Multi-Dimensional Screening with Rich Consumer Data

Mira Frick (Yale)
Ryota Iijima (Yale)
Yuhta Ishii (Penn State)

International Workshop on Misspecified Learning and Beyond
University of Tokyo
February 15, 2024

Motivation

Multi-dimensional screening

Motivation

Multi-dimensional screening

- optimal mechanism: general characterization computationally difficult.

Motivation

Multi-dimensional screening

- optimal mechanism: general characterization computationally difficult.
- mechanisms used in practice are often "simple"
  - although optimal mechanisms are more complicated.

Approach

- Study multi-dimensional screening where seller observes **rich data** about consumer's type
  - e.g., technological advances in data collection/analysis.

Approach

- Study multi-dimensional screening where seller observes **rich data** about consumer's type
  - e.g., technological advances in data collection/analysis.
- Study **convergence rate** of mechanisms: how fast does the seller's revenue approximate first-best with rich data
  - natural efficiency measure of mechanisms in data-rich settings

Main Questions:

1. What is the **optimal** convergence rate?
2. Can simple mechanisms achieve the optimal convergence rate?

Main Results

1. Optimal convergence rate determined by **Fisher information**

Main Results

1. Optimal convergence rate determined by **Fisher information**

2. Optimal convergence rate achieved by **pure bundling**

Main Results

1. Optimal convergence rate determined by **Fisher information**
2. Optimal convergence rate achieved by **pure bundling**
   - but **not** by separate sales

Related literature (incomplete)

Multi-dimensional screening:

- Optimal mechanisms:
  - Wilson (1993), Armstrong (1996), Rochet and Chone (1998), Manelli and Vincent (2006), Daskalakis, Deckelbaum, Tzamos (2017)
  - Bundling mechanisms: Haghpanah and Hartline (2021), Ghili (2023), Yang (2023), etc.
- Robust optimality of simple mechanisms:
  - Carroll (2017), Deb and Roesler (2023), Che and Zhong (2021)
- Revenue guarantee/approximate optimality:
  - Hart and Nisan (2012), Cai et al. (2016), etc.
  - many products: Armstrong (1999), etc.

Model

Model

Seller:

- endowed with a finite set, G, of indivisible goods.

Single buyer of unknown type $\theta \in \Theta \subseteq \mathbb{R}^G_{++}$:

- Θ: compact set with non-empty interior,
- θ drawn according to prob. with density g,
- consuming goods $B \subseteq G$ with transfer t yields:

$$\mathbf{1}^B \cdot \theta - t.$$

- $\mathbf{1}^B_\ell := 1_{\ell \in B}$.

Seller's Information

Before sale, seller observes a sequence of n signals, $x^n = (x_1, \ldots, x_n)$:

- $x_1, \ldots, x_n$: drawn iid $P_\theta \in \Delta(X)$ with density $f(\cdot, \theta)$.
- n parametrizes richness of seller's data.

Seller's Information

Before sale, seller observes a sequence of n signals, $x^n = (x_1, \ldots, x_n)$:

- $x_1, \ldots, x_n$: drawn iid $P_\theta \in \Delta(X)$ with density $f(\cdot, \theta)$.
- n parametrizes richness of seller's data.

Seller's Information

Before sale, seller observes a sequence of n signals, $x^n = (x_1, \ldots, x_n)$:

- $x_1, \ldots, x_n$: drawn iid $P_\theta \in \Delta(X)$ with density $f(\cdot, \theta)$.
- n parametrizes richness of seller's data.

Technical regularity assumptions on g and f throughout:

- more details

Seller's Information

Before sale, seller observes a sequence of n signals, $x^n = (x_1, \ldots, x_n)$:

- $x_1, \ldots, x_n$: drawn iid $P_\theta \in \Delta(X)$ with density $f(\cdot, \theta)$.
- n parametrizes richness of seller's data.

Technical regularity assumptions on g and f throughout:

- more details

After observing x^n, seller commits to a direct mechanism, (q, t):

- $q : \Theta \to \Delta(2^G)$,
- $t : \Theta \to \mathbb{R}$.

Seller's Problem

$$R^{SB}(x^n) := \sup_{(q,t)} \mathbb{E}\left[t(\theta) \mid x^n\right]$$

such that for all θ,

$$\sum_{B \subseteq G} q(B \mid \theta)\left(\mathbf{1}^B \cdot \theta\right) - t(\theta) = \max_{\theta' \in \Theta} \sum_{B \subseteq G} q(B \mid \theta')(\mathbf{1}^B \cdot \theta) - t(\theta'), \quad \text{(IC)}$$

$$\sum_{B \subseteq G} q(B \mid \theta)(\mathbf{1}^B \cdot \theta) - t(\theta) \geq 0. \quad \text{(IR)}$$

