Online Appendix for "Do We All Coordinate in the Long Run?"

Manja Gärtner Robert Östling Sebastian Tebbe

Abstract

This Online Appendix provides supplementary material for the paper "Do We All Coordinate in the Long Run?". Sections A and B contain additional tables and figures referenced in the main text. Section C contains a theoretical analysis of repeated-game strategies. Section D contains a power analysis. Sections E and F contain additional details about the preference elicitation and construction of the preference indices. Section G contains additional analyses of the correlation with socio-demographic characteristics and behavioral traits as well as analyses of the open-ended questions in the post-experimental survey. Section H contains complete instructions for the experiment.

A Online Appendix Tables

Table A1: Differences in experimental design

	Reviewed studies		
	This study	Berninghaus and Ehrhart (1998)	Engelmann and Normann (2010)
Period length	10/50	10/30/90	10
Stake size per round [in €]	(1)+(2): 0-0.60 (3)+(4): 0-0.12	(1): 0.23-1.46 (2): 0.08-0.49 (3): 0.03-0.16	(1)-(5): 0.13-1.74
Benefit-cost ratio $\left(\frac{a}{c}\right)$	1.5	2	2
Fixed-variable payoff $\left(\frac{E}{a}\right)$	3.67	5	3
Subject pool	University of Copenhagen	University of Karlsruhe	University of Copenhagen
Default option	No default	Previous period	No default
Feedback mechanism	Minimum effort	Minimum effort	All/Minimum effort
Number of participants	312	144	170
Group size	4	8	4/6/9
Scale of action values	1-7	1-9	1-7
Number of treatments	4	3	5
Number of groups (per treatment)	17-21	6	4-8
Total duration [in min]	45	40	45
Year of experiment	2013	Before 1998	2006
Decision periods [in sec]	15	90/30/10	-
Revisions allowed	Once	Infinite	-
Average payoff [in €]	17.12	10.23	17.50

Table A2: Number of subjects across treatments and sessions

	Short horizon		Long horizon	
Session	High stakes	Low stakes	High stakes	Low stakes
0	24	-	28	-
1	12	12	8	12
2	8	12	8	12
3	12	8	12	8
4	-	12	12	8
5	12	12	-	24
6	-	-	16	12
7	16	12	-	-
Total	84	68	84	76

Table A3: Donations and profits across sessions (in DKK)

Session	Donation	Average donation	Total profit	Average profit	Observations
0	619.25	11.91	8140.75	156.55	52
1	641.70	14.58	4804.40	109.19	44
2	322.20	8.06	4874.60	121.87	40
3	310.70	7.77	5914.30	147.86	40
4	281.00	8.78	3988.40	124.64	32
5	386.15	8.04	5047.75	105.16	48
6	312.35	11.16	3430.65	122.52	28
7	385.10	13.75	3704.10	132.29	28

Table A4: Distribution of group minimum effort ($\underline{e}$) in the last period

Treatment	Distribution of $\underline{e}$							Average
	1	2	3	4	5	6	7	$\underline{e}$
Short horizon - Low stakes (17)	13	1	1	0	0	0	2	1.88
Short horizon - High stakes (21)	15	0	2	0	2	1	1	2.10
Long horizon - Low stakes (19)	13	0	3	0	1	0	2	2.16
Long horizon - High stakes (21)	14	1	1	2	0	1	2	2.24

Notes: The table illustrates the distribution of the groups' final-period minimum efforts and the average minimum effort level across subsets of treatments. The minimum effort level is taken from period 10 for the short horizon and from period 50 for the long horizon. The number in parentheses after each treatment represents the respective number of groups.
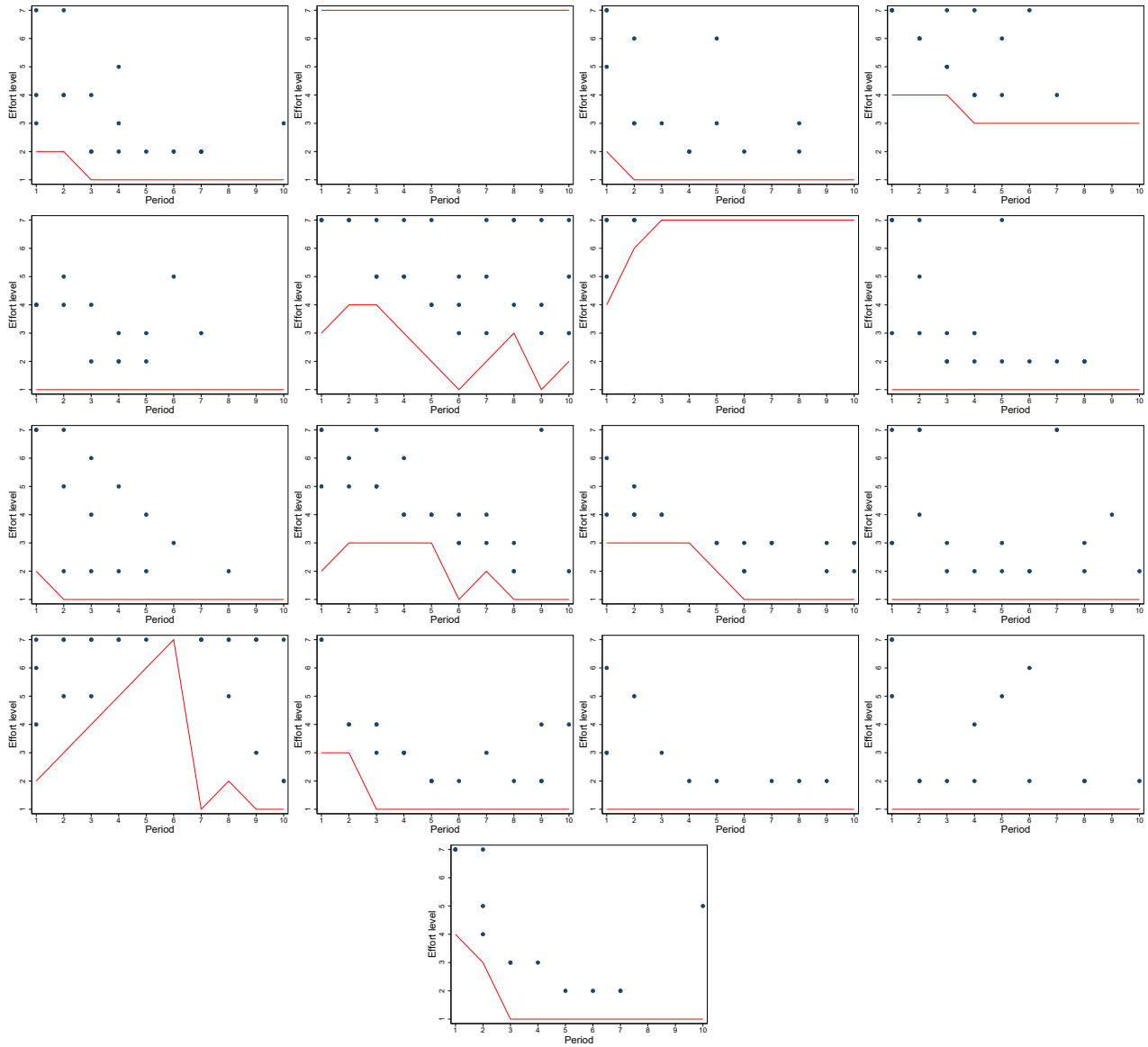
B   Online Appendix Figures

Figure A1: Short horizon - low stakes

Notes: The above figure displays the group minimum, indicated by the red line, for the treatment with short horizon and low stakes. Individual effort deviations from the minimum are represented by blue dots.
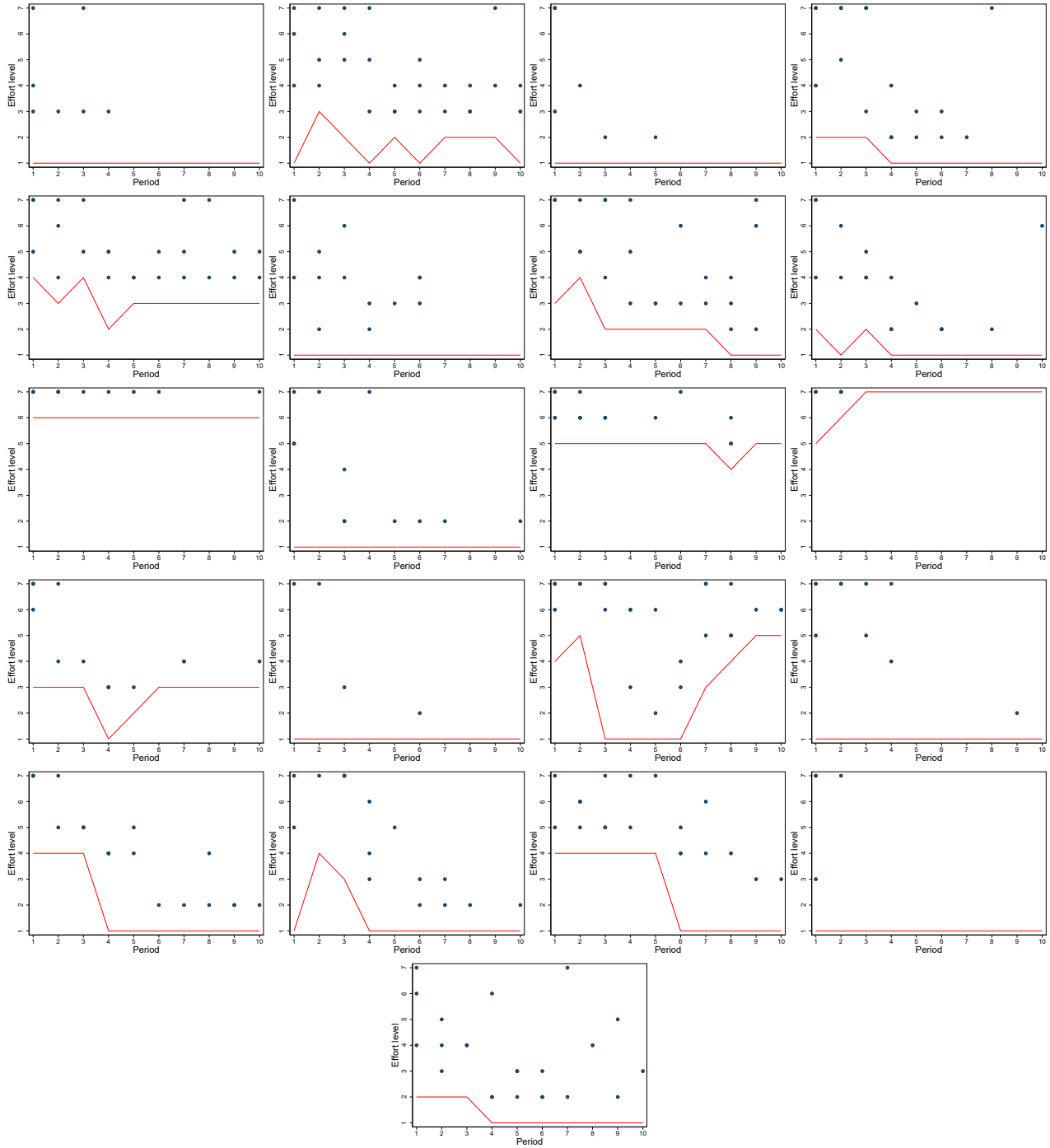
Figure A2: Short horizon - high stakes

