Strong truthfulness in Peer Prediction with Overlapping Tasks

David C. Parkes

Computer Science
John A. Paulson School of Engineering
Harvard University

November 18, 2015

Joint work with Victor Shnayder and Rafael Frongillo

Information Elicitation Without Verification

Illustrative examples:

- Participatory sensing
- Emotional response to content
- Consumer surveys
- Algorithm feedback
- Peer grading in MOOCs

Effort is costly. Need to reward informative responses, but without any ground truth; avoid unintended equilibria, collusion.

How to do this?

Information Elicitation Without Verification

Illustrative examples:

- Participatory sensing
- Emotional response to content
- Consumer surveys
- Algorithm feedback
- Peer grading in MOOCs

Effort is costly. Need to reward informative responses, but without any ground truth; avoid unintended equilibria, collusion.

How to do this?

Information Elicitation Without Verification

Illustrative examples:

- Participatory sensing
- Emotional response to content
- Consumer surveys
- Algorithm feedback
- Peer grading in MOOCs

Effort is costly. Need to reward informative responses, but without any ground truth; avoid unintended equilibria, collusion.

How to do this?

Two Kinds of Mechanisms

- Minimal mechanisms : ask agents for information ('signal')
- Non-minimal: ask agents for signal, along with belief about signal of another agent.

Two Kinds of Mechanisms

- **Minimal mechanisms**: ask agents for information ('signal')
- Non-minimal: ask agents for signal, along with belief about signal of another agent.

Simple Output agreement

- Two agents, joint *signal distribution* $P(X_1, X_2)$
- Signals $i, j \in \{1, \ldots, m\}$
- Take *reports* $\{r_1, r_2\}$, provide payment to each agent:

$$\text{report } r_1 \begin{array}{c} 1 \\ 2 \end{array} \overset{\text{report } r_2}{\overset{\begin{array}{cc} 1 & 2 \end{array}}{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}}}$$

(Strict) proper: truthful reporting is a Bayes-Nash equilibrium.

Need $P(X_2 = 1 | X_1 = 1) > P(X_2 = 2 | X_1 = 1)$*; and there are uninformative equilibria with greater payment.*

Simple Output agreement

- Two agents, joint *signal distribution* $P(X_1, X_2)$
- Signals $i, j \in \{1, \ldots, m\}$
- Take *reports* $\{r_1, r_2\}$, provide payment to each agent:

$$\text{report } r_1 \quad \begin{matrix} \\ 1 \\ 2 \end{matrix} \quad \overset{\text{report } r_2}{\begin{matrix} 1 & 2 \\ \end{matrix}} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

(Strict) proper: truthful reporting is a Bayes-Nash equilibrium.

Need $P(X_2 = 1|X_1 = 1) > P(X_2 = 2|X_1 = 1)$; and there are uninformative equilibria with greater payment.

Simple Output agreement

- Two agents, joint *signal distribution* $P(X_1, X_2)$
- Signals $i, j \in \{1, \ldots, m\}$
- Take *reports* $\{r_1, r_2\}$, provide payment to each agent:

$$\text{report } r_1 \begin{matrix} 1 \\ 2 \end{matrix} \overset{\text{report } r_2}{\overset{\begin{matrix} 1 & 2 \end{matrix}}{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}}}$$

(Strict) proper: truthful reporting is a Bayes-Nash equilibrium.

Need $P(X_2 = 1 | X_1 = 1) > P(X_2 = 2 | X_1 = 1)$; and there are uninformative equilibria with greater payment.

The Peer Prediction Method

Miller, Resnick and Zeckhauser, 2005:

- Receive report r_1, and form belief $b_1 = P(X_2 | X_1 = r_1)$
- Use proper scoring rule $t_1(b_1, r_2)$

Strict proper.

Problems: *(1) designer needs model; (2) uninformative equilibria with greater payment.*

The Peer Prediction Method

Miller, Resnick and Zeckhauser, 2005:

- Receive report r_1, and form belief $b_1 = P(X_2|X_1 = r_1)$
- Use proper scoring rule $t_1(b_1, r_2)$

Strict proper.

Problems: *(1) designer needs model; (2) uninformative equilibria with greater payment.*

The Peer Prediction Method

Miller, Resnick and Zeckhauser, 2005:

- Receive report r_1, and form belief $b_1 = P(X_2|X_1 = r_1)$
- Use proper scoring rule $t_1(b_1, r_2)$

Strict proper.

Problems: *(1) designer needs model; (2) uninformative equilibria with greater payment.*

1/Prior, Shadowing Method

Faltings et al. (2012); Witkowski and Parkes (2012):

- Assume knowledge of *marginal probability*, $P(X)$

$$\text{report } r_1 \quad \begin{array}{c} \\ 1 \\ 2 \end{array} \begin{array}{c} \text{report } r_2 \\ \begin{array}{cc} 1 & 2 \end{array} \\ \begin{pmatrix} \frac{1}{P(1)} & 0 \\ 0 & \frac{1}{P(2)} \end{pmatrix} \end{array}$$

Strict proper (under restrictions for $m > 2$ signals).

Problems: *(1) designer needs marginal probabilities; (2) uninformative equilibria with greater payment*.

1/Prior, Shadowing Method

Faltings et al. (2012); Witkowski and Parkes (2012):

- Assume knowledge of *marginal probability*, $P(X)$

$$\begin{array}{ccc} & & \text{report } r_2 \\ & & \begin{array}{cc} 1 & 2 \end{array} \\ \text{report } r_1 & \begin{array}{c} 1 \\ 2 \end{array} & \begin{pmatrix} \frac{1}{P(1)} & 0 \\ 0 & \frac{1}{P(2)} \end{pmatrix} \end{array}$$

Strict proper (under restrictions for $m > 2$ signals).

Problems: *(1) designer needs marginal probabilities; (2) uninformative equilibria with greater payment.*

1/Prior, Shadowing Method

Faltings et al. (2012); Witkowski and Parkes (2012):

- Assume knowledge of *marginal probability*, $P(X)$

$$\text{report } r_1 \quad \begin{array}{c} \\ 1 \\ 2 \end{array} \begin{array}{c} \text{report } r_2 \\ \begin{array}{cc} 1 & 2 \end{array} \\ \begin{pmatrix} \frac{1}{P(1)} & 0 \\ 0 & \frac{1}{P(2)} \end{pmatrix} \end{array}$$

Strict proper (under restrictions for $m > 2$ signals).

