In TWO, a laggard has non-positive payoffs from Period 1, whereas those of the leader are positive; inequity aversion thus suggests a willingness to pay to reduce future payoffs of the opponent. In both treatments, being a laggard suggests high opponent efforts in Period 1; reciprocity would correspond to reacting to this unkind behavior with unkind behavior (high second-period efforts).

[39] One might assume further that joy of winning depends on the size of the lag, reflecting social preferences or greater pleasure from obtaining a favorable outcome in a more difficult situation. We did not impose this additional restriction as we do not require it for our main arguments.

Assumption 2: (i) *In Period 2, each player i has an increasing belief function $p(e_{i2})$, which can depend on his identity, the treatment and the history of play.*
(ii) *For any e_{i2}, increases in W and reductions in Δs_{i1} shift the belief function down (reduce p weakly).*

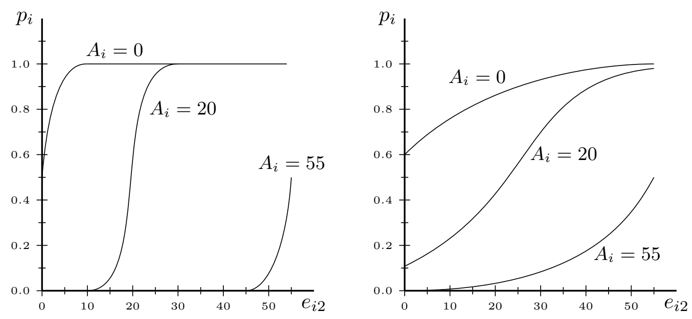
This assumption maintains the idea that higher efforts increase winning chances. An example is the *equilibrium belief function.* This function is generated by assuming that the opponent chooses the second-period equilibrium and that errors are distributed as in Section 2. It does not depend on the player's identity. Furthermore, in treatment TWO, it is independent of history: Player i believes he wins with the probability of the event that $\Delta\epsilon_{j2} \leq e_{i2} - e^*$, where $e^* = e_2^*$ is the symmetric time and history independent equilibrium effort in the one-period game; thus the belief function is the corresponding cumulative (normal) distribution function of the error term. In treatment ONE however, the equilibrium belief function depends on Δs_{i1}: It is the given by the probability that $\Delta\epsilon_{j2} \leq e_{i2} + \Delta s_{i1} - e_2^*(\Delta s_{i1})$, where $e_2^*(\Delta s_{i1})$ is the symmetric equilibrium effort given as in (1).

Crucially though, Assumption 2 provides additional flexibility. The probability assessment can be subjective, allowing for belief heterogeneity. Moreover, the effects of the history of the game can differ from those under equilibrium beliefs, for instance, because players partly attribute Δs_{i1} to hidden characteristics of the opponent. Given own past effort e_{i1}, having a low Δs_{i1} (and, in particular, being a laggard) is a signal of high past opponent effort and thus high opponent joy of winning. If player i believes that joy of winning is a persistent trait across periods, he rationally expects that winning chances are increasing in Δs_{i1}. This is because a low Δs_{i1} makes high second-period efforts of the opponent more likely, shifting $p(e_{i2})$ downwards. This effect could be present in both treatments and should not be confused with the equilibrium effect that greater asymmetry reduces the winning chances of laggards; it arises in addition to the standard equilibrium effect. For leaders, the reasoning implies that a greater positive performance differential should increase winning chances (even beyond the standard equilibrium logic that is present in ONE).

To explain Observations 3 and 4, we will specify the belief function further, assuming that $e_{i2} \in [0, 55]$ as in the experiment. We postulate the logistic form

$$p_i(e_{i2}) = \frac{1}{1 + \exp(-r_i(e_{i2} - A_i))}. \tag{4}$$

Here $r_i > 0$ and A_i are constants that characterize a player's subjective assessment of winning chances given the treatment, the size of the prize and the history of the game. Before linking the shape of the belief function more directly to our assumptions and using



(a) $r_i = 0.5$

(b) $r_i = 0.1$

Figure 4: Belief Functions

it to explain our observations, we show in Figure 4 how the parameters affect the resulting belief functions. We focus on $r_i = 0.5$ and $r_i = 0.1$. In the former case, the player has strong opinions about which efforts are required to win, with p_i close to 0 or close to 1 for most effort levels. In the latter case, the player is less certain about the relation between efforts and winning probability.[40]

When explaining second-period observations below, we will emphasize the role of A_i, which we think of as capturing a player's perception of winning chances. For fixed r_i, an increase in A_i reduces p_i, reflecting growing pessimism. In line with Assumption 2, A_i can have various determinants: Clearly, even with equilibrium beliefs it should be higher for a player in ONE who lags further behind, reflecting lower winning chances. In addition, A_i will take up individual heterogeneity in pessimism. Moreover, as we lay out below, it can reflect differences in information about the other player's likely behavior.

5.2 Explaining Second-Period Observations

We now explain Observations 1 to 4, assuming that players maximize expected net payoffs $\pi_i(e_{i2}) = p_i\,(e_{i2})\,\Pi_i - K(e_{i2})$.

Observation 1 can be explained by joy of winning (Assumption 1(i)). All else equal, greater joy of winning increases $\pi_i'(e_{i2})$, the marginal effect of effort on expected net payoff, fostering higher efforts. Thus, for fixed beliefs, the optimal second-period effort is higher than with pure monetary preferences.[41]

Observation 2 states that the average efforts of laggards in TWO are higher than those of leaders. The reasoning is similar as for Observation 1: By Assumption 1(ii), joy

[40] The value of r_i may capture both how confident players are about the other player and to which extent they are aware of the noise in the performance measure.

[41] A caveat is that players should also expect that the opponent increases efforts because of joy of winning. Therefore, they should become more pessimistic about their winning chances, potentially reducing the direct positive effect of joy of winning on efforts (see the discussion below).

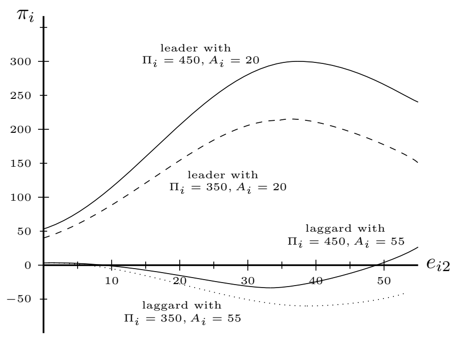


Figure 5: Expected leader and laggard payoffs for different degrees of optimism and joy of winning

of winning is higher for laggards than for leaders, which works in favor of the observation. In line with the motivation of this assumption, we therefore attribute the observation to a *revenge effect*:[42] Laggards exert higher efforts than leaders because they want to win at least once – a motivation that is absent for a leader who is already partly saturated.

To discuss the remaining two observations, which concern treatment ONE, we use a parametric specification. As in the experiment, we set $K(e_{i2}) = k(e_{i2})^2$ with $k = 0.066$. We suppose that $p(e_{i2})$ has the logistic shape (4). To capture a situation with substantial subjective uncertainty, we fix $r_i = 0.1$. To capture heterogeneity in pessimism, we will somewhat arbitrarily set $A_i = 20$ or $A_i = 55$. Such differences could reflect exogenous heterogeneity in players' assessments. However, we focus on the interpretation that $A_i = 20$ corresponds to a situation where Δs_{i1} is relatively high (e.g., for a leader), while $A_i = 55$ corresponds to a situation where it is low (e.g., for a laggard facing this leader). Finally, we allow Π_i to take values 350 and 450, reflecting differences in joy of winning (for the given monetary price $W = 300$).

Figure 5 illustrates how Observation 3 can be explained. As argued above, higher joy of winning increases the effort of both players. For the laggard, this effect can be much larger than for the leader, potentially causing discontinuities in the optimal effort choice. Intuitively, the laggard reasons that very high efforts will be necessary to have a winning chance. This is not worthwhile with low joy of winning, and he gives up. In this case, where being a laggard makes a player so pessimistic that the optimal effort falls close to zero, we speak of an *intimidation effect*. By contrast, the laggard with high joy of winning chooses the highest possible effort, reasoning that very high efforts will be necessary to have a winning chance. Unlike the laggard with low joy of winning, he deems this effort

[42]Again, the evidence suggests that the joy-of-winning effect dominates any potential countervailing effect.

worthwhile – pessimism then induces a *wake-up effect.* The two lower lines in Figure 5 illustrate these two possibilities.[43]

Observation 4 is a statement on average behavior. We have argued that heterogeneity in joy of winning can explain individual differences in behavior, with a pronounced tendency for laggards to choose boundary solutions. The fact that average efforts of laggards are lower than those of leaders in ONE suggests that the intimidation effect on laggards with low joy of winning prevails over the wake-up effect on those with high joy of winning, leading to an aggregate intimidation effect.[44]

To sum up, we attribute the asymmetries in the behavior of leaders and laggards in Period 2 to the interplay of non-monetary payoffs and non-equilibrium beliefs with player heterogeneity. We postulate that joy of winning differs across players and is strongest when players have previously lost – a revenge effect.[45] Moreover, we argue that players take the history of play as informative of the opponent's preferences and thus of his likelihood to choose high efforts. This exacerbates the negative effect of a performance disadvantage on perceived winning chances and leads to an intimidation effect in the aggregate. While the effects are potentially present in both treatments, joy of winning dominates only in TWO where laggards exert higher efforts than leaders. Intuitively, joy of winning is likely to be higher for a player who has previously lost; moreover being a laggard in treatment TWO does not directly reduce winning chances. By contrast, in treatment ONE, the role of the performance differential on perceived winning chances is particularly salient: Many players react to the pessimism resulting from a performance gap by either giving up completely or by exerting maximal effort to keep their winning chances alive.

In Subsection 8.3 in the Appendix, we strengthen the above arguments by showing the consistency of the behavior of players in the two treatments. We find that even for otherwise identical players who exerted the same first-period effort, the player in the role of a laggard will exert higher effort than a leader, whereas the converse statement applies in ONE.

[43]While our informational argument means that being a laggard should make a player more pessimistic even in TWO, the effect should be smaller than in ONE where it arises on top of the direct negative effect on winning chances. This explains why extreme choices are much less common in TWO.

[44]Assuming that an increase in $|s_{i1}|$ results in higher pessimism, the above arguments not only help to understand the differences in the behavior of leaders and laggards in ONE, they also help to explain why greater performance differentials lead to more polarization and to more pronounced effort differences between leaders and laggards (See the comparison between strong and weak laggards in Figure 3).

[45]We should, however, point out that there appears to be no statistically significant relation between the RCI and behavior in Periods 1 and 2. This could reflect the fact that the RCI asks abstract questions about whether individuals like competition, without directly addressing whether they like to win if forced to compete (see Table 6 in the Web Appendix).

5.3 Remarks on First-Period Behavior

Next, we sketch how we think about behavior in Period 1. Reflecting the independence of the two contests, we assume that players in TWO maximize expected first-period payoffs. Adapting Assumption 1(i), joy of winning can explain excess efforts. In ONE, there is no first-period prize, but first-period efforts improve the player's position in Period 2: Higher effort today increases future winning chances and therefore pushes up the belief function tomorrow. A joy-of-winning component increases the benefits from increasing the probability of winning tomorrow, thus explaining above-equilibrium first-period efforts in ONE.[46] A more subtle question is why excess spending in Period 1 is small in ONE (and in INT) compared to TWO. One possible reason could be that it is hard to understand the rather complex positive effect of higher efforts today on the prospects of winning tomorrow. A plausible conjecture would be that players are *strategically myopic*, meaning that they underestimate the beneficial effect of today's effort on strategic interaction tomorrow (the upward push in the chances of success) relative to the costs which are more directly visible and therefore more salient. As a result of such strategic myopia, excess efforts are small in spite of joy of winning.

5.4 Discussion

The preceding analysis leaves several questions for future research.

First, to explain the differences between laggards and leaders in ONE, we argued that players would infer from bad news that their opponent has high joy of winning, making them more pessimistic about their future winning chances. With the current design, we cannot rule out alternative explanations. For instance, suppose that subjects follow the equilibrium logic in their probability assessment, understanding that, according to (1), they should exert more effort when $f(\eta \Delta s_{i1})$ is higher. However suppose that – lacking a correct understanding of the normal distribution – subjects do not understand that the density f of this term is single-peaked at zero and instead think that it is strictly increasing. In this situation, which may well be in line with Assumptions 1 and 2, (1) implies that leaders will exert higher efforts than laggards in ONE, thereby providing an alternative explanation of the observed behavior. To limit the scope for such confusion, one could simplify the nature of the noise in the performance structure, for instance by replacing the normal distribution with a discrete distribution where the performance difference can be biased by the same amount in each direction with probability 1/2. In such a setting, confusion would seem less likely. If the excess spending of laggards is

[46]Again, players can differ with respect to joy of winning as well as the assessment of winning chances.

still lower than for leaders, this would lend additional credibility to the informational argument provided in Section 5.2.

Second, this informational argument assumes that players make inferences about the opponent's effort in Period 2 from behavior in Period 1. To validate this argument, one could consider a setting where such inferences are unlikely to be relevant. For instance, suppose some players are assigned to fresh opponents in Period 2 which they have not previously interacted with (while maintaining the performance differential corresponding to the first period), whereas other players play against their first-period opponents. If there is asymmetry in the second-period behavior of leaders and laggards in the latter case but not in the former, it would strengthen the informational argument.