Simple Mechanisms

(Pure) bundling: post a single price, p_G, for **grand bundle**, G

- $R^{\mathrm{bd}}(x^n)$: optimal revenue under bundling mechanisms

Simple Mechanisms

(Pure) bundling: post a single price, p_G, for **grand bundle**, G

- $R^{\mathrm{bd}}(x^n)$: optimal revenue under bundling mechanisms

Separate sales: post a price, p_g, for each good $g \in G$

- $R^{\mathrm{sep}}(x^n)$: optimal revenue under separate sales mechanisms

Main Questions

Clear that

$$\mathbb{E}\left[R^{\mathrm{SB}}(x^n)\right], \mathbb{E}\left[R^{\mathrm{bd}}(x^n)\right], \mathbb{E}\left[R^{\mathrm{sep}}(x^n)\right] \to \underbrace{\mathbb{E}\left[\mathbf{1}^G \cdot \theta\right]}_{\text{first best}}.$$

Main Questions

Clear that

$$\mathbb{E}\left[R^{\mathrm{SB}}(x^n)\right], \mathbb{E}\left[R^{\mathrm{bd}}(x^n)\right], \mathbb{E}\left[R^{\mathrm{sep}}(x^n)\right] \to \underbrace{\mathbb{E}\left[\mathbf{1}^G \cdot \theta\right]}_{\text{first best}}.$$

1. What are the rates of convergence of the above?

Main Questions

Clear that

$$\mathbb{E}\left[R^{\mathrm{SB}}(x^n)\right], \mathbb{E}\left[R^{\mathrm{bd}}(x^n)\right], \mathbb{E}\left[R^{\mathrm{sep}}(x^n)\right] \to \underbrace{\mathbb{E}\left[\mathbf{1}^G \cdot \theta\right]}_{\text{first best}}.$$

1. What are the rates of convergence of the above?
2. How do they compare?

Main Results

Fisher Information

Informativeness of signals determines how much surplus seller can capture.

- What is the right measure of informativeness?

Fisher Information

Informativeness of signals determines how much surplus seller can capture.

- What is the right measure of informativeness?

Fisher information at θ:

$$I(\theta) := \left(-\mathbb{E}\left[\frac{\partial^2}{\partial\theta_g \partial\theta_{g'}} \ln f(x_1, \theta) \mid \theta \right] \right)_{g,g' \in G}$$

- differential analogue of Kullback-Leibler divergence:

$$\mathrm{KL}(P_\theta, P_{\theta'}) = (\theta - \theta') \cdot I(\theta)(\theta - \theta') + o\left(\|\theta - \theta'\|^2 \right)$$

Fisher Information

Informativeness of signals determines how much surplus seller can capture.

- What is the right measure of informativeness?

Fisher information at θ:

$$I(\theta) := \left(-\mathbb{E}\left[\frac{\partial^2}{\partial\theta_g \partial\theta_{g'}} \ln f(x_1, \theta) \mid \theta\right]\right)_{g,g' \in G}$$

- differential analogue of Kullback-Leibler divergence:

$$\mathrm{KL}(P_\theta, P_{\theta'}) = (\theta - \theta') \cdot I(\theta)(\theta - \theta') + o\left(\|\theta - \theta'\|^2\right)$$

Assume henceforth that $I(\theta)$ is positive definite for all θ

- ensures complete learning about θ for large n.

Fisher Information and Bernstein-von Mises

Bernstein-von Mises Theorem:

- under regularity, conditional on θ, beliefs of seller $\approx N(\theta, \frac{1}{n} I(\theta)^{-1})$.

Fisher Information and Bernstein-von Mises

Bernstein-von Mises Theorem:

- under regularity, conditional on θ, beliefs of seller $\approx N(\theta, \frac{1}{n} I(\theta)^{-1})$.

For each bundle $B \subseteq G$, define:

$$\lambda^B(\theta) := \sqrt{\mathbf{1}^B \cdot I(\theta)^{-1} \mathbf{1}^B}.$$

Fisher Information and Bernstein-von Mises

Bernstein-von Mises Theorem:

- under regularity, conditional on θ, beliefs of seller $\approx N(\theta, \frac{1}{n} I(\theta)^{-1})$.

