Belief Elicitation and Behavioral Incentive Compatibility

David Danz, Lise Vesterlund, Alistair J. Wilson

Sun Fengfei

Renmin University of China

September 18, 2022

Outline

Outline

Authors

Department of Economics, University of Pittsburgh

- David Danz
- Lise Vesterlund
- Alistair J. Wilson

Outline

1. **Backgroud Introduction**
   - About Authors
   - Belief Elicitation Mechanisms
   - Main Purpose of the Paper
2. **Experimental Design**
   - Baseline Treatment
   - RCL & No-Information Treatment
   - Feedback Treatment
3. **Impact & Implication**
   - Discussion 1: Impact of Center-biased Belief
   - Discussion 2: Implication for Belief Elicitations
4. **Contribution & Discussion**

Incentive Compatible Rules

- Report **true** belief
- Truth-telling elicitation methods
  - Theoretical incentive compatible
  - Behavioral incentive compatible
- Scoring rules

  payment functions that depend on the report of the subject and the realization of the event

[1] Schlag, K. H., Tremewan, J., Van der Weele, J. J. (2015). A penny for your thoughts: A survey of methods for eliciting beliefs. Experimental Economics, 18(3), 457-490.

Basic Framework

Preliminary Framework

$\arg\max_{r\in\Theta} Eu(S(r,X))$, where
$Eu(S(r,\ X))=\sum_{x\in\mathcal{X}} u(S(r,x))P(X=x)$

$\Rightarrow \{\theta(X)\} = \arg\max_{r\in\Theta} Eu(S(r,X))$ for all $\mathrm{u} \in \mathrm{U}$ and all $\mathrm{P_X} \in \mathcal{P}_\mathrm{X}$

We say that θ can be elicited for the subjects with utility belonging to U if there is a scoring rule S that is a truth-telling rule for U

Common Belief Elicitation Mechanisms

- Proper scoring rules: risk neutrality
  - Quadratic Scoring Rule, QSR
  - ......
- No Requirement of risk neutrality of participants
  - Binarized Scoring Rule, BSR (*seem more precise to elicit true belief*)
  - Becker- DeGroot -Marschak Mechanism, BDM
  - Frequency Method & Interval Method
  - ......

Example of BDM

- Initial Method: Becker et al.(1964)
- Karni Method: Hard to fully understand (Karni, 2009)
- Revised Frequency Method (Schlag and Tremewan, 2021)

[1]Becker, G. M., Degroot, M. H., Marschak, J. (1964). Measuring utility by a single-response sequential method. Behavioral Science, 9(3), 226–232. doi:10.1002/bs.3830090304

[2]Karni, E. (2009). A mechanism for eliciting probabilities. Econometrica, 77(2), 603–606

[3]Schlag, K., Tremewan, J. (2021). Simple belief elicitation: An experimental evaluation. Journal of Risk and Uncertainty, 62(2), 137–155. doi:10.1007/s11166-021-09349-6

Instruction of Frequency or Interval Method

In this part of the experiment we will randomly select 20 participants from Part 1, excluding the participant you were matched with in that Part. How many of these participants do you think chose Option B. You will earn 1 Point if your guess is correct.
Remember: If a participant chose Option A, he/she would receive 10 Points no matter what the participant with whom they were matched chose. If a participant chose Option B, he/she would receive 15 Points if the participant with whom they were matched also chose Option B, and nothing if the participant with whom they were matched also chose Option A.
How many of the 20 randomly selected participants do you think chose Option B?

Example of Continuous Scoring Rule

Participants had a financial incentive for correct beliefs, but it was small, to avoid hedging.
If their estimation was exactly right, subjects received three experimental money units (0.8 dollar) in addition to their other experimental earnings.
They received two additional money units if their estimation deviated by only one point from the other group members' actual average contribution, one money unit if they deviated by two points, and no additional money if their estimation was off the actual contribution by more than three points.

[1] Fischbacher, U., Gächter, S. (2010). Social Preferences, Beliefs, and the Dynamics of Free Riding in Public Goods Experiments. American Economic Review, 100(1), 541–556. doi:10.1257/aer.100.1.541

Outline

Main Purpose of the Paper

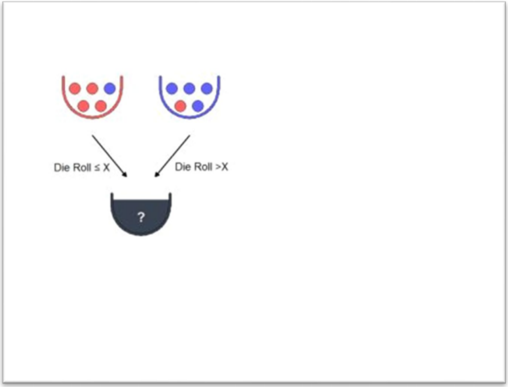
- To test whether incentives offered in BSR elicit **truthful** belief of participants
- To propose and check weak conditions for **behavioral** incentive compatible elicitations

Outline

Main Instructions

Slides	Script (read out loud by experimenter)
	We now summarize the task in each scenario. To begin with the computer fills the two urns. Each urn is filled with five balls, which are either blue or red. The red urn is the urn with more red balls in it.