Third, to disentangle the role of social preferences from other sources of joy of winning, one could replace the contest with a single-agent decision problem where each subject exerts an effort in each period which is measured with noise. Subjects would either get one prize at the end of each period where performance in that period is above a threshold, or one prize at the end if the sum of performance levels is above a threshold. If excess spending in Period 2 is still present for subjects who did not win in the previous period in the case with two prizes, it would suggest that the pronounced excess spending of laggards in TWO may not just reflect social preferences, but a more general relation between previous success and joy of winning.[47]

6 Relation to the Literature

The theoretical literature on multi-stage contests is large. Even the more specific design issues concerning prize structure and the role of past performance have received considerable attention.[48] The experimental literature on contests is vast as well.[49] We therefore focus on experiments that relate closely to the design of two-stage contests and to the observations in our own experiment.

Comparison of Multi-Stage Contests Schmitt et al. (2004) compare repeated Tullock contests with multi-stage contests where efforts in one period increase winning

[47]Similarly, if subjects whose effort is low after bad first-period performance in a treatment with just one prize at the end exert lower effort than those whose effort is high, then inferences about the opponent are not likely to be the only reason behind the low efforts of laggard in treatment ONE.

[48]Apart from Klein and Schmutzler (2017), Möller (2012), Clark et al. (2012), Clark and Nilssen (2013) and Kubitz (2020) discuss prize structure; Meyer (1992), Harbaugh and Ridlon (2011), Ridlon and Shin (2013), and Denter and Sisak (2015, 2016) address the effects of incorporating past performance.

[49]Dechenaux et al. (2015) provide a comprehensive survey. Most closely related, Harbring and Irlenbusch (2003), Orrison et al. (2004) and Chen et al. (2011) analyze the prize structure in static contests; Delfgaauw et al. (2015) and Stracke et al. (2014) do this for elimination contests.

chances in future periods, resembling our treatments TWO and INT. Contrary to our rank-order setting, total equilibrium efforts are the same in both cases. Somewhat reminiscent of our results, the authors nevertheless observe a relative effort reduction in the latter case with effort carry-over.[50]

Laggard Behavior in Multi-Stage Games In our rank-order contest, laggards exert lower second-period efforts than leaders, even though there is no difference in equilibrium. By contrast, Konrad (2012) discusses models in which laggards' equilibrium efforts are lower (a *discouragement effect*), such as Harris and Vickers (1985), where players repeatedly carry out tournaments, and the first contestant who has won sufficiently often gets the final prize.[51] Whereas Zizzo (2002) does not find support for the theoretical results in such races, Mago et al. (2013) show that leaders exert more effort than laggards. Contrary to these authors who consider Tullock contests, Ederer and Fehr (2017) analyze a rank-order contest (like our treatment ONE). Their paper focuses on very different questions than ours, in particular, on the effects of dishonest feedback by principals.[52]

In real-effort experiments, without a clear theoretical benchmark, the distinction between discouragement and intimidation effects becomes blurred. In Eriksson et al. (2009), even laggards with small winning chances exert effort; in Casas-Arce and Martínez-Jerez (2009) agents only reduce effort levels when the distance to the leader is high. Berger and Pope (2011) and Goldman and Rao (2017) show that basketball teams who are slightly behind their opponents at half time exert more effort and have higher winning chances (similar to our marginal laggards in ONE). However, Gill and Prowse (2014) show that, in a repeated real-effort task, female subjects reduce effort after previous losses. This resembles our intimidation effect with the twist that it arises even though, as in our treatment TWO, there is no negative effect of past performance on current winning chances.[53] On a related note, Buser (2016) shows how success (luck) in one real-effort contest influences the willingness to engage in competition thereafter even when previous success is uninformative about future chances of success. While the aggregate effect of losing on the willingness to seek challenges is zero, this masks a positive effect on men and a negative

[50] Tong and Leung (2002) compare repeated rank-order tournaments with dynamic tournaments where past performance affects future winning chances. However, in particular in the latter case, the analysis is not clearly related to an equilibrium prediction for the underlying dynamic game.

[51] Similar discouragement effects arise in elimination tournaments.

[52] As in our model, performance asymmetry reduces second-period efforts. The behavior of laggards seems consistent with our intimidation effect. The paper does not consider independent contests (TWO).

[53] However, in a real effort task, past effort may affect future effort costs.

effect on women.[54] In our lab experiment, there is no intimidation effect for females. In ONE, however, we find a (gender-independent) intimidation effect.

Discouragement in Asymmetric Static Contests A complementary literature deals with asymmetric static contests. March and Sahm (2017) analyze static Tullock contests where some players have an exogenous handicap. Contrary to their equilibrium prediction, only laggards reduce their efforts as the asymmetry increases. Llorente-Saguer et al. (2016) find a discouragement effect in the all-pay auction setting of Baye et al. (1996). Müller and Schotter (2010) consider static all-pay auctions with asymmetric abilities. Depending on their costs, individuals either "drop out" or "become workaholics". This is related to the polarization among laggards in Period 2 of our treatment ONE.[55]

Behavioral Contest Design The literature has dealt with behavioral aspects of contest design theoretically and empirically, mainly focusing on static contests. Confirming standard theory, Sheremeta (2011) shows experimentally that efforts are higher with a grand contest than with multiple prizes and contests with two subcontests. Lim (2010) observes that social comparisons may lead to higher effort when there is a higher proportion of winners than losers. Mermer (2017) shows theoretically that, with expectation-based preferences as in Kőszegi and Rabin (2006, 2007), a designer may want to use multiple prizes. These last two results are in line with our result that models with two prizes perform better than expected, except that they refer to a static setting.

7 Conclusion

This study asks whether efforts should be rewarded with frequent small prizes or whether a designer should give infrequent large rewards, which are based on a longer performance history. A grand prize incentivizes efforts in every previous period, but softens competition later on. While the former incentive effect dominates over the latter in our simple model, one might wonder about additional motivational effects. With several small prizes, revenge motives may lead losers in early periods to compete aggressively to make up for past losses; with a single large prize, laggards may believe that they are facing a tough opponent, so that they might regard it as pointless to exert high effort with a large performance lag.

[54] In a related real effort experiment by Azmat and Iriberri (2016) information on relative performance in a piece-rate setting affects effort (positively), even though it should have no effect. However, contrary to our observations, it does not matter whether the information is positive or negative.

[55] More broadly, our findings are related to some recent observations on peer effects. Feld and Zölitz (2017) have found that the grades of students with low ability (GPA) deteriorate in the presence of high ability peers in the classroom, which appears similar to the intimidation effect in our experiment.

These motivational effects both appear to matter. Even though overall efforts are higher with a grand prize than with two prizes, the motivational effects tend to reduce the difference. These observations relate to the agents' reaction to the first-period relative performance: The low second-period effort in the single-prize treatment reflects an intimidation effect for laggards, whereas the high effort in the two-prize contest corresponds to a revenge effect. However, we observe considerable heterogeneity in behavior. Laggards who are far behind the leader choose minimal efforts disproportionately often, whereas laggards who are relatively close choose maximal efforts disproportionately often (though in each case, a non-negligible part chooses the other extreme).

The analysis suggests that the frequency of prizes and past performance weights affect efforts. Frequent small prizes are advisable if efforts are complementary, so that laggards' behavior matters. More tentatively, our analysis suggests giving infrequent large prizes, but putting more emphasis on recent efforts (to avoid intimidation). Obviously, a direct test of this claim would require further treatments. More broadly, our analysis also begs the question to which extent our insights are specific to contests. Would they also survive for more general forms of relative performance evaluation? Would they even be relevant in settings with individual efforts, in particular, for real-effort tasks? The answer to these questions is not obvious, but we believe it would be worthwhile to study intertemporal reward structures in organizations more broadly along the lines suggested by our work.

Our analysis potentially has wider implications as a microeconomic explanation of the emergence of economic inequality. If we think of agents as interacting during their education and their professional career in a sequence of contests, our experiment highlights the necessity of maintaining a level playing field, so as to avoid the intimidation effect, while at the same time harvesting the benefits of the revenge effect. More broadly, our study contributes to a discussion on the microfoundations of inequality. Real-world incentive systems, in particular, in labor markets, take past performance into account to some extent. Theoretical models such as the one underlying our experiment show that this may result in low efforts after large performance differences. The implications of our laboratory experiment are potentially more troubling: Because of the intimidation effect, initial performance differences could become self-perpetuating, as many laggards become less motivated and essentially give up competition. There are two conceivable policy conclusions. First, incentive systems should not rely too heavily on the distant past. Second, the heterogeneity of responses to first-period losses suggests that interventions enabling agents to cope with potentially upsetting experiences might be helpful.

8 Appendix

8.1 Figures and Tables

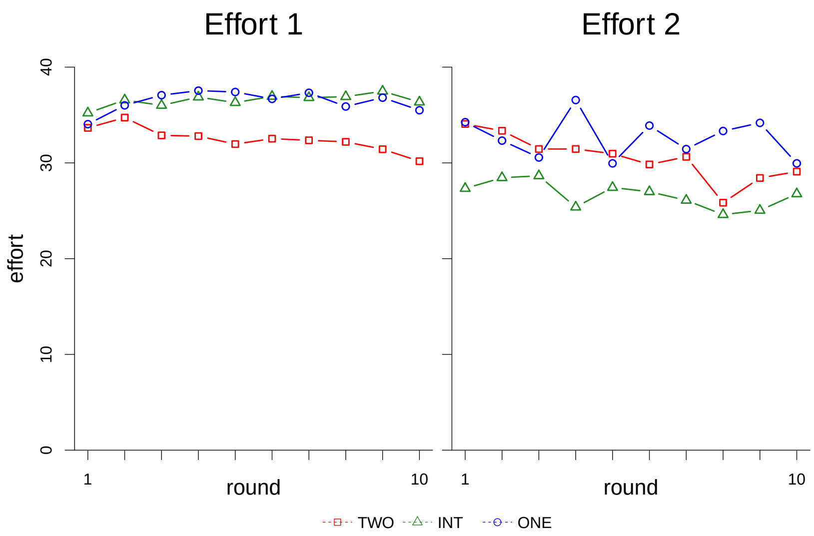
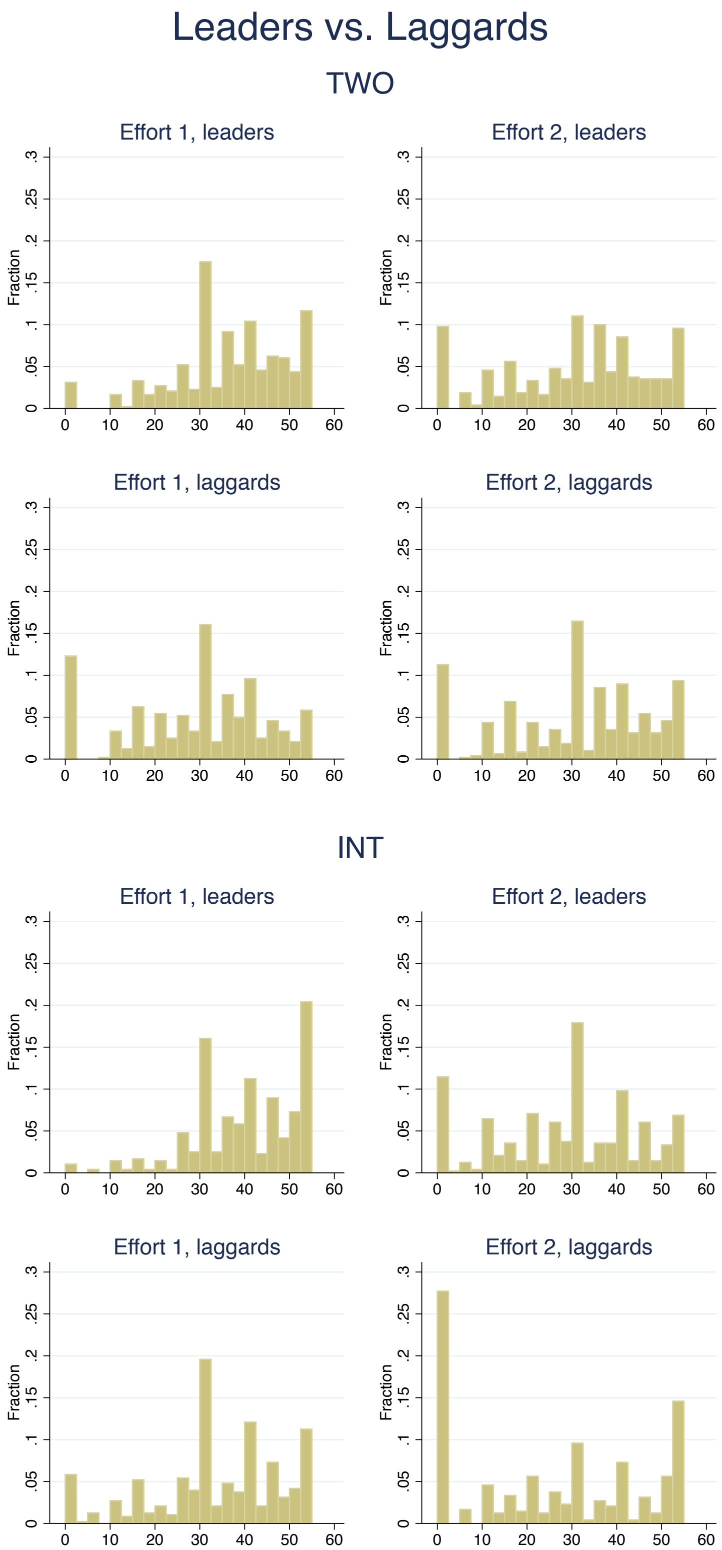


Figure A1: Means of efforts over rounds. $N = 96$ per round and policy. Sample: All participants.



