Robust scoring rules

Elias Tsakas

Maastricht University

August 2018 / EEA-ESEM, Cologne

Roadmap

Roadmap

Background

General problem: How to elicit latent subjective beliefs?

- **Usual answer:** Proper scoring rules.
- **Methodological problem (Heisenberg):** Monetary incentives (provided by the scoring rule) may affect the beliefs.

Background

General problem: How to elicit latent subjective beliefs?

- **Usual answer:** Proper scoring rules.
- **Methodological problem (Heisenberg):** Monetary incentives (provided by the scoring rule) may affect the beliefs.

Background

General problem: How to elicit latent subjective beliefs?

- **Usual answer:** Proper scoring rules.
- **Methodological problem (Heisenberg):** Monetary incentives (provided by the scoring rule) may affect the beliefs.

Example: Eliciting population beliefs

- Three steps to estimate the population beliefs:
  1. draw a representative sample from the population,
  2. elicit individual beliefs from each subject in the sample,
  3. uses frequency of elicited beliefs as an estimate for population.
- **Problem:** Subjects in the sample respond to the incentives we provide them, by putting extra thought to their answer.
- Thus, we obtain biased estimate of population beliefs.

Example: Eliciting population beliefs

- Three steps to estimate the population beliefs:
  1. draw a representative sample from the population,
  2. elicit individual beliefs from each subject in the sample,
  3. uses frequency of elicited beliefs as an estimate for population.
- **Problem:** Subjects in the sample respond to the incentives we provide them, by putting extra thought to their answer.
- Thus, we obtain biased estimate of population beliefs.

Example: Eliciting population beliefs

- Three steps to estimate the population beliefs:
  1. draw a representative sample from the population,
  2. elicit individual beliefs from each subject in the sample,
  3. uses frequency of elicited beliefs as an estimate for population.
- **Problem:** Subjects in the sample respond to the incentives we provide them, by putting extra thought to their answer.
- Thus, we obtain biased estimate of population beliefs.

Research question and main result

- **Question:** Can we elicit the beliefs that the subject would have had, *if we had not asked him?*
- **Main Theorem:** YES, under standard assumptions!

Research question and main result

- **Question:** Can we elicit the beliefs that the subject would have had, *if we had not asked him?*
- **Main Theorem:** YES, under standard assumptions!

Literature(s)

- **Scoring rules**
  - Early contributions: Brier (1950), Good (1952), Savage (1971)
  - Recent theoretical contributions: Karni (1999, 2009), Chambers (2008), Offerman et al. (2009), Hossain & Okui (2013), Chambers & Lambert (2017), Chambers, Healy & Lambert (2017), ...
  - Recent experimental contributions: Selten et al. (1999), Nyarko & Schotter (2002), Offerman et al. (2009), Armantier & Treich (2013), ...
  - Review articles: Schlag et al. (2013), Schotter & Trevino (2014)
  - Applications: Thomson (1979), Hanson (2003), Ostrovsky (2012), ...
  - Other disciplines: CS, education, finance, medicine, politics, ...
- **Rational inattention**
  - Early contributions: Sims (2003, 2006)
  - Axiomatic foundations: De Oliveira et al. (2017), Ellis (2018)
  - Revealed preference: Caplin & Dean (2015) Chambers, Liu & Rehbeck (2017), Caplin et al. (2017)
  - Experimental contributions: Caplin et al. (2010), Dean & Neligh (2017)
  - Applications: Bartoš et al. (2016), Matejka (2016), Matejka & Tabellini (2016), ...

Roadmap

Preliminaries

FUNDAMENTALS:

- Binary state space: $\Omega = \{\omega_0, \omega_1\}$
- Latent subjective belief (of ω_0 occurring): $\mu \in [0, 1]$
- (Non-verifiable) self-report: $r \in [0, 1]$

ELICITATION MECHANISM:

- Scoring rule: $\boxed{S : [0, 1] \times \Omega \to \mathbb{R}}$
- Payment depends on self-report and state realization.
- Payment is in monetary payoffs.

Preliminaries

FUNDAMENTALS:

- Binary state space: $\Omega = \{\omega_0, \omega_1\}$
- Latent subjective belief (of ω_0 occurring): $\mu \in [0, 1]$
- (Non-verifiable) self-report: $r \in [0, 1]$

ELICITATION MECHANISM:

- Scoring rule: $\boxed{S : [0, 1] \times \Omega \to \mathbb{R}}$
- Payment depends on self-report and state realization.
- Payment is in monetary payoffs.

Preliminaries

FUNDAMENTALS:

- Binary state space: $\Omega = \{\omega_0, \omega_1\}$
- Latent subjective belief (of ω_0 occurring): $\mu \in [0, 1]$
- (Non-verifiable) self-report: $r \in [0, 1]$

ELICITATION MECHANISM:

- Scoring rule: $\boxed{S : [0, 1] \times \Omega \to \mathbb{R}}$
- Payment depends on self-report and state realization.
- Payment is in monetary payoffs.

Preliminaries

FUNDAMENTALS:

- Binary state space: $\Omega = \{\omega_0, \omega_1\}$
- Latent subjective belief (of ω_0 occurring): $\mu \in [0, 1]$
- (Non-verifiable) self-report: $r \in [0, 1]$

ELICITATION MECHANISM:

- Scoring rule: $\boxed{S : [0, 1] \times \Omega \to \mathbb{R}}$
- Payment depends on self-report and state realization.
- Payment is in monetary payoffs.

Preliminaries

FUNDAMENTALS:

- Binary state space: $\Omega = \{\omega_0, \omega_1\}$
- Latent subjective belief (of ω_0 occurring): $\mu \in [0, 1]$
- (Non-verifiable) self-report: $r \in [0, 1]$

ELICITATION MECHANISM:

- Scoring rule: $\boxed{S : [0, 1] \times \Omega \to \mathbb{R}}$
- Payment depends on self-report and state realization.
- Payment is in monetary payoffs.

