Information Design Perspective on Calibration

PART II – APPLICATIONS

WEI TANG, CHINESE UNIVERSITY OF HONG KONG

JOINT TUTORIAL WITH YIDING FENG, HKUST

WINE 2025

Recap

Definition [Dawid, JASA'82][Foster Vohra, Biometrika'98]**.**

Predictor F is **calibrated** if for **every** prediction $q \in [0,1]$

$$\mathbb{E}[Y|F = q] = q$$

Calibrated predictor can be viewed as a signaling scheme

- **Data distribution** $D \in \Delta(\mathcal{X} \times \{0, 1\})$
  - State Space: $\{p_X\}_{X\in\mathcal{X}}$ where $p_X = \mathbb{P}(Y = 1 \mid x = X)$
  - Prior: $\{\mathbb{P}_{x\sim D}(x = X)\}_{X\in\mathcal{X}}$
- **Predictor:** $F: \mathcal{X} \mapsto \Delta([0, 1])$ (or equivalently $\{p_X\}_{X\in\mathcal{X}} \to \Delta([0, 1])$) is a **signaling scheme over posterior means**
  - The prediction q is a signal
  - Calibration requires that $q =$ induced **posterior mean**

Applications

- InfoGap to compare different predictors

- Calibrated signaling in digital auctions

Key technical component: formulating the optimization problem as an optimal transportation problem.

Motivation: Comparison between Multiple Predictors

In reality, platform choose between **multiple predictors** **for their downstream users**

Amazon Forecast Algorithms

PDF | RSS

An Amazon Forecast predictor uses an algorithm to train a model with your time series datasets. The trained model is then used to generate metrics and predictions.

Amazon Forecast provides six built-in algorithms for you to choose from. These range from commonly used statistical algorithms like Autoregressive Integrated Moving Average (ARIMA), to complex neural network algorithms like CNN-QR and DeepAR+.

CNN-QR → Convolutional Neural Network

`arn:aws:forecast:::algorithm/CNN-QR`

Amazon Forecast CNN-QR, Convolutional Neural Network - Quantile Regression, is a proprietary machine learning algorithm for forecasting time series using causal convolutional neural networks (CNNs). CNN-QR works best with large datasets containing hundreds of time series. It accepts item metadata, and is the only Forecast algorithm that accepts related time series data without future values.

DeepAR+ → Recurrent Neural Network

`arn:aws:forecast:::algorithm/Deep_AR_Plus`

Amazon Forecast DeepAR+ is a proprietary machine learning algorithm for forecasting time series using recurrent neural networks (RNNs). DeepAR+ works best with large datasets containing hundreds of feature time series. The algorithm accepts forward-looking related time series and item metadata.

Prophet → Time-series forecaster

`arn:aws:forecast:::algorithm/Prophet`

Prophet is a time series forecasting algorithm based on an additive model where non-linear trends are fit with yearly, weekly, and daily seasonality. It works best with time series with strong seasonal effects and several seasons of historical data.

Motivation: Comparison between Multiple Predictors

In reality, platform choose between **multiple predictors** **for their downstream users**

Show 10 entries Search:

Rank	Org	Model	Overall
1		Superforecaster median forecast	0.093
2		Public median forecast	0.129
3		Gemini-2.5-Flash-Preview-04-17 (zero shot)	0.130
4		O3-2025-04-16 (zero shot)	0.135
5		GPT-4.5-Preview-2025-02-27 (zero shot)	0.136
5		Claude-3-7-Sonnet-20250219 (scratchpad)	0.136
5		O3-2025-04-16 (scratchpad)	0.136
8		GPT-4.5-Preview-2025-02-27 (scratchpad)	0.140
9		DeepSeek-R1 (scratchpad)	0.143
9		Grok-beta (zero shot)	0.143

Showing 1 to 10 of 83 entries
last updated 2025-12-07
Previous Next

ForecastBench

Market Return

Average Return measures the decision value of a probabilistic prediction by simulating the expected profit of an optimal betting strategy based on the prediction, under the market conditions at the time of prediction and a specified level of risk aversion.

Rank	Model	Provider	Events	Average Return	Confidence Interval (90% CI)
1	GPT-4o	OpenAI	2,011	**90.87%**	±0.1240
2	o3	OpenAI	2,069	**90.30%**	±0.1310
3	GPT-5 (high)	OpenAI	2,037	**88.19%**	±0.0500
4	Qwen 3 235B	Qwen	2,051	**87.67%**	±0.1300
5	GPT-4.1	OpenAI	2,008	**86.95%**	±0.0467
6	Claude Sonnet 4 (Thinking)	Anthropic	2,102	**86.92%**	±0.1301

Prophet Arena

Desiderata for Comparisons

- A decision-making task is specified by $(\mathcal{A}, u)$
  - $\mathcal{A}$ is action set,
  - Decision maker's utility function: $u: \mathcal{A} \times \{0,1\} \to \mathbb{R}$

- Can we say a predictor is **always** useful than another predictor?
  - DM's Payoff from **acting by trusting** a predictor F: $U(F) = \mathbb{E}_{Y,q\sim F}[u(a^*(q), Y)]$
  - Ordinal comparisons (Partial order)

- Can we bound, to what extent, how much a predictor is **worse** than another predictor?
  - Cardinal comparison

Example I

Suppose $Y \sim \mathrm{Bern}(0.5)$.

	Predictors	ECE	"Usefulness for decision making"
	Predict mean $\mathbb{E}[Y]$	0	
	Predict 1 iff $Y = 1$	0	
	Predict 1, 0 uniformly	0.5	
	Predict 0 iff $Y = 1$	1	

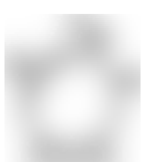
Example I

Takeaways

Even with **same ECE**, one predictor may **dominate** another predictor for decision making.

Example II:

Suppose $Y \sim \mathrm{Bern}(0.5)$.

	Predictors	ECE	"Usefulness for decision making"
	Predict mean $\mathbb{E}[Y]$	0	
	Predict 1 iff $\color{green}{Y = 1}$	0	
	Predict 1 w.p. 0.99 if $\color{green}{Y = 1}$ Predict 0 w.p. 0.99 if $\color{red}{Y = \mathbf{0}}$ Predict $\mathbb{E}[Y] - 0.001$ otherwise	> 0	

Example II:

Takeaways

One **miscalibrated** predictor may **dominate** another calibrated predictor for decision making.

