The Selective Disclosure of Evidence

An Experiment

Agata Farina
NYU

Guillaume Fréchette
NYU

Alessandro Lizzeri
Princeton

Jacopo Perego
Columbia

February 2024

Motivation

In many settings, agents communicate by disclosing *selected* evidence:

- *Journalists* select which facts to report in their articles
- *Managers* select which results to discuss in performance reports
- *Job candidates* select which achievements to list on their CVs

Motivation

In many settings, agents communicate by disclosing *selected* evidence:

- *Journalists* select which facts to report in their articles
- *Managers* select which results to discuss in performance reports
- *Job candidates* select which achievements to list on their CVs

These are examples of **selective disclosure**:

- A sender can disclose K pieces of verifiable evidence, which she selects from a privately observed pool of N pieces

Motivation

In many settings, agents communicate by disclosing *selected* evidence:

- *Journalists* select which facts to report in their articles
- *Managers* select which results to discuss in performance reports
- *Job candidates* select which achievements to list on their CVs

These are examples of **selective disclosure**:

- A sender can disclose K pieces of verifiable evidence, which she selects from a privately observed pool of N pieces

A pervasive force in communication, e.g., a principal source of news media bias

("filtering," Gentzkow et al. '14)

This Paper

An experimental study on **selective disclosure**

We build on a small theoretical literature on selective disclosure of noisy evidence

Milgrom ('81, Bell), Fishman and Hagerty ('90, QJE), Di Tillio et al ('21, Ecma)

Our model generates rich comparative statics in N and K for outcomes such as:

- Which evidence do senders disclose?
- How much information do they transmit to receivers?
- Do receivers account the selection in the evidence they see?

These comparative statics inform a novel experimental design and provide a rigorous test of the theory

Preview of Results

Our data corroborates the key qualitative predictions of the theory

- Validation of **selective disclosure** as a force in communication that is behaviorally descriptive

We document two novel *quantitative* deviations from the theory:

- A form of **deception aversion** in senders leads to **overcommunication**
- Evidence of **selection neglect** in a *strategic* setting

Policy implications: Mandating disclosure can be ineffective (and possibly detrimental) when selection opportunities are large

Related Literature: Theory

Classic disclosure models focus on **rich** evidence (e.g., Grossman, '81; Milgrom, '81; Jovanovic, '82; Okuno-Fujiwara et al., '90)

- Senders can verifiably disclose their type $\rightsquigarrow$ *unravelling* results

Related Literature: Theory

Classic disclosure models focus on **rich** evidence (e.g., Grossman, '81; Milgrom, '81; Jovanovic, '82; Okuno-Fujiwara et al., '90)

- Senders can verifiably disclose their type $\rightsquigarrow$ *unravelling* results

We focus on settings where evidence is **not rich** and, thus, unravelling does not occur Fishman and Hagerty ('90, QJE), Di Tillio et al ('21, Ecma)

- This enables nontrivial comparative statistics, which are instrumental for testing the theory

Related Literature: Theory

Classic disclosure models focus on **rich** evidence (e.g., Grossman, '81; Milgrom, '81; Jovanovic, '82; Okuno-Fujiwara et al., '90)

- Senders can verifiably disclose their type $\rightsquigarrow$ *unravelling* results

We focus on settings where evidence is **not rich** and, thus, unravelling does not occur Fishman and Hagerty ('90, QJE), Di Tillio et al ('21, Ecma)

- This enables nontrivial comparative statistics, which are instrumental for testing the theory

Related but less connected settings: Glazer and Rubinstein ('04, Ecma; '06, TE), Shin ('03, Ecma), Dziuda ('11, JET), Haghtalab et al. ('21), Gao ('23)

Related Literature: Empirics

Disclosure settings with verifiable and "rich" evidence — Forsythe et al ('89, RAND), Jin and Leslie ('03, QJE), Jin, Luca and Martin ('22, AEJ: Micro)

Settings with unverifiable evidence, i.e., **cheap talk** — Cai and Wang ('06, GEB)

Recent and related settings with selective disclosure — Degan, Li, Xie ('23, CJE), Penczynski, Koch, Zhang ('23)

Methodologically: (close to Frechette, Lizzeri and Perego (2022, Ecma))

- Exploit comparative statics to test underlying forces in the theory

Outline

1. Model
2. Equilibrium and Testable Predictions
3. Experimental Design
4. Results

model

Model

from Milgrom (1981)

Sender privately observes the state $\theta \in \Theta$:

- Θ finite and ordered, $p \in \Delta(\Theta)$ common prior

Model

from Milgrom (1981)

Sender privately observes the state $\theta \in \Theta$:

- Θ finite and ordered, $p \in \Delta(\Theta)$ common prior

Given θ, Sender draws N i.i.d. **signals**

- Exogenous info structure $f : \Theta \to \Delta(S)$, S finite and ordered, MLRP
- $f(\cdot|\theta)$ has full support for every θ

Model

from Milgrom (1981)

Sender privately observes the state $\theta \in \Theta$:

- Θ finite and ordered, $p \in \Delta(\Theta)$ common prior

Given θ, Sender draws N i.i.d. **signals**

- Exogenous info structure $f : \Theta \to \Delta(S)$, S finite and ordered, MLRP
- $f(\cdot|\theta)$ has full support for every θ

Sender can disclose up to K of the N available signals $\rightsquigarrow$ message

Model

from Milgrom (1981)

Sender privately observes the state $\theta \in \Theta$:

- Θ finite and ordered, $p \in \Delta(\Theta)$ common prior

Given θ, Sender draws N i.i.d. **signals**

- Exogenous info structure $f : \Theta \to \Delta(S)$, S finite and ordered, MLRP
- $f(\cdot|\theta)$ has full support for every θ

Sender can disclose up to K of the N available signals $\rightsquigarrow$ message

Receiver observes the message and takes an action $a \in A$

Payoffs

Given state θ and action a,

- Receiver's payoff is $u(\theta, a) = -(a - \theta)^2$ — wants to guess the state
- Sender's payoff is $v(\theta, a) = a$ — higher actions preferred

Discussion

No message can verifiably reveal $\theta \rightsquigarrow$ failure of richness (Okuno-Fujiwara et al., '90)

Sender does not choose N, i.e., available evidence is **exogenous**

If $K = N$, the sender can disclose **all** her available evidence if so she wants

If $K < N$, sender can cherry pick which evidence to disclose

- $K < N$ captures exogenous communication constraints

Changes in K and N generate rich testable predictions, which we use as a test of the theory

equilibrium

Equilibrium

Unlike typical disclosure games with “rich” evidence structure, our game admits multiple **PBEs** outcomes

Equilibrium

Unlike typical disclosure games with “rich” evidence structure, our game admits multiple **PBEs** outcomes

Focus on sender’s strategies that are pure and weakly increasing in θ

Equilibrium

Unlike typical disclosure games with "rich" evidence structure, our game admits multiple **PBEs** outcomes

Focus on sender's strategies that are pure and weakly increasing in θ

We refine the equilibrium set using **neologism proofness**, adapted to our setting with verifiable information — Farrel ('93, *GEB*)

appendix

Under this refinement, our game admits a **unique** equilibrium outcome

Equilibrium

More formally, consider the following class of sender's strategies:

Definition

A sender's strategy is **maximally selective** if, given the available signals, she discloses the K-highest ones.

Equilibrium

More formally, consider the following class of sender's strategies:

Definition

A sender's strategy is **maximally selective** if, given the available signals, she discloses the K-highest ones.

Proposition 1

There exists a PBE in which the sender plays a maximally selective strategy

(Milgrom 1981)

Moreover, the outcome it induces is unique in the class of neologism-proof PBEs

Informativeness

Main outcome of interest is the **informativeness** the equilibrium strategies

- I.e., how effectively sender and receiver communicate

Informativeness

Main outcome of interest is the **informativeness** the equilibrium strategies

- I.e., how effectively sender and receiver communicate

We measure informativeness as the **correlation** btw θ and a, denoted by $\mathcal{I} = Corr(\theta, a)$ as in Lizzeri, Frechette, Perego ('22, Ecma)

Main Comparative Statics

Rich predictions regarding how informativeness changes in K and N

Proposition 2

Fixing N, Equilibrium informativeness increases in K

Fixing K, equilibrium informativeness can increase for small N but eventually decreases to zero as $N \to \infty$

If $K = N$, equilibrium informativeness increases in N

Main Comparative Statics

Rich predictions regarding how informativeness changes in K and N

Proposition 2

Fixing N, Equilibrium informativeness increases in K

Fixing K, equilibrium informativeness can increase for small N but eventually decreases to zero as $N \to \infty$

If $K = N$, equilibrium informativeness increases in N

That is, giving the sender more discretion can make the receiver better off

Main Comparative Statics

Rich predictions regarding how informativeness changes in K and N

Proposition 2

Fixing N, Equilibrium informativeness increases in K

Fixing K, equilibrium informativeness can increase for small N but eventually decreases to zero as $N \to \infty$

If $K = N$, equilibrium informativeness increases in N

That is, giving the sender more discretion can make the receiver better off

see also Fishman and Hagerty ('90, QJE),
Di Tillio, Ottaviani and Sorensen ('21, Ecma)

Example: Conclusive Good News

Suppose $\Theta = \{\theta_L, \theta_H\}, \quad p(\theta_H) = \frac{1}{2}, \quad S = \{A, B\}, \quad K = 1$

Example: Conclusive Good News

Suppose $\Theta = \{\theta_L, \theta_H\}, \quad p(\theta_H) = \frac{1}{2}, \quad S = \{A, B\}, \quad K = 1$

$f(s\|\theta)$	Signal	
State	A	B
θ_L	0	1
θ_H	γ	$1-\gamma$

Example: Conclusive Good News

Suppose $\Theta = \{\theta_L, \theta_H\}, \quad p(\theta_H) = \frac{1}{2}, \quad S = \{A, B\}, \quad K = 1$

$f(s\|\theta)$	Signal	
State	A	B
θ_L	0	1
θ_H	γ	$1-\gamma$

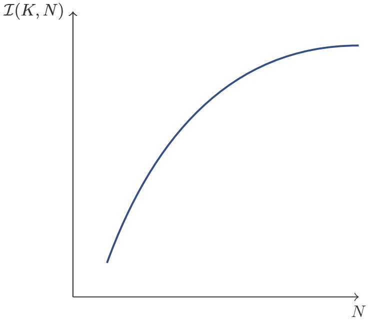
$$\mathcal{I}(K, N) = \frac{1}{4} - \frac{(1-\gamma)^N}{2(1+(1-\gamma)^N)}$$

Example: Conclusive Good News

Suppose $\Theta = \{\theta_L, \theta_H\}$, $p(\theta_H) = \frac{1}{2}$, $S = \{A, B\}$, $K = 1$

$f(s\|\theta)$	Signal	
State	A	B
θ_L	0	1
θ_H	γ	$1-\gamma$

$$\mathcal{I}(K,N) = \frac{1}{4} - \frac{(1-\gamma)^N}{2(1+(1-\gamma)^N)}$$



Example: Inconclusive News

Suppose $\Theta = \{\theta_L, \theta_H\}$, $\quad p(\theta_H) = \frac{1}{2}$, $\quad S = \{A, B\}$, $\quad K = 1$

Example: Inconclusive News

Suppose $\Theta = \{\theta_L, \theta_H\}$, $p(\theta_H) = \frac{1}{2}$, $S = \{A, B\}$, $K = 1$

If $\eta \in (0, \frac{1}{2})$ and $\gamma \in (\frac{1}{2}, 1)$:

$f(s\|\theta)$	Signal	
State	A	B
θ_L	η	$1-\eta$
θ_H	γ	$1-\gamma$

Example: Inconclusive News

Suppose $\Theta = \{\theta_L, \theta_H\}$, $\quad p(\theta_H) = \frac{1}{2}$, $\quad S = \{A, B\}$, $\quad K = 1$

If $\eta \in (0, \frac{1}{2})$ and $\gamma \in (\frac{1}{2}, 1)$:

$f(s\|\theta)$	Signal	
State	A	B
θ_L	η	$1-\eta$
θ_H	γ	$1-\gamma$

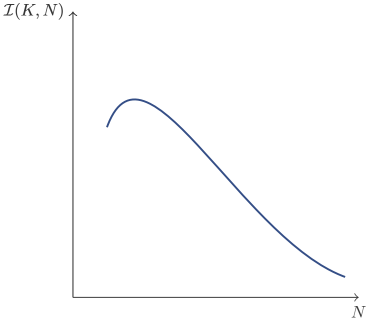
$\mathcal{I}(K, N) =$ more complex

Example: Inconclusive News

Suppose $\Theta = \{\theta_L, \theta_H\}$, $p(\theta_H) = \frac{1}{2}$, $S = \{A, B\}$, $K = 1$

If $\eta \in (0, \frac{1}{2})$ and $\gamma \in (\frac{1}{2}, 1)$:

$f(s\|\theta)$	Signal	
State	A	B
θ_L	η	$1-\eta$
θ_H	γ	$1-\gamma$



$\mathcal{I}(K, N) =$ more complex

experiment

Experimental Design

Binary state: θ_L and θ_H, equal probability

Four possible signals $S = \{A, B, C, D\}$

Information structure f:

	Signal			
State	A	B	C	D
θ_L	10%	20%	25%	45%
θ_H	45%	25%	20%	10%

Receiver's action $a \in [0, 1]$, which makes it equivalent to a belief elicitation task

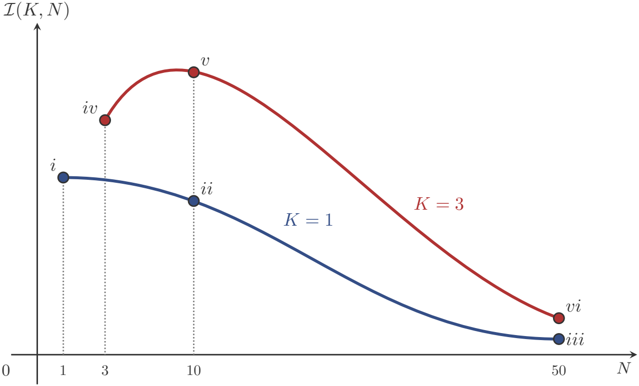
implemented using BSR (Hossain and Okui, '13 *Restud*) Details

Treatments

We vary K and N as follows:

	$N = 1$	$N = 3$	$N = 10$	$N = 50$
$K = 1$	$\checkmark$	$\cdot$	$\checkmark$	$\checkmark$
$K = 3$	$\cdot$	$\checkmark$	$\checkmark$	$\checkmark$

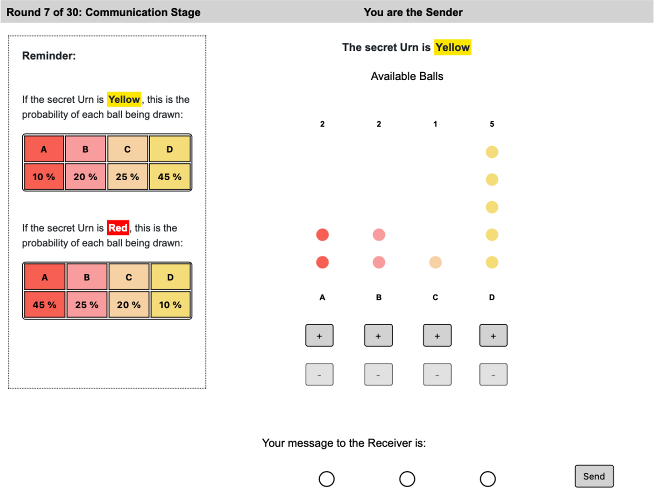
Main Comparative Statics



Experimental Details

- Undergrad population Columbia and NYU: Spring, Summer, Fall 2023
- 420 subjects, between-subject design
- 6 treatments
- 4 sessions per treatment
- 30 rounds per session, random rematching
- 17.5 subjects per sessions on average
- Average payout $30 per subject
- Fixed roles

User Interface: $N = 10$ and $K = 3$



Round 7 of 30: Communication Stage — **You are the Sender**

Reminder:

If the secret Urn is **Yellow**, this is the probability of each ball being drawn:

A	B	C	D
10 %	20 %	25 %	45 %

If the secret Urn is **Red**, this is the probability of each ball being drawn:

A	B	C	D
45 %	25 %	20 %	10 %

The secret Urn is Yellow

Available Balls

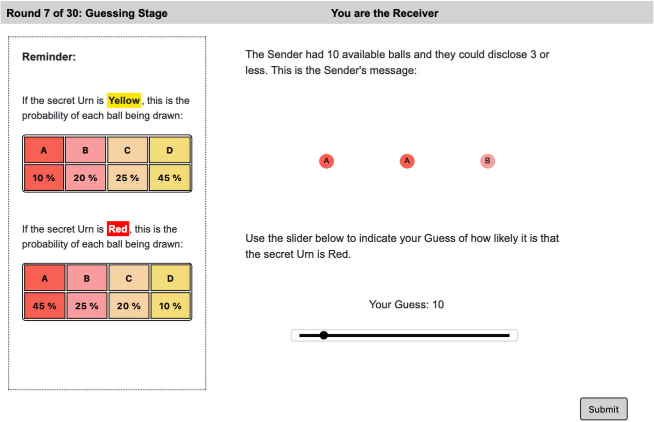
A	B	C	D
2	2	1	5

Your message to the Receiver is:

Send

User Interface: $N = 10$ and $K = 3$

User Interface: $N = 10$ and $K = 3$



results

Outline

Progression of our analysis:

– Which evidence do senders disclose?

– How informative is it?

– How do receivers respond to it?

result 1

(which evidence is disclosed)

Which Evidence is Disclosed?

Question 1: Which evidence do senders disclose?

Which Evidence is Disclosed?

Question 1: Which evidence do senders disclose?

Theory predicts that:

- If N increases, the evidence disclosed should become **more** favorable

  sender can be more selective with larger sample

- If K increases, evidence disclosed should become **less** favorable

  held to higher a standard, sender needs to be less selective

Which Evidence is Disclosed?

Question 1: Which evidence do senders disclose?

Theory predicts that:

- If N increases, the evidence disclosed should become **more** favorable

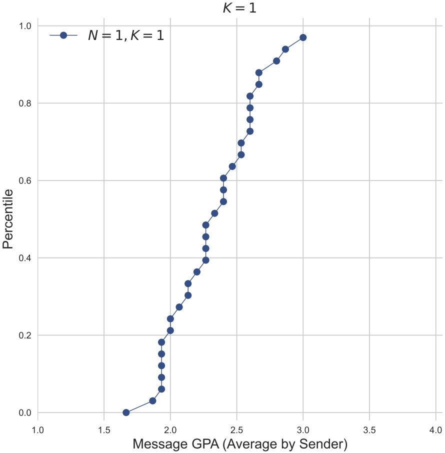
  sender can be more selective with larger sample

- If K increases, evidence disclosed should become **less** favorable

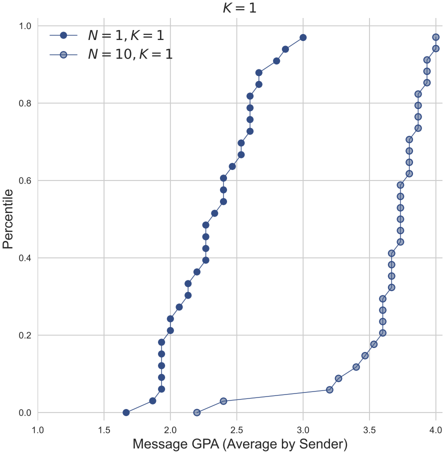
  held to higher a standard, sender needs to be less selective

To test this, we compute the **GPA** of each message ($A \rightsquigarrow 4$, $B \rightsquigarrow 3$, etc) and study how it changes in N and K

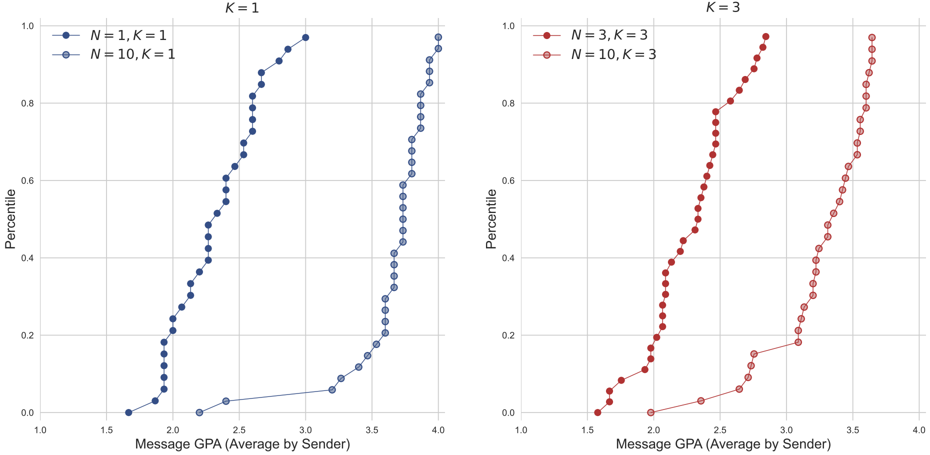
Which Evidence is Disclosed?



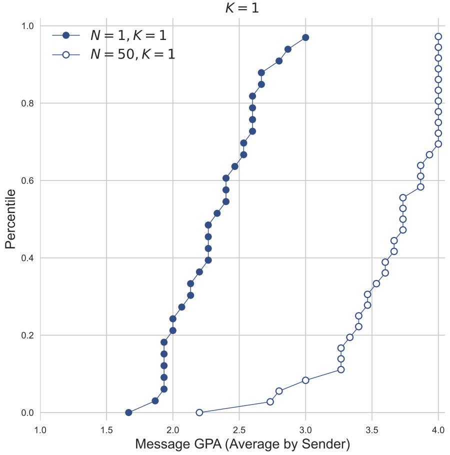
Which Evidence is Disclosed?

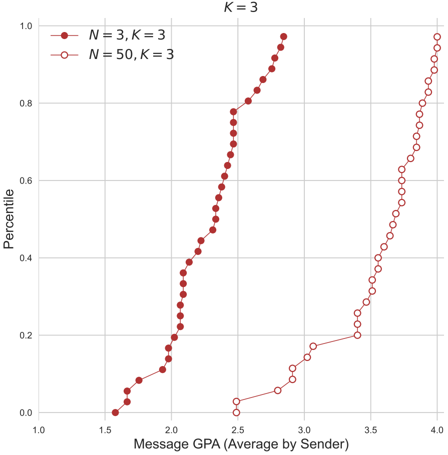


Which Evidence is Disclosed?

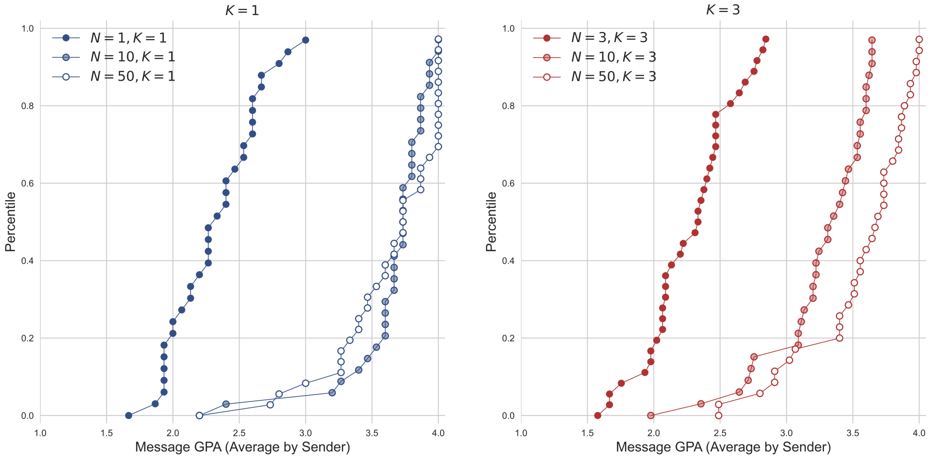


Which Evidence is Disclosed?

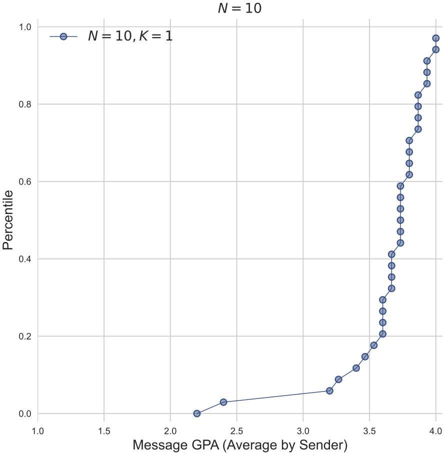




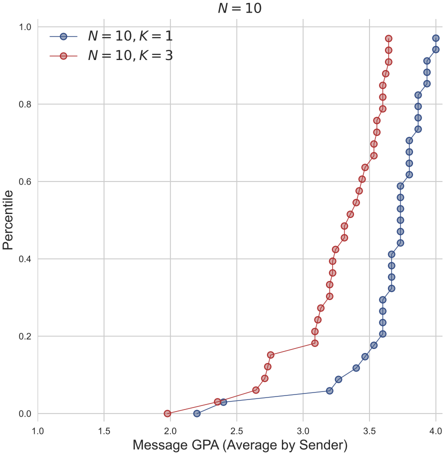
Which Evidence is Disclosed?



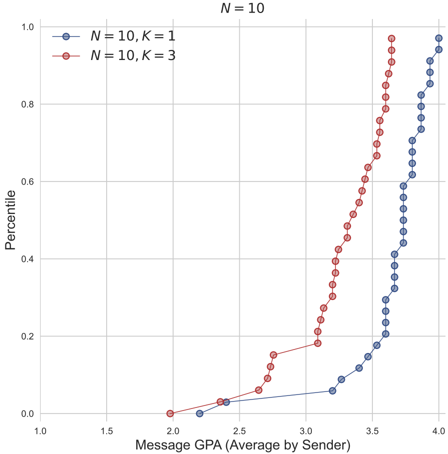
Which Evidence is Disclosed?