Preliminaries

FUNDAMENTALS:

- Binary state space: $\Omega = \{\omega_0, \omega_1\}$
- Latent subjective belief (of ω_0 occurring): $\mu \in [0, 1]$
- (Non-verifiable) self-report: $r \in [0, 1]$

ELICITATION MECHANISM:

- Scoring rule: $\boxed{S : [0, 1] \times \Omega \to \mathbb{R}}$
- Payment depends on self-report and state realization.
- Payment is in monetary payoffs.

Preliminaries

FUNDAMENTALS:

- Binary state space: $\Omega = \{\omega_0, \omega_1\}$
- Latent subjective belief (of ω_0 occurring): $\mu \in [0, 1]$
- (Non-verifiable) self-report: $r \in [0, 1]$

ELICITATION MECHANISM:

- Scoring rule: $\boxed{S : [0, 1] \times \Omega \to \mathbb{R}}$
- Payment depends on self-report and state realization.
- Payment is in monetary payoffs.

Preliminaries

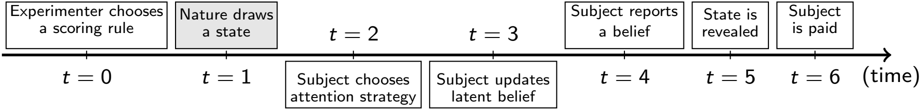
FUNDAMENTALS:

- Binary state space: $\Omega = \{\omega_0, \omega_1\}$
- Latent subjective belief (of ω_0 occurring): $\mu \in [0,1]$
- (Non-verifiable) self-report: $r \in [0,1]$

ELICITATION MECHANISM:

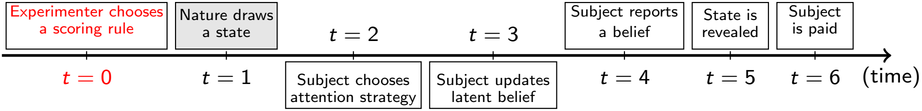
- Scoring rule: $\boxed{S : [0,1] \times \Omega \to \mathbb{R}}$
- Payment depends on self-report and state realization.
- Payment is in monetary payoffs.

Purpose and timeline of our mechanism



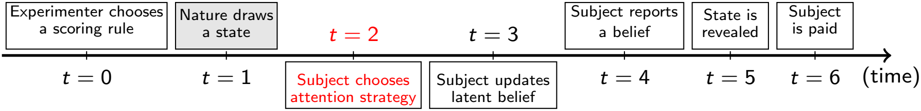
- The experimenter wants to design a scoring rule, such that:
  1. the subject rationally does not update his beliefs, and
  2. the subject reports truthfully.
- We will call this, a robust scoring rule.

Purpose and timeline of our mechanism



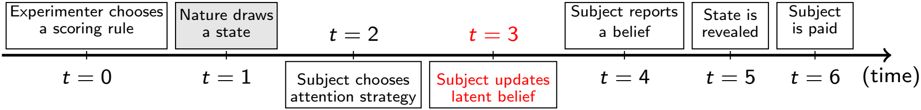
- The experimenter wants to design a scoring rule, such that:
  1. the subject rationally does not update his beliefs, and
  2. the subject reports truthfully.
- We will call this, a robust scoring rule.

Purpose and timeline of our mechanism



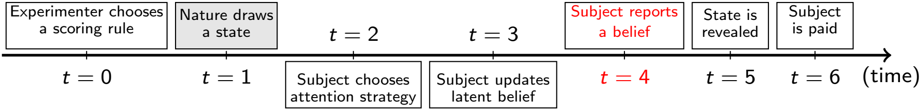
- The experimenter wants to design a scoring rule, such that:
  1. the subject rationally does not update his beliefs, and
  2. the subject reports truthfully.
- We will call this, a robust scoring rule.

Purpose and timeline of our mechanism



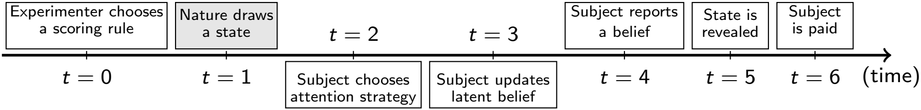
- The experimenter wants to design a scoring rule, such that:
  1. the subject rationally does not update his beliefs, and
  2. the subject reports truthfully.
- We will call this, a robust scoring rule.

Purpose and timeline of our mechanism



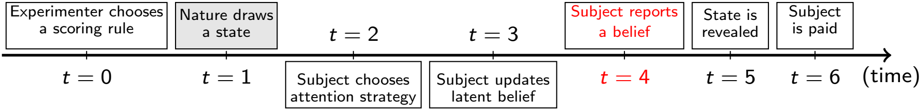
- The experimenter wants to design a scoring rule, such that:
  1. the subject rationally does not update his beliefs, and
  2. the subject reports truthfully.
- We will call this, a robust scoring rule.

Purpose and timeline of our mechanism



- The experimenter wants to design a scoring rule, such that:
  1. the subject rationally does not update his beliefs, and
  2. the subject reports truthfully.
- We will call this, a robust scoring rule.

Proper scoring rules



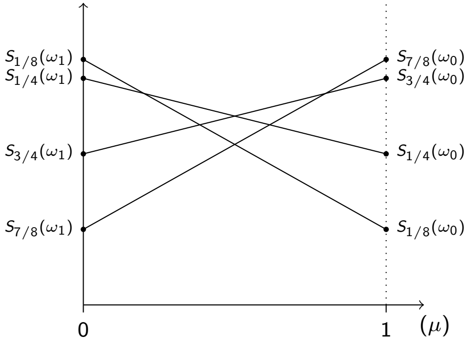
- Begin with the second step.

Proper scoring rules

- Strictly dominant to report truthfully irrespective of the beliefs

$$\boxed{\mathbb{E}_\mu(S_\mu) > \mathbb{E}_\mu(S_r) \text{ for all } r \neq \mu \text{ and all } \mu \in [0,1]}$$

- Expected utility from truthfully reporting: $\phi(\mu) := \mathbb{E}_\mu(S_\mu)$
- **Proper scoring rule**: ϕ is strictly convex and subdifferentiable.