Notes: The above figure displays the group minimum, indicated by the red line, for the treatment with short horizon and high stakes. Individual effort deviations from the minimum are represented by blue dots.
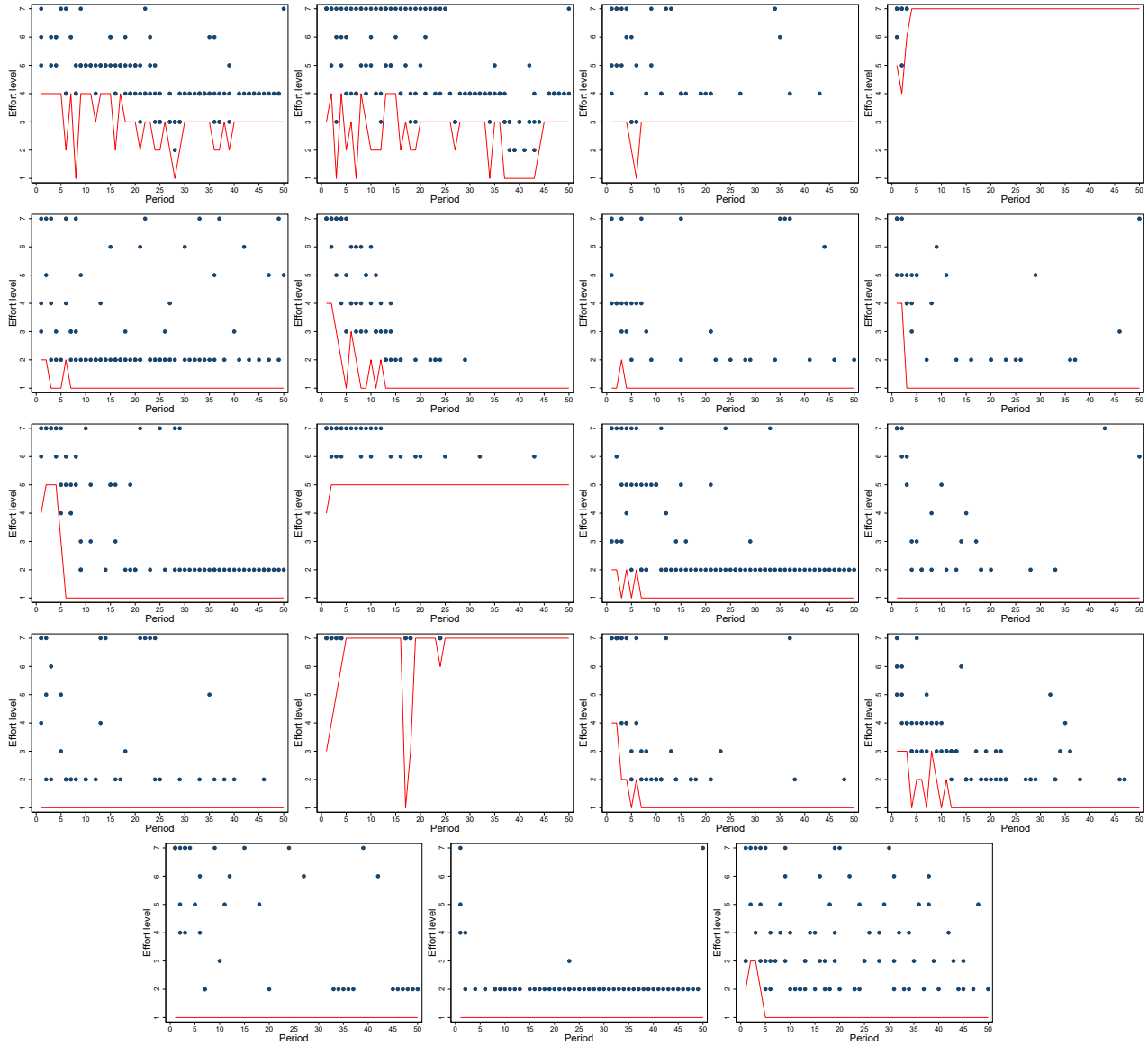
Figure A3: Long horizon - low stakes

Notes: The above figure displays the group minimum, indicated by the red line, for the treatment with long horizon and low stakes. Individual effort deviations from the minimum are represented by blue dots.
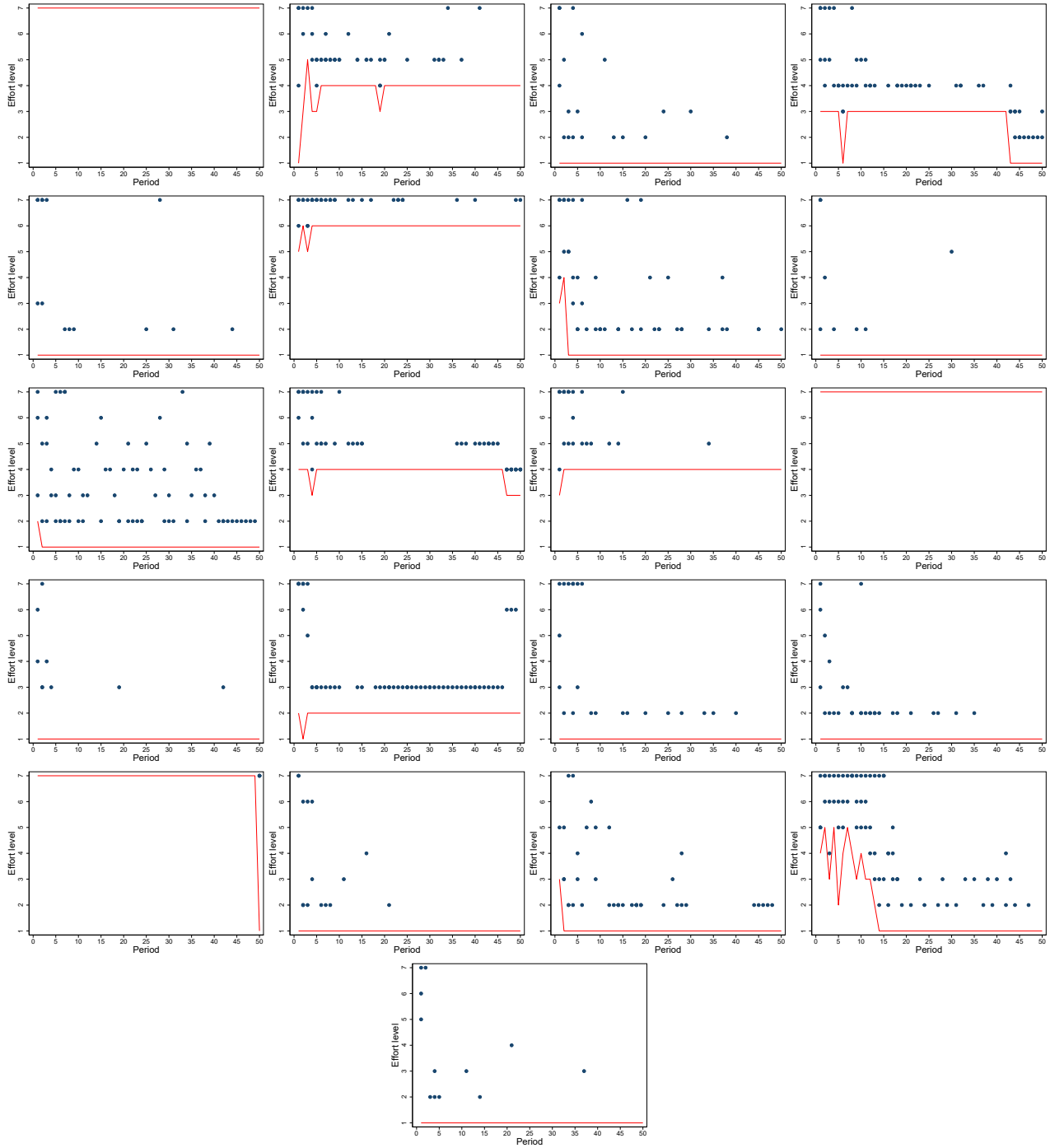
Figure A4: Long horizon - high stakes

Notes: The above figure displays the group minimum, indicated by the red line, for the treatment with long horizon and high stakes. Individual effort deviations from the minimum are represented by blue dots.

