

Blackwell-Monotone Information Costs

ITAM

Xiaoyu Cheng

Florida State University

Yonggyun (YG) Kim

Florida State University

October 25, 2024

Introduction

Introduction

- Agenda: integration of costly information across various fields
- Question: Which information cost function *should* or *could* be used
 - Information cost: a map from a statistical experiment to a cost
- Examples
 - Entropy Costs: Sims (2003); Matějka, McKay (2015)
 - Posterior Separable Costs: Caplin, Dean, Leahy (2022); Denti (2022)
 - Log-Likelihood Ratio Costs: Pomatto, Strack, Tamuz (2023)
- Common Principle: Blackwell Monotonicity
 - More informative in Blackwell's order \Rightarrow higher cost
 - **Minimum requirement** for plausible information costs

Introduction

- Agenda: integration of costly information across various fields
- Question: Which information cost function *should* or *could* be used
 - Information cost: a map from a statistical experiment to a cost
- Examples
 - Entropy Costs: Sims (2003); Matějka, McKay (2015)
 - Posterior Separable Costs: Caplin, Dean, Leahy (2022); Denti (2022)
 - Log-Likelihood Ratio Costs: Pomatto, Strack, Tamuz (2023)
- Common Principle: Blackwell Monotonicity
 - More informative in Blackwell's order \Rightarrow higher cost
 - **Minimum requirement** for plausible information costs

Introduction

- Agenda: integration of costly information across various fields
- Question: Which information cost function *should* or *could* be used
 - Information cost: a map from a statistical experiment to a cost
- Examples
 - Entropy Costs: Sims (2003); Matějka, McKay (2015)
 - Posterior Separable Costs: Caplin, Dean, Leahy (2022); Denti (2022)
 - Log-Likelihood Ratio Costs: Pomatto, Strack, Tamuz (2023)
- Common Principle: Blackwell Monotonicity
 - More informative in Blackwell's order \Rightarrow higher cost
 - **Minimum requirement** for plausible information costs

Introduction

- Agenda: integration of costly information across various fields
- Question: Which information cost function *should* or *could* be used
 - Information cost: a map from a statistical experiment to a cost
- Examples
 - Entropy Costs: Sims (2003); Matějka, McKay (2015)
 - Posterior Separable Costs: Caplin, Dean, Leahy (2022); Denti (2022)
 - Log-Likelihood Ratio Costs: Pomatto, Strack, Tamuz (2023)
- Common Principle: Blackwell Monotonicity
 - More informative in Blackwell's order \Rightarrow higher cost
 - **Minimum requirement** for plausible information costs

Illustration: Blackwell Monotonicity

- Consider consumers seeking to acquire information about their COVID-19 status
- Two tests are available in the competitive market:

		signal	
		n	p
state	—	80%	20%
	+	20%	80%

Test A (\$10)

		signal	
		n	p
state	—	60%	40%
	+	15%	85%

Test B (\$12)

- Test B can be replicated using test A, which costs \$10
 - When n is realized, toss a coin twice and replace it with p if both are heads

Illustration: Blackwell Monotonicity

- Consider consumers seeking to acquire information about their COVID-19 status
- Two tests are available in the competitive market:

		signal	
		n	p
state	−	80%	20%
	+	20%	80%

Test A (\$10)

		signal	
		n	p
state	−	60%	40%
	+	15%	85%

Test B (\$12)

- Test B can be replicated using test A, which costs \$10
 - When n is realized, toss a coin twice and replace it with p if both are heads

Illustration: Blackwell Monotonicity

- Consider consumers seeking to acquire information about their COVID-19 status
- Two tests are available in the competitive market:

		signal		
		$n_{oth.}$	n_{HH}	p
state	—	60%	20%	20%
	+	15%	5%	80%

Test A (\$10)

		signal	
		n	p
state	—	60%	40%
	+	15%	85%

Test B (\$12)

- Test B can be replicated using test A, which costs \$10
 - When n is realized, toss a coin twice and replace it with p if both are heads

Illustration: Blackwell Monotonicity

- Consider consumers seeking to acquire information about their COVID-19 status
- Two tests are available in the competitive market:

		signal		=			signal	
		$n_{oth.}$	p				n	p
state	—	60%	20%		state	—	60%	40%
	+	15%	80%			+	15%	85%
Test A (\$10)					Test B (\$12)			

- Test B can be replicated using test A, which costs \$10
 - When n is realized, toss a coin twice and replace it with p if both are heads

Illustration: Blackwell Monotonicity

- **Blackwell's Theorem**

- A is more informative than $B \Leftrightarrow B$ is a *garbling* of A

- **Blackwell Monotonicity**

- A should be more costly than B whenever A is Blackwell more informative than B

- **This paper**

- identifies elementary necessary and sufficient conditions for Blackwell monotonicity
 - characterizes a practical and tractable class of information cost functions

► Literature

Illustration: Blackwell Monotonicity

- **Blackwell's Theorem**

- A is more informative than $B \Leftrightarrow B$ is a *garbling* of A

- **Blackwell Monotonicity**

- A should be more costly than B whenever A is Blackwell more informative than B

- **This paper**

- identifies elementary necessary and sufficient conditions for Blackwell monotonicity
- characterizes a practical and tractable class of information cost functions

► Literature

Illustration: Blackwell Monotonicity

- **Blackwell's Theorem**

- A is more informative than $B \Leftrightarrow B$ is a *garbling* of A

- **Blackwell Monotonicity**

- A should be more costly than B whenever A is Blackwell more informative than B

- **This paper**

- identifies elementary necessary and sufficient conditions for Blackwell monotonicity
- characterizes a practical and tractable class of information cost functions

► Literature

1. Preliminaries
2. Binary Experiments
3. Finite Experiments: more than two signals
4. Likelihood Separable Costs
5. Application: Bargaining and Information Acquisition

Preliminaries

Experiments

- $\Omega = \{\omega_1, \dots, \omega_n\}$: a finite set of states
- $\mathcal{S} = \{s_1, \dots, s_m\}$: a finite set of signals
- A *statistical experiment* $f : \Omega \rightarrow \Delta(\mathcal{S})$ can be represented by an $n \times m$ matrix:

$$f = \begin{bmatrix} f_1^1 & \dots & f_1^m \\ \vdots & \ddots & \vdots \\ f_n^1 & \dots & f_n^m \end{bmatrix},$$

where $f_i^j = \Pr(s_j | \omega_i)$, thus, $f_i^j \geq 0$ and $\sum_{j=1}^m f_i^j = 1$

- Let f^j denote j -th column vector, namely the likelihood vector of signal j
- $\mathcal{E}_m \subset \mathbb{R}^{n \times m}$: the space of all experiments with m possible signals

Blackwell Informativeness

- $f \succeq_B g$: f is *Blackwell more informative* than g
iff g is a garbling of f : \exists a stochastic matrix M s.t. $g = f M$

- Examples of garbling under binary signal

1. **Signal Replacement**: for some $\epsilon > 0$,

$$M = \begin{bmatrix} 1 - \epsilon & \epsilon \\ 0 & 1 \end{bmatrix}$$

meaning that s_1 is replaced with s_2 with probability ϵ

2. **Permutation**:

$$P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

meaning that signals are relabeled

so $f \circ P \succeq_B f$: relabeling signals does not change the informativeness

Blackwell Informativeness

- $f \succeq_B g$: f is *Blackwell more informative* than g
iff g is a garbling of f : \exists a stochastic matrix M s.t. $g = f M$

- Examples of garbling under binary signal

1. **Signal Replacement**: for some $\epsilon > 0$,

$$M = \begin{bmatrix} 1 - \epsilon & \epsilon \\ 0 & 1 \end{bmatrix}$$

meaning that s_1 is replaced with s_2 with probability ϵ

2. **Permutation**:

$$P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

meaning that signals are relabeled

• $f \simeq_B f P$: relabeling signals does not change the informativeness

Blackwell Informativeness

- $f \succeq_B g$: f is *Blackwell more informative* than g
iff g is a garbling of f : \exists a stochastic matrix M s.t. $g = f M$

- Examples of garbling under binary signal

1. **Signal Replacement**: for some $\epsilon > 0$,

$$M = \begin{bmatrix} 1 - \epsilon & \epsilon \\ 0 & 1 \end{bmatrix}$$

meaning that s_1 is replaced with s_2 with probability ϵ

2. **Permutation**:

$$P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

meaning that signals are relabeled

- $f \simeq_B f P$: relabeling signals does not change the informativeness

Blackwell Informativeness

- $f \succeq_B g$: f is *Blackwell more informative* than g
iff g is a garbling of f : \exists a stochastic matrix M s.t. $g = f M$

- Examples of garbling under binary signal

1. **Signal Replacement**: for some $\epsilon > 0$,

$$M = \begin{bmatrix} 1 - \epsilon & \epsilon \\ 0 & 1 \end{bmatrix}$$

meaning that s_1 is replaced with s_2 with probability ϵ

2. **Permutation**:

$$P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

meaning that signals are relabeled

- $f \simeq_B f P$: relabeling signals does not change the informativeness

Information Costs and Blackwell Monotonicity

- **Information Costs**

- $C : \mathcal{E}_m \rightarrow \mathbb{R}_+$: an information cost function
- \mathcal{C}_m : the set of all absolutely continuous information cost functions defined over \mathcal{E}_m
- Absolute continuity ensures that a derivative exists a.e. and is integrable
- In the talk, assume that C is differentiable and the gradient exists

- **Blackwell Monotonicity**

- An information cost function $C \in \mathcal{C}_m$ is **Blackwell monotone** if for all $f, g \in \mathcal{E}_m$, $C(f) \geq C(g)$ whenever $f \succeq_B g$.