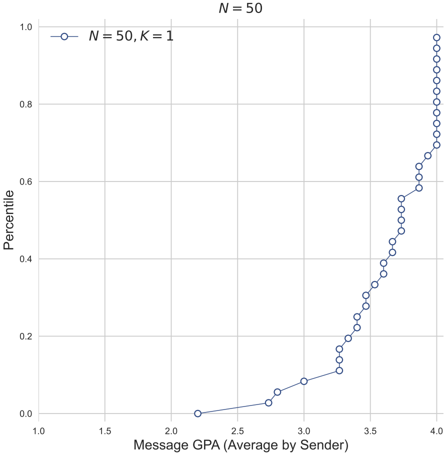


Which Evidence is Disclosed?

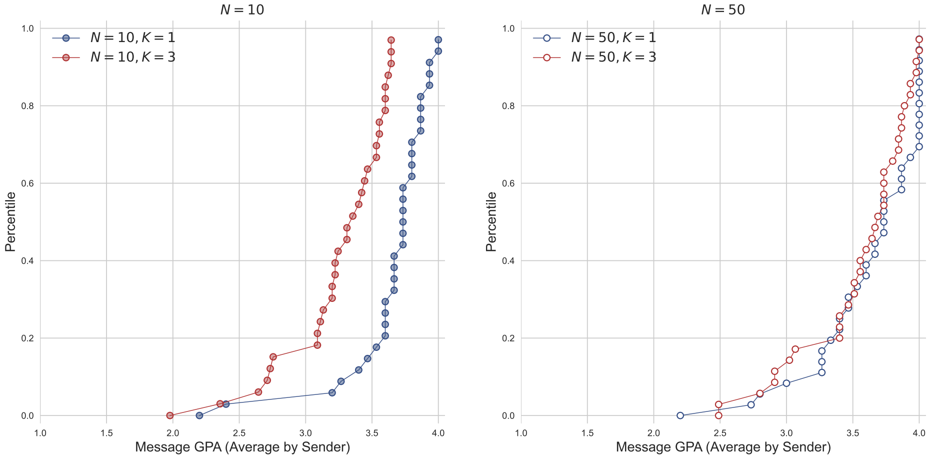


Which Evidence is Disclosed?





Which Evidence is Disclosed?



Which Evidence is Disclosed?

Robustness:

- Theoretical predictions Appendix
- Average treatment effects, statistical tests Appendix
- Raw data Appendix

Summary of Result 1

Result 1. Senders selectively disclose the available evidence in ways that are consistent with the key qualitative predictions of the theory

Note that:

- FOSD rankings are a demanding test for the theory
- Contrasting signs reduce scope for alternative explanations

result 2

(informativeness)

How Much Information is Transmitted?

Result 1 documents that senders engage in selective disclosure:

Question 2: What are the consequences of this selection on how much information is transmitted?

How Much Information is Transmitted?

Result 1 documents that senders engage in selective disclosure:

Question 2: What are the consequences of this selection on how much information is transmitted?

Theory offers rich and contrasting predictions on how informativeness should change in N and K

How Much Information is Transmitted?

Result 1 documents that senders engage in selective disclosure:

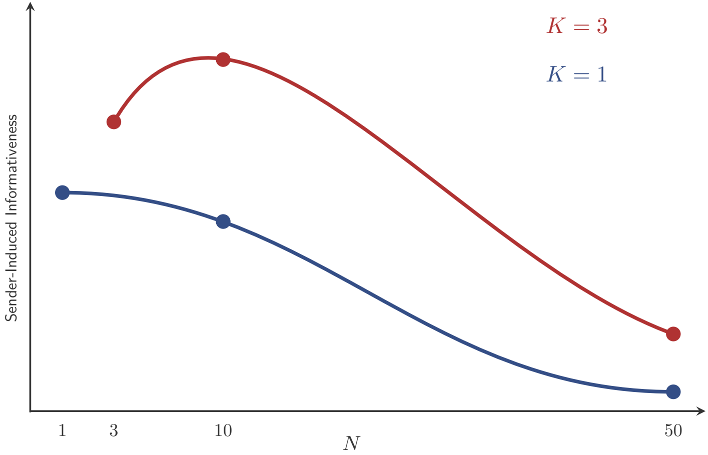
Question 2: What are the consequences of this selection on how much information is transmitted?

Theory offers rich and contrasting predictions on how informativeness should change in N and K

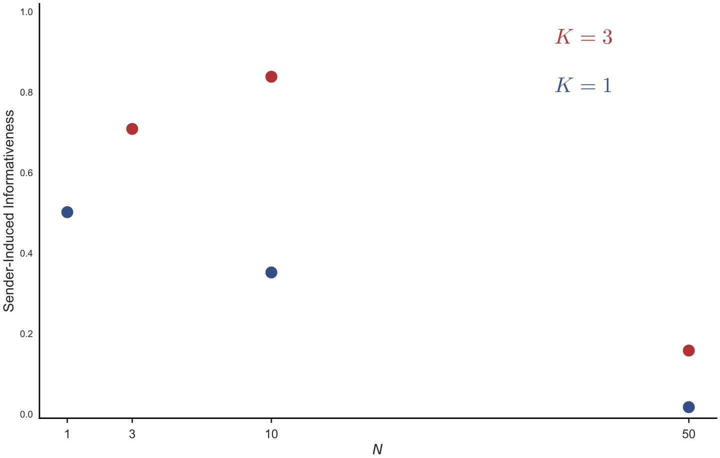
We measure **informativeness** as the correlation between the state θ and the guess a induced by the sender' strategy

Appendix

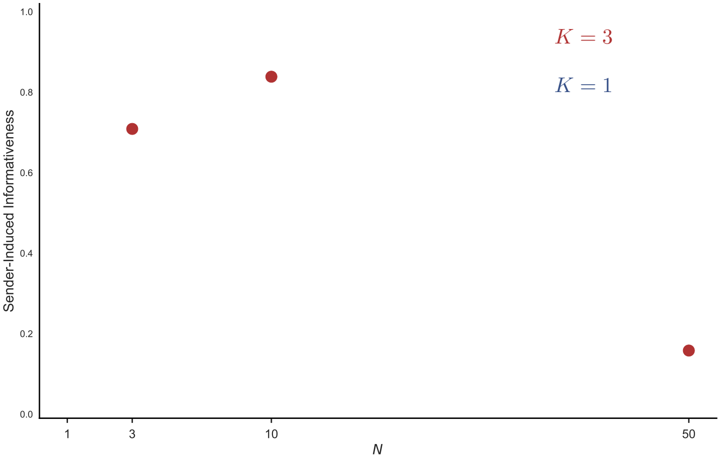
How Much Information is Transmitted?



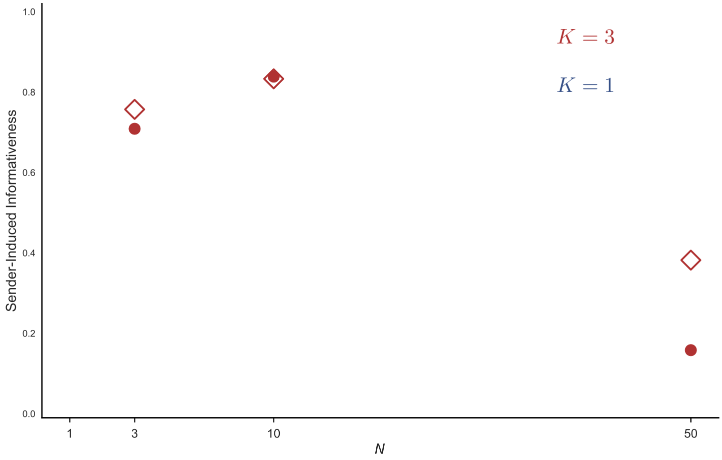
How Much Information is Transmitted?



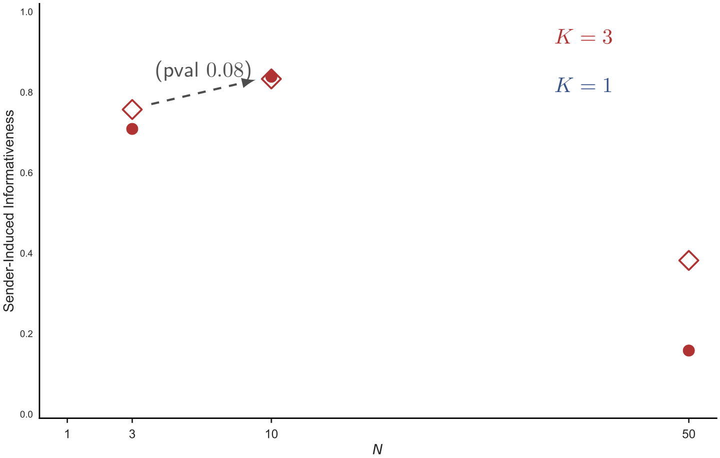
How Much Information is Transmitted?



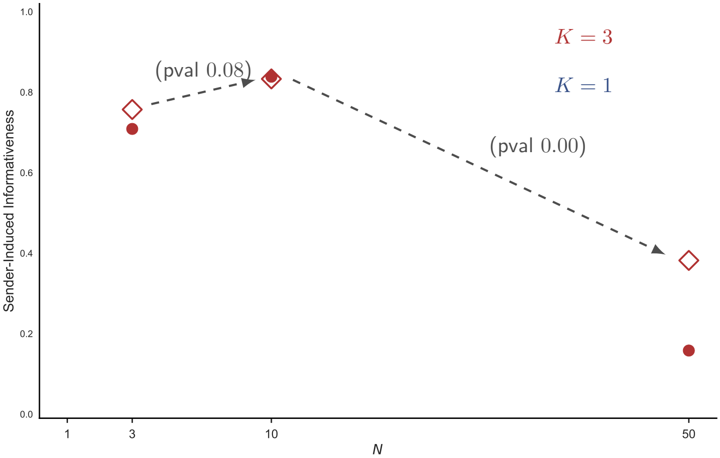
How Much Information is Transmitted?



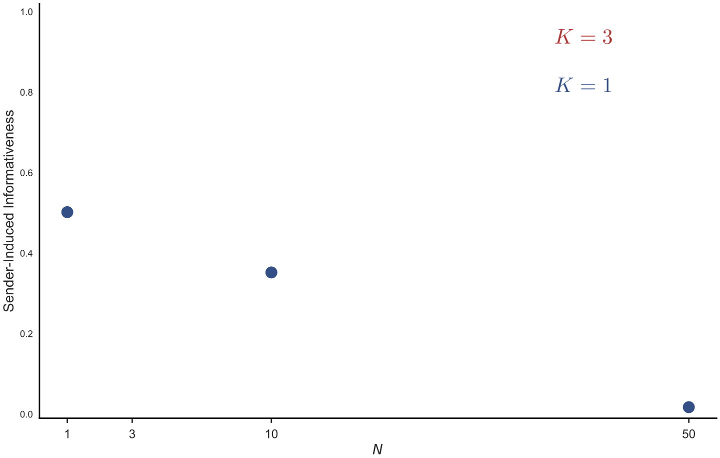
How Much Information is Transmitted?



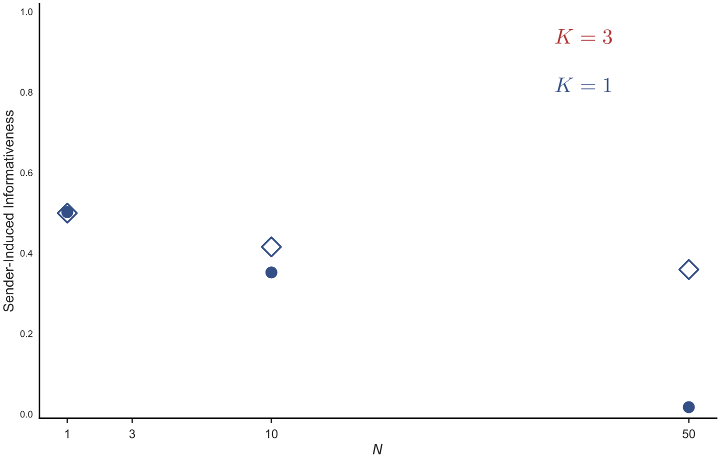
How Much Information is Transmitted?



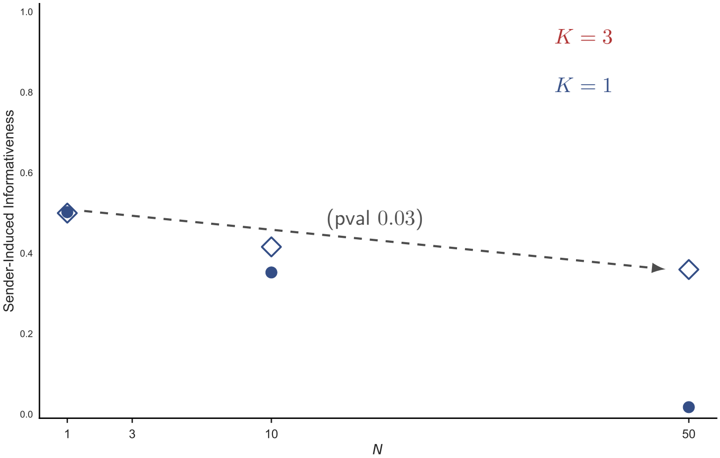
How Much Information is Transmitted?



