Multiple Games Elicitation Experiment: Personal and Social Appropriateness

Wednesday 17th August, 2022

Contents

1 Hypotheses and Tests: Comparing across questions situations **4**
- 1.1 Dictator Games . . . 4
  - 1.1.1 DE with taking option . . . 4
  - 1.1.2 DG with surplus generation . . . 5
  - 1.1.3 DG with exit option . . . 8
- 1.2 Ultimatum games . . . 9
  - 1.2.1 Mini UG . . . 9
  - 1.2.2 UG proposer . . . 10
  - 1.2.3 UG responder . . . 12
- 1.3 Information avoidance game . . . 14
- 1.4 Trust games . . . 15
  - 1.4.1 TG proposer . . . 15
  - 1.4.2 TG responder . . . 16

2 Hypotheses and Tests: gender differences **19**
- 2.1 Dictator games . . . 19
- 2.2 Information avoidance game . . . 20
- 2.3 Ultimatum and trust games . . . 21

3 Hypotheses and Tests: Order effects **22**
- 3.1 Test at level of questions/actions . . . 22
  - 3.1.1 Dictator games . . . 23
  - 3.1.2 Information avoidance game . . . 36
  - 3.1.3 Ultimatum games . . . 37
  - 3.1.4 Trust games . . . 41

4 Figures **47**
- 4.1 Dictator games . . . 48
- 4.2 Information avoidance games . . . 49
- 4.3 Ultimatum and trust game . . . 50
- 4.4 Dictator games ratings accounting for the order of games . . . 53

List of Tables

1 DE stnd. vs DE take: social and personal appropriateness (t-tests) . . . 4
2 DE stnd. vs DE take: social and personal appropriateness (Rank sum test) . . . 5
3 DE stnd. vs DE quiz: social and personal appropriateness (t-tests) . . . 5
4 DE stnd. vs DE quiz: social and personal appropriateness (Rank sum test) . . . 6
5 DE stand. vs DE surplus generation: social and personal appropriateness (t-tests) . . . 6
6 DE stand. vs DE surplus generation: social and personal appropriateness (Rank sum test) . . . 7
7 DE stnd. vs DE surplus generation: social and personal appropriateness (t-tests) . . . 7
8 DE stand. vs DE surplus generation: social and personal appropriateness (Rank sum test) . . . 8
9 DE stnd. vs DE exit: social and personal appropriateness (t-tests) . . . 8
10 DE stnd. vs DE exit: social and personal appropriateness(Rank sum test) . . . 9
11 Mini UG: social and personal appropriateness (t-tests) . . . 9
12 Mini UG: social and personal appropriateness (Rank sum test) . . . 10

13 UG proposer: personal and social appropriateness (t-tests) . . . 11
14 UG proposer: personal and social appropriateness (Rank sum test) . . . 12
15 UG Responder: personal and social appropriateness (t-tests) . . . 13
16 UG responder: personal and social appropriateness (Rank sum test) . . . 14
17 DWK treatment: social and personal appropriateness (t-tests) . . . 15
18 DWK treatment: personal and social appropriateness (Rank sum test) . . . 15
19 DWK treatment: social and personal appropriateness (t-tests) . . . 16
20 TG proposer: personal and social appropriateness (Rank sum test) . . . 16
21 DWK treatment: social and personal appropriateness (t-tests) . . . 17
22 TG responder: personal and social appropriateness (Rank sum test) . . . 18
23 DG standard, with exit option, take 1: gender differences in social appropriateness (t-tests) . . . 19
24 DE standard, with exit option, take 1: gender differences in social appropriateness (Rank sum test) . . . 19
25 DE standard, with exit option, take 1: gender differences in personal appropriateness (t-tests) . . . 20
26 DE standard, with exit option, take 1: gender differences in personal appropriateness (Rank sum test) 20
27 DWK treatment: gender differences in social appropriateness (t-tests) . . . 20
28 DWK treatment: gender differences in social appropriateness (Rank sum test) . . . 21
29 DWK treatment: gender differences in personal appropriateness (t-tests) . . . 21
30 DWK treatment: gender differences in personal appropriateness (Rank sum test) . . . 21
31 UG and TG: gender differences in social appropriateness (t-tests) . . . 21
32 UG and TG: gender differences in social appropriateness (Rank sum test) . . . 22
33 UG and TG: gender differences in personal appropriateness (t-tests) . . . 22
34 UG and TG: gender differences in social appropriateness (Rank sum test) . . . 22
35 Standard DE: order effects in social appropriateness (t-tests) . . . 23
36 Standard DE: order effects in social appropriateness (Rank sum test) . . . 23
37 Standard DE: order effects in personal appropriateness (t-tests) . . . 24
38 Standard DE: order effects in personal appropriateness (Rank sum test) . . . 24
39 DE take 1: order effects in social appropriateness (t-tests) . . . 25
40 DE take 1: order effects in social appropriateness (Rank sum test) . . . 25
41 DE take 1: order effects in personal appropriateness (t-tests) . . . 26
42 DE take 1: order effects in personal appropriateness (Rank sum test) . . . 26
43 DE take 5: order effects in social appropriateness (t-tests) . . . 27
44 DE take 5: order effects in social appropriateness (Rank sum test) . . . 28
45 DE take 5: order effects in personal appropriateness (t-tests) . . . 29
46 DE take 5: order effects in personal appropriateness (Rank sum test) . . . 30
47 DE quiz: order effects in social appropriateness (t-tests) . . . 30
48 DE quiz: order effects in social appropriateness (Rank sum test) . . . 31
49 DE quiz: order effects in personal appropriateness (t-tests) . . . 31
50 DE quiz: order effects in pers appropriateness (Rank sum test) . . . 32
51 DG production: order effects in social appropriateness (t-tests) . . . 32
52 DE production: order effects in social appropriateness (Rank sum test) . . . 33
53 DE production: order effects in personal appropriateness (t-tests) . . . 33
54 DE production: order effects in personal appropriateness (Rank sum test) . . . 34
55 DE with exit option: order effects in social appropriateness (t-tests) . . . 34
56 DE exit option: order effects in social appropriateness (Rank sum test) . . . 35
57 DE with exit option: order effects in personal appropriateness (t-tests) . . . 35
58 DE exit option: order effects in personal appropriateness (Rank sum test) . . . 36
59 DWK treatment: order effects in social appropriateness (t-tests) . . . 36
60 DWK treatment: order effects in social appropriateness (Rank sum test) . . . 36
61 DWK treatment: order effects in personal appropriateness (t-tests) . . . 37
62 DWK treatment: order effects in social appropriateness (Rank sum test) . . . 37
63 Mini UG: order effects in social appropriateness (t-tests) . . . 37
64 Mini UG: order effects in social appropriateness (Rank sum test) . . . 38
65 Mini UG: order effects in personal appropriateness (t-tests) . . . 38
66 Mini UG: order effects in pers appropriateness (Rank sum test) . . . 38
67 UG proposer: order effects in social appropriateness (t-tests) . . . 38

68 UG proposer: order effects in social appropriateness (Rank sum test) . . . 39
69 UG proposer: order effects in personal appropriateness (t-tests) . . . 39
70 UG proposer: order effects in personal appropriateness (Rank sum test) . . . 39
71 UG proposer: order effects in social appropriateness (t-tests) . . . 40
72 UG responder: order effects in personal appropriateness (Rank sum test) . . . 40
73 UG responder: order effects in personal appropriateness (t-tests) . . . 40
74 UF responder: order effects in personal appropriateness (Rank sum test) . . . 41
75 TG proposer: order effects in personal appropriateness (t-tests) . . . 41
76 TG responder: order effects in social appropriateness (Rank sum test) . . . 41
77 TG responder: order effects in personal appropriateness (t-tests) . . . 42
78 TG responder: order effects in personal appropriateness (Rank sum test) . . . 42
79 TG responder 3: order effects in personal appropriateness (t-tests) . . . 42
80 TG responder: order effects in social appropriateness (Runk sum test) . . . 43
81 TG responder 3: order effects in personal appropriateness (t-tests) . . . 43
82 TG responder: order effects in personal appropriateness (Rank sum test) . . . 43
83 TG responder 9: order effects in personal appropriateness (t-tests) . . . 44
84 TG responder: order effects in social appropriateness (Rank sum test) . . . 44
85 TG responder 9: order effects in personal appropriateness (t-tests) . . . 44
86 TG responder: order effects in social appropriateness (Rank sum test) . . . 45
87 TG responder 12: order effects in personal appropriateness (t-tests) . . . 45
88 TG responder: order effects in social appropriateness (Rank sum test) . . . 45
89 TG responder 12: order effects in personal appropriateness (t-tests) . . . 46
90 TG responder: order effects in social appropriateness (Rank sum test) . . . 46
91 TG responder 18: order effects in personal appropriateness (t-tests) . . . 46
92 TG responder: order effects in social appropriateness (Rank sum test) . . . 47
93 TG responder 18: order effects in personal appropriateness (t-tests) . . . 47
94 TG responder: order effects in personal appropriateness (Rank sum test) . . . 47

List of Figures

1 Average ratings in the standard DE and DE with exit option . . . 48
2 Average ratings in the standard DE and DE with take options . . . 48
3 Average ratings in the standard DE and DE with surplus . . . 49
4 Average ratings in the DWK treatment game . . . 49
5 Average ratings in the DWK baseline game . . . 50
6 Average ratings in the mini ultimatum game (FFF 5:5 and 2:8) . . . 50
7 Average ratings in the mini ultimatum game (FFF 2:8) . . . 51
8 Average ratings in the UG proposer game . . . 51
9 Average ratings in the UG responder game . . . 52
10 Average ratings in the TG proposer game . . . 52
11 Average ratings in the TG responder game . . . 53
12 Average ratings in the Standard DE for subject who met this game as their first . . . 53
13 Average ratings in the DE with exit option for subject who met this game as their first . . . 54
14 Average ratings in the surplus DE take 2 for subject who met this game as their first . . . 54
15 Average ratings in the surplus DE take 10 for subject who met this game as their first . . . 55
16 Average ratings in the surplus DE (production) for subject who met this game as their first . . . 55
17 Average ratings in the surplus DE (quiz) for subject who met this game as their first . . . 56

1 Hypotheses and Tests: Comparing across questions situations

1.1 Dictator Games

The tables below show the results from the hypothesis testing about the different dictator games (DE) (standard DE, DE with taking option, DE with surplus generation, DE with Exit option).

1.1.1 DE with taking option

We hypothesise that giving close to zero is more appropriate in the Take-1 DG and Take-5 DG than in the Standard DG. Formally we expect to reject the following H0 in favour of H1, for $0 \leqslant X \leqslant 2$.

H0: AverageGive X (Take-1 DG) = AverageGive X (Standard DG)
H1: AverageGive X (Take-1 DG) > AverageGive X (Standard DG)
H0: AverageGive X (Take-5 DG) = AverageGive X (Standard DG)
H1: AverageGive X (Take-5 DG) > AverageGive X (Standard DG)

Table 1: DE stnd. vs DE take: social and personal appropriateness (t-tests)

	Obs (a)	Obs (b)	Mean (a)	Mean (b)	Diff. (a) - (b)	P-value (H0: Diff. = 0, H1: Diff. < 0)
Social appropriateness						
St. DE (0) - DE take 2 (give 0)	148	157	-.649	.01	-.658	0
St. DE (0) - DE take 10 (give 0)	148	135	-.649	.241	-.889	0
St. DE (1) - DE take 2 (give 1)	148	157	-.476	.073	-.55	0
St. DE (1) - DE take 10 (give 1)	148	135	-.476	.241	-.717	0
St. DE (2) - DE take 2 (give 2)	148	157	-.355	.156	-.511	0
St. DE (2) - DE take 10 (give 2)	148	135	-.355	.289	-.644	0
Personal appropriateness						
St. DE (0) - DE take 2 (give 0)	152	125	-.599	-.02	-.579	0
St. DE (0) - DE take 10 (give 0)	152	171	-.599	.14	-.739	0
St. DE (1) - DE take 2 (give 1)	152	125	-.487	.04	-.527	0
St. DE (1) - DE take 10 (give 1)	152	171	-.487	.164	-.651	0
St. DE (2) - DE take 2 (give 2)	152	125	-.391	.12	-.511	0
St. DE (2) - DE take 10 (give 2)	152	170	-.391	.153	-.544	0

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 2: DE stnd. vs DE take: social and personal appropriateness (Rank sum test)

	Obs (a)	Obs (b)	Z-stat (a)	P-value (b)	P-value (exact) (H0: Diff. = 0, H1: Diff. < 0)
Social appropriateness					
St. DE (0) - DE take 2 (give 0)	148	157	-8.149	0	0
St. DE (0) - DE take 10 (give 0)	148	135	-10.499	0	0
St. DE (1) - DE take 2 (give 1)	148	157	-6.871	0	0
St. DE (1) - DE take 10 (give 1)	148	135	-8.725	0	0
St. DE (2) - DE take 2 (give 2)	148	157	-6.748	0	0
St. DE (2) - DE take 10 (give 2)	148	135	-8.124	0	0
Personal appropriateness					
St. DE (0) - DE take 2 (give 0)	152	125	-7.028	0	0
St. DE (0) - DE take 10 (give 0)	152	171	-9.652	0	0
St. DE (1) - DE take 2 (give 1)	152	125	-6.718	0	0
St. DE (1) - DE take 10 (give 1)	152	171	-8.749	0	0
St. DE (2) - DE take 2 (give 2)	152	125	-6.814	0	0
St. DE (2) - DE take 10 (give 2)	152	170	-7.83	0	0

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

1.1.2 DG with surplus generation

We hypothesise that giving close to zero is more appropriate in the DG with Quiz and the DG with Production than in the Standard DG. Formally we expect to reject the following H0 in favour of H1, for $0 \leqslant X \leqslant 2$.