Problems: *(1) designer needs marginal probabilities; (2) uninformative equilibria with greater payment*.

Mechanism Desiderata

For peer prediction to be used in practice:

- Minimal mechanism
- (Strictly) Proper
- Low knowledge requirements on designer
- Truthful reports maximize expected payments:
  - *Strong-truthfulness* (in case of a tie ⇒ permutation)
  - *Informed-truthfulness* (in case of a tie ⇒ informed strategy)
- Heterogeneous agents (i.e., qualities, tastes)

Can assume multiple (independent) tasks.

What do we know?

- Jurca and Faltings, 2009.
  - $n \geq 4$ agents, knock-out pure, uninformative equil. *Ignore mixed equilibria, binary-signal only, require model.*
- Dasgupta and Ghosh, 2013.
  - Multiple tasks. Strict-proper, strong-truthful. *Binary-signal only.*
- Radonovic and Faltings, 2015.
  - Multiple tasks. Strict-proper, strong-truthful amongst symmetric strategies. *Multi-signal, but results only hold asymptotically, and need homogeneous agents.*
- Kamble et al., 2015.
  - Multiple tasks. Strict-proper, strong-truthful amongst symmetric strategies. *Multi-signal, but results only hold asymptotically, and need homogeneous agents.*
- Cai, Daskalakis and Papadimitriou, 2015.
  - Multiple tasks (with distinct, known context). Non-binary effort. Optimal effort in unique, DSE. *Multi-signal, but ignore misreports (not strong-truthful, not proper.)*

What do we know?

- Jurca and Faltings, 2009.
  - $n \geq 4$ agents, knock-out pure, uninformative equil. *Ignore mixed equilibria, binary-signal only, require model.*
- Dasgupta and Ghosh, 2013.
  - Multiple tasks. Strict-proper, strong-truthful. *Binary-signal only.*
- Radonovic and Faltings, 2015.
  - Multiple tasks. Strict-proper, strong-truthful amongst symmetric strategies. *Multi-signal, but results only hold asymptotically, and need homogeneous agents.*
- Kamble et al., 2015.
  - Multiple tasks. Strict-proper, strong-truthful amongst symmetric strategies. *Multi-signal, but results only hold asymptotically, and need homogeneous agents.*
- Cai, Daskalakis and Papadimitriou, 2015.
  - Multiple tasks (with distinct, known context). Non-binary effort. Optimal effort in unique, DSE. *Multi-signal, but ignore misreports (not strong-truthful, not proper.)*

What do we know?

- Jurca and Faltings, 2009.
  - $n \geq 4$ agents, knock-out pure, uninformative equil. *Ignore mixed equilibria, binary-signal only, require model.*
- Dasgupta and Ghosh, 2013.
  - Multiple tasks. Strict-proper, strong-truthful. *Binary-signal only.*
- Radonovic and Faltings, 2015.
  - Multiple tasks. Strict-proper, strong-truthful amongst symmetric strategies. *Multi-signal, but results only hold asymptotically, and need homogeneous agents.*
- Kamble et al., 2015.
  - Multiple tasks. Strict-proper, strong-truthful amongst symmetric strategies. *Multi-signal, but results only hold asymptotically, and need homogeneous agents.*
- Cai, Daskalakis and Papadimitriou, 2015.
  - Multiple tasks (with distinct, known context). Non-binary effort. Optimal effort in unique, DSE. *Multi-signal, but ignore misreports (not strong-truthful, not proper.)*

What do we know?

- Jurca and Faltings, 2009.
  - $n \geq 4$ agents, knock-out pure, uninformative equil. *Ignore mixed equilibria, binary-signal only, require model.*
- Dasgupta and Ghosh, 2013.
  - Multiple tasks. Strict-proper, strong-truthful. *Binary-signal only.*
- Radonovic and Faltings, 2015.
  - Multiple tasks. Strict-proper, strong-truthful amongst symmetric strategies. *Multi-signal, but results only hold asymptotically, and need homogeneous agents.*
- Kamble et al., 2015.
  - Multiple tasks. Strict-proper, strong-truthful amongst symmetric strategies. *Multi-signal, but results only hold asymptotically, and need homogeneous agents.*
- Cai, Daskalakis and Papadimitriou, 2015.
  - Multiple tasks (with distinct, known context). Non-binary effort. Optimal effort in unique, DSE. *Multi-signal, but ignore misreports (not strong-truthful, not proper.)*

What do we know?

- Jurca and Faltings, 2009.
  - $n \geq 4$ agents, knock-out pure, uninformative equil. *Ignore mixed equilibria, binary-signal only, require model.*
- Dasgupta and Ghosh, 2013.
  - Multiple tasks. Strict-proper, strong-truthful. *Binary-signal only.*
- Radonovic and Faltings, 2015.
  - Multiple tasks. Strict-proper, strong-truthful amongst symmetric strategies. *Multi-signal, but results only hold asymptotically, and need homogeneous agents.*
- Kamble et al., 2015.
  - Multiple tasks. Strict-proper, strong-truthful amongst symmetric strategies. *Multi-signal, but results only hold asymptotically, and need homogeneous agents.*
- Cai, Daskalakis and Papadimitriou, 2015.
  - Multiple tasks (with distinct, known context). Non-binary effort. Optimal effort in unique, DSE. *Multi-signal, but ignore misreports (not strong-truthful, not proper.)*

Experimental evidence

- Gao, Mao, Chen and Adams, 2014.
  - This matters! mTurk experiment (see either *collusion*, or *confusion*.)

Our Contributions: Robust, Multi-Signal Methods

- *OA-mechanism*. Strict-proper and strong-truthful for multi-signal, *categorical domains* (generalizes DG'13).

With some domain knowledge:

- *01-mechanism*. Informed-truthful and proper for general, multi-signal domains (allow heterogeneity).
- *ABCD-mechanism*. Strong-truthful (symmetric) for general, multi-signal domains w/ het. *Not proper, but all equil strictly worse than truth in large system.*

Empirical analysis:

- ∼100 questions across ∼30 exercises in 17 MOOCs. Around 325,000 peer-evaluation responses.

Our Contributions: Robust, Multi-Signal Methods

- *OA-mechanism*. Strict-proper and strong-truthful for multi-signal, *categorical domains* (generalizes DG'13).