- **Permutation Invariance**

- Any Blackwell-monotone information cost function is **permutation invariant**, i.e., $C(f) = C(f P)$ for any permutation matrix P

Information Costs and Blackwell Monotonicity

- **Information Costs**

- $C : \mathcal{E}_m \rightarrow \mathbb{R}_+$: an information cost function
- \mathcal{C}_m : the set of all absolutely continuous information cost functions defined over \mathcal{E}_m
- Absolute continuity ensures that a derivative exists a.e. and is integrable
- In the talk, assume that C is differentiable and the gradient exists

- **Blackwell Monotonicity**

- An information cost function $C \in \mathcal{C}_m$ is **Blackwell monotone** if for all $f, g \in \mathcal{E}_m$, $C(f) \geq C(g)$ whenever $f \succeq_B g$.

- **Permutation Invariance**

- Any Blackwell-monotone information cost function is **permutation invariant**, i.e., $C(f) = C(f P)$ for any permutation matrix P

Information Costs and Blackwell Monotonicity

- **Information Costs**

- $C : \mathcal{E}_m \rightarrow \mathbb{R}_+$: an information cost function
- \mathcal{C}_m : the set of all absolutely continuous information cost functions defined over \mathcal{E}_m
- Absolute continuity ensures that a derivative exists a.e. and is integrable
- In the talk, assume that C is differentiable and the gradient exists

- **Blackwell Monotonicity**

- An information cost function $C \in \mathcal{C}_m$ is **Blackwell monotone** if for all $f, g \in \mathcal{E}_m$, $C(f) \geq C(g)$ whenever $f \succeq_B g$.

- **Permutation Invariance**

- Any Blackwell-monotone information cost function is **permutation invariant**, i.e., $C(f) = C(f P)$ for any permutation matrix P

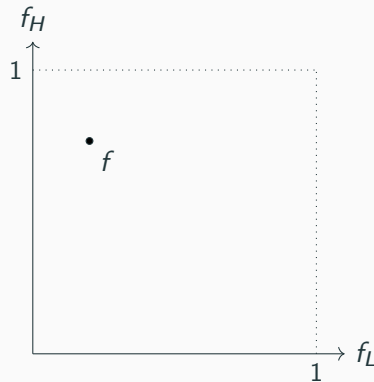
Binary Experiments

Binary Experiments

- Focus on the case where $n = m = 2$
- Any experiment can be represented by $f \equiv (f_L, f_H)^\top \in [0, 1]^2$:

$$[1 - f, f] = \begin{array}{c|cc} & s_L & s_H \\ \hline \omega_L & 1 - f_L & f_L \\ \omega_H & 1 - f_H & f_H \end{array}$$

- $1 - f$ is a permutation of f
- When $f_L = f_H$, it is completely uninformative

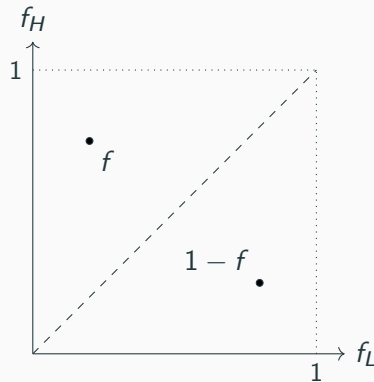


Binary Experiments

- Focus on the case where $n = m = 2$
- Any experiment can be represented by $f \equiv (f_L, f_H)^\top \in [0, 1]^2$:

$$[1 - f, f] = \begin{array}{c|cc} & s_L & s_H \\ \hline \omega_L & 1 - f_L & f_L \\ \omega_H & 1 - f_H & f_H \end{array}$$

- $1 - f$ is a permutation of f
- When $f_L = f_H$, it is completely uninformative

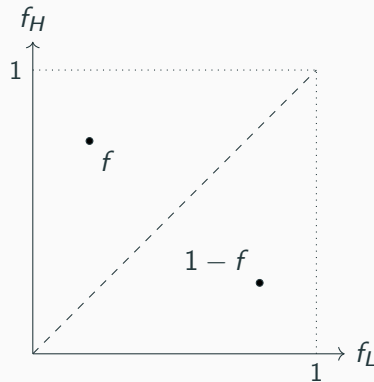


Binary Experiments

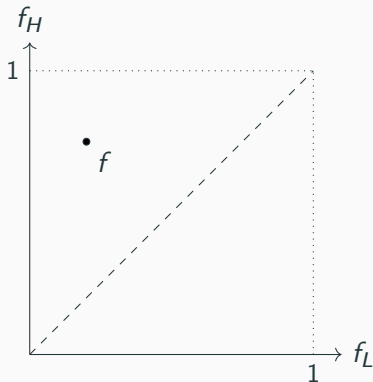
- Focus on the case where $n = m = 2$
- Any experiment can be represented by $f \equiv (f_L, f_H)^\top \in [0, 1]^2$:

$$[1 - f, f] = \begin{array}{c|cc} & s_L & s_H \\ \hline \omega_L & 1 - f_L & f_L \\ \omega_H & 1 - f_H & f_H \end{array}$$

- $1 - f$ is a permutation of f
- When $f_L = f_H$, it is completely uninformative



Blackwell Informativeness under Binary Experiments



- Recall that $f \succeq_B g$ iff

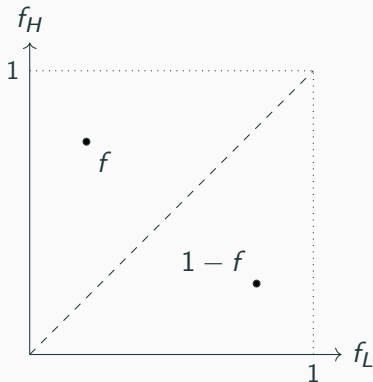
$$[1 - g, g] = [1 - f, f] M$$

for some stochastic matrix M

- Extreme points of M :

$$M_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad M_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$
$$M_3 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \quad M_4 = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$$

Blackwell Informativeness under Binary Experiments



- Recall that $f \succeq_B g$ iff

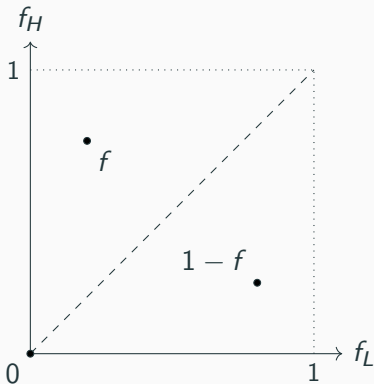
$$[1 - g, g] = [1 - f, f] M$$

for some stochastic matrix M

- Extreme points of M :

$$M_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad M_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$
$$M_3 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \quad M_4 = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$$

Blackwell Informativeness under Binary Experiments



- Recall that $f \succeq_B g$ iff

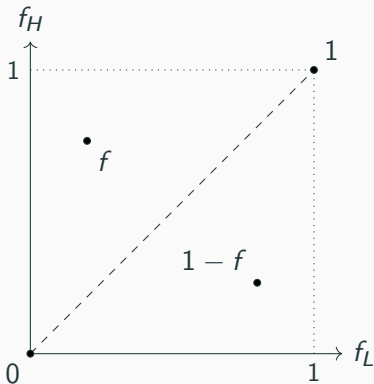
$$[1 - g, g] = [1 - f, f] M$$

for some stochastic matrix M

- Extreme points of M :

$$M_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad M_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$
$$M_3 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \quad M_4 = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$$

Blackwell Informativeness under Binary Experiments



- Recall that $f \succeq_B g$ iff

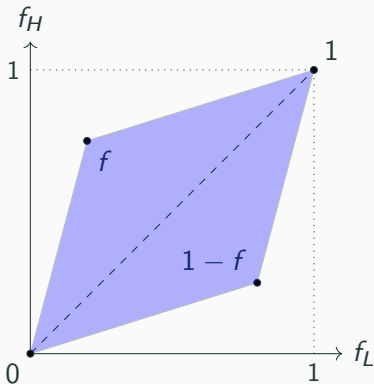
$$[1 - g, g] = [1 - f, f] M$$

for some stochastic matrix M

- Extreme points of M :

$$M_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad M_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$
$$M_3 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \quad M_4 = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$$