$N = 960$. Sample: All participants.

Figure A2: Distribution of efforts under TWO and INT

	TWO			INT			ONE		
	Pred.	Mean	P-val.	Pred.	Mean	P-val.	Pred.	Mean	P-val.
Average effort	22.7	31.5	0.002	27.4	31.6	0.010	32.1	34.5	0.050
Effort 1	22.7	32.5	0.002	38.7	36.5	0.381	32.1	36.4	0.015
Effort 2	22.7	30.5	0.002	16.0	26.7	0.001	32.1	32.6	0.507
Effort 2 $\|\Delta s_{i1}$	–	–	–	15.2	26.7	0.001	30.9	32.6	0.117

Table A1: p-values for level predictions. Signed-rank tests according to Datta and Satten (2008) based on each participant's mean effort under the corresponding policy ($N = 96$). Two-sided hypotheses. Sample: All participants.

Table A2: Effect of first-period outcome on over-expenditure in second-period effort

Model	(1)	(2)	(3)	(4)	(5)	(6)
Dep. var.	Over-expenditure in effort 2					
Policy	TWO	TWO	TWO	ONE	ONE	ONE
EFFORT 1*	0.86***	0.85***	0.85***	0.57***	0.56***	0.57***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
LEADER	4.88***	8.02***	5.92***	2.72**	4.83**	2.91*
	(0.000)	(0.002)	(0.002)	(0.015)	(0.021)	(0.052)
LEADER · PROSOC		-4.90*			-3.35	
		(0.062)			(0.113)	
LEADER · FEMALE			-2.86			-0.53
			(0.231)			(0.753)
LAGGARD	10.91***	10.01***	9.64***	0.70	1.32	0.49
	(0.000)	(0.001)	(0.000)	(0.253)	(0.356)	(0.591)
LAGGARD · PROSOC		1.37			-1.01	
		(0.469)			(0.555)	
LAGGARD · FEMALE			3.27**			0.56
			(0.036)			(0.729)
N	960	960	960	960	960	960
Number of clusters	12	12	12	12	12	12
Adj. R^2				0.24	0.24	0.24
Log-likelihood	-3408.58	-3402.35	-3404.02			
Bootstrap samples	9999	9999	9999	9999	9999	9999

Tobit/ordinary least squares regressions. Dependent variable is calculated as observed second-period effort minus predicted second-period effort (conditional on first-period performance difference for ONE). EFFORT 1* is calculated as deviation from mean effort under the corresponding policy. Bootstrapped p-values given in parentheses, computed using pairs cluster bootstrap-t (see Section 2 in the Web Appendix) with standard errors clustered on matching group. Sample: All participants. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table A3: Effect of first-period outcome (first-period asymmetry) on over-expenditure in second-period effort (on second-period effort) in INT

Model	(1)	(2)	(3)	(4)
Dep. var.	Over-expenditure in effort 2			Effort 2
Policy	INT	INT	INT	INT
EFFORT 1*	0.45***	0.45***	0.47***	0.57***
	(0.000)	(0.000)	(0.000)	(0.000)
LEADER	11.38***	12.98***	10.02***	31.25***
	(0.000)	(0.000)	(0.000)	(0.000)
LEADER · PROSOC		-2.58		
		(0.176)		
LEADER · FEMALE			3.60	
			(0.151)	
LAGGARD	11.57***	12.37***	10.10***	35.87***
	(0.000)	(0.001)	(0.000)	(0.000)
LAGGARD · PROSOC		-1.28		
		(0.586)		
LAGGARD · FEMALE			4.17*	
			(0.055)	
$\|\Delta s_{i1}\| \cdot$ LEADER				-0.16***
				(0.000)
$\|\Delta s_{i1}\| \cdot$ LAGGARD				-0.30***
				(0.000)
N	960	960	960	960
Number of clusters	12	12	12	12
Adj. R^2	0.38	0.38	0.39	
Log-likelihood				-3409.82
Bootstrap samples	9999	9999	9999	9999

Tobit/ordinary least squares regressions. Over-expenditure in second-period effort is calculated as observed second-period effort minus predicted second-period effort (conditional on first-period performance difference for INT). EFFORT 1* is calculated as deviation from mean effort under the corresponding policy. Bootstrapped p-values given in parentheses, computed using pairs cluster bootstrap-t (see Section 2 in the Web Appendix) with standard errors clustered on matching group. Sample: All participants. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table A4: Explanations of over-expenditure in first-period effort

Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dep. var.	Over-expenditure in effort 1								
Policy	TWO	TWO	TWO	INT	INT	INT	ONE	ONE	ONE
CONSTANT	25.35**	24.57*	26.28*	4.72	14.06	13.93	-6.54	-6.05	-3.28
	(0.042)	(0.066)	(0.075)	(0.718)	(0.219)	(0.341)	(0.647)	(0.691)	(0.854)
PROSOC	-3.28*	-3.25**	-3.32*	-2.16	-2.05	-2.30	-3.18	-2.89	-3.25
	(0.067)	(0.032)	(0.070)	(0.432)	(0.449)	(0.416)	(0.185)	(0.234)	(0.187)
SPRISK	0.02	-0.30	0.01	2.73	3.56**	2.91	2.42	2.79	2.52
	(0.991)	(0.832)	(0.993)	(0.169)	(0.029)	(0.130)	(0.267)	(0.228)	(0.249)
λ	-3.65	-3.63	-3.65	-2.34	-2.13	-2.58	-3.30	-3.24	-3.44
	(0.114)	(0.138)	(0.112)	(0.335)	(0.323)	(0.311)	(0.227)	(0.230)	(0.217)
RCI	-0.09	-0.08	-0.09	-0.23	-0.29	-0.26	0.23	0.27	0.23
	(0.493)	(0.583)	(0.479)	(0.232)	(0.122)	(0.176)	(0.170)	(0.155)	(0.199)
FEMALE	2.38	2.61	2.38	-3.63	-4.31	-3.55	0.47	0.87	0.44
	(0.228)	(0.198)	(0.212)	(0.198)	(0.123)	(0.186)	(0.877)	(0.777)	(0.886)
ROUND	-0.57**	-0.49**	-0.57**	0.20	0.12	0.20	0.03	-0.17	0.03
	(0.017)	(0.038)	(0.016)	(0.298)	(0.500)	(0.302)	(0.853)	(0.292)	(0.844)
REL LUCK PREV ROUND	-0.01			-0.02			-0.01*		
	(0.446)			(0.118)			(0.067)		
SHARE GOOD LUCK TREAT		1.47			-19.41***			-5.01	
		(0.799)			(0.003)			(0.262)	
SHARE GOOD LUCK EXP			-1.99			-16.14***			-6.10
			(0.787)			(0.008)			(0.406)
N	919	855	919	918	855	918	918	855	918
Number of clusters	12	12	12	12	12	12	12	12	12
Log-likelihood	-3483.66	-3247.20	-3483.97	-3356.35	-3102.67	-3349.04	-3315.74	-3080.89	-3315.49
Bootstrap samples	9999	9999	9999	9999	9999	9999	9999	9999	9999

REL LUCK PREV ROUND is sum of first- and second-period observation error difference in previous round. SHARE GOOD LUCK TREAT (EXP) is actual relative to highest possible number of positive observation error difference realizations in current treatment (experiment). Tobit regressions. Bootstrapped p-values given in parentheses, computed using pairs cluster bootstrap-t (see Section 2 in the Web Appendix) with standard errors clustered on matching group. Sample: All participants with non-inconsistent choices in lottery tasks. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

8.2 Interpreting the Observations in Treatment INT

The framework of Section 5 can explain several aspects of the behavior in INT. Figure A2 shows that, as in ONE, there is substantial polarization in the second-period behavior of laggards in treatment INT. As in ONE, subjects with a handicap from past performance know that they will have to work hard to have a chance of winning. Whereas some subjects may be sufficiently optimistic and/or value winning enough to "go for it", others will remain skeptical and therefore give up. The difference between INT and ONE in the behavior of laggards in Period 2 (excessive effort in INT, but not in ONE) fits well with the idea that joy of winning is higher for players who have already lost once: Even though an intimidation effect should be present in INT as well as in ONE, the more pronounced joy-of winning effect in INT (reflecting revenge motives due to not having obtained the first prize) suggests that *excess* effort should be higher in this case. (Note that efforts themselves should reasonably be expected to be higher in ONE as subjects are still fighting for the full prize W rather that only for the second-period prize as in INT).[56] Similarly, efforts in Period 1 are slightly below the equilibrium level. This may reflect strategic myopia, that is, the inability of subjects to fully understand the positive effect of investments today on winning chances tomorrow.

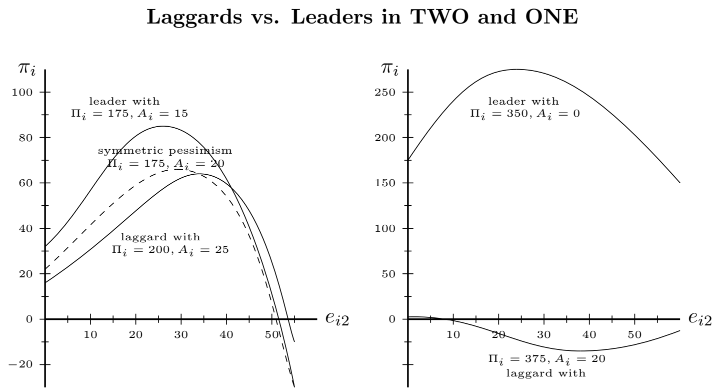
8.3 Interpreting the Differences in Treatments ONE and TWO

In Section 5, we explained the behavior in TWO without reference to the specific parameterization of the belief function that we used to explain the observations in ONE. We now extend the parameterized example to TWO in such a way that it plausibly models the same players in both treatments. We argue that it is perfectly plausible that, when comparing such otherwise identical players who exerted the same first-period effort in TWO, the player in the role of a laggard will exert higher effort than a leader, whereas the converse statement applies in ONE. We continue to maintain that $k = 0.066$ and, in ONE, $W = 300$, whereas $W = 150$ in TWO. In line with the identical first-period effort, we suppose both players are ex-ante symmetric with respect to joy of winning and beliefs about the effectiveness of their efforts.

8.3.1 Treatment TWO

In treatment TWO, we suppose that both players have joy of winning ($V_i = 25$) in Period 1, so that $\Pi_i = 175$. Moreover, we suppose that the belief functions are characterized by $r_i = 0.1$ and $A_i = 20$. Given the initial symmetry, both players choose the same

[56] As to INT and TWO, behavior seems to be similar in Period 2, with substantial excess effort relative to predictions.



(a) Behavior of Leaders and Laggards in TWO

(b) Behavior of Leaders and Laggards in ONE

Figure A3: Laggards vs. Leader in TWO and ONE

initial effort, with the winner (say, Player 1) determined entirely by chance. Suppose that (based on our informational argument that he believes he is facing an opponent with high joy-of-winning) the laggard becomes slightly more pessimistic (say, $A_2 = 25$), whereas the leader becomes slightly more optimistic ($A_1 = 15$). Moreover, suppose that (in line with Assumption 1(ii)) the laggard's joy of winning increases to $V_2 = 50$, while the leader's joy of winning remains $V_1 = 25$.

This leads to expected payoffs as depicted in the left part of Figure A3. The dashed line corresponds to expected payoffs in the benchmark case where both players still have the initial symmetric level of pessimism. The lower solid black line corresponds to expected payoffs of a laggard in Period 2 who, while being more pessimistic, has higher joy of winning than the leader (whose joy of winning remains as in Period 1, but who is more optimistic than the laggard). Hence, the laggard's optimal effort in Period 2 is higher than the leader's.

8.3.2 Treatment ONE

Now consider ONE. As in Section 5 and in line with treatment TWO, we assume that $W = 300$. Further, we take joy of winning to be 50, so that $\Pi_1 = \Pi_2 = 350$ initially. Moreover, we assume that both players are equally optimistic (meaning here that their perceived chance of moving ahead of the other is determined by the same belief function). As in TWO, we assume that $r_i = 0.1$ and $A_i = 20$. Again suppose that, in the first period, player 1 gets lucky and moves ahead of the rival. Unlike in TWO, players in ONE know that the advantage of player 1 increases his winning chances in Period 2. In addition, the informational argument still remains. Together, one should therefore expect the beliefs to be more asymmetric than in TWO. Therefore, suppose for instance that

$A_1 = 0$ and $A_2 = 55$. As in TWO, we assume that the laggard has higher joy of winning after receiving the information (Π_2 increases to 375, while Π_1 remains at 350).

The curves in Part (b) of Figure A3 display the expected payoffs for the optimistic leader and the pessimistic laggard. Even though the latter has slightly higher joy of winning, the effect of pessimism dominates and the laggard chooses much lower effort than the leader (the intimidation effect).

To sum up, with suitable parameter values, we find that laggards exert higher second-period efforts than (otherwise identical) leaders in TWO, but lower efforts in ONE. Even if being a laggard increases joy of winning relative to leaders in both treatments (which fosters relatively high efforts), the adverse effect of being a laggard on pessimism is more substantial in the latter case than in the former and thus dominates.