[Questions]

Can we **compare** any two (possibly miscalibrated) predictors based on how "useful" they are to the **decision-making problems**?

Our Main Results

Our Results [Informal]. We provide a measure, referred to as ***informativeness gap*** between **any** two (**possibly miscalibrated**) predictors, that allows both **ordinal/cardinal comparisons**

Informativeness Gap

Definition. [Feng, Qian & Tang, arXiv'25]

Given two predictors F and G, **informativeness gap** $\mathrm{INFOGAP}[F, G]$ of G relative to F is

$$\mathrm{INFOGAP}[F, G] \stackrel{\text{def}}{=} \sup_{u \in \mathcal{U}} U(F) - U(G)$$

where

- $\mathcal{U}$: all decision tasks with bounded utility differences
- $U(F)$: expected payoff by naively best responding to prediction $p \sim F$

$$U(F) = \mathbb{E}_{Y, q \sim F}[u(a^*(q), Y)]$$

- In general, $\mathrm{INFOGAP}[F, G] \neq \mathrm{INFOGAP}[G, F]$
- If $\mathrm{INFOGAP}[F, G] \to 0 \Longrightarrow$ predictor G is more **useful** than F, or **more informative**
- $\mathrm{UCal}[G] = \mathrm{INFOGAP}\left[\delta_{(\lambda)}, G\right]$ where $\lambda = \mathbb{E}_{p \sim G}[p]$ Kleinberg, Leme, Schneider, Teng COLT'23
- $\mathrm{CDL}[G] = \mathrm{INFOGAP}[G^{\mathrm{Bayes}}, G]$ where G^{Bayes} is the corresponding true distribution induced by G Hu & Wu, FOCS'24

Blackwell's Informativeness

Blackwell's Informativeness: When F, G are **calibrated**, G **Blackwell dominants** F if and only if $U(G) \geq U(F)$ for all decision task u

[Blackwell, Annals of Mathematical Statistics'53]

- When F, G are **calibrated**, $\mathrm{INFOGAP}[F, G] = 0 \Leftrightarrow G$ **Blackwell dominants** F
  - Blackwell order: **ordinal** comparison (**partial order**) over **calibrated** predictors
- InfoGap: **cardinal** comparison over **any two possibly miscalibrated** predictors

Dual Characterization of Infor Gap

Theorem I [Feng, Qian & Tang, arXiv'25] Given two **calibrated** predictors F and G, $\mathrm{INFOGAP}[F,G]$ equals to corresponding **relaxed earth mover's distance**:

$$\mathrm{INFOGAP}[F,G] = \mathrm{REMD}[f,g] \overset{\mathrm{def}}{=} \inf_{\pi\in\Pi(f,g)} \int_0^1 \left| \int_0^1 \pi(p,q)\cdot(p-q)\mathrm{d}q \right| \mathrm{d}p$$

where

- f, g: prediction distribution PDF of predictors F, G
- $\Pi(f,g)$: all couplings (matching marginal) for distributions f, g

- Recall classic earth mover's distance (aka., Wasserstein distance)

$$\mathrm{EMD}[f,g] \overset{\mathrm{def}}{=} \inf_{\pi\in\Pi(f,g)} \int_0^1 \int_0^1 \pi(p,q)\cdot|p-q|\cdot\mathrm{d}q\cdot\mathrm{d}p$$

Thus, $\mathrm{REMD}[f,g] \le \mathrm{EMD}[f,g]$

Closed-Form for REMD

Theorem II. [Feng, Qian & Tang, arXiv'25]

Given two **calibrated** distribution f, g with [0,1] support and identical means,

$$\mathrm{REMD}[f, g] = 2 \cdot \max_{t \in [0,1]} S_f(t) - S_g(t)$$

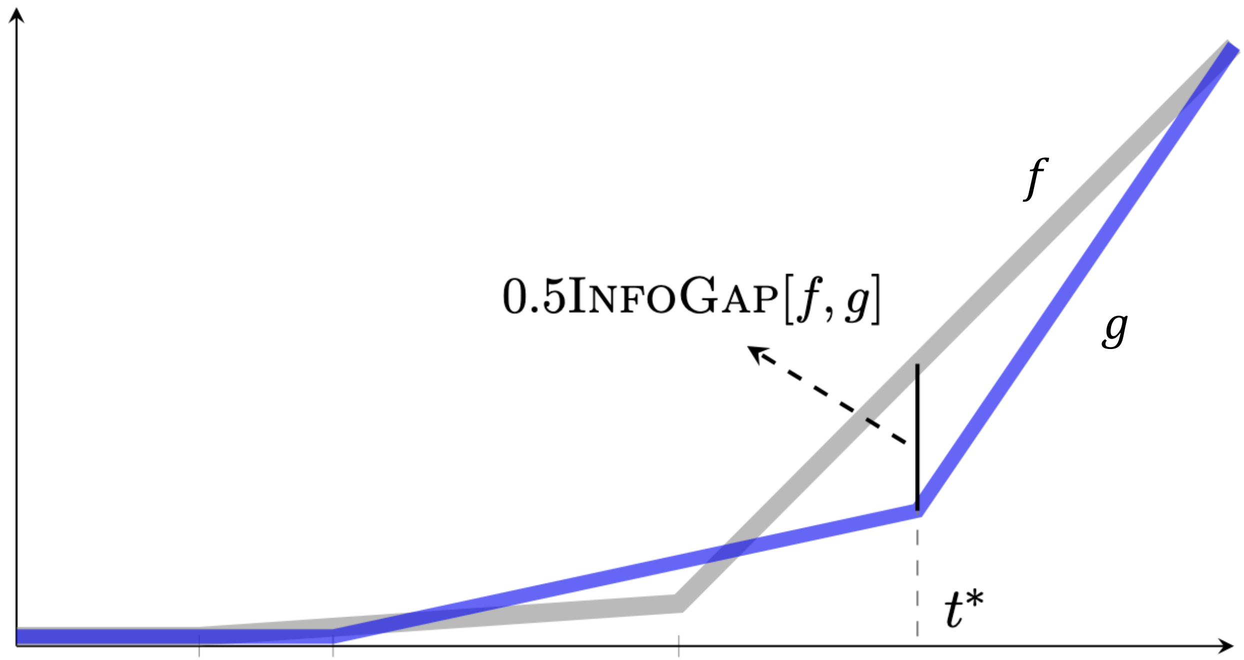
where

- $S_f(t)$ is super-cumulative distribution function (SCDF) defined as

$$S_f(t) \stackrel{\text{def}}{=} \int_0^t \int_0^s f(z) \cdot \mathrm{d}z \cdot \mathrm{d}s$$