Blackwell Informativeness under Binary Experiments



- Recall that $f \succeq_B g$ iff

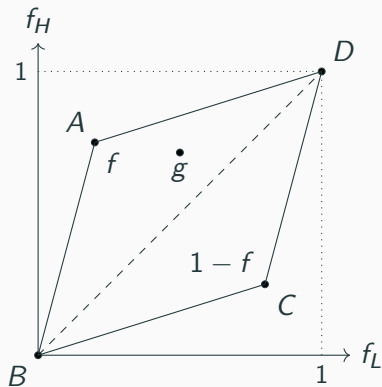
$$[1 - g, g] = [1 - f, f] M$$

for some stochastic matrix M

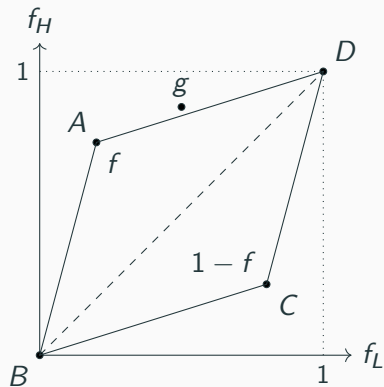
- Extreme points of M :

$$M_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad M_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$
$$M_3 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \quad M_4 = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$$

Blackwell Informativeness under Binary Experiments



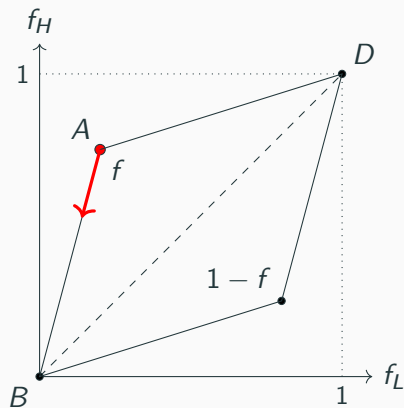
$f \succeq_B g$



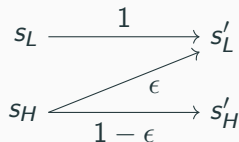
$f \not\succeq_B g$

Necessary Conditions for Blackwell Monotonicity

When an information cost C is Blackwell monotone,



$$1. \langle \nabla C(f), -f \rangle \leq 0$$

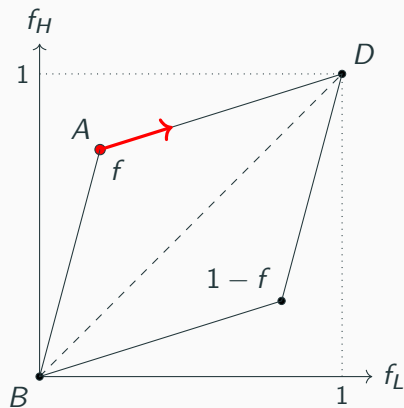


$$2. \langle \nabla C(f), 1 - f \rangle \leq 0$$

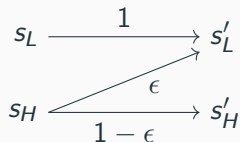


Necessary Conditions for Blackwell Monotonicity

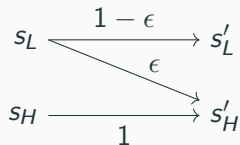
When an information cost C is Blackwell monotone,



$$1. \langle \nabla C(f), -f \rangle \leq 0$$



$$2. \langle \nabla C(f), 1 - f \rangle \leq 0$$



Theorem for Binary Experiments

Theorem 1

$C \in \mathcal{C}_2$ is Blackwell monotone if and only if it is

1. permutation invariant;
2. for all $f \in \mathcal{E}_2$,

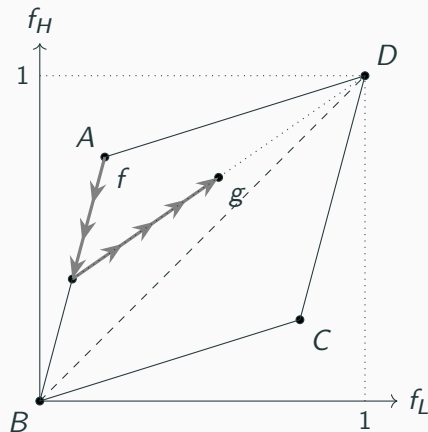
$$\langle \nabla C(f), f \rangle \geq 0 \geq \langle \nabla C(f), 1 - f \rangle. \quad (1)$$

- This theorem holds for the cases with more than two states, but the binary signal assumption is crucial.

► Quiz

Proof for Sufficiency

For any $f \succeq_B g$, we can find a path from f to g (or the permutation of it) along which Blackwell informativeness decreases



Further Characterizations with Binary States

Quiz

Which of the followings (defined over $f_H > f_L$) are Blackwell-monotone information cost functions?

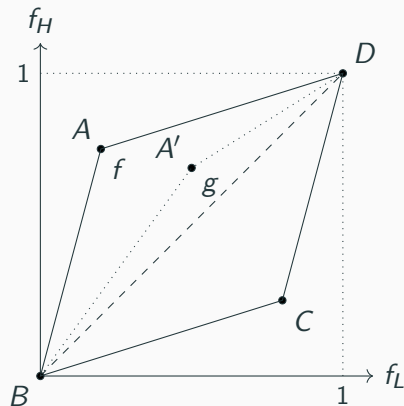
1. $C(f_L, f_H) = \frac{f_H(1 - f_H)}{f_L(1 - f_L)} - 1$

2. $C(f_L, f_H) = \frac{f_H}{f_L} + \frac{1 - f_L}{1 - f_H} - 2$

3. $C(f_L, f_H) = (f_H - f_L)^2$

4. $C(f_L, f_H) = f_H - 2f_L$

Further Characterizations with Binary States



$f \succeq_B g$ is equivalent to:

1. AB steeper than $A'B$:

$$\alpha \equiv \frac{f_H}{f_L} \geq \frac{g_H}{g_L} \equiv \alpha'$$

α : likelihood ratio (LR) of receiving s_H

2. AD shallower than $A'D$:

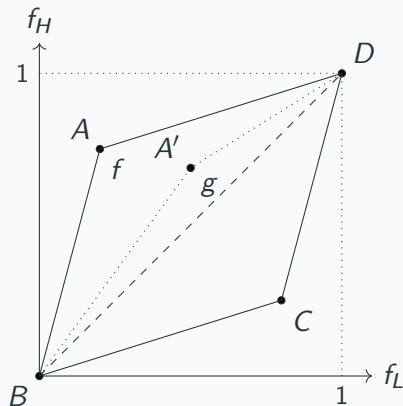
$$\beta \equiv \frac{1 - f_L}{1 - f_H} \geq \frac{1 - g_L}{1 - g_H} \equiv \beta'$$

β : the inverse of LR of receiving s_L

Proposition

C is Blackwell monotone iff it is increasing in α and β after reparametrization

Further Characterizations with Binary States



$f \succeq_B g$ is equivalent to:

1. AB steeper than $A'B$:

$$\alpha \equiv \frac{f_H}{f_L} \geq \frac{g_H}{g_L} \equiv \alpha'$$

α : likelihood ratio (LR) of receiving s_H

2. AD shallower than $A'D$:

$$\beta \equiv \frac{1 - f_L}{1 - f_H} \geq \frac{1 - g_L}{1 - g_H} \equiv \beta'$$

β : the inverse of LR of receiving s_L

Proposition

C is Blackwell monotone iff it is increasing in α and β after reparametrization

Further Characterizations with Binary States

1. $C(f_L, f_H) = \frac{f_H(1 - f_H)}{f_L(1 - f_L)} - 1$ with $1 > f_H > f_L > 0$

$$\tilde{C}(\alpha, \beta) = \frac{\alpha}{\beta} - 1$$

- \tilde{C} is increasing in α but not in β , thus, \tilde{C} is not Blackwell monotone.

2. $C(f_L, f_H) = \frac{f_H}{f_L} + \frac{1 - f_L}{1 - f_H} - 2$ with $1 > f_H > f_L > 0$

$$\tilde{C}(\alpha, \beta) = \alpha + \beta - 2$$

- \tilde{C} is increasing in both α and β , thus, \tilde{C} is **Blackwell monotone**.