References

Anderson, L. R. & Freeborn, B. A. (2010). Varying the intensity of competition in a multiple prize rent seeking experiment. *Public Choice*, 143(1-2), 237–254.

Azmat, G. & Iriberri, N. (2016). The provision of relative performance feedback: An analysis of performance and satisfaction. *Journal of Economics & Management Strategy*, 25(1), 77–110.

Balafoutas, L., Kerschbamer, R., & Sutter, M. (2012). Distributional preferences and competitive behavior. *Journal of Economic Behavior & Organization*, 83(1), 125–135.

Baye, M. R., Kovenock, D., & de Vries, C. G. (1996). The all-pay auction with complete information. *Economic Theory*, 8(2), 291–305.

Berger, J. & Pope, D. (2011). Can losing lead to winning? *Management Science*, 57(5), 817–827.

Brookins, P. & Ryvkin, D. (2014). An experimental study of bidding in contests of incomplete information. *Experimental Economics*, 17(2), 245–261.

Buser, T. (2016). The impact of losing in a competition on the willingness to seek further challenges. *Management Science*, 64, 3439–3449.

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *The Review of Economics and Statistics*, 90(3), 414–427.

Casas-Arce, P. & Martínez-Jerez, F. A. (2009). Relative performance compensation, contests, and dynamic incentives. *Management Science*, 55(8), 1306–1320.

Charness, G., Gneezy, U., & Kuhn, M. A. (2012). Experimental methods: Between-subject and within-subject design. *Journal of Economic Behavior & Organization*, 81(1), 1–8.

Chen, H., Ham, S. H., & Lim, N. (2011). Designing multiperson tournaments with asymmetric contestants: An experimental study. *Management Science*, 57(5), 864–883.

Chlosta, K., Pull, K., & Futagami, S. (2014). Tournament structures in Japan and the U.S.: Why are they different and will they change? *Management and Organizational Studies*, 1(1), 63–71.

Clark, D. J. & Nilssen, T. (2013). Learning by doing in contests. *Public Choice*, 156(1-2), 329–343.

Clark, D. J., Nilssen, T., & Sand, J. Y. (2012). Motivating over time: Dynamic win effects in sequential contests. University of Oslo, Department of Economics Working Paper No. 28/2012.

Cubitt, R. P., Starmer, C., & Sugden, R. (1998). On the validity of the random lottery incentive system. *Experimental Economics*, 1(2), 115–131.

Datta, S. & Satten, G. A. (2008). A signed-rank test for clustered data. *Biometrics*, 64(2), 501–507.

Dechenaux, E., Kovenock, D., & Sheremeta, R. M. (2015). A survey of experimental research on contests, all-pay auctions and tournaments. *Experimental Economics*, 18(4), 609–669.

Delfgaauw, J., Dur, R., Non, A., & Verbeke, W. (2015). The effects of prize spread and noise in elimination tournaments: A natural field experiment. *Journal of Labor Economics*, 33(3), 521–569.

Denter, P. & Sisak, D. (2015). Do polls create momentum in political competition? *Journal of Public Economics*, 130, 1–14.

Denter, P. & Sisak, D. (2016). Head starts in dynamic tournaments? *Economics Letters*, 149, 94–97.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., & Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, 9(3), 522–550.

Dufwenberg, M. & Kirchsteiger, G. (2004). A theory of sequential reciprocity. *Games and Economic Behavior*, 47(2), 268–298.

Ederer, F. & Fehr, E. (2017). Deception and incentives: How dishonesty undermines effort provision. Working paper, October 25, 2017.

Eriksson, T., Poulsen, A., & Villeval, M. C. (2009). Feedback and incentives: Experimental evidence. *Labour Economics*, 16, 679–688.

Falk, A. & Fischbacher, U. (2006). A theory of reciprocity. *Games and Economic Behavior*, 54(2), 293–315.

Fehr, E. & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics*, 114(3), 817–868.

Fehr, E. & Williams, T. (2013). Endogenous emergence of institutions to sustain cooperation. Unpublished working paper.

Feld, J. & Zölitz, U. (2017). Understanding peer effects: on the nature, estimation, and channels of peer effects. *Journal of Labor Economics*, 35(2), 387–428.

Fischbacher, U. (2007). z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2), 171–178.

Gächter, S., Johnson, E. J., & Herrmann, A. (2010). Individual-level loss aversion in riskless and risky choices. CeDEx Discussion Paper No. 2010-20.

Gill, D. & Prowse, V. (2014). Gender differences and dynamics in competition: The role of luck. *Quantitative Economics*, 5, 351–376.

Goldman, M. & Rao, J. M. (2017). Loss aversion around a fixed reference point in highly experienced agents. *Available at SSRN 2782110*.

Harbaugh, R. & Ridlon, R. W. (2011). Handicapping under uncertainty in an all-pay auction. Working paper, February 2011.

Harbring, C. & Irlenbusch, B. (2003). An experimental study on tournament design. *Labour Economics*, 10(4), 443–464.

Harris, C. & Vickers, J. (1985). Perfect equilibrium in a model of a race. *The Review of Economic Studies*, 52(2), 193–209.

Harris, P. B. & Houston, J. M. (2010). A reliability analysis of the revised competitiveness index. *Psychological Reports*, 106(3), 870–874.

Hernandez-Lagos, P., Minor, D., & Sisak, D. (2017). Do people who care about others cooperate more? experimental evidence from relative incentive pay. *Experimental Economics*, (pp. 1–27).

Houston, J., Harris, P., McIntire, S., & Francis, D. (2002). Revising the competitiveness index using factor analysis. *Psychological Reports*, 90(1), 31–34.

Klein, A. H. & Schmutzler, A. (2017). Optimal effort incentives in dynamic tournaments. *Games and Economic Behavior*, 103, 199–224.

Kong, X. (2008). Loss aversion and rent-seeking: An experimental study. CeDEx Discussion Paper No. 2008-13.

Konrad, K. A. (2012). Dynamic contests and the discouragement effect. *Revue d'économie politique*, 122, 233–256.

Konrad, K. A. & Kovenock, D. (2009). Multi-battle contests. *Games and Economic Behavior*, 66(1), 256–274.

Kőszegi, B. & Rabin, M. (2006). A model of reference-dependent preferences. *The Quarterly Journal of Economics*, 121(4), 1133–1165.

Kőszegi, B. & Rabin, M. (2007). Reference-dependent risk attitudes. *American Economic Review*, 97(4), 1047–1073.

Kubitz, G. (2020). Sharing cost information in dynamic oligopoly. Unpublished working paper.

Lim, N. (2010). Social loss aversion and optimal contest design. *Journal of Marketing Research*, 47(4), 777–787.

Llorente-Saguer, A., Sheremeta, R. M., & Szech, N. (2016). Designing contests between heterogenous contestants: An experimental study of tie-breaks and bid-caps in all-pay auctions. Working paper, May 8, 2016.

Mago, S. D., Sheremeta, R. M., & Yates, A. (2013). Best-of-three contest experiments: Strategic versus psychological momentum. *International Journal of Industrial Organization*, 31(3), 287–296.

March, C. & Sahm, M. (2017). Asymmetric discouragement in asymmetric contests. *Economics Letters*, 151, 23–27.

McGregor, J. (2013). For whom the bell curve tolls. *The Washington Post – On Leadership* [weblog]. Retrieved 2017/10/18, from http://www.washingtonpost.com/blogs/on-leadership/wp/2013/11/20/for-whom-the-bell-curve-tolls/.

Mermer, A. G. (2017). Effort provision and optimal prize structure in contests with loss-averse players. Working paper, April 5, 2017.

Meyer, M. A. (1992). Biased contests and moral hazard: Implications for career profiles. *Annales d'Économie et de Statistique*, 25/26, 165–187.

Millner, E. L. & Pratt, M. D. (1991). Risk aversion and rent-seeking: An extension and some experimental evidence. *Public Choice*, 69(1), 81–92.

Möller, M. (2012). Incentives versus competitive balance. *Economics Letters*, 117(2), 505–508.

Müller, W. & Schotter, A. (2010). Workaholics and dropouts in organizations. *Journal of the European Economic Association*, 8(4), 717–743.

Murphy, R. O., Ackermann, K. A., & Handgraaf, M. J. J. (2011). Measuring social value orientation. *Judgment and Decision Making*, 6(8), 771–781.

Orrison, A., Schotter, A., & Weigelt, K. (2004). Multiperson tournaments: An experimental examination. *Management Science*, 50(2), 268–279.

Price, C. R. & Sheremeta, R. M. (2011). Endowment effects in contests. *Economics Letters*, 111(3), 217–219.

Price, C. R. & Sheremeta, R. M. (2015). Endowment origin, demographic effects, and individual preferences in contests. *Journal of Economics & Management Strategy*, 24(3), 597–619.

Ridlon, R. & Shin, J. (2013). Favoring the winner or loser in repeated contests. *Marketing Science*, 32(5), 768–785.

Schmitt, P., Shupp, R., Swope, K., & Cadigan, J. (2004). Multi-period rent-seeking contests with carryover: Theory and experimental evidence. *Economics of Governance*, 5(3), 187–211.

Sheremeta, R. M. (2010). Experimental comparison of multi-stage and one-stage contests. *Games and Economic Behavior*, 68(2), 731–747.

Sheremeta, R. M. (2011). Contest design: An experimental investigation. *Economic Inquiry*, 49(2), 573–590.

Sheremeta, R. M., Masters, W. A., & Cason, T. N. (2017). Winner-take-all and proportional-prize contests: Theory and experimental results. Working paper, August 15, 2017.

Sheremeta, R. M. & Zhang, J. (2010). Can groups solve the problem of over-bidding in contests? *Social Choice and Welfare*, 35(2), 175–197.

Shupp, R., Sheremeta, R. M., Schmidt, D., & Walker, J. (2013). Resource allocation contests: Experimental evidence. *Journal of Economic Psychology*, 39, 257–267.

Stracke, R., Höchtl, W., Kerschbamer, R., & Sunde, U. (2014). Optimal prizes in dynamic elimination contests: Theory and experimental evidence. *Journal of Economic Behavior & Organization*, 102, 43–58.

Tong, K. & Leung, K. (2002). Tournament as a motivational strategy: Extension to dynamic situations with uncertain duration. *Journal of Economic Psychology*, 23, 399–420.

Tversky, A. & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4), 297–323.

Welch, J. (2005). The vitality curve. *Leadership Excellence*, 22(9), 4–5.

Zizzo, D. J. (2002). Racing with uncertainty: a patent race experiment. *International Journal of Industrial Organization*, 20(6), 877–902.

Web Appendix

9 Details on pre-experimental questionnaire

9.1 Details on measure for social value orientation

The SVO Slider Measure developed by Murphy et al. (2011) consists of a sequence of six dictator games. In each dictator game, the participants have to choose one of nine allocations in terms of payoffs for themselves and another participant (see Table A5 and Figure A4). As an example, game 5 involves the distribution of a total surplus of 150 points between oneself and the other.[57] The allocations are constructed in a way that in each dictator game, each of the classical types of social value orientation either strictly prefers exactly one of the allocations, or is wholly indifferent between all of them.

Table A5: Possible allocation choices in SVO Slider Measure

Game #	Receiver	Allocation #								
		1	2	3	4	5	6	7	8	9
1	oneself	85	85	85	85	85	85	85	85	85
	other	85	76	68	59	50	41	33	24	15
2	oneself	85	87	89	91	93	94	96	98	100
	other	15	19	24	28	33	37	41	46	50
3	oneself	50	54	59	63	68	72	76	81	85
	other	100	98	96	94	93	91	89	87	85
4	oneself	50	54	59	63	68	72	76	81	85
	other	100	89	79	68	58	47	36	26	15
5	oneself	100	94	88	81	75	69	63	56	50
	other	50	56	63	69	75	81	88	94	100
6	oneself	100	98	96	94	93	91	89	87	85
	other	50	54	59	63	68	72	76	81	85

Source: Adapted from Murphy et al. (2011).

The categorization of the participants into one of the orientation types is based on their choices in these dictator games. Let $\overline{A}_o$ be the average of what the participant

[57] Note that in this dictator game, the price of giving is 1. In the other games, the price of giving is not equal to 1, so that the total surplus varies between choices.

allocated to the other across the six games and $\overline{A}_s$ the average of what the participant allocated to him/herself. Murphy et al. (2011) then define a participant's SVO index as

$$\text{SVO} = \arctan\left(\frac{\overline{A}_o - 50}{\overline{A}_s - 50}\right).$$

Intuitively, in Figure A4, the index corresponds to the angle at vertex $(50, 50)$ between the average allocation chosen and the horizontal. Note that each classical type would generate a particular index value when faced with the dictator games. Following Murphy et al. (2011), we determine the participants' social value orientation type as the classical orientation type whose index value their own index value is closest to. Table A6 contains the index values implied by orientation types and the resulting intervals for the empirical categorization.[58]

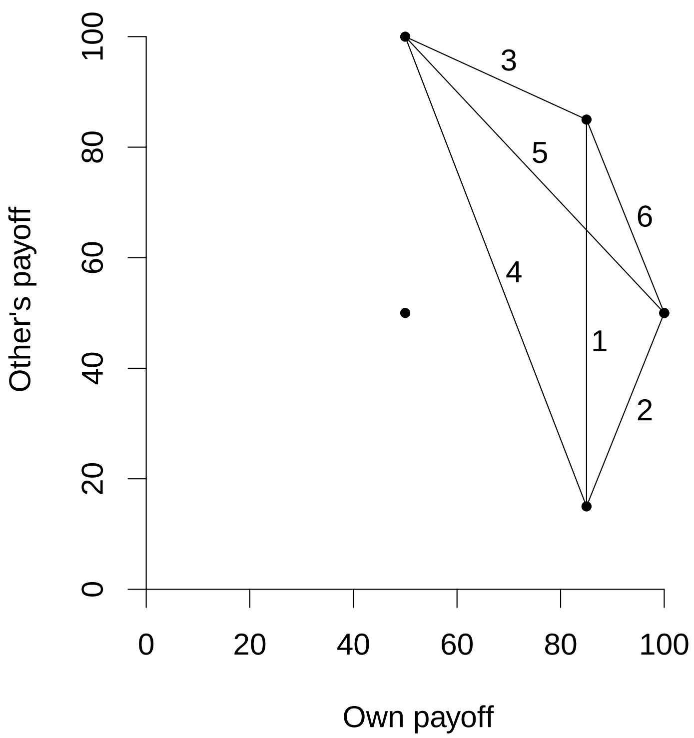


Figure A4: Allocations in the SVO Slider Measure. The lines connect the allocations in the six dictator games given in Table A5. Source: Adapted from Fehr and Williams (2013).

