Eliciting Experts' Advice

Francisco Poggi

May 25, 2024

Motivation

- Decision makers rely on experts to make informed decisions.
- DM would like to know the expertise of different experts to properly weight their information.
- However, experts might be reluctant to reveal their (lack of) expertise (e.g. due to career concerns).
- **Question**: How can a communication protocol be designed to aggregate experts' information according to their expertise?

In this talk

- Define a property of communication protocols that captures **robustness to outside communication**.
  - Related to the concept of neologism-proofness in cheap talk.

- Apply this property to the problem of designing communication mechanisms to elicit experts' advice when
  - experts share a common language with the DM and can communicate outside of the mechanism.
  - experts have career concerns.

In this talk

- Define a property of communication protocols that captures **robustness to outside communication**.
  - Related to the concept of neologism-proofness in cheap talk.

- Apply this property to the problem of designing communication mechanisms to elicit experts' advice when
  - experts share a common language with the DM and can communicate outside of the mechanism.
  - experts have career concerns.

Neologism-proofness

- Sender-receiver framework.
- Equally likely states.

	State 1	**State 2**
Action 1	(3, 3)	(0, 0)
Action 2	(0, 0)	(3, 3)
Action 3	(2, 2)	(2, 2)

- Cheap talk equilibria:
  - **Informative equilibrium**: Action matches the state.
  - **Babbling equilibrium**: Action 3 independent of the state.

Neologism-proofness

- Sender-receiver framework.
- Equally likely states.

	State 1	**State 2**
Action 1	**(3, 3)**	(0, 0)
Action 2	(0, 0)	**(3, 3)**
Action 3	**(2, 2)**	**(2, 2)**

- Cheap talk equilibria:
  - **Informative equilibrium**: Action matches the state.
  - **Babbling equilibrium**: Action 3 independent of the state.

Neologism-proofness

- Two senders:
  - Sender 2 distinguishes {{A1,B1}, {A2,B2}} gets a constant payoff.
  - Sender 1 distinguishes {{A1,A2}, {B1,B2}}. Payoff in matrix, aligned with receiver.

	State A1	**State A2**	**State B1**	**State B2**
Action 1	(3, 3)	(0, 0)	(0, 0)	(1, 1)
Action 2	(0, 0)	(3, 3)	(1, 1)	(0, 0)
Action 3	(2, 2)	(2, 2)	(0, 0)	(0, 0)

- Mechanism Γ perfectly reveals the information of Sender 1.
- However, if S1 and R share a *common language* they can do better.

Neologism-proofness

- Two senders:
  - Sender 2 distinguishes {{A1,B1}, {A2,B2}} gets a constant payoff.
  - Sender 1 distinguishes {{A1,A2}, {B1,B2}}. Payoff in matrix, aligned with receiver.

	State A1	**State A2**	**State B1**	**State B2**
Action 1	(3, 3)	(0, 0)	(0, 0)	**(1, 1)**
Action 2	(0, 0)	(3, 3)	**(1, 1)**	(0, 0)
Action 3	**(2, 2)**	**(2, 2)**	(0, 0)	(0, 0)

- Mechanism Γ perfectly reveals the information of Sender 1.
- However, if S1 and R share a *common language* they can do better.

Framework

- $n+1$ players. One Receiver (player 0), and n Senders.
- Receiver must take an action a from set A.
- Senders have a type $\theta_i \in \Theta_i$ drawn from $\mu \in \Delta(\Theta)$, where Θ is the Cartesian product of senders' types.
- Final payoffs of players given by

$$\pi_i : A \times \Theta \to R$$

Framework

- $n+1$ players. One Receiver (player 0), and n Senders.
- Receiver must take an action a from set A.
- Senders have a type $\theta_i \in \Theta_i$ drawn from $\mu \in \Delta(\Theta)$, where Θ is the Cartesian product of senders' types.
- Final payoffs of players given by