Further Characterizations with Binary States

1. $C(f_L, f_H) = \frac{f_H(1 - f_H)}{f_L(1 - f_L)} - 1$ with $1 > f_H > f_L > 0$

$$\tilde{C}(\alpha, \beta) = \frac{\alpha}{\beta} - 1$$

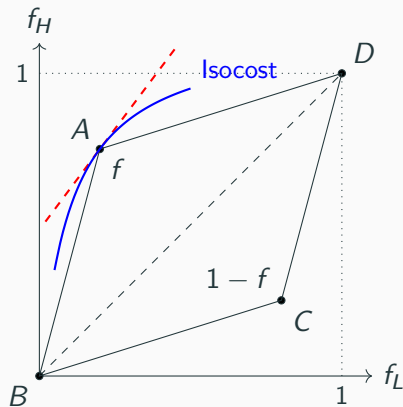
- \tilde{C} is increasing in α but not in β , thus, \tilde{C} is not Blackwell monotone.

2. $C(f_L, f_H) = \frac{f_H}{f_L} + \frac{1 - f_L}{1 - f_H} - 2$ with $1 > f_H > f_L > 0$

$$\tilde{C}(\alpha, \beta) = \alpha + \beta - 2$$

- \tilde{C} is increasing in both α and β , thus, \tilde{C} is **Blackwell monotone**.

Further Characterizations with Binary States

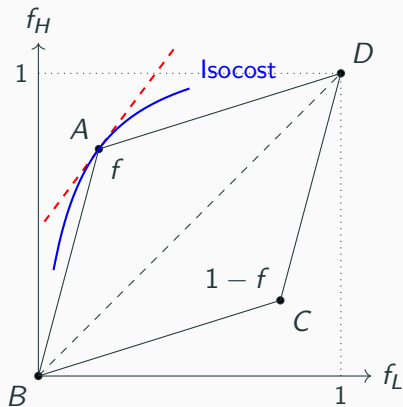


$\langle \nabla C(f), f \rangle \geq 0 \geq \langle \nabla C(f), 1 - f \rangle$
is equivalent to:

$$\underbrace{\frac{f_H}{f_L}}_{\text{the slope of } \overline{AB}} \geq \underbrace{-\frac{\partial C / \partial f_L}{\partial C / \partial f_H}}_{\text{the slope of the isocost curve}} \geq \underbrace{\frac{1 - f_H}{1 - f_L}}_{\text{the slope of } \overline{AD}}$$

- **Interpretation:** a *marginal rate of information transformation* (MRIT) lies between the two likelihood ratios provided by the experiment.

Further Characterizations with Binary States



$\langle \nabla C(f), f \rangle \geq 0 \geq \langle \nabla C(f), 1 - f \rangle$
 is equivalent to:

$$\underbrace{\frac{f_H}{f_L}}_{\text{the slope of } \overline{AB}} \geq \underbrace{-\frac{\partial C / \partial f_L}{\partial C / \partial f_H}}_{\text{the slope of the isocost curve}} \geq \underbrace{\frac{1 - f_H}{1 - f_L}}_{\text{the slope of } \overline{AD}}$$

- **Interpretation:** a *marginal rate of information transformation* (MRIT) lies between the two likelihood ratios provided by the experiment.

Further Characterizations with Binary States

3. $C(f_L, f_H) = (f_H - f_L)^2$ with $1 > f_H > f_L > 0$

$$\frac{f_H}{f_L} \geq -\frac{\partial C / \partial f_L}{\partial C / \partial f_H} = 1 \geq \frac{1 - f_H}{1 - f_L}$$

- The above inequalities hold for all $1 > f_H > f_L > 0$, thus, it is **Blackwell monotone**.

4. $C(f_L, f_H) = f_H - 2f_L$ with $1 > f_H > f_L > 0$

$$\frac{f_H}{f_L} \geq -\frac{\partial C / \partial f_L}{\partial C / \partial f_H} = 2 \geq \frac{1 - f_H}{1 - f_L}$$

- The above inequalities does not always hold, e.g., $f_L = .5$ and $f_H = .6$, thus, it is not Blackwell monotone.

Further Characterizations with Binary States

3. $C(f_L, f_H) = (f_H - f_L)^2$ with $1 > f_H > f_L > 0$

$$\frac{f_H}{f_L} \geq -\frac{\partial C / \partial f_L}{\partial C / \partial f_H} = 1 \geq \frac{1 - f_H}{1 - f_L}$$

- The above inequalities hold for all $1 > f_H > f_L > 0$, thus, it is **Blackwell monotone**.

4. $C(f_L, f_H) = f_H - 2f_L$ with $1 > f_H > f_L > 0$

$$\frac{f_H}{f_L} \geq -\frac{\partial C / \partial f_L}{\partial C / \partial f_H} = 2 \geq \frac{1 - f_H}{1 - f_L}$$

- The above inequalities does not always hold, e.g., $f_L = .5$ and $f_H = .6$, thus, it is not Blackwell monotone.

Answer for the Quiz

Which of the followings are Blackwell-monotone information cost functions?

1. $C(f_L, f_H) = \frac{f_H(1 - f_H)}{f_L(1 - f_L)} - 1$

2. $C(f_L, f_H) = \frac{f_H}{f_L} + \frac{1 - f_L}{1 - f_H} - 2$

3. $C(f_L, f_H) = (f_H - f_L)^2$

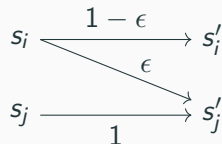
4. $C(f_L, f_H) = f_H - 2f_L$

Finite Experiments: more than two signals

Necessary Conditions for Blackwell Monotonicity

Now assume that there are more than two signals.

- Permutation invariance is still necessary
- For any pair (i, j) , the following garbling worsens the informativeness:



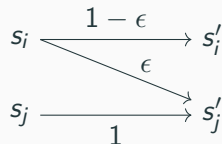
- This gives us $\langle \nabla^{i \rightarrow j} C(f), f^i \rangle \leq 0$, where

$$\langle \nabla^{i \rightarrow j} C(f), f^i \rangle \equiv - \sum_{s=1}^n \frac{\partial C}{\partial f_s^i} \cdot f_s^i + \sum_{s=1}^n \frac{\partial C}{\partial f_s^j} \cdot f_s^i$$

Necessary Conditions for Blackwell Monotonicity

Now assume that there are more than two signals.

- Permutation invariance is still necessary
- For any pair (i, j) , the following garbling worsens the informativeness:



- This gives us $\langle \nabla^{i \rightarrow j} C(f), f^i \rangle \leq 0$, where

$$\langle \nabla^{i \rightarrow j} C(f), f^i \rangle \equiv - \sum_{s=1}^n \frac{\partial C}{\partial f_s^i} \cdot f_s^i + \sum_{s=1}^n \frac{\partial C}{\partial f_s^j} \cdot f_s^i$$

Sufficient Conditions for Blackwell Monotonicity

- For binary experiments, sufficiency was established by finding a path between two experiments along which informativeness decreases
- However, when $m \geq 3$, there may not exist such path [▶ Illustrations](#)
- To overcome this issue, we impose quasiconvexity on C : [▶ Illustration](#)

$$C(\lambda f + (1 - \lambda)g) \leq \max\{C(f), C(g)\}.$$

With quasiconvexity, the first-order condition serves as a sufficient condition for Blackwell monotonicity

• Remarks

- Quasiconvexity is not a necessary condition for Blackwell Monotonicity [▶ Details](#)
- We found a weaker (but less standard) version of Quasiconvexity serving as a necessary condition for Blackwell monotonicity [▶ Details](#)

Sufficient Conditions for Blackwell Monotonicity

- For binary experiments, sufficiency was established by finding a path between two experiments along which informativeness decreases
- However, when $m \geq 3$, there may not exist such path [▶ Illustrations](#)
- To overcome this issue, we impose quasiconvexity on C : [▶ Illustration](#)

$$C(\lambda f + (1 - \lambda)g) \leq \max\{C(f), C(g)\}.$$

With quasiconvexity, the first-order condition serves as a sufficient condition for Blackwell monotonicity

• Remarks

- Quasiconvexity is not a necessary condition for Blackwell Monotonicity [▶ Details](#)
- We found a weaker (but less standard) version of Quasiconvexity serving as a necessary condition for Blackwell monotonicity [▶ Details](#)

Sufficient Conditions for Blackwell Monotonicity

- For binary experiments, sufficiency was established by finding a path between two experiments along which informativeness decreases
- However, when $m \geq 3$, there may not exist such path [▶ Illustrations](#)
- To overcome this issue, we impose quasiconvexity on C : [▶ Illustration](#)

$$C(\lambda f + (1 - \lambda)g) \leq \max\{C(f), C(g)\}.$$

With quasiconvexity, the first-order condition serves as a sufficient condition for Blackwell monotonicity

• Remarks

- Quasiconvexity is not a necessary condition for Blackwell Monotonicity [▶ Details](#)
- We found a weaker (but less standard) version of Quasiconvexity serving as a necessary condition for Blackwell monotonicity [▶ Details](#)