With some domain knowledge:

- *01-mechanism*. Informed-truthful and proper for general, multi-signal domains (allow heterogeneity).
- *ABCD-mechanism*. Strong-truthful (symmetric) for general, multi-signal domains w/ het. *Not proper, but all equil strictly worse than truth in large system.*

Empirical analysis:

- ∼100 questions across ∼30 exercises in 17 MOOCs. Around 325,000 peer-evaluation responses.

Our Contributions: Robust, Multi-Signal Methods

- *OA-mechanism*. Strict-proper and strong-truthful for multi-signal, *categorical domains* (generalizes DG'13).

With some domain knowledge:

- *01-mechanism*. Informed-truthful and proper for general, multi-signal domains (allow heterogeneity).
- *ABCD-mechanism*. Strong-truthful (symmetric) for general, multi-signal domains w/ het. *Not proper, but all equil strictly worse than truth in large system.*

Empirical analysis:

- ∼100 questions across ∼30 exercises in 17 MOOCs. Around 325,000 peer-evaluation responses.

Basic Set-up

- Agents 1, 2
- Tasks k (≥ 3); Signals $i, j \in \{1, \ldots, m\}$ (require effort)
- Joint distribution $P(X_1 = i, X_2 = j)$ (possibly asymmetric)
- Overlapping tasks: *shared* K_s, *agent 1* K_1, *agent 2* K_2.
- *Multi-task peer prediction*: for each $k \in K_s$, payment $\{r_1^k, r_2^k, r_1^{K_1}, r_2^{K_2}\} \mapsto \mathbb{R}$
- *Strategies*: $F_{ir} = P(r_1 = r | X_1 = i)$ $\quad G_{jr} = P(r_2 = r | X_2 = j)$
  - *Informed strategy*: $F_{ir} \neq F_{jr}$, some $i \neq j$, some r
  - *Truthful strategy*: F^*

Basic Set-up

- Agents 1, 2
- Tasks k (≥ 3); Signals $i, j \in \{1, \ldots, m\}$ (require effort)
- Joint distribution $P(X_1 = i, X_2 = j)$ (possibly asymmetric)
- Overlapping tasks: *shared* K_s, *agent 1* K_1, *agent 2* K_2.
- *Multi-task peer prediction*: for each $k \in K_s$, payment $\{r_1^k, r_2^k, r_1^{K_1}, r_2^{K_2}\} \mapsto \mathbb{R}$
- *Strategies*: $F_{ir} = P(r_1 = r | X_1 = i) \quad G_{jr} = P(r_2 = r | X_2 = j)$
  - *Informed strategy*: $F_{ir} \neq F_{jr}$, some $i \neq j$, some r
  - *Truthful strategy*: F^*

Basic Set-up

- Agents 1, 2
- Tasks k (≥ 3); Signals $i, j \in \{1, \ldots, m\}$ (require effort)
- Joint distribution $P(X_1 = i, X_2 = j)$ (possibly asymmetric)
- Overlapping tasks: *shared* K_s, *agent 1* K_1, *agent 2* K_2.
- *Multi-task peer prediction*: for each $k \in K_s$, payment $\{r_1^k, r_2^k, r_1^{K_1}, r_2^{K_2}\} \mapsto \mathbb{R}$
- *Strategies*: $F_{ir} = P(r_1 = r | X_1 = i)$ $\quad G_{jr} = P(r_2 = r | X_2 = j)$
  - *Informed strategy*: $F_{ir} \neq F_{jr}$, some $i \neq j$, some r
  - *Truthful strategy*: F^*

Basic Set-up

- Agents 1, 2
- Tasks k (≥ 3); Signals $i, j \in \{1, \ldots, m\}$ (require effort)
- Joint distribution $P(X_1 = i, X_2 = j)$ (possibly asymmetric)
- Overlapping tasks: *shared* K_s, *agent 1* K_1, *agent 2* K_2.
- *Multi-task peer prediction*: for each $k \in K_s$, payment $\{r_1^k, r_2^k, r_1^{K_1}, r_2^{K_2}\} \mapsto \mathbb{R}$
- *Strategies*: $F_{ir} = P(r_1 = r | X_1 = i) \quad G_{jr} = P(r_2 = r | X_2 = j)$
  - *Informed strategy*: $F_{ir} \neq F_{jr}$, some $i \neq j$, some r
  - *Truthful strategy*: F^*

Solution Concepts

- $E(F, G)$: *expected payment* for a shared task
- *Bayes-Nash equil.*
- *(Strict) Proper*: $E(F^*, G^*) \geq E(F, G^*)$, for all $F \neq F^*$
- *Strong-truthful*: $E(F^*, G^*) \geq E(F, G)$, for all F, G (tie $\Rightarrow$ permutation); also strict proper
- *Informed-truthful*: $E(F^*, G^*) \geq E(F, G)$, for all F, G (tie $\Rightarrow$ informed); also proper

Solution Concepts

- $E(F, G)$: *expected payment* for a shared task
- *Bayes-Nash equil.*
- *(Strict) Proper*: $E(F^*, G^*) \geq E(F, G^*)$, for all $F \neq F^*$
- *Strong-truthful*: $E(F^*, G^*) \geq E(F, G)$, for all F, G (tie $\Rightarrow$ permutation); also strict proper
- *Informed-truthful*: $E(F^*, G^*) \geq E(F, G)$, for all F, G (tie $\Rightarrow$ informed); also proper

Solution Concepts

- $E(F, G)$: *expected payment* for a shared task
- *Bayes-Nash equil.*
- *(Strict) Proper*: $E(F^*, G^*) \geq E(F, G^*)$, for all $F \neq F^*$
- *Strong-truthful*: $E(F^*, G^*) \geq E(F, G)$, for all F, G (tie $\Rightarrow$ permutation); also strict proper
- *Informed-truthful*: $E(F^*, G^*) \geq E(F, G)$, for all F, G (tie $\Rightarrow$ informed); also proper

Family of Mechanisms

Parameterized by score $S : \{1, \ldots, m\} \times \{1, \ldots, m\} \mapsto \mathbb{R}$

- Assign agents to tasks, get reports
- For shared $k \in K_s$, pay both 1 and 2

$$S(r_1^k, r_2^k) - S(r_1^\ell, r_2^m),$$

for $\ell \in K_1$ and $m \in K_2$ (can also take empirical average)

Idea: Reward 'excess agreement' not 'default agreement.'