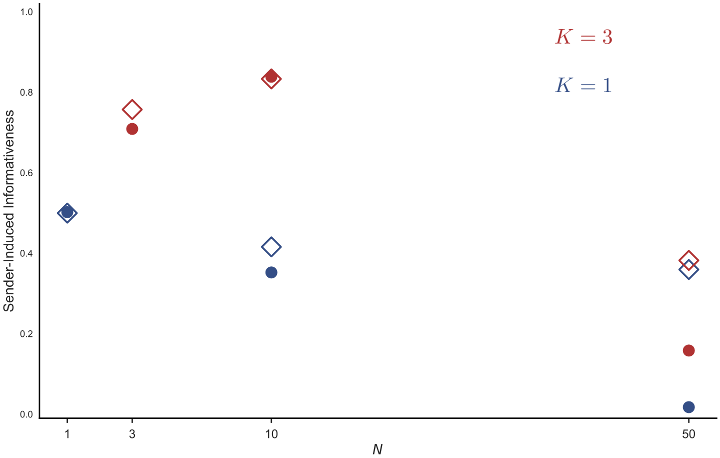
How Much Information is Transmitted?



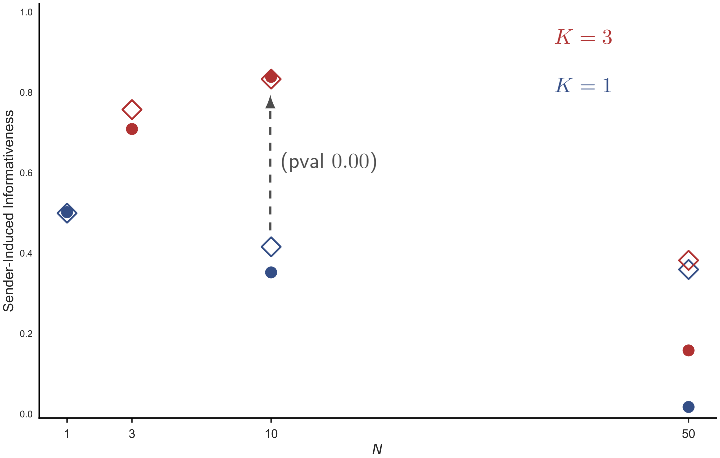
How Much Information is Transmitted?



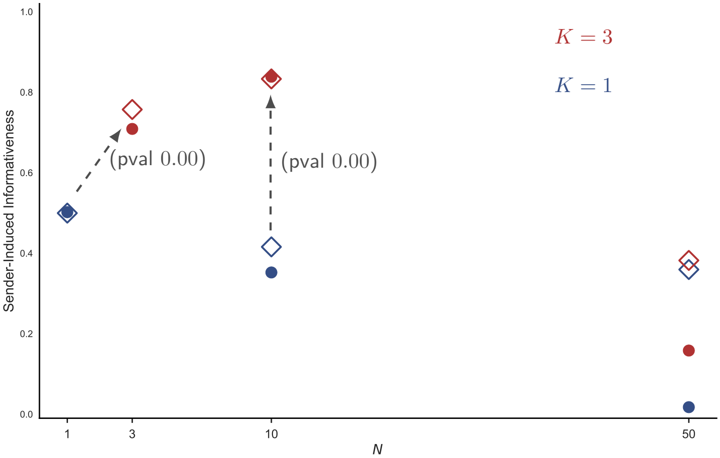
How Much Information is Transmitted?



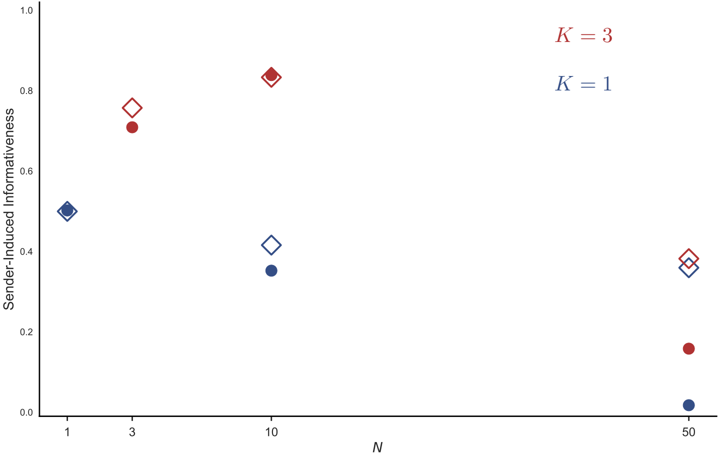
How Much Information is Transmitted?



How Much Information is Transmitted?



How Much Information is Transmitted?



Summary of Result 2

Array of contrasting predictions sets high bar for the theory

Result 2. Informativeness changes in ways that is consistent with the key qualitative predictions of the theory

Overall, as a force in communication, selective disclosure seems behaviorally descriptive

Summary of Result 2

Array of contrasting predictions sets high bar for the theory

Result 2. Informativeness changes in ways that is consistent with the key qualitative predictions of the theory

Overall, as a force in communication, selective disclosure seems behaviorally descriptive

Yet, our results also reveal substantial **quantitative** deviations

- Senders transmit (weakly) more information than predicted. That is, they **overcommunicate**

result 3

(overcommunication)

Explaining Overcommunication

This finding is at odds with existing experimental literature on disclosure

e.g., Forsythe, Isaac and Palfrey ('89, Rand), Jin and Leslie ('03, QJE), Jin, Luca, and Martin ('22, AEJ: Micro), Lizzeri, Frechette, Perego ('22, Ecma)

Explaining Overcommunication

This finding is at odds with existing experimental literature on disclosure

e.g., Forsythe, Isaac and Palfrey ('89, Rand), Jin and Leslie ('03, QJE), Jin, Luca, and Martin ('22, AEJ: Micro), Lizzeri, Frechette, Perego ('22, Ecma)

- These papers consistently find that senders **undercommunicate**
- Failure of the "unraveling principle" $\rightsquigarrow$ Senders fail to disclose evidence when it is sufficiently unfavorable
- They offer empirical support to policies that mandate disclosure in the marketplace

Explaining Overcommunication

This finding is at odds with existing experimental literature on disclosure

e.g., Forsythe, Isaac and Palfrey ('89, Rand), Jin and Leslie ('03, QJE), Jin, Luca, and Martin ('22, AEJ: Micro), Lizzeri, Frechette, Perego ('22, Ecma)

- These papers consistently find that senders **undercommunicate**
- Failure of the "unraveling principle" $\rightsquigarrow$ Senders fail to disclose evidence when it is sufficiently unfavorable
- They offer empirical support to policies that mandate disclosure in the marketplace

Question 3. Then why do we observe **overcommunication**?

Explaining Overcommunication

In our setting, full unraveling is not an equilibrium:

Why?

- Informativeness is always predicted to be interior $\mathcal{I} \in (0, 1)$
- In contrast, literature largely focused on an extreme prediction: $\mathcal{I} = 1$

Explaining Overcommunication

In our setting, full unraveling is not an equilibrium:

Why?

- Informativeness is always predicted to be interior $\mathcal{I} \in (0, 1)$
- In contrast, literature largely focused on an extreme prediction: $\mathcal{I} = 1$

This is a novel and essential feature of our approach:

- It enables nontrivial comparative statistics, which are instrumental for testing the theory
- It allows theoretical predictions to fail from both directions: **over** and **under** communication)

Explaining Overcommunication

Why?

In our setting, full unraveling is not an equilibrium:

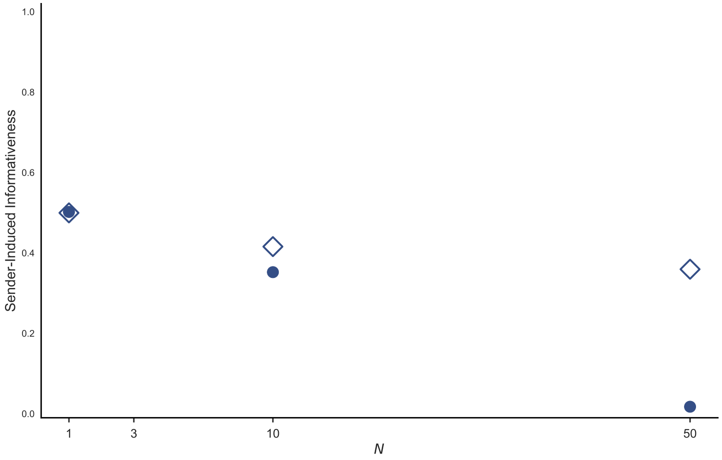
- Informativeness is always predicted to be interior $\mathcal{I} \in (0, 1)$
- In contrast, literature largely focused on an extreme prediction: $\mathcal{I} = 1$

This is a novel and essential feature of our approach:

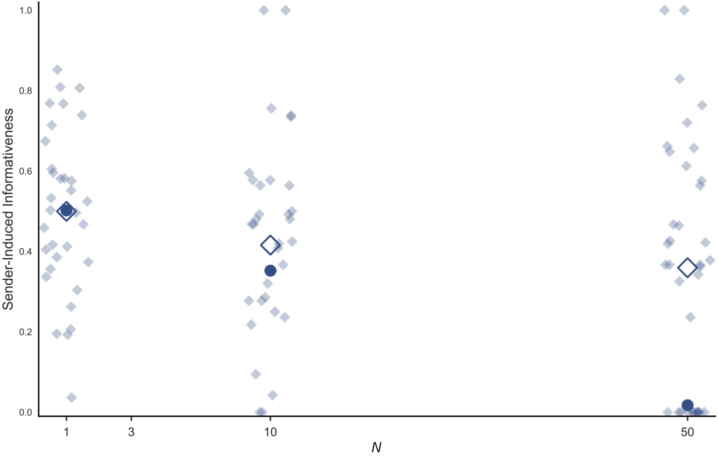
- It enables nontrivial comparative statistics, which are instrumental for testing the theory
- It allows theoretical predictions to fail from both directions: **over** and **under** communication)

Our findings suggest that **undercommunication** may not be a robust behavioral feature in disclosure games

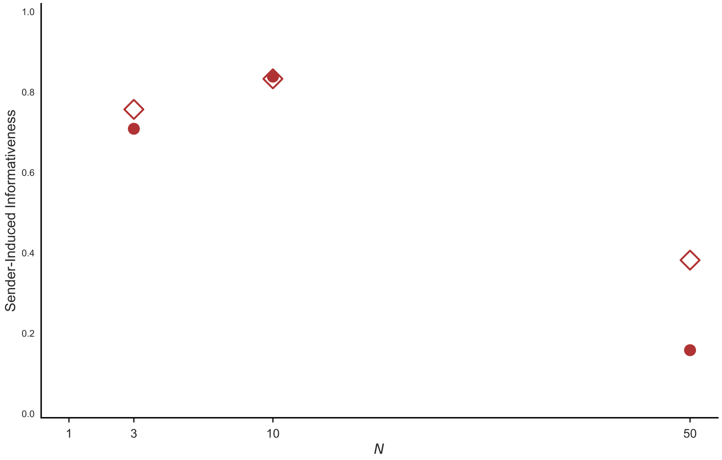
Explaining Overcommunication



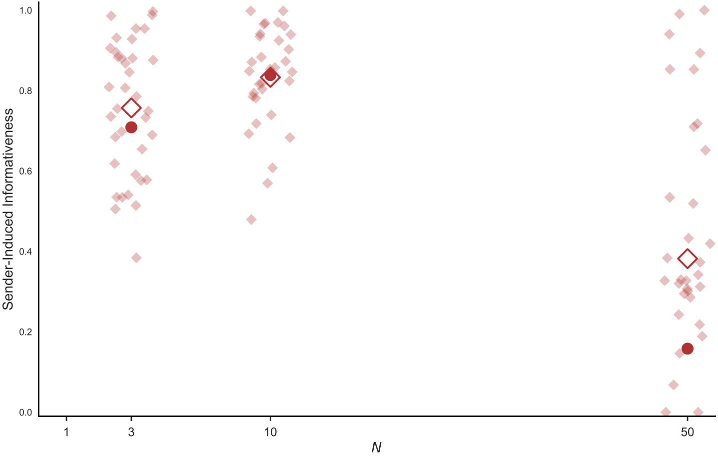
Explaining Overcommunication



Explaining Overcommunication



Explaining Overcommunication



Explaining Overcommunication

More concretely, how does overcommunication come about?

Explaining Overcommunication

More concretely, how does overcommunication come about?

Conditional on evidence available, sender's behavior should not depend on θ

– Equilibrium strategy: disclose K-best signals, regardless of θ

In contrast, we find that some senders adopt state-dependent strategies

Explaining Overcommunication

More concretely, how does overcommunication come about?