$$\pi_i : A \times \Theta \to R$$

Framework

- $n+1$ players. One Receiver (player 0), and n Senders.
- Receiver must take an action a from set A.
- Senders have a type $\theta_i \in \Theta_i$ drawn from $\mu \in \Delta(\Theta)$, where Θ is the Cartesian product of senders' types.
- Final payoffs of players given by

$$\pi_i : A \times \Theta \to R$$

Framework

- $n+1$ players. One Receiver (player 0), and n Senders.
- Receiver must take an action a from set A.
- Senders have a type $\theta_i \in \Theta_i$ drawn from $\mu \in \Delta(\Theta)$, where Θ is the Cartesian product of senders' types.
- Final payoffs of players given by

$$\pi_i : A \times \Theta \to R$$

Direct Revelation Mechanisms

Definition

a Direct Revelation Mechanism (DRM) is a mapping $\Gamma : \Theta \to \Delta(A)$.

- by the Revelation Principle, we can focus on DRM that are *obedient* and *truthful*.

Obedience

DRM Γ is *obedient* if the Receiver finds it optimal to follow the recommendation assuming truthful reporting.

Truthful

Γ is BIC if reporting truthfully is a BNE (assuming obedience).

Direct Revelation Mechanisms

Definition

a Direct Revelation Mechanism (DRM) is a mapping $\Gamma : \Theta \to \Delta(A)$.

- by the Revelation Principle, we can focus on DRM that are *obedient* and *truthful*.

Obedience

DRM Γ is *obedient* if the Receiver finds it optimal to follow the recommendation assuming truthful reporting.

Truthful

Γ is BIC if reporting truthfully is a BNE (assuming obedience).

Robustness

Confession

A *confession* given Γ is a tuple (i, T, τ) where $i \in I$, $T \subseteq \Theta_i$, $\tau : A \to A$.

Credible Confession

A confession given Γ is *credible* iff, given that all senders other than i report truthfully,

- Type θ_i benefits from the transformation τ iff $\theta_i \in T$.
- Let $\theta_i \in T$. There is an report $\hat{\theta}_i$
  - report is optimal for θ_i given τ.
  - τ is optimal for the receiver given $(\theta_i, \hat{\theta}_i)$.

NP Mechanisms

DRM Γ is Neologism Proof if there is no credible confession given Γ.

Robustness

Confession

A *confession* given Γ is a tuple (i, T, τ) where $i \in I$, $T \subseteq \Theta_i$, $\tau : A \to A$.

Credible Confession

A confession given Γ is *credible* iff, given that all senders other than i report truthfully,

- Type θ_i benefits from the transformation τ iff $\theta_i \in T$.
- Let $\theta_i \in T$. There is an report $\hat{\theta}_i$
  - report is optimal for θ_i given τ.
  - τ is optimal for the receiver given $(\theta_i, \hat{\theta}_i)$.

NP Mechanisms

DRM Γ is Neologism Proof if there is no credible confession given Γ.

Robustness

Confession

A *confession* given Γ is a tuple (i, T, τ) where $i \in I$, $T \subseteq \Theta_i$, $\tau : A \to A$.

Credible Confession

A confession given Γ is *credible* iff, given that all senders other than i report truthfully,

- Type θ_i benefits from the transformation τ iff $\theta_i \in T$.
- Let $\theta_i \in T$. There is an report $\hat{\theta}_i$
  - report is optimal for θ_i given τ.
  - τ is optimal for the receiver given $(\theta_i, \hat{\theta}_i)$.

NP Mechanisms

DRM Γ is Neologism Proof if there is no credible confession given Γ.

Robustness

Confession

A *confession* given Γ is a tuple (i, T, τ) where $i \in I$, $T \subseteq \Theta_i$, $\tau : A \to A$.

Credible Confession

A confession given Γ is *credible* iff, given that all senders other than i report truthfully,

- Type θ_i benefits from the transformation τ iff $\theta_i \in T$.
- Let $\theta_i \in T$. There is an report $\hat{\theta}_i$
  - report is optimal for θ_i given τ.
  - τ is optimal for the receiver given $(\theta_i, \hat{\theta}_i)$.