- Zero payment if say '1' all the time, or random report.
- For S as the identify (output-agreement) matrix, this is multi-signal generalization of DG'13.

Family of Mechanisms

Parameterized by score $S : \{1, \ldots, m\} \times \{1, \ldots, m\} \mapsto \mathbb{R}$

- Assign agents to tasks, get reports
- For shared $k \in K_s$, pay both 1 and 2

$$S(r_1^k, r_2^k) - S(r_1^\ell, r_2^m),$$

for $\ell \in K_1$ and $m \in K_2$ (can also take empirical average)

Idea: Reward 'excess agreement' not 'default agreement.'

- Zero payment if say '1' all the time, or random report.
- For S as the identify (output-agreement) matrix, this is multi-signal generalization of DG'13.

Family of Mechanisms

Parameterized by score $S : \{1, \ldots, m\} \times \{1, \ldots, m\} \mapsto \mathbb{R}$

- Assign agents to tasks, get reports
- For shared $k \in K_s$, pay both 1 and 2

$$S(r_1^k, r_2^k) - S(r_1^\ell, r_2^m),$$

for $\ell \in K_1$ and $m \in K_2$ (can also take empirical average)

Idea: Reward 'excess agreement' not 'default agreement.'

- Zero payment if say '1' all the time, or random report.
- For S as the identify (output-agreement) matrix, this is multi-signal generalization of DG'13.

Analysis: Expected Payment

For identity score-matrix:

$$E(F,G) = \sum_{ij} P(i,j) \sum_r F_{ir}G_{jr} - \sum_{ij} P(i)P(j) \sum_r F_{ir}G_{jr}$$
$$= \sum_{ij} \Delta_{ij} \sum_r F_{ir}G_{jr}.$$

Delta matrix:

- $\Delta_{ij} = P(i,j) - P(i)P(j)$; if $\Delta_{ij} > 0$ then $P(j|i) > P(j)$

Example: P: $\begin{pmatrix} 0.4 & 0.15 \\ 0.15 & 0.3 \end{pmatrix}$ $\Delta \approx \begin{pmatrix} 0.1 & -0.1 \\ -0.1 & 0.1 \end{pmatrix}$ or, $\begin{pmatrix} + & - \\ - & + \end{pmatrix}$

For general S:

$$E(F,G) = \sum_{ij} \Delta_{ij} \sum_{r_1,r_2} S_{r_1,r_2} F_{ir_1} G_{jr_2}$$

Analysis: Expected Payment

For identity score-matrix:

$$E(F,G) = \sum_{ij} P(i,j) \sum_{r} F_{ir} G_{jr} - \sum_{ij} P(i)P(j) \sum_{r} F_{ir} G_{jr}$$
$$= \sum_{ij} \Delta_{ij} \sum_{r} F_{ir} G_{jr}.$$

Delta matrix:

- $\Delta_{ij} = P(i,j) - P(i)P(j)$; if $\Delta_{ij} > 0$ then $P(j|i) > P(j)$

Example: P: $\begin{pmatrix} 0.4 & 0.15 \\ 0.15 & 0.3 \end{pmatrix}$ $\Delta \approx \begin{pmatrix} 0.1 & -0.1 \\ -0.1 & 0.1 \end{pmatrix}$ or, $\begin{pmatrix} + & - \\ - & + \end{pmatrix}$

For general S:

$$E(F,G) = \sum_{ij} \Delta_{ij} \sum_{r_1,r_2} S_{r_1,r_2} F_{ir_1} G_{jr_2}$$

Analysis: Expected Payment

For identity score-matrix:

$$E(F,G) = \sum_{ij} P(i,j) \sum_{r} F_{ir} G_{jr} - \sum_{ij} P(i)P(j) \sum_{r} F_{ir} G_{jr}$$
$$= \sum_{ij} \Delta_{ij} \sum_{r} F_{ir} G_{jr}.$$

Delta matrix:

- $\Delta_{ij} = P(i,j) - P(i)P(j)$; if $\Delta_{ij} > 0$ then $P(j|i) > P(j)$

Example: P: $\begin{pmatrix} 0.4 & 0.15 \\ 0.15 & 0.3 \end{pmatrix}$ $\Delta \approx \begin{pmatrix} 0.1 & -0.1 \\ -0.1 & 0.1 \end{pmatrix}$ or, $\begin{pmatrix} + & - \\ - & + \end{pmatrix}$

For general S:

$$E(F,G) = \sum_{ij} \Delta_{ij} \sum_{r_1,r_2} S_{r_1,r_2} F_{ir_1} G_{jr_2}$$

Deterministic Strategies

Lemma 1

Deterministic F, G maximize $E(F, G)$.

$$E(F, G) = \max_F \max_G h(F, G) = \max_F OBJ(F),$$

where $h(F, G)$ is linear in either argument. Fixing F, opt G is deterministic. $OBJ(F)$ is convex, and opt F is deterministic.

Can *focus on deterministic strategies*:

- sufficient to prove strong-truthful, or and informed-truthful.
- sufficient to check for deviations from truthful

Deterministic Strategies

Lemma 1

Deterministic F, G maximize $E(F, G)$.

$$E(F, G) = \max_F \max_G \; h(F, G) = \max_F OBJ(F),$$

where $h(F, G)$ is linear in either argument. Fixing F, opt G is deterministic. $OBJ(F)$ is convex, and opt F is deterministic.

Can *focus on deterministic strategies*:

- sufficient to prove strong-truthful, or and informed-truthful.
- sufficient to check for deviations from truthful

Deterministic Strategies

Lemma 1

Deterministic F, G maximize $E(F, G)$.

$$E(F, G) = \max_F \max_G \; h(F, G) = \max_F OBJ(F),$$

where $h(F, G)$ is linear in either argument. Fixing F, opt G is deterministic. $OBJ(F)$ is convex, and opt F is deterministic.

Can *focus on deterministic strategies*:

- sufficient to prove strong-truthful, or and informed-truthful.
- sufficient to check for deviations from truthful

Simplified analysis

Deterministic strategies $F(i), G(j)$.