Conditional on evidence available, sender's behavior should not depend on θ

- Equilibrium strategy: disclose K-best signals, regardless of θ

In contrast, we find that some senders adopt state-dependent strategies

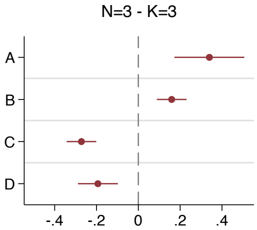
To illustrate, we estimate an OLS regression model:

details

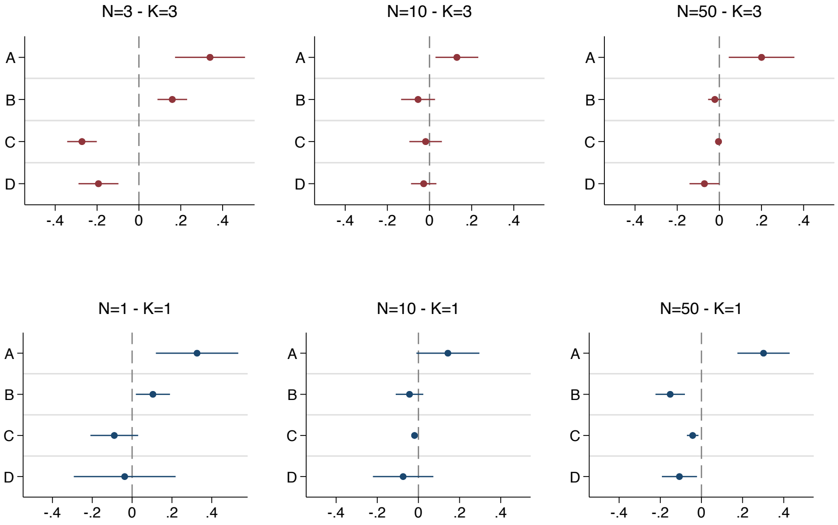
$$\text{Prob}(s \text{ is disclosed}) = \alpha + \beta_s \cdot \theta + \gamma \cdot X + \varepsilon$$

where X is a set of regressors that controls for senders' available evidence

Explaining Overcommunication



Explaining Overcommunication



Summary of Result 3

We find consistent patterns across all treatments:

- When the state is low (relative to when is high), senders under-disclose good evidence and over-disclose bad evidence

Also: effects on GPA

Summary of Result 3

We find consistent patterns across all treatments:

- When the state is low (relative to when is high), senders under-disclose good evidence and over-disclose bad evidence Also: effects on GPA

Result 3. Senders exhibit a form of **deception aversion**

State-dependent behavior generates overcommunication

Discussion:

- Senders can't lie in our setting, yet some avoid being deceptive Sobel, '23, *JPE*
- Never a best response to observed receivers behavior appendix

result 4

(receivers' beliefs)

Do Receivers Account for Selection?

The previous results have established that (modulo overcommunication) the evidence receivers see is **endogenously selected**

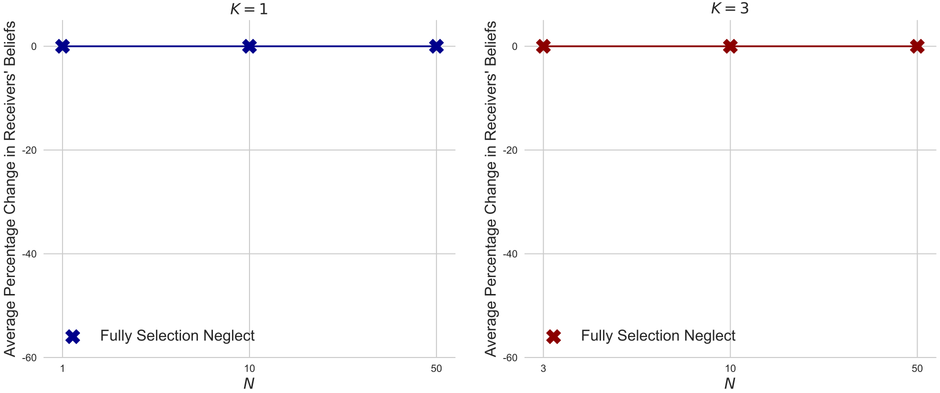
Question 4. To what extent do receivers account this selection in their responses?

To test this, we exploit the following prediction of the theory:

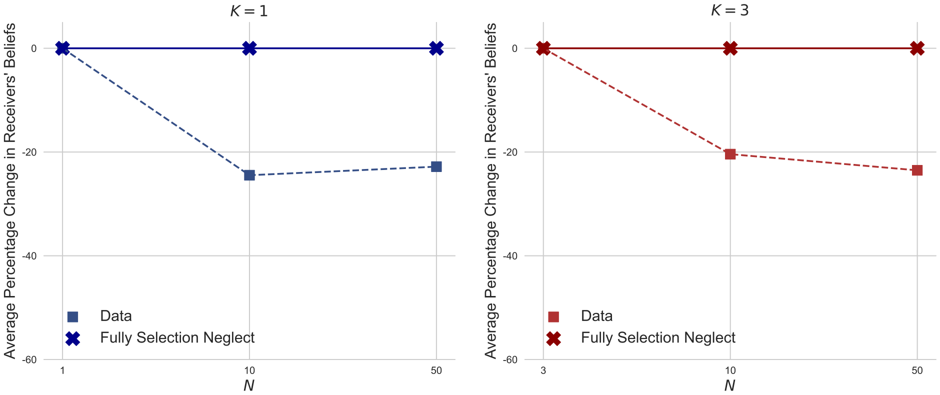
- Given any message, as N increases, receivers' beliefs about the state being high should decrease

We report the percentage change in receiver' beliefs averaged out across all messages and receivers

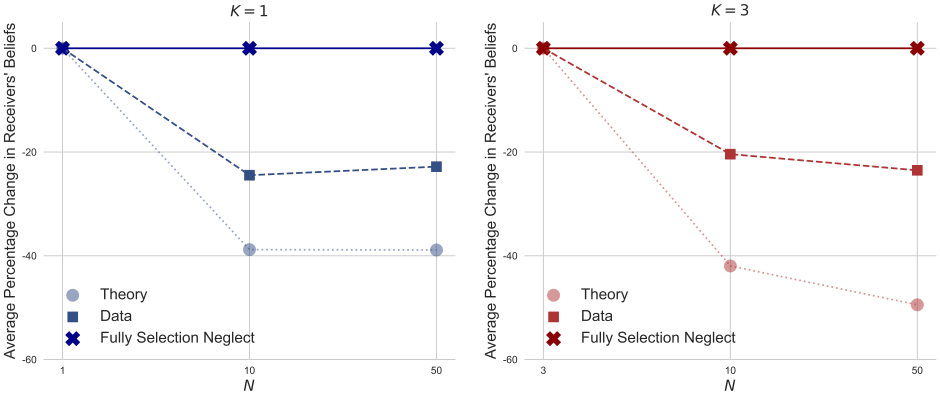
Do Receivers Account for Selection?



Do Receivers Account for Selection?



Do Receivers Account for Selection?



Summary of Result 4

Result 4. On average, receivers are increasingly skeptical of the evidence they see as it becomes more selected, as predicted by the theory

Summary of Result 4

Result 4. On average, receivers are increasingly skeptical of the evidence they see as it becomes more selected, as predicted by the theory

Yet, quantitatively, they fail to fully account for selection

First evidence of *selection neglect* in **communication**, a setting where selection arises endogenously, as an equilibrium outcome

Recent literature has documented selection neglect in non-strategic settings, where selection is exogenous

Esponda, Vespa ('18, QE), Enke ('20, QJE), Barron, Huck, Jehiel (2023, AEJ:Micro)

Summary of Result 4

Result 4. On average, receivers are increasingly skeptical of the evidence they see as it becomes more selected, as predicted by the theory

Yet, quantitatively, they fail to fully account for selection

First evidence of *selection neglect* in **communication**, a setting where selection arises endogenously, as an equilibrium outcome

Recent literature has documented selection neglect in non-strategic settings, where selection is exogenous

Esponda, Vespa ('18, QE), Enke ('20, QJE), Barron, Huck, Jehiel (2023, AEJ:Micro)

In progress: Identify receivers' "types"

result 5

(receivers' accuracy)

Do Receivers Account for Selection?

Question 5. The previous result focused on beliefs, but how costly are these mistakes in terms of payoffs?

Do Receivers Account for Selection?

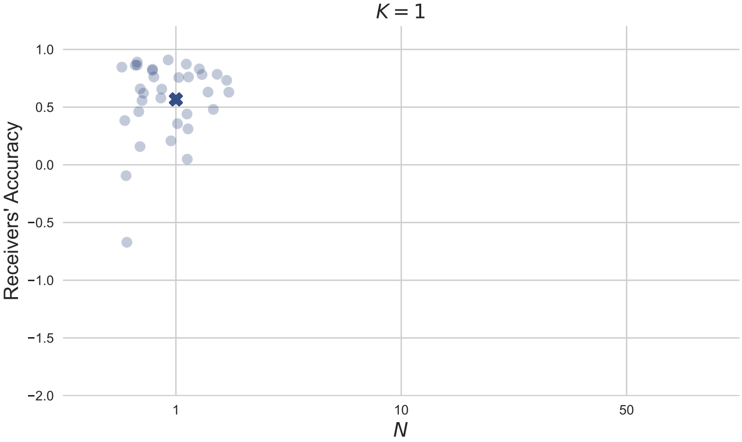
Question 5. The previous result focused on beliefs, but how costly are these mistakes in terms of payoffs?

Define **accuracy** as the percentage of the payoff the receiver obtains relative to what a Bayesian would have obtained

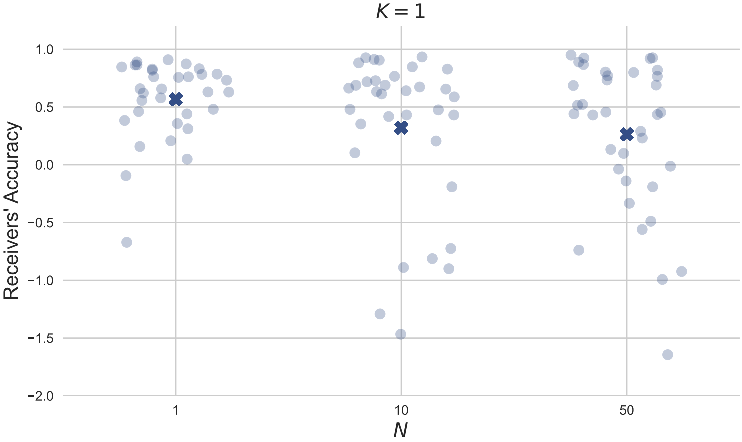
We normalize accuracy by the accuracy a receiver would obtain if she behaved at random

The theory predicts accuracy is equal to 1 in all treatments

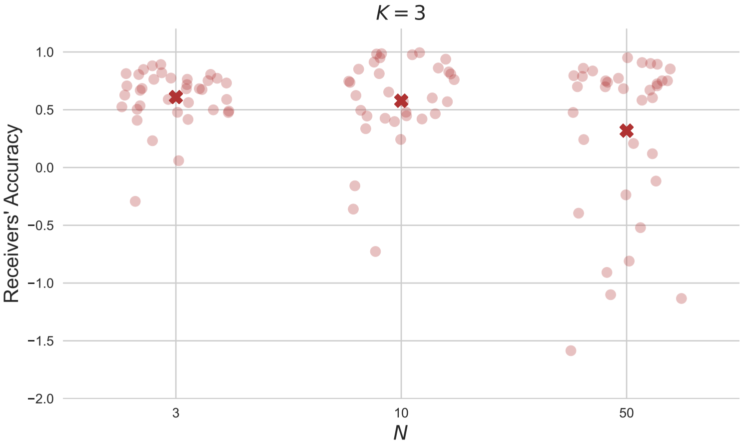
Receivers’ Accuracy



Receivers' Accuracy



Receivers' Accuracy



Summary of Result 5

Result 5. Receivers become less accurate as N increases

Especially puzzling given that the receiver' problem becomes "easier" as N increases

In progress: Is it selection-neglect the driver of loss in payoff/accuracy?

conclusion

Conclusion

A comprehensive experimental study on **selective disclosure**

We exploit comparative statics to inform a novel experimental design

Our data corroborates the key qualitative predictions of the theory

- Validation of **selective disclosure** as a force in communication that is behaviorally descriptive

We detect two main *quantitative* deviations from the theory:

- A form of **deception aversion** in senders leads to **overcommunicate**
- We find evidence of **selection neglect** in a strategic setting

The Selective Disclosure of Evidence
An Experiment

Agata Farina
NYU

Guillaume Fréchette
NYU

Alessandro Lizzeri
Princeton

Jacopo Perego
Columbia

thank you

Appendix

How Much Information is Transmitted?

Question: How informative are senders' strategies?

How Much Information is Transmitted?

Question: How informative are senders' strategies?

We measure informativeness $\mathcal{I}(K, N)$ as correlation between:

- The state θ
- The receiver's guess a

How Much Information is Transmitted?

Question: How informative are senders' strategies?