Sufficient Conditions for Blackwell Monotonicity

- For binary experiments, sufficiency was established by finding a path between two experiments along which informativeness decreases
- However, when $m \geq 3$, there may not exist such path [▶ Illustrations](#)
- To overcome this issue, we impose quasiconvexity on C : [▶ Illustration](#)

$$C(\lambda f + (1 - \lambda)g) \leq \max\{C(f), C(g)\}.$$

With quasiconvexity, the first-order condition serves as a sufficient condition for Blackwell monotonicity

• Remarks

- Quasiconvexity is not a necessary condition for Blackwell Monotonicity [▶ Details](#)
- We found a weaker (but less standard) version of Quasiconvexity serving as a necessary condition for Blackwell monotonicity [▶ Details](#)

Theorem for Finite Experiments

Theorem 2

Suppose that $C \in \mathcal{C}_m$ is absolutely continuous and quasiconvex. Then, C is Blackwell monotone if and only if it is

1. permutation invariant;
2. for all $f \in \mathcal{E}_m$ and $i \neq j$,

$$\langle \nabla^{i \rightarrow j} C(f), f \rangle \leq 0. \quad (2)$$

- $S_B(f)$: the set of experiments that are less Blackwell informative than f
- Two conditions ensure that extreme points of $S_B(f)$ are not more costly than f
- Then, we can apply quasiconvexity

Theorem for Finite Experiments

Theorem 2

Suppose that $C \in \mathcal{C}_m$ is absolutely continuous and quasiconvex. Then, C is Blackwell monotone if and only if it is

1. permutation invariant;
2. for all $f \in \mathcal{E}_m$ and $i \neq j$,

$$\langle \nabla^{i \rightarrow j} C(f), f \rangle \leq 0. \quad (2)$$

- $S_B(f)$: the set of experiments that are less Blackwell informative than f
- Two conditions ensure that extreme points of $S_B(f)$ are not more costly than f
- Then, we can apply quasiconvexity

Likelihood Separable Costs

Likelihood Separable Costs

Likelihood Separable Costs

C is *likelihood separable* if there exist a constant a and an absolutely continuous function $\psi : \mathbb{R}_+^n \rightarrow \mathbb{R}_+$ such that, for all m and $f \in \mathcal{E}_m$,

$$C(f) = \sum_{j=1}^m \psi(f^j) + a,$$

where f^j is the likelihood vector of signal j .

Let \mathcal{C}^{LS} be the class of likelihood separable costs

Groundedness

C is *grounded* if it assigns zero cost to uninformative experiments.

Let \mathcal{C}^G be the class of grounded costs.

Theorem 3

When $C \in \mathcal{C}^{LS}$, C is Blackwell monotone if and only if ψ is sublinear:

► proof

1. positive homogeneity: $\psi(\alpha h) = \alpha\psi(h)$;
2. subadditivity: $\psi(k) + \psi(l) \geq \psi(k + l)$

Corollary: GSLS costs

C is called *grounded sublinear likelihood separable (GSLS)* if there exists a **sublinear** and absolutely continuous function ψ such that

$$C(f) = \sum_{j=1}^m \psi(f^j) - \psi(1).$$

Then,

$$\mathcal{C}^{GSLS} = \mathcal{C}^{LS} \cap \mathcal{C}^G \cap \mathcal{C}^{BM}$$

Theorem 3

When $C \in \mathcal{C}^{LS}$, C is Blackwell monotone if and only if ψ is sublinear:

► proof

1. positive homogeneity: $\psi(\alpha h) = \alpha\psi(h)$;
2. subadditivity: $\psi(k) + \psi(l) \geq \psi(k + l)$

Corollary: GSLS costs

C is called *grounded sublinear likelihood separable (GSLS)* if there exists a **sublinear** and absolutely continuous function ψ such that

$$C(f) = \sum_{j=1}^m \psi(f^j) - \psi(1).$$

Then,

$$\mathcal{C}^{GSLS} = \mathcal{C}^{LS} \cap \mathcal{C}^G \cap \mathcal{C}^{BM}$$

Examples: GSLS Costs

1. Supnorm Costs

$$C(f) = \sum_{j=1}^m \max_i f_i^j - 1,$$

2. Absolute-Linear Costs

$$C(f) = \sum_{j=1}^m |\langle a, f^j \rangle| - |\langle a, 1 \rangle| = \sum_{j=1}^m \left| \sum_{i=1}^n a_i f_i^j \right| - \left| \sum_{i=1}^n a_i \right|.$$

3. Linear ϕ -divergence Costs (including LLR costs of Pomatto, Strack, Tamuz (2023))

$$C(f) = \sum_{j=1}^m \sum_{i,i'} \beta_{ii'} f_{i'}^j \phi_{ii'} \left(\frac{f_i^j}{f_{i'}^j} \right) = \sum_{i,i'} \beta_{ii'} \sum_{j=1}^m f_{i'}^j \phi_{ii'} \left(\frac{f_i^j}{f_{i'}^j} \right), \quad (3)$$

where $\phi_{ii'} : [0, \infty] \rightarrow \mathbb{R} \cup \{+\infty\}$ is a convex function with $\phi_{ii'}(1) = 0$ and $\beta_{ii'} \geq 0$

Posterior Separability

C has a *posterior separable (PS)* representation at a prior belief $\mu \in \Delta(\Omega)$ if there exists a concave and absolutely continuous function $H : \Delta(\Omega) \rightarrow \mathbb{R}$ such that

$$C(f) = H(\mu) - \sum_{j=1}^m \tau_{\mu}(f^j) \cdot H(q_{\mu}(f^j))$$

where $q_{\mu}(f^j)$ is the posterior belief upon receiving s_j and $\tau_{\mu}(f^j)$ is the probability of receiving s_j .

Let C_{μ}^{PS} denote the class of cost functions that have PS representations at μ .

Proposition

For any full support prior $\mu \in \Delta(\Omega)$, $C^{GSLS} = C_{\mu}^{PS}$.

Posterior Separability

C has a *posterior separable (PS)* representation at a prior belief $\mu \in \Delta(\Omega)$ if there exists a concave and absolutely continuous function $H : \Delta(\Omega) \rightarrow \mathbb{R}$ such that

$$C(f) = H(\mu) - \sum_{j=1}^m \tau_{\mu}(f^j) \cdot H(q_{\mu}(f^j))$$

where $q_{\mu}(f^j)$ is the posterior belief upon receiving s_j and $\tau_{\mu}(f^j)$ is the probability of receiving s_j .

Let C_{μ}^{PS} denote the class of cost functions that have PS representations at μ .

Proposition

For any full support prior $\mu \in \Delta(\Omega)$, $C^{GSLS} = C_{\mu}^{PS}$.

Application: Bargaining and Information Acquisition

- Consider a bargaining problem with information acquisition
 - Players: **S**eller and **B**uyer
 - State (**B**'s valuation): $v \in \{L, H\}$ with $H > L > 0$
 - Prior belief: $\pi \equiv \Pr(v = H) \in (0, 1)$
 - Timing of the game
 1. Nature draws v and **S** observes v
 2. **S** offers p
 3. **B** costly acquires information about v and then accepts or rejects
- Chatterjee et al. focus on specific types of information acquisition
- We extend their analysis by allowing **B** to choose information flexibly

- Consider a bargaining problem with information acquisition
 - Players: **S**eller and **B**uyer
 - State (**B**'s valuation): $v \in \{L, H\}$ with $H > L > 0$
 - Prior belief: $\pi \equiv \Pr(v = H) \in (0, 1)$
 - Timing of the game
 1. Nature draws v and **S** observes v
 2. **S** offers p
 3. **B** costly acquires information about v and then accepts or rejects
- Chatterjee et al. focus on specific types of information acquisition
- We extend their analysis by allowing **B** to choose information flexibly

B's cost: $\lambda \cdot c(f_H)$

Result 1: pooling eq'm

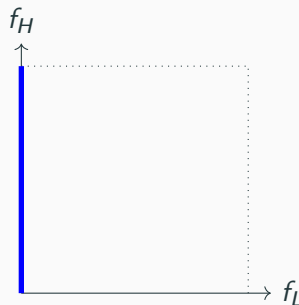
under H-focused signal structure, for any λ , there exists $\epsilon > 0$ such that every equilibrium is a pooling equilibrium where

1. both types of **S** offer $p^* \in [L, L + \epsilon)$;
2. **B** accepts without information acquisition.