For identity-matrix S:

$$E(F, G) = \sum_{ij} \Delta_{ij} \mathbb{I}(F(i) = G(j))$$

For general score matrix S:

$$E(F, G) = \sum_{ij} \Delta_{ij} S_{F(i), G(j)}$$

The game is to find 'which scores to pick' for each (i,j) pair.

Simplified analysis

Deterministic strategies $F(i), G(j)$.

For identity-matrix S:

$$E(F,G) = \sum_{ij} \Delta_{ij} \mathbb{I}(F(i) = G(j))$$

For general score matrix S:

$$E(F,G) = \sum_{ij} \Delta_{ij} S_{F(i),G(j)}$$

The game is to find 'which scores to pick' for each (i,j) pair.

Simplified analysis

Deterministic strategies $F(i), G(j)$.

For identity-matrix S:

$$E(F, G) = \sum_{ij} \Delta_{ij} \mathbb{I}(F(i) = G(j))$$

For general score matrix S:

$$E(F, G) = \sum_{ij} \Delta_{ij} S_{F(i), G(j)}$$

The game is to find 'which scores to pick' for each (i,j) pair.

The OA-Mechanism

S is the identity matrix.

- Categorical domain:

$$\text{sig}(\Delta) : \begin{pmatrix} + & - & - \\ - & + & - \\ - & - & + \end{pmatrix}$$

- Image labeling {swim, fly, walk}, vs. grading {76, 78, 79, . . .}

Theorem 1

The OA-mechanism is strict-proper and strongly-truthful if the world is categorical.

Obtain DG'13 as a corollary. Theorem is tight.

The OA-Mechanism

S is the identity matrix.

- Categorical domain:

$$\mathrm{sig}(\Delta) : \begin{pmatrix} + & - & - \\ - & + & - \\ - & - & + \end{pmatrix}$$

- Image labeling {swim, fly, walk}, vs. grading $\{76, 78, 79, \ldots\}$

Theorem 1

The OA-mechanism is strict-proper and strongly-truthful if the world is categorical.

Obtain DG'13 as a corollary. Theorem is tight.

The OA-Mechanism

S is the identity matrix.

- Categorical domain:

$$\mathrm{sig}(\Delta) : \begin{pmatrix} + & - & - \\ - & + & - \\ - & - & + \end{pmatrix}$$

- Image labeling {swim, fly, walk}, vs. grading $\{76, 78, 79, \ldots\}$

Theorem 1

The OA-mechanism is strict-proper and strongly-truthful if the world is categorical.

Obtain DG'13 as a corollary. Theorem is tight.

The OA-mechanism

Theorem 1

The OA-mechanism is strict-proper and strongly-truthful if the world is categorical.

$$E(F^*, G^*) = \sum_i \Delta_{ii} = \sum_{ij:\Delta_{ij}>0} \Delta_{ij} \geq \sum_{ij} \Delta_{ij} \mathbb{1}(F(i) = G(j)) = E(F, G),$$

for all F, G.

Also need: tie in payment ⇒ permutation strategy.

- Case 1: Not permutation, and symmetric. Must be two i, j ($i \neq j$) that map to r. Assign $\Delta_{ij} < 0$ pair to score $S_{r,r} = 1$. Worse!
- Case 2: Asymmetric, e.g., agent 1 strategy $i \mapsto r$, agent 2 strategy $i \mapsto r'$. Assign $\Delta_{ii} > 0$ pair to score $S_{r,r'} = 0$. Worse!

The OA-mechanism

Theorem 1

The OA-mechanism is strict-proper and strongly-truthful if the world is categorical.

$$E(F^*, G^*) = \sum_i \Delta_{ii} = \sum_{ij:\Delta_{ij}>0} \Delta_{ij} \geq \sum_{ij} \Delta_{ij} \mathbb{1}(F(i) = G(j)) = E(F, G),$$

for all F, G.

Also need: tie in payment ⇒ permutation strategy.

- Case 1: Not permutation, and symmetric. Must be two i, j ($i \neq j$) that map to r. Assign $\Delta_{ij} < 0$ pair to score $S_{r,r} = 1$. Worse!
- Case 2: Asymmetric, e.g., agent 1 strategy $i \mapsto r$, agent 2 strategy $i \mapsto r'$. Assign $\Delta_{ii} > 0$ pair to score $S_{r,r'} = 0$. Worse!

The OA-mechanism

Theorem 1

The OA-mechanism is strict-proper and strongly-truthful if the world is categorical.

$$E(F^*, G^*) = \sum_i \Delta_{ii} = \sum_{ij:\Delta_{ij}>0} \Delta_{ij} \geq \sum_{ij} \Delta_{ij} \mathbb{1}(F(i) = G(j)) = E(F, G),$$

for all F, G.

Also need: tie in payment ⇒ permutation strategy.

- Case 1: Not permutation, and symmetric. Must be two i, j ($i \neq j$) that map to r. Assign $\Delta_{ij} < 0$ pair to score $S_{r,r} = 1$. Worse!
- Case 2: Asymmetric, e.g., agent 1 strategy $i \mapsto r$, agent 2 strategy $i \mapsto r'$. Assign $\Delta_{ii} > 0$ pair to score $S_{r,r'} = 0$. Worse!

The OA-mechanism

Theorem 1

The OA-mechanism is strict-proper and strongly-truthful if the world is categorical.

$$E(F^*, G^*) = \sum_i \Delta_{ii} = \sum_{ij:\Delta_{ij}>0} \Delta_{ij} \geq \sum_{ij} \Delta_{ij} \mathbb{1}(F(i) = G(j)) = E(F, G),$$

for all F, G.

Also need: tie in payment $\Rightarrow$ permutation strategy.

- Case 1: Not permutation, and symmetric. Must be two i, j ($i \neq j$) that map to r. Assign $\Delta_{ij} < 0$ pair to score $S_{r,r} = 1$. Worse!
- Case 2: Asymmetric, e.g., agent 1 strategy $i \mapsto r$, agent 2 strategy $i \mapsto r'$. Assign $\Delta_{ii} > 0$ pair to score $S_{r,r'} = 0$. Worse!