NP Mechanisms

DRM Γ is Neologism Proof if there is no credible confession given Γ.

Robustness

Confession

A *confession* given Γ is a tuple (i, T, τ) where $i \in I$, $T \subseteq \Theta_i$, $\tau : A \to A$.

Credible Confession

A confession given Γ is *credible* iff, given that all senders other than i report truthfully,

- Type θ_i benefits from the transformation τ iff $\theta_i \in T$.
- Let $\theta_i \in T$. There is an report $\hat{\theta}_i$
  - report is optimal for θ_i given τ.
  - τ is optimal for the receiver given $(\theta_i, \hat{\theta}_i)$.

NP Mechanisms

DRM Γ is Neologism Proof if there is no credible confession given Γ.

Application: Experts with Career Concerns

- Payoff-relevant state: $\omega \in \{\text{Left}, \text{Right}\}$.
- Receiver must take an action $a \in \{\text{Left}, \text{Right}\}$.
- Two senders (experts) observe conditionally independent signals $s_i \in \{\text{Left}, \text{Right}\}$.
- *Expertise* of the experts:
  - **Good expert**: Signal matches the state with probability q_H.
  - **Bad expert**: Signal matches the state with probability q_L.
  - There is exactly one good expert, but the identity $\theta \in \{1, 2\}$ is unknown to the receiver.
  - Ex-ante identical probabilities.

- The Receiver must nominate one of the experts for a promotion.

Application: Experts with Career Concerns

- Payoff-relevant state: $\omega \in \{\text{Left}, \text{Right}\}$.
- Receiver must take an action $a \in \{\text{Left}, \text{Right}\}$.
- Two senders (experts) observe conditionally independent signals $s_i \in \{\text{Left}, \text{Right}\}$.
- *Expertise* of the experts:
  - **Good expert**: Signal matches the state with probability q_H.
  - **Bad expert**: Signal matches the state with probability q_L.
  - There is exactly one good expert, but the identity $\theta \in \{1, 2\}$ is unknown to the receiver.
  - Ex-ante identical probabilities.

- The Receiver must nominate one of the experts for a promotion.

Application: Experts with Career Concerns

- Payoff-relevant state: $\omega \in \{\text{Left}, \text{Right}\}$.
- Receiver must take an action $a \in \{\text{Left}, \text{Right}\}$.
- Two senders (experts) observe conditionally independent signals $s_i \in \{\text{Left}, \text{Right}\}$.
- *Expertise* of the experts:
  - **Good expert**: Signal matches the state with probability q_H.
  - **Bad expert**: Signal matches the state with probability q_L.
  - There is exactly one good expert, but the identity $\theta \in \{1, 2\}$ is unknown to the receiver.
  - Ex-ante identical probabilities.
- The Receiver must nominate one of the experts for a promotion.

Application: Experts with Career Concerns

- Payoff-relevant state: $\omega \in \{\text{Left}, \text{Right}\}$.
- Receiver must take an action $a \in \{\text{Left}, \text{Right}\}$.
- Two senders (experts) observe conditionally independent signals $s_i \in \{\text{Left}, \text{Right}\}$.
- *Expertise* of the experts:
  - **Good expert**: Signal matches the state with probability q_H.
  - **Bad expert**: Signal matches the state with probability q_L.
  - There is exactly one good expert, but the identity $\theta \in \{1, 2\}$ is unknown to the receiver.
  - Ex-ante identical probabilities.

- The Receiver must nominate one of the experts for a promotion.

Application: Experts with Career Concerns

- Payoff-relevant state: $\omega \in \{\text{Left}, \text{Right}\}$.
- Receiver must take an action $a \in \{\text{Left}, \text{Right}\}$.
- Two senders (experts) observe conditionally independent signals $s_i \in \{\text{Left}, \text{Right}\}$.
- *Expertise* of the experts:
  - **Good expert**: Signal matches the state with probability q_H.
  - **Bad expert**: Signal matches the state with probability q_L.
  - There is exactly one good expert, but the identity $\theta \in \{1, 2\}$ is unknown to the receiver.
  - Ex-ante identical probabilities.