We implemented the SVO Slider Measure in the following way. Before making decisions in the dictator games, the participants were instructed that after the completion of the questionnaire, they would be randomly paired with another participant, and that

[58]Note that there is a range of possible values for prosocials and individualists. This is due to the fact that both orientation types are indifferent between the allocations in one of the six dictator games each.

Table A6: Characterization of SVO types

Orientation type	Implied index value	Range for characterization
Competitive	$-16.26°$	$\leq -12.04°$
Individualist	$\in [-7.82°, 7.82°]$	$(-12.04°, 22.45°]$
Prosocial	$\in [37.09°, 52.91°]$	$(22.45°, 57.15°]$
Altruist	$61.39°$	$> 57.15°$

Source: Adapted from Murphy et al. (2011).

one of the 12 decisions made by both participants in this pair would then be randomly chosen to determine their payoffs. Then, every game appeared separately on the participants' computer screen. For every participant, the order of presentation was randomly determined. All payoffs were expressed in CHF, using an exchange rate of CHF 1 per 10 points as given in Table A5.

9.2 Details on measure for risk aversion

In the lottery task, the safe payoff varied between CHF $X \in \{2, 3, 4, 5, 6, 7\}$. The lottery yielded CHF 10 or CHF 0 with equal chances. The choice situations were presented on one screen, ordered in decreasing value of the safe payoff (see Table A7).

Table A7: Lottery task to elicit risk aversion

Situation #	Safe payoff	Lottery
1	CHF 7	50%: CHF 10, 50%: CHF 0
2	CHF 6	50%: CHF 10, 50%: CHF 0
3	CHF 5	50%: CHF 10, 50%: CHF 0
4	CHF 4	50%: CHF 10, 50%: CHF 0
5	CHF 3	50%: CHF 10, 50%: CHF 0
6	CHF 2	50%: CHF 10, 50%: CHF 0

We infer the participants' degree of risk aversion from the position at which they switched from choosing the safe amount to choosing the lottery. To this end, we use a similar argument as Dohmen et al. (2011): Since the expected value of the lottery is CHF 5, risk-loving subjects should switch before situation 3 (at which the safe payoff equals the expected value of the lottery), and risk-averse subjects after situation 3. Hence, the later an individual switches, the higher is the underlying degree of risk aversion. We thus define the variable SPRISK as the number of the situation at which the participants chose the lottery for the first time.[59,60]

After the completion of the questionnaire, the computer randomly selected one of the six choice situations for each participant (Cubitt et al., 1998). The participants' payoff then followed from their decision for the selected situation – if they had chosen the safe payoff, the payoff was equal to the safe payoff, while if they had chosen the lottery, the payoff was randomly chosen between CHF 10 and CHF 0.

9.3 Details on measure for loss aversion

In this task, which was developed by Gächter et al. (2010), the participants have to make decisions for six choice situations involving a safe payoff of CHF 0 and a lottery. The lotteries yields, with equal chances, a payoff of CHF 6 or a payoff of CHF $-X$, while $X \in \{2, 3, 4, 5, 6, 7\}$ (see Table A8).

Table A8: Lottery task to elicit loss aversion

Situation #	Safe payoff	Lottery
1	CHF 0	50%: CHF 6, 50%: CHF -2
2	CHF 0	50%: CHF 6, 50%: CHF -3
3	CHF 0	50%: CHF 6, 50%: CHF -4
4	CHF 0	50%: CHF 6, 50%: CHF -5
5	CHF 0	50%: CHF 6, 50%: CHF -6
6	CHF 0	50%: CHF 6, 50%: CHF -7

Source: Adapted from Gächter et al. (2010).

The participants' level of loss aversion follows from the point at which they start rejecting the lottery in favor of the safe amount. Following Gächter et al. (2010), a decision maker is indifferent between accepting and rejecting a lottery that yields a gain of G and a loss of L with equal chances if

$$G = \lambda \cdot L.$$

[59] With this definition, a participant who always chose the lottery – and is thus very risk-loving – receives a value of 1, which is an upper bound for the switching point in the hypothetical case that there were additional situations above situation 1 with a safe payoff of more than CHF 7. For a participant who always chose the lottery – and is thus very risk-averse – we set SPRISK to 7, which is a lower bound for the switching point in the hypothetical case that there were additional situations below situation 6 with a safe payoff of less than CHF 2.

[60] The measure which we use is a linear transformation of the measure of Dohmen et al. (2011).

Gächter et al. (2010) define λ as the coefficient of loss aversion. They argue that $\lambda > 1$ implies loss aversion, as losses are weighted more heavily than equally sized gains.[61] Gächter et al. (2010) then calculate a participant's λ as

$$\lambda = \frac{G^a}{L^a},$$

where G^a and L^a are the gain and the loss of the lottery with the highest loss that is still accepted by a participant. Note that if a participants always (never) rejected the lottery, we can only determine a lower (upper) bound for λ.[62] Furthermore, if a participant was inconsistent and rejected a first lottery but accepted a second that would have yielded a higher loss than the first, it is not possible to determine λ. Table A9 shows the values for λ that follow from the possible choices in the lottery task.

Table A9: Possible choices and implied values for λ

Choice	implied λ
Always reject	> 3
Accept #1, reject #2–#6	3
Accept #1 – #2, reject #3–#6	2
Accept #1 – #3, reject #4–#6	1.5
Accept #1 – #4, reject #5–#6	1.2
Accept #1 – #5, reject #6	1
Never reject	≤ 0.87

Source: Adapted from Gächter et al. (2010).

As for the lottery task to elicit risk aversion, the computer randomly selected one of the six choice situations for each participant after the completion of the questionnaire (Cubitt et al., 1998) and determined the payoffs according to the decisions for the selected situation.

9.4 Details on measure for competitiveness

The Revised Competitiveness Index developed by Houston et al. (2002) consists of 14 statements about competition in daily life contexts (see Table A10). The participants state on a Likert scale from 1 (strongly disagree) to 5 (strongly agree). According to the definition of Houston et al. (2002), a participant's RCI value equals the sum of the

[61] This is a special case of the more general model of Tversky and Kahneman (1992), who allow for probability weighting and nonlinear utility.

[62] In these cases, we set λ to 3 or 0.87, respectively.

individual answers and ranges between 14 and 70.[63] In the pre-experimental questionnaire, the questions appeared separately on the screen in a randomly determined order.

Table A10: Statements in Revised Competitiveness Index

#	Statement
1	I like competition.
2	I am a competitive individual.
3	I enjoy competing against an opponent.
4	I don't like competing against other people.
5	I get satisfaction from competing with others.
6	I find competitive situations unpleasant.
7	I dread competing against other people.
8	I try to avoid competing with others.
9	I often try to out perform others.
10	I try to avoid arguments.
11	I will do almost anything to avoid an argument.
12	I often remain quiet rather than risk hurting another person.
13	I don't enjoy challenging others even when I think they are wrong.
14	In general, I will go along with the group rather than create conflict.

Source: Adapted from Houston et al. (2002).

10 The pairs-cluster bootstrap-t procedure

Note that two issues complicate the estimation: (a) the clustering of observations within matching groups, and (b) the small number of clusters, here equal to the number of matching groups (12). For an unbiased estimation of the variance-covariance matrix, issue (a) requires to use cluster-robust standard errors.[64] It is well-known, however, that with few clusters (issue (b)), the conventional cluster-robust sandwich estimator for the variance-covariance matrix may be biased as well (see, for example, Cameron et al., 2008). As a solution, we determine the p-values of the coefficients with the pairs cluster bootstrap-t procedure. Cameron et al. (2008) show by simulation that in an ordinary least squares estimation, this procedure maintains a reasonably correct size with both clustered observations and a small number of clusters.

The following summary relies on Cameron et al. (2008, p. 427). Suppose there are C clusters. The pairs-cluster bootstrap-t procedure starts with an ordinary least squares

[63] The answers to statements 4, 6, 7, 8, 11, 12, 13 and 14 are reverse-coded in the calculation of the overall score.

[64] Clustering of the observations implies correlation in the error structure, which violates the assumption of uncorrelated errors underlying the standard regression model. As a consequence, the conventional estimation of the variance-covariance matrix may be biased.

estimation of coefficient j, $\hat{\beta}_j$, and a cluster-robust estimation of its standard error, $s_{\hat{\beta}_j}$, based on the whole sample. In the next step, B bootstrap replications are executed. In each replication, the procedure samples C clusters (i.e., all of the observations from that cluster) with replacement from the original sample of clusters and calculates t-statistics based on the resampled data using cluster-robust standard errors. More precisely, let $\hat{\beta}_{j,b}$ be the ordinary least squares estimate of the coefficient in the bth replication and $s_{\hat{\beta}_{j,b}}$ the cluster-robust estimate of its standard error. The bth t-statistic is then defined as

$$t_{j,b} = \frac{\hat{\beta}_{j,b} - \hat{\beta}_j}{s_{\hat{\beta}_{j,b}}}.$$

In the last step, the p-value for parameter j results from comparing the regular t-statistic

$$t_j = \frac{\hat{\beta}_j}{s_{\hat{\beta}_j}}$$

to the empirical distributions of the $t_{j,b}$.[65] Note that in some of the regressions, we determine the estimates of the coefficients with the Tobit model instead of ordinary least squares to take the censoring of effort choices at 0 and 55 into account.

[65] We implemented the bootstraps in R using the AER, censReg, doBy, doMC, doRNG, foreach, lmtest, sampling and sandwich packages.

11 Instructions

This section contains the instructions for session 1. The instructions for session 2 and 3 were analogous.

Instructions

General Information

Three parts, 10 periods each

This is experiment has 3 parts (Part I, Part II, and Part III). Each part is divided into 10 periods. Thus, there are 30 periods in total.

The instructions on this page and on pages 2 and 3 are relevant for all three parts. Instructions which are specific to Part I will follow on page 4. Instructions which are specific to Part II and Part III will be distributed to you before the corresponding part.

In each period, you will generate a payoff. The payoff depends on your decisions in that period and the decisions of others in that period. How you generate a payoff will be explained to you in the following.

Final payoff

Your final payoff from the experiment will be a participation payment of 10 CHF plus the payoff you generated in one randomly chosen period. The period that is randomly chosen will be the same period for all participants. Every period is equally likely to be chosen. During the experiment, you will not know which period will be chosen. Therefore, you should treat each period as if it would be the one that is relevant for your final payoff.

Final payoff = 10 CHF + your payoff from a randomly chosen period

Upon completion of the experiment, you will be paid individually and in private.

Exchange rate

Throughout the experiment, payoffs are expressed in terms of "points". At the end of the experiment, payoffs in points will be converted into payoffs in CHF. The exchange rate is:

10 points = 1 CHF

Rules

If you have any questions during the experiment, please raise your hand and wait for an experimenter to come to you. During the experiment, you are not allowed to communicate with other participants, exclaim, use personal electronic devices, or use the computer in a way not specified by the experimenter. If you are not following these rules, you may be excluded from the experiment and might only receive the participation payment.

What happens in a period

Interaction with randomly chosen participant

In each of the 30 periods, every participant is assigned into a pair with one randomly chosen other participant. In the following, we will refer to the other randomly chosen participant as the "*other*". The participants will never know the identity of the other, nor will the other know their identity.

Two stages per period

Each period consists of two stages. In each stage, the participants must make a decision. In the following, this is explained in further detail.

Stage 1 works in the following way:

Stage 1: input

First, each participant individually chooses an input level between 0 and 55 in increments of 0.5. The numbers are entered in the corresponding field of the computer screen. Choosing a positive input level is costly for the participants. A detailed explanation of the costs the participants have to pay for choosing a particular input level follows below.

Stage 1: output

Second, the computer determines each participant's output level in Stage 1. A participant's output level depends on the participant's input level in Stage 1 and on a random number. This random number is drawn for each participant individually in Stage 1.

output level in Stage 1 = input level in Stage 1 + random number in Stage 1

On average, the random numbers are zero, but they can take up positive and negative values. Positive and negative values are equally likely, and values close to zero are more likely than values further away from zero. On page 2 of the appendix, you find a detailed explanation of the distribution of the random numbers.