Main Instructions



Next the computer selects one of the two urns for the scenario.

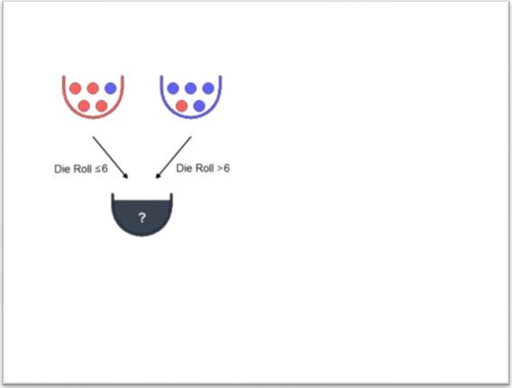
It does this using the rule X and a 10-sided die roll.

If the die roll is equal to less than X the red urn is selected.

If it's greater than X, the blue urn is selected.

Because of this rule, the chance of selecting the red urn is X-in-10.

Main Instructions



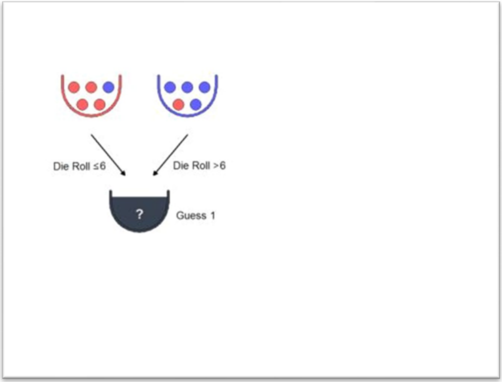
Suppose that X is equal to 6. So for die rolls of 1 to 6 the Red urn is selected.

And for die rolls from 7 to 10 the Blue urn is selected.

So the chance the red urn is selected is 6-in-10, or 60 percent.

The selected urn remains the same for the entire scenario.

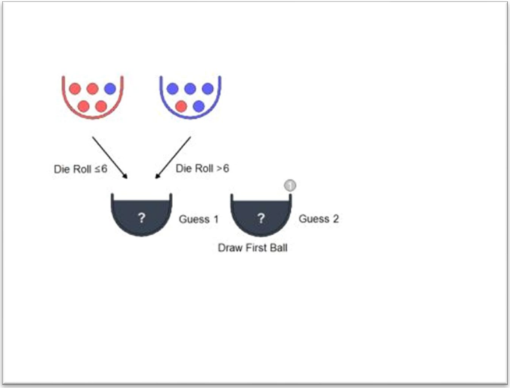
Main Instructions



After the computer has selected one of the two urn you make your first guess.

You make your first guess only knowing the die roll rule (here 6) and how many red and blue balls are in each urn.

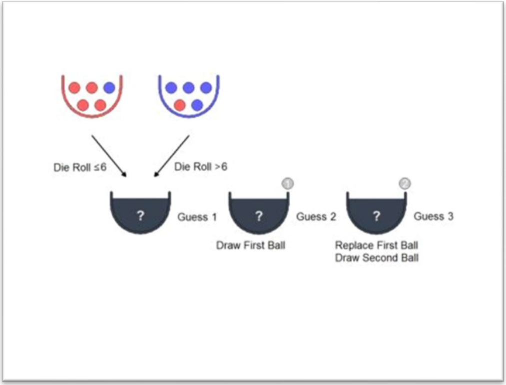
Main Instructions



After you make your first guess, you then get to see a drawn ball from the selected urn. The drawn ball can be either red or blue, where the chance of this depends on which urn was selected for the scenario.

After seeing the color of the drawn ball, you make your second guess.

Main Instructions

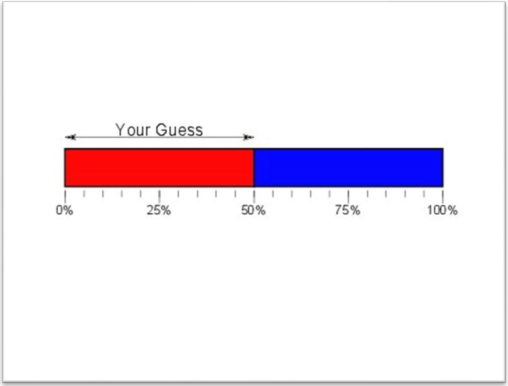


The first ball is put back into the selected urn, and the balls mixed.