For each bundle $B \subseteq G$, define:

$$\lambda^B(\theta) := \sqrt{\mathbf{1}^B \cdot I(\theta)^{-1} \mathbf{1}^B}.$$

Thus,

$$\mathbf{1}^B \cdot \theta \approx N\left(\mathbf{1}^B \cdot \theta, \frac{\lambda^B(\theta)}{\sqrt{n}}\right).$$

Main Theorem

Theorem

Under both the optimal and bundling mechanisms, first-best gap vanishes at the same rate:

$$\mathbb{E}\left[\mathbf{1}^G \cdot \theta\right] - \mathbb{E}\left[R_n^{\mathrm{SB}}(x^n)\right] = \mathbb{E}\left[\lambda^G(\theta)\right]\sqrt{\frac{\ln n}{n}} + o\left(\sqrt{\frac{\ln n}{n}}\right);$$

$$\mathbb{E}\left[\mathbf{1}^G \cdot \theta\right] - \mathbb{E}\left[R_n^{\mathrm{bd}}(x^n)\right] = \mathbb{E}\left[\lambda^G(\theta)\right]\sqrt{\frac{\ln n}{n}} + o\left(\sqrt{\frac{\ln n}{n}}\right).$$

Main Theorem (Cont.): Separate Sales

Theorem

Under separate sales, first-best gap vanishes at a slower rate:

$$\mathbb{E}\left[\mathbf{1}^G \cdot \theta\right] - \mathbb{E}\left[R_n^{\text{sep}}(x^n)\right] = \mathbb{E}\left[\sum_{g \in G} \lambda^g(\theta)\right] \sqrt{\frac{\ln n}{n}} + o\left(\sqrt{\frac{\ln n}{n}}\right).$$

By triangle inequality,

$$\sum_{g \in G} \lambda_d(\theta) = \sum_{g \in G} \sqrt{\mathbf{1}^g \cdot I(\theta)^{-1} \cdot \mathbf{1}^g} \geq \sqrt{\mathbf{1}^G \cdot I(\theta)^{-1} \mathbf{1}^G} = \lambda^G(\theta).$$

Discussion

Optimal and pure bundling converge to first best at same rate.

- Any additional benefit from using more general mechanisms has at most a second-order effect on seller's revenue beyond pure bundling.

Discussion

Optimal and pure bundling converge to first best at same rate.

- Any additional benefit from using more general mechanisms has at most a second-order effect on seller's revenue beyond pure bundling.

To interpret, under pure bundling,

- for any $\varepsilon > 0$, εn extra signals outperforms SB for n suff. large.

Discussion

Optimal and pure bundling converge to first best at same rate.

- Any additional benefit from using more general mechanisms has at most a second-order effect on seller's revenue beyond pure bundling.

To interpret, under pure bundling,

- for any $\varepsilon > 0$, εn extra signals outperforms SB for n suff. large.

In contrast, under separate sales,

- number of extra signals needed to outperform SB for n suff. large is at least

$$\underbrace{\left(\left(\frac{\mathbb{E}\left[\sum_{g \in G} \lambda^g(\theta)\right]}{\mathbb{E}\left[\lambda^G(\theta)\right]} \right)^2 - 1 \right)}_{>0} n.$$

Fisher Information: Gaussian Example

Suppose $|D| = 2$ and

$$x_i \sim N\left(\begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix}, \begin{pmatrix} \sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 \end{pmatrix}\right)$$

Inverse Fisher information:

$$I(\theta)^{-1} = \begin{pmatrix} \sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 \end{pmatrix}$$

Fisher Information: Gaussian Example

Suppose $|D| = 2$ and

$$x_i \sim N\left(\left(\begin{array}{c}\theta_1\\ \theta_2\end{array}\right), \left(\begin{array}{cc}\sigma^2 & \rho\sigma^2\\ \rho\sigma^2 & \sigma^2\end{array}\right)\right)$$

Inverse Fisher information:

$$I(\theta)^{-1} = \left(\begin{array}{cc}\sigma^2 & \rho\sigma^2\\ \rho\sigma^2 & \sigma^2\end{array}\right)$$

Bundling and optimal mechanism:

$$\mathbb{E}\left[\mathbf{1}^G \cdot \theta\right] - \mathbb{E}\left[R^{\text{bd}}(x^n)\right] \approx \underbrace{\sigma\sqrt{2(1+\rho)}}_{\lambda^G(\theta)}\sqrt{\frac{\ln n}{n}}.$$

Separate sales:

$$\mathbb{E}\left[\mathbf{1}^G \cdot \theta\right] - \mathbb{E}\left[R^{\text{sep}}(x^n)\right] \approx \underbrace{2\sigma}_{\lambda^1(\theta)+\lambda^2(\theta)}\sqrt{\frac{\ln n}{n}}.$$

Proof

Proof Outline

1. Reduction to normally distributed types;
2. Convergence rates of pure bundling and separate sales;
3. General mechanisms cannot improve the convergence rate.

Proof Outline

1. Reduction to normally distributed types;
2. Convergence rates of pure bundling and separate sales;
3. General mechanisms cannot improve the convergence rate;

Key Idea: Bernstein-von Mises Theorem

Define $R^{\mathrm{SB}}(\mu, \Sigma), R^{\mathrm{bd}}(\mu, \Sigma), R^{\mathrm{sep}}(\mu, \Sigma)$:

- corresponding optimal revenues when $\theta' \sim N(\mu, \Sigma)$

Lemma

For any $\theta \in \Theta$ and for all $i \in \{\mathrm{SB}, \mathrm{bd}, \mathrm{sep}\}$,

$$\mathbb{E}\left[R_n^i(x^n) \mid \theta\right] - R^i\left(\theta, \frac{1}{n} I(\theta)^{-1}\right) = o\left(\sqrt{\frac{\ln n}{n}}\right).$$

Proof Outline

1. Reduction to normal types;
2. Convergence rates of pure bundling and separate sales;
3. General mechanisms cannot improve the convergence rate.