H0: AverageGive X (DG with Quiz) = AverageGive X (Standard DG)
H1: AverageGive X (DG with Quiz) > AverageGive X (Standard DG)
H0: AverageGive X (DG with Production) = AverageGive X (Standard DG)
H1: AverageGive X (DG with Production) > AverageGive X (Standard DG)

Table 3: DE stnd. vs DE quiz: social and personal appropriateness (t-tests)

	Obs (a)	Obs (b)	Mean (a)	Mean (b)	Diff. (a) - (b)	P-value (H0: Diff. = 0, H1: Diff. < 0)
Social appropriateness						
St. DE (0) - DE Quiz (0)	148	155	-.649	-.348	-.3	0
St. DE (1) - DE Quiz (1)	148	155	-.476	-.165	-.312	0
St. DE (2) - DE Quiz (2)	148	155	-.355	-.058	-.297	0
Personal appropriateness						
St. DE (0) - DE Quiz (0)	152	150	-.599	-.213	-.385	0
St. DE (1) - DE Quiz (1)	152	150	-.487	-.12	-.367	0
St. DE (2) - DE Quiz (2)	152	150	-.391	-.07	-.321	0

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 4: DE stnd. vs DE quiz: social and personal appropriateness (Rank sum test)

	Obs (a)	Obs (b)	Z-stat (a)	P-value (b)	P-value (exact) (H0: Diff. = 0, H1: Diff. < 0)
Social appropriateness					
St. DE (0) - DE Quiz (0)	108	155	-3.001	.001	.001
St. DE (1) - DE Quiz (1)	108	155	-3.512	0	0
St. DE (2) - DE Quiz (2)	108	155	-2.84	.002	.002
Personal appropriateness					
St. DE (0) - DE Quiz (0)	114	150	-3.577	0	0
St. DE (1) - DE Quiz (1)	114	150	-3.78	0	0
St. DE (2) - DE Quiz (2)	114	150	-3.406	0	0

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 5: DE stand. vs DE surplus generation: social and personal appropriateness (t-tests)

	Obs (a)	Obs (b)	Mean (a)	Mean (b)	Diff. (a) - (b)	P-value (H0: Diff. = 0, H1: Diff. < 0)
Social appropriateness						
St. DE (0) - DE Prod (0)	148	148	-.649	-.209	-.439	0
St. DE (1) - DE Prod (1)	148	148	-.476	-.084	-.392	0
St. DE (2) - DE Prod (2)	148	148	-.355	-.041	-.314	0
Personal appropriateness						
St. DE (0) - DE Prod (0)	152	152	-.599	.043	-.641	0
St. DE (1) - DE Prod (1)	152	152	-.487	.013	-.5	0
St. DE (2) - DE Prod (2)	152	152	-.391	.053	-.444	0

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 6: DE stand. vs DE surplus generation: social and personal appropriateness (Rank sum test)

	Obs (a)	Obs (b)	Z-stat (a)	P-value (b)	P-value (exact) (H0: Diff. = 0, H1: Diff. < 0)
Social appropriateness					
St. DE (0) - DE Prod (0)	108	148	-4.038	0	0
St. DE (1) - DE Prod (1)	108	148	-3.487	0	0
St. DE (2) - DE Prod (2)	108	148	-3.258	.001	.001
Personal appropriateness					
St. DE (0) - DE Prod (0)	114	152	-6.122	0	0
St. DE (1) - DE Prod (1)	114	152	-5.673	0	0
St. DE (2) - DE Prod (2)	114	152	-5.297	0	0

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 7: DE stnd. vs DE surplus generation: social and personal appropriateness (t-tests)

	Obs (a)	Obs (b)	Mean (a)	Mean (b)	Diff. (a) - (b)	P-value (H0: Diff. = 0, H1: Diff. < 0)
Social appropriateness						
DE Quiz (0) - DE Prod (0)	155	148	-.348	-.209	-.139	.048
DE Quiz (1) - DE Prod (1)	155	148	-.165	-.084	-.08	.153
DE Quiz (2) - DE Prod (2)	155	148	-.058	-.041	-.018	.405
Personal appropriateness						
DE Quiz (0) - DE Prod (0)	150	152	-.213	.043	-.256	.002
DE Quiz (1) - DE Prod (1)	150	152	-.12	.013	-.133	.042
DE Quiz (2) - DE Prod (2)	150	152	-.07	.053	-.123	.047

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 8: DE stand. vs DE surplus generation: social and personal appropriateness (Rank sum test)

	Obs (a)	Obs (b)	Z-stat (a)	P-value (b)	P-value (exact) (H0: Diff. = 0, H1: Diff. < 0)
Social appropriateness					
DE Quiz (0) - DE Prod (0)	155	148	-1.515	.065	.065
DE Quiz (1) - DE Prod (1)	155	148	-.854	.197	.197
DE Quiz (2) - DE Prod (2)	155	148	-.141	.444	.444
Personal appropriateness					
DE Quiz (0) - DE Prod (0)	150	152	-2.719	.003	.003
DE Quiz (1) - DE Prod (1)	150	152	-1.667	.048	.047
DE Quiz (2) - DE Prod (2)	150	152	-1.691	.045	.045

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

1.1.3 DG with exit option

We hypothesise that taking the exit option (giving 9 for the dictator and 0 for the recipient) is rated as more appropriate than staying in the game and giving 1 to recipient (leaving 9 for the dictator). Formally we expect to reject the following H0 in favour of H1.

H0: AverageExit (DG with Exit) = AverageStay and giv \$1 (DG with Exit)
H1: AverageExit (DG with Exit) > AverageStay and giv \$1 (DG with Exit)

Table 9: DE stnd. vs DE exit: social and personal appropriateness (t-tests)

	Obs (a)	Obs (b)	Mean (a)	Mean (b)	Diff. (a) - (b)	P-value (H0: Diff. = 0, H1: Diff. < 0)
Social appropriateness						
DE ex. (give 0) - DE ex. (exit)	161	161	-.73	-.149	-.581	0
DE ex. (give 1) - DE ex. (exit)	161	161	-.475	-.149	-.326	0
Personal appropriateness						
DE ex. (give 0) - DE ex. (exit)	152	152	-.625	-.125	-.5	0
DE ex. (give 1) - DE ex. (exit)	152	152	-.319	-.125	-.194	.004

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 10: DE stnd. vs DE exit: social and personal appropriateness(Rank sum test)

	Obs (a)	Obs (b)	Z-stat (a)	P-value (b)	P-value (exact) (H0: Diff. = 0, H1: Diff. < 0)
Social appropriateness					
DE ex. (give 0) - DE ex. (exit)	161	161	-8.881	0	0
DE ex. (give 1) - DE ex. (exit)	161	161	-5.153	0	0
Personal appropriateness					
DE ex. (give 0) - DE ex. (exit)	152	152	-7.884	0	0
DE ex. (give 1) - DE ex. (exit)	152	152	-3.324	0	0

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

1.2 Ultimatum games

1.2.1 Mini UG

We hypothesise that the act of rejecting an offer of 8 for the Proposer and 2 for the Responder is rated as more appropriate when the outside option is an equal split (5 each) than when the outside option is unequal (2 for the Proposer and 8 for the Responder). Formally we expect to reject the following H0 in favour of H1.

H0: AverageReject (outside option 5:5) = AverageReject (outside option 2:8)
H1: AverageReject (outside option 5:5) > AverageReject (outside option 2:8)

Table 11: Mini UG: social and personal appropriateness (t-tests)

	Obs (a)	Obs (b)	Mean (a)	Mean (b)	Diff. (a) - (b)	P-value (H0: Diff. = 0, H1: Diff. < 0)
Social appropriateness						
Reject FFF28 - Reject FFF55	132	158	-.182	-.054	-.128	.045
Personal appropriateness						
Reject FFF28 - Reject FFF55	140	144	-.186	.101	-.286	0

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 12: Mini UG: social and personal appropriateness (Rank sum test)

	Obs (a)	Obs (b)	Z-stat (a)	P-value (b)	P-value (exact) (H0: Diff. = 0, H1: Diff. < 0)
Social appropriateness					
Reject FFF28 - Reject FFF55	132	158	-1.61	.054	.054
Personal appropriateness					
Reject FFF28 - Reject FFF55	140	144	-3.964	0	0

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

1.2.2 UG proposer

We hypothesise that offering close to zero is considered inappropriate. Formally we expect to reject the following H0 in favour of H1, for $0 \leqslant X \leqslant 2$.

H0: Average Propose X for responder = 0
H1: Average Propose X for responder < 0

Table 13: UG proposer: personal and social appropriateness (t-tests)

	Obs	Mean	P-value (H0: Mean= 0, H1: Mean < 0)
Social appropriateness			
B gets 0	166	-.822	0
B gets 1	166	-.693	0
B gets 2	166	-.59	0
B gets 3	166	-.373	0
B gets 4	166	-.03	.237
B gets 5	166	.786	1
B gets 6	166	.431	1
B gets 7	166	.283	1
B gets 8	166	.169	.999
B gets 9	166	.099	.947
B gets 10	166	.09	.918
Personal appropriateness			
B gets 0	167	-.781	0
B gets 1	167	-.662	0
B gets 2	167	-.605	0
B gets 3	167	-.425	0
B gets 4	167	-.114	.005
B gets 5	167	.799	1
B gets 6	167	.225	1
B gets 7	167	.006	.55
B gets 8	167	-.132	.005
B gets 9	167	-.204	0
B gets 10	167	-.269	0

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 14: UG proposer: personal and social appropriateness (Rank sum test)

	Obs (a)	Obs (b)	Z-stat (a)	P-value (b)	P-value (exact) (H0: Diff. = 0, H1: Diff. < 0)
Social appropriateness					
B gets 0	166	166	-11.12	0	0
B gets 1	166	166	-10.288	0	0
B gets 2	166	166	-9.642	0	0
B gets 3	166	166	-7.68	0	0
B gets 4	166	166	-.549	.292	.287
B gets 5	166	166	11.225	1	1
B gets 6	166	166	7.933	1	1
B gets 7	166	166	5.316	1	1
B gets 8	166	166	2.991	.999	.999
B gets 9	166	166	1.592	.944	.946
B gets 10	166	166	1.307	.904	.904
Personal appropriateness					
B gets 0	167	167	-10.803	0	0
B gets 1	167	167	-9.85	0	0
B gets 2	167	167	-9.695	0	0
B gets 3	167	167	-7.991	0	0
B gets 4	167	167	-2.546	.005	.005
B gets 5	167	167	11.536	1	1
B gets 6	167	167	4.973	1	1
B gets 7	167	167	.102	.54	.545
B gets 8	167	167	-2.583	.005	.005
B gets 9	167	167	-3.667	0	0
B gets 10	167	167	-4.595	0	0

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

1.2.3 UG responder

We hypothesise that rejecting an offer of 5 or more is considered inappropriate. Formally we expect to reject the following H0 in favour of H1, for $6 \leqslant X \leqslant 11$.

H0: Average of rejecting all offers less than X = 0
H1: Average of rejecting all offers less than X < 0

Table 15: UG Responder: personal and social appropriateness (t-tests)

	Obs	Mean	P-value (H0: Mean= 0, H1: Mean < 0)
Social appropriateness			
B reject offers less than 0	145	.345	1
B reject offers less than 1	145	.297	1
B reject offers less than 2	145	.238	1
B reject offers less than 3	145	.166	1
B reject offers less than 4	145	.031	.743
B reject offers less than 5	145	-.066	.109
B reject offers less than 6	145	-.445	0
B reject offers less than 7	145	-.538	0
B reject offers less than 8	145	-.624	0
B reject offers less than 9	145	-.655	0
B reject offers less than 10	145	-.676	0
B reject all offers	145	-.748	0
Personal appropriateness			
B reject offers less than 0	149	.302	1
B reject offers less than 1	149	.393	1
B reject offers less than 2	149	.285	1
B reject offers less than 3	149	.131	.996
B reject offers less than 4	149	-.037	.216
B reject offers less than 5	149	-.258	0
B reject offers less than 6	149	-.527	0
B reject offers less than 7	149	-.601	0
B reject offers less than 8	149	-.668	0
B reject offers less than 9	149	-.685	0
B reject offers less than 10	149	-.718	0
B reject all offers	149	-.752	0

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 16: UG responder: personal and social appropriateness (Rank sum test)

	Obs (a)	Obs (b)	Z-stat (a)	P-value (b)	P-value (exact) (H0: Diff. = 0, H1: Diff. < 0)
Social appropriateness					
B reject offers less than 0	145	145	5.537	1	1
B reject offers less than 1	145	145	5.008	1	1
B reject offers less than 2	145	145	4.289	1	1
B reject offers less than 3	145	145	3.442	1	1
B reject offers less than 4	145	145	.76	.776	.776
B reject offers less than 5	145	145	-1.289	.099	.1
B reject offers less than 6	145	145	-7.566	0	0
B reject offers less than 7	145	145	-8.562	0	0
B reject offers less than 8	145	145	-8.999	0	0
B reject offers less than 9	145	145	-9.098	0	0
B reject offers less than 10	145	145	-9.074	0	0
B reject all offers	145	145	-9.9	0	0
Personal appropriateness					
B reject offers less than 0	149	149	5.406	1	1
B reject offers less than 1	149	149	6.334	1	1
B reject offers less than 2	149	149	4.894	1	1
B reject offers less than 3	149	149	2.714	.997	.997
B reject offers less than 4	149	149	-.657	.256	.257
B reject offers less than 5	149	149	-4.778	0	0
B reject offers less than 6	149	149	-8.551	0	0
B reject offers less than 7	149	149	-9.18	0	0
B reject offers less than 8	149	149	-9.56	0	0
B reject offers less than 9	149	149	-9.539	0	0
B reject offers less than 10	149	149	-9.536	0	0
B reject all offers	149	149	-10.383	0	0

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

1.3 Information avoidance game

We hypothesise that choosing to become informed, learning that one is in the unaligned interest state, and choosing Y (yielding 5,5), is more appropriate than choosing to stay uninformed. Moreover, we hypothesise that choosing to become informed, learning that one is in the unaligned interest state, and choosing X (yielding 6,1), is more inappropriate than choosing to stay uninformed. Formally we expect to reject the following H0 in favour of H1.

H0: AverageBecome informed, chose Y in unaligned state = AverageStay uninformed
H1: AverageBecome informed, chose Y in unaligned state > AverageStay uninformed
H0: AverageBecome informed, chose X in unaligned state = AverageStay uninformed
H1: AverageBecome informed, chose X in unaligned state < AverageStay uninformed

Table 17: DWK treatment: social and personal appropriateness (t-tests)

	Obs (a)	Obs (b)	Mean (a)	Mean (b)	Diff. (a) - (b)	P-value (H0: Diff. = 0, H1: Diff. < 0)
Social appropriateness						
Uninf. - A chooses Y, inf.	150	150	.135	.823	-.688	0
A chooses X, inf. - Uninf.	150	150	-.513	.135	-.648	0
Personal appropriateness						
Uninf. - A chooses Y, inf.	150	150	.267	.787	-.52	0
A chooses X, inf. - Uninf.	150	150	-.493	.267	-.76	0

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 18: DWK treatment: personal and social appropriateness (Rank sum test)

	Obs (a)	Obs (b)	Z-stat (a)	P-value (b)	P-value (exact) (H0: Diff. = 0, H1: Diff. < 0)
Social appropriateness					
Uninf. - A chooses Y, inf.	150	150	-9.213	0	0
A chooses X, inf. - Uninf.	150	150	-8.625	0	0
Personal appropriateness					
Uninf. - A chooses Y, inf.	150	150	-8.324	0	0
A chooses X, inf. - Uninf.	150	150	-8.987	0	0

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

1.4 Trust games

1.4.1 TG proposer

We hypothesise that sending less than half is considered less appropriate than sending more than half. Formally, let

$$\text{AverageSend } <3 = (\text{AverageSend } 0 + \text{AverageSend } 1 + \text{AverageSend } 2)/3$$
$$\text{AverageSend } >3 = (\text{AverageSend } 4 + \text{AverageSend } 5 + \text{AverageSend } 6)/3$$

We expect to reject the following H0 in favour of H1.

$$\text{H0: AverageSend } <3 = \text{AverageSend } >3$$
$$\text{H1: AverageSend } <3 < \text{AverageSend } >3$$

Table 19: DWK treatment: social and personal appropriateness (t-tests)

	Obs (a)	Obs (b)	Mean (a)	Mean (b)	Diff. (a) - (b)	P-value (H0: Diff. = 0, H1: Diff. < 0)
Social appropriateness						
Sending<3 - Sending>3	157	157	-.031	.246	-.277	0
Personal appropriateness						
Sending<3 - Sending>3	144	144	.025	.093	-.067	.148

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 20: TG proposer: personal and social appropriateness (Rank sum test)

	Obs (a)	Obs (b)	Z-stat (a)	P-value (b)	P-value (exact) (H0: Diff. = 0, H1: Diff. < 0)
Social appropriateness					
Sending<3 - Sending>3	157	157	-4.114	0	0
Personal appropriateness					
Sending<3 - Sending>3	144	144	-1.018	.154	.155

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

1.4.2 TG responder

We hypothesise that sending back less than half is considered less appropriate than sending back more than half. Formally, let

$$\text{AverageReturn} <1/2 = (\text{AverageReturn } 0 + \text{AverageReturn } 1/4 + \text{AverageReturn } 1/3)/3$$
$$\text{AverageReturn} >1/2 = (\text{AverageReturn } 2/3 + \text{AverageReturn } 3/4)/2$$

We expect to reject the following H0 in favour of H1,

$$\text{H0: AverageReturn} <1/2 = \text{AverageReturn} >1/2$$
$$\text{H1: AverageReturn} <1/2 < \text{AverageReturn} >1/2$$

Table 21: DWK treatment: social and personal appropriateness (t-tests)

	Obs (a)	Obs (b)	Mean (a)	Mean (b)	Diff. (a) - (b)	P-value (H0: Diff. = 0, H1: Diff. < 0)
Social appropriateness						
A gets 3						
Sending $< 1/2$ - sending $> 1/2$	149	149	-.163	.648	-.811	0
A gets 9						
Sending $< 1/2$ - sending $> 1/2$	149	149	-.504	.674	-1.179	0
A gets 12						
Sending $< 1/2$ - sending $> 1/2$	149	149	-.518	.666	-1.184	0
A gets 18						
Sending $< 1/2$ - sending $> 1/2$	149	149	-.633	.653	-1.286	0
Personal appropriateness						
A gets 3						
Sending $< 1/2$ - sending $> 1/2$	149	149	-.186	.545	-.731	0
A gets 9						
Sending $< 1/2$ - sending $> 1/2$	149	149	-.449	.572	-1.021	0
A gets 12						
Sending $< 1/2$ - sending $> 1/2$	149	149	-.5	.529	-1.029	0
A gets 18						
Sending $< 1/2$ - sending $> 1/2$	149	149	-.591	.545	-1.136	0