You then draw a second ball from the urn and see what color it is.

After seeing the color, you make your third and final guess.

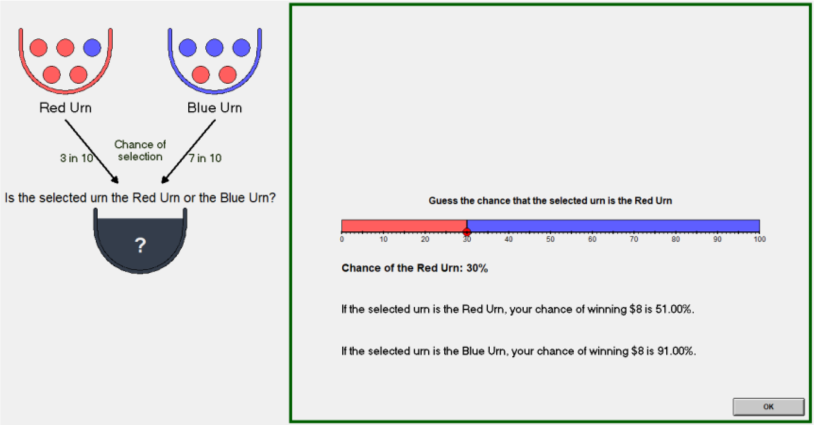
Main Instructions



You enter Your Guesses by clicking the response bar on your screen.

The width of the red part of the bar indicates your percentage chance that the red urn was selected.

Main Instructions



Incentive Lotteries

- Objective prior: π_0
- State - contingent lottery pair
  - $1-(1-q)^2$: winning 8 dollars of Red urn was selected
  - $1-q^2$: winning 8 dollars of Blue urn was selected
- Dominant strategy: truthful reporting

"The payment rule is designed so that you can secure the largest chance of winning the prize by reporting your most-accurate guess."

Illustration of a subset of lotteries

Submitted Belief on Red	Chance of receiving \$8 if: Urn is Red	Urn is Blue
1.0	100%	0%
0.9	99%	19%
0.8	96%	36%
0.7	91%	51%
0.6	84%	64%
0.5	75%	75%

Dominant strategy

Subject's problem: Max U(q)

$U(q)=\pi_0\left[1-(1-q)^2\right]+(1-\pi_0)\left(1-q^2\right), q\in[0,1]$

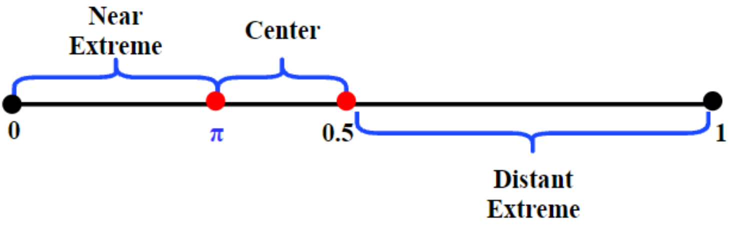
$\Rightarrow q^* = \pi_0$

$\Rightarrow$ **True belief** is the **induced prior** π_0

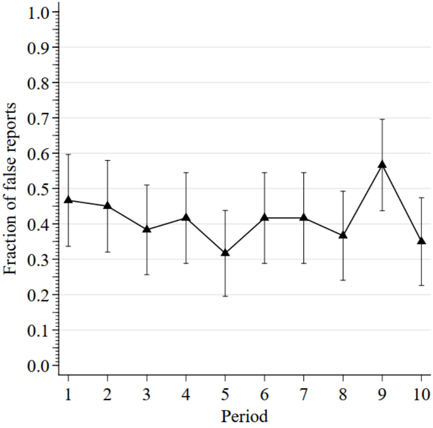
True belief is belief reports of induced prior