- Implication [Blackwell'53]: g Blackwell dominates f iff $S_g(t) \geq S_f(t)$ for all t
- Implication: $\mathrm{REMD}[f, g]$ (and $\mathrm{INFOGAP}[F, G]$) admits **polynomial** (in $1/\epsilon$) **time complexity** and **sample complexity**

A graphic illustration of $\mathrm{REMD}[f, g]$



$$\mathrm{REMD}[f, g] = 2 \cdot \max_{t \in [0,1]} S_f(t) - S_g(t)$$

- We also generalize **Theorem I** and **Theorem II** when F, G are possibly miscalibrated

Characterization of Infor Gap

Theorem III [Feng, Qian & Tang, arXiv'25] Given two **miscalibrated** predictors F and G,

$$\mathrm{INFOGAP}[F,G] = \mathrm{REMD}^{\mathrm{Mis}C}[f,g,\kappa_f,\kappa_g] \stackrel{\text{def}}{=}$$

$$\inf_{\pi\in\overline{\Pi}(f,g)} \int_0^1 \left| \int_0^1 \pi(p,q)\cdot(p-q)\mathrm{d}q + \big(\kappa_f(p)-p\big)\cdot f(p) - \big(\kappa_g(p)-p\big)\cdot g(p) \right| \mathrm{d}p$$

where

- $\overline{\Pi}(f,g)$: all **flow coupling** for distributions f,g

$$\overline{\Pi}(f,g) = \left\{\pi \in \Delta([0,1]\times[0,1]) : f(p) - g(p) - \int_0^1 \pi(p,q)\mathrm{d}q + \int_0^1 \pi(q,p)dq = 0\right\}$$

- $\kappa_f(p) = \mathbb{E}_f[Y \mid p]$: true probability underlying prediction p

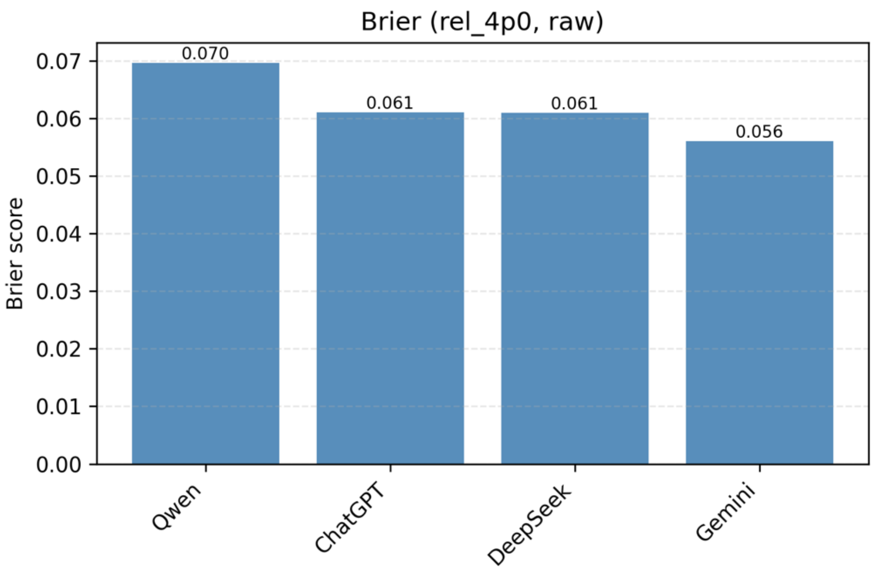
Theorem IV [Feng, Qian & Tang, arXiv'25]

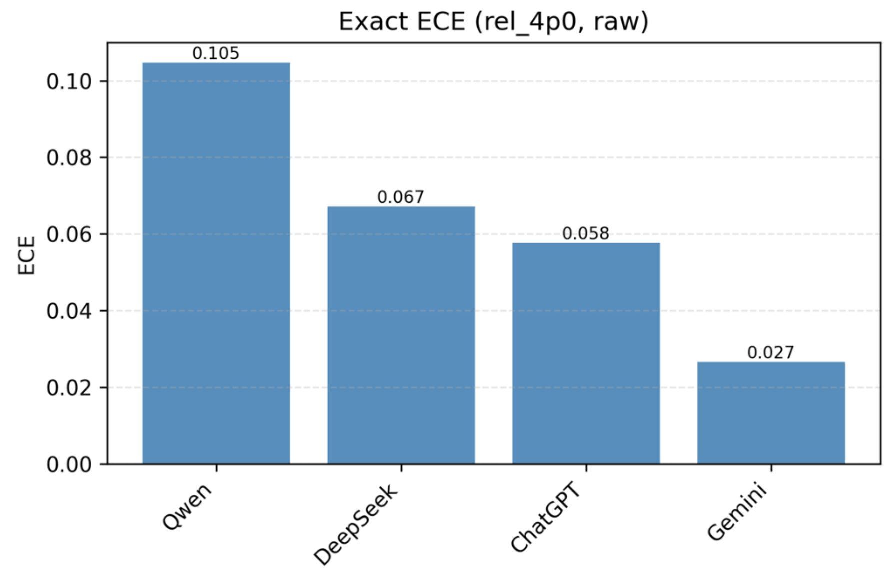
Given two **miscalibrated** distribution f,g with [0,1] support and identical means,

$$\mathrm{REMD}^{\mathrm{MisC}}[f,g] = 2\cdot \max_{t\in[0,1]} \left(S_f(t) + \int_0^t \big(p-\kappa_f(p)\big)\cdot f(p)dp\right) - \left(S_g(t) + \int_0^t \big(p-\kappa_g(p)\big)\cdot f(p)\mathrm{dp}\right)$$