C Repeated-game Strategies in the Presence of Strategic Uncertainty

The finitely repeated weak-link game has a very large set of equilibria. In this section, we illustrate how the number of repetitions can affect the expected payoff of different repeated-game strategies in the presence of strategic uncertainty. For simplicity we focus on the following two strategies. Both strategies prescribe play of the same effort level $e < 7$ in the first period, but differ in what players do in the following periods. One strategy, the *minimum strategy*, prescribes play of the previous period's minimum from the second period to the end of the game. The other strategy, the *above-minimum strategy* prescribes play of 7 from the second round for $s < T - 2$ rounds, and then play of the previous period's minimum until the end of the game. Note that both strategies constitute equilibria of the repeated game because both will result in play of a stage game equilibrium in all periods if all players follow the same strategy.

Let $\pi(e_i|\underline{e})$ be the payoff of the stage game to a player when playing e_i and the group minimum is $\underline{e}$. We assume that there are two types of players corresponding to each of the two equilibrium strategies. The probability that a player plays the above-minimum strategy is p and players are sampled independently from an infinite underlying population.

The minimum strategy results in a payoff of $\pi(e|e)$ in every round irrespective of which of the two strategies the other players are playing. The payoff of the above-minimum strategy depends on what the other players do, but because the payoff only depends on the group minimum, there are only two relevant states: at least one other player plays the minimum strategy or all other players play the above-minimum strategy. The probability of each state is $(1 - p^3)$ and p^3, respectively. The expected payoff from playing the above-minimum strategy therefore is

$$(1 - p^3)\left[\pi(e|e) + \pi(7|e)s + \pi(e|e)(T - 1 - s)\right] + p^3\left[\pi(e|e) + \pi(7|7)(T - 1)\right].$$

In words, when faced with three above-minimum players, the outcome is the efficient equilibrium from the second period to the end of the game. When at least one of the other players is a minimum player, however, the payoff is instead the deviation payoff $\pi(7|e)$ for s periods until play returns to the initial-period equilibrium.

The expected payoff of the above-minimum strategy is higher than the minimum strategy whenever

$$(1 - p^3)\left[\pi(e|e) + \pi(7|e)s + \pi(e|e)(T - 1 - s)\right] + p^3\left[\pi(e|e) + \pi(7|7)(T - 1)\right] > \pi(e|e)T.$$

Solving for p^3 gives:
$$p^3 > \frac{\left[\pi(e|e) - \pi(7|e)\right]s}{\pi(7|7)(T - 1) - \left[\pi(7|e)s + \pi(e|e)(T - 1 - s)\right]}.$$

The expected payoff of the above-minimum strategy is higher than the minimum strategy if the probability that others play the above-minimum strategy is sufficiently high. The numerator shows that the threshold depends on how costly it is to deviate from the first-period equilibrium relative to the first-period payoff multiplied by the number of rounds of possible deviations (s). Rewriting the denominator as

$$\left[\pi(7|7) - \pi(7|e)\right]s + \left[\pi(7|7) - \pi(e|e)\right](T - 1 - s),$$

we see that the denominator is increasing in T, showing that the threshold p is decreasing in T. This implies

that the above-minimum strategy becomes relatively more attractive the longer the stage game is repeated in the sense that it is a better reply than the minimum strategy for a larger set of initial beliefs about the opponents. Note also that the threshold does not depend on stakes as multiplying all stage-game payoffs by a constant does not change the threshold.

D   Power and Optimal Sample Size

We calculate statistical power at the group level and we focus on our short horizon vs long horizon treatments. Under several simplifying assumptions, List et al. (2011) show that the minimal detectable treatment effect ($\hat{\delta}$) from a design with two treatments, significance level α and power $1 - \beta$ is

$$\hat{\delta} = (t_{\alpha/2} + t_\beta)\sqrt{\frac{\sigma_0^2}{n_0} + \frac{\sigma_1^2}{n_1}}. \tag{1}$$

Based on a significance level of $\alpha = 0.05$ and setting statistical power to $1 - \beta = 0.80$, we have $t_{\alpha/2} = 1.96$ and $t_\beta = 0.84$ using standard normal tables. We can also use the equation (1) to back out the statistical power of our experiment for effect size δ:

$$t_\beta = \frac{\delta}{\sqrt{\frac{\sigma_0^2}{n_0} + \frac{\sigma_1^2}{n_1}}} - t_{\alpha/2}. \tag{2}$$

To calculate the statistical power of our experiment, we retrieved data from Table I in Berninghaus and Ehrhart (1998). They report three different summary statistics at the group level: best minimum action, average minimum action and the average action. Since the long horizon in our experiment lies in-between the medium (M) and long horizon (L) treatments of Berninghaus and Ehrhart (1998), we use their results for the medium and long horizons (compared to their short horizon treatment) as bounds on expected treatment effects in our experiment. The relevant data from Berninghaus and Ehrhart (1998) and our power calculations are displayed in Table A5.

Given the sample size in our experiment, we can detect the effect size from Berninghaus and Ehrhart (1998) when increasing the horizon from 10 to 30 rounds ($\delta_M = 4.5$) with at least 94% power and the effect size when comparing 10 to 90 rounds ($\delta_L = 5.66$) with more than 99.9% power. Table A5 also show the minimal effect size we can detect with 80% power. The minimal detectable effect size lies between 1.26 to 1.77 depending on which measure of group effort is used and which treatments from Berninghaus and Ehrhart (1998) are used to compute the standard deviations.

Because Berninghaus and Ehrhart (1998) simultaneously varied both stakes and the horizon, a more conservative approach is to only compare our short-horizon-high-stakes treatment to our treatment with long horizon and low stakes. This results in a statistical power of at least 99.9% based on the effect size from the 90-round treatment, and a power of 71% to more than 99.9% for the 30-round treatment depending on the measure of group effort that is used.

Table A5: Summary of statistical power calculations

	Medium horizon (30 repetitions)			Long horizon (90 repetitions)		
	Best minimum effort $\underline{e}^B$	Average group effort $\bar{e}$	Minimum group effort $\underline{e}$	Best minimum effort $\underline{e}^B$	Average group effort $\bar{e}$	Minimum group effort $\underline{e}$
Mean μ_S	2.67	3.91	2.32	2.67	3.91	2.32
Mean $\mu_{M/L}$	7.17	5.84	4.84	8.33	7.48	6.72
Effect $\delta_{M/L}$	4.50	1.93	2.52	5.66	3.57	4.40
St. Dev. σ_S	3.14	2.33	2.80	3.14	2.33	2.80
St. Dev. $\sigma_{M/L}$	2.23	2.52	2.77	1.63	1.55	1.75
Statistical power $1-\beta$	1.00	0.94	0.98	1.00	1.00	1.00
Minimal detectable effect size $\hat{\delta}$	1.74	1.54	1.77	1.60	1.26	1.49

E   Design of Preference Elicitation

Risk preferences: Following Dohmen et al. (2011) individual risk attitude questionnaire, we elicit two variables of subjects' risk aversion. The first measure (Q2) asks subjects about their general willingness to take a risk on a scale from zero to 10 (*Risk aversion I: General Risk Attitude*), whereas the second (Q3) is a hypothetical lottery question, which lets subjects decide how much of their 100,000 DKK lottery win they would invest in an asset that either doubles or only returns half of the initial investment with equal probability (*Risk aversion II: Investment risk level*). Both variables are designed such that a higher value is associated with more risk-aversion.

Patience: Using a slightly modified version of the multiple price list presented in Andersen et al. (2008), participants faced a choice between 1,000 DKK in one month (A) and (1,000+Y) DKK in seven months (B). The amount of (1,000+Y) DKK began at 1,025 DKK and rose in 50 DKK increments to 1,275 DKK with six rows in total. Our patience measure (*Patience*) reflects the number of items, in which the participant prefers the option further away in the future.[10] A high patience measure represents an individual that is willing to postpone their payoff for little monetary incentives.

Negative reciprocity: Subjects' inclination for (negative) reciprocity is measured twofold. Based on Dohmen et al. (2009) survey on behavioral evidence, the first proxy (Q5b, Q5c, Q5e) asked subjects on a 7-point scale how well different statements apply to their personal inclination of negative reciprocity (*Negative Reciprocity),* while the second measure approximates reciprocity through the minimum accepted amount in an Ultimatum Game (*Minimum Accept*). In the ultimatum game, subjects are asked to specify how to split 40 DKK ($\approx$ \$6.31) between themselves and another participant and to specify the minimum amount that they would accept in the role of the recipient. Pairs of subjects and ultimatum game roles are randomly determined and subjects are paid accordingly.

Altruism: Utilizing the responses to the payoff-relevant tasks, we measure the subjects' altruism through the donation to a medical humanitarian organization (*Donation*) and the proposed offer in the Ultimatum Game (*Offer*). The direct influence of individuals' altruism on the probability to agree upon an efficient equilibrium in coordination games is as yet an open question.

Trust: Adapted from Fehr et al. (2003), the level of generalized trust (*Trust*) is based on a three-item list, for which subjects have to indicate their level of agreement on a 4-point scale. People are asked about their general trust in people (Q4a), their ability to rely on others (Q4b) and trust in strangers (Q4c). The trust variable is then simply computed as the average of the reported trust level across the three responses.