False belief is belief reports deviated from induced prior

Illustration of Types of False Reports

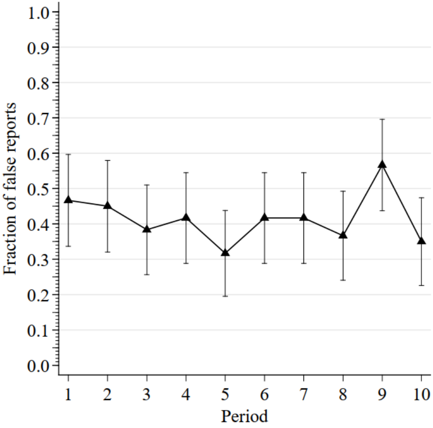


Basic Results

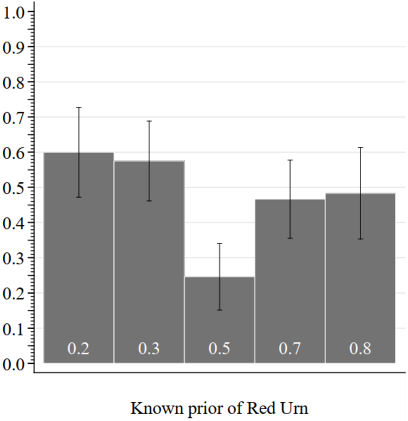


(A) By Period

Basic Results

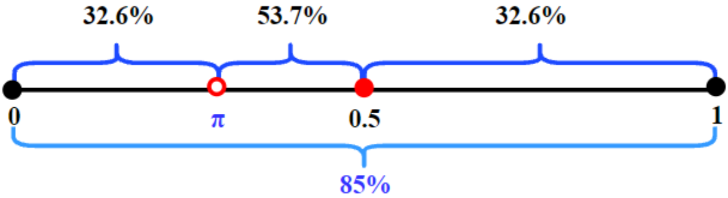


(A) By Period



(B) By Prior

Basic Results



Why so many participants misreporting the prior?

Potential Source of False Reporting

- Fail to understand the task
- **Fail to reduce compound lottery of BSR** ($incentive\ mechanism$)
- **Other motivation related to BSR** (e.g. hedging motive[1])

[1] preferring a substantial increase in winning on the unlikely event in return for a slight decrease in the chance of winning on the likely event

Additional Treatments

- **RCL Treatment**(Reduction of Compound Lotteries Treatment)
  $\Rightarrow$ Limit misunderstanding of compound lottery
- **No-Information Treatment**
  $\Rightarrow$ Identify other factors(not the incentives)

Outline

1 **Backgroud Introduction**
- About Authors
- Belief Elicitation Mechanisms
- Main Purpose of the Paper

2 **Experimental Design**
- Baseline Treatment
- RCL & No-Information Treatment
- Feedback Treatment

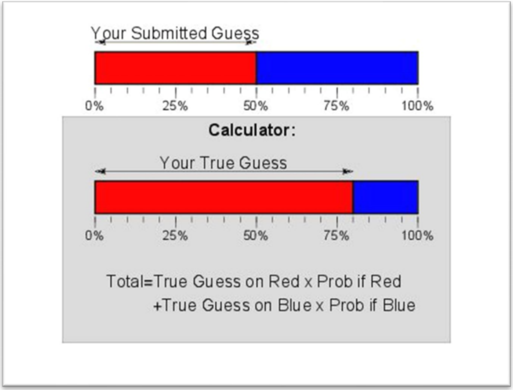
3 **Impact & Implication**
- Discussion 1: Impact of Center-biased Belief
- Discussion 2: Implication for Belief Elicitations

4 **Contribution & Discussion**

Treatment Designs

- RCL treatment:providing a calculator to **compute** the total chance of winning 8 dollars
  $\Rightarrow$ Assess the Source 2: the extent of inability to reduce the compound lotteries
- No-information treatment: removing all **quantitative** information on the incentives
  $\Rightarrow$ Assess the Source 1: factors other than incentives (e.g. confusion)
- Treatment Difference
  $\Rightarrow$ Assess the Source 1: factors related to BSR incentives

RCL Treatment Description



In addition to the bar where you enter Your Submitted Guess, we also provide you with a calculator.

To use the calculator, you enter *Your True Best* guess.

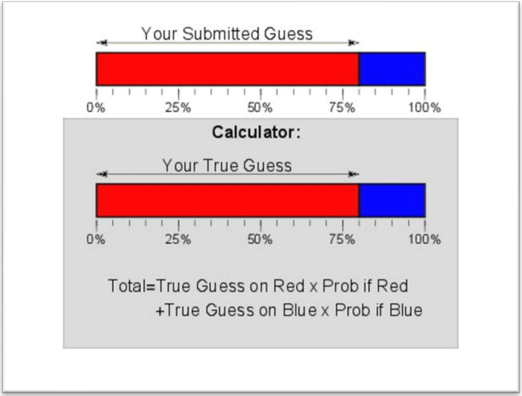
For any selection of Your Submitted Guess and Your True best Guess the calculator will provide you with your total chance of winning.