- The Receiver must nominate one of the experts for a promotion.

Application: Payoffs

- Experts and DM want the action to match the state.
- The expert that is promoted obtains a bonus B.
- The DM prefers to promote the good expert.

$$\pi_0 = 1_{\{a=\omega\}} + C \cdot 1_{\{m=\theta\}}$$

$$\pi_i = 1_{\{a=\omega\}} + B \cdot 1_{\{m=i\}}$$

Application: Payoffs

- Experts and DM want the action to match the state.
- The expert that is promoted obtains a bonus B.
- The DM prefers to promote the good expert.

$$\pi_0 = 1_{\{a=\omega\}} + C \cdot 1_{\{m=\theta\}}$$
$$\pi_i = 1_{\{a=\omega\}} + B \cdot 1_{\{m=i\}}$$

Communication Design to Match the State

- Optimal action given experts' information: s_θ.
  - Probability of matching the state with optimal action: q_H.
- Is it possible to implement this action?

- **DRM**: experts report their information $(\hat{\theta}_i, \hat{s}_i)$, and the mechanism recommends an action $\hat{a}$ and an expert $\hat{i}$ to promote.

Family of DRM and IC

- Consider the following family of DRM:
  - If reports coincide, $\hat{\theta}_1 = \hat{\theta}_2$,
    - mechanism recommends action $\hat{s}_{\hat{\theta}}$.
    - mechanism recommends to promote $\hat{\theta}$ with probability $g \geq 1/2$.
  - If reports do not coincide in the good expert, mechanism recommends
    - action equal to the reported signals when these coincide.
    - a random action when reported signals don't coincide.
    - a random promotion.

- Let p_L be the probability of matching the state when signals are aggregated with the same weight. The bad expert reports truthfully if

$$q_H + (1-g)\cdot B \geq p_L + \frac{1}{2}\cdot B \qquad \Rightarrow \qquad g \leq \frac{1}{2} + \frac{q_H - p_L}{B}$$

Family of DRM and IC

- Consider the following family of DRM:
  - If reports coincide, $\hat{\theta}_1 = \hat{\theta}_2$,
    - mechanism recommends action $\hat{s}_{\hat{\theta}}$.
    - mechanism recommends to promote $\hat{\theta}$ with probability $g \geq 1/2$.
  - If reports do not coincide in the good expert, mechanism recommends
    - action equal to the reported signals when these coincide.
    - a random action when reported signals don't coincide.
    - a random promotion.

- Let p_L be the probability of matching the state when signals are aggregated with the same weight. The bad expert reports truthfully if

$$q_H + (1-g)\cdot B \geq p_L + \frac{1}{2}\cdot B \qquad \Rightarrow \qquad g \leq \frac{1}{2} + \frac{q_H - p_L}{B}$$

Family of DRM and IC

- Consider the following family of DRM:
  - If reports coincide, $\hat{\theta}_1 = \hat{\theta}_2$,
    - mechanism recommends action $\hat{s}_{\hat{\theta}}$.
    - mechanism recommends to promote $\hat{\theta}$ with probability $g \geq 1/2$.
  - If reports do not coincide in the good expert, mechanism recommends
    - action equal to the reported signals when these coincide.
    - a random action when reported signals don't coincide.
    - a random promotion.

- Let p_L be the probability of matching the state when signals are aggregated with the same weight. The bad expert reports truthfully if

$$q_H + (1-g)\cdot B \geq p_L + \frac{1}{2}\cdot B \qquad \Rightarrow \qquad g \leq \frac{1}{2} + \frac{q_H - p_L}{B}$$

Family of DRM and IC

- Consider the following family of DRM:
  - If reports coincide, $\hat{\theta}_1 = \hat{\theta}_2$,
    - mechanism recommends action $\hat{s}_{\hat{\theta}}$.
    - mechanism recommends to promote $\hat{\theta}$ with probability $g \geq 1/2$.
  - If reports do not coincide in the good expert, mechanism recommends
    - action equal to the reported signals when these coincide.
    - a random action when reported signals don't coincide.
    - a random promotion.