Selfishness: The selfishness proxy (*Selfishness*) asks subjects if they primarily think about their payoff or the group payoff as a whole (Q14) on a 7-point scale.

Cognitive ability: We adopted the three-item "Cognitive Reflection Test" (CRT hereafter; Q6-Q8) from Frederick (2005) as a measure for cognitive processing (*Cognitive ability*).[11] The CRT is designed to quantify

---

[10] An alternative way of measuring patience is discussed in Brockhoff et al. (2015).

[11] Note that the CRT is predictive of various measures of mental heuristics including the gambler's fallacy, understanding of the regression to the mean, the sunk cost fallacy as well as for many measures of economic thinking such as temporal discounting,

people's tendency to override an erroneous gut feeling and engage in more extensive reflection to detect a more deliberate, correct answer. Thus, the correct response requires the activation of system 2 to overrule the immediate intuitive system 1, which would predict an incorrect answer. In our experimental setting, the CRT score simply reflects the percentage of correctly solved questions.

risk preferences, numeracy and gambling preferences (Hoppe and Kusterer, 2011; Oechssler et al., 2009; Szaszi et al., 2017; Toplak et al., 2011).

F Construction of Preference Indices

Given that most behavioral characteristics include multiple questions related to a single behavior and to keep the number of control variables limited, we construct indices for several groups following the classification of behavioral variables in Falk et al. (2016) and Falk et al. (2018). The computation of the behavioral preference indicx weights follows the procedure proposed in Anderson (2008) and is reproduced below:

First, each behavioral trait b_{jk}, where the index j represents the single behavioral preference and k the questions related to the behavior, is re-coded such that higher factors correspond to a positive output.

Next, the covariance matrix $\hat{\sum}_j$ within each outcome group of the behavioral traits j is calculated, which consists of the entries:

$$\hat{\sum}_{jmn} = \sum_{i=1}^{N_{jmn}} \frac{b_{ijm} - \bar{b}_{jm}}{\sigma_{jm}^b} \frac{b_{ijn} - \bar{b}_{jn}}{\sigma_{jn}^b}, \tag{3}$$

where N_{jmn} is the number of non-missing observations for questions m and n of the behavioral preference j. The means for outcomes m and n for the behavioral preference j are denoted by $\bar{b}_{jm}$ and $\bar{b}_{jn}$ and respectively, the standard deviations are σ_{jm}^b and σ_{jn}^b for the same outcomes.

In a third step, the covariance matrix $\hat{\sum}_j$ is inverted and the Anderson weight w_{jk} is defined for each question k of the behavioral preference group j by summing the elements in the row of the inverted covariance matrix according to that outcome:

$$\hat{\sum}_j^{-1} = \begin{bmatrix} c_{j11} & c_{j12} & \cdots & c_{j1K_j} \\ c_{j21} & c_{j22} & \cdots & \cdots \\ \vdots & \vdots & \ddots & \\ c_{jK_j1} & \vdots & & c_{jK_jK_j} \end{bmatrix}, \tag{4}$$

$$w_{jk} = \sum_{l=1}^{K_j} c_{jkl}, \tag{5}$$

where K_j is the total number of questions related to the behavioral preference group j. Then, each behavioral preference is constructed by subtracting its mean, dividing by the groups' standard deviation and weighting it with the respective weights obtained through the procedure described above. The preference indices are denoted by $\hat{b}_{ij}$ as the transformation results in a generalized least squares estimator (Anderson, 2008).

$$\hat{b}_{ij} = \left(\sum_{k \in K_{ij}} w_{jk} \right)^{-1} \sum_{k \in K_{ij}} w_{jk} \frac{y_{ijk} - \bar{y}_{jk}}{\sigma_{jk}^b}, \tag{6}$$

where K_{ij} indicates the set of non-missing outcomes for observation i in the behavioral preference group j. Finally, we standardize each preference indix $\hat{b}_{ij}$.

G   Additional Analyses of Behavioral Correlates and Open-ended Responses

G.1   Socio-demographic characteristics and behavioral traits

Unless players coordinate in the first period, successful coordination requires that players play above the minimum effort level from the previous period. To analyze this in more depth, Figure A5 displays the share of effort choices e_i^t below, equal to, or above the group minimum of the previous period ($\underline{e}^{t-1}$) across the four treatment groups over time. The pattern during the first 10 or 50 rounds is similar across the four treatments. The share of attempts to raise the minimum falls quickly while the share of choices equal to the previous period's minimum increases. Very few choices are below the previous period's minimum. The pattern is repeated each time the game is restarted. If a larger fraction of players were willing to play above the minimum, successful coordination would have been more likely.

(a) Short horizon - Low stakes

(b) Short horizon - High stakes

(c) Long horizon - Low stakes
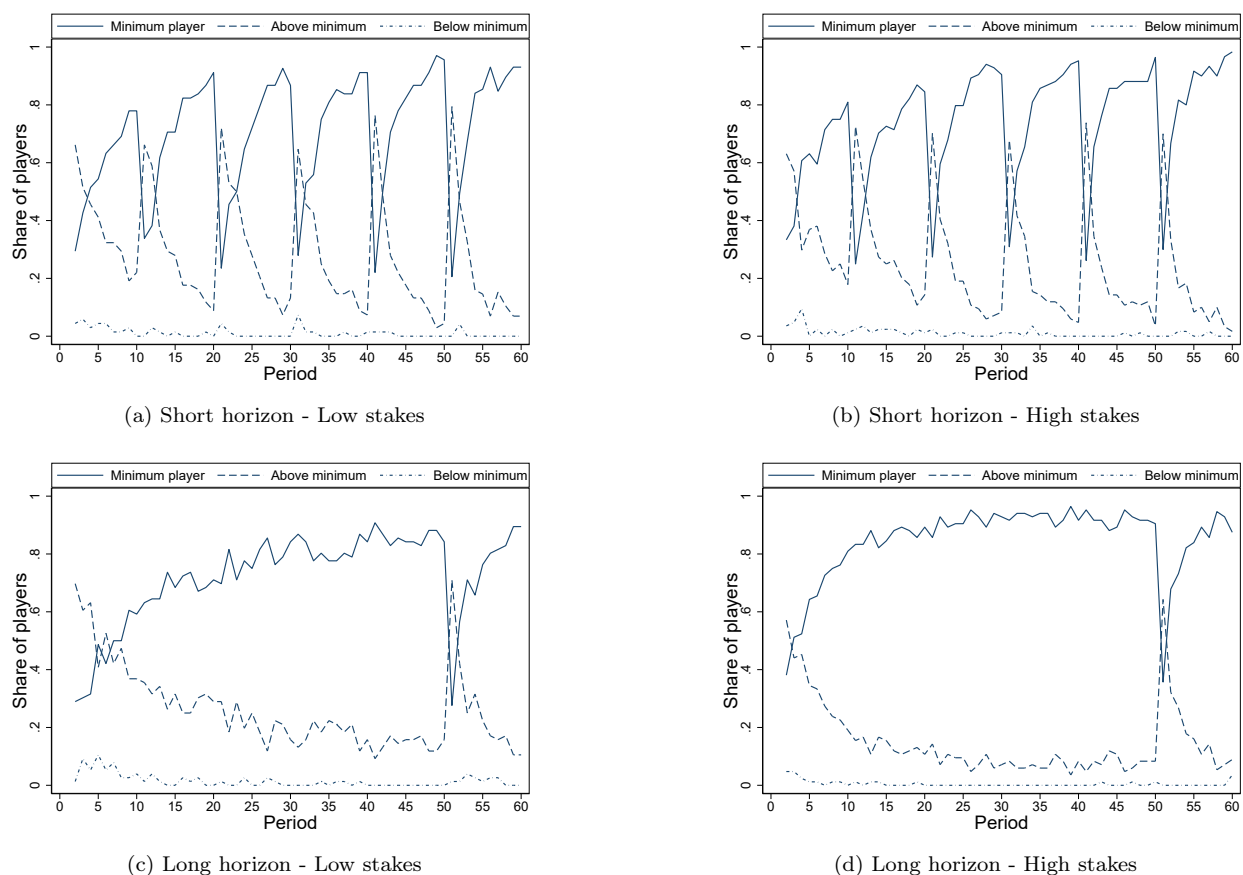
(d) Long horizon - High stakes

Figure A5: Share of effort choices above, below or at the previous period's minimum

Because play above the previous group minimum is required for successful coordination, we further analyze which individual characteristics predict persistent play above the group minimum by regressing an indicator variable for above-minimum play on individual characteristics. The results are shown in Table A6. The dependent variables indicate whether a player played above the previous period's minimum in all

periods up to t. For example, the dependent variable in the second column takes the value 1 if $e_i^2 > \underline{e}^1$ and $e_i^3 > \underline{e}^2$, and 0 otherwise. Among the behavioral traits, risk aversion appears most robustly correlated with persistent above minimum play. More risk averse subjects are less likely to play above minimum, although the coefficient varies across specifications. Trust is predictive for subjects to play above the minimum from period 2 to 5. Intuitively, more risk-averse subjects want to avoid the risk of choosing a higher effort level and more trusting subjects may have greater faith in the other group members to also increase their effort level.