Your total chance of winning is calculated as

Your True Best Guess on Red times the Likelihood that you Win if Red is Selected, given Your Submitted Guess

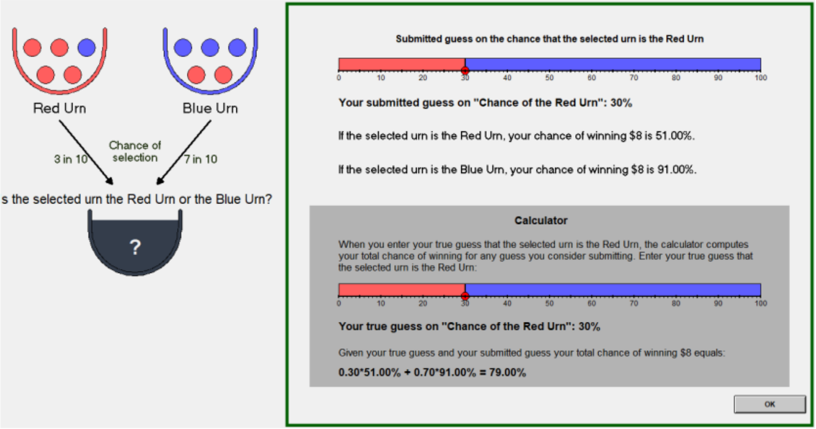
+ Your True Best Guess on Blue times the Likelihood that you Win if Blue is Selected, given Your Submitted Guess

RCL Treatment Description

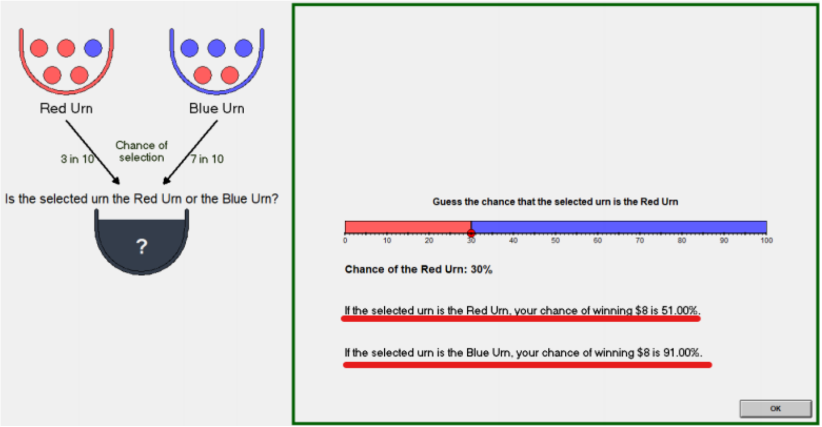


The calculator allows you to verify that whatever your True Best Guess might be, the payment rule ensures that you will maximize your total chance of winning by setting your submitted guess equal to your True Guess.

