Supplementary Material for
Fairness and Risk in Ultimatum Bargaining

By KYLE HYNDMAN AND MATTHEW J. WALKER

This appendix contains further information on the experimental design as well as additional tables and more details about the econometric procedure.

S1. Additional Details on Experimental Design

Attrition is a significant challenge for interactive experiments conducted online (see Arechar et al., 2018, for a discussion). To address this problem, we used on-screen timers on each decision page and informed subjects that failure to decide in time would result in their removal from the experiment without payment. We kept the task instructions as concise as possible and not role-specific. Any subject could opt to leave the study (and claim a notional bonus of $0.25) if asked to wait more than five minutes for a match to arrive. Since HITs on AMT move down the advertisement list over time, entry rates decline. To mitigate this and so reduce wait times, we only invited subjects to join the experiment within 25 minutes of posting the HIT. Recruitment rates were high, and subjects usually entered within 10 minutes of posting.[1] On average, the experiment took 15 minutes to complete.

To improve the likelihood that participants completed the tasks carefully, we set the required number of previous approved HITs at greater than 100, and the approval rating on these HITs at greater than 95% (see Peer et al., 2014, on the benefits of recruiting workers with high approval rating for data quality). We also provided decision support tables that summarized the payoffs (and for the Lottery UG also the likelihoods) associated with each action (screenshots are presented in

[1] Full details of the recruitment procedure are contained in the pre-registration document at the AEA RCT registry.

a later section). To prevent retakes, we granted a custom qualification to each subject on entering the experiment (regardless of completion). To reduce the possibility of variable cultural norms influencing bargaining behaviour, we also restricted subject location to the United States.[2]

S2. Inconsistent responder choices in strategy method treatments

TABLE S.I—OBSERVED CHOICE SEQUENCES FOR INCONSISTENT RESPONSES

Standard UG (N = 8)	\$0	\$1	\$2	\$3	\$4	\$5	\$6				
	0	0	1	0	1	1	1				
	0	0	1	0	1	1	1				
	1	1	0	1	1	0	0				
	0	1	0	1	1	1	1				
	0	0	1	0	1	1	1				
	0	0	1	0	1	1	1				
	0	1	1	1	1	0	1				
	0	1	1	1	1	0	1				
Lottery UG (N = 3)	0	10	20	30	40	50	60	70	80	90	100
	1	0	1	1	0	0	1	0	0	1	1
	1	0	0	1	1	0	1	1	1	0	1
	1	0	0	1	1	0	1	1	1	0	1

Notes: Table columns present the acceptance decision (coded as 1 for accept and 0 for reject) for all possible level 1 offers in each task. N denotes the number of observations.

S3. Structural estimation of responder strategy sequences

In this section of the appendix, we structurally estimate the posterior probabilities that a responder strategy sequence in the Risk-SM treatment falls into either the ex ante or ex post fairness type. We assume that the utility of responder i from payoff y to him/herself and payoff x to his/her matched proposer is given by:

$$u_i = y - \alpha_i \cdot \max\{x - y, 0\}, \tag{S1}$$

[2] In the post-experiment questionnaire, 90% of subjects reported (North) American nationality.

which corresponds to the linear inequity averse preferences of Fehr and Schmidt (1999).[3] When there is uncertainty about the final payoff distribution, implied by the lottery $F(x,y)$, evaluation of utility is either by expected utility maximization (ex post) or as the utility of expected outcomes (ex ante, see main text for details).

To construct the posteriors, we first calculate the degree of inequality aversion implied by the Lottery UG, separately based on the ex post $(\alpha_{i,ep}^{L})$ and ex ante $(\alpha_{i,ea}^{L})$ assumptions, respectively. We then estimate a mixture model on responder strategy sequences from the Standard UG.

To permit decision errors, we add idiosyncratic and independent error terms ϵ_{ij} and assume that the difference between any two error terms follows a logistic distribution, with noise-to-signal ratio parameter λ.[4] This parameter drives the sub-optimality of responder acceptance decisions. Thus, the responder accepts any offer j in the Standard UG when $u_{ij} + \lambda\epsilon_{ij} > 0$.

Given the logistic error structure, the probability that responder i accepts offer j conditional on α_i^L is:

$$P_{ij}(\alpha_i^L) = \frac{1}{1+\exp(-u_{ij}/\lambda)}. \tag{S2}$$

For each of the 45 responders in the Risk-SM treatment, we observe a sequence of seven level 1 decisions for the Standard UG task, where $d_{ij}=1$ denotes acceptance and $d_{ij}=0$ rejection by responder i of offer j. For 39 of these 45 responders, we also observe a sequence of nine level 2 decisions. One responder reported a threshold of zero for both ultimatum game tasks and so is excluded from

[3] Initially, we also included the advantageous inequality aversion parameter β_i in the structural estimation model. We found no evidence that this parameter was significantly different from zero, however, and so in the interests of parsimony we removed this parameter from the model.

[4] The results of the estimation are robust to using the alternative Fechner error specification.

the analysis. This yields a total of 659 observations and $T_i \in \{7,16\}$ decisions per responder.

The two fairness types are represented in the model by the mixing proportions p_{ep} and p_{ea}. The marginal likelihood of the strategy sequence of responder i is:

(S3)

$$L_i = p_{ep}\prod_{j=1}^{T_i}\left[P_{ij}(\alpha^L_{i,ep})\right]^{d_{ij}}\left[1-P_{ij}(\alpha^L_{i,ep})\right]^{1-d_{ij}} + p_{ea}\prod_{j=1}^{T_i}\left[P_{ij}(\alpha^L_{i,ea})\right]^{d_{ij}}\left[1-P_{ij}(\alpha^L_{i,ea})\right]^{1-d_{ij}}.$$

Based on equation (S3), the sample log-likelihood is given by:

(S4)
$$Log\,L = \sum_{i=1}^{44}\log(L_i),$$

on maximization of which we obtain likelihood estimates for the rationality parameter λ and one of the two mixing proportions (from which the other is calculated using the delta method).

The estimation was conducted in Stata 16, using the d0 estimator to account for the panel structure of the data. A starting value for the mixing proportion was obtained from the classification of types in the results section of the main text. The results of this estimation are presented in Table S.II. The proportion of ex post fairness types is estimated to be 51.9%, which lies in the middle of the 40 to 60% range estimated using the non-parametric and least squares methods.

TABLE S.II—MAXIMUM LIKELIHOOD ESTIMATES FROM MIXTURE MODEL

Rationality Parameter		
λ	1.963***	(0.172)
Mixing proportions		
p_{ep}	0.519***	(0.0946)
p_{ea}	0.481***	(0.0946)
Number of Observations	659	
Log-likelihood	-284.749	

Note: The table reports maximum likelihood estimates from responder strategy sequences in the Risk-SM treatment. Asymptotic standard errors in parentheses. *** denotes significance at the 1% level.

Posterior probabilities for the two types can be calculated as:

$$P(i=ep \mid d_{i1},...,d_{iT_i}) = \frac{p_{ep}\prod_{j=1}^{T_i}\left[P_{ij}(\alpha_{i,ep}^{L})\right]^{d_{ij}}\left[1-P_{ij}(\alpha_{i,ep}^{L})\right]^{1-d_{ij}}}{L_i};$$

(S5)

$$P(i=ea \mid d_{i1},...,d_{iT_i}) = \frac{p_{ea}\prod_{j=1}^{T_i}\left[P_{ij}(\alpha_{i,ea}^{L})\right]^{d_{ij}}\left[1-P_{ij}(\alpha_{i,ea}^{L})\right]^{1-d_{ij}}}{L_i}.$$

In Figure S1, we present a histogram of posterior probabilities of being classified as an ex post fairness type, split according to responders' non-parametric strategy comparisons between the Standard and Lottery UG. The posterior estimates are consistent with the more crude classifications in the main text. Responders who tend to report a higher acceptance threshold in the Lottery UG, tend to have a larger posterior probability of being classified as an ex post type, with a discrepancy of roughly 20% of responders who are mis-classified as an ex ante type. Those responders who report an equal threshold in the two UG tasks are unambiguously classified as an ex ante type. Unlike with the crude classifications, we can now additionally infer that those responders who report a lower threshold in the Lottery UG are not well-classified by either fairness model, as the theory predicts.