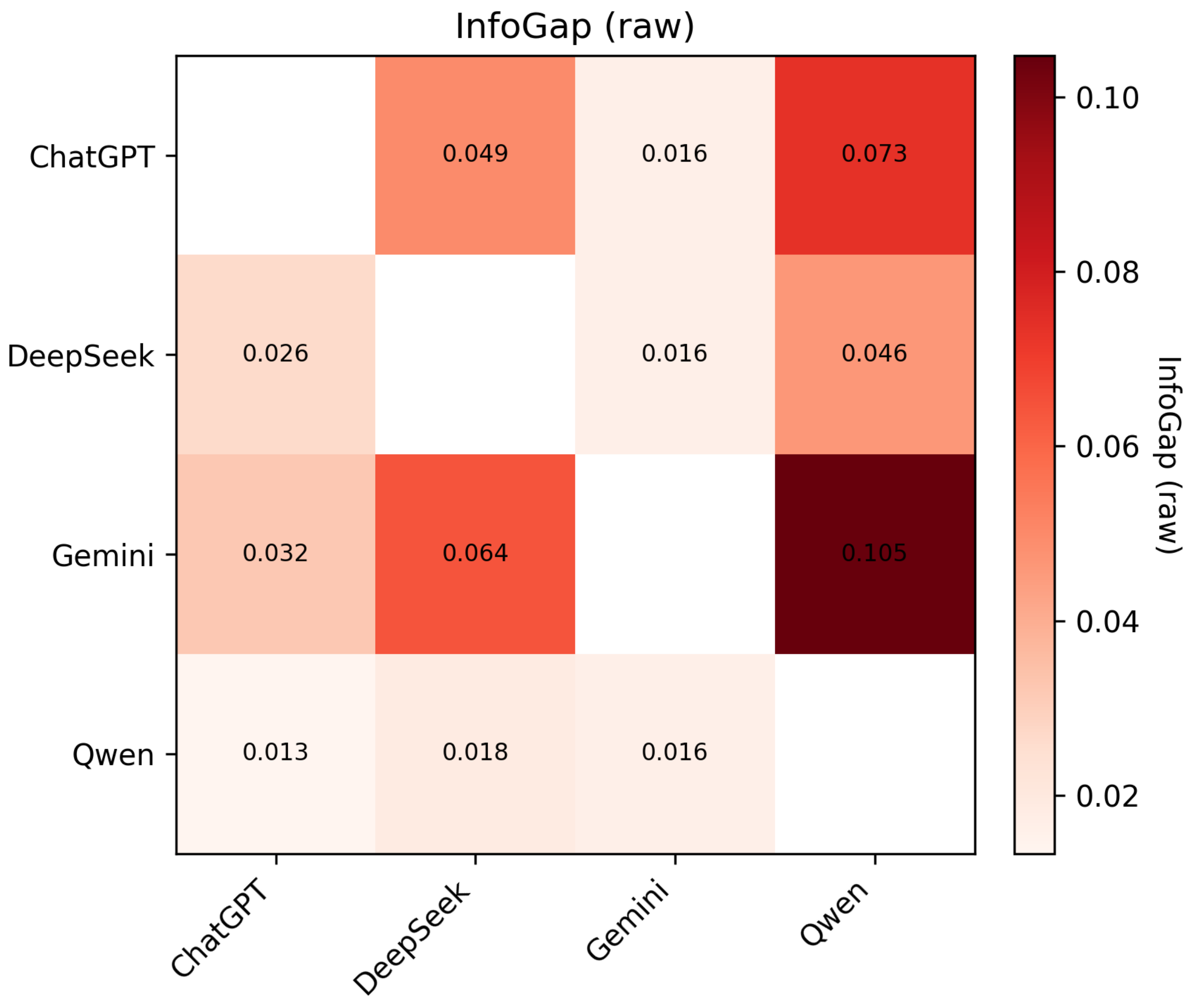
InfoGap is a More Informative Criterion

- Prediction task: daily Bitcoin closing price increase by 4% vs previous day?
- LLM models: DeepSeek, Qwen, Gemini (2.0-flash-lite), ChatGPT (4o-mini)