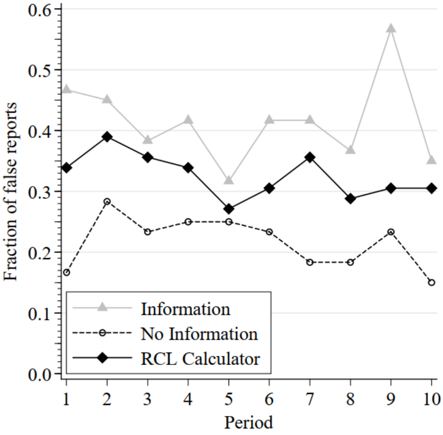
RCL Treatment Description



No-Information Treatment Description

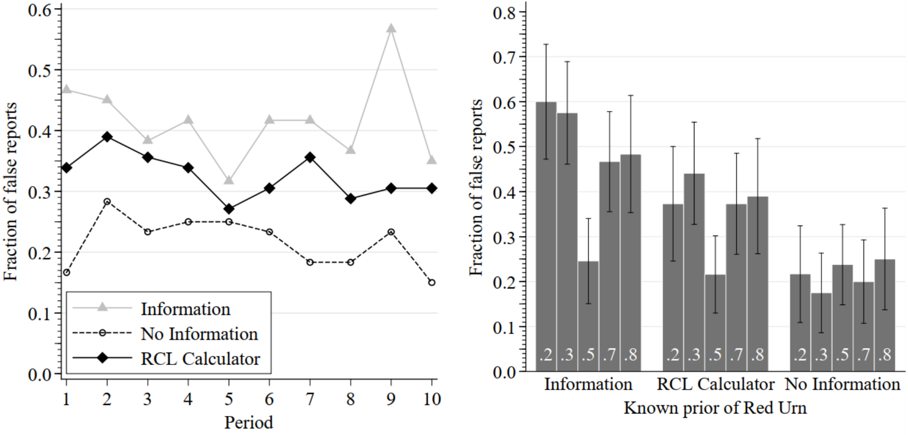


Basic Results



(A) By Period

Basic Results



(A) By Period

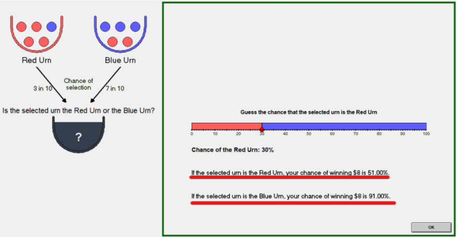
(B) By Prior

Outline

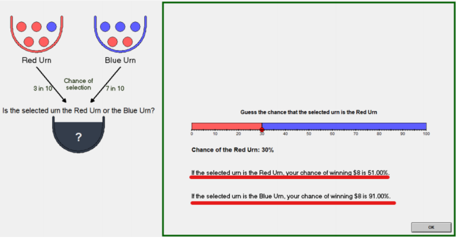
Basic Description

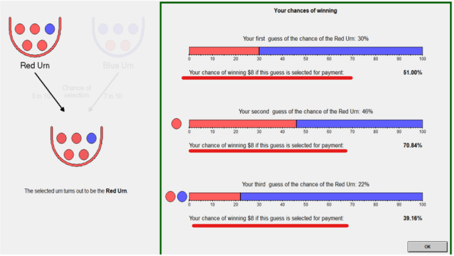
- Identify the effect from information on quantitative
- Replicate No-Information instructions and main decision screen
- After belief elicitation, provide end-of-period feedback

Interface Screenshots

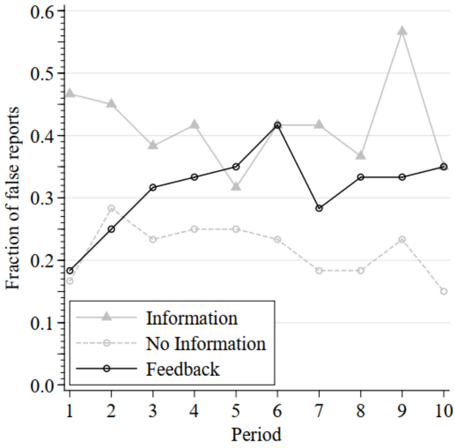


Interface Screenshots



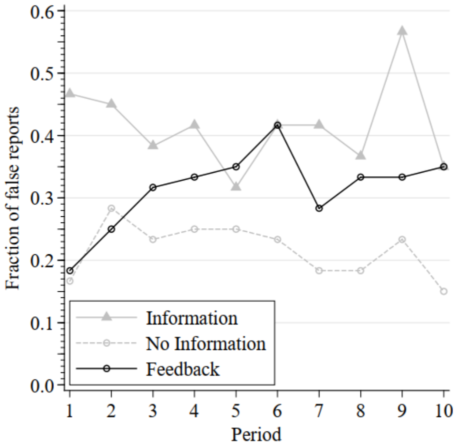


Results

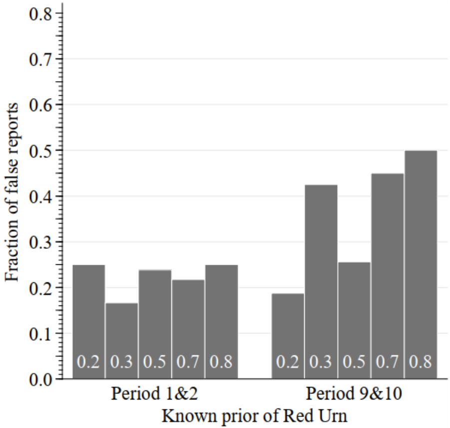


(A) By Period

Results



(A) By Period



(B) By Prior

Outline

Background Information

- Examine the study of Niederle and Vesterlund(2007)
- Mirror decision to perform under a non-competitive piece-rate or a competitive tournament.
- Belief elicitation: Information v.s. No-information design
  - BSR elicitation with information on quantative incentives
  - BSR elicitation without that information
- Purpose:
  - To test whether there is empirical evidence for center-bias effect
  - To evaluate potential inferential impact of using elicited beliefs fro analysis

[1] Niederle, M., Vesterlund, L. (2007). Do women shy away from competition? Do men compete too much?. The quarterly journal of economics, 122(3), 1067-1101.