Table A6: Determinants of above-minimum play

	Playing Above Minimum			
	(1)	(2)	(3)	(4)
Treatment				
High stakes	-0.026	0.075	-0.029	-0.083
	[0.096]	[0.105]	[0.090]	[0.069]
Long horizon	0.029	0.047	0.108	0.047
	[0.105]	[0.114]	[0.104]	[0.085]
(High stakes) x (Long horizon)	-0.064	-0.177	-0.055	0.031
	[0.142]	[0.142]	[0.129]	[0.103]
Socio-demographics				
Danish	-0.098	-0.042	-0.068	-0.081
	[0.071]	[0.079]	[0.070]	[0.051]
Male	0.052	0.175***	0.159***	0.135***
	[0.057]	[0.054]	[0.046]	[0.037]
Economics	-0.083	-0.105	-0.115**	-0.157***
	[0.058]	[0.063]	[0.054]	[0.044]
Experiment	-0.043	-0.005	0.059	0.101**
	[0.060]	[0.061]	[0.050]	[0.045]
Game theory	-0.068	-0.099	-0.089	-0.015
	[0.061]	[0.071]	[0.057]	[0.047]
Age	0.061**	0.024	0.037	0.020
	[0.025]	[0.033]	[0.032]	[0.026]
Parental education	0.007	0.036	0.032	0.010
	[0.030]	[0.032]	[0.028]	[0.021]
Cognitive ability	0.033	0.008	0.039	0.031
	[0.035]	[0.034]	[0.028]	[0.022]
Behavioral traits				
Risk aversion	-0.057**	-0.017	-0.056**	-0.008
	[0.027]	[0.027]	[0.024]	[0.020]
Negative reciprocity	-0.023	0.040	0.003	0.003
	[0.028]	[0.032]	[0.024]	[0.022]
Altruism	0.010	-0.015	0.007	0.043
	[0.029]	[0.028]	[0.029]	[0.027]
Trust	0.013	0.036	0.026	0.051**
	[0.031]	[0.031]	[0.027]	[0.024]
Patience	0.015	-0.021	-0.017	-0.020
	[0.032]	[0.031]	[0.027]	[0.020]
Treatment level	Individual	Individual	Individual	Individual
Period	2	3	4	5
Observations	295	295	295	295
Adjusted R^2	0.015	0.022	0.066	0.103

Notes: OLS regression for consistently playing an effort level above the previous period's minimum, in all periods up to 2, 3, 4, and 5 (in columns 1, 2, 3, and 4, respectively), conditioning the specified set of treatment variables, socio-demographic factors and behavioral traits on the individual level. "Economics", "Experiment" and "Game theory" refer to dummy variables indicating whether the subject reported having taken courses in economics, had previously participated in similar experiments or had taken classes in game theory. Significance level of $p<0.1$, $p<0.05$ and $p<0.01$ are indicated by */**/***. Standard errors clustered on the group level are reported within parentheses.

Whereas risk aversion and trust are correlated both with initial period choices and adjustments in the expected direction, the pattern is less consistent for socio-demographic characteristics. For example, whereas male subjects tended to make lower initial effort choices, they are more likely to play above the minimum. There is no statistically significant relationship between above-minimum play and cognitive ability. Unsurprisingly given the graphical evidence shown above, there are no statistically significant differences across treatments. We obtain very similar results when regressing the above-minimum dummies on socio-demographic factors (Table A7) and behavioral traits (Table A8) separately. We also find that the socio-demographic characteristics have considerably higher explanatory power than behavioral traits in explaining above-minimum play.

Table A7: Socio-demographic factors of playing above minimum

| | Playing Above Minimum | | | |
	(1)	(2)	(3)	(4)
Treatment				
High stakes	-0.016	0.062	-0.027	-0.095
	[0.094]	[0.106]	[0.090]	[0.074]
Long horizon	0.024	0.050	0.102	0.029
	[0.098]	[0.112]	[0.103]	[0.086]
(High stakes) x (Long horizon)	-0.084	-0.175	-0.068	0.027
	[0.139]	[0.140]	[0.126]	[0.102]
Socio-demographics				
Danish	-0.068	-0.046	-0.061	-0.050
	[0.067]	[0.072]	[0.066]	[0.045]
Male	0.069	0.193***	0.178***	0.139***
	[0.058]	[0.054]	[0.045]	[0.038]
Economics	-0.084	-0.111*	-0.118**	-0.165***
	[0.058]	[0.062]	[0.053]	[0.045]
Experiment	-0.066	-0.010	0.042	0.082*
	[0.057]	[0.060]	[0.049]	[0.046]
Game theory	-0.064	-0.098	-0.094*	-0.045
	[0.064]	[0.071]	[0.057]	[0.047]
Age	0.066***	0.026	0.044	0.028
	[0.024]	[0.032]	[0.031]	[0.026]
Parental education	0.008	0.031	0.031	0.010
	[0.030]	[0.032]	[0.028]	[0.021]
Cognitive ability	0.035	0.008	0.037	0.035
	[0.035]	[0.034]	[0.029]	[0.022]
Treatment level	Individual	Individual	Individual	Individual
Period	2	3	4	5
Observations	298	298	298	298
Adjusted R^2	0.016	0.029	0.064	0.090

Notes: OLS regression for consistently playing an effort level above the previous period's minimum, in all periods up to 2, 3, 4, and 5 (in columns 1, 2, 3, and 4, respectively), conditioning the specified set of treatment variables and socio-demographic factors on the individual level. "Economics", "Experiment" and "Game theory" refer to dummy variables indicating whether the subject reported having taken courses in economics, had previously participated in similar experiments or had taken classes in game theory. Significance level of $p<0.1$, $p<0.05$ and $p<0.01$ are indicated by */**/***. Standard errors clustered on the group level are reported within parentheses.

Table A8: Behavioral traits of playing above minimum

	Playing Above Minimum			
	(1)	(2)	(3)	(4)
Treatment				
High stakes	-0.025	0.051	-0.009	-0.060
	[0.091]	[0.095]	[0.083]	[0.066]
Long horizon	0.049	0.035	0.145	0.080
	[0.096]	[0.102]	[0.095]	[0.085]
(High stakes) x (Long horizon)	-0.096	-0.142	-0.075	-0.003
	[0.135]	[0.131]	[0.117]	[0.099]
Behavioral traits				
Risk aversion	-0.060**	-0.020	-0.053**	-0.008
	[0.028]	[0.028]	[0.026]	[0.023]
Negative reciprocity	-0.027	0.040	0.003	0.008
	[0.028]	[0.033]	[0.027]	[0.022]
Altruism	0.011	-0.012	0.005	0.038
	[0.027]	[0.028]	[0.029]	[0.026]
Trust	0.010	0.053*	0.033	0.047**
	[0.029]	[0.028]	[0.026]	[0.022]
Patience	0.005	-0.026	-0.024	-0.024
	[0.031]	[0.031]	[0.028]	[0.022]
Treatment level	Individual	Individual	Individual	Individual
Period	2	3	4	5
Observations	308	308	308	308
Adjusted R^2	0.002	0.000	0.015	0.023

Notes: OLS regression for consistently playing an effort level above the previous period's minimum, in all periods up to 2, 3, 4, and 5 (in columns 1, 2, 3, and 4, respectively), conditioning the specified set of treatment variables and behavioral traits on the individual level. Significance level of $p<0.1$, $p<0.05$ and $p<0.01$ are indicated by */**/***. Standard errors clustered on the group level are reported within parentheses.

Finally, we regress the final-period group minimum (period 10 in the short horizon and 50 in the long horizon treatment) on behavioral and socio-demographic variables. We perform these regressions at the group level because there is a strong interdependence of choices within groups after the initial periods. Table A9 shows the results when regressing the final-period group minimum on individual characteristics averaged at the group level. The strongest predictor of final-period group minimum effort is the average cognitive ability of the group. A group with a high average cognitive ability is more likely to coordinate on a better equilibrium, which is consistent with higher initial-period choices of cognitively able subjects. To get an indication of the magnitude, the coefficient in the fourth column implies that a one standard deviation increase of all group member's cognitive ability (approximately an increase of the average cognitive ability of four standard deviations) raises the group effort by 1.6 points.

Table A9: Multiple regression on group final period effort

| | Final Period Effort | | | |
	(1)	(2)	(3)	(4)
Treatment				
High stakes	0.052	-0.329	0.119	-0.248
	[0.639]	[0.654]	[0.619]	[0.635]
Long horizon	0.265	-0.261	0.188	-0.247
	[0.654]	[0.703]	[0.749]	[0.772]
(High stakes) x (Long horizon)	-0.123	0.288	-0.086	0.236
	[0.917]	[0.975]	[0.979]	[1.040]
Socio-demographics				
Danish		-1.362		-1.450
		[1.124]		[1.251]
Male		1.024		1.147
		[1.041]		[1.122]
Economics		-1.225		-1.311
		[0.966]		[1.081]
Experiment		0.253		0.427
		[1.026]		[1.249]
Game theory		0.459		1.110
		[0.896]		[1.058]
Age		0.088		0.081
		[0.526]		[0.596]
Parental education		-0.815*		-0.772
		[0.469]		[0.494]
Cognitive ability		1.478***		1.598***
		[0.470]		[0.500]
Behavioral traits				
Risk aversion			-0.009	0.062
			[0.497]	[0.564]
Negative reciprocity			0.095	0.502
			[0.397]	[0.447]
Altruism			-0.104	0.267
			[0.401]	[0.433]
Trust			0.338	0.203
			[0.535]	[0.575]
Patience			-0.382	-0.135
			[0.485]	[0.456]
Treatment level	Group	Group	Group	Group
N(Groups)	78	78	78	78
Period	10/50	10/50	10/50	10/50
Adjusted R^2	-0.038	0.034	-0.099	-0.018

Notes: OLS regressions of the final period effort level conditioning on the treatment variables (column 1), in addition to the socio-demographic factors (column 2), behavioral attributes (column 3) and all of the controls (column 4) on the group level. "Economics", "Experiment" and "Game theory" refer to dummy variables indicating whether the subject reported having taken courses in economics, had previously participated in similar experiments or had taken classes in game theory. Significance level of p<0.1, p<0.05 and p<0.01 are indicated by */**/***. Heteroscedasticity-robust standard errors are reported within parentheses.