Moreover, $\epsilon \rightarrow 0$ as $\lambda \rightarrow 0$, thus, **B** extracts full surplus as $\lambda \rightarrow 0$

H-focused Information

	s_L	s_H
L	1	0
H	$1 - f_H$	f_H



B's cost: $\lambda \cdot c(1 - f_L)$

Result 2: almost-separating eq'm

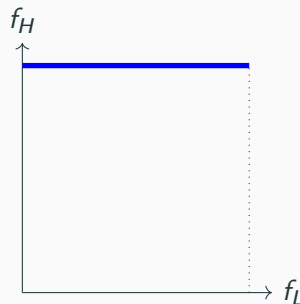
under L-focused signal structure, for any small enough λ , there exists an equilibrium such that

1. type H **S** offers $p^* \approx H$;
2. type L **S** offers L with prob. $1 - \epsilon$,
 p^* with prob. ϵ ;
3. **B** acquires information and conditions her purchase decision on the signal realization

Moreover, **S's** payoff is close to v and **B's** payoff is close to zero

L-focused Information

	s_L	s_H
L	$1 - f_L$	f_L
H	0	1



Flexible Information Acquisition

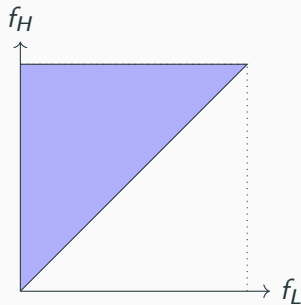
- We extend to the full domain and consider $\lambda|f_2 - f_1|$ and $\lambda(f_2 - f_1)^2$

Result 1': when $C(f_1, f_2) = \lambda|f_2 - f_1|$, the unique equilibrium is the pooling equilibrium, and as $\lambda \rightarrow 0$, **B** extracts full surplus

Result 2': when $C(f_1, f_2) = \lambda(f_2 - f_1)^2$, there exists an almost-separating equilibrium, and **S**'s payoff is close to v and **B**'s payoff is close to zero

Flexible Information

	s_L	s_H
L	$1 - f_L$	f_L
H	$1 - f_H$	f_H



Conclusion

Conclusion

- We identify necessary and sufficient conditions for Blackwell Monotonicity.
- Under likelihood separability, we show that the sublinearity of the component function is equivalent to Blackwell Monotonicity.
- **Applications:** we apply our results to extend
 1. Costly Persuasion (Gentzkow, Kamenica, 2014)
 2. Bargaining and Information Acquisition (Chatterjee, Dong, Hoshino, 2024)
- Future Research: Lehmann-Monotone Information Costs based on Kim [2023]

► Costly Persuasion

Conclusion

- We identify necessary and sufficient conditions for Blackwell Monotonicity.
- Under likelihood separability, we show that the sublinearity of the component function is equivalent to Blackwell Monotonicity.
- **Applications:** we apply our results to extend
 1. Costly Persuasion (Gentzkow, Kamenica, 2014)
 2. Bargaining and Information Acquisition (Chatterjee, Dong, Hoshino, 2024)
- Future Research: Lehmann-Monotone Information Costs based on Kim [2023]

► Costly Persuasion

Conclusion

- We identify necessary and sufficient conditions for Blackwell Monotonicity.
- Under likelihood separability, we show that the sublinearity of the component function is equivalent to Blackwell Monotonicity.
- **Applications:** we apply our results to extend
 1. Costly Persuasion (Gentzkow, Kamenica, 2014)
 2. Bargaining and Information Acquisition (Chatterjee, Dong, Hoshino, 2024)
- Future Research: Lehmann-Monotone Information Costs based on Kim [2023]

► Costly Persuasion

- We identify necessary and sufficient conditions for Blackwell Monotonicity.
- Under likelihood separability, we show that the sublinearity of the component function is equivalent to Blackwell Monotonicity.
- **Applications:** we apply our results to extend
 1. Costly Persuasion (Gentzkow, Kamenica, 2014)
 2. Bargaining and Information Acquisition (Chatterjee, Dong, Hoshino, 2024)
- Future Research: Lehmann-Monotone Information Costs based on Kim [2023]

► Costly Persuasion

Thank You!

- **Posterior-based information costs**

- Entropy cost: Sims [2003]; Matějka, McKay [2015]
- Decision theory: Caplin, Dean [2015]; Caplin, Dean, Leahy [2022]; Chambers, Liu, Rehbeck [2020]; Denti [2022]
- Applications: Ravid [2020]; Zhong [2022]; Gentzkow, Kamenica [2014]

- **Experiment-based information costs**

- LLR cost: Pomatto, Strack, Tamuz [2023];
- Applications: Denti, Marinacci, Rustichini [2022]; Ramos-Mercado [2023]

Sufficient Conditions for Blackwell Monotonicity

When $m \geq 3$, there may not exist a path along which informativeness decreases

Proposition

Let

$$g = \begin{bmatrix} 4/5 & 1/5 & 0 \\ 0 & 4/5 & 1/5 \\ 1/5 & 0 & 4/5 \end{bmatrix} \in \mathcal{E}_3.$$

If $f \succeq_B g$ and $f \in \mathcal{E}_3$, then f is a permutation of I_3 or g .

- I_3 is Blackwell more informative than g , but we cannot find a path from I_3 to g along which Blackwell informativeness decreases

Sufficient Conditions for Blackwell Monotonicity

When $m \geq 3$, there may not exist a path along which informativeness decreases

Proposition

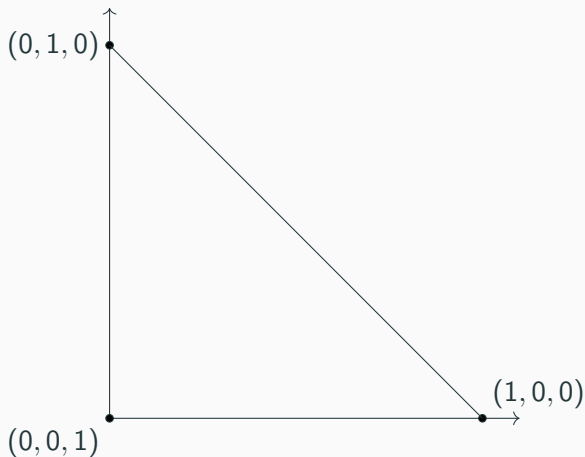
Let

$$g = \begin{bmatrix} 4/5 & 1/5 & 0 \\ 0 & 4/5 & 1/5 \\ 1/5 & 0 & 4/5 \end{bmatrix} \in \mathcal{E}_3.$$

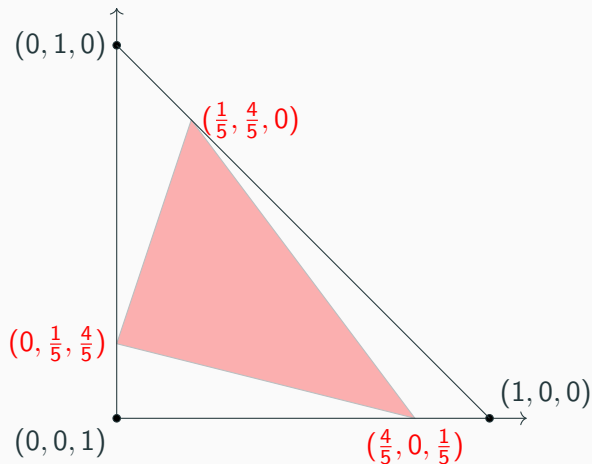
If $f \succeq_B g$ and $f \in \mathcal{E}_3$, then f is a permutation of I_3 or g .

- I_3 is Blackwell more informative than g , but we cannot find a path from I_3 to g along which Blackwell informativeness decreases

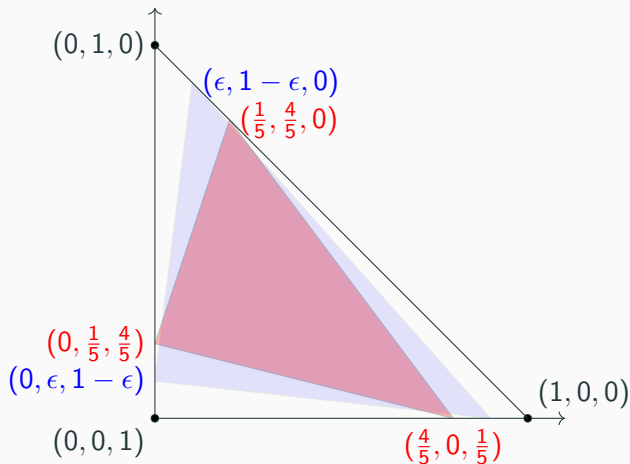
- When $n = m = 3$, $f \succeq_B g$ iff the triangle generated by f^1, f^2, f^3 includes the one generated by g^1, g^2, g^3



- When $n = m = 3$, $f \succeq_B g$ iff the triangle generated by f^1, f^2, f^3 includes the one generated by g^1, g^2, g^3



- When $n = m = 3$, $f \succeq_B g$ iff the triangle generated by f^1, f^2, f^3 includes the one generated by g^1, g^2, g^3