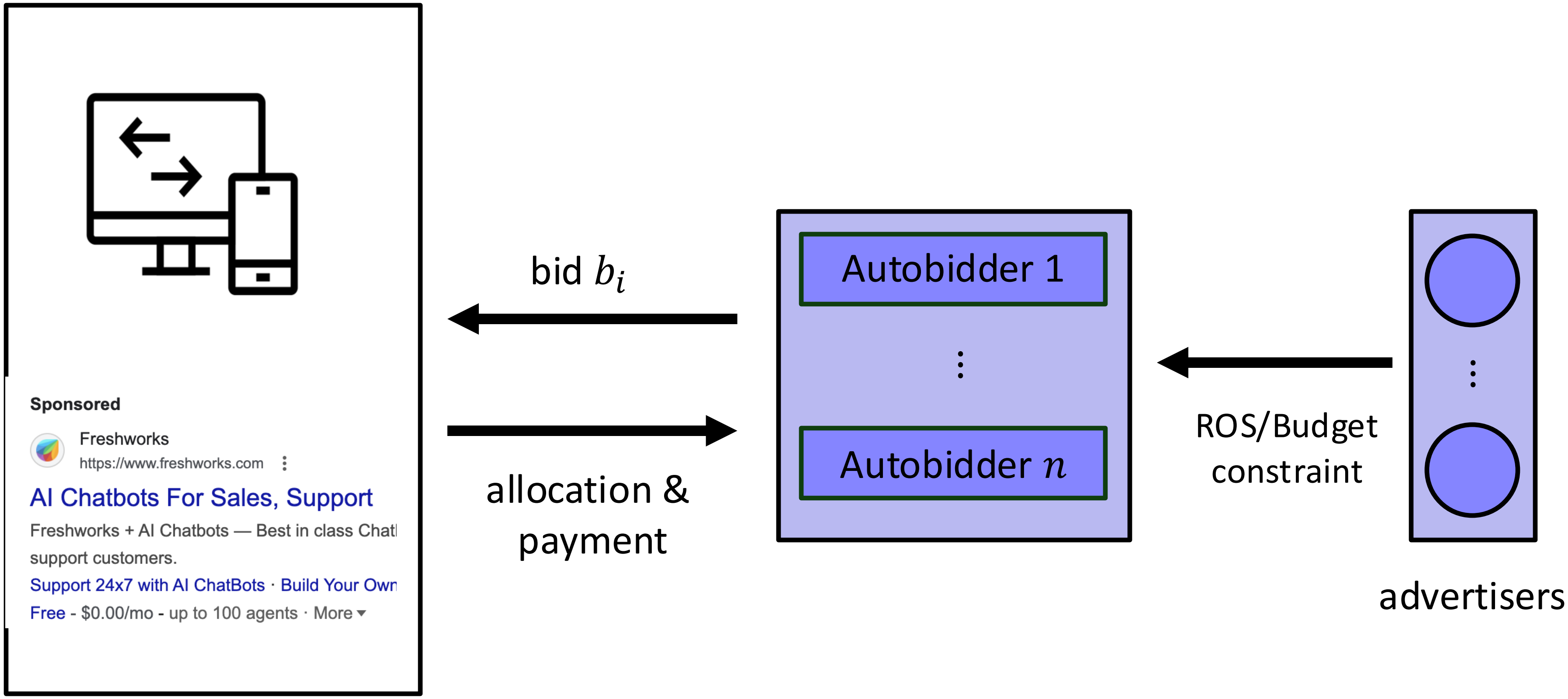
InfoGap is a More Informative Criterion



Applications

- InfoGap to compare different predictors
- Calibrated signaling in digital auctions

Digital Auctions for Ad Impression



2nd –price auction

A Simple yet Effective Bidding Strategy

Uniform bidding strategy $b_i \propto c_i \cdot v_i$

- Optimality in many (truthful) auction format [BBW MS'15, ABM WINE'19, DMMZZ WWW'23]

- Performs robustly well v.s optimal non-uniform bidding in many ad auctions [FPMS EC'07, BG MS'19, BFMW EC'14, DLMZ WWW'20, DMMZ WWW'21]

- **Challenge One**: Value v_i typically depends on CTR, and it's unknown to autobidder
  - Solution: Platform needs to inform CTRs to autobidders

How to Credibly Inform CTRs?

Google Ads R
Glitch Inflates
Through Rate

Free PPC Audit

May 13, 2025

Table of Contents

Google Ads Reporting Glitch: Google has recently acknowledged a significant reporting bug within its Google Ads platform that has resulted in inflated click-through rates (CTR) for some advertisers. This anomaly has raised concerns among users who rely on accurate data for their advertising strategies. The issue was highlighted by industry expert Brad Geddes, who shared his findings on LinkedIn, prompting a response from Google Ads Liaison Ginny Marvin.

Meta Platforms must face advertisers' class action, US appeals court says

By **Jonathan Stempel**

March 22, 2024 9:05 AM GMT+8 · Updated March 22, 2024

Aa

March 21 (Reuters) - A divided U.S. appeals court said Meta Platforms (META.O) must face a class action by advertisers that accused the Facebook and Instagram owner of overcharging them by fraudulently inflating the number of people their ads might reach.

In a 2-1 decision on Thursday, the 9th U.S. Circuit Court of Appeals in San Francisco said advertisers could sue for damages as a group over Meta's claims about the "potential reach" of their ads.

- **Challenge Two**: How to calculate CTRs is typically generated through platform's complex internal machine learning algorithms, which are usually considered as trade secrets.
  - Solution: Calibrated signaling

Calibrated Signaling in 2nd-Price Auction

Seller:

- Single-item 2nd-price auction for a finite of n bidders
- Click outcome $\vec{o} = (o_1, o_2, \cdots, o_n) \sim \lambda \in \Delta(\{0,1\}^n)$ **(product dist.)**
  - Seller knows λ, designs a **calibrated signaling $\boldsymbol{\pi}$**:
    - Given outcomes $\vec{o} \sim \lambda$, sends signals $\vec{s} = (s_1, s_2, \cdots, s_n) \sim \pi(\cdot | \vec{o})$
  - Each bidder **privately** receives a signal s_i

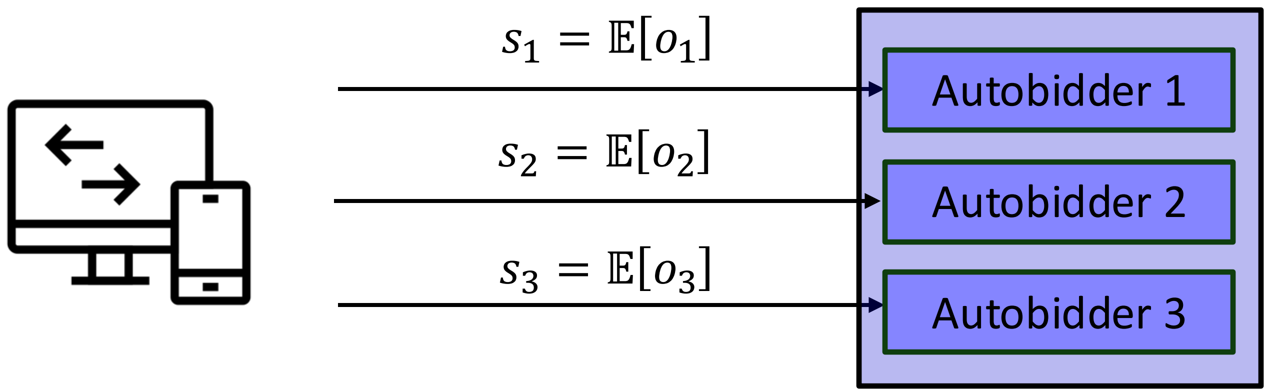
Calibrated signaling: A signaling is calibrated iff:
for each bidder i and every possible signal s_i: $\mathbb{E}(\, o_i \mid s_i \,) = s_i$,

Autobidders:

- Prior-free: neither know λ nor the signaling details π
- Knowing π is calibrated, simply bid $c_i \cdot s_i$
  - Normalize all $c_i \equiv 1$