Convergence Rates of Simple Mechanisms

By Lemma, suffices to show for each θ^*:

$$\mathbf{1}^G \cdot \theta^* - R^{\mathrm{bd}}\left(\theta^*, \frac{1}{n} I(\theta^*)^{-1}\right) = \lambda^G(\theta^*)\sqrt{\frac{\ln n}{n}} + o\left(\sqrt{\frac{\ln n}{n}}\right),$$

$$\mathbf{1}^G \cdot \theta^* - R^{\mathrm{sep}}\left(\theta^*, \frac{1}{n} I(\theta^*)^{-1}\right) = \left(\sum_{g \in G} \lambda^g(\theta^*)\right)\sqrt{\frac{\ln n}{n}} + o\left(\sqrt{\frac{\ln n}{n}}\right).$$

Warm-Up Exercise: Single-good Monopoly

Pure bundling & separate selling: single-good mech design

- single posted price optimal.

Warm-Up Exercise: Single-good Monopoly

Pure bundling & separate selling: single-good mech design

- single posted price optimal.

Consider a type distribution, F_n, for a **single** good: $N(\mu, \sigma^2/n)$.

- assume $\mu > 0$.

Warm-Up Exercise: Single-good Monopoly

Pure bundling & separate selling: single-good mech design

- single posted price optimal.

Consider a type distribution, F_n, for a **single** good: $N(\mu, \sigma^2/n)$.

- assume $\mu > 0$.

What is the optimal profit, Π_n^*, under uniform monopoly pricing?

$$\Pi_n^* = \max_p p(1 - F_n(p)), \; p_n^* = \arg\max_p p(1 - F_n(p)).$$

Warm-up Exercise: Single-good Monopoly

Proposition (Single-Good Monopoly)

$$\mu - \Pi_n^* = \sigma \sqrt{\frac{\ln n}{n}} + o\left(\sqrt{\frac{\ln n}{n}}\right).$$

Moreover,

$$p_n^* = \mu - \sigma \sqrt{\frac{\ln n}{n}} + o\left(\sqrt{\frac{\ln n}{n}}\right), F_n(p_n^*) = o\left(\sqrt{\frac{\ln n}{n}}\right)$$

Proof of Proposition

$$\mu - \Pi_n^* = \underbrace{\mu - p_n^*}_{\text{intensive margin}} + \underbrace{\mu F_n(p_n^*)}_{\text{extensive margin}} - \underbrace{(\mu - p_n^*) F_n(p_n^*)}_{\text{smaller order terms}}.$$

Key idea: suppose that $p_n = \mu - \alpha\sqrt{\frac{\ln n}{n}}$ for some $\alpha > 0$.

$$\mu - \Pi_n^* \approx \underbrace{\alpha\sqrt{\ln n} \cdot n^{-\frac{1}{2}}}_{\text{intensive margin}} + \underbrace{\frac{\mu}{\sqrt{2\pi}} \frac{\sigma}{\alpha\sqrt{\ln n}} \cdot n^{-\frac{\alpha^2}{2\sigma^2}}}_{\text{extensive margin}}$$

Proof of Proposition

$$\mu - \Pi_n^* = \underbrace{\mu - p_n^*}_{\text{intensive margin}} + \underbrace{\mu F_n(p_n^*)}_{\text{extensive margin}} - \underbrace{(\mu - p_n^*)F_n(p_n^*)}_{\text{smaller order terms}}.$$

Key idea: suppose that $p_n = \mu - \alpha\sqrt{\frac{\ln n}{n}}$ for some $\alpha > 0$.

$$\mu - \Pi_n^* \approx \underbrace{\alpha\sqrt{\ln n} \cdot n^{-\frac{1}{2}}}_{\text{intensive margin}} + \underbrace{\frac{\mu}{\sqrt{2\pi}} \frac{\sigma}{\alpha\sqrt{\ln n}} \cdot n^{-\frac{\alpha^2}{2\sigma^2}}}_{\text{extensive margin}}$$

At $\alpha = \sigma$, intensive margin dominates and

$$\mu - \Pi_n^* \approx \underbrace{\sigma\sqrt{\ln n} \cdot n^{-\frac{1}{2}}}_{\text{intensive margin}}.$$

Proof of Proposition

$$\mu - \Pi_n^* = \underbrace{\mu - p_n^*}_{\text{intensive margin}} + \underbrace{\mu F_n(p_n^*)}_{\text{extensive margin}} - \underbrace{(\mu - p_n^*)F_n(p_n^*)}_{\text{smaller order terms}}.$$