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 22: TG responder: personal and social appropriateness (Rank sum test)

	Obs (a)	Obs (b)	Z-stat (a)	P-value (b)	P-value (exact) (H0: Diff. = 0, H1: Diff. < 0)
Social appropriateness					
A gets 3					
Sending $< 1/2$ - sending $> 1/2$	149	149	-9.536	0	0
A gets 9					
Sending $< 1/2$ - sending $> 1/2$	149	149	-10.315	0	0
A gets 12					
Sending $< 1/2$ - sending $> 1/2$	149	149	-10.198	0	0
A gets 18					
Sending $< 1/2$ - sending $> 1/2$	149	149	-10.413	0	0
Personal appropriateness					
A gets 3					
Sending $< 1/2$ - sending $> 1/2$	149	149	-9.275	0	0
A gets 9					
Sending $< 1/2$ - sending $> 1/2$	149	149	-10.133	0	0
A gets 12					
Sending $< 1/2$ - sending $> 1/2$	149	149	-10.223	0	0
A gets 18					
Sending $< 1/2$ - sending $> 1/2$	149	149	-10.213	0	0

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

2 Hypotheses and Tests: gender differences

We hypothesise that women will be more accurate at guessing the social appropriateness rating of others in each game. While we hypothesise that it is not the case that women's personal ratings being more likely (than men's personal ratings)

H0: Correctness of Women = correctness of Men
H1: Correctness of Wome > correctness of Me

2.1 Dictator games

Table 23: DG standard, with exit option, take 1: gender differences in social appropriateness (t-tests)

	Obs.	Mean (men)	Mean (women)	Diff.	P-value (H0: diff =0, H1: diff<0)
Standard DE	148	0.4460	0.4300	0.0160	0.6747
DE Take 1	157	0.4308	0.3715	0.0593	0.9425
DE take 5	135	0.4573	0.4830	-0.0257	0.2657
DE production	148	0.3165	0.3267	-0.0103	0.3448
DE quiz	155	0.3562	0.3137	0.0424	0.8939
DE with exit option	161	0.4247	0.3910	0.0337	0.8636

* p<0.05, ** p<0.01, *** p<0.001

Table 24: DE standard, with exit option, take 1: gender differences in social appropriateness (Rank sum test)

	Obs (Men) (a)	Obs (Women) (b)	Z-stat (a)	P-value (b)	P-value (exact) (H0: Diff. = 0, H1: Diff. < 0)
Standard DG	85	63	.401	.656	.656
DE Take 1	75	82	1.428	.923	.923
DE take 5	58	77	-.575	.283	.284
DE production	79	69	-.484	.314	.315
DE quiz	73	82	1.432	.924	.924
DE with exit option	83	78	1.039	.851	.85

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 25: DE standard, with exit option, take 1: gender differences in personal appropriateness (t-tests)

	Obs.	Mean (men)	Mean (women)	Diff.	P-value (H0: diff =0, H1: diff<0)
Standard DE	152	0.4033	0.3648	0.0386	0.9270
DE Take 1	125	0.3462	0.3704	-0.0243	0.1986
DE take 5	170	0.4216	0.4462	-0.0246	0.2276
DE production	152	0.2909	0.2857	0.0052	0.5941
DE quiz	150	0.2727	0.3034	-0.0307	0.1130
DE with exit option	152	0.3513	0.3733	-0.0220	0.2148

* p<0.05, ** p<0.01, *** p<0.001

Table 26: DE standard, with exit option, take 1: gender differences in personal appropriateness (Rank sum test)

	Obs (Men) (a)	Obs (Women) (b)	Z-stat (a)	P-value (b)	P-value (exact) (H0: Diff. = 0, H1: Diff. < 0)
Standard DE	71	81	1.462	.928	.928
DE Take 1	60	65	-.929	.176	.177
DE take 5	89	81	-.538	.295	.296
DE production	75	77	-.041	.483	.484
DE quiz	76	74	-1.247	.106	.107
DE with exit option	79	73	-.678	.249	.25

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

2.2 Information avoidance game

Table 27: DWK treatment: gender differences in social appropriateness (t-tests)

	Obs.	Mean (men)	Mean (women)	Diff.	P-value (H0: diff =0, H1: diff!=0)
DWK treatment	150	0.4383	0.5036	-0.0654	0.1501

* p<0.05, ** p<0.01, *** p<0.001

Table 28: DWK treatment: gender differences in social appropriateness (Rank sum test)

	Obs (Men) (a)	Obs (Women) (b)	Z-stat (a)	P-value (b)	P-value (exact) (H0: Diff. = 0, H1: Diff. < 0)
DWK treatment	81	69	-1.279	.1	.101

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 29: DWK treatment: gender differences in personal appropriateness (t-tests)

	Obs.	Mean (men)	Mean (women)	Diff.	P-value (H0: diff =0, H1: diff!=0)
DWK treatment	150	0.4414	0.5181	-0.0768	0.0995

* p<0.05, ** p<0.01, *** p<0.001

Table 30: DWK treatment: gender differences in personal appropriateness (Rank sum test)

	Obs (Men) (a)	Obs (Women) (b)	Z-stat (a)	P-value (b)	P-value (exact) (H0: Diff. = 0, H1: Diff. < 0)
DWK treatment	81	69	-1.867	.031	.031

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

2.3 Ultimatum and trust games

Table 31: UG and TG: gender differences in social appropriateness (t-tests)

	Obs.	Mean (men)	Mean (women)	Diff.	P-value (H0: diff =0, H1: diff<0)
Mini UG FFF28	132	0.3800	0.3509	0.0291	0.6600
Mini UG FFF55	158	0.3171	0.3947	-0.0777	0.0584
UG proposer	166	0.4773	0.4567	0.0206	0.7392
UG responder	145	0.4701	0.4588	0.0112	0.6190

* p<0.05, ** p<0.01, *** p<0.001

Table 32: UG and TG: gender differences in social appropriateness (Rank sum test)

	Obs (Men) (a)	Obs (Women) (b)	Z-stat (a)	P-value (b)	P-value (exact) (H0: Diff. = 0, H1: Diff. < 0)
Mini UG FFF28	75	57	.351	.637	.644
Mini UG FFF55	82	76	-1.609	.054	.058
UG proposer	80	86	.825	.795	.795
UG responder	64	81	.355	.639	.639

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 33: UG and TG: gender differences in personal appropriateness (t-tests)

	Obs.	Mean (men)	Mean (women)	Diff.	P-value (H0: diff =0, H1: diff<0)
Mini UG FFF28	140	0.3194	0.4118	-0.0923	0.0801
Mini UG FFF55	144	0.4058	0.4533	-0.0475	0.2039
UG proposer	167	0.4368	0.4481	-0.0113	0.3523
UG responder	149	0.4470	0.4688	-0.0218	0.2378

* p<0.05, ** p<0.01, *** p<0.001

Table 34: UG and TG: gender differences in social appropriateness (Rank sum test)

	Obs (Men) (a)	Obs (Women) (b)	Z-stat (a)	P-value (b)	P-value (exact) (H0: Diff. = 0, H1: Diff. < 0)
Mini UG FFF28	72	68	-1.513	.065	.066
Mini UG FFF55	69	75	-.806	.21	.223
UG proposer	82	85	-.482	.315	.316
UG responder	77	72	-.519	.302	.303

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

3 Hypotheses and Tests: Order effects

3.1 Test at level of questions/actions

For each action in each game, we compared the (i) average rating by subjects who saw that game as the first game in the whole experiment, with (ii) average rating by all other subjects (i.e. all subjects who did not see that game as the first game in the whole experiment), under the same mode of appropriateness. Use a z-test to compare the average ratings of these two groups.

3.1.1 Dictator games

Table 35: Standard DE: order effects in social appropriateness (t-tests)

	Obs.	Mean (others)	Mean (seen first)	Diff.	P-value
St. DE (0)	148	-0.6806	-0.5625	-0.1181	0.2436
St. DE (1)	148	-0.4954	-0.4250	-0.0704	0.5149
St. DE (2)	148	-0.3750	-0.3000	-0.0750	0.4784
St. DE (3)	148	-0.1713	-0.0500	-0.1213	0.2371
St. DE (4)	148	0.1574	0.1875	-0.0301	0.7525
St. DE (5)	148	0.7361	0.7375	-0.0014	0.9847
St. DE (6)	148	0.4722	0.5000	-0.0278	0.7772
St. DE (7)	148	0.3333	0.4000	-0.0667	0.5568
St. DE (8)	148	0.2454	0.4125	-0.1671	0.1901
St. DE (9)	148	0.2083	0.3750	-0.1667	0.2212
St. DE (10)	148	0.2130	0.3000	-0.0870	0.5502

* p<0.05, ** p<0.01, *** p<0.001

Table 36: Standard DE: order effects in social appropriateness (Rank sum test)

	Obs (a)	Obs (b)	Z-stat (a)	P-value (b)	P-value (exact) (H0: Diff. = 0, H1: Diff. < 0)
Standard DG	108	40	-.873	.191	.194
St. DE (1)	108	40	-.521	.301	.301
St. DE (2)	108	40	-.526	.299	.301
St. DE (3)	108	40	-1.281	.1	.101
St. DE (4)	108	40	-.487	.313	.314
St. DE (5)	108	40	-.043	.483	.488
St. DE (6)	108	40	-.142	.443	.446
St. DE (7)	108	40	-.591	.277	.28
St. DE (8)	108	40	-1.123	.131	.131
St. DE (9)	108	40	-1	.159	.16
St. DE (10)	108	40	-.427	.335	.336

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 37: Standard DE: order effects in personal appropriateness (t-tests)

	Obs.	Mean (others)	Mean (seen first)	Diff.	P-value
St. DE (0)	152	-0.6398	-0.4559	-0.1839	0.0887
St. DE (1)	152	-0.5339	-0.3235	-0.2104	0.0562
St. DE (2)	152	-0.4449	-0.2059	-0.2390*	0.0231
St. DE (3)	152	-0.2712	-0.0147	-0.2565*	0.0181
St. DE (4)	152	-0.0127	0.2353	-0.2480*	0.0248
St. DE (5)	152	0.7415	0.7353	0.0062	0.9342
St. DE (6)	152	0.3093	0.3529	-0.0436	0.6752
St. DE (7)	152	0.0890	0.1029	-0.0140	0.9034
St. DE (8)	152	-0.0169	0.0441	-0.0611	0.6308
St. DE (9)	152	-0.0847	0.0000	-0.0847	0.5311
St. DE (10)	152	-0.1441	-0.0147	-0.1294	0.3651

* p<0.05, ** p<0.01, *** p<0.001

Table 38: Standard DE: order effects in personal appropriateness (Rank sum test)

	Obs (a)	Obs (b)	Z-stat (a)	P-value (b)	P-value (exact) (H0: Diff. = 0, H1: Diff. < 0)
Standard DG	118	34	-2.079	.019	.02
St. DE (1)	118	34	-1.892	.029	.03
St. DE (2)	118	34	-1.951	.026	.026
St. DE (3)	118	34	-2.318	.01	.01
St. DE (4)	118	34	-2.264	.012	.012
St. DE (5)	118	34	-.201	.42	.427
St. DE (6)	118	34	-.557	.289	.29
St. DE (7)	118	34	-.208	.418	.418
St. DE (8)	118	34	-.594	.276	.278
St. DE (9)	118	34	-.615	.269	.271
St. DE (10)	118	34	-.9	.184	.185

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 39: DE take 1: order effects in social appropriateness (t-tests)

	Obs.	Mean (others)	Mean (seen first)	Diff.	P-value
DE take 2 (-2)	157	-0.5000	-0.5833	0.0833	0.7765
DE take 2 (-1)	157	-0.6783	-0.5357	-0.1425	0.0462
DE take 2 (0)	157	-0.2826	0.0595	-0.3421**	0.0017
DE take 2 (1)	157	-0.1348	0.4048	-0.5395***	0.0000
DE take 2 (2)	157	-0.0565	0.4286	-0.4851***	0.0000
DE take 2 (3)	157	0.0478	0.4524	-0.4046***	0.0002
DE take 2 (4)	157	0.3087	0.5119	-0.2032*	0.0160
DE take 2 (5)	157	0.8000	0.7381	0.0619	0.8078
DE take 2 (6)	157	0.4391	0.4405	-0.0013	0.4944
DE take 2 (7)	157	0.3174	0.3690	-0.0517	0.3174
DE take 2 (8)	157	0.2478	0.3095	-0.0617	0.3006
DE take 2 (9)	157	0.1609	0.2500	-0.0891	0.2513
DE take 2 (10)	157	0.1478	0.2143	-0.0665	0.3201

* p<0.05, ** p<0.01, *** p<0.001

Table 40: DE take 1: order effects in social appropriateness (Rank sum test)

	Obs (a)	Obs (b)	Z-stat (a)	P-value (b)	P-value (exact) (H0: Diff. = 0, H1: Diff. < 0)
DE take 2 (-2)	115	42	.635	.737	.737
DE take 2 (-1)	115	42	-1.565	.059	.06
DE take 2 (0)	115	42	-2.866	.002	.002
DE take 2 (1)	115	42	-4.203	0	0
DE take 2 (2)	115	42	-3.848	0	0
DE take 2 (3)	115	42	-3.5	0	0
DE take 2 (4)	115	42	-1.829	.034	.033
DE take 2 (5)	115	42	1.194	.884	.886
DE take 2 (6)	115	42	.168	.567	.568
DE take 2 (7)	115	42	-.388	.349	.351
DE take 2 (8)	115	42	-.51	.305	.306
DE take 2 (9)	115	42	-.635	.263	.267
DE take 2 (10)	115	42	-.482	.315	.316

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 41: DE take 1: order effects in personal appropriateness (t-tests)

	Obs.	Mean (others)	Mean (seen first)	Diff.	P-value
DE take 2 (-2)	125	-0.4890	-0.5294	0.0404	0.6401
DE take 2 (-1)	125	-0.5989	-0.4118	-0.1871	0.0381
DE take 2 (0)	125	-0.0879	0.1618	-0.2497	0.0379
DE take 2 (1)	125	-0.0604	0.3088	-0.3693**	0.0019
DE take 2 (2)	125	0.0330	0.3529	-0.3200**	0.0036
DE take 2 (3)	125	0.0934	0.2794	-0.1860	0.0548
DE take 2 (4)	125	0.3132	0.3529	-0.0398	0.3507
DE take 2 (5)	125	0.6868	0.5294	0.1574	0.9475
DE take 2 (6)	125	0.3022	0.0735	0.2287*	0.9772
DE take 2 (7)	125	0.0549	0.0294	0.0255	0.5811
DE take 2 (8)	125	0.0000	-0.0735	0.0735	0.7173
DE take 2 (9)	125	-0.0604	-0.1471	0.0866	0.7357
DE take 2 (10)	125	-0.1648	-0.1765	0.0116	0.5303

* p<0.05, ** p<0.01, *** p<0.001

Table 42: DE take 1: order effects in personal appropriateness (Rank sum test)

	Obs (a)	Obs (b)	Z-stat (a)	P-value (b)	P-value (exact) (H0: Diff. = 0, H1: Diff. < 0)
DE take 10 (-10)	91	34	-.464	.321	.323
DE take 2 (-1)	91	34	-2.479	.007	.007
DE take 2 (-1)	91	34	-2.479	.007	.007
DE take 2 (0)	91	34	-1.795	.036	.035
DE take 2 (1)	91	34	-2.895	.002	.002
DE take 2 (2)	91	34	-2.677	.004	.004
DE take 2 (3)	91	34	-1.513	.065	.066
DE take 2 (4)	91	34	-.457	.324	.327
DE take 2 (5)	91	34	1.447	.926	.927
DE take 2 (6)	91	34	2.005	.978	.978
DE take 2 (7)	91	34	.287	.613	.614
DE take 2 (8)	91	34	.634	.737	.738
DE take 2 (9)	91	34	.522	.699	.698
DE take 2 (10)	91	34	.017	.507	.509