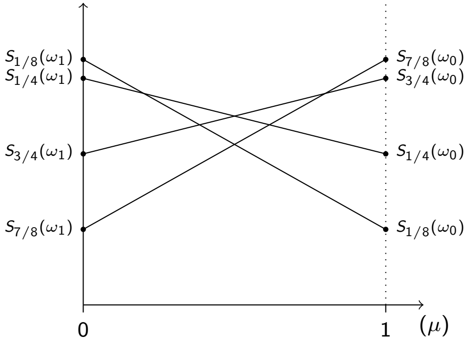


Proper scoring rules

- Strictly dominant to report truthfully irrespective of the beliefs

$$\mathbb{E}_\mu(S_\mu) > \mathbb{E}_\mu(S_r) \text{ for all } r \neq \mu \text{ and all } \mu \in [0,1]$$

- Expected utility from truthfully reporting: $\phi(\mu) := \mathbb{E}_\mu(S_\mu)$
- **Proper scoring rule**: ϕ is strictly convex and subdifferentiable.

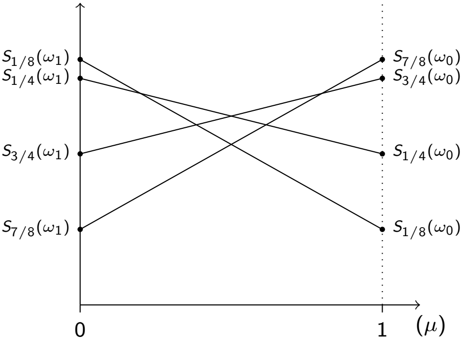


Proper scoring rules

- Strictly dominant to report truthfully irrespective of the beliefs

$$\mathbb{E}_\mu(S_\mu) > \mathbb{E}_\mu(S_r) \text{ for all } r \neq \mu \text{ and all } \mu \in [0,1]$$

- Expected utility from truthfully reporting: $\phi(\mu) := \mathbb{E}_\mu(S_\mu)$
- **Proper scoring rule**: ϕ is strictly convex and subdifferentiable.

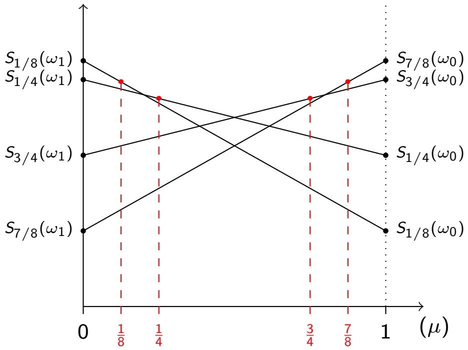


Proper scoring rules

- Strictly dominant to report truthfully irrespective of the beliefs

$$\boxed{\mathbb{E}_\mu(S_\mu) > \mathbb{E}_\mu(S_r) \text{ for all } r \neq \mu \text{ and all } \mu \in [0,1]}$$

- Expected utility from truthfully reporting: $\phi(\mu) := \mathbb{E}_\mu(S_\mu)$
- **Proper scoring rule**: ϕ is strictly convex and subdifferentiable.

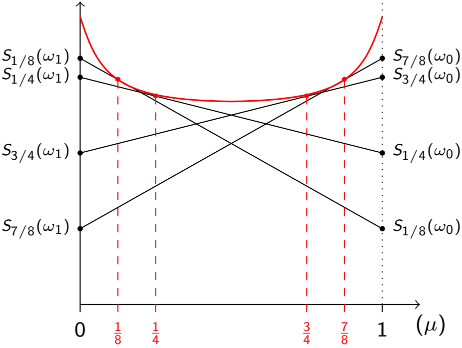


Proper scoring rules

- Strictly dominant to report truthfully irrespective of the beliefs

$$\boxed{\mathbb{E}_\mu(S_\mu) > \mathbb{E}_\mu(S_r) \text{ for all } r \neq \mu \text{ and all } \mu \in [0,1]}$$

- Expected utility from truthfully reporting: $\phi(\mu) := \mathbb{E}_\mu(S_\mu)$
- **Proper scoring rule**: ϕ is strictly convex and subdifferentiable.

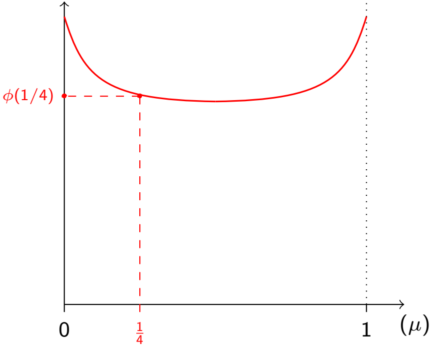


Proper scoring rules

- Strictly dominant to report truthfully irrespective of the beliefs

$$\boxed{\mathbb{E}_\mu(S_\mu) > \mathbb{E}_\mu(S_r) \text{ for all } r \neq \mu \text{ and all } \mu \in [0,1]}$$

- Expected utility from truthfully reporting: $\phi(\mu) := \mathbb{E}_\mu(S_\mu)$
- **Proper scoring rule**: ϕ is strictly convex and subdifferentiable.

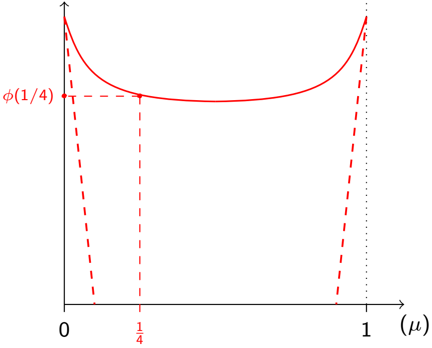


Proper scoring rules

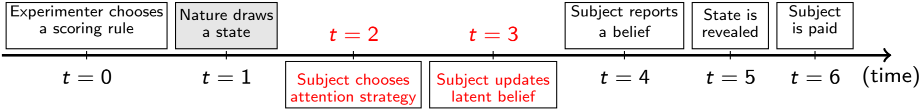
- Strictly dominant to report truthfully irrespective of the beliefs

$$\boxed{\mathbb{E}_\mu(S_\mu) > \mathbb{E}_\mu(S_r) \text{ for all } r \neq \mu \text{ and all } \mu \in [0,1]}$$

- Expected utility from truthfully reporting: $\phi(\mu) := \mathbb{E}_\mu(S_\mu)$
- **Proper scoring rule**: ϕ is strictly convex and subdifferentiable.