The OA-mechanism

Theorem 1

The OA-mechanism is strict-proper and strongly-truthful if the world is categorical.

$$E(F^*, G^*) = \sum_i \Delta_{ii} = \sum_{ij:\Delta_{ij}>0} \Delta_{ij} \geq \sum_{ij} \Delta_{ij} \mathbb{1}(F(i) = G(j)) = E(F, G),$$

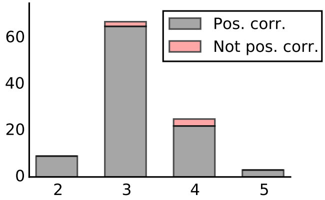
for all F, G.

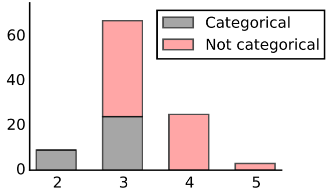
Also need: tie in payment $\Rightarrow$ permutation strategy.

- Case 1: Not permutation, and symmetric. Must be two i, j ($i \neq j$) that map to r. Assign $\Delta_{ij} < 0$ pair to score $S_{r,r} = 1$. Worse!
- Case 2: Asymmetric, e.g., agent 1 strategy $i \mapsto r$, agent 2 strategy $i \mapsto r'$. Assign $\Delta_{ii} > 0$ pair to score $S_{r,r'} = 0$. Worse!

Are we done? Let's look at some data.

- Peer-evaluation responses to 100 questions across 30 exercises in 17 MOOCs
- Vast majority of questions have $m \in \{2, 3, 4\}$.
  - Example rubric element: "Not much of a style at all", "Communicative style", and "Strong, flowing writing style".





The 01-mechanism

$$S_{ij} = \begin{cases} 1 & \text{, if } \Delta_{ij} > 0 \\ 0 & \text{o.w.} \end{cases} \quad \Delta = \begin{pmatrix} + & + & - \\ + & + & - \\ - & - & + \end{pmatrix} \quad S = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Theorem 2

For general domains, the 01-mechanism is informed truthful (and proper).

$$E(F^*, G^*) = \sum_{ij} \Delta_{ij} S_{ij} = \sum_{ij:\Delta_{ij}>0} \Delta_{ij} \geq E(F, G), \quad \forall F, G$$

Also need to show uninformed $\Rightarrow$ strictly less payment.
Fix G, consider uninformed F (e.g., $F(i) = `1'$, for all i). Have $E(F, G) = \sum_{ij} \Delta_{ij} S_{1,G(j)} < E(F^*, G^*)$.

Note: indifference between '1 → 1, 2 → 2' and '1 → 1, 2 → 1'.

The 01-mechanism

$$S_{ij} = \begin{cases} 1 & \text{, if } \Delta_{ij} > 0 \\ 0 & \text{o.w.} \end{cases} \quad \Delta = \begin{pmatrix} + & + & - \\ + & + & - \\ - & - & + \end{pmatrix} S = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Theorem 2

For general domains, the 01-mechanism is informed truthful (and proper).

$$E(F^*, G^*) = \sum_{ij} \Delta_{ij} S_{ij} = \sum_{ij:\Delta_{ij}>0} \Delta_{ij} \geq E(F, G), \quad \forall F, G$$

Also need to show uninformed ⇒ strictly less payment.
Fix G, consider uninformed F (e.g., $F(i) = \text{`}1\text{'}$, for all i). Have $E(F, G) = \sum_{ij} \Delta_{ij} S_{1,G(j)} < E(F^*, G^*)$.

Note: indifference between '1 → 1, 2 → 2' and '1 → 1, 2 → 1'.

The 01-mechanism

$$S_{ij} = \begin{cases} 1 & \text{, if } \Delta_{ij} > 0 \\ 0 & \text{o.w.} \end{cases} \quad \Delta = \begin{pmatrix} + & + & - \\ + & + & - \\ - & - & + \end{pmatrix} S = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Theorem 2

For general domains, the 01-mechanism is informed truthful (and proper).

$$E(F^*, G^*) = \sum_{ij} \Delta_{ij} S_{ij} = \sum_{ij:\Delta_{ij}>0} \Delta_{ij} \geq E(F, G), \quad \forall F, G$$

Also need to show uninformed $\Rightarrow$ strictly less payment.
Fix G, consider uninformed F (e.g., $F(i) = \text{'1'}$, for all i). Have $E(F, G) = \sum_{ij} \Delta_{ij} S_{1,G(j)} < E(F^*, G^*)$.

Note: indifference between '1 → 1, 2 → 2' and '1 → 1, 2 → 1'.

The 01-mechanism

$$S_{ij} = \begin{cases} 1 & \text{, if } \Delta_{ij} > 0 \\ 0 & \text{o.w.} \end{cases} \quad \Delta = \begin{pmatrix} + & + & - \\ + & + & - \\ - & - & + \end{pmatrix} \quad S = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Theorem 2

For general domains, the 01-mechanism is informed truthful (and proper).

$$E(F^*, G^*) = \sum_{ij} \Delta_{ij} S_{ij} = \sum_{ij:\Delta_{ij}>0} \Delta_{ij} \geq E(F, G), \quad \forall F, G$$

Also need to show uninformed ⇒ strictly less payment.

Fix G, consider uninformed F (e.g., $F(i) = \text{‘}1\text{’}$, for all i). Have $E(F, G) = \sum_{ij} \Delta_{ij} S_{1,G(j)} < E(F^*, G^*)$.

Note: indifference between ‘1 → 1, 2 → 2’ and ‘1 → 1, 2 → 1’.

The 01-mechanism

$$S_{ij} = \begin{cases} 1 & \text{, if } \Delta_{ij} > 0 \\ 0 & \text{o.w.} \end{cases} \quad \Delta = \begin{pmatrix} + & + & - \\ + & + & - \\ - & - & + \end{pmatrix} \quad S = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Theorem 2

For general domains, the 01-mechanism is informed truthful (and proper).

$$E(F^*, G^*) = \sum_{ij} \Delta_{ij} S_{ij} = \sum_{ij:\Delta_{ij}>0} \Delta_{ij} \geq E(F, G), \quad \forall F, G$$

Also need to show uninformed ⇒ strictly less payment.
Fix G, consider uninformed F (e.g., $F(i) = \text{'}1\text{'}$, for all i). Have $E(F, G) = \sum_{ij} \Delta_{ij} S_{1,G(j)} < E(F^*, G^*)$.

Note: indifference between '1 → 1, 2 → 2' and '1 → 1, 2 → 1'.