Model of Inferential Effects

Simplifying Model

$\mathsf{q}_i = \mu_q + \delta_q \cdot \text{ Female }_i + \epsilon_q$

ϵ_q: independent mean finite variance

μ_q : true mean of male belief$_q$:estimated gender gap in confidence

Difference in confidence

$\mathsf{q}_i = \mu_q + \delta_q \cdot \text{ Female }_i + \epsilon_q$
$\mathsf{q}_i = \alpha \cdot \mathrm{c} + (1-\alpha) \cdot q^*$
(1) When $\alpha = 0 : E(\hat{\delta} q) = \delta_q$, thus unbiased
(2) When $\alpha > 0 : E(\hat{\delta} q) = (1-\alpha) \cdot \delta_q$, thus biased

Basic Derivation

$$\hat{\delta}_q = \frac{\sum(x_i - \bar{x})y_i}{\sum(x_i - \bar{x})^2}$$
$$= \frac{\sum(x_i - \bar{x})[\alpha c + (1-\alpha)(\mu_q + \delta_q x_i + \varepsilon_q)]}{\sum(x_i - \bar{x})^2}$$
$$= \frac{(1-\alpha)\delta_q(x_i - \bar{x})x_i + (1-\alpha)\sum(x_i - \bar{x})\varepsilon_q}{\sum(x_i - \bar{x})^2}$$
$$= (1-\partial)\delta q + \frac{(1-\alpha)\Sigma(x_i - \bar{x})\cdot\varepsilon_q}{\sum(x_i - \bar{x})^2}$$

$$E\left(\hat{\delta}_q\right) = \text{plim}(\hat{\delta} q) = (1-\alpha)\delta q$$
$\Rightarrow$ Asymptotic bias $= -\alpha \cdot \delta q$, biased in the estimator

Tournament Entry Decision

Simplifying Model

$\mathrm{y}_i = \mu_y + \delta_y \cdot \mathrm{Female}_i + \beta_q \cdot q_i + v_i$

v_i: independent mean finite variance

$\hat{\delta}_y$: estimated gender gap in tournament entry decision

Basic Inference

$$q_i = \mu_q + \delta_q x_i + \varepsilon_i = \begin{cases} \mu_q + \delta_q + \varepsilon_i & i = \text{ Female} \\ \mu_q + \varepsilon_i & i = \text{ Male} \end{cases}$$

$$\begin{aligned} y_i &= \mu_y + \delta_y \cdot x_i + \beta_q \cdot q_i + v_i \\ &= \begin{cases} \mu_y + \delta_y + \beta_q \left(\mu_q + \delta_q + \varepsilon_i\right) + v_i & i = \text{ Female} \\ \mu_y + \beta_q \left(\mu_q + \varepsilon_i\right) + v_i & i = \text{ Male} \end{cases} \end{aligned}$$

$$\Rightarrow \text{ Diff } = \delta_y + \beta_q \cdot \delta_q$$

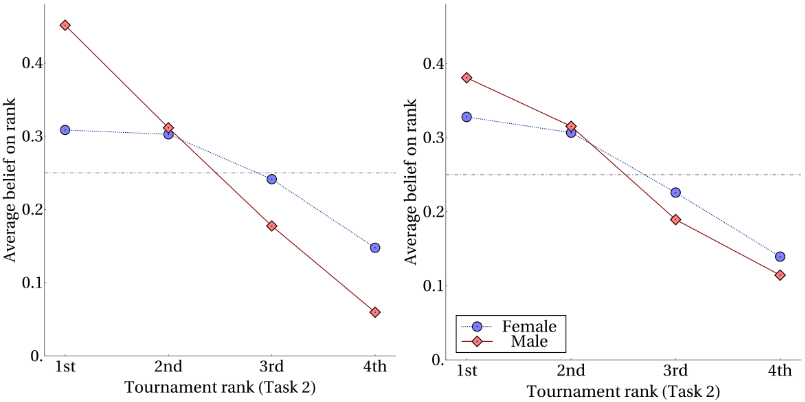
Tournament Entry Decision

Gender Difference in Entry Decision

$\text{Diff} = \delta_y + \beta_q \cdot \delta_q$

β_q: positive effect (confidence have positive effect on entry)
δ_q: negative effect (men are more confident than women)
$\beta_q \cdot \delta_q$: negative effect (marginal effect of confidence on tournament entry)
$\Rightarrow \hat{\delta_y}$: being **overestimated**
$\Rightarrow$:Center biased beliefs would **overestimate** the size of the gender gap in tournament entry.