We measure informativeness $\mathcal{I}(K, N)$ as correlation between:

- The state θ
- The receiver's guess a

$$\text{state } \theta \xrightarrow{\textit{Sender provides information}} \text{message } m \xrightarrow{\textit{Receiver processes information}} \text{guess } a$$

How Much Information is Transmitted?

Question: How informative are senders' strategies?

We measure informativeness $\mathcal{I}(K, N)$ as correlation between:

- The state θ
- The ~~receiver's guess a~~ guess of a hypothetical Bayesian receiver, a^B

$$\text{state } \theta \xrightarrow[\text{information}]{\substack{\textit{Sender} \\ \textit{provides}}} \text{message } m \xrightarrow[\text{information}]{\substack{\textit{Receiver} \\ \textit{processes}}} \text{guess } a$$

How Much Information is Transmitted?

Question: How informative are senders' strategies?

We measure informativeness $\mathcal{I}(K, N)$ as correlation between:

- The state θ
- The ~~receiver's guess a~~ guess of a hypothetical Bayesian receiver, a^B

$$\text{state } \theta \;\overset{\textit{Sender provides information}}{\rightsquigarrow}\; \text{message } m \;\xrightarrow{\textit{Receiver processes information}}\; \text{guess } a^B$$

How Much Information is Transmitted?

Question: How informative are senders' strategies?

We measure informativeness $\mathcal{I}(K, N)$ as correlation between:

- The state θ
- The ~~receiver's guess a~~ guess of a hypothetical Bayesian receiver, a^B

$$\text{state } \theta \xrightarrow{\textit{Sender provides information}} \text{message } m \xrightarrow{\textit{Receiver processes information}} \text{guess } a^B$$

We refer to $\mathrm{Corr}(\theta, a^B)$ as the **sender-induced correlation**

Back

Richness

Theoretically, our setting differs from typical disclosure model because evidence structure is not "rich"

Evidence structure is rich if, for all θ, sender can send message that verifiably reveals θ

In our setting, evidence is noisy, and K and N are finite. No message can verifiably reveal θ

Richness drives unraveling results (Okuno-Fujiwara et al., 1990, *Restud*)

Strong assumption in many practical settings

Back

OLS Regression Model

We restrict attention to the observations in which s is available:

$$\text{Prob}(s \text{ is disclosed}) = \alpha + \beta_s \cdot \theta + \sum_{s \in S} \gamma_s \cdot \min\{k, \text{av}_s\} + \varepsilon$$

Senders' random effects; Standard errror clustered at the session level

Regressors $\{\min\{k, \text{av}_s\}\}_{s \in S}$ controls for senders available evidence

Results robust to controlling for set of available messages

Back

Some Notation: Strategies and Beliefs

Denote $\mathcal{M}$ the space of all messages

Sender's Strategy

pure and θ-independent

- $\sigma : \Omega^N \to \mathcal{M}$ s.t. $\sigma(\bar{s}) \in M(\bar{s})$, for all $\bar{s}$

where $M(\bar{s})$ is the space of available messages given $\bar{s}$

Receiver's Beliefs and Strategy

- $\mu : \mathcal{M} \to \Delta(S^N)$
- $a : \mathcal{M} \to \Delta(A)$

Given μ, receiver's optimal strategy given by

$$a(m) = \mathbb{E}(\theta|m) = \sum_{\bar{s}} \mu(\bar{s}|m)\mathbb{E}(\theta|\bar{s}) \qquad \forall\, m$$

Sequential Equilibrium

A **Sequential Equilibrium** is a pair (σ^*, μ^*) s.t.

1. For all $\bar{s} \in \Omega^N$, $\sigma^*(\bar{s}) \in M(\bar{s})$ and

$$\sum_{\bar{s}'} \mu^*(\bar{s}'|\sigma^*(\bar{s}))\mathbb{E}(\theta|\bar{s}') \geq \sum_{\bar{s}'} \mu^*(\bar{s}'|m')\mathbb{E}(\theta|\bar{s}') \qquad m' \in M(\bar{s})$$

2. For all m, supp $\mu^*(\cdot|m) \subseteq C(m) = \{\bar{s} \in S^N : m \in M(\bar{s})\}$. In particular, if $m \in \sigma^*(S^N)$,

$$\mu^*(\bar{s}|m) = q(\bar{s}|{\sigma^\star}^{-1}(m)) \quad \forall\, \bar{s}$$

   where $q(\bar{s}) = \sum_\theta p(\theta) f(\bar{s}|\theta)$

Multiplicity and Neologism Proofness

Unlike classic disclosure games, the sequential equilibrium outcome is **not unique** when $K < N$.

- Off-path beliefs can support other equilibrium outcome.
- Refinements for signalling games (e.g., Cho-Kreps '87, Banks-Sobel '87) have no force here.
- Refinements for cheap talk games: Farrel (1993)'s **Neologism Proofness**.

Multiplicity and Neologism Proofness

$\Theta = \{0, 1\}$ and $p(1) = \frac{1}{2}$. $N = 2$ and $K = 1$.

$\Omega = \{A, B\}$, $f(A|\theta_H) = 1$ and $f(A|\theta_L) = \frac{1}{2}$.

θ	$\bar{s}$	$M(\bar{s})$	$\sigma^*(\bar{s})$
1	(A, A)	$\{\varnothing, A\}$	A
0	(A, B)	$\{\varnothing, A, B\}$	A
	(B, B)	$\{\varnothing, B\}$	B

$$\mathbb{E}[\theta|m = A] = \frac{4}{7} \text{ and } \mathbb{E}[\theta|m = B] = \mathbb{E}[\theta|m = \varnothing] = 0 \implies$$

No incentive to deviate

Multiplicity and Neologism Proofness

$\Theta = \{0, 1\}$ and $p(1) = \frac{1}{2}$. $N = 2$ and $K = 1$.

$\Omega = \{A, B\}$, $f(A|\theta_H) = 1$ and $f(A|\theta_L) = \frac{1}{2}$.

θ	$\bar{s}$	$M(\bar{s})$	$\sigma^*(\bar{s})$
1	(A, A)	$\{\varnothing, A\}$	$\varnothing$
0	(A, B)	$\{\varnothing, A, B\}$	$\varnothing$
	(B, B)	$\{\varnothing, B\}$	$\varnothing$

$$\mathbb{E}[\theta|m = \varnothing] = \tfrac{1}{2} \text{ and } \mathbb{E}[\theta|m = A] = \mathbb{E}[\theta|m = B] = 0 \implies$$

No incentive to deviate

Neologism Proofness

A **neologism** is a a pair (m, C), $C \subseteq \{\bar{s} \in S^N : m \in M(\bar{s})\}$

Literal meaning of (m, C) $\rightsquigarrow$ "My type $\bar{s}$ belongs to C and I can prove it by sending message m"

A neologism (m, C) is **credible** relative to equilibrium (σ^*, μ^*) if

1. $\sum_{\bar{s}'} q(\bar{s}'|C)\mathbb{E}(\theta|\bar{s}') > \sum_{\bar{s}'} \mu^*(\bar{\omega}'|\sigma^*(\bar{s}))\mathbb{E}(\theta|\bar{s}')$ for all $\bar{s} \in C$
2. $\sum_{\bar{s}'} q(\bar{s}'|C)\mathbb{E}(\theta|\bar{s}') \leq \sum_{\bar{s}'} \mu^*(\bar{s}'|\sigma^*(\bar{s}))\mathbb{E}(\theta|\bar{s}')$ for all $\bar{s} \notin C$

The equilibrium is **Neologism Proof** if no neologism is credible.

Neologism Proofness

θ	$\bar{s}$	$M(\bar{s})$	$\sigma^*(\bar{s})$
1	(A, A)	$\{\varnothing, A\}$	$\varnothing$
0	(A, B)	$\{\varnothing, A, B\}$	$\varnothing$
	(B, B)	$\{\varnothing, B\}$	$\varnothing$

$$m = A \text{ and } C = \{(A, A), (A, B)\} \implies$$

$$\mathbb{E}[\theta | m = A] = \tfrac{4}{7} > \mathbb{E}[\theta | m = \varnothing] = \tfrac{1}{2}$$

Since neologism (m, C) is credible, this PBE is not neologism proof equilibrium

Back

Neologism Proofness

Proposition

The equilibrium with maximal selective disclosure is Neologism Proof.

Neologism Proofness delivers outcome uniqueness

An equilibrium (σ, μ) induces an outcome $x : S^N \to A$,

$$x(\bar{s}) = \sum_{\bar{s}'} \mu(\bar{s}' | \sigma(\bar{s})) \mathbb{E}(\theta | \bar{s}') \qquad \forall \, \bar{s}.$$

Belief Elicitation

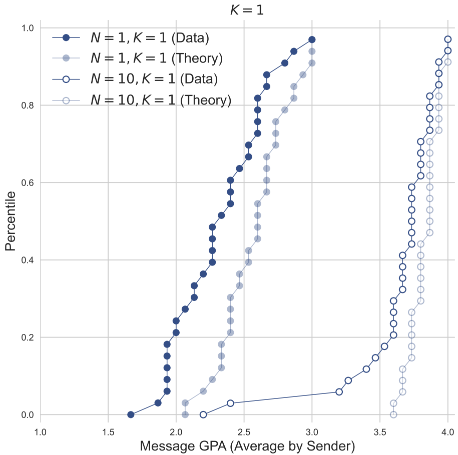
Since Θ is binary and $u(a, \theta) = -(a - \theta)^2$, the receiver's task is equivalent to eliciting her beliefs via a quadratic scoring rule (QSR)

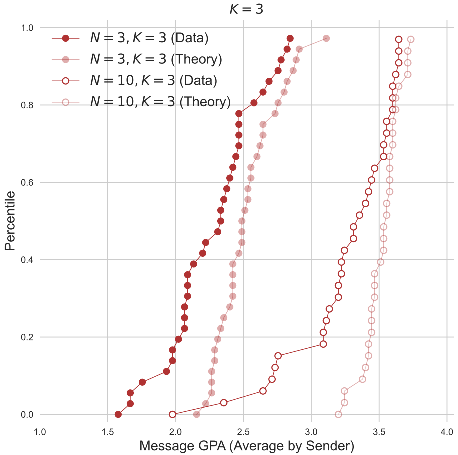
A large literature on belief elicitation has shown that QSR can be biased when subjects are not risk-neutral

To avoid this issue, we implement a binarized scoring rule *a la* Hossain and Okui (2013), which is robust to various risk preferences

Back

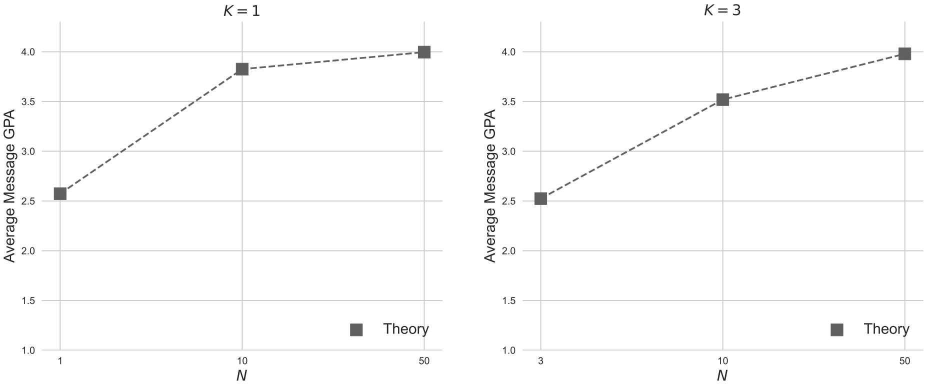
Which Evidence is Disclosed?



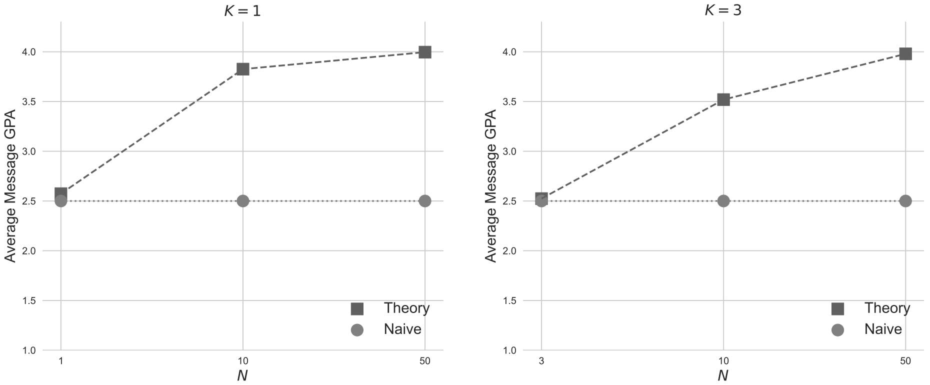


Back

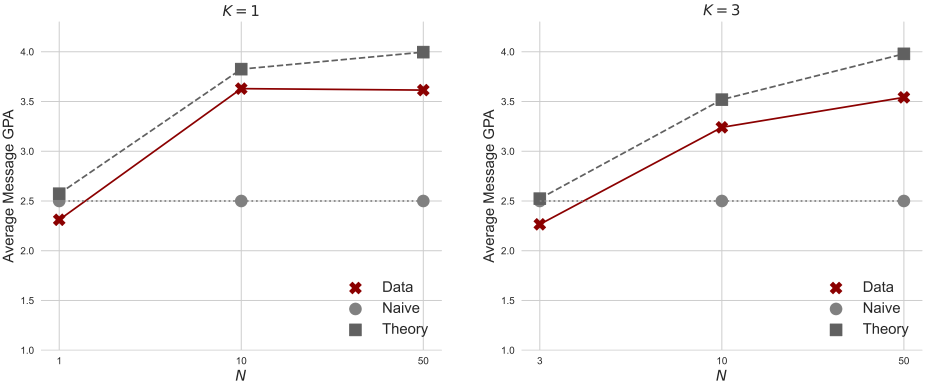
Which Evidence is Disclosed?