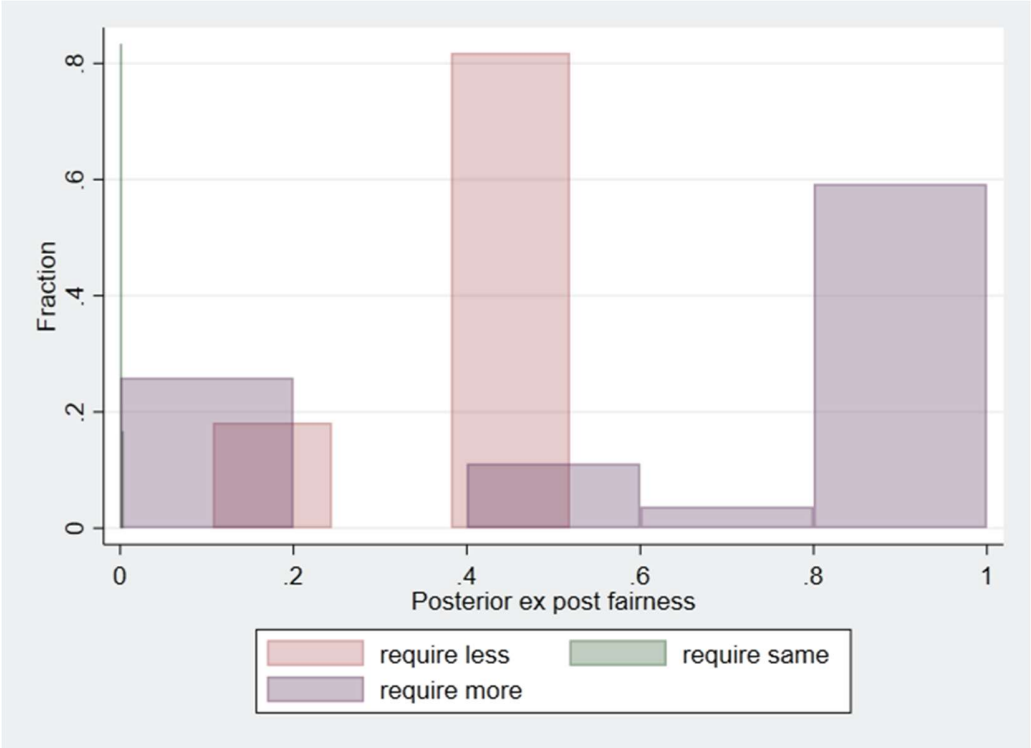


FIGURE S1. POSTERIOR PROBABILITIES OF FAIRNESS TYPES BY RESPONDER STRATEGY COMPARISON

Notes: The posterior probabilities are estimated from the mixture model based on the equations in (S7). The responder strategy comparisons are taken from the raw classification of types presented in Section III.A of the main text. Require more (less) [same] indicates the responder reported a higher (lower) [equal] acceptance threshold in the Lottery UG as compared to the Standard UG. Data is from responder strategy sequences in the Risk-SM treatment.

S4. AMT recruitment outcomes, subject pool characteristics and randomization check

We recruited 1,029 subjects on the Amazon Mechanical Turk (AMT) platform in May and June 2020. For transparency, in Table S.IV we report the full breakdown of recruitment outcomes. Figure S2 displays the description of the HIT on AMT as seen by subjects.

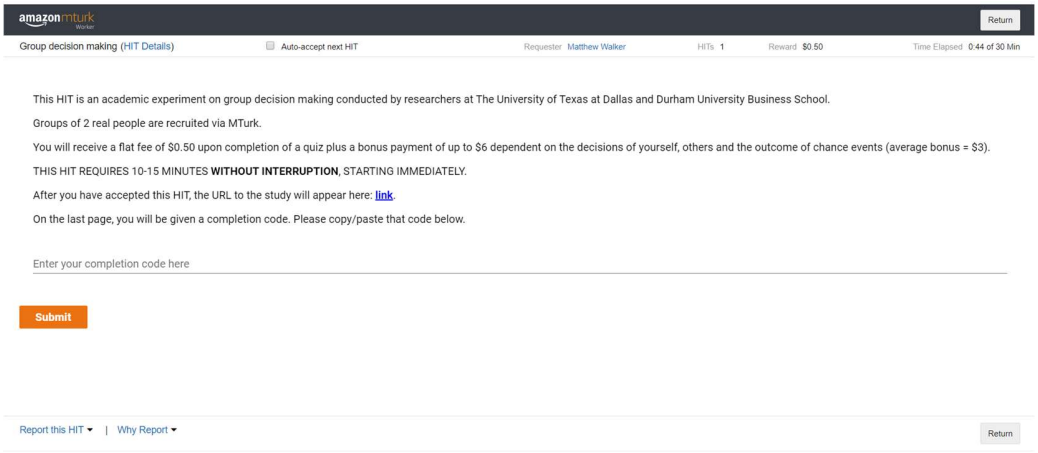


FIGURE S2. AMT RECRUITMENT SCREEN.

In total, 503 recruits successfully completed the experiment. As noted in the main text, 10 of these subjects displayed inconsistent behaviour in either the Risk-SM or Intent-SM treatment and so were excluded from the analysis in accordance with the pre-registration protocol. This left us with a final sample size of 493 subjects across the four treatments.

Non-completes can be classified into one of six categories: failed the Standard or Lottery UG comprehension quiz; opted to leave the study after waiting more than five minutes for a match to arrive (and claim a notional bonus of $0.25 in addition to the participation fee); timed out in one of the tasks; or dropped out for some unobserved reason.

TABLE S.IV —MTURK RECRUITMENT OUTCOMES

	Completes	Non-completes				
Treatment		Fail quiz Standard UG	Fail quiz Lottery UG	Leave after >5 mins wait	Timeout	Dropout
Risk-DR	112	22	12	11	2	32
Risk-SM	94	29	14	11	5	37
Intent-DR	156	33	22	24	5	62
Intent-SM	141	57	45	15	5	83

Note: To obtain the final sample size of 493, exclude 5 completes in the Risk-SM treatment and 5 completes in the Intent-SM treatment who displayed inconsistent behavior, pre-defined as switching from accept to reject, or reject to accept, more than twice in an ordered list of proposals.

Below we summarize characteristics of our sample, based on information elicited from the post-experiment questionnaire. The summary statistics are calculated based on the full sample of 493 consistent completes.

Age: Interval variable. Years.

Mean 36.52, Median 34, Standard deviation 10.62, Minimum 20, Maximum 74

Gender: Categorical variable.

Male 60.21%; Female 38.76%; Other 0.21%; Prefer not to Say 0.82%.

Nationality: Categorical variable:

Central and Eastern Asia 1.63%; Central and Western Africa 0.20%; Central, South America and the Caribbean 1.02%; Europe (excl. UK) 4.07%; Middle East and North Africa 0.20%; North America 91.06%; South-East Asia 1.22%; Southern Asia 0.41%; UK 0.20%.

Employment Status: Categorical variable:

Full-time employed 68.43%; Part-time employed 7.13%; Self-employed 9.16%; Unemployed 9.78%; Retired 1.83%; Student 2.65%; Other 1.02%.

COVID-19 Impact: Categorical variable:

Laid off 5.48%; Pay cut and/or reduction in hours 16.63%; Work from home 39.96%; No impact 35.90%; Positive impact 2.03%.

Risk Index: Based on two questions from Dohmen et al. (2011). Likert scale from 0 “Completely unwilling to take risks” to 10 “Completely willing to take risks”.

1) Are you generally a person who is fully willing to take risks or do you try to avoid taking risks?

Mean 4.51, Median 4, Standard deviation 2.69, Minimum 0, Maximum 10

2) How would you rate your willingness to take risks in financial matters?

Mean 4.26, Median 4, Standard deviation 2.69, Minimum 0, Maximum 10

Competitiveness Index: Based on question from Buser et al. (2020). Likert scale from 0 “Not competitive at all” to 10 “Very competitive”. How competitive do you consider yourself to be?

Mean 5.50, Median 6, Standard deviation 2.70, Minimum 0, Maximum 10

We also conducted a randomization check on age and a dummy for being female. The results are presented in Table S.V. We observe no significant differences between treatments (for age, $p > 0.20$ for all 6 comparisons based on pairwise t-tests with p-values adjusted using Bonferroni correction; for female, $p = 0.756$ based on Pearson's Chi-squared test).