Attention



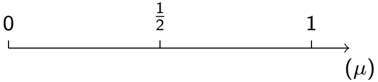
- Continue with the first step.

Attention strategies

- An **attention strategy** is modelled with a signal, $\sigma : \Omega \to \Delta(S)$.
- It is in fact a cognitive experiment by the subject.
- Given a prior $\mu \in [0, 1]$, each feasible attention strategy is characterized by a unique distribution of posteriors:

$$\boxed{\pi \in \Delta([0,1]) \text{ such that } \mu = \mathbb{E}_\pi(\nu)}.$$

- The set of feasible attention strategies is denoted by $\Pi(\mu)$.
- Important special cases:
  - No-attention strategy: $\hat{\mu} \in \Pi(\mu)$ puts probability 1 to μ.
  - Perfectly informative strategy: $\pi \in \Pi(\mu)$ puts probability 1 to $\{0, 1\}$.
- Attention has benefits and costs.

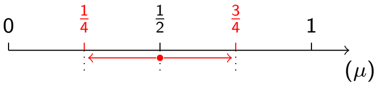


Attention strategies

- An **attention strategy** is modelled with a signal, $\sigma : \Omega \to \Delta(S)$.
- It is in fact a cognitive experiment by the subject.
- Given a prior $\mu \in [0, 1]$, each feasible attention strategy is characterized by a unique distribution of posteriors:

$$\boxed{\pi \in \Delta([0,1]) \text{ such that } \mu = \mathbb{E}_\pi(\nu)}.$$

- The set of feasible attention strategies is denoted by $\Pi(\mu)$.
- Important special cases:
  - No-attention strategy: $\hat{\mu} \in \Pi(\mu)$ puts probability 1 to μ.
  - Perfectly informative strategy: $\pi \in \Pi(\mu)$ puts probability 1 to $\{0, 1\}$.
- Attention has benefits and costs.

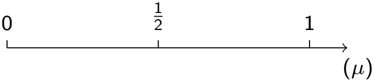


Attention strategies

- An **attention strategy** is modelled with a signal, $\sigma : \Omega \to \Delta(S)$.
- It is in fact a cognitive experiment by the subject.
- Given a prior $\mu \in [0,1]$, each feasible attention strategy is characterized by a unique distribution of posteriors:

$$\boxed{\pi \in \Delta([0,1]) \text{ such that } \mu = \mathbb{E}_\pi(\nu)}.$$

- The set of feasible attention strategies is denoted by $\Pi(\mu)$.
- Important special cases:
  - No-attention strategy: $\hat{\mu} \in \Pi(\mu)$ puts probability 1 to μ.
  - Perfectly informative strategy: $\pi \in \Pi(\mu)$ puts probability 1 to $\{0,1\}$.
- Attention has benefits and costs.

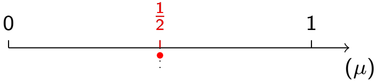


Attention strategies

- An **attention strategy** is modelled with a signal, $\sigma : \Omega \to \Delta(S)$.
- It is in fact a cognitive experiment by the subject.
- Given a prior $\mu \in [0,1]$, each feasible attention strategy is characterized by a unique distribution of posteriors:

$$\boxed{\pi \in \Delta([0,1]) \text{ such that } \mu = \mathbb{E}_\pi(\nu)}.$$

- The set of feasible attention strategies is denoted by $\Pi(\mu)$.
- Important special cases:
  - No-attention strategy: $\hat{\mu} \in \Pi(\mu)$ puts probability 1 to μ.
  - Perfectly informative strategy: $\pi \in \Pi(\mu)$ puts probability 1 to $\{0,1\}$.
- Attention has benefits and costs.

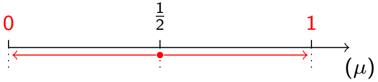


Attention strategies

- An **attention strategy** is modelled with a signal, $\sigma : \Omega \to \Delta(S)$.
- It is in fact a cognitive experiment by the subject.
- Given a prior $\mu \in [0,1]$, each feasible attention strategy is characterized by a unique distribution of posteriors:

$$\boxed{\pi \in \Delta([0,1]) \text{ such that } \mu = \mathbb{E}_\pi(\nu)}.$$

- The set of feasible attention strategies is denoted by $\Pi(\mu)$.
- Important special cases:
  - No-attention strategy: $\hat{\mu} \in \Pi(\mu)$ puts probability 1 to μ.
  - Perfectly informative strategy: $\pi \in \Pi(\mu)$ puts probability 1 to $\{0,1\}$.
- Attention has benefits and costs.

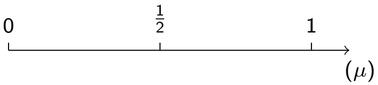


Attention strategies

- An **attention strategy** is modelled with a signal, $\sigma : \Omega \to \Delta(S)$.
- It is in fact a cognitive experiment by the subject.
- Given a prior $\mu \in [0, 1]$, each feasible attention strategy is characterized by a unique distribution of posteriors:

$$\boxed{\pi \in \Delta([0,1]) \text{ such that } \mu = \mathbb{E}_\pi(\nu)}.$$

- The set of feasible attention strategies is denoted by $\Pi(\mu)$.
- Important special cases:
  - No-attention strategy: $\hat{\mu} \in \Pi(\mu)$ puts probability 1 to μ.
  - Perfectly informative strategy: $\pi \in \Pi(\mu)$ puts probability 1 to $\{0, 1\}$.
- Attention has benefits and costs.