This means that on average, a participant's output level corresponds exactly to this participant's input level. Thus, a higher input level results on average in a higher output level. However, depending on the realization of the random number (positive or negative), the output level can positively or negatively deviate from the chosen input level. Positive and negative deviations are equally likely, and small deviations are more likely than large deviations.

Stage 1: information

At the end of Stage 1, the computer screen displays the following information to each participant: the participant's own output level in Stage 1, the other's output level in Stage 1, and the difference between the participant's own output level and the other's output level. Note that the participants will not know the other's input level, nor will the other know their input level.

Stage 2

Stage 2 works in the same way: Each participant chooses an input level. Choosing a positive input level is costly to the participants. Then, the computer draws another random number for each participant and determines each participant's output level in Stage 2, which is the sum of the participant's input level in Stage 2 and the participant's random number in Stage 2. Finally, the computer displays the corresponding information.

Payments At the end of each period, the participants receive payments depending on their own and the other's output in Stage 1 and in Stage 2. The rules according to which these payments are made differ between the three parts of the experiment. The payment rules in each part will be explained in the instructions specific to this part, i.e., the payment rules for Part I are explained on page 4.

Payoff A participant's payoff is calculated in the same way in all three parts: It is equal to the difference between the participant's total payments and the costs for the participant's inputs in both stages, plus a fixed payment of 200 points.

payoff from a period = total payments
- costs for input in Stage 1
- costs for input in Stage 2
+ 200

A participant's payoff is therefore higher when the participant's payments are higher and the participant's costs for the inputs are lower. Note that a participant's payoff will never be negative.

Costs On page 1 of the appendix, you find a table and a graph showing which costs the participants have to pay for choosing a particular input level in a stage. The costs are increasing in the input level chosen by a participant. That is, the higher is the chosen input level, the higher are the costs the participant has to pay.

No predictions possible Note that it is not possible to make predictions about future draws of random numbers from draws of random numbers in the past. The random numbers are newly drawn for every participant in every stage of every period, and these draws are independent from each other.

Instructions specific to Part I

In Part I, the **payment rules** after Stage I and Stage II are the following:

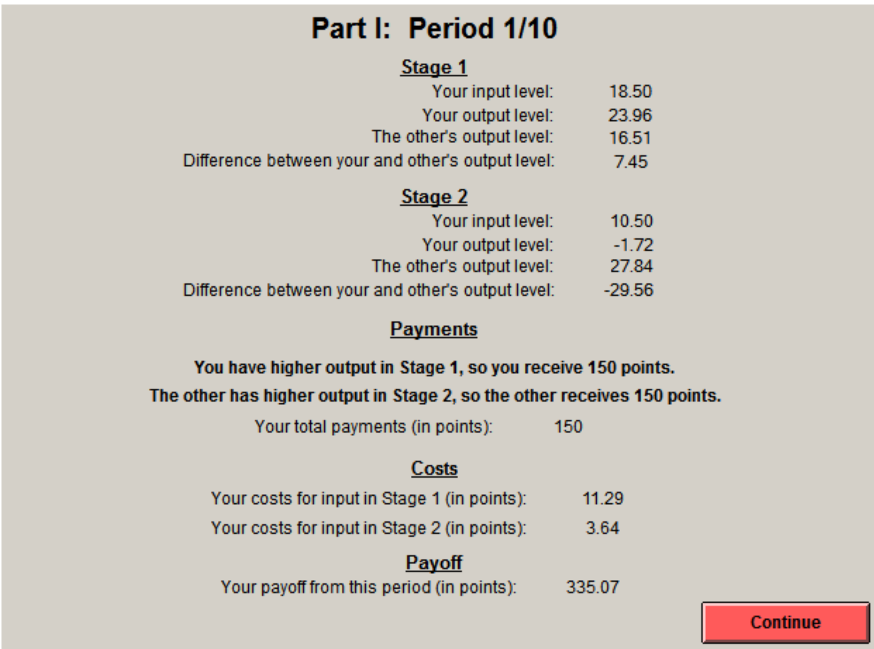
- A first payment of 150 points is given to the participant in the pair who has higher output in Stage 1. The participant with lower output in Stage 1 receives no payment. If both outputs in Stage 1 are equal, the participant who receives the first payment is randomly chosen.
- A second payment of 150 points is given to the participant in the pair who has higher output in Stage 2. The participant with lower output in Stage 2 receives no payment. If both outputs in Stage 2 are equal, the participant who receives the second payment is randomly chosen.

150 points to participant with higher output in Stage 1
150 points to participant with higher output in Stage 2

This means that output in Stage 1 only counts for the first payment, and output in Stage 2 only counts for the second payment.

Example – Part I

In the following, you take up the perspective of some participant in one of the 10 periods of Part I. Below, you see a picture of the computer screen after Stage 1 and Stage 2. Note that the numbers only serve as an example to illustrate the rules of Part I, and are not a recommendation towards what you should do.



Part I: Period 1/10

Stage 1

Your input level:	18.50
Your output level:	23.96
The other's output level:	16.51
Difference between your and other's output level:	7.45

Stage 2

Your input level:	10.50
Your output level:	-1.72
The other's output level:	27.84
Difference between your and other's output level:	-29.56

Payments

You have higher output in Stage 1, so you receive 150 points.
The other has higher output in Stage 2, so the other receives 150 points.

Your total payments (in points): 150

Costs

Your costs for input in Stage 1 (in points):	11.29
Your costs for input in Stage 2 (in points):	3.64

Payoff

Your payoff from this period (in points): 335.07

Continue

In **Stage 1**, you chose an input level of 18.5. The computer then determined your output level in Stage 1 as 23.96. This means that your random number was +5.46 (23.96 + 5.46 = 18.5). The other's output level in Stage 1 was 16.51. The computer thus displays the difference between your and the other's output level as 7.45 (23.96 - 16.51 = 7.45).

In **Stage 2**, you chose an input level of 10.50. The computer then determined your output level in Stage 2 as -1.72. This means that your random number was -12.22 (10.5 - 12.22 = -1.72). The other's output level in Stage 2 was 27.84. The computer thus displays the difference between your and the other's output level as -29.56 (-1.72 - 27.84 = -29.56).

Then, you and the other received **payments** depending on your and the other's output in Stage 1 and in Stage 2. First, since your output in Stage 1 (23.96) was higher than the other's output in Stage 1 (16.51), a payment of 150 points was given to you. Second, since your output in Stage 2 (-1.72) was lower than the other's output in Stage 2 (27.84), a payment of 150 points was given to the other. Thus, your total payments were 150 points (150 + 0 = 150).

The **costs** for your input level of 18.5 in Stage 1 were 11.29 points, and the costs for your input level of 10.5 in Stage 2 were 3.64 points.

Your **payoff from this period** are your total payments (150 points) minus your costs for input in Stage 1 (11.29 points) and for input in Stage 2 (3.64 points), plus the fixed payment of 200 points. Your payoff from this period is thus 335.07 points (150 - 11.29 - 3.64 + 200 = 335.07).

Suppose this period would be randomly chosen to be relevant for your **final payoff**. Your final payoff would then be 45.51 CHF, which is the sum of the 10 CHF participation payment and the payoff from this period converted into 33.51 CHF (= 355.07/10).

Instructions specific to Part II

In Part II, the **payment rules** after Stage I and Stage II are the following:

- A first payment of 150 points is given to the participant in the pair who has higher output in Stage 1. The participant with lower output in Stage 1 receives no payment. If both outputs in Stage 1 are equal, the participant who receives the first payment is randomly chosen.
- A second payment of 150 points is given to the participant in the pair whose sum of output in Stage 1 and output in Stage 2 is higher. The participant whose sum of output in Stage 1 and output in Stage 2 is lower receives no payment. If both participants have the same sum of output in Stage 1 and output in Stage 2, the participant who receives the second payment is randomly chosen.

150 points to participant with higher output in Stage 1
150 points to participant with higher sum of output

This means that output in Stage 1 counts both for the first payment and for the second payment, while output in Stage 2 only counts for the second payment.

Example – Part II

Reconsider the example from Part I. You had an output level of 23.96 in Stage 1 and of -1.72 in Stage 2. The sum of your output in Stage 1 and output in Stage 2 is then 22.24 (23.96 + (-1.72) = 22.24).

The other had an output level of 16.51 in Stage 1 and of 27.84 in Stage 2. The other's sum of output in Stage 1 and output in Stage 2 is then 44.35 (16.51 + 27.84 = 44.35).

With the rules of Part II, payments are as follows:

First, since your output in Stage 1 (23.96) is higher than the other's output in Stage 1 (16.51), a payment of 150 points is given to you. Second, since your sum of output (22.24) is smaller than the other's sum of output (44.35), a payment of 150 points is given to the other. Your total payments are thus 150 points (150 + 0 = 150).

Suppose you chose the same input levels as in the example for Part I. Your costs for input are thus 11.29 for Stage 1, and 3.64 for Stage 2. Your payoff from this period is thus 335.07 points (150 - 11.29 - 3.64 + 200 = 335.07).

Instructions specific to Part III

In Part III, the **payment rules** after Stage I and Stage II are the following:

- A payment of 300 points is given to the participant whose sum of output in Stage 1 and output in Stage 2 is higher. The participant whose sum of output in Stage 1 and output in Stage 2 is lower receives no payment. If both participants have the same sum of output in Stage 1 and output in Stage 2, the participant who receives the second payment is randomly chosen.

300 points to participant with higher sum of output

This means that output in Stage 1 and output in Stage 2 both count for the payment.

Example – Part III

Reconsider the example from Part I. You had an output level of 23.96 in Stage 1 and of -1.72 in Stage 2. The sum of your output in Stage 1 and output in Stage 2 is then 22.24 (23.96 + (-1.72) = 22.24).

The other had an output level of 16.51 in Stage 1 and of 27.84 in Stage 2. The other's sum of output in Stage 1 and output in Stage 2 is then 44.35 (16.51 + 27.84 = 44.35).

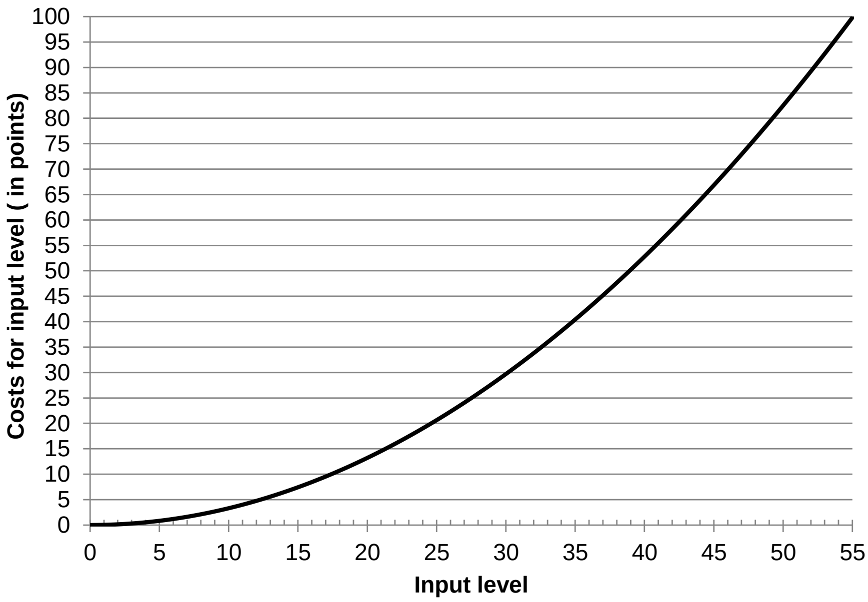
With the rules of Part III, payments are as follows:

Since your sum of output (22.24) is smaller than the other's sum of output (44.35), a payment of 300 points is given to the other. Your total payments are thus 0 points (0 + 0 = 0).

Suppose you chose the same input levels as in the example for Part I. Your costs for input are thus 11.29 for Stage 1, and 3.64 for Stage 2. Your payoff from this period is thus 185.07 points (0 - 11.29 - 3.64 + 200 = 185.07).