TABLE S.V —TREATMENT RANDOMIZATION CHECK

Treatment	Age	Female
Risk-DR	35.05 (9.86)	0.39 (0.49)
Risk-SM	36.33 (10.34)	0.37 (0.49)
Intent-DR	37.85 (11.18)	0.35 (0.48)
Intent-SM	36.34 (10.70)	0.41 (0.49)

Note: Mean (SD) values in table.

S5. Experimental instructions and decision screens

The experimental instructions and comprehension quiz for each treatment and ultimatum game task are presented in Figures S3 to S8. Each box corresponds to a separate screen. Note the instructions below are for sessions in which the Standard UG was presented first (minor changes for the alternate order are self-evident).

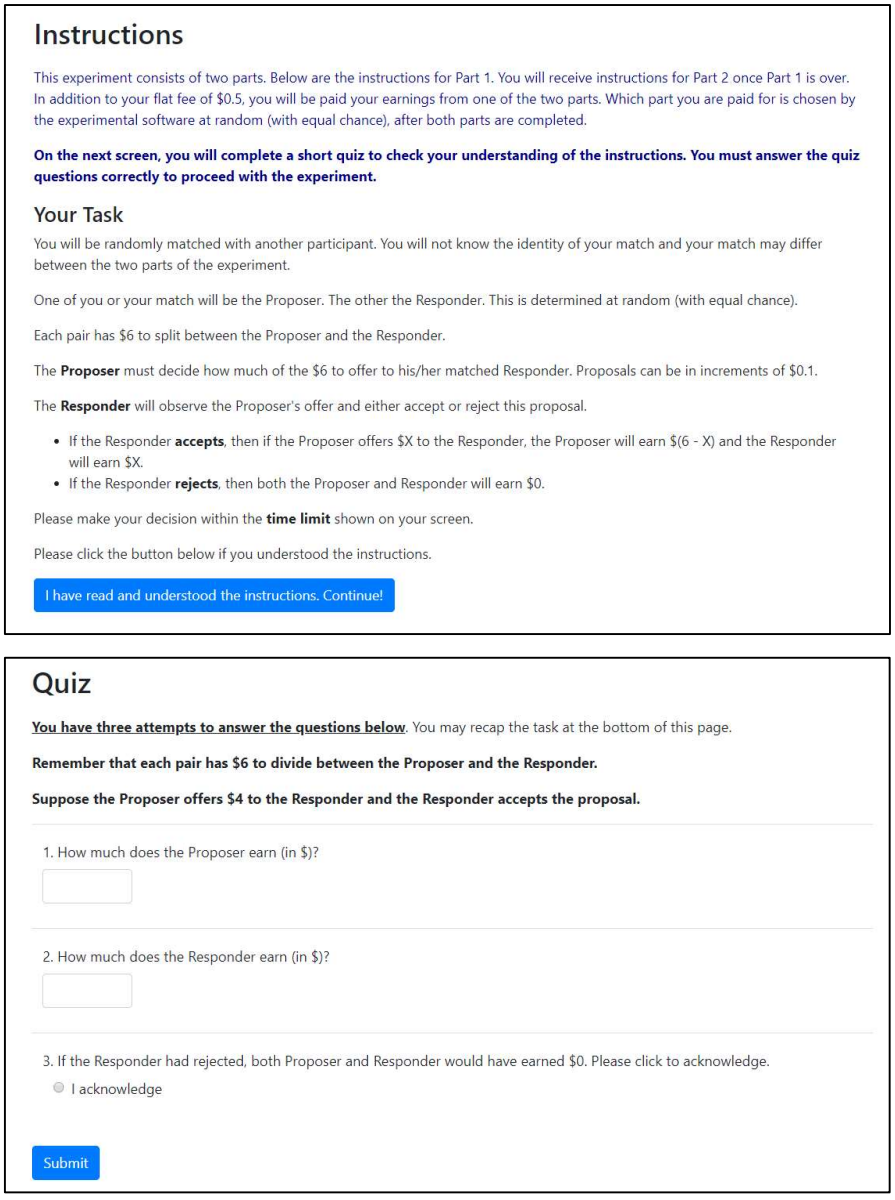


FIGURE S3. EXPERIMENTAL INSTRUCTIONS AND COMPREHENSION QUIZ —RISK-DR AND INTENT-DR (STANDARD UG)

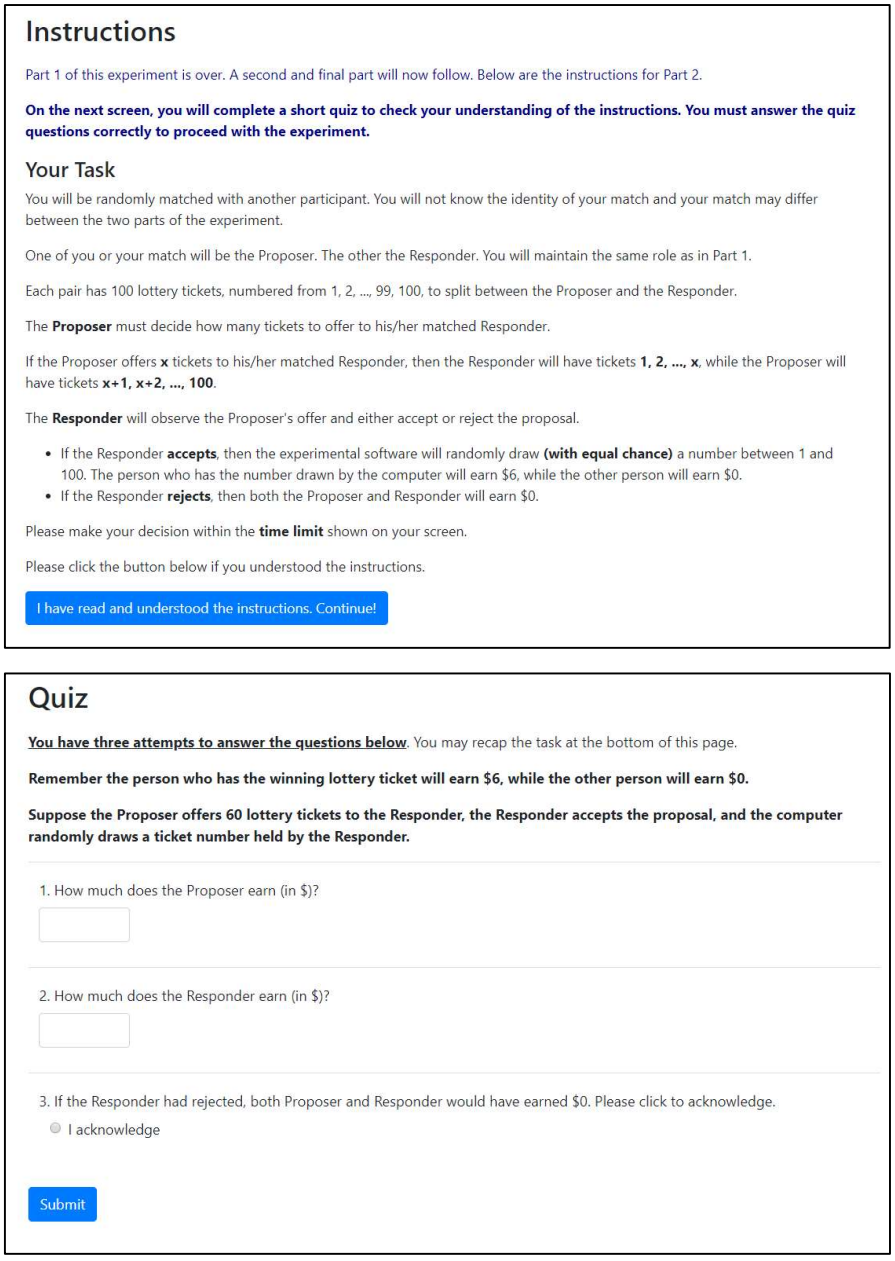


FIGURE S4. EXPERIMENTAL INSTRUCTIONS AND COMPREHENSION QUIZ —RISK-DR (LOTTERY UG).