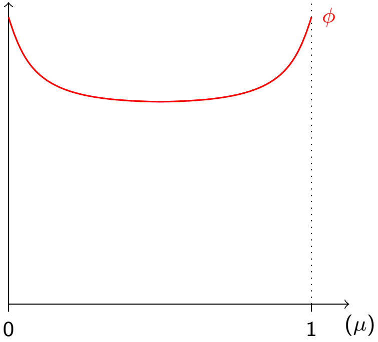


Benefit of attention

- Fix a proper scoring rule ϕ.
- For prior μ and attention $\pi \in \Pi(\mu)$, the expected benefit is

$$\boxed{B_\phi(\pi) = \langle \phi, \pi \rangle - \phi(\mu)}$$

- Every attention strategy yields a strictly positive expected benefit

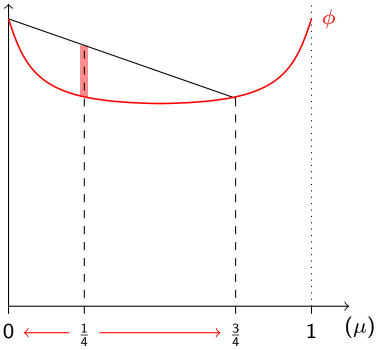


Benefit of attention

- Fix a proper scoring rule ϕ.
- For prior μ and attention $\pi \in \Pi(\mu)$, the expected benefit is

$$B_\phi(\pi) = \langle \phi, \pi \rangle - \phi(\mu)$$

- Every attention strategy yields a strictly positive expected benefit

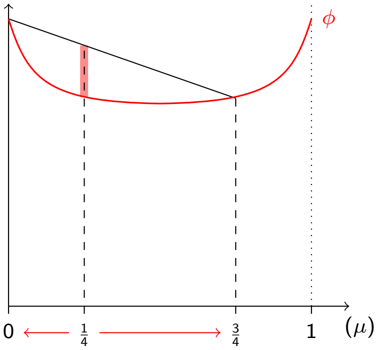


Benefit of attention

- Fix a proper scoring rule ϕ.
- For prior μ and attention $\pi \in \Pi(\mu)$, the expected benefit is

$$\boxed{B_\phi(\pi) = \langle \phi, \pi \rangle - \phi(\mu)}$$

- Every attention strategy yields a strictly positive expected benefit



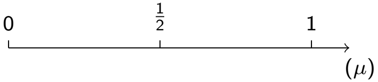
Cost of attention

$$\boxed{C : \Delta([0,1]) \to \mathbb{R}_+}$$

- Our **regularity assumptions**:
  - (C_1) NORMALIZATION: $C(\hat{\mu}) = 0$ for all $\mu \in [0,1]$
  - (C_2) ATTENTION IS COSTLY: $C(\pi) > 0$ for all $\pi \in \hat{\Pi}(\mu) := \Pi(\mu) \setminus \{\hat{\mu}\}$ and all $\mu \in [0,1]$
  - (C_3) DYNAMIC CONSISTENCY: If $\sigma : [0,1] \to \Delta([0,1])$ satisfies $\sigma(\mu) \in \Pi(\mu)$ for all $\mu \in [0,1]$, then

$$C(\mathbb{E}_\pi(\sigma)) = C(\pi) + \mathbb{E}_\pi(C \circ \sigma) \tag{1}$$

    for all $\pi \in \Delta([0,1])$.
- **Interpretation:** New information is always costly and the order of information does not matter (additive separability)



Cost of attention

$$\boxed{C : \Delta([0,1]) \to \mathbb{R}_+}$$

- Our **regularity assumptions**:

  (C_1) NORMALIZATION: $C(\hat{\mu}) = 0$ for all $\mu \in [0,1]$

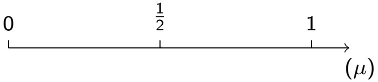
  (C_2) ATTENTION IS COSTLY: $C(\pi) > 0$ for all $\pi \in \hat{\Pi}(\mu) := \Pi(\mu) \setminus \{\hat{\mu}\}$ and all $\mu \in [0,1]$

  (C_3) DYNAMIC CONSISTENCY: If $\sigma : [0,1] \to \Delta([0,1])$ satisfies $\sigma(\mu) \in \Pi(\mu)$ for all $\mu \in [0,1]$, then

  $$C(\mathbb{E}_\pi(\sigma)) = C(\pi) + \mathbb{E}_\pi(C \circ \sigma) \tag{1}$$

  for all $\pi \in \Delta([0,1])$.

- **Interpretation:** New information is always costly and the order of information does not matter (additive separability)



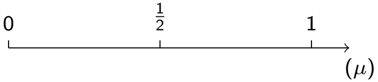
Cost of attention

$$\boxed{C : \Delta([0,1]) \to \mathbb{R}_+}$$

- Our **regularity assumptions**:
  - (C_1) NORMALIZATION: $C(\hat{\mu}) = 0$ for all $\mu \in [0,1]$
  - (C_2) ATTENTION IS COSTLY: $C(\pi) > 0$ for all $\pi \in \hat{\Pi}(\mu) := \Pi(\mu) \setminus \{\hat{\mu}\}$ and all $\mu \in [0,1]$
  - (C_3) DYNAMIC CONSISTENCY: If $\sigma : [0,1] \to \Delta([0,1])$ satisfies $\sigma(\mu) \in \Pi(\mu)$ for all $\mu \in [0,1]$, then

$$C(\mathbb{E}_\pi(\sigma)) = C(\pi) + \mathbb{E}_\pi(C \circ \sigma) \qquad (1)$$

    for all $\pi \in \Delta([0,1])$.
- **Interpretation:** New information is always costly and the order of information does not matter (additive separability)



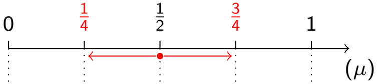
Cost of attention

$$\boxed{C : \Delta([0,1]) \to \mathbb{R}_+}$$

- Our **regularity assumptions**:
  - (C_1) NORMALIZATION: $C(\hat{\mu}) = 0$ for all $\mu \in [0,1]$
  - (C_2) ATTENTION IS COSTLY: $C(\pi) > 0$ for all $\pi \in \hat{\Pi}(\mu) := \Pi(\mu) \setminus \{\hat{\mu}\}$ and all $\mu \in [0,1]$
  - (C_3) DYNAMIC CONSISTENCY: If $\sigma : [0,1] \to \Delta([0,1])$ satisfies $\sigma(\mu) \in \Pi(\mu)$ for all $\mu \in [0,1]$, then

$$C(\mathbb{E}_\pi(\sigma)) = C(\pi) + \mathbb{E}_\pi(C \circ \sigma) \tag{1}$$

    for all $\pi \in \Delta([0,1])$.
- **Interpretation:** New information is always costly and the order of information does not matter (additive separability)