Quasiconvexity

- Observe that there is a permutation of l_3 such that

$$g = \frac{4}{5} \cdot l_3 + \frac{1}{5} \cdot (l_3 \cdot P).$$

- If we impose **quasiconvexity**, with permutation invariance, we have

$$C(l_3) = C(l_3 \cdot P) \geq C\left(\frac{4}{5} \cdot l_3 + \frac{1}{5} \cdot l_3 \cdot P\right) = C(g).$$

Quasiconvexity

- Observe that there is a permutation of l_3 such that

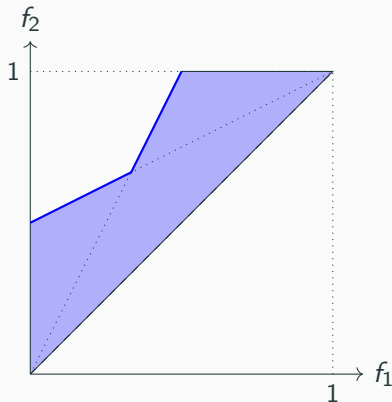
$$g = \frac{4}{5} \cdot l_3 + \frac{1}{5} \cdot (l_3 \cdot P).$$

- If we impose **quasiconvexity**, with permutation invariance, we have

$$C(l_3) = C(l_3 \cdot P) \geq C\left(\frac{4}{5} \cdot l_3 + \frac{1}{5} \cdot l_3 \cdot P\right) = C(g).$$

Quasiconvexity

- The following information cost function for binary experiments is not quasiconvex



$$C(f_1, f_2) = \min \left\{ \frac{f_2}{f_1}, \frac{1-f_1}{1-f_2} \right\}$$
$$= \min\{\alpha, \beta\}$$

Garbling Quasiconvexity

Garbling Quasiconvexity

$C \in \mathcal{C}_m$ is *garbling-quasiconvex* if for all $f \in \mathcal{E}_m$, any finite collection of its garblings, $\{g_1, \dots, g_n\}$, and $\lambda_0, \dots, \lambda_n \in [0, 1]$ with $\sum_{i=0}^n \lambda_i = 1$,

$$C\left(\lambda_0 f + \sum_{i=1}^n \lambda_i g_i\right) \leq \max\{C(f), C(g_1), \dots, C(g_n)\}$$

Theorem 4

$C \in \mathcal{C}_m$ is Blackwell monotone if and only if (i) C is permutation invariant; (ii) C is garbling-quasiconvex; and (iii) for all $f \in \mathcal{E}_m$,

[▶ Go Back](#)

$$\langle \nabla^{i \rightarrow j} C(f), f \rangle \leq 0.$$

Garbling Quasiconvexity

Garbling Quasiconvexity

$C \in \mathcal{C}_m$ is *garbling-quasiconvex* if for all $f \in \mathcal{E}_m$, any finite collection of its garblings, $\{g_1, \dots, g_n\}$, and $\lambda_0, \dots, \lambda_n \in [0, 1]$ with $\sum_{i=0}^n \lambda_i = 1$,

$$C\left(\lambda_0 f + \sum_{i=1}^n \lambda_i g_i\right) \leq \max \{C(f), C(g_1), \dots, C(g_n)\}$$

Theorem 4

$C \in \mathcal{C}_m$ is Blackwell monotone if and only if (i) C is permutation invariant; (ii) C is garbling-quasiconvex; and (iii) for all $f \in \mathcal{E}_m$,

► Go Back

$$\langle \nabla^{i \rightarrow j} C(f), f \rangle \leq 0.$$

Proof of Theorem 3

[Sublinearity \Rightarrow Blackwell Monotonicity]

- From sublinearity, we can show that C is convex.
- Consider the garbling of replacing s_j to s_k with prob. ϵ :

$$\begin{aligned}\Delta C &= \psi(f^k + \epsilon \cdot f^j) + \psi((1 - \epsilon)f^j) - [\psi(f^k) + \psi(f^j)] \\ &= \psi(f^k + \epsilon \cdot f^j) + (1 - \epsilon) \cdot \psi(f^j) - \psi(f^k) - \psi(f^j) \\ &= \psi(f^k + \epsilon \cdot f^j) - \psi(f^k) -\end{aligned}$$

Proof of Theorem 3

[Sublinearity \Rightarrow Blackwell Monotonicity]

- From sublinearity, we can show that C is convex.
- Consider the garbling of replacing s_j to s_k with prob. ϵ :

$$\begin{aligned}\Delta C &= \psi(f^k + \epsilon \cdot f^j) + \psi((1 - \epsilon)f^j) - [\psi(f^k) + \psi(f^j)] \\ &= \psi(f^k + \epsilon \cdot f^j) + (1 - \epsilon) \cdot \psi(f^j) - \psi(f^k) - \psi(f^j) \\ &= \psi(f^k + \epsilon \cdot f^j) - \psi(f^k) -\end{aligned}$$

Proof of Theorem 3

[Sublinearity \Rightarrow Blackwell Monotonicity]

- From sublinearity, we can show that C is convex.
- Consider the garbling of replacing s_j to s_k with prob. ϵ :

$$\begin{aligned}\Delta C &= \psi(f^k + \epsilon \cdot f^j) + \psi((1 - \epsilon)f^j) - [\psi(f^k) + \psi(f^j)] \\ &= \psi(f^k + \epsilon \cdot f^j) + (1 - \epsilon) \cdot \psi(f^j) - \psi(f^k) - \psi(f^j) \\ &= \psi(f^k + \epsilon \cdot f^j) - \psi(f^k) - \epsilon \cdot \psi(f^j)\end{aligned}$$

Proof of Theorem 3

[Sublinearity \Rightarrow Blackwell Monotonicity]

- From sublinearity, we can show that C is convex.
- Consider the garbling of replacing s_j to s_k with prob. ϵ :

$$\begin{aligned}\Delta C &= \psi(f^k + \epsilon \cdot f^j) + \psi((1 - \epsilon)f^j) - [\psi(f^k) + \psi(f^j)] \\ &= \psi(f^k + \epsilon \cdot f^j) + (1 - \epsilon) \cdot \psi(f^j) - \psi(f^k) - \psi(f^j) \\ &= \psi(f^k + \epsilon \cdot f^j) - \psi(f^k) - \psi(\epsilon \cdot f^j) \leq 0\end{aligned}$$

Proof of Theorem 3

[Blackwell Monotonicity \Rightarrow Sublinearity]

1. **Positive homogeneity:** Note that $\psi(0) = 0$. For any $k \in \mathbb{N}$,

$$[\hat{f}, 0, \dots, 0, 1 - \hat{f}] \sim_B [\hat{f}/k, \hat{f}/k, \dots, \hat{f}/k, 1 - \hat{f}] \Rightarrow \psi(\hat{f}) = k \psi(\hat{f}/k).$$

Then, for any $(k, l) \in \mathbb{N}^2$, we also have

$$\frac{l}{k} \psi(\hat{f}) = l \psi\left(\frac{\hat{f}}{k}\right) = \psi\left(\frac{l}{k} \hat{f}\right)$$

By density of \mathbb{Q} in \mathbb{R} and the continuity of ψ , $\psi(\alpha \hat{f}) = \alpha \psi(\hat{f})$ for all $\alpha \in \mathbb{R}_+$

2. **Subadditivity:**

$$[\hat{f}, \hat{g}, 1 - \hat{f} - \hat{g}] \succeq_B [\hat{f} + \hat{g}, 0, 1 - \hat{f} - \hat{g}] \Rightarrow \psi(\hat{f}) + \psi(\hat{g}) \geq \psi(\hat{f} + \hat{g})$$

Illustration: Quasiconvexity

- Consider COVID tests again
- Test C' is a convex combination of A' and B'

	n	p
—	60%	40%
+	20%	80%

Test A' (\$10)

	n	p
—	80%	20%
+	40%	60%

Test B' (\$12)

	n	p
—	70%	30%
+	30%	70%

Test C' ($\leq \$12$?)

- Pay someone \$12 to perform the following process:
 - toss a coin and run Test A' for heads, Test B' for tails
 - report the result without specifying which test was run

Illustration: Quasiconvexity

- Consider COVID tests again
- Test C' is a convex combination of A' and B'

	n	p
—	60%	40%
+	20%	80%

Test A' (\$10)

	n	p
—	80%	20%
+	40%	60%

Test B' (\$12)

	n	p
—	70%	30%
+	30%	70%

Test C' (\leq \$12?)