- Let p_L be the probability of matching the state when signals are aggregated with the same weight. The bad expert reports truthfully if

$$q_H + (1-g) \cdot B \geq p_L + \frac{1}{2} \cdot B \qquad \Rightarrow \quad g \leq \frac{1}{2} + \frac{q_H - p_L}{B}$$

Good Expert Might Sabotage the Mechanism

Consider the following deviation of the good expert i which obtains $s_i = A$.

- Instead of sending report (i, A), he sends (i, B).
- He approaches the DM and says: "I am the good expert and gave an incorrect report, thus you should not follow the recommended action from the mechanism. By the way, there is no strategy of the bad expert that benefits from you switching the action, so you should trust that I'm the good expert."

Good Expert Might Sabotage the Mechanism

Consider the following deviation of the good expert i which obtains $s_i = A$.

- Instead of sending report (i, A), he sends (i, B).
- He approaches the DM and says: "I am the good expert and gave an incorrect report, thus you should not follow the recommended action from the mechanism. By the way, there is no strategy of the bad expert that benefits from you switching the action, so you should trust that I'm the good expert."

Good Expert Might Sabotage the Mechanism

- Is it beneficial for the good expert? **YES.**

$$q_H + B \qquad \text{vs} \qquad q_H + g \cdot B$$

- Is it be beneficial for the bad expert? **Depends.**

$$(1 - \underline{p}) + B \qquad \text{vs} \qquad q_H + (1 - g) \cdot B$$

- $\underline{p}$: minimum probability of having the mechanism recommend the correct action that can be induced by the bad expert.
- Beneficial iff $g > \frac{q_H - (1 - \underline{p})}{B}$

Good Expert Might Sabotage the Mechanism

- Is it beneficial for the good expert? **YES.**

$$q_H + B \qquad \text{vs} \qquad q_H + g \cdot B$$

- Is it be beneficial for the bad expert? **Depends.**

$$(1 - \underline{p}) + B \qquad \text{vs} \qquad q_H + (1 - g) \cdot B$$

- $\underline{p}$: minimum probability of having the mechanism recommend the correct action that can be induced by the bad expert.
- Beneficial iff $g > \frac{q_H - (1 - \underline{p})}{B}$

When can the DM take optimal actions

Proposition

The optimal recommendation can be implemented with a NP mechanism iff the career concerns are not too high:
$B \leq 2 \cdot (p_L - (1 - \underline{p}))$.

- We just need that it exists a g small enough so that bad expert want to report his expertise truthfully, but high enough so that bad expert would like to sabotage the mechanism.

$$\frac{q_H - (1 - \underline{p})}{B} > \frac{1}{2} + \frac{q_H - p_L}{B}$$

When can the DM take optimal actions

Proposition

The optimal recommendation can be implemented with a NP mechanism iff the career concerns are not too high:
$B \leq 2 \cdot (p_L - (1 - \underline{p}))$.

- We just need that it exists a g small enough so that bad expert want to report his expertise truthfully, but high enough so that bad expert would like to sabotage the mechanism.

$$\frac{q_H - (1 - \underline{p})}{B} > \frac{1}{2} + \frac{q_H - p_L}{B}$$

Conclusion

- When mechanism participants have a common language, mechanism designers have to account for the incentives to confess deviations.
- We study the problem of designing communication mechanisms to elicit experts' information when
  - Experts have career concerns.
  - Experts share a common language with the DM and can communicate outside of the mechanism.
- We find that
  - To induce an optimal action, a mechanism must aggregates the experts recommendation, and not promote the good expert too often. However, good experts have incentives to sabotage the mechanism in an attempt to signal their type.
  - When career concerns are sufficiently high, it is not possible to implement the optimal action in a way that is robust to out-of-mechanism communication.