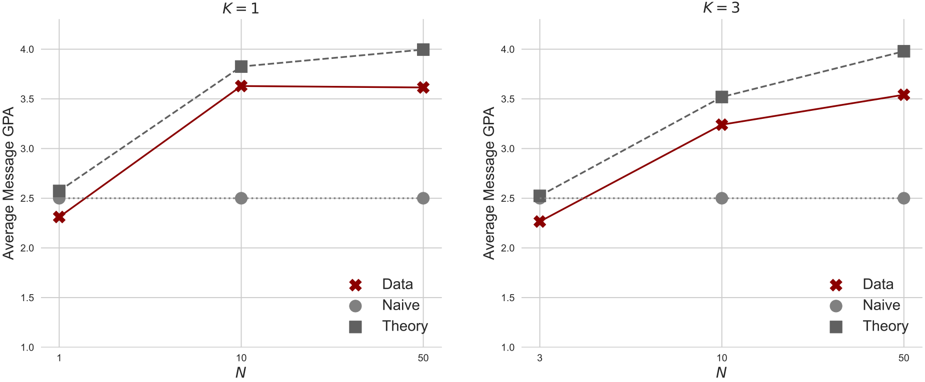
Which Evidence is Disclosed?



Which Evidence is Disclosed?

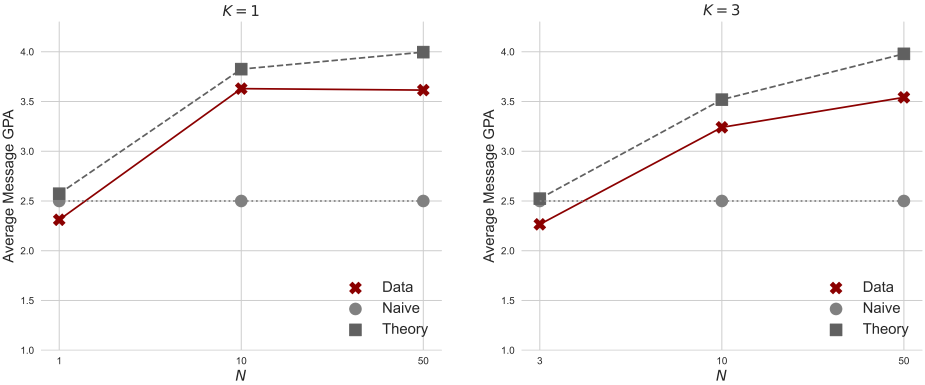


Which Evidence is Disclosed?



Qualitative predictions are corroborated by the data (pvals ~ 0.01)

Which Evidence is Disclosed?

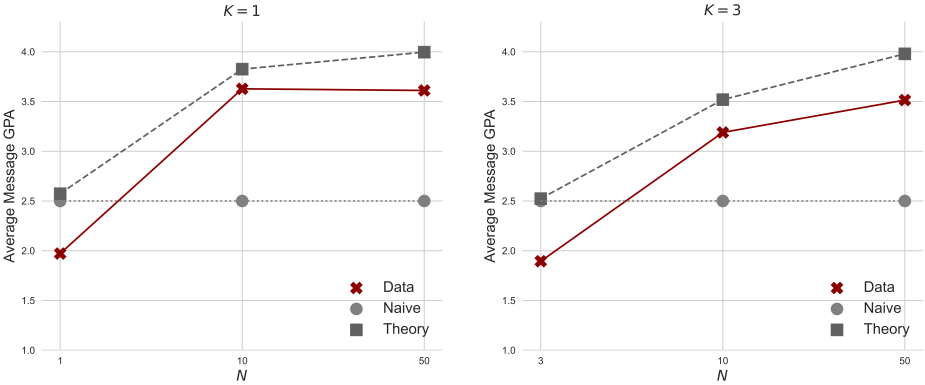


Qualitative predictions are corroborated by the data (pvals ~ 0.01)

Quantitatively, senders select less than theory predicts (pvals < 0.05)

Back

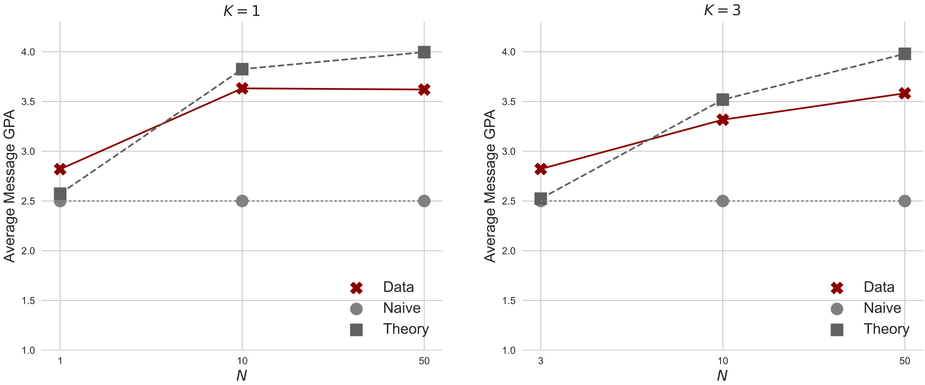
Alternative GPA: Empty = 0



Qualitative predictions are corroborated by the data (pvals ~ 0.01)

Quantitatively, senders select less than theory predicts (pvals < 0.05)

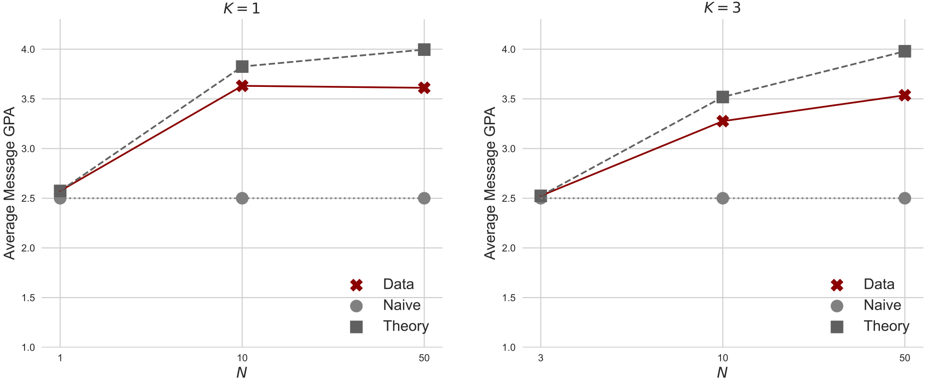
Alternative GPA: Empty = 2.5



Qualitative predictions are corroborated by the data (pvals ~ 0.01)

Quantitatively, senders select less than theory predicts (pvals < 0.05)

Alternative GPA: Empty = Avg Undisclosed

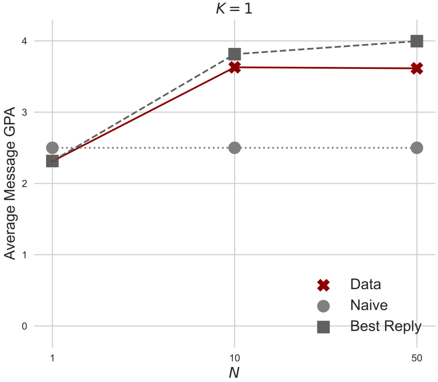


Qualitative predictions are corroborated by the data (pvals ~ 0.01)

Quantitatively, senders select less than theory predicts (pvals < 0.05)

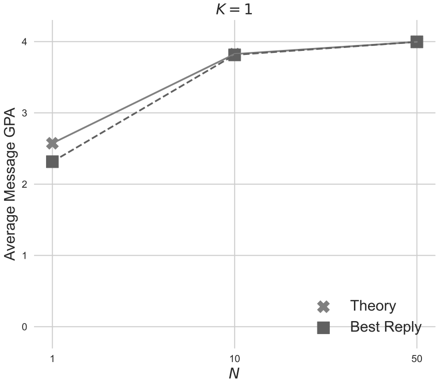
Which Evidence is Disclosed?

For $K = 1$ we can compare the observed message GPA with the one that would arise from an optimal empirical behavior of the sender: $\varnothing$ better than C and D



Quantitatively, GPA smaller than theory predicts for $N > K$ (pvals < 0.05)

Best Reply vs Theoretical Predictions



Quantitatively, best reply and theory are different for $N < 50$ (pvals < 0.01)

The behavior is the same 74% of the times for $N = 1$, 99% of the times for $N = 10$ and 100% of the times for $N = 50$

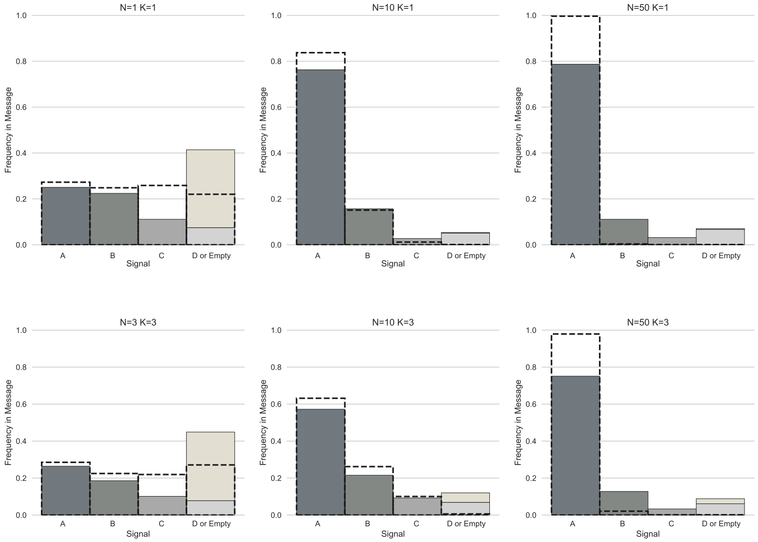
Empirical Best Response vs Equilibrium

Given observed receivers' behavior, senders' best response coincides with equilibirum strategy

- 74% of the times in treatment $(N = 1, K = 1)$
- 99% of the times in treatment $(N = 10, K = 1)$
- 100% of the times in treatment $(N = 50, K = 1)$

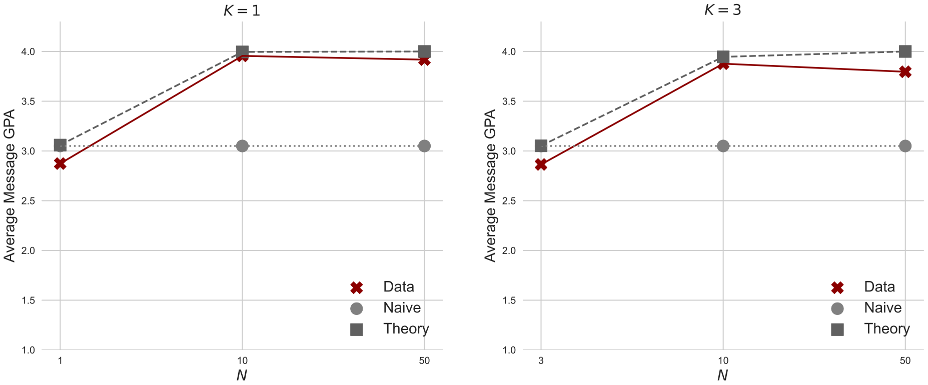
Back

Frequency of Signals in Sender's Message



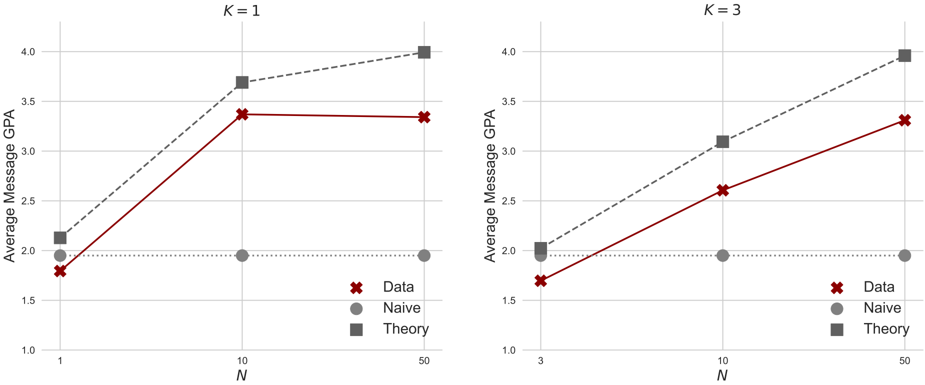
Back

Which Evidence is Disclosed? High Type



Quantitatively, senders select less than theory predicts (pvals < 0.1)