To further illustrate the correlation between individual characteristics and final-period group coordination and to verify that Table A9 does not mask any strong non-linear effects, Figure A6 shows the distribution of final-period group minima plotted against treatment indicators and individual characteristics. Consistent with the findings from Table A9, it is only cognitive ability that appears related to more successful coordination ($\rho = .259$, $p = .022$). Notably, there is no positive correlation between the share of Danes in a group and the final-period group minimum, which contrasts with Engelmann and Normann (2010). However, their results were largely driven by groups with at least 75% Danes, whereas we had very few such groups (no group with 100% Danes and six groups with 75% Danes, but all the latter had a group minimum of 1 in the last period).
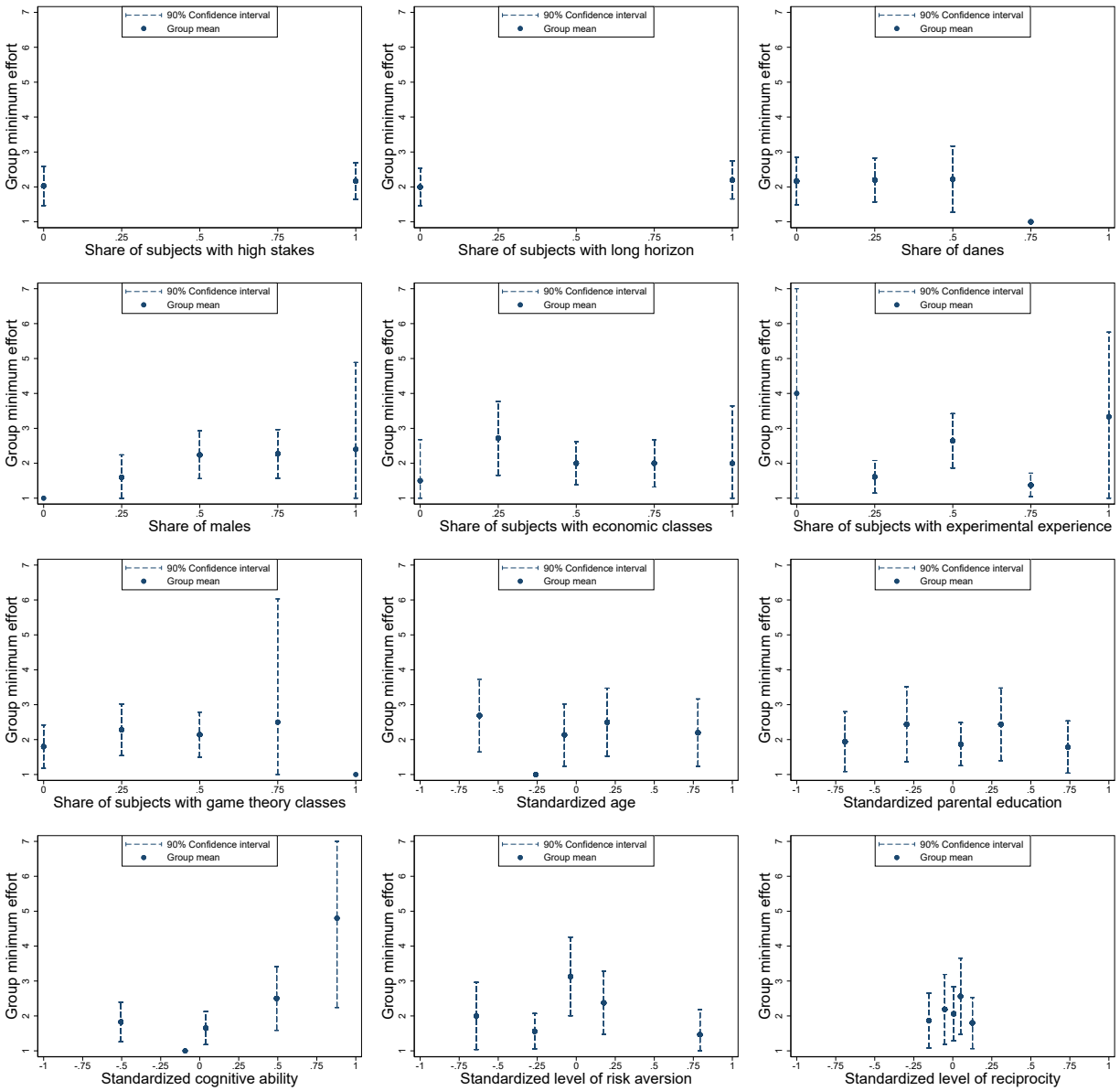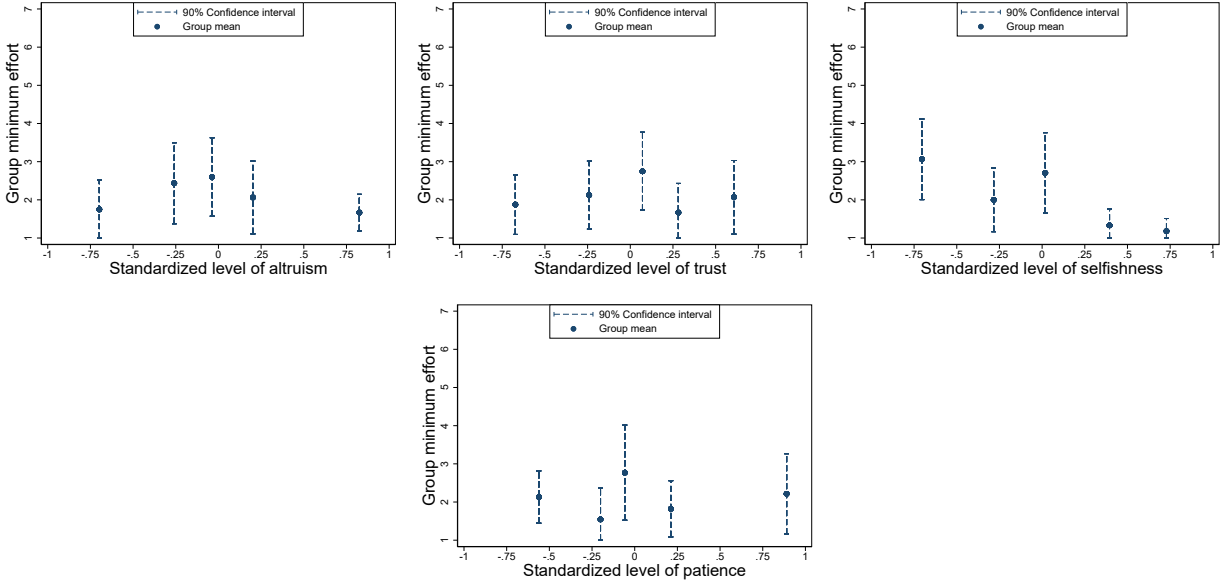
Figure A6: Individual characteristics and final-period group minima

Notes: The treatment variables are shown unmodified as binary, the binary socio-demographic factors are categorized into quintiles and the standardized variables are grouped into quintiles. The last period minimum effort level shows period 10 for the short horizon and from period 50 for the long horizon.

Finally, to help interpreting the findings above, Table A10 displays pairwise correlations between the socio-demographic characteristics and behavioral traits that we found to be most robustly linked to behavior.

Table A10: Pairwise correlations between selected socio-demographics and behavioral traits

	Danish	Male	Cognitive ability	Risk aversion	Trust
Danish	1.00				
Male	0.13	1.00			
Cognitive ability	0.29	0.21	1.00		
Risk aversion	0.01	-0.12	0.06	1.00	
Trust	0.27	0.01	0.15	-0.05	1.00

G.2 Textual analysis of coordination behavior

To complement the analysis of subjects' behavior during the experiment, we also conducted exploratory analyses of subjects' answers to some of the open-ended questions in our post-experimental questionnaire.

Figure A7 shows how subjects motivated their choice in the initial period (left panel) and their explanation for any revisions of their effort choices during the experiment (right panel). We manually coded answers into six categories. Whenever subjects provided multiple answers, we identified the most relevant answer based on their effort choices during the experiment.
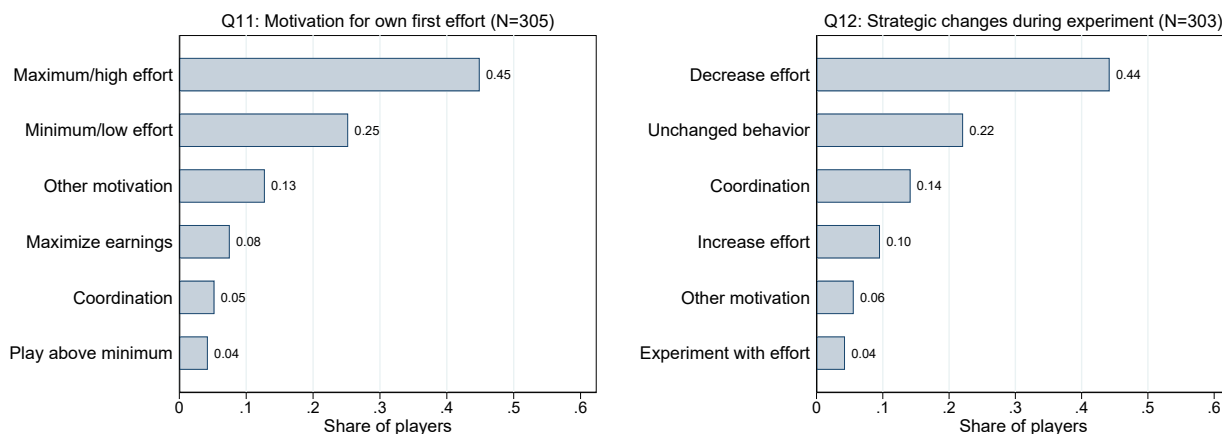
Figure A7: Self-expressed motivations

Notes: This figure presents data on subjects motivations for their behavior based on textual responses to two open-ended questions in the post-experimental questionnaire. Motives of the stated answers are manually categorized and represented in percentage terms. Maximum and high effort is defined to range from effort level 5 to 7, medium and low effort from 1 to 4.