Examples of Calibrated Signaling

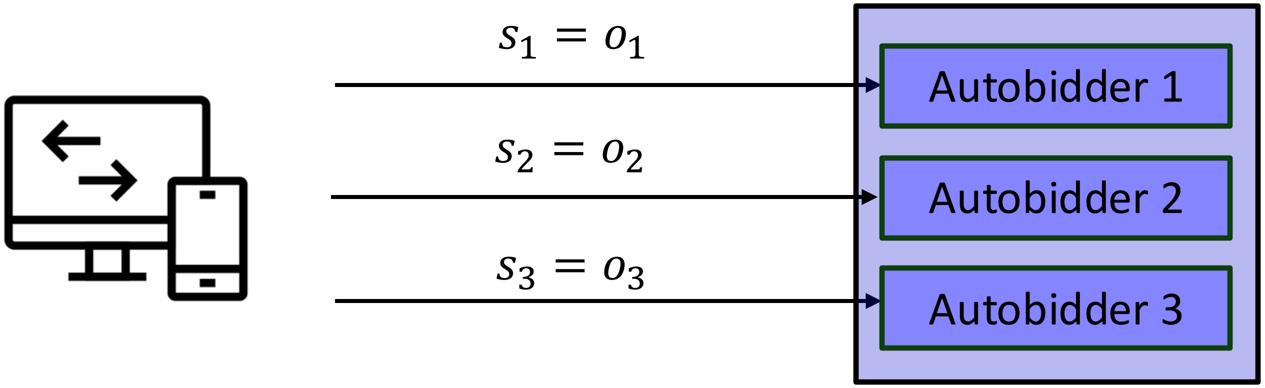
- Let $n = 3, \mathrm{Prob}[o_i = 1] = 0.5$
  - No information signaling: $s_i = \mathbb{E}[o_i]$ for all i



$\mathrm{Rev}(\text{No information signaling}) = \mathbb{E}[o_i] = 0.5$

Examples of Calibrated Signaling

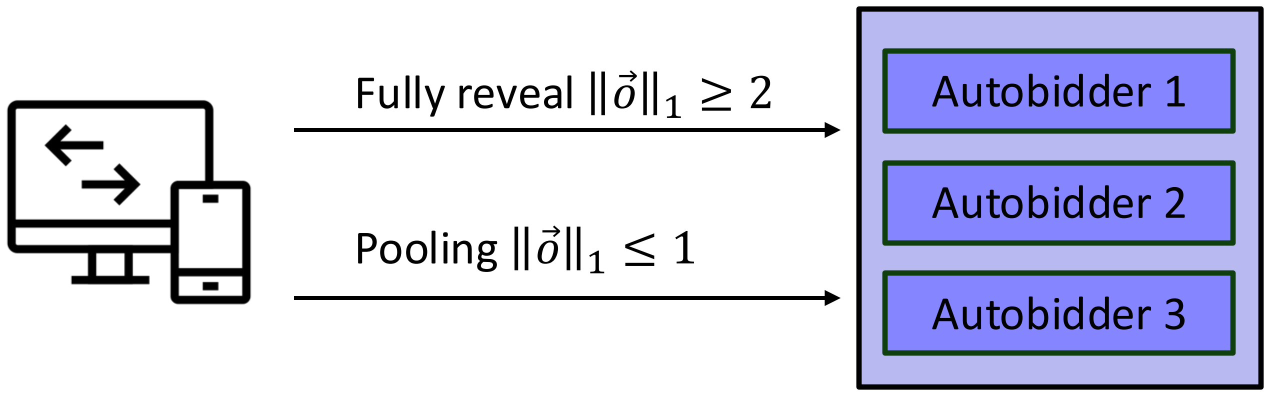
- Let $n = 3$, $\mathrm{Prob}[o_i = 1] = 0.5$
  - Fully information signaling: $s_i = o_i$ for all i



$\mathrm{Rev}(\text{Fully information signaling}) = \mathrm{Prob}(\|\vec{o}\|_1 \geq 2) = 0.5$

Examples of Calibrated Signaling

- Let $n = 3$, $\text{Prob}[o_i = 1] = 0.5$
  - For $\|\vec{o}\|_1 \geq 2$: $\pi(\vec{s} \mid \vec{o}) = 1$ where $\vec{s} = \vec{o}$
  - For $\|\vec{o}\|_1 \leq 1$: $\pi(\vec{s} \mid \vec{o}) = 1$ where $\vec{s} = \left(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right)$



$$\text{Rev(Fully +Pooling)} = \text{Prob}(\|\vec{o}\|_1 \geq 2) \cdot 1 + \text{Prob}(\|\vec{o}\|_1 \leq 1) \cdot \frac{1}{4}$$
$$= 0.5 \cdot 1 + 0.5 \cdot \frac{1}{4} > 0.5$$

$$\text{Rev(Optimal calibrated signaling)} = 0.729$$

Seller's Problem

Revenue Maximizing:

$$\pi^\star = \arg \sup_{\text{calibrated } \pi} \mathbb{E}_{\vec{o}\sim\lambda, \vec{s}\sim\pi(\cdot|\vec{o})}[\text{secmax}(\vec{s})].$$

- secmax($\vec{s}$): second-highest value
- calibration constraint:

$$s = \frac{\sum_{\vec{o}:o_i=1} \lambda(\vec{o}) \int_{s_{-i}} \pi\left((s, s_{-i}) \mid \vec{o}\right) \mathrm{d}s_{-i}}{\sum_{\vec{o}\in\{0,1\}^n} \lambda(\vec{o}) \int_{s_{-i}} \pi\left((s, s_{-i}) \mid \vec{o}\right) \mathrm{d}s_{-i}}, \quad i \in [n], s \in [0,1].$$

- Infinite-dimensional Linear program
- Private information design with n receivers
  - Every $\vec{o}$ is a state
  - Exponential number of states

Related Work

Autobidding: [Survey by Aggarwal et al. SIGecom Exchanges'24] …

Signaling in Auctions:

- In (generalized) 2nd -price auction [Bro Miltersen, Sheffet. EC'12], [Emek et al. TEAC'14], [Badanidiyuru, Bhawalkar, Xu. SODA'18], [Bergemann et al. AER Insights'22], [Bergemann, Duetting, Paes Leme, Zuo. WWW'22], [Chen et al. ICALP'24] …
  - [Bergemann et al. AER Insights'22] considers independent signaling, our work extends to general signaling.
- Joint design of auction and signaling [Bergemann, Pesendorfer. JET'07], [Cai, Li, Wu. EC'24] …

Private information design: [Dughmi and Xu, EC'17], [Arieli, Babichenko. ITCS'19 & JET'22] …

Feasible joint posterior belief: [Morris, 2020],[Brooks et al., ECMA'22], [Arieli et al. EC'20 & JPE'21], [Arieli, Babichenko. EC'22], [Arieli, Babichenko, Sandomirskiy. EC'22], [He, Sandomirskiy, Tamuz. JPE'25], [Yang and Yang. EC'25]

  - As noted in [Arieli et al. EC'20 & JPE'21] , characterizing extreme points of feasible joint posterior belief still remains an open question.