Cost of attention

$$\boxed{C : \Delta([0,1]) \to \mathbb{R}_+}$$

- Our **regularity assumptions**:

  (C_1) NORMALIZATION: $C(\hat{\mu}) = 0$ for all $\mu \in [0,1]$

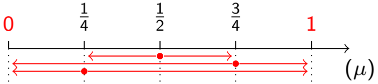
  (C_2) ATTENTION IS COSTLY: $C(\pi) > 0$ for all $\pi \in \hat{\Pi}(\mu) := \Pi(\mu) \setminus \{\hat{\mu}\}$ and all $\mu \in [0,1]$

  (C_3) DYNAMIC CONSISTENCY: If $\sigma : [0,1] \to \Delta([0,1])$ satisfies $\sigma(\mu) \in \Pi(\mu)$ for all $\mu \in [0,1]$, then

$$C(\mathbb{E}_\pi(\sigma)) = C(\pi) + \mathbb{E}_\pi(C \circ \sigma) \tag{1}$$

  for all $\pi \in \Delta([0,1])$.

- **Interpretation:** New information is always costly and the order of information does not matter (additive separability)



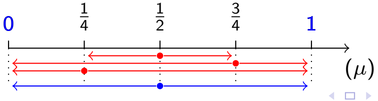
Cost of attention

$$\boxed{C : \Delta([0,1]) \to \mathbb{R}_+}$$

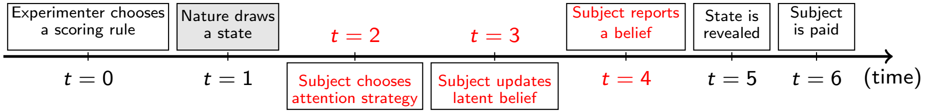
- Our **regularity assumptions**:
  - (C_1) NORMALIZATION: $C(\hat{\mu}) = 0$ for all $\mu \in [0,1]$
  - (C_2) ATTENTION IS COSTLY: $C(\pi) > 0$ for all $\pi \in \hat{\Pi}(\mu) := \Pi(\mu) \setminus \{\hat{\mu}\}$ and all $\mu \in [0,1]$
  - (C_3) DYNAMIC CONSISTENCY: If $\sigma : [0,1] \to \Delta([0,1])$ satisfies $\sigma(\mu) \in \Pi(\mu)$ for all $\mu \in [0,1]$, then

$$C(\mathbb{E}_\pi(\sigma)) = C(\pi) + \mathbb{E}_\pi(C \circ \sigma) \tag{1}$$

for all $\pi \in \Delta([0,1])$.
- **Interpretation:** New information is always costly and the order of information does not matter (additive separability)



Eliciting the prior belief



- Put the two steps together.

Robust scoring rule

- Value of attention given a scoring rule: $\boxed{V_\phi(\pi) := B_\phi(\pi) - C(\pi)}$
- For every prior, it is simultaneously strictly dominant to:
  1. not to update prior, and
  2. report truthfully.

$$\boxed{V_\phi(\hat{\mu}) > V_\phi(\pi) \text{ for all } \pi \in \hat{\Pi}(\mu), \text{ with } \phi \text{ proper}}$$

- Intuitively, "*ϕ must be strictly convex, but not too convex*".

Robust scoring rule

- Value of attention given a scoring rule: $\boxed{V_\phi(\pi) := B_\phi(\pi) - C(\pi)}$
- For every prior, it is simultaneously strictly dominant to:
  1. not to update prior, and
  2. report truthfully.

$$\boxed{V_\phi(\hat{\mu}) > V_\phi(\pi) \text{ for all } \pi \in \hat{\Pi}(\mu), \text{ with } \phi \text{ proper}}$$

- Intuitively, *"ϕ must be strictly convex, but not too convex"*.

Robust scoring rule

- Value of attention given a scoring rule: $\boxed{V_\phi(\pi) := B_\phi(\pi) - C(\pi)}$
- For every prior, it is simultaneously strictly dominant to:
  1. not to update prior, and
  2. report truthfully.

$$\boxed{V_\phi(\hat{\mu}) > V_\phi(\pi) \text{ for all } \pi \in \hat{\Pi}(\mu), \text{ with } \phi \text{ proper}}$$

- Intuitively, "*ϕ must be strictly convex, but not too convex*".

Robust scoring rule

- Value of attention given a scoring rule: $\boxed{V_\phi(\pi) := B_\phi(\pi) - C(\pi)}$
- For every prior, it is simultaneously strictly dominant to:
  1. not to update prior, and
  2. report truthfully.

$$\boxed{V_\phi(\hat{\mu}) > V_\phi(\pi) \text{ for all } \pi \in \hat{\Pi}(\mu), \text{ with } \phi \text{ proper}}$$

- Intuitively, *"ϕ must be strictly convex, but not too convex"*.

Roadmap

Main result

Theorem (Existence)

If the cost function satisfies $(C_1) - (C_3)$, there exists a robust scoring rule.

- Proof is constructive (at the end of presentation).
- Not necessarily true for quadratic scoring rules.

Proposition (QSR under entropic costs*)

If costs are entropic, there is a quadratic scoring rule.

Main result

Theorem (Existence)

If the cost function satisfies $(C_1) - (C_3)$, there exists a robust scoring rule.

- Proof is constructive (at the end of presentation).
- Not necessarily true for quadratic scoring rules.

Proposition (QSR under entropic costs*)

If costs are entropic, there is a quadratic scoring rule.

Main result

Theorem (Existence)

If the cost function satisfies $(C_1) - (C_3)$, there exists a robust scoring rule.