The figure in the left panel shows that the most common response was some kind of motivation for choosing a high number in the initial period (44.9%). But a large fraction of subjects (25.2%) also thought it was better to choose a low number and a common motivation for the low choice was that it was safer to play low. Some subjects (4.3%) stated that they wanted to avoid being the minimum player, others emphasized that they wanted to coordinate on the same number as others in the group (5.2%) and some simply stated that they wanted to maximize earnings (7.5%). The residual category "Other Motivation" (12.8%) includes subjects that were uncertain about what to do, whose explanation indicated a lack of understanding (despite having answered the control questions correctly) and some subjects simply stating that they chose effort randomly.

The right panel shows motivations for potential behavioral changes over the course of the experiment. The most common answer (44.2%) was that they decreased their effort choices during the experiment, which is in line with the behavioral pattern. A sizable fraction (22.1%) indicated that they did not change their behavior. The third most common answer (14.2%) was that subjects tried to coordinate and choose the same effort choice as other group members. Some players indicated that they tried to raise the minimum by switching to a higher effort level (9.6%) and some tried to experiment with different effort choices (4.3%). Finally, a small fraction of subjects (5.6%) provided motivations that were incomplete or did not show that they had understood the game.

Overall, the analysis of the open-ended questions aligns well with the behavioral patterns in the data. A considerable fraction chooses lower numbers initially, either deliberately or due to an incomplete understanding. This is likely to result in a low group minimum in many groups. Few players state that they tried to raise the minimum or experiment in later periods, which can explain why almost no group managed to "climb the ladder" to higher effort levels (see Figure A1 to A4).

H   Experimental instructions

H.1   Instructions for weak-link game

Instructions [Sheet of paper – High stakes treatment]

Welcome to this experiment! Please read the instructions on this sheet and on your screen thoroughly so that you understand the structure of the experiment. Do not talk to other participants during the experiment. Whenever you have any questions, please raise your hand and wait for one of us to come to you.

The experiment consists of several tasks. After those tasks are completed, there will be a questionnaire that we kindly ask you to answer as carefully as possible. The length of the experiment will not exceed 60 minutes.

You will be paid for your participation and the size of your reward will depend on your own decisions as well as on the decisions of other participants. All amounts stated in the experiment are in Danish kroner (DKK) and your earnings will be paid at the end of the experiment in private and in cash. Your decisions and the decisions of other participants will remain anonymous.

In the first and main part of the experiment you will interact with three other participants during a number of rounds. In each round, every participant chooses a number from 1 to 7. Thus, the numbers you may choose are 1, 2, 3, 4, 5, 6, and 7. You can earn between 0 DKK and 4.50 DKK in each round.

The table below illustrates how your earnings depend on your choice and the choices of the three other participants you will be matched with. The first column on the left indicates the number you chose. The first row on top indicates the smallest number chosen by the other three participants and you. For example:

- Suppose you chose 3 and the smallest number chosen by the other three participants is 4. Then the minimum is 3 and you will earn 3.50 DKK.

- Suppose you chose 4 and the smallest number chosen by the other three participants is 2. In this case the minimum is 2 and you will earn 2.25 DKK.

Table A11: Payoff matrices with high stakes (in DKK)

	Minimum effort level $\underline{e}$						
	7	6	5	4	3	2	1
7	4.50	3.75	3.00	2.25	1.50	0.75	0.00
6		4.25	3.50	2.75	2.00	1.25	0.50
5			4.00	3.25	2.50	1.75	1.00
4				3.75	3.00	2.25	1.50
3					3.50	2.75	2.00
2						3.25	2.50
1							3.00

Individual effort e_i

In each round, you will be asked to choose one number within 15 seconds. Each round represents the situation described above and the payoff table is the same in every round. A box on the lower half of the screen will show your choice, the smallest number chosen and your earnings in past rounds.

Before starting the first task, we will ask a number of control questions. Once everybody has correctly answered the control questions, the experiment will begin. Instructions for the tasks following the first task will be displayed on your screen.

Instructions [Sheet of paper – Low stakes treatment]

Welcome to this experiment! Please read the instructions on this sheet and on your screen thoroughly so that you understand the structure of the experiment. Do not talk to other participants during the experiment. Whenever you have any questions, please raise your hand and wait for one of us to come to you.

The experiment consists of several tasks. After those tasks are completed, there will be a questionnaire that we kindly ask you to answer as carefully as possible. The length of the experiment will not exceed 60 minutes.

You will be paid for your participation and the size of your reward will depend on your own decisions as well as on the decisions of other participants. All amounts stated in the experiment are in Danish kroner (DKK) and your earnings will be paid at the end of the experiment in private and in cash. Your decisions and the decisions of other participants will remain anonymous.

In the first and main part of the experiment you will interact with three other participants during a number of rounds. In each round, every participant chooses a number from 1 to 7. Thus, the numbers you may choose are 1, 2, 3, 4, 5, 6, and 7.

You can earn between 0 DKK and 0.90 DKK (90 øre) in each round. The table below illustrates how your earnings depend on your choice and the choices of the three other participants you will be matched with. The first column on the left indicates the number you chose. The first row on top indicates the smallest number chosen by the other three participants and you. For example:

- Suppose you chose 3 and the smallest number chosen by the other three participants is 4. Then the minimum is 3 and you will earn 0.70 DKK.

- Suppose you chose 4 and the smallest number chosen by the other three participants is 2. In this case the minimum is 2 and you will earn 0.45 DKK.

Table A12: Payoff matrices with low stakes (in DKK)

	\underline{e} = 7	6	5	4	3	2	1
Minimum effort level $\underline{e}$							
Individual effort e_i = 7	0.90	0.75	0.60	0.45	0.30	0.15	0.00
6		0.85	0.70	0.55	0.40	0.25	0.10
5			0.80	0.65	0.50	0.35	0.20
4				0.75	0.60	0.45	0.30
3					0.70	0.55	0.40
2						0.65	0.50
1							0.60

In each round, you will be asked to choose one number within 15 seconds. Each round represents the situation described above and the payoff table is the same in every round. A box on the lower half of the screen will show your choice, the smallest number chosen and your earnings in past rounds.

Before starting the first task, we will ask a number of control questions. Once everybody has correctly answered the control questions, the experiment will begin. Instructions for the tasks following the first task will be displayed on your screen.

Instructions continued [ON SCREEN]

General instructions As explained in your handout, you will interact with three other participants in this task. In each round, every participant chooses a number from 1 to 7. Thus, the numbers you may choose are 1, 2, 3, 4, 5, 6, and 7. Please read the handout carefully.

The table below illustrates how your earnings depend on your choice and the choices of the three other participants. The first column on the left indicates the number you chose. The first row on top indicates the smallest number chosen by the other three participants and you. All amounts stated are in DKK.

[PAYOFF TABLE HERE – DIFFERS BY TREATMENT]

If you have no further questions, press OK to start the control questions.

Control questions [3] Suppose you chose YY and the smallest number chosen by the other three participants is ZZ.

Then the minimum is: [SUBJECT ENTRY]

and you will earn: [SUBJECT ENTRY]

All amounts are stated in DKK. Please use "." as delimiter, i.e. 1.00, 1.10, 1.25, etc. Press OK to check your answer.

Instructions first weak link game

You answered all control questions correctly. Next, the task will start.

You can earn between 0 DKK and 4.50 DKK [0.90 DKK] in each round. You will be matched randomly with three other participants. You will interact with the same participants in every round. **You will interact for 50 rounds [10 rounds]**. In each round, you will be asked to promptly choose one number. Each round represents the situation described above and the payoff table is the same in every round. A box on the lower half of the screen will show your choice, the smallest number chosen and your earnings in past rounds.

We kindly ask you to make your decision in every round within 15 seconds. If you have no further questions, press OK to start the first round.

Decision screens

[Effort level choice – SUBJECT ENTRY]

[Payoff table display]

[History reporting own effort choice, minimum effort chosen, own payoff for all previous rounds of that game]

Announcement of next weak link games – ONLY TREATMENTS WITH SHORT HORIZON Below you can see a summary of the last 10 rounds.

You will now carry out the previous task for another 40 rounds. Each round represents the same situation and payoff table as described in the previous task.

Now and every 10 rounds, you will be matched randomly with three participants, i.e. you will interact with three other participants for 10 rounds before you are rematched.

[PAYOFF TABLE HERE – DIFFERS BY TREATMENT]

If you have no further questions, press OK to start the first round.

Every 10 rounds:

Below you can see a summary of the last 10 rounds.

You will now again be randomly matched and interact with three other participants for another 10 rounds.

Press OK to continue.

Announcement of final weak link game – ALL TREATMENTS, ONLY LAST SESSIONS [When stakes were low before] Below you can see a summary of the last rounds.

Next, you will carry out the previous task for another 10 rounds. You will interact with the same participants as in the last 50 rounds [10 rounds], which is known to each of them.

Each round represents the s**ame situation and payoff table** as described in the previous task. [Each round represents the same situation as described in the previous task, **but you and the other participants face another payoff table, as shown on the right hand side.**] Recall that all amounts stated are in DKK and that you will be paid according to your decisions and the decisions of the other three participants. **You can earn between 0 DKK and 4.50 DKK in each round.**

[PAYOFF TABLE WITH HIGH STAKES HERE]

If you have no further questions, press OK to start the first round.