Which Evidence is Disclosed? Low Type



Quantitatively, senders select less than theory predicts (pvals < 0.05)

Back

Heterogeneity: Senders

Challenge

- Large number of urn / balls / message combinations
- Specific behavior of interest varies across treatments
  - Number of balls sent ($K = 1$ vs $K = 3$)
  - Balls sent vs balls available ($N = K$ vs $N > K$)

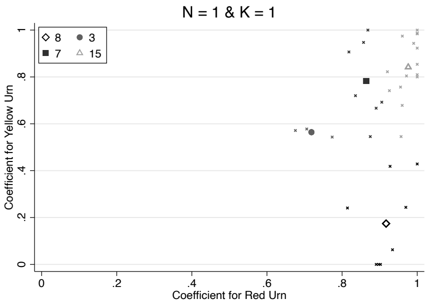
$\rightarrow$ Precludes a unified approach using those variables

Heterogeneity: Senders

Solution

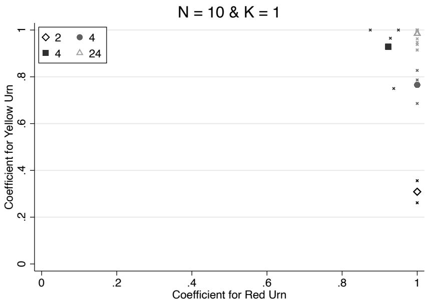
- Transform balls and messages to numbers ($B^{\#}$ and $M^{\#}$)
- Regress $M^{\#}$ on $B^{\#}$|yellow urn and $B^{\#}$|red urn
- Cluster the coefficient estimates
- Describe behavior along key dimensions of interest

Heterogeneity: Senders



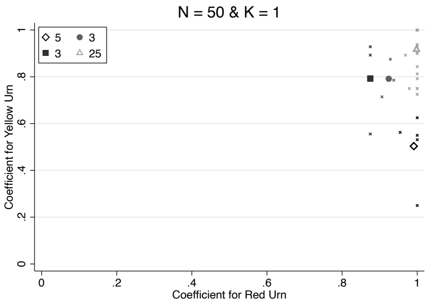
Cluster	Obs (33)	Urn	K	A	D
Triangle	15				
		Red	0.91	1	0.38
		Yellow	0.64	1	0.27
Square	7				
		Red	0.73	1	0.25
		Yellow	0.51	1	0.21
Circle	3				
		Red	0.5	0.92	n/a
		Yellow	0.54	0.67	0.49
Diamond	8				
		Red	0.71	1	0.20
		Yellow	0.30	0	0.46

Heterogeneity: Senders



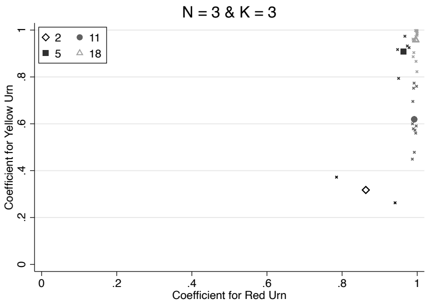
Cluster	Obs (34)	Urn	K	A	D
Triangle	24				
		Red	1	1	0
		Yellow	1	0.97	0.02
Square	4				
		Red	1	0.81	0.08
		Yellow	1	0.88	0.07
Circle	4				
		Red	1	1	0
		Yellow	1	0.46	0.14
Diamond	2				
		Red	1	1	0
		Yellow	1	0	0.89

Heterogeneity: Senders



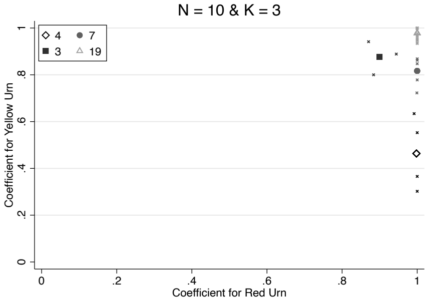
Cluster	Obs (36)	Urn	K	A	D
Triangle	25				
		Red	1	0.99	0
		Yellow	1	0.74	0.03
Square	3				
		Red	0.96	0.82	0.1
		Yellow	1	0.51	0.15
Circle	3				
		Red	1	0.78	0
		Yellow	1	0.63	0.18
Diamond	5				
		Red	1	0.96	0
		Yellow	0.95	0.26	0.46

Heterogeneity: Senders



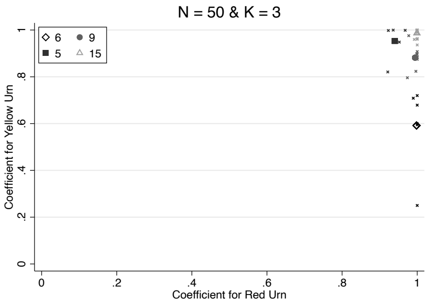
Cluster	Obs (36)	Urn	K	A	D
Triangle	18				
		Red	0.58	1	0.15
		Yellow	0.18	1	0.12
Square	5				
		Red	0.29	1	0
		Yellow	0.10	0.88	0.05
Circle	11				
		Red	0.26	1	0.06
		Yellow	0.15	0.23	0.60
Diamond	2				
		Red	0	1	0
		Yellow	0.06	0.25	0.50

Heterogeneity: Senders



Cluster	Obs (33)	Urn	K	A	D
Triangle	19				
		Red	0.99	0.99	0
		Yellow	0.88	0.96	0.01
Square	3				
		Red	1	0.46	0.17
		Yellow	1	0.43	0.04
Circle	7				
		Red	1	0.94	0
		Yellow	0.74	0.66	0.10
Diamond	4				
		Red	0.92	0.83	0
		Yellow	0.76	0.28	0.43

Heterogeneity: Senders



Cluster	Obs (35)	Urn	K	A	D
Triangle	15				
		Red	1	0.88	0
		Yellow	0.94	0.80	0
Square	5				
		Red	0.89	0.17	0
		Yellow	0.87	0.32	0
Circle	9				
		Red	0.97	0.70	0
		Yellow	0.94	0.31	0.04
Diamond	6				
		Red	1	0.86	0.03
		Yellow	0.95	0.31	0.41

Heterogeneity: Senders

Equilibrium type (56%)

- Most common
- $N > K$: Mostly report best balls independently of the state
- $N = K$: Disclose fewer than K balls

Deception Averse Type (17%)

- A's reported more often when the state is high
- D's reported more often when the state is low
- $N = K$: Disclose fewer than K balls

Others (27%)

- Similar to <u>equilibrium types</u> when the state is high
- Report A's less but do not report D's when the state is low
- Some low rates of A's when the state is high [confusion]

Heterogeneity: Receivers

Challenge

- Large number of messages
- Different messages across treatments
- Some messages have very few observations

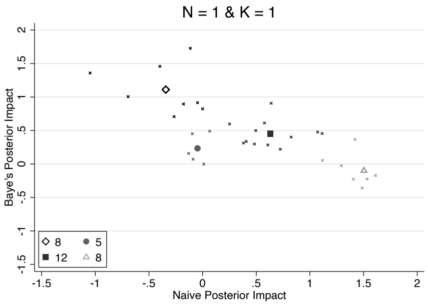
$\rightarrow$ Precludes a unified approach using that variable

Heterogeneity: Receivers

Solution

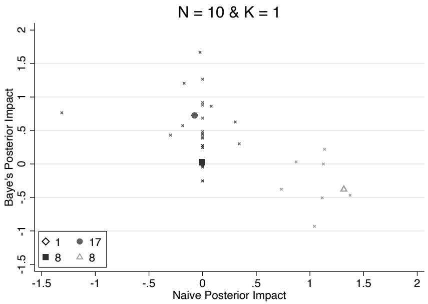
- Compute equilibrium update following each message
- Compute the update of someone who ignores selection: naive update
- Regress guesses on a constant (α) and the equilibrium and naive updates
- Cluster the coefficient estimates
- Describe behavior along key dimensions of interest

Heterogeneity: Receivers



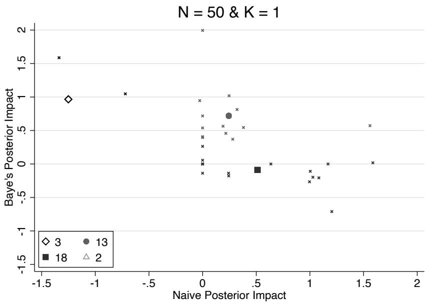
Cluster	Obs (33)	A	B	∅	C
Diamond	8				
$\alpha = 0.23$		0.87	0.67	0.23	0.47
Circle	5				
$\alpha = 0.39$		0.56	0.49	0.41	0.37
Square	12				
$\alpha = 0.02$		0.86	0.73	0.41	0.38
Triangle	8				
$\alpha = -0.23$		0.90	0.67	0.51	0.23

Heterogeneity: Receivers



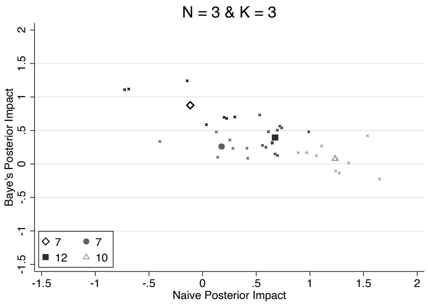
Cluster	Obs (34)	A	B	$\emptyset$	D
Diamond $\alpha = 4.20$	1	0.60*	0.23*	0.60*	n/a
Circle $\alpha = 0.28$	17	0.66	0.26	n/a	0.11
Square $\alpha = 0.56$	8	0.58	0.60	n/a	0.60
Triangle $\alpha = -0.23$	8	0.62	0.52	n/a	0.11

Heterogeneity: Receivers



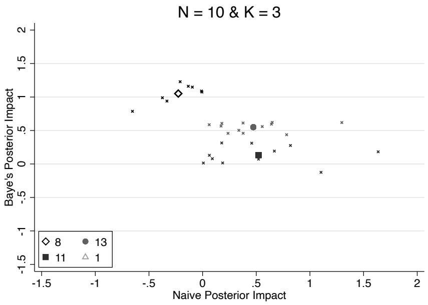
Cluster	Obs (36)	A	B	∅	D
Diamond	3				
$\alpha = 0.89$		0.35	0.17	0.21*	0.75
Circle	13				
$\alpha = 0.15$		0.71	0.29	0.46*	0.11
Square	18				
$\alpha = 0.26$		0.63	0.53	n/a	0.19
Triangle	2				
$\alpha = -1.15$		0.69	0.41	n/a	n/a

Heterogeneity: Receivers



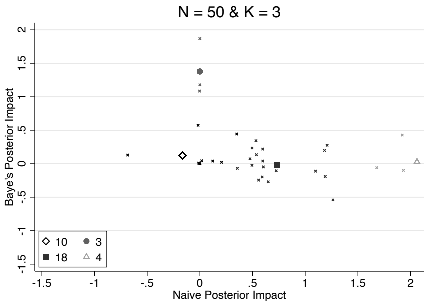
Cluster	Obs (36)	AAA	AAB	AA	AB
Diamond	7				
$\alpha = 0.20$		0.92*	0.86	0.86	0.62
Circle	7				
$\alpha = 0.30$		0.72	0.66	0.63	0.68
Square	12				
$\alpha = -0.04$		0.88	0.92	0.91	0.86
Triangle	10				
$\alpha = -0.24$		1	0.97	0.96	0.90

Heterogeneity: Receivers



Cluster	Obs (33)	AAA	AAB	AA	ABB
Diamond $\alpha = 0.19$	8	0.95	0.11	0.02	0.03
Circle $\alpha = -0.07$	13	0.89	0.70	0.24	0.26
Square $\alpha = 0.10$	11	0.74	0.70	n/a	0.61
Triangle $\alpha = -3.98$	1	1*	0.54*	n/a	0.02*

Heterogeneity: Receivers



Cluster	Obs (35)	AAA	AAB	AA	DDD
Diamond	10				
$\alpha = 0.64$		0.54	0.49	0.33	0.32
Circle	3				
$\alpha = 0.11$		0.84	0.01*	n/a	0.07
Square	18				
$\alpha = -0.04$		0.67	0.69	0.57	0.12
Triangle	4				
$\alpha = -1.16$		0.89	0.80	0.91*	n/a

Heterogeneity: Receivers

- Variation in updating strategies
  - Extent they account for selection
- Being closer to equilibrium $\not\to$ higher payoffs
- However, in many treatments, groups better at accounting for selection are among the highest
- With $N = 50$, few differences in payoffs

Summary

Senders

- The majority:
  - Select the better balls to send.
  - Behave similarly for both urns.
- Some convey more information by conditioning on the type.

$\rightarrow$ More information transmitted than predicted.

Receivers

- Many do not fully account for selection.
- Some are not very responsive.

$\rightarrow$ Less information received than predicted.