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 43: DE take 5: order effects in social appropriateness (t-tests)

	Obs.	Mean (others)	Mean (seen first)	Diff.	P-value
DE take 10 (-10)	135	-0.8474	-0.8500	0.0026	0.5147
DE take 10 (-9)	135	-0.8105	-0.8250	0.0145	0.5786
DE take 10 (-8)	135	-0.8053	-0.8250	0.0197	0.6169
DE take 10 (-7)	135	-0.7105	-0.8000	0.0895	0.8712
DE take 10 (-6)	135	-0.6895	-0.7375	0.0480	0.7208
DE take 10 (-5)	135	-0.5947	-0.4750	-0.1197	0.1157
DE take 10 (-4)	135	-0.5579	-0.5500	-0.0079	0.4667
DE take 10 (-3)	135	-0.5263	-0.5000	-0.0263	0.3954
DE take 10 (-2)	135	-0.4000	-0.4500	0.0500	0.6742
DE take 10 (-1)	135	-0.3895	-0.3625	-0.0270	0.4051
DE take 10 (0)	135	0.0895	0.4000	-0.3105**	0.0043
DE take 10 (1)	135	0.2053	0.3250	-0.1197	0.1474
DE take 10 (2)	135	0.2105	0.3125	-0.1020	0.1863
DE take 10 (3)	135	0.2526	0.3750	-0.1224	0.1366
DE take 10 (4)	135	0.3632	0.3875	-0.0243	0.4076
DE take 10 (5)	135	0.6526	0.7125	-0.0599	0.2712
DE take 10 (6)	135	0.3737	0.4500	-0.0763	0.2423
DE take 10 (7)	135	0.3053	0.3875	-0.0822	0.2441
DE take 10 (8)	135	0.2211	0.3000	-0.0789	0.2666
DE take 10 (9)	135	0.2000	0.2625	-0.0625	0.3220
DE take 10 (10)	135	0.1263	0.2625	-0.1362	0.1735

* $p<0.05$, ** $p<0.01$, *** $p<0.001$

Table 44: DE take 5: order effects in social appropriateness (Rank sum test)

	Obs (a)	Obs (b)	Z-stat (a)	P-value (b)	P-value (exact) (H0: Diff. = 0, H1: Diff. < 0)
DE take 10 (-10)	95	40	.072	.529	.527
DE take 10 (-9)	95	40	.321	.626	.62
DE take 10 (-8)	95	40	.321	.626	.622
DE take 10 (-7)	95	40	1.067	.857	.863
DE take 10 (-6)	95	40	1.019	.846	.843
DE take 10 (-5)	95	40	-.929	.176	.177
DE take 10 (-4)	95	40	0	.5	.503
DE take 10 (-3)	95	40	-.39	.348	.35
DE take 10 (-2)	95	40	.245	.597	.597
DE take 10 (-1)	95	40	-.263	.396	.398
DE take 10 (0)	95	40	-2.771	.003	.003
DE take 10 (1)	95	40	-1.14	.127	.128
DE take 10 (2)	95	40	-1.01	.156	.159
DE take 10 (3)	95	40	-1.157	.124	.124
DE take 10 (4)	95	40	-.488	.313	.315
DE take 10 (5)	95	40	.243	.596	.6
DE take 10 (6)	95	40	-.679	.249	.253
DE take 10 (7)	95	40	-.743	.229	.23
DE take 10 (8)	95	40	-.62	.268	.271
DE take 10 (9)	95	40	-.442	.329	.331
DE take 10 (10)	95	40	-.917	.18	.181

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 45: DE take 5: order effects in personal appropriateness (t-tests)

	Obs.	Mean (others)	Mean (seen first)	Diff.	P-value
DE take 10 (-10)	171	-0.8202	-0.8158	-0.0044	0.4728
DE take 10 (-9)	170	-0.8009	-0.8246	0.0237	0.6455
DE take 10 (-8)	171	-0.7719	-0.8070	0.0351	0.7006
DE take 10 (-7)	171	-0.7412	-0.7632	0.0219	0.6295
DE take 10 (-6)	171	-0.7193	-0.7281	0.0088	0.5518
DE take 10 (-5)	171	-0.6360	-0.5965	-0.0395	0.3051
DE take 10 (-4)	171	-0.5877	-0.5175	-0.0702	0.1937
DE take 10 (-3)	171	-0.5395	-0.4825	-0.0570	0.2485
DE take 10 (-2)	171	-0.5175	-0.4298	-0.0877	0.1588
DE take 10 (-1)	171	-0.4474	-0.3772	-0.0702	0.2293
DE take 10 (0)	171	0.1579	0.3947	-0.2368*	0.0179
DE take 10 (1)	171	0.0921	0.2368	-0.1447	0.0750
DE take 10 (2)	171	0.1140	0.2632	-0.1491	0.0656
DE take 10 (3)	170	0.1150	0.2281	-0.1130	0.1211
DE take 10 (4)	171	0.1667	0.2632	-0.0965	0.1555
DE take 10 (5)	171	0.4912	0.5614	-0.0702	0.2127
DE take 10 (6)	171	0.1535	0.2018	-0.0482	0.3186
DE take 10 (7)	171	0.0439	0.1491	-0.1053	0.1595
DE take 10 (8)	171	-0.0263	0.1316	-0.1579	0.0791
DE take 10 (9)	171	-0.0965	0.0877	-0.1842	0.0582
DE take 10 (10)	171	-0.1579	0.1140	-0.2719*	0.0109

* $p<0.05$, ** $p<0.01$, *** $p<0.001$

Table 46: DE take 5: order effects in personal appropriateness (Rank sum test)

	Obs (a)	Obs (b)	Z-stat (a)	P-value (b)	P-value (exact) (H0: Diff. = 0, H1: Diff. < 0)
DE take 10 (-10)	114	57	.577	.718	.716
DE take 10 (-9)	113	57	1.042	.851	.852
DE take 10 (-8)	114	57	.927	.823	.826
DE take 10 (-7)	114	57	.577	.718	.718
DE take 10 (-6)	114	57	.074	.529	.532
DE take 10 (-5)	114	57	-1.019	.154	.155
DE take 10 (-4)	114	57	-1.394	.082	.082
DE take 10 (-3)	114	57	-.943	.173	.174
DE take 10 (-2)	114	57	-.957	.169	.17
DE take 10 (-1)	114	57	-.933	.175	.176
DE take 10 (0)	114	57	-1.96	.025	.025
DE take 10 (1)	114	57	-1.649	.05	.05
DE take 10 (2)	114	57	-1.732	.042	.042
DE take 10 (3)	113	57	-1.406	.08	.08
DE take 10 (4)	114	57	-1.174	.12	.121
DE take 10 (5)	114	57	-.808	.21	.211
DE take 10 (6)	114	57	-.584	.28	.281
DE take 10 (7)	114	57	-1.023	.153	.153
DE take 10 (8)	114	57	-1.417	.078	.078
DE take 10 (9)	114	57	-1.553	.06	.064
DE take 10 (10)	114	57	-2.229	.013	.013

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 47: DE quiz: order effects in social appropriateness (t-tests)

	Obs.	Mean (others)	Mean (seen first)	Diff.	P-value
DE Quiz (0)	155	-0.3099	-0.4853	0.1754	0.9030
DE Quiz (1)	155	-0.1364	-0.2647	0.1283	0.8485
DE Quiz (2)	155	-0.0331	-0.1471	0.1140	0.8386
DE Quiz (3)	155	0.0744	0.0000	0.0744	0.7541
DE Quiz (4)	155	0.2355	0.2500	-0.0145	0.4408
DE Quiz (5)	155	0.5744	0.6324	-0.0580	0.2522
DE Quiz (7)	155	0.2603	0.3529	-0.0926	0.2079
DE Quiz (8)	155	0.1860	0.2941	-0.1082	0.1926
DE Quiz (9)	155	0.0909	0.1618	-0.0709	0.3019
DE Quiz (10)	155	0.0455	0.1765	-0.1310	0.1920

* p<0.05, ** p<0.01, *** p<0.001

Table 48: DE quiz: order effects in social appropriateness (Rank sum test)

	Obs (a)	Obs (b)	Z-stat (a)	P-value (b)	P-value (exact) (H0: Diff. = 0, H1: Diff. < 0)
DE Quiz (0)	121	34	1.065	.857	.856
DE Quiz (0)	121	34	1.065	.857	.856
DE Quiz (1)	121	34	.898	.815	.815
DE Quiz (2)	121	34	.996	.84	.841
DE Quiz (3)	121	34	.734	.768	.768
DE Quiz (4)	121	34	.167	.566	.566
DE Quiz (5)	121	34	-.301	.382	.381
DE Quiz (7)	121	34	-.714	.238	.24
DE Quiz (8)	121	34	-.881	.189	.19
DE Quiz (9)	121	34	-.593	.277	.278
DE Quiz (10)	121	34	-.812	.208	.209

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 49: DE quiz: order effects in personal appropriateness (t-tests)

	Obs.	Mean (others)	Mean (seen first)	Diff.	P-value
DE Quiz (0)	155	-0.3099	-0.4853	0.1754	0.9030
DE Quiz (1)	155	-0.1364	-0.2647	0.1283	0.8485
DE Quiz (2)	155	-0.0331	-0.1471	0.1140	0.8386
DE Quiz (3)	155	0.0744	0.0000	0.0744	0.7541
DE Quiz (4)	155	0.2355	0.2500	-0.0145	0.4408
DE Quiz (5)	155	0.5744	0.6324	-0.0580	0.2522
DE Quiz (7)	155	0.2603	0.3529	-0.0926	0.2079
DE Quiz (8)	155	0.1860	0.2941	-0.1082	0.1926
DE Quiz (9)	155	0.0909	0.1618	-0.0709	0.3019
DE Quiz (10)	155	0.0455	0.1765	-0.1310	0.1920

* p<0.05, ** p<0.01, *** p<0.001

Table 50: DE quiz: order effects in pers appropriateness (Rank sum test)

	Obs (a)	Obs (b)	Z-stat (a)	P-value (b)	P-value (exact) (H0: Diff. = 0, H1: Diff. < 0)
DE Quiz (0)	121	29	-.352	.362	.363
DE Quiz (0)	121	29	-.352	.362	.363
DE Quiz (1)	121	29	.743	.771	.771
DE Quiz (2)	121	29	1.159	.877	.873
DE Quiz (3)	121	29	-.049	.48	.481
DE Quiz (4)	121	29	.253	.6	.6
DE Quiz (5)	121	29	.434	.668	.667
DE Quiz (7)	121	29	.465	.679	.682
DE Quiz (8)	121	29	.144	.557	.557
DE Quiz (9)	121	29	.093	.537	.538
DE Quiz (10)	121	29	.472	.682	.68

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 51: DG production: order effects in social appropriateness (t-tests)

	Obs.	Mean (others)	Mean (seen first)	Diff.	P-value
DE Prod (0)	148	-0.1157	-0.4625	0.3468*	0.9936
DE Prod (1)	148	-0.0463	-0.1875	0.1412	0.8546
DE Prod (3)	148	0.0509	0.0500	0.0009	0.5032
DE Prod (4)	148	0.2731	0.1875	0.0856	0.8037
DE Prod (5)	148	0.5231	0.5375	-0.0144	0.4411
DE Prod (6)	148	0.2917	0.1875	0.1042	0.8427
DE Prod (7)	148	0.1759	0.1125	0.0634	0.7106
DE Prod (8)	148	0.0833	0.1000	-0.0167	0.4492
DE Prod (9)	148	0.0370	0.0500	-0.0130	0.4630
DE Prod (10)	148	-0.0463	0.0625	-0.1088	0.2329

* p<0.05, ** p<0.01, *** p<0.001

Table 52: DE production: order effects in social appropriateness (Rank sum test)

	Obs (a)	Obs (b)	Z-stat (a)	P-value (b)	P-value (exact) (H0: Diff. = 0, H1: Diff. < 0)
DE Prod (0)	108	40	2.278	.989	.989
DE Prod (1)	108	40	.901	.816	.815
DE Prod (3)	108	40	.073	.529	.531
DE Prod (4)	108	40	.787	.784	.784
DE Prod (5)	108	40	.317	.624	.624
DE Prod (6)	108	40	1.268	.898	.894
DE Prod (7)	108	40	.637	.738	.738
DE Prod (8)	108	40	-.124	.451	.453
DE Prod (9)	108	40	-.115	.454	.452
DE Prod (10)	108	40	-.749	.227	.23

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 53: DE production: order effects in personal appropriateness (t-tests)

	Obs.	Mean (others)	Mean (seen first)	Diff.	P-value
DE Prod (0)	152	0.1154	-0.2000	0.3154*	0.9786
DE Prod (1)	152	0.0598	-0.1429	0.2027	0.9320
DE Prod (3)	152	0.1368	-0.0286	0.1653	0.9183
DE Prod (4)	152	0.1966	0.1000	0.0966	0.8082
DE Prod (5)	152	0.4359	0.4714	-0.0355	0.3798
DE Prod (6)	152	0.1410	0.0571	0.0839	0.7598
DE Prod (7)	152	-0.0342	-0.0571	0.0230	0.5741
DE Prod (8)	152	-0.1325	-0.1714	0.0389	0.6181
DE Prod (9)	152	-0.2094	-0.2857	0.0763	0.7141
DE Prod (10)	152	-0.2821	-0.3000	0.0179	0.5497

* p<0.05, ** p<0.01, *** p<0.001

Table 54: DE production: order effects in personal appropriateness (Rank sum test)

	Obs (a)	Obs (b)	Z-stat (a)	P-value (b)	P-value (exact) (H0: Diff. = 0, H1: Diff. < 0)
DE Prod (0)	117	35	1.981	.976	.976
DE Prod (1)	117	35	1.505	.934	.939
DE Prod (3)	117	35	1.451	.927	.926
DE Prod (4)	117	35	.931	.824	.824
DE Prod (5)	117	35	-.101	.46	.462
DE Prod (6)	117	35	.625	.734	.734
DE Prod (7)	117	35	.079	.531	.533
DE Prod (8)	117	35	.139	.555	.555
DE Prod (9)	117	35	.466	.679	.679
DE Prod (10)	117	35	.165	.566	.565

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 55: DE with exit option: order effects in social appropriateness (t-tests)

	Obs.	Mean (others)	Mean (seen first)	Diff.	P-value
DE exit (0)	161	-0.2051	0.0000	-0.2051	0.0858
DE exit (1)	161	-0.7564	-0.6591	-0.0973	0.2537
DE exit (2)	161	-0.5513	-0.2727	-0.2786**	0.0039
DE exit (3)	161	-0.4188	-0.1364	-0.2824**	0.0020
DE exit (4)	161	-0.2094	-0.0227	-0.1867*	0.0325
DE exit (5)	161	0.1068	0.1932	-0.0863	0.3227
DE exit (6)	161	0.7179	0.7614	-0.0434	0.5442
DE exit (7)	161	0.4188	0.5455	-0.1267	0.1547
DE exit (8)	161	0.2650	0.4318	-0.1669	0.1109
DE exit (9)	161	0.1752	0.4205	-0.2452*	0.0377
DE exit (10)	161	0.0855	0.4205	-0.3350*	0.0110
DE exit (11)	161	0.0171	0.3750	-0.3579*	0.0103

* p<0.05, ** p<0.01, *** p<0.001

Table 56: DE exit option: order effects in social appropriateness (Rank sum test)

	Obs (a)	Obs (b)	Z-stat (a)	P-value (b)	P-value (exact) (H0: Diff. = 0, H1: Diff. < 0)
DE exit (0)	117	44	-1.781	.037	.038
DE exit (1)	117	44	-2.096	.018	.02
DE exit (2)	117	44	-3.15	.001	.001
DE exit (3)	117	44	-2.78	.003	.003
DE exit (4)	117	44	-1.886	.03	.031
DE exit (5)	117	44	-1.159	.123	.124
DE exit (6)	117	44	-.759	.224	.224
DE exit (7)	117	44	-1.165	.122	.123
DE exit (8)	117	44	-1.401	.081	.081
DE exit (9)	117	44	-1.913	.028	.028
DE exit (10)	117	44	-2.323	.01	.01
DE exit (11)	117	44	-2.462	.007	.007