The 01-mechanism

$$S_{ij} = \begin{cases} 1 & \text{, if } \Delta_{ij} > 0 \\ 0 & \text{o.w.} \end{cases} \quad \Delta = \begin{pmatrix} + & + & - \\ + & + & - \\ - & - & + \end{pmatrix} \; S = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Theorem 2

For general domains, the 01-mechanism is informed truthful (and proper).

$$E(F^*, G^*) = \sum_{ij} \Delta_{ij} S_{ij} = \sum_{ij:\Delta_{ij}>0} \Delta_{ij} \geq E(F, G), \quad \forall F, G$$

Also need to show uninformed ⇒ strictly less payment.
Fix G, consider uninformed F (e.g., $F(i) = `1'$, for all i). Have $E(F, G) = \sum_{ij} \Delta_{ij} S_{1,G(j)} < E(F^*, G^*)$.

Note: indifference between '1 → 1, 2 → 2' and '1 → 1, 2 → 1'.

The ABCD-mechanism

Parameterized $0 \leq a < b < c < d$

Scores: $S_{ii} = \begin{cases} b & \text{, if } \Delta_{ii} \leq 0 \\ c & \text{o.w.} \end{cases}$ $\quad S_{ij} = \begin{cases} a & \text{, if } \Delta_{ij} \leq 0 \\ d & \text{o.w.} \end{cases}$

Theorem 3

For general domains, the ABCD-mechanism is strong-truthful amongst symmetric strategies.

Expected payment $E(F^*, G^*) = \sum_{ij} \Delta_{ij} S_{ij}$
Consider a non-permutation, symmetric strategy. Must be i, j ($i \neq j$) that map to same r. Assigns score $S_{r,r} \in \{b, c\}$ to (i, j) and (j, i), worse because $a < \{b, c\} < d$.

Not proper. But, any equil. in a large economy has *strictly less payment*. Prisoner's dilemma!

The ABCD-mechanism

Parameterized $0 \leq a < b < c < d$

Scores: $S_{ii} = \begin{cases} b & \text{, if } \Delta_{ii} \leq 0 \\ c & \text{o.w.} \end{cases}$ $\quad S_{ij} = \begin{cases} a & \text{, if } \Delta_{ij} \leq 0 \\ d & \text{o.w.} \end{cases}$

Theorem 3

For general domains, the ABCD-mechanism is strong-truthful amongst symmetric strategies.

Expected payment $E(F^*, G^*) = \sum_{ij} \Delta_{ij} S_{ij}$

Consider a non-permutation, symmetric strategy. Must be i, j ($i \neq j$) that map to same r. Assigns score $S_{r,r} \in \{b, c\}$ to (i, j) and (j, i), worse because $a < \{b, c\} < d$.

Not proper. But, any equil. in a large economy has *strictly less payment*. Prisoner's dilemma!

The ABCD-mechanism

Parameterized $0 \leq a < b < c < d$

Scores: $S_{ii} = \begin{cases} b & \text{, if } \Delta_{ii} \leq 0 \\ c & \text{o.w.} \end{cases}$ $S_{ij} = \begin{cases} a & \text{, if } \Delta_{ij} \leq 0 \\ d & \text{o.w.} \end{cases}$

Theorem 3

For general domains, the ABCD-mechanism is strong-truthful amongst symmetric strategies.

Expected payment $E(F^*, G^*) = \sum_{ij} \Delta_{ij} S_{ij}$

Consider a non-permutation, symmetric strategy. Must be i, j $(i \neq j)$ that map to same r. Assigns score $S_{r,r} \in \{b, c\}$ to (i, j) and (j, i), worse because $a < \{b, c\} < d$.

Not proper. But, any equil. in a large economy has *strictly less payment*. Prisoner's dilemma!

The ABCD-mechanism

Parameterized $0 \leq a < b < c < d$

Scores: $S_{ii} = \begin{cases} b & \text{, if } \Delta_{ii} \leq 0 \\ c & \text{o.w.} \end{cases}$ $S_{ij} = \begin{cases} a & \text{, if } \Delta_{ij} \leq 0 \\ d & \text{o.w.} \end{cases}$

Theorem 3

For general domains, the ABCD-mechanism is strong-truthful amongst symmetric strategies.

Expected payment $E(F^*, G^*) = \sum_{ij} \Delta_{ij} S_{ij}$
Consider a non-permutation, symmetric strategy. Must be i, j ($i \neq j$) that map to same r. Assigns score $S_{r,r} \in \{b, c\}$ to (i, j) and (j, i), worse because $a < \{b, c\} < d$.

Not proper. But, any equil. in a large economy has *strictly less payment*. Prisoner's dilemma!

The ABCD-mechanism

Parameterized $0 \leq a < b < c < d$

Scores: $S_{ii} = \begin{cases} b & \text{, if } \Delta_{ii} \leq 0 \\ c & \text{o.w.} \end{cases}$ $\quad S_{ij} = \begin{cases} a & \text{, if } \Delta_{ij} \leq 0 \\ d & \text{o.w.} \end{cases}$

Theorem 3

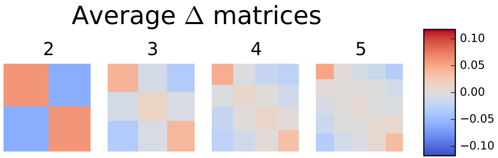
For general domains, the ABCD-mechanism is strong-truthful amongst symmetric strategies.

Expected payment $E(F^*, G^*) = \sum_{ij} \Delta_{ij} S_{ij}$
Consider a non-permutation, symmetric strategy. Must be i, j ($i \neq j$) that map to same r. Assigns score $S_{r,r} \in \{b, c\}$ to (i, j) and (j, i), worse because $a < \{b, c\} < d$.

Not proper. But, any equil. in a large economy has *strictly less payment*. Prisoner's dilemma!

Delta matrices: MOOC Data

- 17 courses, 104 questions, ∼325,000 reports.

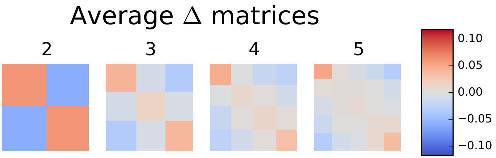


Positive correlation.

For models of size 4 and 5, see failure of categorical (e.g., score 2 is +ve correlated with score 3.) *Ordinal domain*.