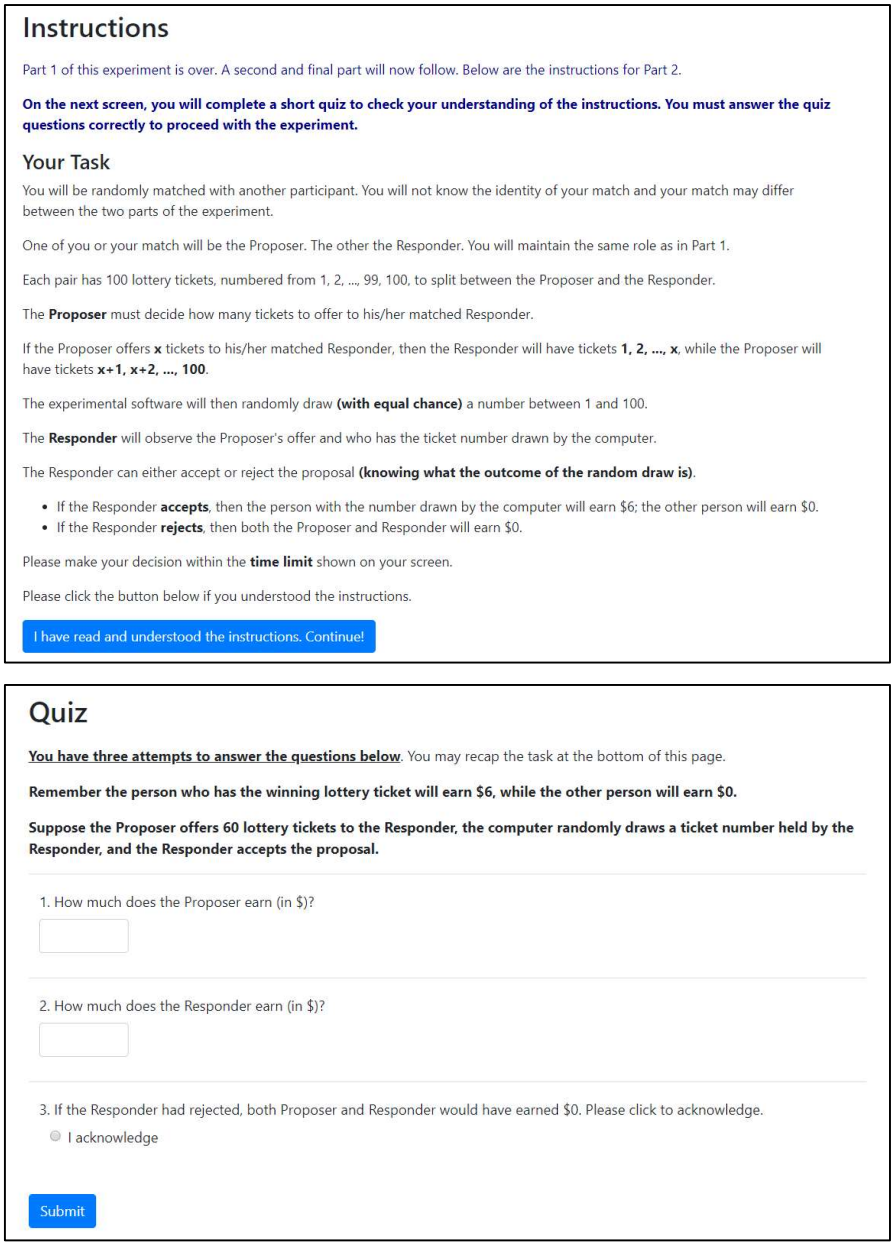


FIGURE S5. EXPERIMENTAL INSTRUCTIONS AND COMPREHENSION QUIZ —INTENT-DR (LOTTERY UG).

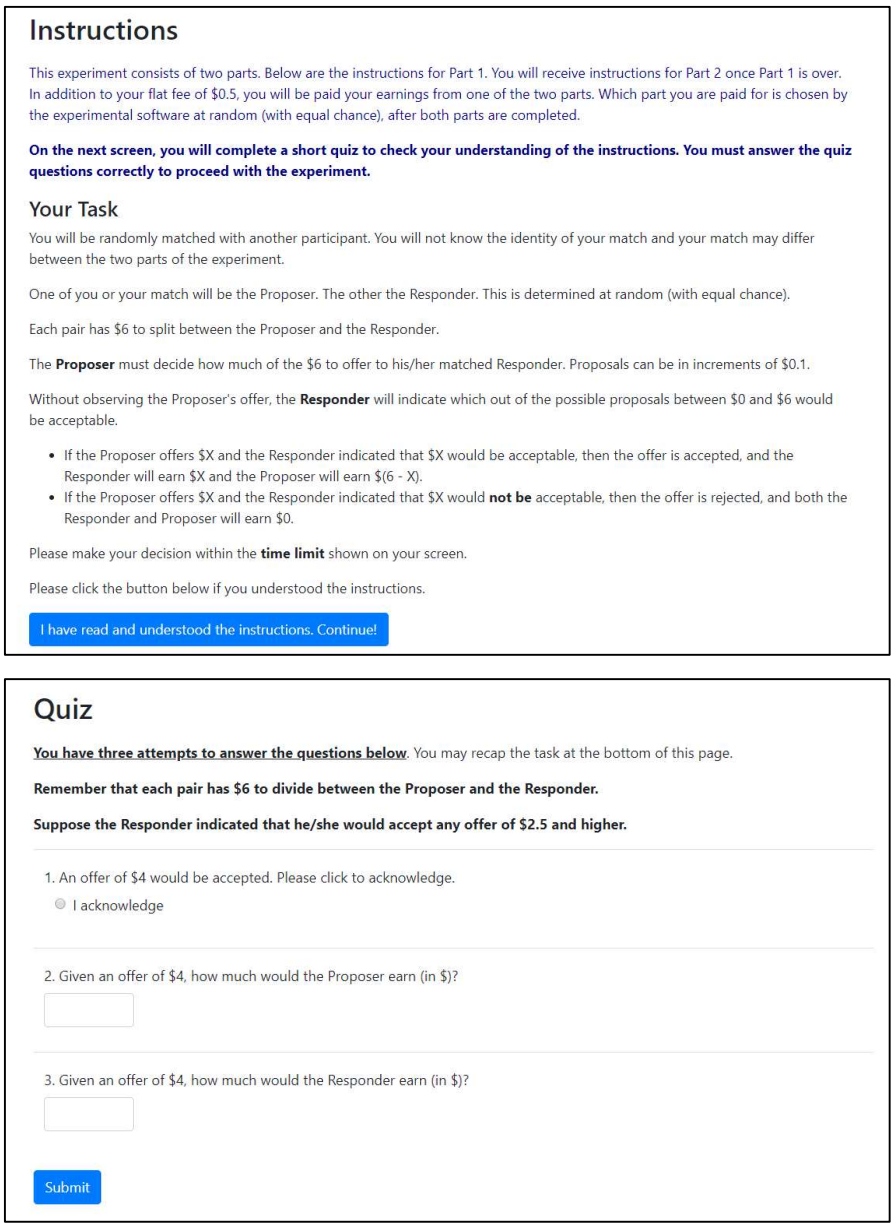


FIGURE S6. EXPERIMENTAL INSTRUCTIONS AND COMPREHENSION QUIZ —RISK-SM AND INTENT-SM (STANDARD UG).