- Pay someone \$12 to perform the following process:
 - toss a coin and run Test A' for heads, Test B' for tails
 - report the result without specifying which test was run

Application: Costly Persuasion

Gentzkow, Kamenica (2014) Revisited

- Consider a costly persuasion problem with the standard example
 - State: $\{innocent, guilty\}$
 - Receiver's action: **A**cquit or **C**onvict
 - Sender's payoff: $u_S(C) = 1, u_S(A) = 0$
 - Receiver's payoff: $u_R(A, innocent) = u_R(C, guilty) = 1$
 $u_R(C, innocent) = u_R(A, guilty) = 0$
 - Sender commits to an experiment at some cost
- GK focuses on posterior separable costs (e.g., entropy cost) to utilize concavification technique
- Can we solve this problem with any Blackwell-monotone information cost function?

Gentzkow, Kamenica (2014) Revisited

- Consider a costly persuasion problem with the standard example
 - State: $\{innocent, guilty\}$
 - Receiver's action: **A**cquit or **C**onvict
 - Sender's payoff: $u_S(C) = 1, u_S(A) = 0$
 - Receiver's payoff: $u_R(A, innocent) = u_R(C, guilty) = 1$
 $u_R(C, innocent) = u_R(A, guilty) = 0$
 - Sender commits to an experiment at some cost
- GK focuses on posterior separable costs (e.g., entropy cost) to utilize concavification technique
- Can we solve this problem with any Blackwell-monotone information cost function?

Gentzkow, Kamenica (2014) Revisited

- Consider a costly persuasion problem with the standard example
 - State: $\{innocent, guilty\}$
 - Receiver's action: **A**cquit or **C**onvict
 - Sender's payoff: $u_S(C) = 1, u_S(A) = 0$
 - Receiver's payoff: $u_R(A, innocent) = u_R(C, guilty) = 1$
 $u_R(C, innocent) = u_R(A, guilty) = 0$
 - Sender commits to an experiment at some cost
- GK focuses on posterior separable costs (e.g., entropy cost) to utilize concavification technique
- Can we solve this problem with any Blackwell-monotone information cost function?

Costly Persuasion with Blackwell-Monotone Information Cost

- It is without loss to consider binary experiments since \mathbf{R} 's action is binary
 - $f_2 = \Pr(C|guilty)$ and $f_1 = \Pr(C|innocent)$
- When the prior is p , the sender's problem is

$$\max_{0 \leq f_1 \leq f_2 \leq 1} pf_2 + (1 - p)f_1 - C(f_1, f_2)$$

subject to

$$\frac{pf_2}{pf_2 + (1 - p)f_1} \geq \frac{1}{2}.$$

- When $p \geq 1/2$, the solution is $f_1 = f_2 = 1$: always convict costlessly

Costly Persuasion with Blackwell-Monotone Information Cost

- It is without loss to consider binary experiments since \mathbf{R} 's action is binary
 - $f_2 = \Pr(C|guilty)$ and $f_1 = \Pr(C|innocent)$
- When the prior is p , the sender's problem is

$$\max_{0 \leq f_1 \leq f_2 \leq 1} pf_2 + (1 - p)f_1 - C(f_1, f_2)$$

subject to

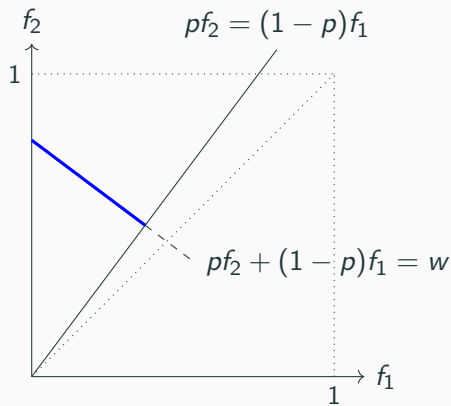
$$\frac{pf_2}{pf_2 + (1 - p)f_1} \geq \frac{1}{2}.$$

- When $p \geq 1/2$, the solution is $f_1 = f_2 = 1$: always convict costlessly

Cost Minimization

- Suppose $p < 1/2$.
- Cost minimization problem under $pf_2 + (1 - p)f_1 = w$:

$$\min C(f_1, f_2) \quad \text{s.t.} \quad \begin{aligned} pf_2 + (1 - p)f_1 &= w, \\ pf_2 &\geq (1 - p)f_1 \end{aligned}$$

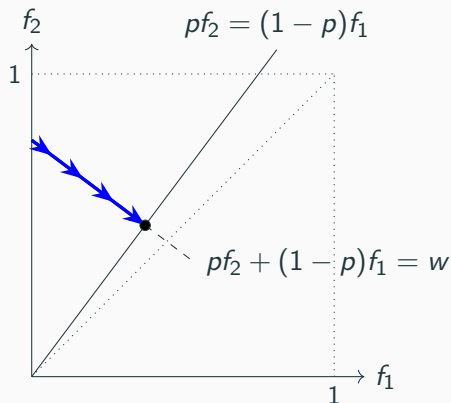


- **Proposition:** for any Blackwell-monotone information cost function, the cost is minimized when $pf_2 = (1 - p)f_1$

Cost Minimization

- Suppose $p < 1/2$.
- Cost minimization problem under $pf_2 + (1 - p)f_1 = w$:

$$\min C(f_1, f_2) \quad \text{s.t.} \quad \begin{aligned} pf_2 + (1 - p)f_1 &= w, \\ pf_2 &\geq (1 - p)f_1 \end{aligned}$$



- **Proposition:** for any Blackwell-monotone information cost function, the cost is minimized when $pf_2 = (1 - p)f_1$

Sender's Problem

- When $pf_2 + (1 - p)f_1 = w$, the cost is minimized at

$$f_2 = \frac{w}{2p} \quad \text{and} \quad f_1 = \frac{w}{2(1-p)}.$$

- Now the sender's problem is

$$\max_{0 \leq w \leq 2p} w - C\left(\frac{w}{2(1-p)}, \frac{w}{2p}\right) \quad (4)$$

- From here on, a specific cost function is needed

► Go Back

Sender's Problem

- When $pf_2 + (1 - p)f_1 = w$, the cost is minimized at

$$f_2 = \frac{w}{2p} \quad \text{and} \quad f_1 = \frac{w}{2(1-p)}.$$

- Now the sender's problem is

$$\max_{0 \leq w \leq 2p} w - C\left(\frac{w}{2(1-p)}, \frac{w}{2p}\right) \tag{4}$$

- From here on, a specific cost function is needed

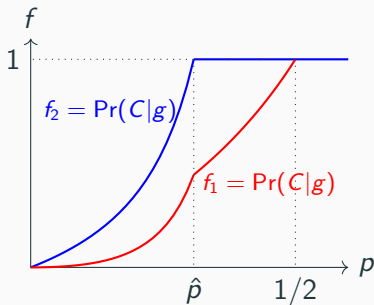
► Go Back

Costly Persuasion with Non-Posterior-Separable Cost

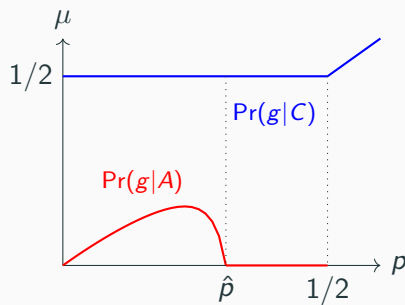
- When $C(f_1, f_2) = (f_2 - f_1)^2$, the solution for $p < 1/2$ is

$$f_2(p) = \min \left\{ 1, \frac{(1-p)^2 p}{(1-2p)^2} \right\} \quad \text{and} \quad f_1(p) = \frac{p}{1-p} \cdot f_2(p).$$

► Entropy



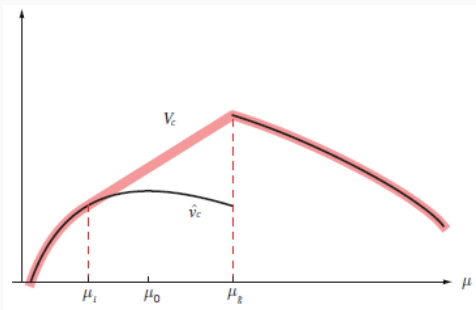
Optimal Experiments



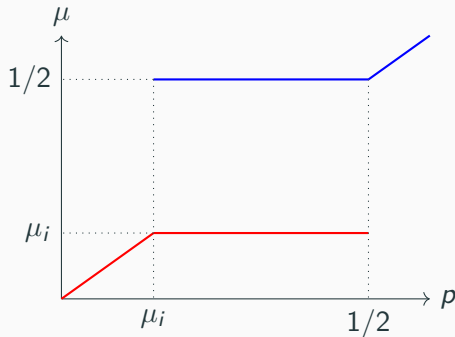
Posteriors

Gentzkow, Kamenica (2014) Revisited

- Entropy cost: $k \cdot \mathbb{E}_{\pi|p}[H(p) - H(\mu_s)]$ where $H(\mu) \equiv -\sum_{\omega} \mu(\omega) \log(\mu(\omega))$
 - p is prior, and μ_i and μ_g are posteriors from an experiment π



Concavification



Posteriors