Structural Characterization of $\pi^\star$

Theorem [Du, Tang, Wang & Zhang, arXiv'25] The seller-optimal calibrated signaling $\pi^\star$ satisfies that:

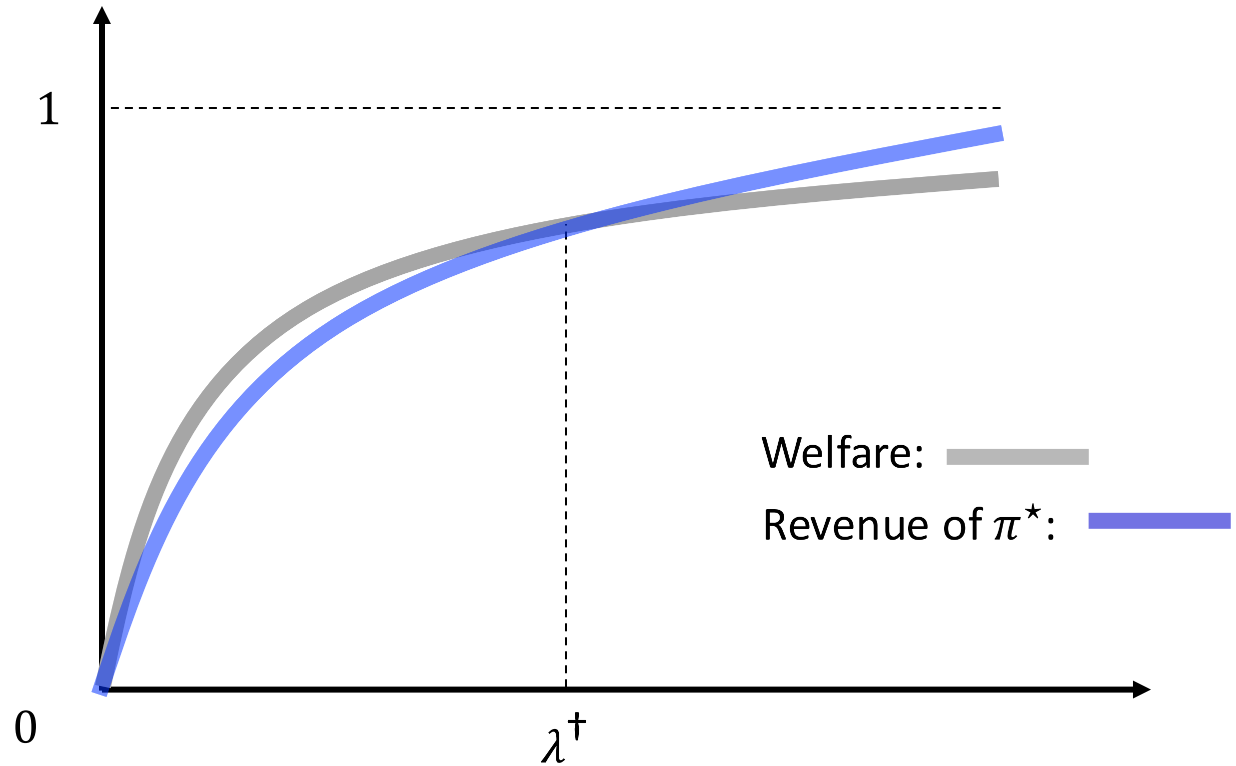
1. **[Equal highest- and 2nd-highest bid]** its every signal profile $\vec{s} \in \mathrm{supp}(\pi^\star)$ has the same highest and second-highest bid
2. **[Four signals suffice]** the induced 2nd-highest bid dist. satisfies:

$$\mathrm{Prob}_{\vec{s}\sim\pi^\star(\cdot|\vec{o})}(\mathrm{secmax}(\vec{s})) = \begin{cases} \delta_{(1)}, & \|\vec{o}\|_1 \geq 2; \\ \delta_{(t_1^*)}, & \|\vec{o}\|_1 = 1; \\ \delta_{(t_0^*)}, & \|\vec{o}\|_1 = 0. \end{cases}$$

where t_1^*, t_0^* are solved via a **linear system** $\mathrm{poly}(n, \lambda)$

- We also characterize optimal calibrated signaling with additional IR constraint

Revenue Characterization



Proof Ideas for Computing $\pi^\star$

Step 1: Symmetrizing calibrated signaling

Lemma: any calibrated signaling π can be **symmetrized** to be $\bar{\pi}$ satisfying following property without hurting any revenue: For any $\|\vec{o}\|_1 = k$, we have

$$\mathrm{Prob}_{\vec{s}\sim\bar{\pi}(\cdot|\vec{o})}(s_i = s) = f_{k,o_i}(s)$$

$\vec{o} = (1,0, \dots, 1)$
signal
$\vec{s}$
$\vec{s}'$
$\vdots$

$\vec{o} = (0,0, \dots, 1)$
signals
$\vec{s}$
$\vec{s}'$
$\vdots$

$\vec{o} = (0,1, \dots, 1)$	
signals	$\pi(\vec{s} \mid \vec{o})$
$\vec{s}$	0.2
$\vec{s}'$	0.6
$\vdots$	$\vdots$

…

For any $\vec{o}$ with $\|\vec{o}\|_1 = k$		
bid	$f_{k,o_i=1}(s)$	$f_{k,o_i=0}(s)$
s	0.3	0.5
s'	0.1	0.2
$\vdots$	$\vdots$	$\vdots$

$\pi: \{0,1\}^n \to \Delta([0,1]^n)$ $\longrightarrow$ $\bar{\pi}: [n] \to \Delta([0,1]) \times \Delta([0,1])$