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 57: DE with exit option: order effects in personal appropriateness (t-tests)

	Obs.	Mean (others)	Mean (seen first)	Diff.	P-value
DE exit (0)	152	-0.1916	0.0333	-0.2249*	0.0499
DE exit (1)	152	-0.6636	-0.5333	-0.1302	0.1583
DE exit (2)	152	-0.3738	-0.1889	-0.1849	0.0839
DE exit (3)	152	-0.3131	-0.1667	-0.1464	0.1687
DE exit (4)	152	-0.1402	-0.0778	-0.0624	0.5433
DE exit (5)	152	0.1121	0.1444	-0.0323	0.7458
DE exit (6)	152	0.6449	0.7000	-0.0551	0.4792
DE exit (7)	152	0.2710	0.3333	-0.0623	0.5213
DE exit (8)	152	0.1215	0.1111	0.0104	0.9194
DE exit (9)	152	-0.0047	0.0556	-0.0602	0.5969
DE exit (10)	152	-0.0421	0.0222	-0.0643	0.6037
DE exit (11)	152	-0.1355	-0.0111	-0.1244	0.3553

* p<0.05, ** p<0.01, *** p<0.001

Table 58: DE exit option: order effects in personal appropriateness (Rank sum test)

	Obs (a)	Obs (b)	Z-stat (a)	P-value (b)	P-value (exact) (H0: Diff. = 0, H1: Diff. < 0)
DE exit (0)	107	45	-2.197	.014	.014
DE exit (1)	107	45	-1.586	.056	.058
DE exit (2)	107	45	-1.914	.028	.028
DE exit (3)	107	45	-1.489	.068	.069
DE exit (4)	107	45	-.815	.207	.208
DE exit (5)	107	45	-.221	.412	.415
DE exit (6)	107	45	-.464	.321	.321
DE exit (7)	107	45	-.398	.345	.346
DE exit (8)	107	45	.191	.576	.576
DE exit (9)	107	45	-.558	.288	.289
DE exit (10)	107	45	-.71	.239	.239
DE exit (11)	107	45	-1.052	.146	.147

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

3.1.2 Information avoidance game

Table 59: DWK treatment: order effects in social appropriateness (t-tests)

	Obs.	Mean (others)	Mean (seen first)	Diff.	P-value
A chooses X, uninf.	150	0.1010	0.1413	-0.0403	0.7104
A chooses Y, uninf.	150	0.1731	0.1196	0.0535	0.6191
A chooses X, inf.	150	-0.5192	-0.5000	-0.0192	0.8562
A chooses Y, inf.	150	0.8125	0.8478	-0.0353	0.6283

* p<0.05, ** p<0.01, *** p<0.001

Table 60: DWK treatment: order effects in social appropriateness (Rank sum test)

	Obs (a)	Obs (b)	Z-stat (a)	P-value (b)	P-value (exact) (H0: Diff. = 0, H1: Diff. < 0)
A chooses X, uninf.	104	46	-.382	.351	.353
A chooses Y, uninf.	104	46	.589	.722	.721
A chooses X, inf.	104	46	-.232	.408	.41
A chooses Y, inf.	104	46	.066	.526	.526

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 61: DWK treatment: order effects in personal appropriateness (t-tests)

	Obs.	Mean (others)	Mean (seen first)	Diff.	P-value
A chooses X, uninf.	150	0.2062	0.3774	-0.1712	0.0553
A chooses Y, uninf.	150	0.2165	0.3585	-0.1420	0.1279
A chooses X, inf.	150	-0.5412	-0.4057	-0.1356	0.1115
A chooses Y, inf.	150	0.8041	0.7547	0.0494	0.4699

* p<0.05, ** p<0.01, *** p<0.001

Table 62: DWK treatment: order effects in social appropriateness (Rank sum test)

	Obs (a)	Obs (b)	Z-stat (a)	P-value (b)	P-value (exact) (H0: Diff. = 0, H1: Diff. < 0)
A chooses X, uninf.	97	53	-1.852	.032	.032
A chooses Y, uninf.	97	53	-1.419	.078	.079
A chooses X, inf.	97	53	-1.311	.095	.096
A chooses Y, inf.	97	53	-.409	.341	.344

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

3.1.3 Ultimatum games

Table 63: Mini UG: order effects in social appropriateness (t-tests)

	Obs.	Mean (others)	Mean (seen first)	Diff.	P-value
Accept	132	0.2222	0.3939	-0.1717	0.1533
Reject FFF28	132	-0.1566	-0.2576	0.1010	0.4067
Accept	158	0.2009	0.1413	0.0596	0.5659
Reject FFF55	158	-0.1027	0.0652	-0.1679	0.1511

* p<0.05, ** p<0.01, *** p<0.001

Table 64: Mini UG: order effects in social appropriateness (Rank sum test)

	Obs (a)	Obs (b)	Z-stat (a)	P-value (b)	P-value (exact) (H0: Diff. = 0, H1: Diff. < 0)
B reject offers less than 0	99	33	-1.227	.11	.11
Reject FFF28	99	33	.655	.744	.744
Accept	112	46	.454	.675	.674
Reject FFF55	112	46	-1.467	.071	.067

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 65: Mini UG: order effects in personal appropriateness (t-tests)

	Obs.	Mean (others)	Mean (seen first)	Diff.	P-value
Accept	140	0.3039	0.2763	0.0276	0.7949
Reject FFF28	140	-0.2059	-0.1316	-0.0743	0.5151
Accept	144	0.1105	0.0612	0.0493	0.6001
Reject FFF55	144	0.0526	0.1939	-0.1412	0.1914

* p<0.05, ** p<0.01, *** p<0.001

Table 66: Mini UG: order effects in pers appropriateness (Rank sum test)

	Obs (a)	Obs (b)	Z-stat (a)	P-value (b)	P-value (exact) (H0: Diff. = 0, H1: Diff. < 0)
Accept	102	38	.313	.623	.624
Reject FFF28	102	38	-.709	.239	.241
Accept	95	49	.6	.726	.725
Reject FFF55	95	49	-1.04	.149	.15

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 67: UG proposer: order effects in social appropriateness (t-tests)

	Obs.	Mean (others)	Mean (seen first)	Diff.	P-value
Accept	132	0.2222	0.3939	-0.1717	0.1533
Reject FFF28	132	-0.1566	-0.2576	0.1010	0.4067
Accept	158	0.2009	0.1413	0.0596	0.5659
Reject FFF55	158	-0.1027	0.0652	-0.1679	0.1511

* p<0.05, ** p<0.01, *** p<0.001

Table 68: UG proposer: order effects in social appropriateness (Rank sum test)

	Obs (a)	Obs (b)	Z-stat (a)	P-value (b)	P-value (exact) (H0: Diff. = 0, H1: Diff. < 0)
B reject offers less than 0	116	50	1.518	.935	.935
B gets 1	116	50	1.56	.941	.943
B gets 2	116	50	1.004	.842	.844
B gets 3	116	50	.778	.782	.781
B gets 4	116	50	.536	.704	.703
B gets 5	116	50	-.921	.179	.183
B gets 6	116	50	.467	.68	.68
B gets 7	116	50	.033	.513	.514
B gets 8	116	50	.646	.741	.741
B gets 9	116	50	.436	.669	.669
B gets 10	116	50	.647	.741	.741

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 69: UG proposer: order effects in personal appropriateness (t-tests)

	Obs.	Mean (others)	Mean (seen first)	Diff.	P-value
Accept	140	0.3039	0.2763	0.0276	0.7949
Reject FFF28	140	-0.2059	-0.1316	-0.0743	0.5151
Accept	144	0.1105	0.0612	0.0493	0.6001
Reject FFF55	144	0.0526	0.1939	-0.1412	0.1914

* p<0.05, ** p<0.01, *** p<0.001

Table 70: UG proposer: order effects in personal appropriateness (Rank sum test)

	Obs (a)	Obs (b)	Z-stat (a)	P-value (b)	P-value (exact) (H0: Diff. = 0, H1: Diff. < 0)
Accept	121	46	.662	.746	.74
B gets 1	121	46	1.161	.877	.88
B gets 2	121	46	.948	.828	.827
B gets 3	121	46	-.174	.431	.431
B gets 4	121	46	.798	.787	.787
B gets 5	121	46	-.831	.203	.211
B gets 6	121	46	-.052	.479	.481
B gets 7	121	46	.329	.629	.628
B gets 8	121	46	.451	.674	.674
B gets 9	121	46	-.099	.46	.461
B gets 10	121	46	-.268	.394	.395

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 71: UG proposer: order effects in social appropriateness (t-tests)

	Obs.	Mean (others)	Mean (seen first)	Diff.	P-value
Accept	132	0.2222	0.3939	-0.1717	0.1533
Reject FFF28	132	-0.1566	-0.2576	0.1010	0.4067
Accept	158	0.2009	0.1413	0.0596	0.5659
Reject FFF55	158	-0.1027	0.0652	-0.1679	0.1511

* p<0.05, ** p<0.01, *** p<0.001

Table 72: UG responder: order effects in personal appropriateness (Rank sum test)

	Obs (a)	Obs (b)	Z-stat (a)	P-value (b)	P-value (exact) (H0: Diff. = 0, H1: Diff. < 0)
B reject offers less than 0	91	54	.676	.751	.752
B reject offers less than 1	91	54	-1.692	.045	.046
B reject offers less than 2	91	54	-2.334	.01	.01
B reject offers less than 3	91	54	-1.545	.061	.061
B reject offers less than 4	91	54	-.936	.175	.174
B reject offers less than 5	91	54	-.103	.459	.459
B reject offers less than 6	91	54	1.193	.884	.884
B reject offers less than 7	91	54	1.03	.848	.849
B reject offers less than 8	91	54	.915	.82	.82
B reject offers less than 9	91	54	.921	.821	.822
B reject offers less than 10	91	54	1.319	.906	.907
B reject all offers	91	54	1.271	.898	.897

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 73: UG responder: order effects in personal appropriateness (t-tests)

	Obs.	Mean (others)	Mean (seen first)	Diff.	P-value
Accept	140	0.3039	0.2763	0.0276	0.7949
Reject FFF28	140	-0.2059	-0.1316	-0.0743	0.5151
Accept	144	0.1105	0.0612	0.0493	0.6001
Reject FFF55	144	0.0526	0.1939	-0.1412	0.1914

* p<0.05, ** p<0.01, *** p<0.001

Table 74: UF responder: order effects in personal appropriateness (Rank sum test)

	Obs (a)	Obs (b)	Z-stat (a)	P-value (b)	P-value (exact) (H0: Diff. = 0, H1: Diff. < 0)
B reject offers less than 0	102	47	.038	.515	.52
B reject offers less than 1	102	47	-.308	.379	.381
B reject offers less than 2	102	47	-.265	.395	.397
B reject offers less than 3	102	47	.156	.562	.563
B reject offers less than 4	102	47	.993	.84	.84
B reject offers less than 5	102	47	1.532	.937	.938
B reject offers less than 6	102	47	1.581	.943	.942
B reject offers less than 7	102	47	1.091	.862	.862
B reject offers less than 8	102	47	1.124	.87	.869
B reject offers less than 9	102	47	1.329	.908	.908

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

3.1.4 Trust games

Table 75: TG proposer: order effects in personal appropriateness (t-tests)

	Obs.	Mean (others)	Mean (seen first)	Diff.	P-value
Accept	132	0.2222	0.3939	-0.1717	0.1533
Reject FFF28	132	-0.1566	-0.2576	0.1010	0.4067
Accept	158	0.2009	0.1413	0.0596	0.5659
Reject FFF55	158	-0.1027	0.0652	-0.1679	0.1511

* p<0.05, ** p<0.01, *** p<0.001

Table 76: TG responder: order effects in social appropriateness (Rank sum test)

	Obs (a)	Obs (b)	Z-stat (a)	P-value (b)	P-value (exact) (H0: Diff. = 0, H1: Diff. < 0)
A keep 6, B send 0	108	49	-.268	.394	.393
A keep 5, B send 1	108	49	-.135	.446	.44
A keep 4, B send 2	108	49	.004	.502	.504
A keep 3, B send 3	108	49	-1.171	.121	.122
A keep 2, B send 4	108	49	-1.709	.044	.044
A keep 1, B send 5	108	49	-1.482	.069	.07
A keep 0, B send 6	108	49	-1.141	.127	.128

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 77: TG responder: order effects in personal appropriateness (t-tests)

	Obs.	Mean (others)	Mean (seen first)	Diff.	P-value
Accept	140	0.3039	0.2763	0.0276	0.7949
Reject FFF28	140	-0.2059	-0.1316	-0.0743	0.5151
Accept	144	0.1105	0.0612	0.0493	0.6001
Reject FFF55	144	0.0526	0.1939	-0.1412	0.1914

* p<0.05, ** p<0.01, *** p<0.001

Table 78: TG responder: order effects in personal appropriateness (Rank sum test)

	Obs (a)	Obs (b)	Z-stat (a)	P-value (b)	P-value (exact) (H0: Diff. = 0, H1: Diff. < 0)
A keep 6, B send 0	102	42	-1.11	.133	.134
A keep 6, B send 0	102	42	-1.11	.133	.134
A keep 5, B send 1	102	42	-2.305	.011	.01
A keep 4, B send 2	102	42	-2.411	.008	.008
A keep 3, B send 3	102	42	-1.841	.033	.033
A keep 2, B send 4	102	42	-1.838	.033	.033
A keep 1, B send 5	102	42	-1.58	.057	.058
A keep 0, B send 6	102	42	-1.772	.038	.041

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 79: TG responder 3: order effects in personal appropriateness (t-tests)

	Obs.	Mean (others)	Mean (seen first)	Diff.	P-value
Accept	132	0.2222	0.3939	-0.1717	0.1533
Reject FFF28	132	-0.1566	-0.2576	0.1010	0.4067
Accept	158	0.2009	0.1413	0.0596	0.5659
Reject FFF55	158	-0.1027	0.0652	-0.1679	0.1511

* p<0.05, ** p<0.01, *** p<0.001

Table 80: TG responder: order effects in social appropriateness (Runk sum test)

	Obs (a)	Obs (b)	Z-stat (a)	P-value (b)	P-value (exact) (H0: Diff. = 0, H1: Diff. < 0)
A gets 3, B gets 15	117	32	3.238	.999	1
A gets 8.25, B gets 5.75	117	32	2.189	.986	.986
A gets 8, B gets 6	117	32	2.147	.984	.984
A gets 7.50, B gets 6.50	117	32	.879	.81	.81
A gets 7, B gets 7	117	32	-1.756	.04	.037
A gets 6.75, B gets 7.25	117	32	.226	.589	.591

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 81: TG responder 3: order effects in personal appropriateness (t-tests)

	Obs.	Mean (others)	Mean (seen first)	Diff.	P-value
Accept	140	0.3039	0.2763	0.0276	0.7949
Reject FFF28	140	-0.2059	-0.1316	-0.0743	0.5151
Accept	144	0.1105	0.0612	0.0493	0.6001
Reject FFF55	144	0.0526	0.1939	-0.1412	0.1914

* p<0.05, ** p<0.01, *** p<0.001

Table 82: TG responder: order effects in personal appropriateness (Rank sum test)

	Obs (a)	Obs (b)	Z-stat (a)	P-value (b)	P-value (exact) (H0: Diff. = 0, H1: Diff. < 0)
A gets 3, B gets 15	110	39	-.395	.347	.348
A gets 9, B gets 5	110	39	-.395	.347	.348
A gets 8.25, B gets 5.75	110	39	-1.18	.119	.121
A gets 8, B gets 6	110	39	-.666	.253	.254
A gets 7.50, B gets 6.50	110	39	-.158	.437	.438
A gets 7, B gets 7	110	39	.414	.661	.668
A gets 6.75, B gets 7.25	110	39	1.279	.9	.899

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 83: TG responder 9: order effects in personal appropriateness (t-tests)