- Proof is constructive (at the end of presentation).
- Not necessarily true for quadratic scoring rules.

Proposition (QSR under entropic costs*)

If costs are entropic, there is a quadratic scoring rule.

Main result

Theorem (Existence)

If the cost function satisfies $(C_1) - (C_3)$, there exists a robust scoring rule.

- Proof is constructive (at the end of presentation).
- Not necessarily true for quadratic scoring rules.

Proposition (QSR under entropic costs*)

If costs are entropic, there is a quadratic scoring rule.

Additional results: Approximate robustness

- A scoring rule is ε-robust if it elicits a belief within ε from the prior.
- Approximate robustness allows us to use "simple scoring rules".

Proposition (Approximately robust QSR under polynomial costs*)

If costs satisfy $(C_1) - (C_3)$ and are polynomial, then for every $\varepsilon > 0$ there is an ε-robust quadratic scoring rule.

Proposition (Approximately robust discrete scoring rules)

If costs satisfy $(C_1) - (C_3)$, then for every $\varepsilon > 0$ there is an ε-robust discrete scoring rule.

Additional results: Approximate robustness

- A scoring rule is ε-robust if it elicits a belief within ε from the prior.
- Approximate robustness allows us to use "simple scoring rules".

Proposition (Approximately robust QSR under polynomial costs*)

If costs satisfy $(C_1) - (C_3)$ and are polynomial, then for every $\varepsilon > 0$ there is an ε-robust quadratic scoring rule.

Proposition (Approximately robust discrete scoring rules)

If costs satisfy $(C_1) - (C_3)$, then for every $\varepsilon > 0$ there is an ε-robust discrete scoring rule.

Additional results: Approximate robustness

- A scoring rule is ε-robust if it elicits a belief within ε from the prior.
- Approximate robustness allows us to use "simple scoring rules".

Proposition (Approximately robust QSR under polynomial costs*)

If costs satisfy $(C_1) - (C_3)$ *and are polynomial, then for every* $\varepsilon > 0$ *there is an* ε*-robust quadratic scoring rule.*

Proposition (Approximately robust discrete scoring rules)

If costs satisfy $(C_1) - (C_3)$*, then for every* $\varepsilon > 0$ *there is an* ε*-robust discrete scoring rule.*

Additional results: Cost uncertainty

- The experimenter cannot calibrate exactly the cost function.
- She estimates the cost function with a probability distribution.
- A scoring rule is (ε, δ)-robust if it elicits a belief within ε from the prior with probability at least $1 - \delta$.

Theorem (Approximate robustness under cost uncertainty)

If costs satisfy $(C_1) - (C_3)$ with probability $p > 0$, then for every $\varepsilon > 0$ and $\delta > 1 - p$ there is an (ε, δ)-robust quadratic scoring rule.

Additional results: Cost uncertainty

- The experimenter cannot calibrate exactly the cost function.
- She estimates the cost function with a probability distribution.
- A scoring rule is (ε, δ)-robust if it elicits a belief within ε from the prior with probability at least $1 - \delta$.

Theorem (Approximate robustness under cost uncertainty)

If costs satisfy $(C_1) - (C_3)$ with probability $p > 0$, then for every $\varepsilon > 0$ and $\delta > 1 - p$ there is an (ε, δ)-robust quadratic scoring rule.

Roadmap

Important step: Characterization result

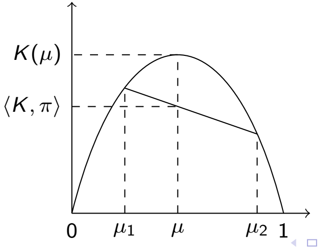
Lemma

The cost function satisfies $(C_1) - (C_3)$ *iff it satisfies*

POSTERIOR-SEPARABILITY: *There is a strictly concave function* $K : [0,1] \to \mathbb{R}$ *such that*

$$C(\pi) = K(\mu) - \langle K, \pi \rangle$$

for every $\pi \in \Pi(\mu)$ *and every* $\mu \in [0,1]$.



Important step: Characterization result

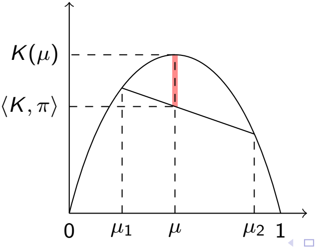
Lemma

The cost function satisfies $(C_1) - (C_3)$ *iff it satisfies*

POSTERIOR-SEPARABILITY: *There is a strictly concave function* $K : [0,1] \to \mathbb{R}$ *such that*

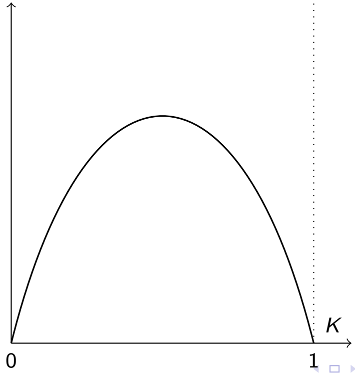
$$C(\pi) = K(\mu) - \langle K, \pi \rangle$$

for every $\pi \in \Pi(\mu)$ *and every* $\mu \in [0,1]$.



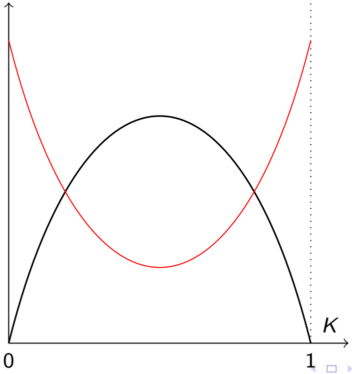
Graphical sketch of the proof

1. Start with K and then take candidate scoring rule $a - K$.
2. Benefits equal costs.
3. Take $f := b(a - K)$ for some $b \in (0, 1)$.
4. Costs offset benefits.
5. Question remaining: is f subdifferentiable?