Key idea: suppose that $p_n = \mu - \alpha\sqrt{\frac{\ln n}{n}}$ for some $\alpha > 0$.

$$\mu - \Pi_n^* \approx \underbrace{\alpha\sqrt{\ln n} \cdot n^{-\frac{1}{2}}}_{\text{intensive margin}} + \underbrace{\frac{\mu}{\sqrt{2\pi}} \frac{\sigma}{\alpha\sqrt{\ln n}} \cdot n^{-\frac{\alpha^2}{2\sigma^2}}}_{\text{extensive margin}}$$

At $\alpha = \sigma$, intensive margin dominates and

$$\mu - \Pi_n^* \approx \underbrace{\sigma\sqrt{\ln n} \cdot n^{-\frac{1}{2}}}_{\text{intensive margin}}.$$

Moreover, $\alpha = \sigma$ is optimal because,

1. If $\alpha > \sigma$, intensive margin increases.
2. If $\alpha < \sigma$, extensive margin dominates.

Implications of Proposition

Single-good monopoly proposition implies:

$$\mathbf{1}^G \cdot \theta^* - R^{\mathrm{bd}}\left(\theta^*, \frac{1}{n} I(\theta^*)^{-1}\right) = \underbrace{\sqrt{\mathbf{1}^G \cdot I(\theta^*)^{-1}\mathbf{1}^G}}_{\lambda^G(\theta)} \sqrt{\frac{\ln n}{n}} + o\left(\sqrt{\frac{\ln n}{n}}\right)$$

Moreover,

$$\underbrace{\mathbf{1}^G \cdot \theta^* - p_n^*}_{\text{intensive margin}} \approx \lambda^G(\theta^*)\sqrt{\frac{\ln n}{n}}, \quad \underbrace{F_n(p_n^*)}_{\text{extensive margin}} = o\left(\sqrt{\frac{\ln n}{n}}\right)$$

- Gap to first-best under optimal bundling is of the same order as the intensive margin.

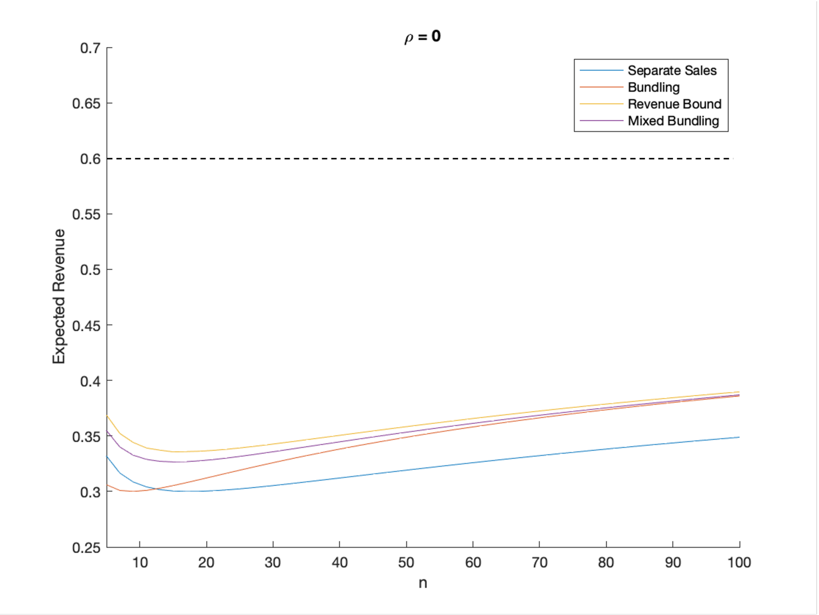
Implications of Proposition (Cont.)

Single-good monopoly proposition also implies:

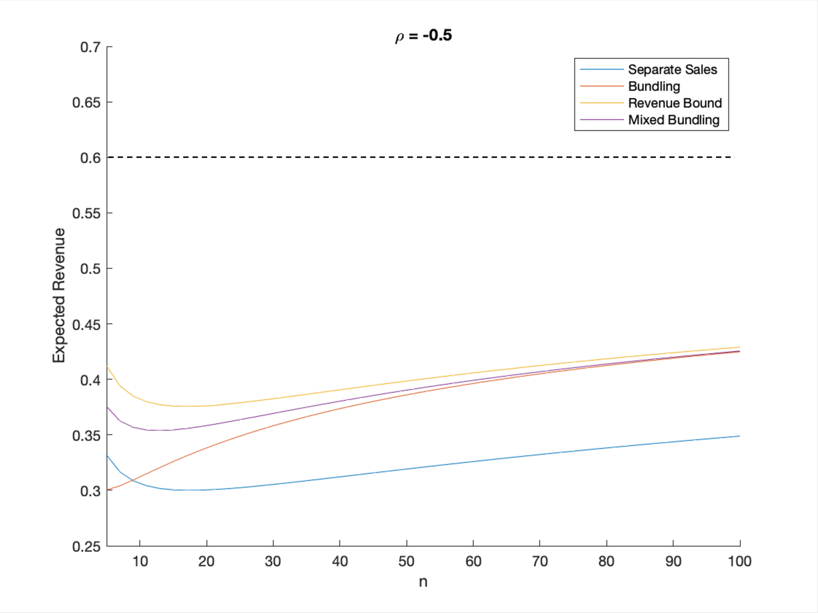
$$\mathbf{1}^G \cdot \theta^* - R^{\mathrm{sep}}\left(\theta^*, \frac{1}{n} I(\theta^*)^{-1}\right) = \sum_{g \in G} \underbrace{\sqrt{\mathbf{1}^g \cdot I(\theta^*)^{-1} \mathbf{1}^g}}_{\lambda^g(\theta^*)} \sqrt{\frac{\ln n}{n}} + o\left(\sqrt{\frac{\ln n}{n}}\right)$$

- $\Rightarrow$ suboptimality of separate sales.

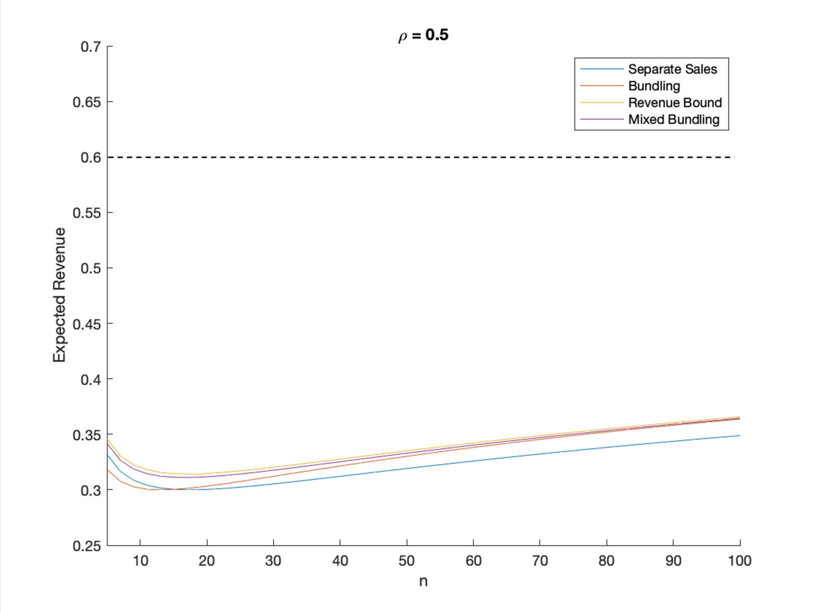
Numerical Examples



Numerical Examples



Numerical Examples



Proof Outline

1. Reduction to normal types;
2. Convergence rates of pure bundling and separate sales;
3. General mechanisms cannot improve the convergence rate.

Pure Bundling vs. "Simple" Mixed Bundling Mechanisms

First consider deterministic mechanisms:

- Offer a collection of deterministic bundles:

$$S_1 = G, S_2, \ldots, S_m \subsetneq G.$$

Pure Bundling vs. “Simple” Mixed Bundling Mechanisms

Consider first the optimal bundling mechanism with price p_G^*.

Pure Bundling vs. "Simple" Mixed Bundling Mechanisms

Consider first the optimal bundling mechanism with price p_G^*.

Recall:

1. First-best gap from optimal bundling mechanism:

$$\mathbf{1}^G \cdot \theta^* - R^{\mathrm{bd}}\left(\theta^*, \frac{1}{n} I(\theta^*)^{-1}\right) \approx \underbrace{\lambda^G(\theta^*)\sqrt{\frac{\ln n}{n}}}_{\text{intensive margin}}.$$

Pure Bundling vs. "Simple" Mixed Bundling Mechanisms

Consider first the optimal bundling mechanism with price p_G^*.

Recall:

1. First-best gap from optimal bundling mechanism:

$$\mathbf{1}^G \cdot \theta^* - R^{\mathrm{bd}}\left(\theta^*, \frac{1}{n} I(\theta^*)^{-1}\right) \approx \underbrace{\lambda^G(\theta^*)\sqrt{\frac{\ln n}{n}}}_{\text{intensive margin}}.$$

2. Extensive margin:

$$F_n(p_G^*) = o\left(\sqrt{\frac{\ln n}{n}}\right).$$

Pure Bundling vs. "Simple" Mixed Bundling Mechanisms

Keeping fixed p_G^*, benefit of additionally offering $S_2, \ldots, S_m$ only materializes if G is rejected.

