

Data and Code for: Belief Elicitation and Behavioral Incentive Compatibility

Overview

This repository contains data and code for the paper “Belief Elicitation and Behavioral Incentive Compatibility.” All data is collected from laboratory experiments conducted by the authors. Code is contained in one master file `/do/StataCode.do` that uses STATA to replicate the core tables and inferences. An additional file `/math/MathematicaCode.m` generates some figures and tables for the paper. The replicator should be able to run both of these files in 10 minutes on a standard desktop computer.

Data Availability and Provenance Statements

- ☒ This paper does not involve analysis of external data (i.e., no data are used or the only data are generated by the authors via simulation in their code).

Statement about Rights

- ☒ I certify that the author(s) of the manuscript have legitimate access to and permission to use the data used in this manuscript.

Summary of Availability

- ☒ All data **are** publicly available.

Details on each Data Source

The data here represents three core experiments:

1. Objective Prior and Posterior Beliefs - The data from experiments examining the elicitation of beliefs in a standard Bayesian updating task (`data/data-bsr-qsr.dta`)
2. Subjective Beliefs - The data from a replication of Niederle & Vesterlund (2007), using the BSR (`data/data-nv.dta`)
3. Incentives only - data from a module at the end of another experiment collecting choices from a choice menu (`data/data-incentives-only.dta`)

Below we outline the contents of each data file in more detail

Computational requirements

The core replication can be conducted on a modern laptop or desktop computer in a few minutes.

Software Requirements

- STATA (code was last run on STATA/SE 17.0)
 - `estout` package
- Mathematica (code was last run on Mathematica 12.3)

Instructions to Replicators

All scripts are written assuming the base directory is the base directory of this repository.

The provided code reproduces:

- ☒ All numbers provided in text in the paper
- ☒ All tables and figures in the paper

Respository Folders:

- `data/` : Raw data in CSV and STATA (dta) formats
- `do/` : STATA scripts to create Figures and Tables
- `figures/` : Folder containing Figure output by file type
- `math/` : Mathematica notebooks to generate Figures

- `simulations/` : Simulation output data files and STATA scripts

Analysis Files

Stata

All STATA code was run using a standard installation of STATA/SE 17.0 (where some of the output use the `estout` module, Jann 2005, 2007, which can be installed via the command `ssc install estout`, we used the `st0085_2` version). The STATA code is contained in the `do/` folder, which contains the script `StataCode.do` (and an annotated version with both the input and output in HTML). This code replicates:

- All analysis in the main text, footnotes, and the online appendix.
- All tables in the paper and online appendix.

Note that some of the figures in the paper are generated with Mathematica (see below). When running `StataCode.do`, the simulation results in Tables B.2-B.5 are retrieved from stored simulations (located in `simulations/simulation-results/`). To rerun the simulations, the `simulations/` folder contains two STATA scripts. To re-run the simulations simply **adjust the location of the root folder in line 10 of each file**.

- `AER-2020-1248_simulations.do` : Runs the simulations for Table B2 (“Simulation of center bias: Beliefs as a LHS Variable”) and Table B3 (“Simulation of center bias: Beliefs as a RHS control”).
- `AER-2020-1248_simulations-reverse.do` : Runs the simulations for Table B4 (“Reverse simulation: Belief on LHS”) and Table B5 (“Reverse simulation: Belief on RHS”).

The `simulations/` folder also contains some auxiliary files for this:

- `sim_center_bias.ado` : Performs one simulation of center bias.
- `sim_center_bias_reverse.ado` : Performs one simulation of attempted reversal of center bias (recovery of unbiased data).
- `store_estimates.ado` : Stores simulation results as STATA estimates in the folder `simulations/simulation-results/`.

Mathematica

Code was run using a standard installation of Mathematica 12.3, where this analysis is in the `math/` folder. This folder contains a Mathematica script `MathematicaCode.m` (and an annotated version with both the input and output in HTML) which replicates:

- Figure 5A and 5B
- Figure 7
- Figure 8A and 8B
- Figure 9A and 9B
- Figures A.2 and A.3 from the Online Appendix

Data Files:

All files are stored in both STATA format (`*.dta`) and in CSV (where the CSV files are written by the `do/StataCode.do` script). There are three original datasets here, where we describe each variable and provide a codebook for it if categorical:

1. Objective Prior and Posterior Beliefs

The following file pair in `data/` provide the raw data for our objective prior and posterior reports:

- `data-bsr-qsr.dta` a STATA-format file
- `data-bsr-qsr.csv` a comma-separated variable file

The two files contain identical data, where the variables are as follows:

Variable	Description	Labels
<code>id</code>	Observation ID	
<code>sessionid</code>	Session ID	
<code>scoringrule</code>	Scoring rule	1. BSR, 2. QSR