	Obs.	Mean (others)	Mean (seen first)	Diff.	P-value
Accept	132	0.2222	0.3939	-0.1717	0.1533
Reject FFF28	132	-0.1566	-0.2576	0.1010	0.4067
Accept	158	0.2009	0.1413	0.0596	0.5659
Reject FFF55	158	-0.1027	0.0652	-0.1679	0.1511

* p<0.05, ** p<0.01, *** p<0.001

Table 84: TG responder: order effects in social appropriateness (Rank sum test)

	Obs (a)	Obs (b)	Z-stat (a)	P-value (b)	P-value (exact) (H0: Diff. = 0, H1: Diff. < 0)
A gets 3, B gets 15	108	41	.753	.774	.773
A gets 3, B gets 15	108	41	.753	.774	.773
A gets 12.75, B gets 5.25	108	41	-.095	.462	.462
A gets 12, B gets 6	108	41	-.821	.206	.207
A gets 10.50, B gets 7.50	108	41	-1.253	.105	.106
A gets 9, B gets 9	108	41	-1.804	.036	.035
A gets 8.25, B gets 9.75	108	41	-.484	.314	.317

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 85: TG responder 9: order effects in personal appropriateness (t-tests)

	Obs.	Mean (others)	Mean (seen first)	Diff.	P-value
Accept	140	0.3039	0.2763	0.0276	0.7949
Reject FFF28	140	-0.2059	-0.1316	-0.0743	0.5151
Accept	144	0.1105	0.0612	0.0493	0.6001
Reject FFF55	144	0.0526	0.1939	-0.1412	0.1914

* p<0.05, ** p<0.01, *** p<0.001

Table 86: TG responder: order effects in social appropriateness (Rank sum test)

	Obs (a)	Obs (b)	Z-stat (a)	P-value (b)	P-value (exact) (H0: Diff. = 0, H1: Diff. < 0)
A gets 3, B gets 15	118	31	1.722	.957	.959
A gets 12.75, B gets 5.25	118	31	.949	.829	.832
A gets 12, B gets 6	118	31	.676	.75	.75
A gets 10.50, B gets 7.50	118	31	.31	.622	.624
A gets 9, B gets 9	118	31	-1.085	.139	.145
A gets 8.25, B gets 9.75	118	31	.336	.632	.636

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 87: TG responder 12: order effects in personal appropriateness (t-tests)

	Obs.	Mean (others)	Mean (seen first)	Diff.	P-value
Accept	132	0.2222	0.3939	-0.1717	0.1533
Reject FFF28	132	-0.1566	-0.2576	0.1010	0.4067
Accept	158	0.2009	0.1413	0.0596	0.5659
Reject FFF55	158	-0.1027	0.0652	-0.1679	0.1511

* p<0.05, ** p<0.01, *** p<0.001

Table 88: TG responder: order effects in social appropriateness (Rank sum test)

	Obs (a)	Obs (b)	Z-stat (a)	P-value (b)	P-value (exact) (H0: Diff. = 0, H1: Diff. < 0)
A gets 18, B gets 2	106	43	1.557	.94	.947
A gets 15, B gets 5	106	43	.728	.767	.767
A gets 14, B gets 6	106	43	.449	.673	.671
A gets 12, B gets 8	106	43	.559	.712	.713
A gets 10, B gets 10	106	43	-1.144	.126	.133
A gets 9, B gets 11	106	43	-.209	.417	.418

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 89: TG responder 12: order effects in personal appropriateness (t-tests)

	Obs.	Mean (others)	Mean (seen first)	Diff.	P-value
Accept	140	0.3039	0.2763	0.0276	0.7949
Reject FFF28	140	-0.2059	-0.1316	-0.0743	0.5151
Accept	144	0.1105	0.0612	0.0493	0.6001
Reject FFF55	144	0.0526	0.1939	-0.1412	0.1914

* p<0.05, ** p<0.01, *** p<0.001

Table 90: TG responder: order effects in social appropriateness (Rank sum test)

	Obs (a)	Obs (b)	Z-stat (a)	P-value (b)	P-value (exact) (H0: Diff. = 0, H1: Diff. < 0)
A gets 18, B gets 2	97	52	-.309	.379	.378
A gets 15, B gets 5	97	52	-.821	.206	.207
A gets 14, B gets 6	97	52	-.409	.341	.342
A gets 12, B gets 8	97	52	-1.554	.06	.06
A gets 10, B gets 10	97	52	-1.138	.128	.127
A gets 9, B gets 11	97	52	.547	.708	.709

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 91: TG responder 18: order effects in personal appropriateness (t-tests)

	Obs.	Mean (others)	Mean (seen first)	Diff.	P-value
Accept	132	0.2222	0.3939	-0.1717	0.1533
Reject FFF28	132	-0.1566	-0.2576	0.1010	0.4067
Accept	158	0.2009	0.1413	0.0596	0.5659
Reject FFF55	158	-0.1027	0.0652	-0.1679	0.1511

* p<0.05, ** p<0.01, *** p<0.001

Table 92: TG responder: order effects in social appropriateness (Rank sum test)

	Obs (a)	Obs (b)	Z-stat (a)	P-value (b)	P-value (exact) (H0: Diff. = 0, H1: Diff. < 0)
A gets 24, B gets 0	106	43	2.286	.989	.992
A gets 19.50, B gets 4.50	106	43	.962	.832	.833
A gets 18, B gets 6	106	43	-.322	.374	.375
A gets 15, B gets 9	106	43	.048	.519	.521
A gets 12, B gets 12	106	43	-1.198	.115	.12
A gets 10.50, B gets 13.50	106	43	-.509	.305	.307

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 93: TG responder 18: order effects in personal appropriateness (t-tests)

	Obs.	Mean (others)	Mean (seen first)	Diff.	P-value
Accept	140	0.3039	0.2763	0.0276	0.7949
Reject FFF28	140	-0.2059	-0.1316	-0.0743	0.5151
Accept	144	0.1105	0.0612	0.0493	0.6001
Reject FFF55	144	0.0526	0.1939	-0.1412	0.1914

* p<0.05, ** p<0.01, *** p<0.001

Table 94: TG responder: order effects in personal appropriateness (Rank sum test)

	Obs (a)	Obs (b)	Z-stat (a)	P-value (b)	P-value (exact) (H0: Diff. = 0, H1: Diff. < 0)
A gets 24, B gets 0	97	52	.317	.624	.617
A gets 19.50, B gets 4.50	97	52	.346	.635	.638
A gets 18, B gets 6	97	52	-1.084	.139	.14
A gets 15, B gets 9	97	52	-.129	.449	.45
A gets 12, B gets 12	97	52	-.951	.171	.17
A gets 10.50, B gets 13.50	97	52	.221	.587	.588

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

4 Figures

This section shows the average rating for each payoff in every game of the experiment.

4.1 Dictator games

Figure 1: Average ratings in the standard DE and DE with exit option

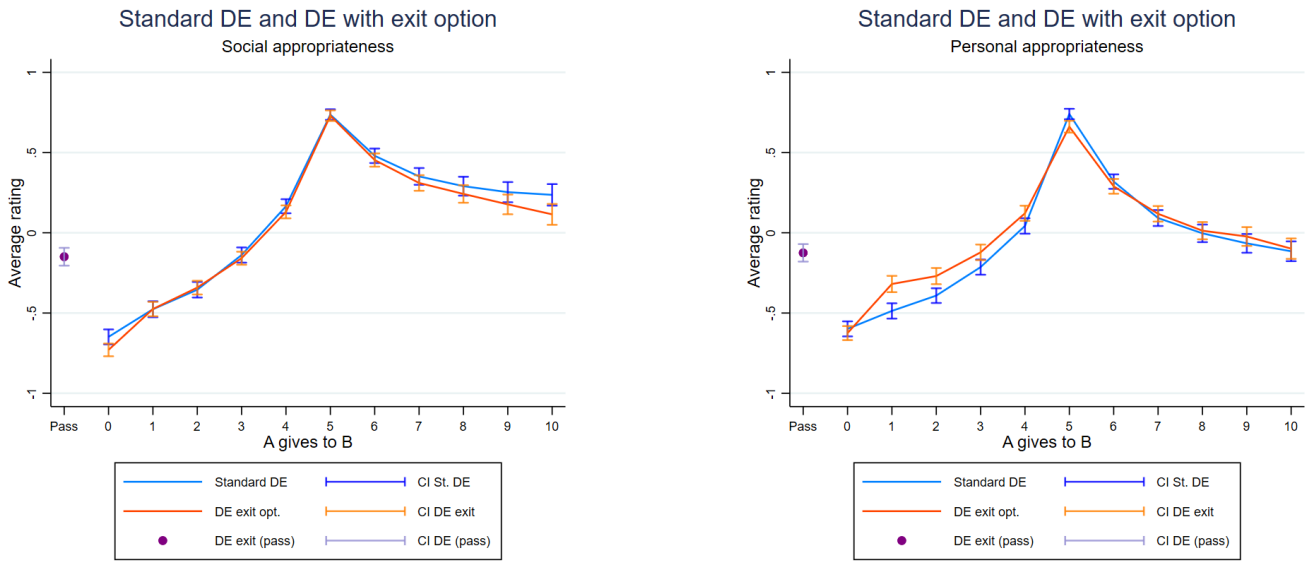


Figure 2: Average ratings in the standard DE and DE with take options

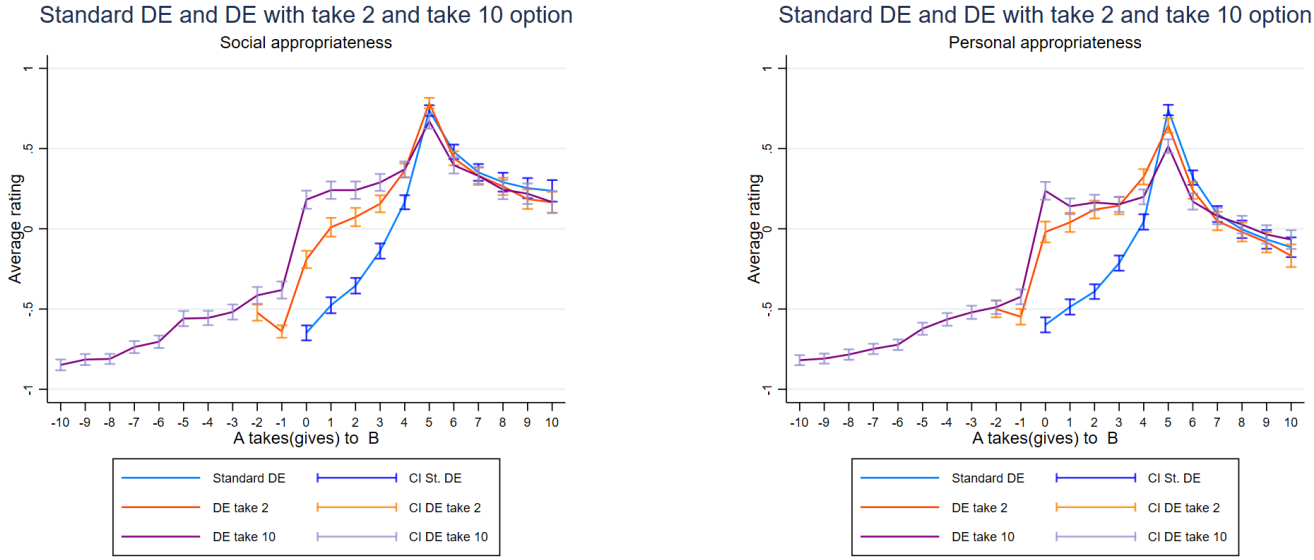
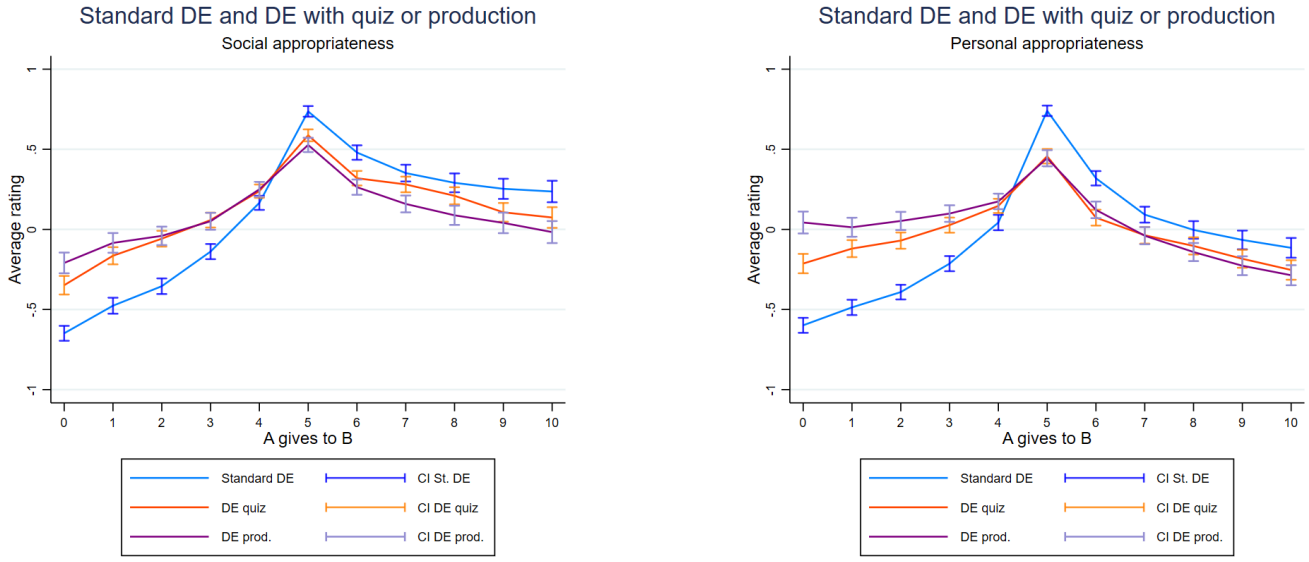