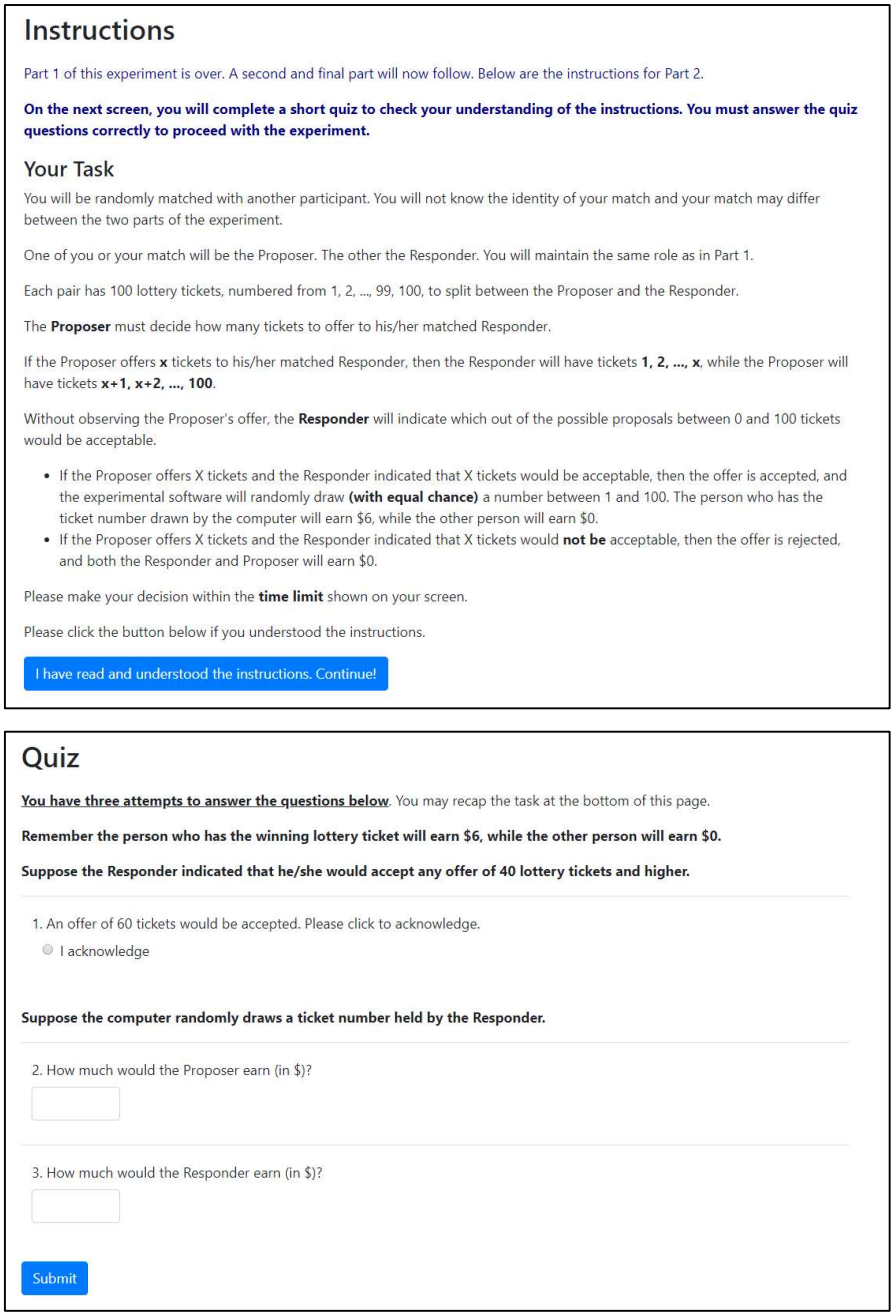


FIGURE S7. EXPERIMENTAL INSTRUCTIONS AND COMPREHENSION QUIZ —RISK-SM (LOTTERY UG).

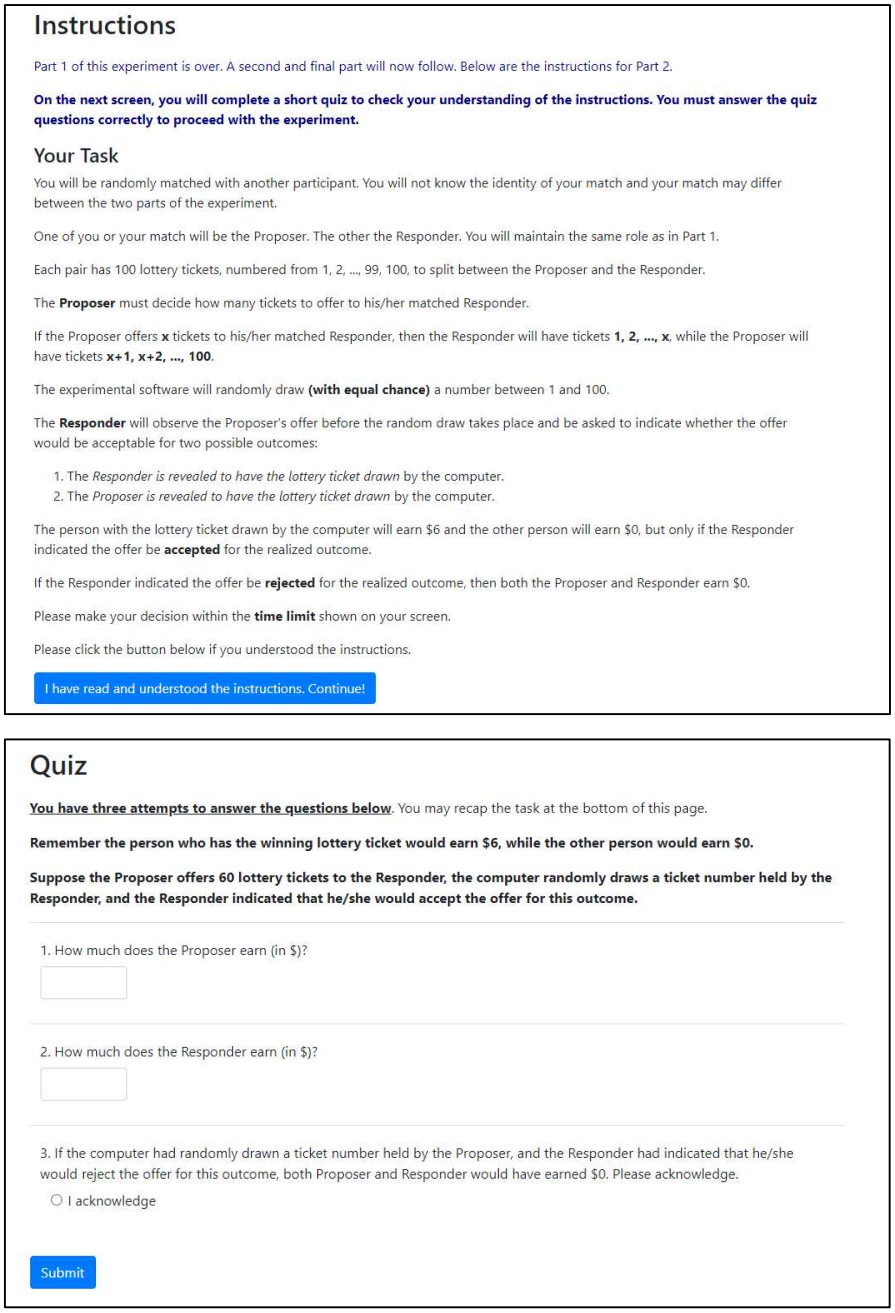


FIGURE S8. EXPERIMENTAL INSTRUCTIONS AND COMPREHENSION QUIZ —INTENT-SM (LOTTERY UG).

Example decision screens for the proposer and responder in each treatment and task are presented in Figures S9 to S14.