Elicited Likelihood of Performance Rank



(A) NV-no-information

(B) NV-information

Replication Regression Results

	DEPENDENT VARIABLE (CF. TABLE V IN NV 2007)		INDEPENDENT VARIABLE (CF. TABLE II AND VI IN NV 2007)			
	Belief on 1st rank (OLS)		Tournament entry (Probit)			
	No-Inform.	Information	No-Information		Information	
	(1)	(2)	(3)	(4)	(5)	(6)
Female	−0.148	−0.038	−0.258	−0.146	−0.357	−0.382
	(0.051)	(0.059)	(0.080)	(0.115)	(0.118)	(0.124)
Tournament	0.008	0.018	0.022	0.017	0.006	−0.011
	(0.004)	(0.005)	(0.012)	(0.014)	(0.011)	(0.014)
Tournament–piece rate	0.017	−0.021	0.001	−0.015	−0.023	−0.004
	(0.008)	(0.010)	(0.017)	(0.021)	(0.021)	(0.022)
Constant	0.305	0.059				
	(0.097)	(0.101)				
Belief weight on 1st rank				1.275		0.994
				(0.432)		(0.329)
N	74	68	74	74	68	68
R^2/adj. R^2	0.273	0.187	0.157	0.303	0.093	0.208

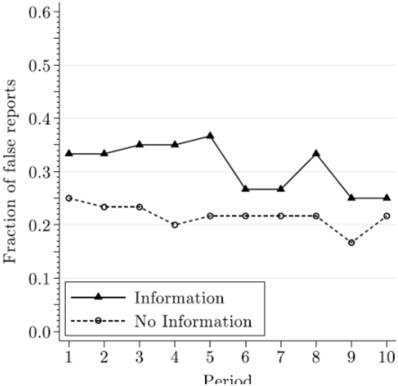
Outline

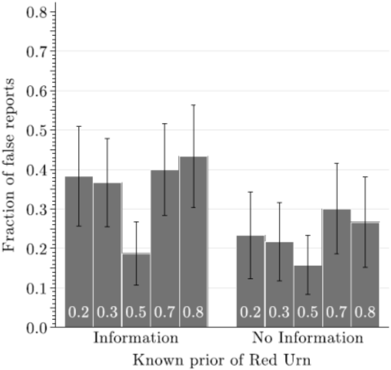
Questions Behind the Results

- Would quantitative information **only** distort reports under **BSR**?
  $\Rightarrow$ **QSR** **Information** v.s. **No-information** Treatments ▸ Result
- Would description of the mechanism's implementation rule influence rate of false reports?
  $\Rightarrow$ **Description** Treatment ▸ result

[1] Quadratic Scoring Rule

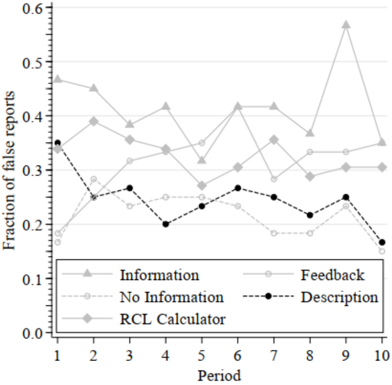
Result of QSR Two Treatments



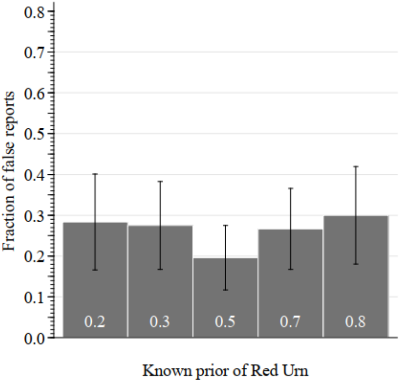


Return

Result of Description Treatment



(A) By Period



(B) By Prior

Behavioral Incentive Compatibility

- Information on deployed incentives increases truthful revelation
  $\Rightarrow$ Being examined by the experiments

Behavioral Incentive Compatibility

- Information on deployed incentives increases truthful revelation
  $\Rightarrow$ Being examined by the experiments
- Most participants select the outcome uniquely maximizing the outcome
  $\Rightarrow$ **Incentive-only** Treatment
  - Provide the chance over the set of lottery pairs underlying the BSR *instead of* elicitation framing
  - Violating of this weak condition, most participants fail to select the lottery thought to be maximized

Brief Summary

- To test whether incentives offered in BSR elicit **truthful** belief of participants
  $\Rightarrow$ Quantitative information on incentives **increases** deviations from truthful reporting and causes systematic **bias** toward the center
- To propose and check violations of weak conditions for **behavioral** incentive compatible elicitations
  - Information on deployed incentives increases truthful revelation
  - Most participants select the outcome uniquely maximizing the outcome

Innovation

- Establish a method to identify truth-reporting behavior and the sources behind center-biased beliefs
- Provide experimental evidence to show that BSR may NOT outperform other incentive mechanism.
- Provide empirical evidence of impact on inference of using center-biased reports under BSR

Further Questions

- What the reasons for formation of **non-center-biased** belief?
- Does cognitive capability of participants (cognitive effect) matter in the experiments?
- How to measure the extent of justification of behavior through stated belief?

Thanks for listening!