Figure 3: Average ratings in the standard DE and DE with surplus



4.2 Information avoidance games

Figure 4: Average ratings in the DWK treatment game

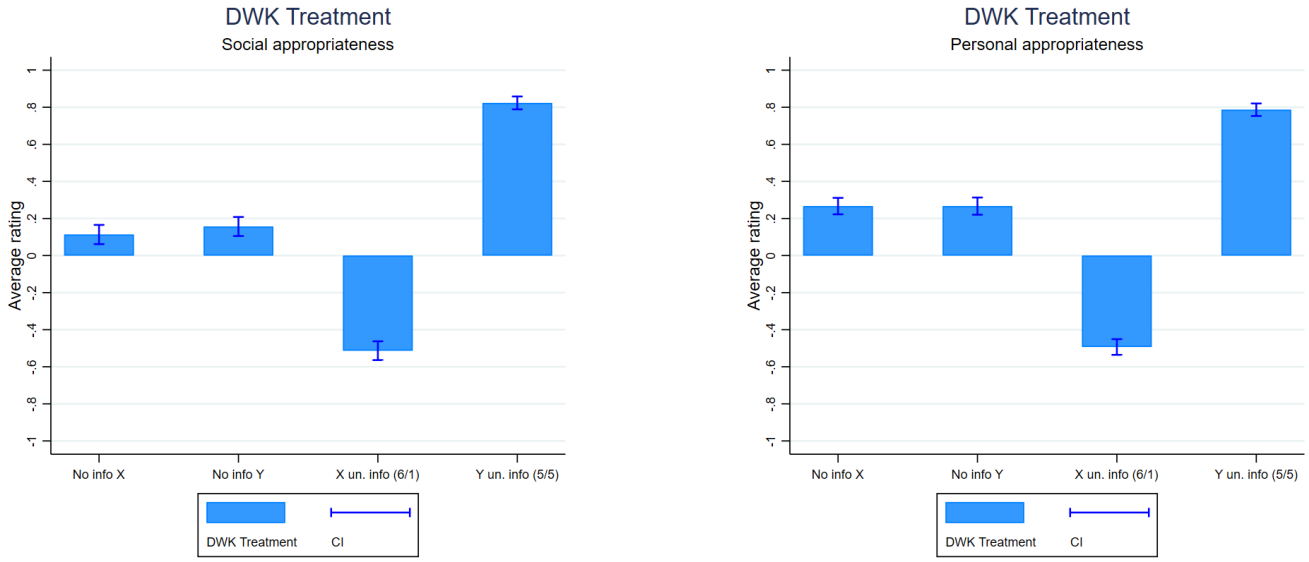
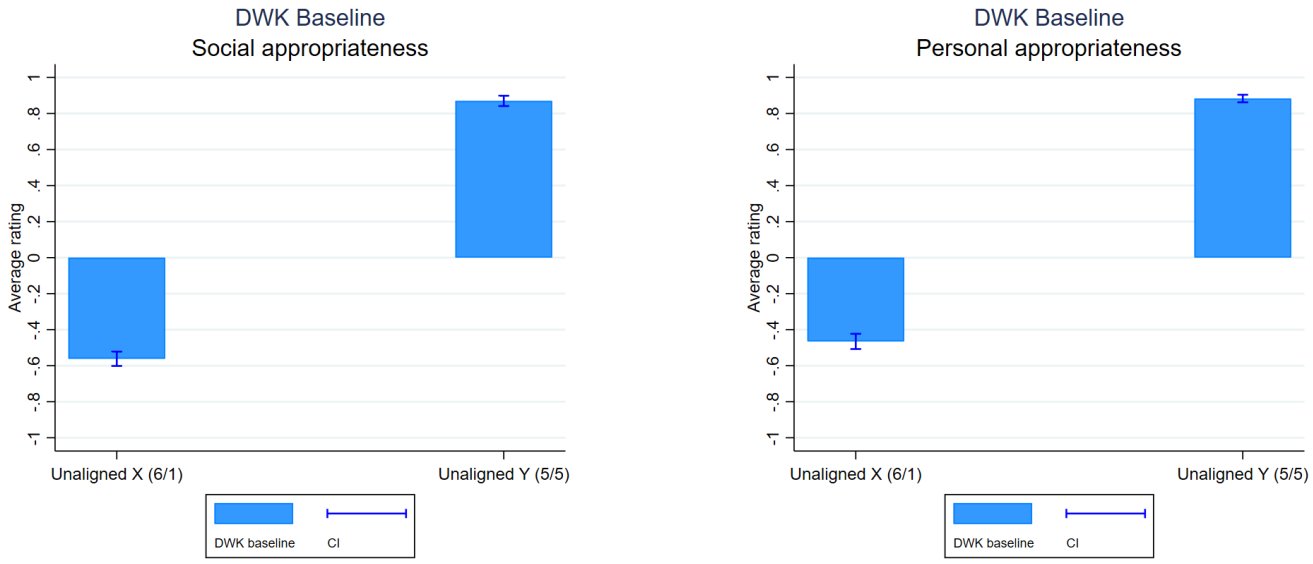


Figure 5: Average ratings in the DWK baseline game



4.3 Ultimatum and trust game

Figure 6: Average ratings in the mini ultimatum game (FFF 5:5 and 2:8)

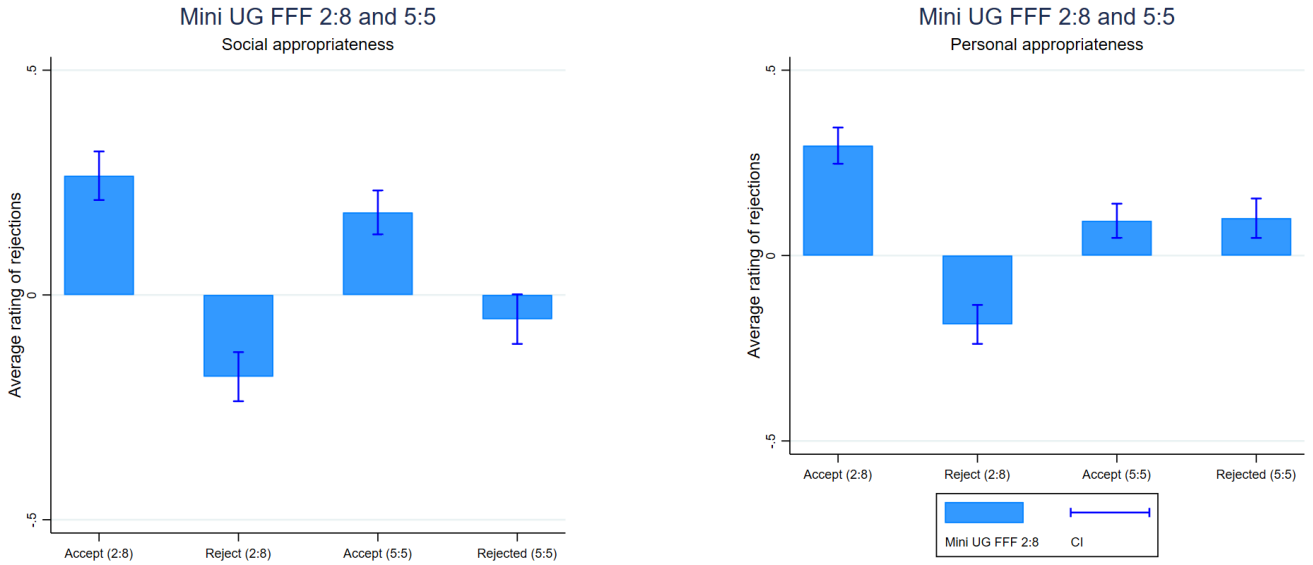


Figure 7: Average ratings in the mini ultimatum game (FFF 2:8)

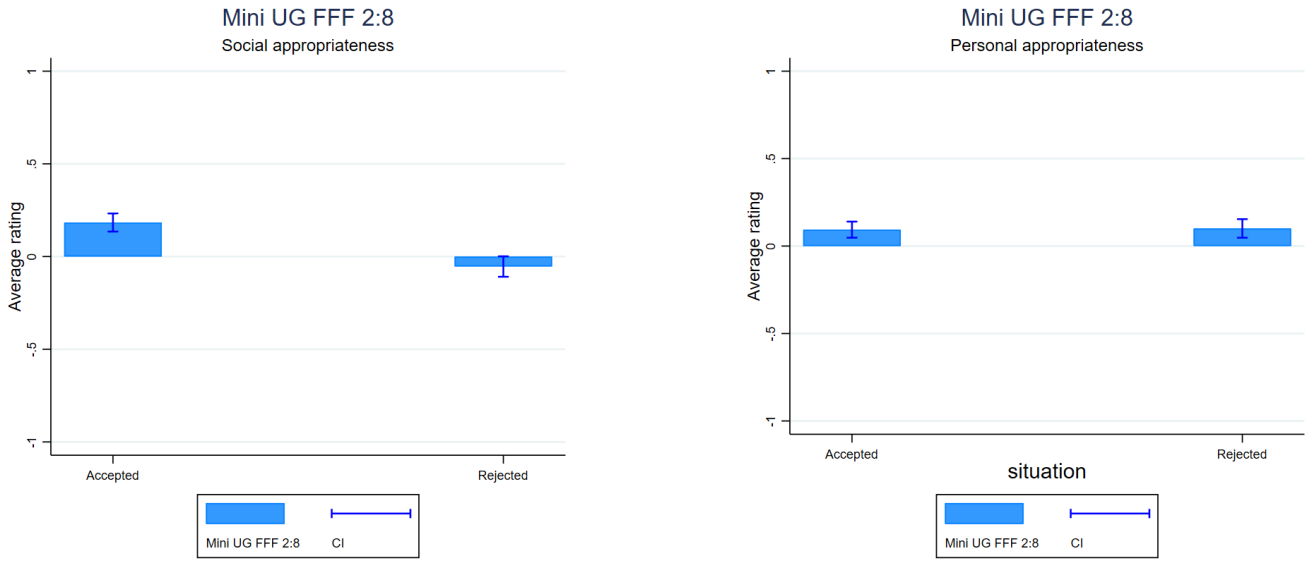


Figure 8: Average ratings in the UG proposer game

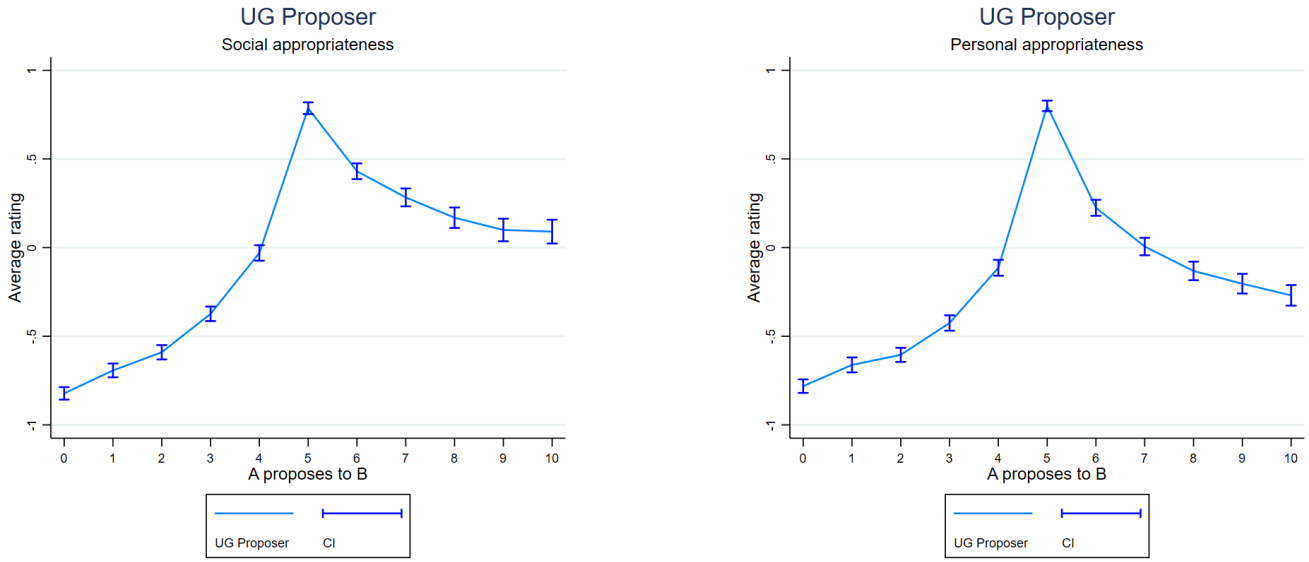


Figure 9: Average ratings in the UG responder game

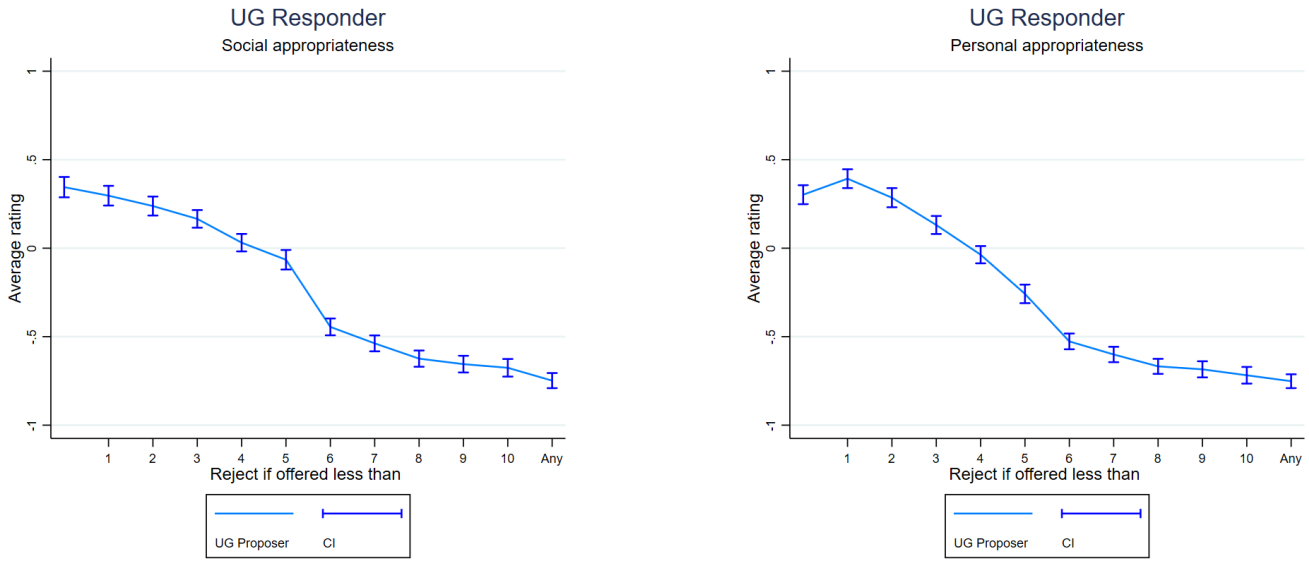


Figure 10: Average ratings in the TG proposer game

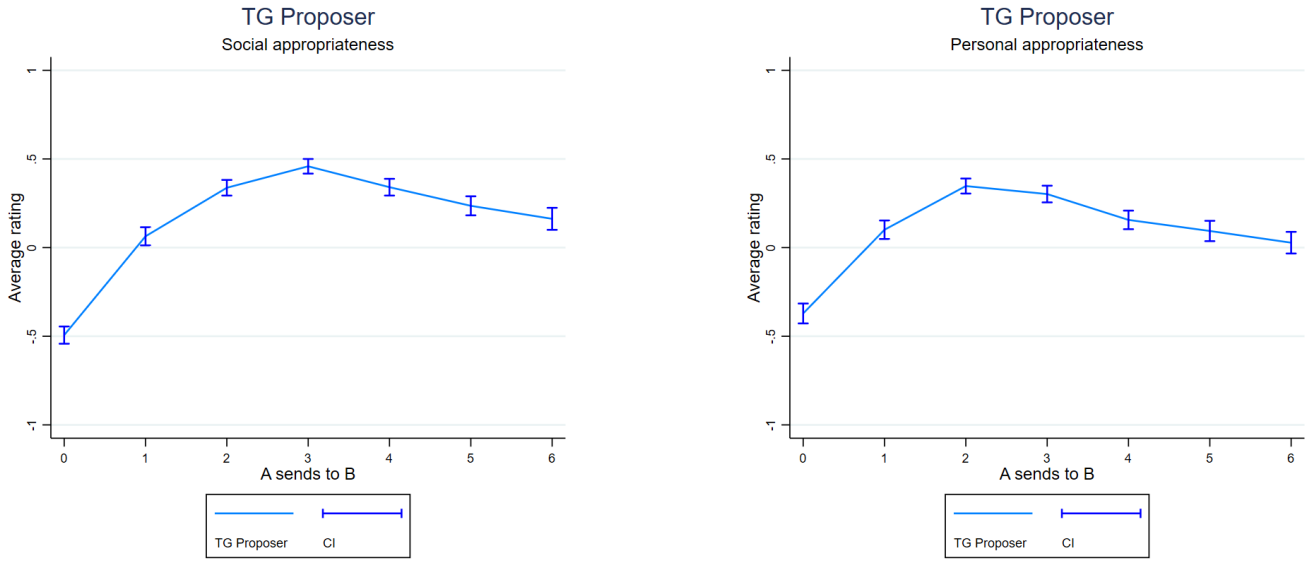
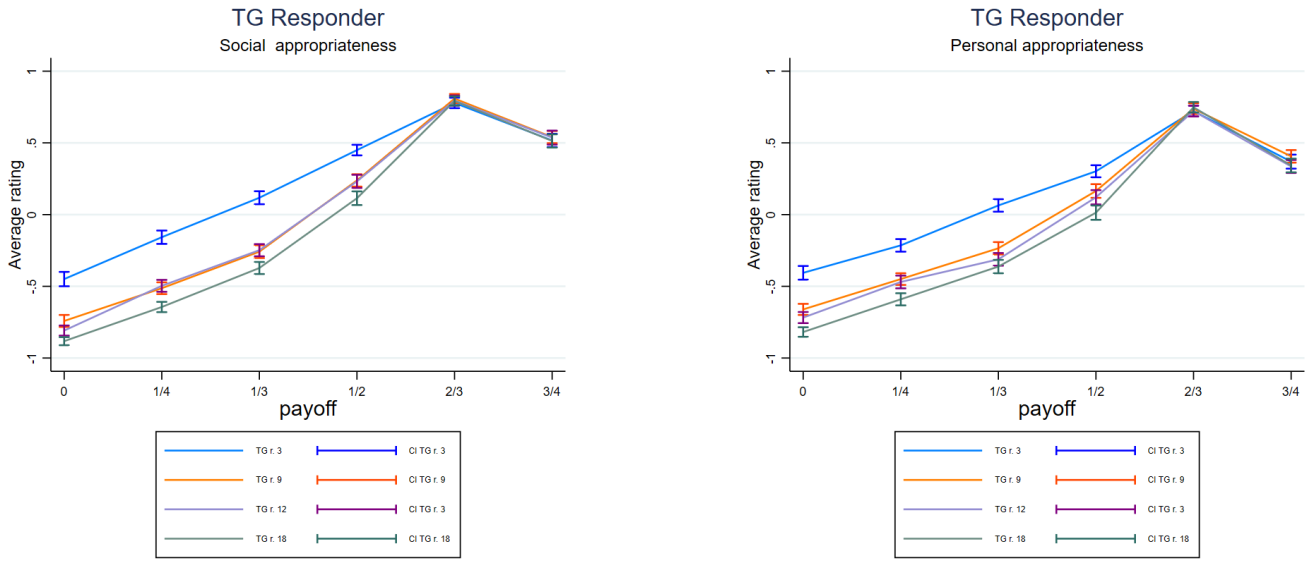