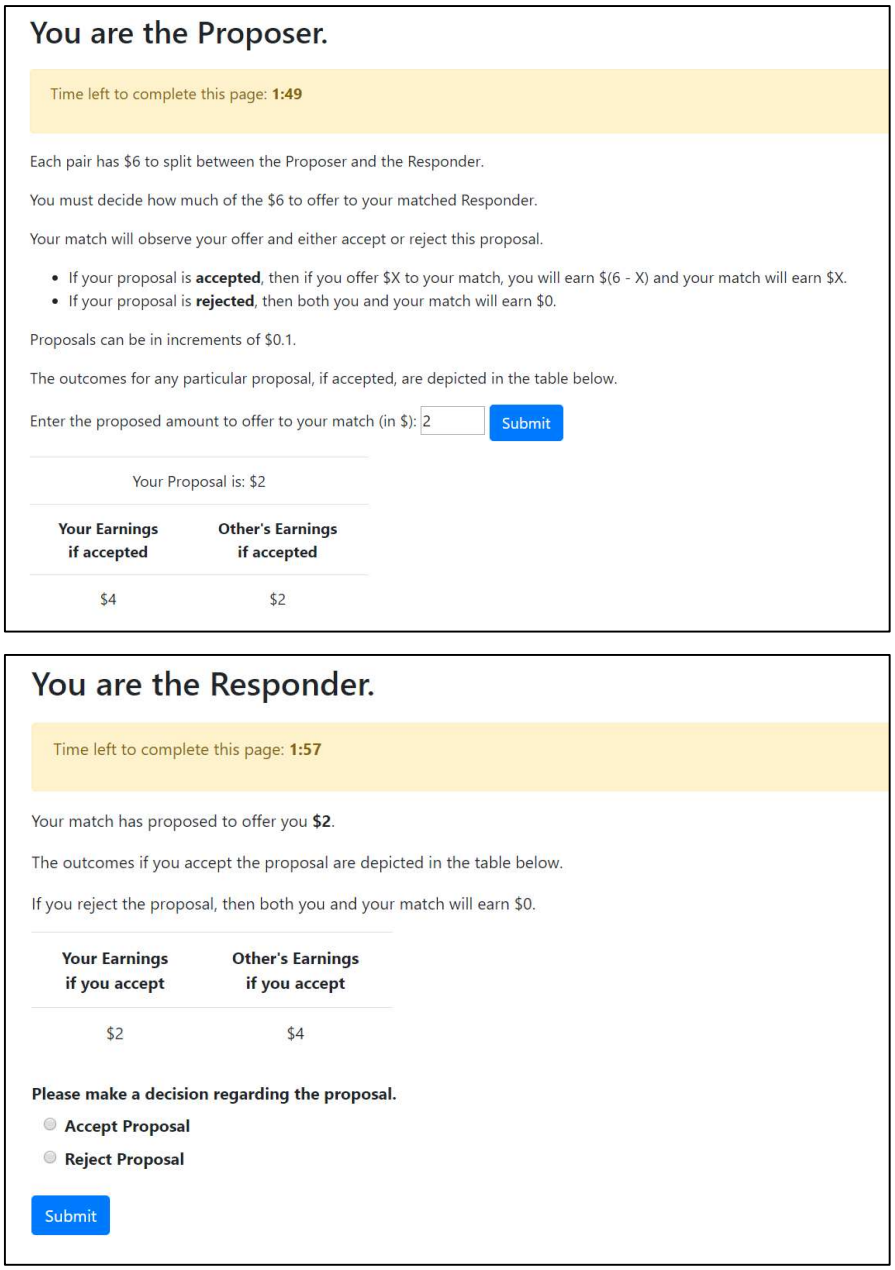
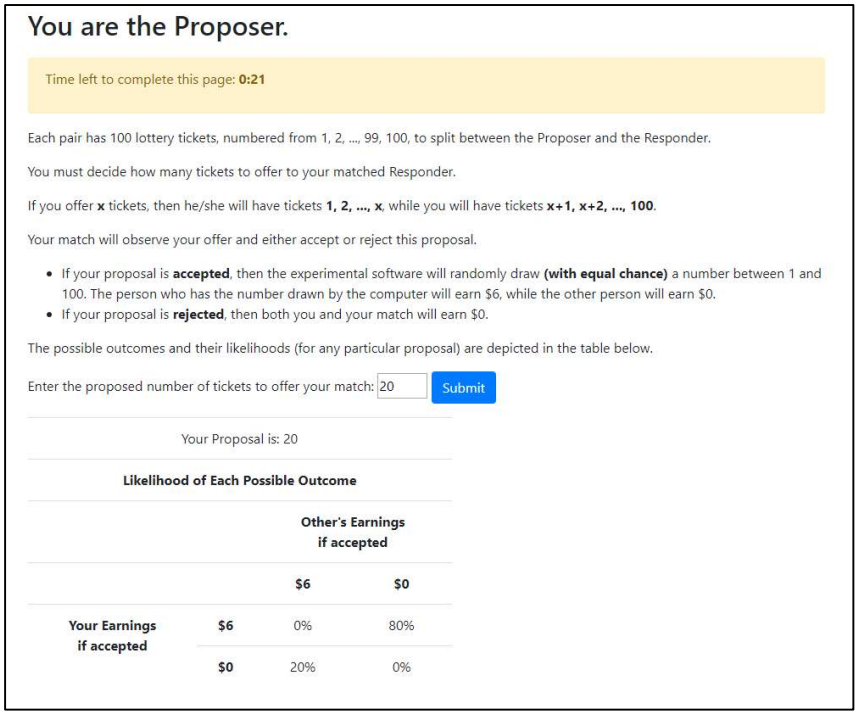


FIGURE S9. DECISION SCREENS —RISK-DR AND INTENT-DR (STANDARD UG).



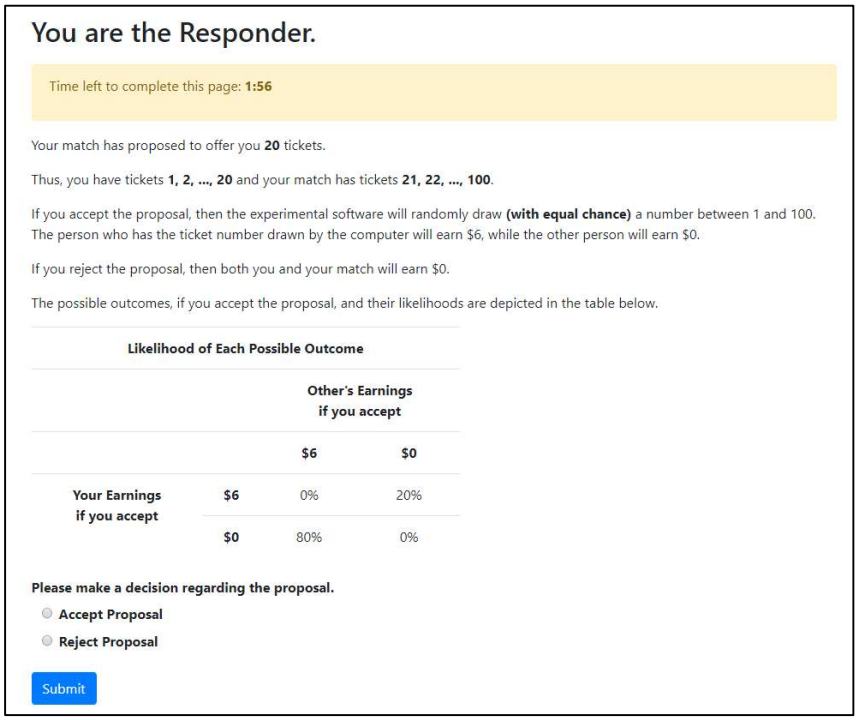


FIGURE S10. DECISION SCREENS —RISK-DR (LOTTERY UG).