Variable	Description	Labels
treatment	Treatment	1. Information, 2. RCL Calculator, 3. No Information, 4. Feedback, 5.Description
subjectid	Subject ID	
scenario_order	Scenario order (session level)	
scenarioid	Scenario ID	
period	Period	
pur	Prior probability red urn (percent)	
p1r	Fraction of red balls in blue urn (percent)	
p2r	Fraction of red balls in red urn (percent)	
urn	Selected urn	0. blue, 1. red
draw1	First draw from selected urn	0. blue, 1. red
draw2	Second draw from selected urn	0. blue, 1. red
belief1	Prior elicitation	
belief2	Posterior after 1st draw	
belief3	Posterior after 2nd draw	
post1	Bayesian prior	

Variable	Description	Labels
post2	Bayesian posterior after 1st draw	
post3	Bayesian posterior after 2nd draw	
crt	CRT score	
risk_switch_pe	Risk switchpoint, PE method	
risk_switch_ce	Risk switchpoint, CE method	
female	Female (10 no response)	0. male, 1. female
agree1	I understood how my pay would be calculated	1. Strongly disagree, 2. Disagree, 3. Neither agree nor disagree, 4. Agree, 5. Strongly agree
agree2	I understood how my stated belief of the probability affected my pay	1. Strongly disagree, 2. Disagree, 3. Neither agree nor disagree, 4. Agree, 5. Strongly agree

Variable	Description	Labels
agree3	I always reported my most accurate guess on the Red urn being the selected urn	1. Strongly disagree, 2. Disagree, 3. Neither agree nor disagree, 4. Agree, 5. Strongly agree

2. Subjective Beliefs (NV-replication)

The following file pair in `data/` provide our replication of Niederle & Vesterlund (2007):

- `data-nv.dta` a STATA-format file
- `data-nv.csv` a comma-separated variable file

The two files contain identical data, where the variables are as follows:

Variable	Description	Labels
id		
sessionid	Session identifier	
subjectid	Participant identifier	
info	Information Treatment over BSR incentives	0. No Information, 1. Information
score1	Correct sums in round 1	
score2	Correct sums in round 2	
score3	Correct sums in round 3	
choice	Round 3, choice: 0=piece rate, 1=tournament	0. Piece rate, 1. Tournament

Variable	Description	Labels
b_to_1	Belief weight on tournament rank 1 (round 2)	
b_to_2	Belief weight on tournament rank 2 (round 2)	
b_to_3	Belief weight on tournament rank 3 (round 2)	
b_to_4	Belief weight on tournament rank 4 (round 2)	
female	What gender do you identify with?	0. Male, 1. Female

3. Incentives Only

The following file pair in `data/` provide the raw data for our incentives only treatment:

- `data-incentives-only.dta` a STATA-format file
- `data-incentives-only.csv` a comma-separated variable file

The two files contain identical data, where the variables are as follows:

Variable	Description	Labels
id	Observation ID	
choice_p20	Choice when $P(\text{Red})=0.2$	0. Lottery K (0%) 10. Lottery J (10%) 20. Lottery I (20%) 30. Lottery H (30%) 40. Lottery G (40%) 50. Lottery F (50%) 60. Lottery E (60%) 70. Lottery D (70%) 80. Lottery C (80%) 90. Lottery B (90%) 100. Lottery A (100%)

Variable	Description	Labels
choice_p30	Choice when $P(\text{Red})=0.3$	0. Lottery K (0%) 10. Lottery J (10%) 20. Lottery I (20%) 30. Lottery H (30%) 40. Lottery G (40%) 50. Lottery F (50%) 60. Lottery E (60%) 70. Lottery D (70%) 80. Lottery C (80%) 90. Lottery B (90%) 100. Lottery A (100%)
first_choice	Choice offered to subject first	20. 20% probability 30. 30% probability

Additional Data

We additionally include data from Holt & Smith (2016) downloaded from the paper's Replication data repository (see data citation).

References

- Holt, Charles A. and Smith, Angela M., "Belief Elicitation with a Synchronized Lottery Choice Menu That Is Invariant to Risk Attitudes," American Economic Journal: Microeconomics 8, 1 (2016), pp. 110–39. [Link](#)
- Holt, Charles A. and Smith, Angela M., "Replication data for: Belief Elicitation with a Synchronized Lottery Choice Menu That Is Invariant to Risk Attitudes" Nashville, TN: American Economic Association [publisher], 2016. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2019-10-12. [Link](#)
- Jann, Ben (2005). Making regression tables from stored estimates. The Stata Journal 5(3): 288-308. [Link](#)
- Jann, Ben (2007). Making regression tables simplified. The Stata Journal 7(2): 227-244 [Link](#)

- Niederle, Muriel and Vesterlund, Lise, "Do Women Shy Away From Competition? Do Men Compete Too Much?", The Quarterly Journal of Economics 122, 3, (2007), pp. 1067-1101. [Link](#)