Figure 11: Average ratings in the TG responder game



4.4 Dictator games ratings accounting for the order of games

Figure 12: Average ratings in the Standard DE for subject who met this game as their first

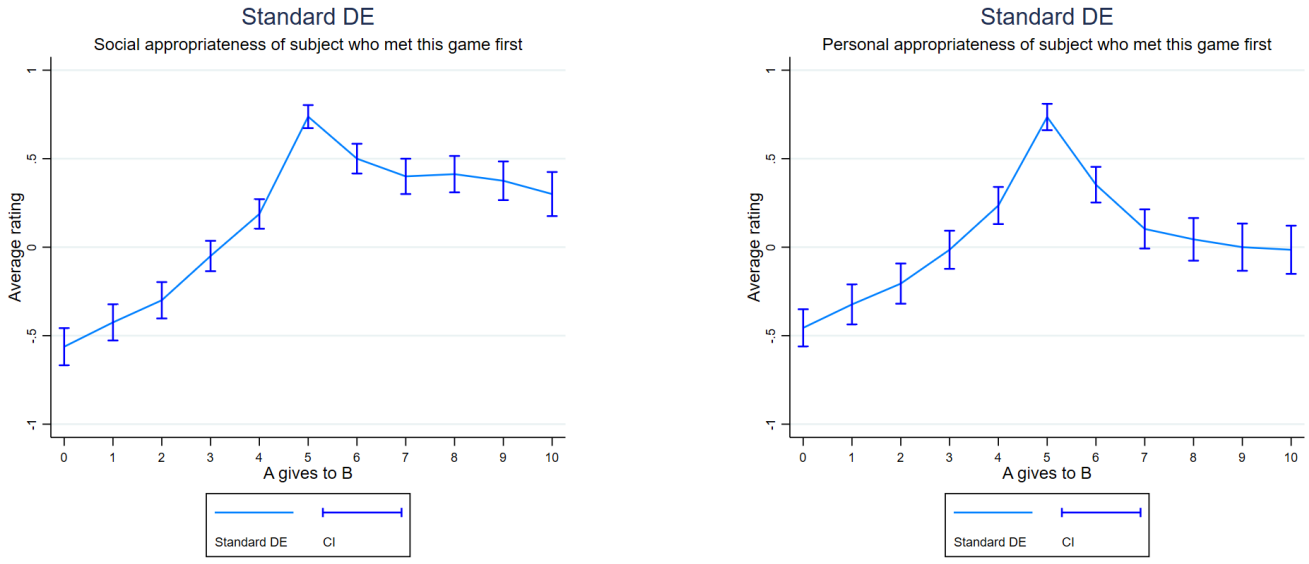


Figure 13: Average ratings in the DE with exit option for subject who met this game as their first

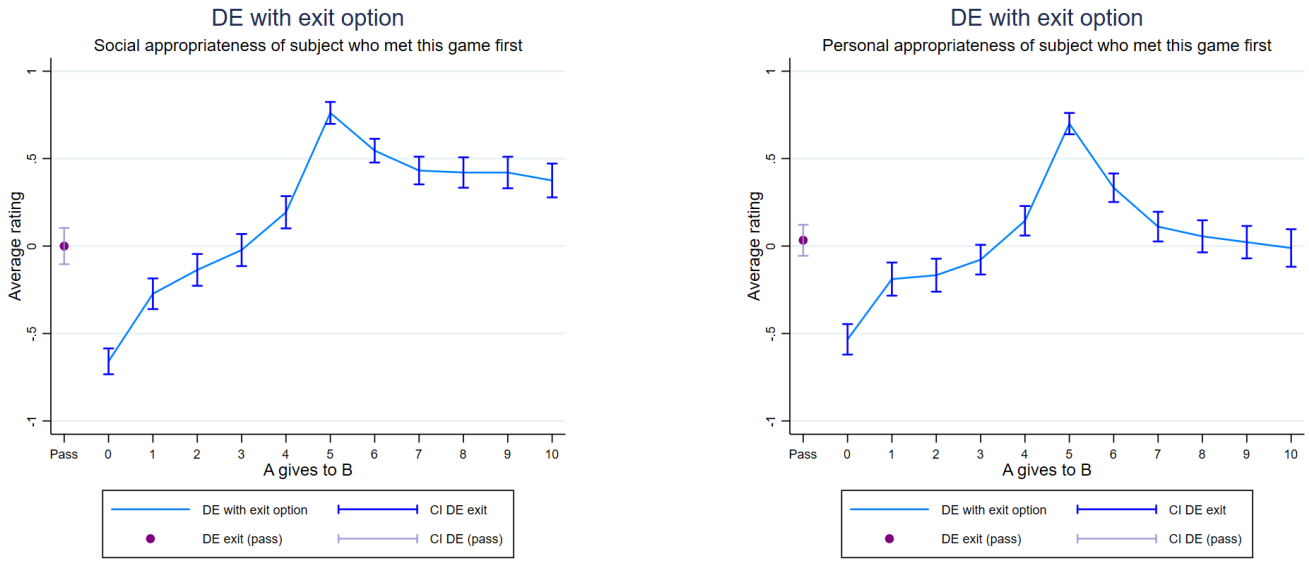


Figure 14: Average ratings in the surplus DE take 2 for subject who met this game as their first

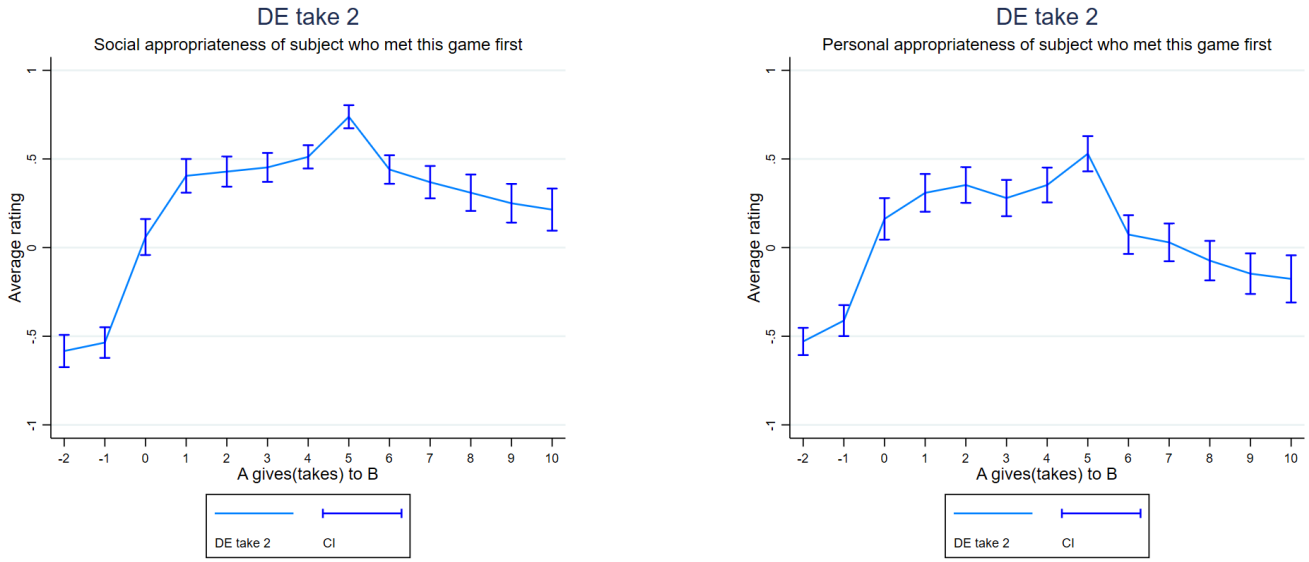


Figure 15: Average ratings in the surplus DE take 10 for subject who met this game as their first

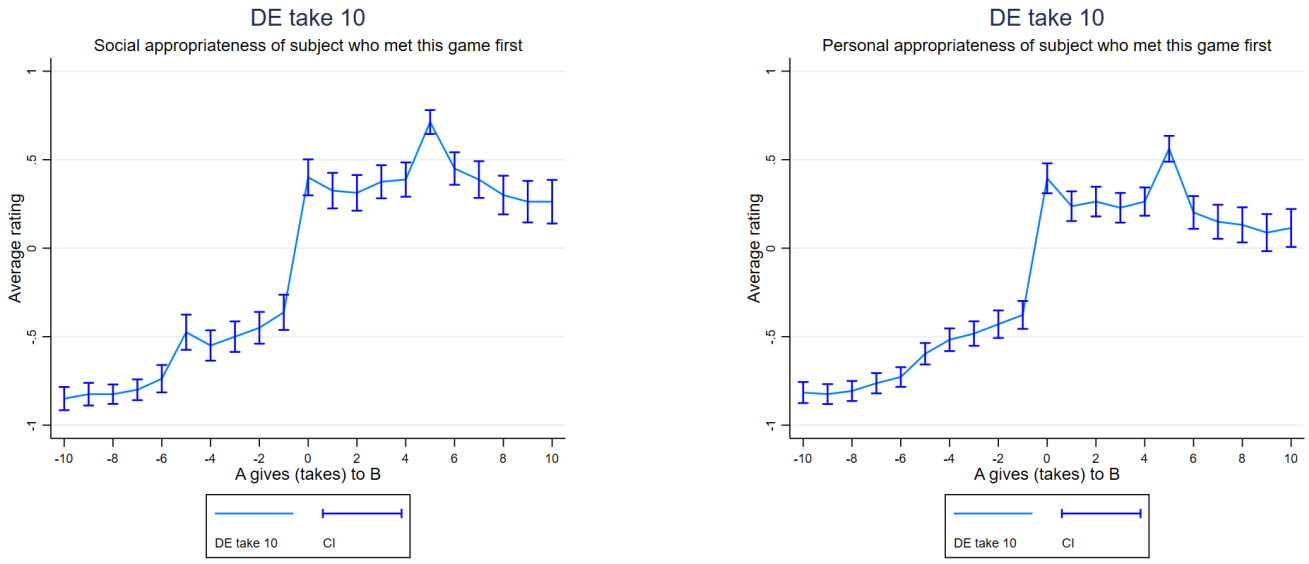


Figure 16: Average ratings in the surplus DE (production) for subject who met this game as their first

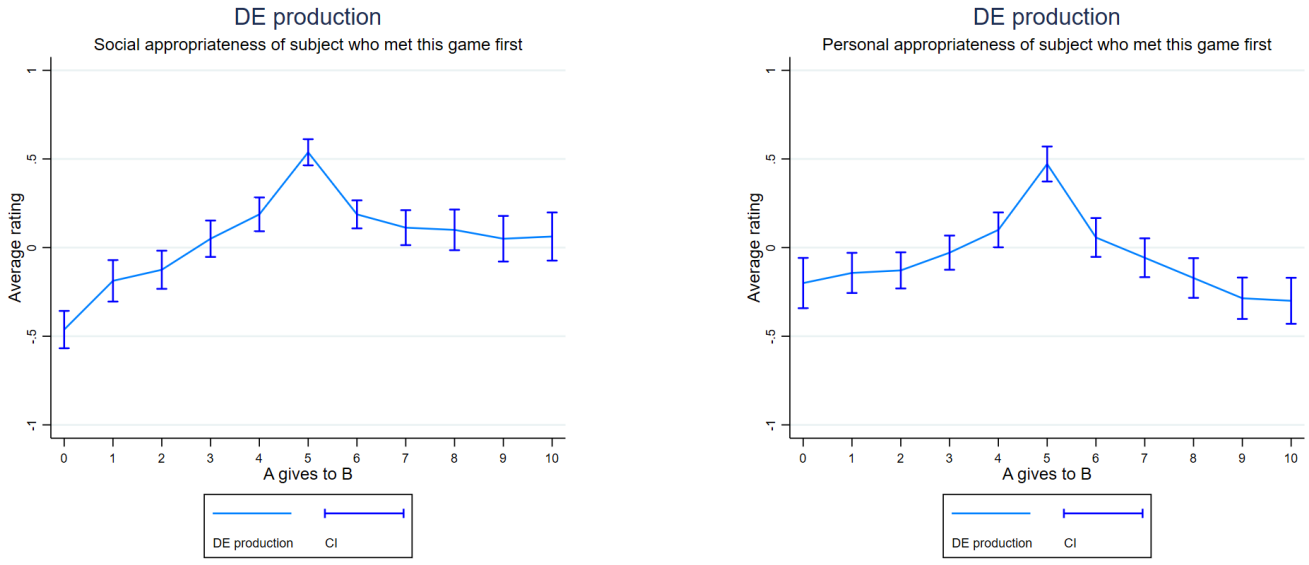
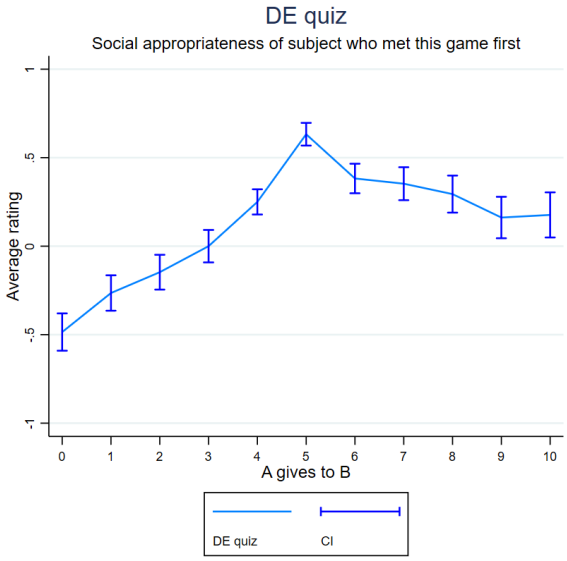
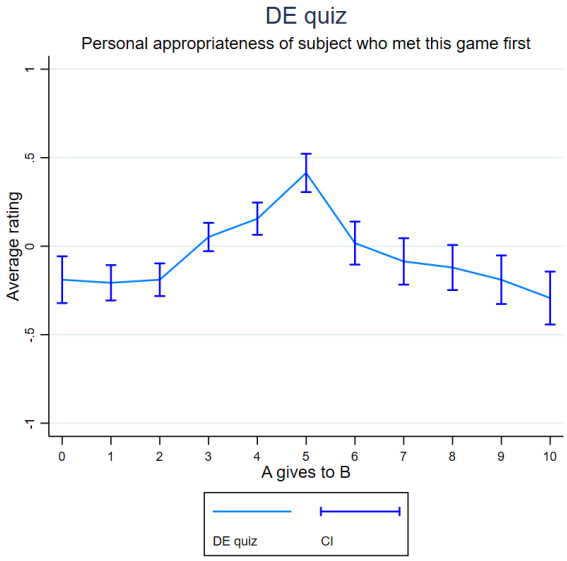



Figure 17: Average ratings in the surplus DE (quiz) for subject who met this game as their first