Costs for input level depending on choice of input level

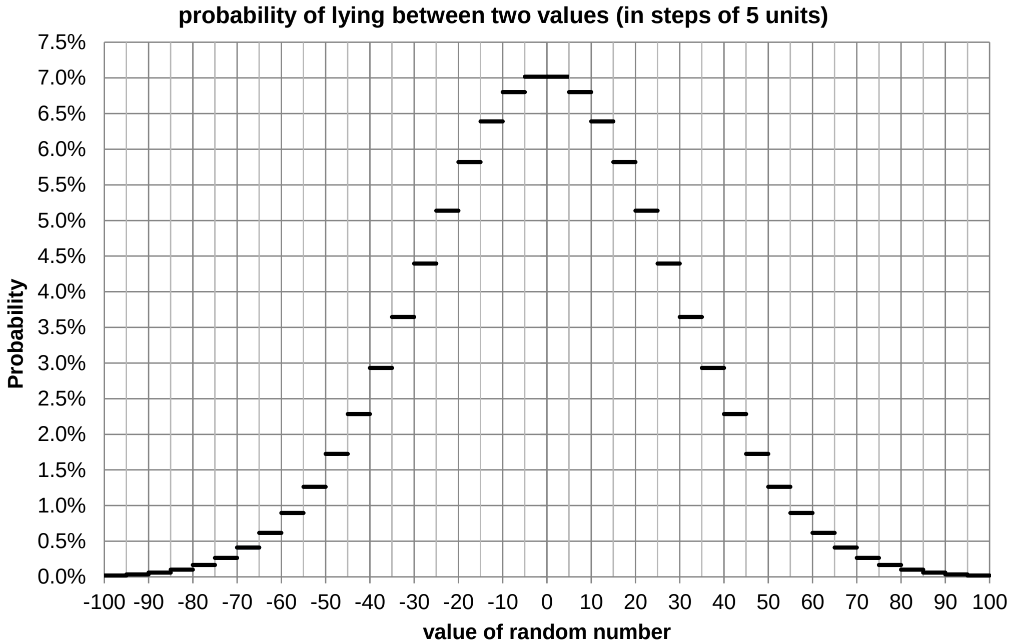
input level	costs for input level (in points)	input level	costs for input level (in points)	input level	costs for input level (in points)	input level	costs for input level (in points)
0.0	0.00	14.0	6.47	28.0	25.87	42.0	58.21
0.5	0.01	14.5	6.94	28.5	26.80	42.5	59.61
1.0	0.03	15.0	7.43	29.0	27.75	43.0	61.02
1.5	0.07	15.5	7.93	29.5	28.72	43.5	62.44
2.0	0.13	16.0	8.45	30.0	29.70	44.0	63.89
2.5	0.21	16.5	8.98	30.5	30.70	44.5	65.35
3.0	0.30	17.0	9.54	31.0	31.71	45.0	66.83
3.5	0.40	17.5	10.11	31.5	32.74	45.5	68.32
4.0	0.53	18.0	10.69	32.0	33.79	46.0	69.83
4.5	0.67	18.5	11.29	32.5	34.86	46.5	71.35
5.0	0.83	19.0	11.91	33.0	35.94	47.0	72.90
5.5	1.00	19.5	12.55	33.5	37.03	47.5	74.46
6.0	1.19	20.0	13.20	34.0	38.15	48.0	76.03
6.5	1.39	20.5	13.87	34.5	39.28	48.5	77.62
7.0	1.62	21.0	14.55	35.0	40.43	49.0	79.23
7.5	1.86	21.5	15.25	35.5	41.59	49.5	80.86
8.0	2.11	22.0	15.97	36.0	42.77	50.0	82.50
8.5	2.38	22.5	16.71	36.5	43.96	50.5	84.16
9.0	2.67	23.0	17.46	37.0	45.18	51.0	85.83
9.5	2.98	23.5	18.22	37.5	46.41	51.5	87.52
10.0	3.30	24.0	19.01	38.0	47.65	52.0	89.23
10.5	3.64	24.5	19.81	38.5	48.91	52.5	90.96
11.0	3.99	25.0	20.63	39.0	50.19	53.0	92.70
11.5	4.36	25.5	21.46	39.5	51.49	53.5	94.45
12.0	4.75	26.0	22.31	40.0	52.80	54.0	96.23
12.5	5.16	26.5	23.17	40.5	54.13	54.5	98.02
13.0	5.58	27.0	24.06	41.0	55.47	55.0	99.83
13.5	6.01	27.5	24.96	41.5	56.83		



Distribution of the random numbers

In every stage of every period, a participant's random number is drawn from a normal distribution with expected value 0 and standard deviation 28.28.

The following graph shows the probabilities that the random number lies between two values (in steps of 5 units):



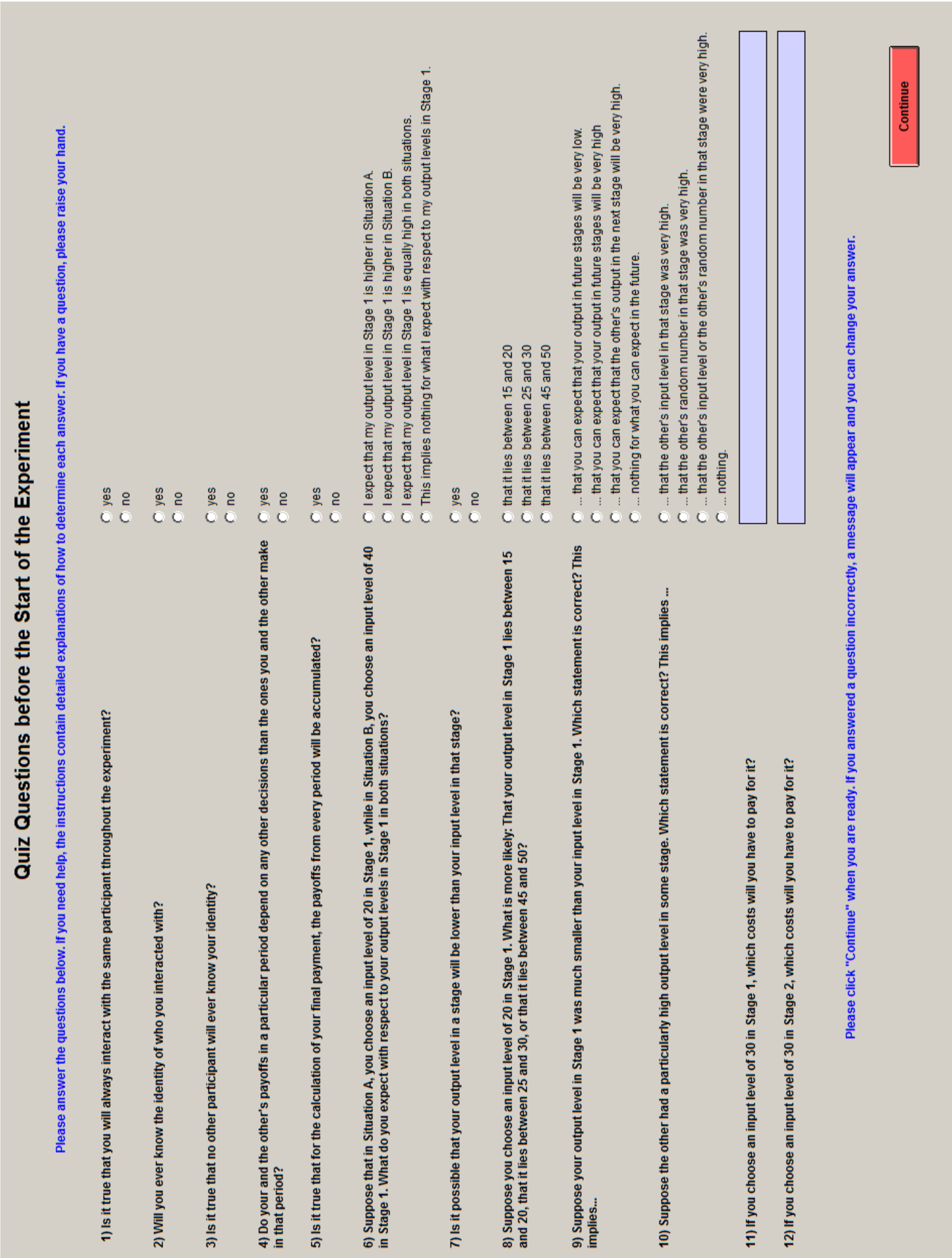
How to read the graph: The horizontal bars each represent a certain range of possible values for the random number. On the horizontal axis, you can read the left and the right boundary of a range. On the vertical axis, you can read the probability that the random number lies within this range.

Example: The probability that the random number lies between +5 and +10 is about 6.8%. This is equal to the probability that the random number lies between -10 and -5. This means that if you choose an input level of, say, 20, the probability that your output level lies between 25 and 30 is about 6.8%, and the probability that your output level lies between 10 and 15 is also 6.8%.

Note that by adding up the probabilities of the four bars to the right of zero (7, 6.8, 6.4, 5.8) and of the four bars to the left of zero (5.8, 6.4, 6.8, 7), you learn that the probability that the output level is within +/- 20 units around your input level is about 52%.

12 Control questions

This section contains the control questions that the participants had to solve at the beginning of the experiment and before treatment *TWO* (here shown for session 1). The control questions for treatments *(INT)* and *(ONE)* were analogous to those for *(TWO)*.



Quiz Questions before the Start of the Experiment

Please answer the questions below. If you need help, the instructions contain detailed explanations of how to determine each answer. If you have a question, please raise your hand.

1) Is it true that you will always interact with the same participant throughout the experiment?
- yes
- no

2) Will you ever know the identity of who you interacted with?
- yes
- no

3) Is it true that no other participant will ever know your identity?
- yes
- no

4) Do your and the other's payoffs in a particular period depend on any other decisions than the ones you and the other make in that period?
- yes
- no

5) Is it true that for the calculation of your final payment, the payoffs from every period will be accumulated?
- yes
- no

6) Suppose that in Situation A, you choose an input level of 20 in Stage 1, while in Situation B, you choose an input level of 40 in Stage 1. What do you expect with respect to your output levels in Stage 1 in both situations?
- I expect that my output level in Stage 1 is higher in Situation A.
- I expect that my output level in Stage 1 is higher in Situation B.
- I expect that my output level in Stage 1 is equally high in both situations.
- This implies nothing for what I expect with respect to my output levels in Stage 1.

7) Is it possible that your output level in a stage will be lower than your input level in that stage?
- yes
- no

8) Suppose you choose an input level of 20 in Stage 1. What is more likely: That your output level in Stage 1 lies between 15 and 20, that it lies between 25 and 30, or that it lies between 45 and 50?
- that it lies between 15 and 20
- that it lies between 25 and 30
- that it lies between 45 and 50

9) Suppose your output level in Stage 1 was much smaller than your input level in Stage 1. Which statement is correct? This implies...
- ... that you can expect that your output in future stages will be very low.
- ... that you can expect that your output in future stages will be very high
- ... that you can expect that the other's output in the next stage will be very high.
- ... nothing for what you can expect in the future.

10) Suppose the other had a particularly high output level in some stage. Which statement is correct? This implies ...
- ... that the other's input level in that stage was very high.
- ... that the other's random number in that stage was very high.
- ... that the other's input level or the other's random number in that stage were very high.
- ... nothing.

11) If you choose an input level of 30 in Stage 1, which costs will you have to pay for it?

12) If you choose an input level of 30 in Stage 2, which costs will you have to pay for it?

Please click "Continue" when you are ready. If you answered a question incorrectly, a message will appear and you can change your answer.

Continue

Quiz Questions specific to Part I

Please answer the questions below. If you need help, the instructions contain detailed explanations of how to determine each answer. If you have a question, please raise your hand.

Suppose you are in some period of Part I.

1) With the input level you choose in Stage 1, you can affect...

- ... only who gets the first payment of 150 points.
- ... only who gets the second payment of 150 points.
- ... who gets the first payment of 150 points AND who gets the second payment of 150 points.
- ... nothing.

2) With the input level you choose in Stage 2, you can affect...

- ... only who gets the first payment of 150 points.
- ... only who gets the second payment of 150 points.
- ... who gets the first payment of 150 points AND who gets the second payment of 150 points.
- ... nothing.

3) Is it true that if your output in Stage 1 is much lower than the other's output in Stage 1, your output in Stage 2 must necessarily be much higher than the other's output in Stage 2 to receive the second payment of 150 points?

- yes
- no

Suppose your output level in Stage 1 is 10, and your output level in Stage 2 is 20. Suppose the other's output level in Stage 1 is -5, and the other's output level in Stage 2 is 45.

4) What are your total payments in this period (in points)?

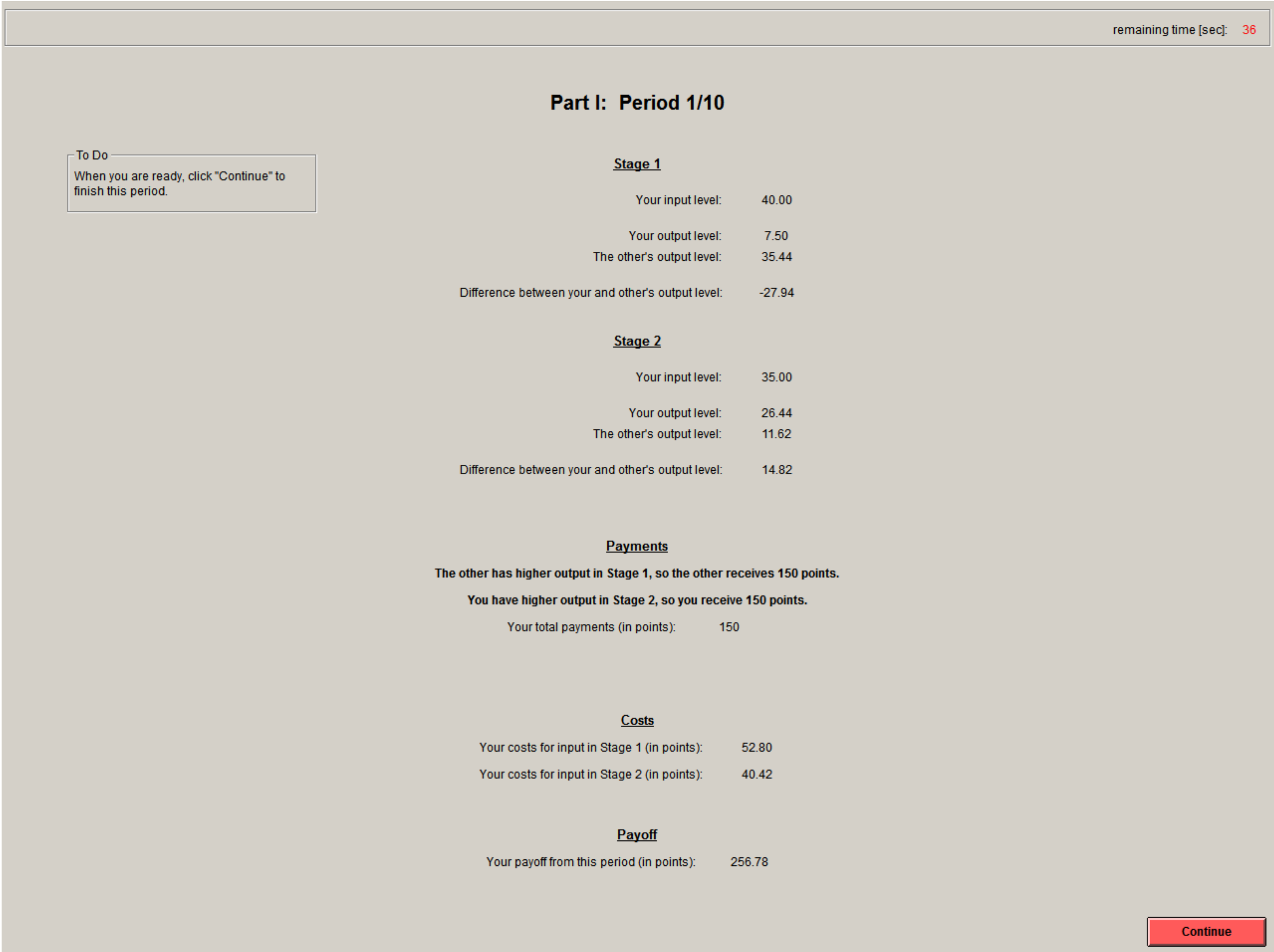
5) What the other's total payments in this period (in points)?