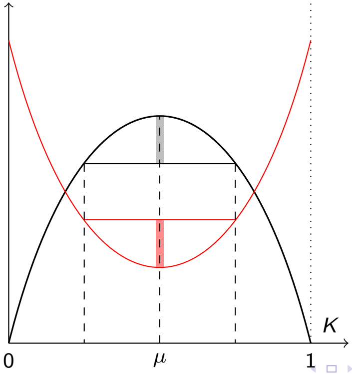
Graphical sketch of the proof

1. Start with K and then take candidate scoring rule $a - K$.
2. Benefits equal costs.
3. Take $f := b(a - K)$ for some $b \in (0, 1)$.
4. Costs offset benefits.
5. Question remaining: is f subdifferentiable?



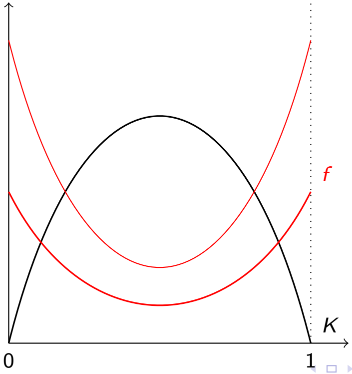
Graphical sketch of the proof

1. Start with K and then take candidate scoring rule $a - K$.
2. Benefits equal costs.
3. Take $f := b(a - K)$ for some $b \in (0, 1)$.
4. Costs offset benefits.
5. Question remaining: is f subdifferentiable?



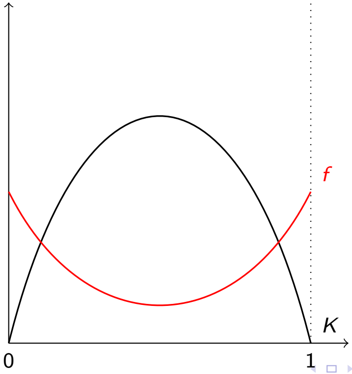
Graphical sketch of the proof

1. Start with K and then take candidate scoring rule $a - K$.
2. Benefits equal costs.
3. Take $f := b(a - K)$ for some $b \in (0, 1)$.
4. Costs offset benefits.
5. Question remaining: is f subdifferentiable?



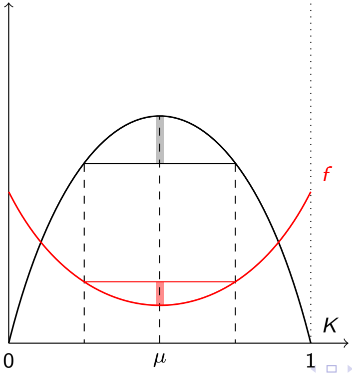
Graphical sketch of the proof

1. Start with K and then take candidate scoring rule $a - K$.
2. Benefits equal costs.
3. Take $f := b(a - K)$ for some $b \in (0, 1)$.
4. Costs offset benefits.
5. Question remaining: is f subdifferentiable?



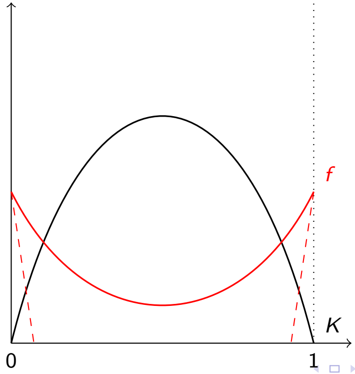
Graphical sketch of the proof

1. Start with K and then take candidate scoring rule $a - K$.
2. Benefits equal costs.
3. Take $f := b(a - K)$ for some $b \in (0, 1)$.
4. Costs offset benefits.
5. Question remaining: is f subdifferentiable?



Graphical sketch of the proof

1. Start with K and then take candidate scoring rule $a - K$.
2. Benefits equal costs.
3. Take $f := b(a - K)$ for some $b \in (0, 1)$.
4. Costs offset benefits.
5. Question remaining: is f subdifferentiable?



Sketch of the proof: boundary problem

Lemma

For every strictly convex function $f : [0,1] \to \mathbb{R}$, there exists some strictly convex $g : [0,1] \to \mathbb{R}$ such that

1. *$f - g$ is (weakly) convex ("g is less convex than f")*
2. *$\partial g(\nu) \neq \emptyset$ for all $\nu \in \{0,1\}$ ("g is subdifferentiable")*

- By (1), g yields even weaker benefits than f.
- By (2), g is a well-defined scoring rule.
- Hence, $\phi := g$ is a robust scoring rule (QED).

Sketch of the proof: boundary problem

Lemma

For every strictly convex function $f : [0,1] \to \mathbb{R}$, there exists some strictly convex $g : [0,1] \to \mathbb{R}$ such that

1. *$f - g$ is (weakly) convex ("g is less convex than f")*
2. *$\partial g(\nu) \neq \emptyset$ for all $\nu \in \{0,1\}$ ("g is subdifferentiable")*

- By (1), g yields even weaker benefits than f.
- By (2), g is a well-defined scoring rule.
- Hence, $\phi := g$ is a robust scoring rule (QED).

Sketch of the proof: boundary problem

Lemma

For every strictly convex function $f : [0,1] \to \mathbb{R}$, there exists some strictly convex $g : [0,1] \to \mathbb{R}$ such that

1. *$f - g$ is (weakly) convex ("g is less convex than f")*
2. *$\partial g(\nu) \neq \emptyset$ for all $\nu \in \{0,1\}$ ("g is subdifferentiable")*

- By (1), g yields even weaker benefits than f.
- By (2), g is a well-defined scoring rule.
- Hence, $\phi := g$ is a robust scoring rule (QED).

Sketch of the proof: boundary problem

Lemma

For every strictly convex function $f : [0,1] \to \mathbb{R}$, there exists some strictly convex $g : [0,1] \to \mathbb{R}$ such that

1. $f - g$ *is (weakly) convex ("g is less convex than f")*
2. $\partial g(\nu) \neq \emptyset$ *for all* $\nu \in \{0,1\}$ *("g is subdifferentiable")*

- By (1), g yields even weaker benefits than f.
- By (2), g is a well-defined scoring rule.
- Hence, $\phi := g$ is a robust scoring rule (QED).

Thanks for listening!!!