H.2   Ultimatum game instructions

In this task, two participants, Person A and Person B, interact. Person A receives 40 kroner. Person B does not initially receive anything.

Person A then proposes how to split the 40 kroner received between Person A and Person B. Person B can choose whether to accept or reject the proposal from Person A. If Person B accepts, Person A will earn 40 kroner minus the amount proposed to Person B, and Person B will earn the amount proposed by Person A. If Person B rejects, both Person A and Person B get nothing.

You will be paired randomly with another participant. Below, we ask you to indicate your choice for both scenarios, i.e. for the case that you are Person A and for the case that you are Person B. After that, you will be randomly assigned to be either Person A or Person B. You will be paid according to your choices and the choices of the other participant.

Indicate the amount between 0 and 40 kroner that you propose to give to Person B in case you are randomly assigned to be Person A.

[SUBJECT ENTRY here]

Indicate the minimum amount between 0 and 40 kroner that you would accept from Person A in case you are randomly assigned to be Person B. All proposals to you as Person B that are lower than what you state here will hence be rejected and in that case both Person A and Person B will not earn anything from this task.

[SUBJECT ENTRY here]

After you made your decisions, press "Submit" to continue.

H.3 Donation instructions

The total amount that you have earned throughout this experiment: ...

You now have the option to give any share of your earnings to Læger uden Grænser (Medecins sans frontiers). Læger uden Grænser is an international medical humanitarian organization that delivers emergency aid to people affected by armed conflict, epidemics, natural disasters and exclusion from healthcare.

We faithfully insure that your donation reaches Læger uden Grænser.

I would like to give [make choice here]

Press OK to continue.

H.4 Questionnaire

General questions

1. Imagine that you have the option between receiving some money in a month or in seven months from now. Please indicate for each of the following situations whether you would prefer option A or B.

A. 1000 DKK in a month   B. 1025 DKK in seven months
A. 1000 DKK in a month   B. 1075 DKK in seven months
A. 1000 DKK in a month   B. 1125 DKK in seven months
A. 1000 DKK in a month   B. 1175 DKK in seven months
A. 1000 DKK in a month   B. 1225 DKK in seven months
A. 1000 DKK in a month   B. 1275 DKK in seven months

[Adapted from Andersen et al. (2008)]

2. Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks? Please tick a box on a scale between 0 and 10 where 0 means "unwilling to take risks" and 10 means "fully prepared to take risk".

[Adapted from Dohmen et al. (2011)]

3. Imagine you had won 100,000 DKK in a lottery. Almost immediately after you collect the money, you receive the following financial offer from a reputable bank, the conditions of which are as follows: There is the chance to double the money within two years. It is equally possible that you could lose half of the amount invested. What share of your lottery winnings would you be prepared to invest in this financially risky, yet lucrative investment?

   0, 20,000, 40,000, 60,000 80,000, or 100,000 DKK

[Adapted from Dohmen et al. (2011)]

4. To what extent do you agree or disagree with the following statements:

A) In general, one can trust people
B) These days you can't rely on anybody else
C) When dealing with strangers it is better to be careful before you trust them.
Agree fully, agree somewhat, disagree somewhat, fully disagree.

[Adapted from Fehr et al. (2003)]

5. Please indicate on a 7-point scale how well each of the following six statements applies to you personally:
A) If someone does me a favor, I am prepared to return it.
B) If I suffer a serious wrong, I will take revenge as soon as possible, no matter what the cost.
C) If somebody puts me in a difficult position, I will do the same to him/her.
D) I go out of my way to help somebody who has been kind to me before.
E) If somebody offends me, I will offend him/her back.
F) I am ready to undergo personal costs to help somebody who helped me before.
1 "does not apply to me at all" to 7 "applies to me perfectly".

[Adapted from Dohmen et al. (2009)]

6. A bat and a ball cost $1.10 in total. The bat costs $1.00 more than the ball. How much does the ball cost?

7. If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets?

8. In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake?

[CRT test, adapted from Frederick (2005)]

Questions about the experiment

9. In the following we would like you to think of the first and main part of the experiment. Please describe in words how you thought when you made up your mind about what action to choose.

10. What were your initial expectations about what others would do in the experiment?

11. How did you think when you chose the first action/number?

12. Did you start to think differently about how to behave during the course of the experiment?

(9-12 Only in later sessions)

13. If you played a higher number than the minimum and continued to play that higher number in the next round, what was your main reason for doing so?

1. I hoped that the other players would switch to higher numbers.
2. I wanted to make sure that I would not be the one to lower the minimum in the next round.
3. Other reason, please state:
4. I never played a higher number than the minimum in the previous round.

14. When deciding which action to choose, did you primarily think about getting a higher payoff for yourself or did you primarily worry about the group as a whole?

1. "Thought primarily about myself"
7. "Thought primarily about the group as a whole"

15. How certain were you generally about what action to pick?

I was very certain what action to take. . .
1. in all rounds
3. during roughly half of the rounds
5. in no round

16. Suppose that three people pick numbers between 1 and 7 completely randomly, i.e. with the same probability for each number. How likely is it that the lowest number picked by the group will be 1?

1. About 15 percent
2. About 25 percent
3. About 33 percent
4. About 40 percent
5. About 50 percent
5. About 75 percent

(15.-16. Only in later sessions)

Background questions

17. Finally, we would like to ask you some personal questions. Do you have experience from previous similar experiments?

Yes/No

18. Have you taken courses in economics?

Yes/No

19. Have you taken courses in game theory?

Yes/No

20. What is your gender?

Male/Female

21. What is your age?

22. Which subject are you or do you plan to major in (if any)?

23. Which country were you born in?

24. Which country were your parents born in?

(23.-24. Only in later sessions)

25. How would you primarily classify your social background?

1. Working class
2. Middle class
3. Upper middle class
4. Upper class

26. What is/was your father's highest education?

1. Primary school
2. Secondary school
3. University, at most 3 years.
4. University, more than 3 years.

27. What is/was your mother's highest education?

1. Primary school
2. Secondary school
3. University, at most 3 years.
4. University, more than 3 years.

28. Was there anything in this experiment that you found difficult/confusing/strange that you think that the researchers responsible for this experiment should think about in future experiments?

References

Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2008). Eliciting risk and time preferences. *Econometrica*, *76*(3), 583–618. https://doi.org/10.1111/j.1468-0262.2008.00848.x

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. *Journal of the American statistical Association*, *103*(484), 1481–1495. https://doi.org/10.1198/016214508000000841

Berninghaus, S. K., & Ehrhart, K.-M. (1998). Time horizon and equilibrium selection in tacit coordination games: Experimental results. *Journal of Economic Behavior & Organization*, *37*(2), 231–248. https://doi.org/10.1016/S0167-2681(98)00086-9

Brockhoff, K., Margolin, M., & Weber, J. (2015). Towards empirically measuring patience. *Universal Journal of Management*, *3*(5), 169–178. https://doi.org/10.13189/ujm.2015.030501

Dohmen, T., Falk, A., Huffman, D., & Sunde, U. (2009). Homo reciprocans: Survey evidence on behavioural outcomes. *Economic Journal*, *119*(536), 592–612. https://doi.org/10.1111/j.1468-0297.2008.02242.x

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., & Wagner, G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, *9*(3), 522–550. https://doi.org/10.1111/j.1542-4774.2011.01015.x

Engelmann, D., & Normann, H.-T. (2010). Maximum effort in the minimum-effort game. *Experimental Economics*, *13*(3), 249–259. https://doi.org/10.1007/s10683-010-9239-3

Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D., & Sunde, U. (2018). Global evidence on economic preferences. *Quarterly Journal of Economics*, *133*(4), 1645–1692. https://doi.org/10.1093/qje/qjy013

Falk, A., Becker, A., Dohmen, T., Huffman, D., & Sunde, U. (2016). The preference survey module: A validated instrument for measuring risk, time, and social preferences. *IZA Discussion Paper*.

Fehr, E., Fischbacher, U., Von Rosenbladt, B., Schupp, J., & Wagner, G. (2003). A nation-wide laboratory: Examining trust and trustworthiness by integrating behavioral experiments into representative survey. *CESifo Working Paper Series*.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, *19*(4), 25–42. https://doi.org/10.1257/089533005775196732

Hoppe, E. I., & Kusterer, D. J. (2011). Behavioral biases and cognitive reflection. *Economics Letters*, *110*(2), 97–100. https://doi.org/10.1016/j.econlet.2010.11.015

List, J. A., Sadoff, S., & Wagner, M. (2011). So you want to run an experiment, now what? Some simple rules of thumb for optimal experimental design. *Experimental Economics*, *14*(4), 439–457. https://doi.org/10.1007/s10683-011-9275-7

Oechssler, J., Roider, A., & Schmitz, P. W. (2009). Cognitive abilities and behavioral biases. *Journal of Economic Behavior & Organization*, *72*(1), 147–152. https://doi.org/10.1016/j.jebo.2009.04.018

Szaszi, B., Szollosi, A., Palfi, B., & Aczel, B. (2017). The cognitive reflection test revisited: Exploring the ways individuals solve the test. *Thinking & Reasoning*, *23*(3), 207–234. https://doi.org/10.1080/13546783.2017.1292954

Toplak, M. E., West, R. F., & Stanovich, K. E. (2011). The cognitive reflection test as a predictor of performance on heuristics-and-biases tasks. *Memory & cognition*, *39*(7), 1275–1289. https://doi.org/10.3758/s13421-011-0104-1