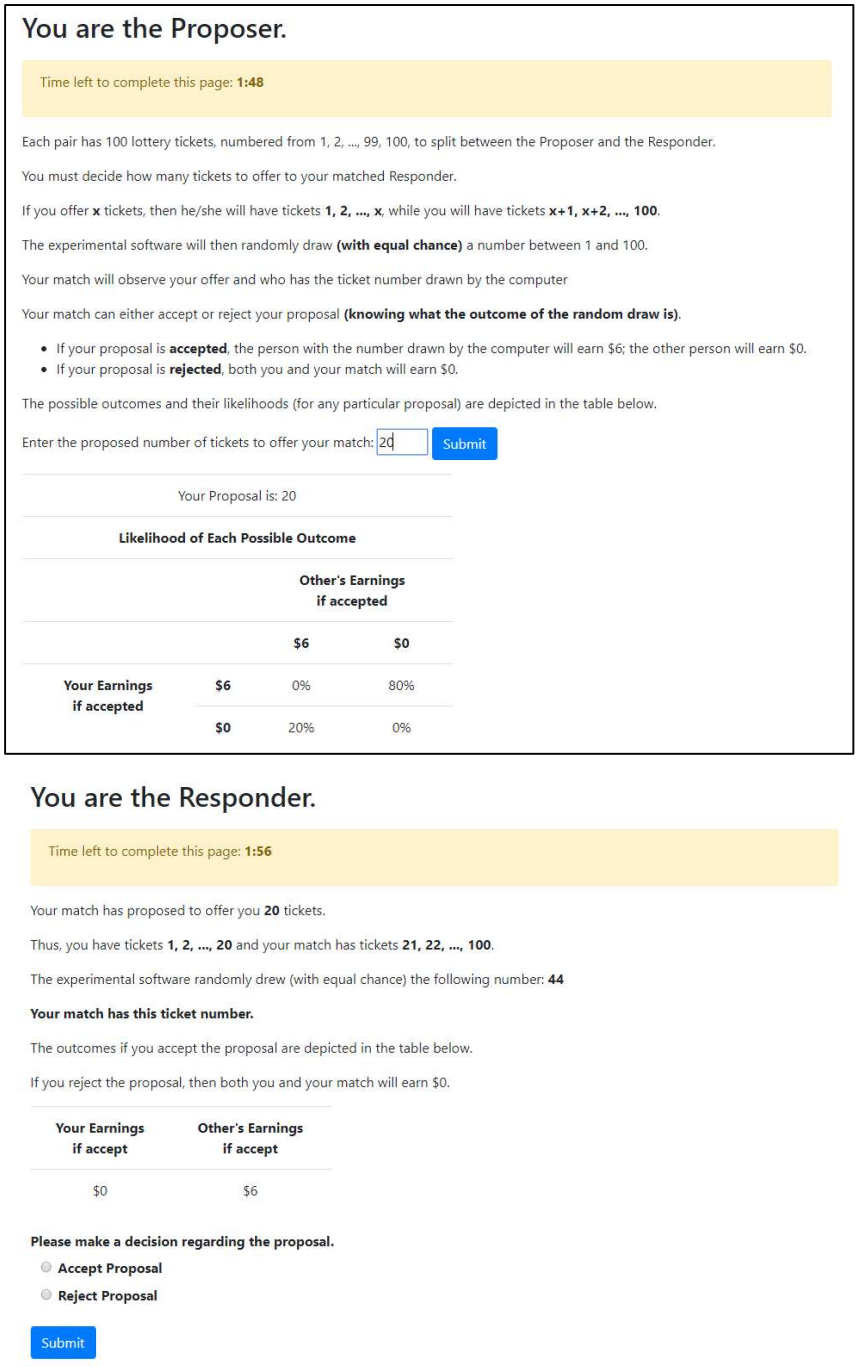


FIGURE S11. DECISION SCREENS —INTENT-DR (LOTTERY UG).

You are the Proposer.

Time left to complete this page: **0:07**

Each pair has $6 to split between the Proposer and the Responder.

You must decide how much of the $6 to offer to your matched Responder.

Without observing your offer, your match will indicate which out of the possible proposals between $0 and $6 is acceptable.

- If you offer $X and your match indicated that $X would be acceptable, then the offer is accepted, and your match will earn $X and you will earn $(6 - X).
- If you offer $X and your match indicated that $X would **not be** acceptable, then the offer is rejected, and both you and your match will earn $0.

Proposals can be in increments of $0.1.

The outcomes for any particular proposal, if accepted, are depicted in the table below.

Enter the proposed amount to offer to your match (in $): 2 Submit

Your Proposal is: $2

Your Earnings if accepted	Other's Earnings if accepted
$4	$2

You are the Responder.

Time left to complete this page: **1:43**

Your match is deciding how much of the $6 to offer to you.

Please indicate which out of the proposals below you would accept or reject.

For each proposal, the table depicts the outcome associated with acceptance, if this is what the Proposer offers.

Rejection of any of these proposals would result in both you and your match earning $0, if this is what the Proposer offers.

		If this is what the Proposer offers,	
Proposal	**Your Decision**	**if Accept, you earn:**	**if Accept, your match earns:**
$0	○ Accept ◉ Reject	$0	$6
$1	○ Accept ◉ Reject	$1	$5
$2	◉ Accept ○ Reject	$2	$4
$3	◉ Accept ○ Reject	$3	$3
$4	◉ Accept ○ Reject	$4	$2
$5	◉ Accept ○ Reject	$5	$1
$6	◉ Accept ○ Reject	$6	$0

Submit

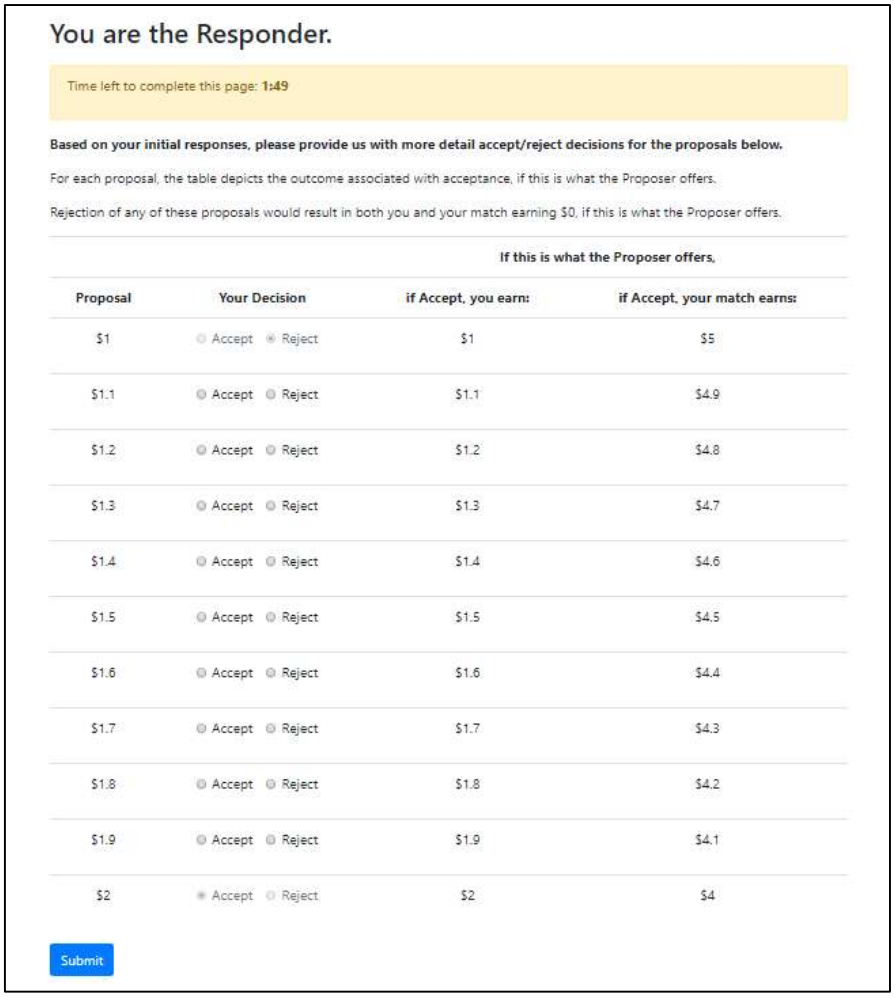


FIGURE S12. DECISION SCREENS —RISK-SM AND INTENT-SM (STANDARD UG).

You are the Proposer.

Time left to complete this page: **1:19**

Each pair has 100 lottery tickets, numbered from 1, 2, ..., 99, 100, to split between the Proposer and the Responder.

You must decide how many tickets to offer to your matched Responder.

If you offer **x** tickets, then he/she will have tickets **1, 2, ..., x**, while you will have tickets **x+1, x+2, ..., 100**.

Without observing your offer, your match will indicate which out of the possible proposals between 0 and 100 tickets would be acceptable.

- If you offer X tickets and your match indicated that X tickets would be acceptable, then the offer is accepted, and the experimental software will randomly draw **(with equal chance)** a number between 1 and 100. The person who has the ticket number drawn by the computer will earn $6, while the other person will earn $0.
- If you offer X tickets and your match indicated that X tickets would **not be** acceptable, then the offer is rejected, and both you and your match will earn $0.

The possible outcomes and their likelihoods (for any particular proposal) are depicted in the table below.

Enter the proposed number of tickets to offer your match: 20 Submit

Your Proposal is: 20

Likelihood of Each Possible Outcome

		Other's Earnings if accepted	
		$6	**$0**
Your Earnings if accepted	**$6**	0%	80%
	$0	20%	0%

You are the Responder.

Time left to complete this page: **0:48**

Your match is deciding how many lottery tickets to offer you.

Please indicate which out of the proposals below you would accept or reject.

The possible outcomes and their likelihoods, if you accept the proposal and this is what the Proposer offers, are depicted below.

Rejection of any of these proposals would result in both you and your match earning $0, if this is what the Proposer offers.