Please click "Continue" when you are ready. If you answered a question incorrectly, a message will appear and you can change your answer.

Continue

13 Computer interface


Part I: Period 1/10
To Do
Please make your input choice for Stage 2.
When you are ready, click "Continue" to proceed to the results of Stage 2 and of the whole period.
Stage 1
Your input level: 40.00
Your output level: 7.50
The other's output level: 35.44
Difference between your and other's output level: -27.94
Stage 2
Your input level in Stage 2:
Continue


remaining time [sec]: 36
Part I: Period 1/10
To Do
When you are ready, click "Continue" to finish this period.
Stage 1
Your input level: 40.00
Your output level: 7.50
The other's output level: 35.44
Difference between your and other's output level: -27.94
Stage 2
Your input level: 35.00
Your output level: 26.44
The other's output level: 11.62
Difference between your and other's output level: 14.82
Payments
The other has higher output in Stage 1, so the other receives 150 points.
You have higher output in Stage 2, so you receive 150 points.
Your total payments (in points): 150
Costs
Your costs for input in Stage 1 (in points): 52.80
Your costs for input in Stage 2 (in points): 40.42
Payoff
Your payoff from this period (in points): 256.78
Continue

References

Anderson, L. R. & Freeborn, B. A. (2010). Varying the intensity of competition in a multiple prize rent seeking experiment. *Public Choice*, 143(1-2), 237–254.

Azmat, G. & Iriberri, N. (2016). The provision of relative performance feedback: An analysis of performance and satisfaction. *Journal of Economics & Management Strategy*, 25(1), 77–110.

Balafoutas, L., Kerschbamer, R., & Sutter, M. (2012). Distributional preferences and competitive behavior. *Journal of Economic Behavior & Organization*, 83(1), 125–135.

Baye, M. R., Kovenock, D., & de Vries, C. G. (1996). The all-pay auction with complete information. *Economic Theory*, 8(2), 291–305.

Berger, J. & Pope, D. (2011). Can losing lead to winning? *Management Science*, 57(5), 817–827.

Brookins, P. & Ryvkin, D. (2014). An experimental study of bidding in contests of incomplete information. *Experimental Economics*, 17(2), 245–261.

Buser, T. (2016). The impact of losing in a competition on the willingness to seek further challenges. *Management Science*, 64, 3439–3449.

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *The Review of Economics and Statistics*, 90(3), 414–427.

Casas-Arce, P. & Martínez-Jerez, F. A. (2009). Relative performance compensation, contests, and dynamic incentives. *Management Science*, 55(8), 1306–1320.

Charness, G., Gneezy, U., & Kuhn, M. A. (2012). Experimental methods: Between-subject and within-subject design. *Journal of Economic Behavior & Organization*, 81(1), 1–8.

Chen, H., Ham, S. H., & Lim, N. (2011). Designing multiperson tournaments with asymmetric contestants: An experimental study. *Management Science*, 57(5), 864–883.

Chlosta, K., Pull, K., & Futagami, S. (2014). Tournament structures in Japan and the U.S.: Why are they different and will they change? *Management and Organizational Studies*, 1(1), 63–71.

Clark, D. J. & Nilssen, T. (2013). Learning by doing in contests. *Public Choice*, 156(1-2), 329–343.

Clark, D. J., Nilssen, T., & Sand, J. Y. (2012). Motivating over time: Dynamic win effects in sequential contests. University of Oslo, Department of Economics Working Paper No. 28/2012.

Cubitt, R. P., Starmer, C., & Sugden, R. (1998). On the validity of the random lottery incentive system. *Experimental Economics*, 1(2), 115–131.

Datta, S. & Satten, G. A. (2008). A signed-rank test for clustered data. *Biometrics*, 64(2), 501–507.

Dechenaux, E., Kovenock, D., & Sheremeta, R. M. (2015). A survey of experimental research on contests, all-pay auctions and tournaments. *Experimental Economics*, 18(4), 609–669.

Delfgaauw, J., Dur, R., Non, A., & Verbeke, W. (2015). The effects of prize spread and noise in elimination tournaments: A natural field experiment. *Journal of Labor Economics*, 33(3), 521–569.

Denter, P. & Sisak, D. (2015). Do polls create momentum in political competition? *Journal of Public Economics*, 130, 1–14.

Denter, P. & Sisak, D. (2016). Head starts in dynamic tournaments? *Economics Letters*, 149, 94–97.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., & Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, 9(3), 522–550.

Dufwenberg, M. & Kirchsteiger, G. (2004). A theory of sequential reciprocity. *Games and Economic Behavior*, 47(2), 268–298.

Ederer, F. & Fehr, E. (2017). Deception and incentives: How dishonesty undermines effort provision. Working paper, October 25, 2017.

Eriksson, T., Poulsen, A., & Villeval, M. C. (2009). Feedback and incentives: Experimental evidence. *Labour Economics*, 16, 679–688.

Falk, A. & Fischbacher, U. (2006). A theory of reciprocity. *Games and Economic Behavior*, 54(2), 293–315.

Fehr, E. & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics*, 114(3), 817–868.

Fehr, E. & Williams, T. (2013). Endogenous emergence of institutions to sustain cooperation. Unpublished working paper.

Feld, J. & Zölitz, U. (2017). Understanding peer effects: on the nature, estimation, and channels of peer effects. *Journal of Labor Economics*, 35(2), 387–428.

Fischbacher, U. (2007). z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2), 171–178.

Gächter, S., Johnson, E. J., & Herrmann, A. (2010). Individual-level loss aversion in riskless and risky choices. CeDEx Discussion Paper No. 2010-20.

Gill, D. & Prowse, V. (2014). Gender differences and dynamics in competition: The role of luck. *Quantitative Economics*, 5, 351–376.

Goldman, M. & Rao, J. M. (2017). Loss aversion around a fixed reference point in highly experienced agents. *Available at SSRN 2782110*.

Harbaugh, R. & Ridlon, R. W. (2011). Handicapping under uncertainty in an all-pay auction. Working paper, February 2011.

Harbring, C. & Irlenbusch, B. (2003). An experimental study on tournament design. *Labour Economics*, 10(4), 443–464.

Harris, C. & Vickers, J. (1985). Perfect equilibrium in a model of a race. *The Review of Economic Studies*, 52(2), 193–209.

Harris, P. B. & Houston, J. M. (2010). A reliability analysis of the revised competitiveness index. *Psychological Reports*, 106(3), 870–874.

Hernandez-Lagos, P., Minor, D., & Sisak, D. (2017). Do people who care about others cooperate more? experimental evidence from relative incentive pay. *Experimental Economics*, (pp. 1–27).

Houston, J., Harris, P., McIntire, S., & Francis, D. (2002). Revising the competitiveness index using factor analysis. *Psychological Reports*, 90(1), 31–34.

Klein, A. H. & Schmutzler, A. (2017). Optimal effort incentives in dynamic tournaments. *Games and Economic Behavior*, 103, 199–224.

Kong, X. (2008). Loss aversion and rent-seeking: An experimental study. CeDEx Discussion Paper No. 2008-13.

Konrad, K. A. (2012). Dynamic contests and the discouragement effect. *Revue d'économie politique*, 122, 233–256.

Konrad, K. A. & Kovenock, D. (2009). Multi-battle contests. *Games and Economic Behavior*, 66(1), 256–274.

Kőszegi, B. & Rabin, M. (2006). A model of reference-dependent preferences. *The Quarterly Journal of Economics*, 121(4), 1133–1165.

Kőszegi, B. & Rabin, M. (2007). Reference-dependent risk attitudes. *American Economic Review*, 97(4), 1047–1073.

Kubitz, G. (2020). Sharing cost information in dynamic oligopoly. Unpublished working paper.

Lim, N. (2010). Social loss aversion and optimal contest design. *Journal of Marketing Research*, 47(4), 777–787.

Llorente-Saguer, A., Sheremeta, R. M., & Szech, N. (2016). Designing contests between heterogenous contestants: An experimental study of tie-breaks and bid-caps in all-pay auctions. Working paper, May 8, 2016.

Mago, S. D., Sheremeta, R. M., & Yates, A. (2013). Best-of-three contest experiments: Strategic versus psychological momentum. *International Journal of Industrial Organization*, 31(3), 287–296.

March, C. & Sahm, M. (2017). Asymmetric discouragement in asymmetric contests. *Economics Letters*, 151, 23–27.

McGregor, J. (2013). For whom the bell curve tolls. *The Washington Post – On Leadership* [weblog]. Retrieved 2017/10/18, from http://www.washingtonpost.com/blogs/on-leadership/wp/2013/11/20/for-whom-the-bell-curve-tolls/.

Mermer, A. G. (2017). Effort provision and optimal prize structure in contests with loss-averse players. Working paper, April 5, 2017.

Meyer, M. A. (1992). Biased contests and moral hazard: Implications for career profiles. *Annales d'Économie et de Statistique*, 25/26, 165–187.

Millner, E. L. & Pratt, M. D. (1991). Risk aversion and rent-seeking: An extension and some experimental evidence. *Public Choice*, 69(1), 81–92.

Möller, M. (2012). Incentives versus competitive balance. *Economics Letters*, 117(2), 505–508.

Müller, W. & Schotter, A. (2010). Workaholics and dropouts in organizations. *Journal of the European Economic Association*, 8(4), 717–743.

Murphy, R. O., Ackermann, K. A., & Handgraaf, M. J. J. (2011). Measuring social value orientation. *Judgment and Decision Making*, 6(8), 771–781.

Orrison, A., Schotter, A., & Weigelt, K. (2004). Multiperson tournaments: An experimental examination. *Management Science*, 50(2), 268–279.

Price, C. R. & Sheremeta, R. M. (2011). Endowment effects in contests. *Economics Letters*, 111(3), 217–219.

Price, C. R. & Sheremeta, R. M. (2015). Endowment origin, demographic effects, and individual preferences in contests. *Journal of Economics & Management Strategy*, 24(3), 597–619.

Ridlon, R. & Shin, J. (2013). Favoring the winner or loser in repeated contests. *Marketing Science*, 32(5), 768–785.

Schmitt, P., Shupp, R., Swope, K., & Cadigan, J. (2004). Multi-period rent-seeking contests with carryover: Theory and experimental evidence. *Economics of Governance*, 5(3), 187–211.

Sheremeta, R. M. (2010). Experimental comparison of multi-stage and one-stage contests. *Games and Economic Behavior*, 68(2), 731–747.

Sheremeta, R. M. (2011). Contest design: An experimental investigation. *Economic Inquiry*, 49(2), 573–590.

Sheremeta, R. M., Masters, W. A., & Cason, T. N. (2017). Winner-take-all and proportional-prize contests: Theory and experimental results. Working paper, August 15, 2017.

Sheremeta, R. M. & Zhang, J. (2010). Can groups solve the problem of over-bidding in contests? *Social Choice and Welfare*, 35(2), 175–197.

Shupp, R., Sheremeta, R. M., Schmidt, D., & Walker, J. (2013). Resource allocation contests: Experimental evidence. *Journal of Economic Psychology*, 39, 257–267.

Stracke, R., Höchtl, W., Kerschbamer, R., & Sunde, U. (2014). Optimal prizes in dynamic elimination contests: Theory and experimental evidence. *Journal of Economic Behavior & Organization*, 102, 43–58.

Tong, K. & Leung, K. (2002). Tournament as a motivational strategy: Extension to dynamic situations with uncertain duration. *Journal of Economic Psychology*, 23, 399–420.

Tversky, A. & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4), 297–323.

Welch, J. (2005). The vitality curve. *Leadership Excellence*, 22(9), 4–5.

Zizzo, D. J. (2002). Racing with uncertainty: a patent race experiment. *International Journal of Industrial Organization*, 20(6), 877–902.