Delta matrices: MOOC Data

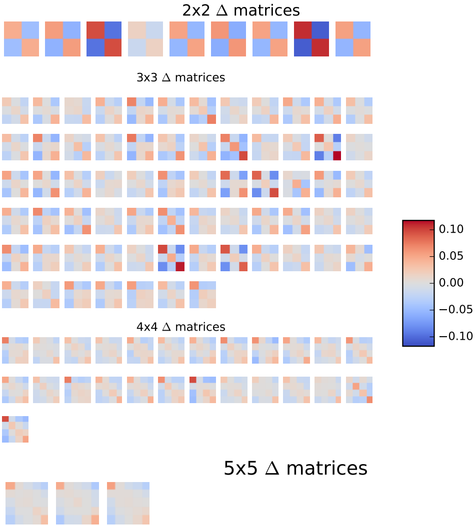
- 17 courses, 104 questions, ∼325,000 reports.



Positive correlation.

For models of size 4 and 5, see failure of categorical (e.g., score 2 is +ve correlated with score 3.) *Ordinal domain*.

Delta matrices: MOOC Data



Empirical observations

34 of 104 worlds are categorical:

- All mechanisms are strong-truthful and strict-proper.

In the other 70 worlds:

- OA-mechanism not strong-truthful or proper.
- 01-mechanism is informed-truthful and proper. It is also *strict-proper* in 19/70 worlds.
- ABCD-mechanism is strong-truthful (symmetric). It is also *strict-proper* in 12/70 worlds.
- An incomplete, heuristic search for score matrices yields *strong-truthful* mechanisms in 49/70 worlds.

Empirical observations

34 of 104 worlds are categorical:

- All mechanisms are strong-truthful and strict-proper.

In the other 70 worlds:

- OA-mechanism not strong-truthful or proper.
- 01-mechanism is informed-truthful and proper. It is also *strict-proper* in 19/70 worlds.
- ABCD-mechanism is strong-truthful (symmetric). It is also *strict-proper* in 12/70 worlds.
- An incomplete, heuristic search for score matrices yields *strong-truthful* mechanisms in 49/70 worlds.

Empirical observations

34 of 104 worlds are categorical:

- All mechanisms are strong-truthful and strict-proper.

In the other 70 worlds:

- OA-mechanism not strong-truthful or proper.
- 01-mechanism is informed-truthful and proper. It is also *strict-proper* in 19/70 worlds.
- ABCD-mechanism is strong-truthful (symmetric). It is also *strict-proper* in 12/70 worlds.
- An incomplete, heuristic search for score matrices yields *strong-truthful* mechanisms in 49/70 worlds.

Empirical observations

34 of 104 worlds are categorical:

- All mechanisms are strong-truthful and strict-proper.

In the other 70 worlds:

- OA-mechanism not strong-truthful or proper.
- 01-mechanism is informed-truthful and proper. It is also *strict-proper* in 19/70 worlds.
- ABCD-mechanism is strong-truthful (symmetric). It is also *strict-proper* in 12/70 worlds.
- An incomplete, heuristic search for score matrices yields *strong-truthful* mechanisms in 49/70 worlds.

Empirical observations

34 of 104 worlds are categorical:

- All mechanisms are strong-truthful and strict-proper.

In the other 70 worlds:

- OA-mechanism not strong-truthful or proper.
- 01-mechanism is informed-truthful and proper. It is also *strict-proper* in 19/70 worlds.
- ABCD-mechanism is strong-truthful (symmetric). It is also *strict-proper* in 12/70 worlds.
- An incomplete, heuristic search for score matrices yields *strong-truthful* mechanisms in 49/70 worlds.

Empirical observations

34 of 104 worlds are categorical:

- All mechanisms are strong-truthful and strict-proper.

In the other 70 worlds:

- OA-mechanism not strong-truthful or proper.
- 01-mechanism is informed-truthful and proper. It is also *strict-proper* in 19/70 worlds.
- ABCD-mechanism is strong-truthful (symmetric). It is also *strict-proper* in 12/70 worlds.
- An incomplete, heuristic search for score matrices yields *strong-truthful* mechanisms in 49/70 worlds.

Review: Our Results

- *OA-mechanism* (generalizes DG'13) is *strong-truthful* (and strict-proper) for *categorical domains*.
- *01-mechanism* is *informed-truthful* (and proper) for general domains. Needs knowledge of sign structure of correlations.
- *ABCD-mechanism* is *strong-truthful (symmetric)* for general domains, may not be proper. Needs knowledge of sign structure of correlations.
- *Empirical analysis* supports the need for these mechanisms.

Review: Our Results

- *OA-mechanism* (generalizes DG'13) is *strong-truthful* (and strict-proper) for *categorical domains*.
- *01-mechanism* is *informed-truthful* (and proper) for general domains. Needs knowledge of sign structure of correlations.
- *ABCD-mechanism* is *strong-truthful (symmetric)* for general domains, may not be proper. Needs knowledge of sign structure of correlations.
- *Empirical analysis* supports the need for these mechanisms.

Review: Our Results

- *OA-mechanism* (generalizes DG'13) is *strong-truthful* (and strict-proper) for *categorical domains*.
- *01-mechanism* is *informed-truthful* (and proper) for general domains. Needs knowledge of sign structure of correlations.
- *ABCD-mechanism* is *strong-truthful (symmetric)* for general domains, may not be proper. Needs knowledge of sign structure of correlations.
- *Empirical analysis* supports the need for these mechanisms.

Review: Our Results

- *OA-mechanism* (generalizes DG'13) is *strong-truthful* (and strict-proper) for *categorical domains*.
- *01-mechanism* is *informed-truthful* (and proper) for general domains. Needs knowledge of sign structure of correlations.
- *ABCD-mechanism* is *strong-truthful (symmetric)* for general domains, may not be proper. Needs knowledge of sign structure of correlations.
- *Empirical analysis* supports the need for these mechanisms.

Discussion

- Is there a proper and strong-truthful (symmetric) mechanism for general domains? Perhaps leveraging two S matrices?
- Can heterogeneity be handled (e.g., “pushing” reports towards categorical)?
- Prior-free design: can we use observed data to design and then apply a score matrix?
- Population learning: does strong- or informed-truthful promote convergence to truthful equilibrium?
- Richer models of effort.
- Experiments and applications.

Thank you