Proof Ideas for Computing $\pi^\star$

Step 2: Reformulated as a two-stage optimization with **optimal transport**

$$\mathrm{REV}\left[\pi^\star\right] = \max_{(f_{k,1}, f_{k,0})_{k\in[n]_0} \in \mathcal{F}} \sum\nolimits_{k\in[n]_0} \lambda_k \cdot \mathrm{REV}^{\mathrm{CoRR}}\left[f_{k,1}, f_{k,0}\right]$$

Solving for
optimal marginals

Solving for
optimal correlation

The calibration constraint on $\bar{\pi}$ can be reduced to a calibration constraint on $\left(f_{k,1}, f_{k,0}\right)_{k\in[n]}$

$$\left.\begin{array}{r} k \text{ copies } f_{k,1} \in \Delta([0,1]) \\ n-k \text{ copies } f_{k,0} \in \Delta([0,1]) \end{array}\right\} \longrightarrow \phi^{\mathrm{CoRR}} \in \Delta([0,1]^n)$$

$$\mathrm{REV}^{\mathrm{CoRR}}\left[f_{k,1}, f_{k,0}\right] := \max_{\phi^{\mathrm{Corr}}} \mathbb{E}_{\vec{s}\sim\phi^{\mathrm{CoRR}}}\left[\mathrm{secmax}(\vec{s})\right]$$

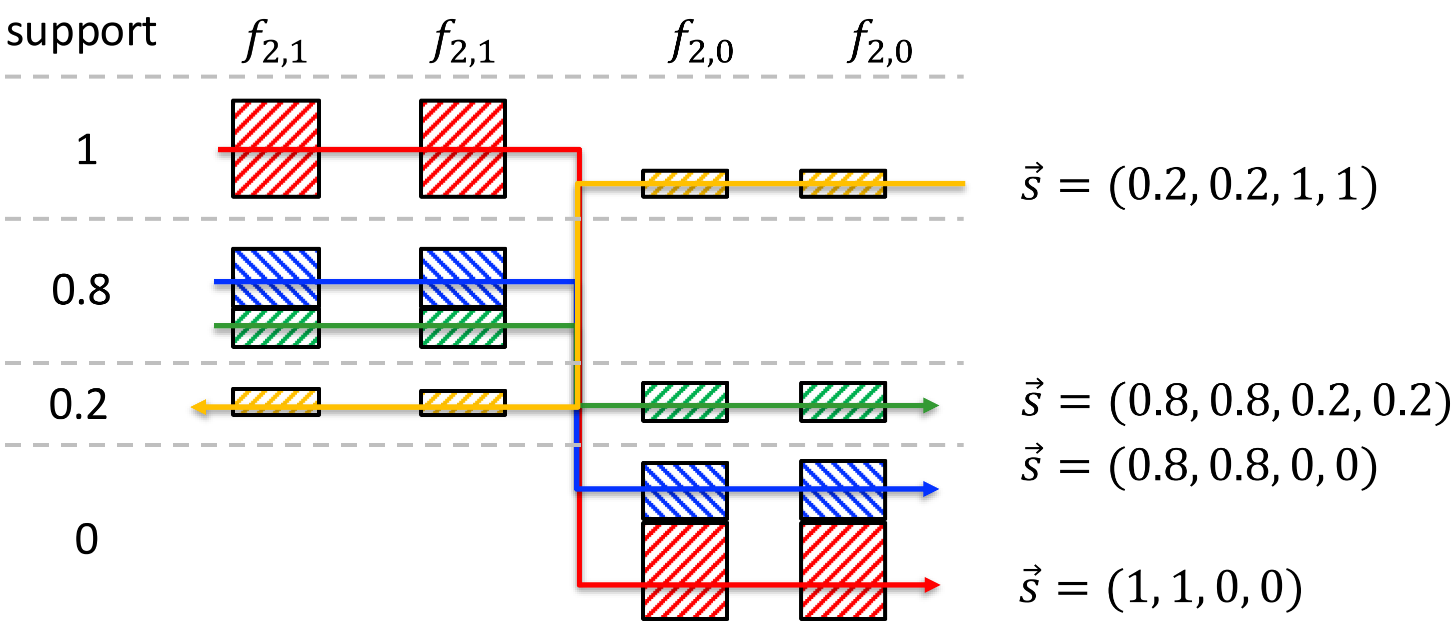
Proof Ideas for Computing $\pi^\star$

Step 3: Solving the optimal correlation:

Lemma: Fix any $k \in [n]$, $\exists$ a "greedy algo." that finds optimal correlation $\phi^{\mathrm{CoRR}} \in \Delta([0,1]^n)$ s.t. it maximizes

$$\mathrm{REV}^{\mathrm{CoRR}}[f_{k,1}, f_{k,0}] \coloneqq \max_{\phi^{\mathrm{CoRR}}} \mathbb{E}_{\vec{s} \sim \phi^{\mathrm{CoRR}}}[\mathrm{secmax}(\vec{s})]$$

$n = 4$
$k = 2$



Proof Ideas for Computing $\pi^\star$

Step 3: Solving the optimal correlation:

Lemma: Fix any $k \in [n]$, $\exists$ a "greedy algo." that finds optimal correlation $\phi^{\text{CoRR}} \in \Delta([0,1]^n)$ s.t. it maximizes

$$\text{REV}^{\text{CoRR}}\left[f_{k,1}, f_{k,0}\right] \coloneqq \max_{\phi^{\text{CoRR}}} \mathbb{E}_{\vec{s}\sim\phi^{\text{CoRR}}}[\text{secmax}(\vec{s})]$$

Step 4: With optimal correlation, then solve for optimal marginals:

Lemma: Given the optimal correlation plan, the optimal marginals can be solved via a **linear system** with size $\text{poly}(n, \sum_i \text{bit}(\lambda_k))$ where $\lambda_k = \text{Prob}(\|\vec{o}\|_1 = k)$, and $\text{bit}(\cdot)$ denotes the bit complexity.

Summary

- An intrinsic connection between **calibration** and **information design**
  - **Comparison principle**: How to **compare** different predictors?
  - **Design principle**: How to **design** predictors?

- Many interesting questions:
  - More properties of InfoGap for deterministic predictors?
  - More combinatorial structure on the space of predictors?
  - Beyond binary outcome?

Thanks!

Questions?

Please send us an email for any questions/comments:

wtang2359@gmail.com
ydfeng@ust.hk