		If this is what the Proposer offers,	
Proposal	**Your Decision**	**if Accept, you earn:**	**if Accept, your match earns:**
0	○ Accept ◉ Reject	$0 for sure	$6 for sure
10	○ Accept ◉ Reject	$0 with 90% chance $6 with 10% chance	$6 with 90% chance $0 with 10% chance
20	○ Accept ◉ Reject	$0 with 80% chance $6 with 20% chance	$6 with 80% chance $0 with 20% chance
30	○ Accept ◉ Reject	$0 with 70% chance $6 with 30% chance	$6 with 70% chance $0 with 30% chance
40	◉ Accept ○ Reject	$0 with 60% chance $6 with 40% chance	$6 with 60% chance $0 with 40% chance
50	◉ Accept ○ Reject	$0 with 50% chance $6 with 50% chance	$6 with 50% chance $0 with 50% chance
60	◉ Accept ○ Reject	$0 with 40% chance $6 with 60% chance	$6 with 40% chance $0 with 60% chance
70	◉ Accept ○ Reject	$0 with 30% chance $6 with 70% chance	$6 with 30% chance $0 with 70% chance
80	◉ Accept ○ Reject	$0 with 20% chance $6 with 80% chance	$6 with 20% chance $0 with 80% chance
90	◉ Accept ○ Reject	$0 with 10% chance $6 with 90% chance	$6 with 10% chance $0 with 90% chance
100	◉ Accept ○ Reject	$6 for sure	$0 for sure

Submit

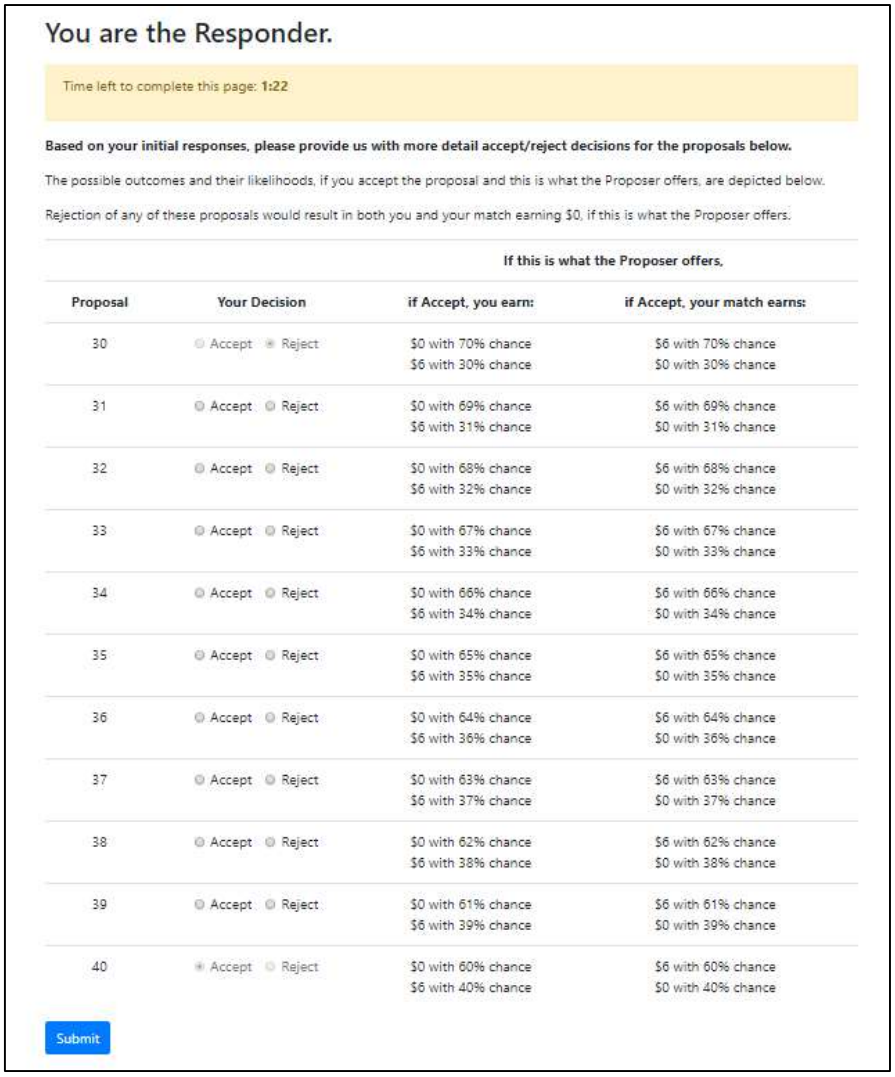


FIGURE S13. DECISION SCREENS —RISK-SM (LOTTERY UG).

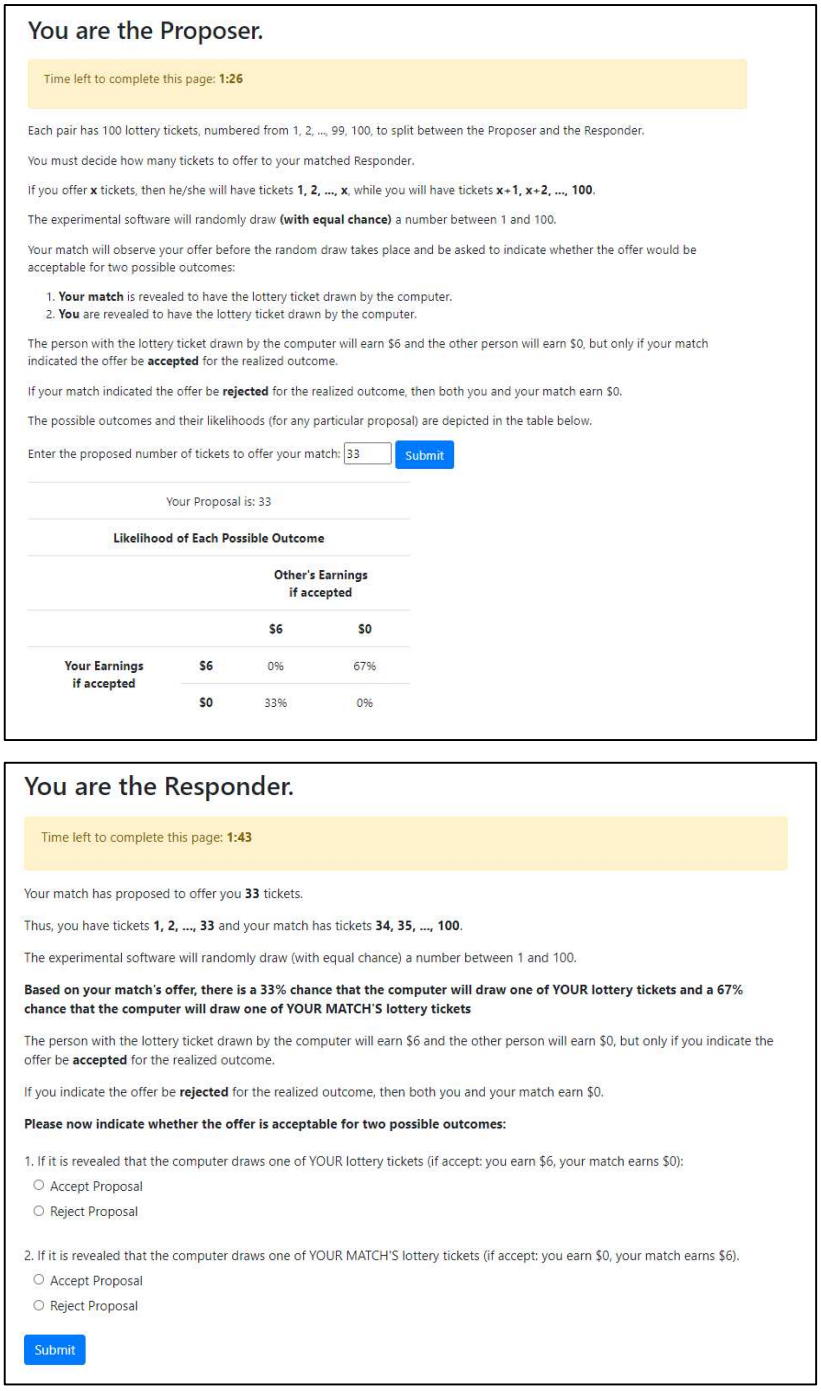


FIGURE S14. DECISION SCREENS —INTENT-SM (LOTTERY UG).

REFERENCES

Arechar, Antonio A., Simon Gächter, and Lucas Molleman. 2018. "Conducting interactive experiments online." *Experimental Economics* 21 (1):99-131. doi: 10.1007/s10683-017-9527-2.

Fehr, Ernst, and Klaus M. Schmidt. 1999. "A Theory of Fairness, Competition, and Cooperation*." *Quarterly Journal of Economics* 114 (3):817-868. doi: 10.1162/003355399556151.

Peer, Eyal, Joachim Vosgerau, and Alessandro Acquisti. 2014. "Reputation as a sufficient condition for data quality on Amazon Mechanical Turk." *Behavior Research Methods* 46 (4):1023-1031. doi: 10.3758/s13428-013-0434-y.