- but this benefit is small: $F_n(p_G^*) = o\left(\sqrt{\frac{\ln n}{n}}\right)$.

Pure Bundling vs. "Simple" Mixed Bundling Mechanisms

Keeping fixed p_G^*, benefit of additionally offering $S_2, \ldots, S_m$ only materializes if G is rejected.

- but this benefit is small: $F_n(p_G^*) = o\left(\sqrt{\frac{\ln n}{n}}\right)$.

Only way to improve profits substantially is to raise $p_G \gg p_G^*$.

- But this makes the extensive margin too large: $F_n(p_G) \gg \sqrt{\frac{\ln n}{n}}$.
- Any surplus from sales of $S_2, \ldots, S_m$ bounded away from $\mathbf{1}^G \cdot \theta^*$.

General Mechanisms

Using single-good monopoly proposition, straightforward to compare convergence rates for "simple" mechanisms.

- pure bundling vs. separate sales.
- pure bundling vs. mixed bundling.

General Mechanisms

Using single-good monopoly proposition, straightforward to compare convergence rates for "simple" mechanisms.

- pure bundling vs. separate sales.
- pure bundling vs. mixed bundling.

However, general mechanisms can be substantially more complex:

- Offer random bundles;
- Offer continuum of random bundles.

General Mechanisms

Using single-good monopoly proposition, straightforward to compare convergence rates for "simple" mechanisms.

- pure bundling vs. separate sales.
- pure bundling vs. mixed bundling.

However, general mechanisms can be substantially more complex:

- Offer random bundles;
- Offer continuum of random bundles.

Question: How to bound the convergence rate for all mechanisms?

General Mechanisms

Using single-good monopoly proposition, straightforward to compare convergence rates for “simple” mechanisms.

- pure bundling vs. separate sales.
- pure bundling vs. mixed bundling.

However, general mechanisms can be substantially more complex:

- Offer random bundles;
- Offer continuum of random bundles.

Question: How to bound the convergence rate for all mechanisms?

- **Challenge:** solving for optimal mechanism for n is intractable.

Key Idea: Relaxed Problem

Fix some $(|D|-1) \times |D|$ matrix, A, with full row rank.

Relaxed Problem:

$$\bar{R}^{\mathrm{SB}}\left(\theta^*, \frac{1}{n} I(\theta^*)^{-1}\right) := \sup_{(q,t)} \mathbb{E}\left[t(\theta)\right]$$

such that for all θ,

$$\sum_{B \subseteq G} q(B \mid \theta)\left(\mathbf{1}^B \cdot \theta\right) - t(\theta) = \max_{\theta' : A\theta' = A\theta} \sum_{B \subseteq G} q(B \mid \theta')(\mathbf{1}^B \cdot \theta) - t(\theta'), \quad \text{(IC)}$$

$$\sum_{B \subseteq G} q(B \mid \theta)(\mathbf{1}^B \cdot \theta) - t(\theta) \geq 0. \quad \text{(IR)}$$

Rewriting Relaxed Problem

Relaxed Problem:

$$\bar{R}^{\mathrm{SB}}\left(\theta^*, \frac{1}{n} I(\theta^*)^{-1} \mid y\right) = \sup_{(q,t)} \mathbb{E}\left[t(\theta) \mid A\theta = y\right]$$

such that for all $\theta \in A^{-1}(y)$,

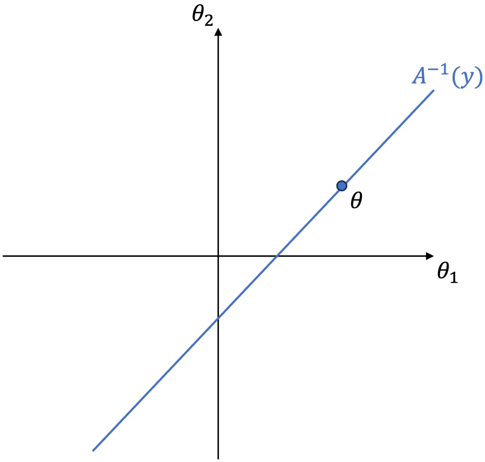
$$\sum_{B \subseteq G} q(B \mid \theta)\left(\mathbf{1}^B \cdot \theta\right) - t(\theta) = \max_{\theta' \in A^{-1}(y)} \sum_{B \subseteq G} q(B \mid \theta')(\mathbf{1}^B \cdot \theta) - t(\theta'), \quad \text{(IC)}$$

$$\sum_{B \subseteq G} q(B \mid \theta)(\mathbf{1}^B \cdot \theta) - t(\theta) \geq 0. \quad \text{(IR)}$$

Note:

$$\bar{R}^{\mathrm{SB}}\left(\theta^*, \frac{1}{n} I(\theta^*)^{-1}\right) = \mathbb{E}\left[R^{\mathrm{SB}}\left(\theta^*, \frac{1}{n} I(\theta^*)^{-1} \mid y\right)\right].$$

Relaxed Problem



Proof Overview

1. In the relaxed problem, $\bar{R}^{\mathrm{SB}}\left(\theta^*, \frac{1}{n} I(\theta^*)^{-1} \mid y\right)$, type space is one-dimensional.
   - We prove that there exist optimal mechanisms that offer only deterministic bundles or a single random bundle.
     - ⋆ Rules out "complex" mechanisms that offer a continuum of random bundles.

Proof Overview

1. In the relaxed problem, $\bar{R}^{\mathrm{SB}}\left(\theta^*, \frac{1}{n} I(\theta^*)^{-1} \mid y\right)$, type space is one-dimensional.
   - We prove that there exist optimal mechanisms that offer only deterministic bundles or a single random bundle.
     - ⋆ Rules out "complex" mechanisms that offer a continuum of random bundles.
2. Similar arguments to comparison of pure bundling to deterministic mixed bundling $\Rightarrow$ approx. optimality of pure bundling after all y.

Proof Overview

1. In the relaxed problem, $\bar{R}^{\mathrm{SB}}\left(\theta^*, \frac{1}{n} I(\theta^*)^{-1} \mid y\right)$, type space is one-dimensional.
   - We prove that there exist optimal mechanisms that offer only deterministic bundles or a single random bundle.
     - ★ Rules out "complex" mechanisms that offer a continuum of random bundles.
2. Similar arguments to comparison of pure bundling to deterministic mixed bundling $\Rightarrow$ approx. optimality of pure bundling after all y.
3. $\Rightarrow$ pure bundling after all y is approximately optimal in relaxed problem.

How to Choose A?

Can choose A such that $y = A\theta$ is orthogonal to $\mathbf{1}^G \cdot \theta$:

$$(\mathbf{1}^G \cdot \theta \mid A\theta = y) \sim (\mathbf{1}^G \cdot \theta) \; \forall y.$$

How to Choose A?

Can choose A such that $y = A\theta$ is orthogonal to $\mathbf{1}^G \cdot \theta$:

$$(\mathbf{1}^G \cdot \theta \mid A\theta = y) \sim (\mathbf{1}^G \cdot \theta) \; \forall y.$$

Ensures that optimal bundling mechanism in relaxed problem is the same as the optimal bundling mechanism in the original problem.

- Set price at original grand bundle price after all y.

How to Choose A?

Can choose A such that $y = A\theta$ is orthogonal to $\mathbf{1}^G \cdot \theta$:

$$(\mathbf{1}^G \cdot \theta \mid A\theta = y) \sim (\mathbf{1}^G \cdot \theta) \; \forall y.$$

Ensures that optimal bundling mechanism in relaxed problem is the same as the optimal bundling mechanism in the original problem.

- Set price at original grand bundle price after all y.

$\therefore$ optimal mechanism does not improve (in terms of convergence rate) on pure bundling. □

Concluding Remarks

Optimal convergence rate to first best achieved by pure bundling:

- While analysis is conducted asymptotically, numerical examples suggest pure bundling $\approx \mathrm{SB}$ well before $\mathrm{SB} \approx \mathrm{FB}$.

Concluding Remarks

Optimal convergence rate to first best achieved by pure bundling:

- While analysis is conducted asymptotically, numerical examples suggest pure bundling $\approx \mathrm{SB}$ well before $\mathrm{SB} \approx \mathrm{FB}$.

Ongoing work:

- Incorporate costs, general non-additive utilities.

Concluding Remarks

Optimal convergence rate to first best achieved by pure bundling:

- While analysis is conducted asymptotically, numerical examples suggest pure bundling $\approx \mathrm{SB}$ well before $\mathrm{SB} \approx \mathrm{FB}$.

Ongoing work:

- Incorporate costs, general non-additive utilities.

Approach of analyzing convergence rates seems fruitful in other applications:

- Large markets: Rustichini, Satterthwaite, and Williams (1994), Satterthwaite and Williams (2002), Hong and Shum (2004)
- Moral hazard contracts: Frick, Iijima, Ishii (2023)

Thank you!

Regularity Assumptions

1. g is strictly positive and locally Lipschitz continuous.
2. For each $x \in X$, $f(x, \cdot) > 0$ and twice-differentiable in θ.
3. There exists L such that

$$\sup_{\theta, g, g', x} \left| \frac{\partial^2 \log f(x, \theta)}{\partial \theta_g \partial \theta_{g'}} \right| \leq L,$$

$$\sup_{\theta, \theta', (d, d'), x} \left| \frac{\partial^2 \log f(x, \theta)}{\partial \theta_g \partial \theta_{g'}} - \frac{\partial^2 \log f(x, \theta')}{\partial \theta_g \partial \theta_{g'}} \right| \leq L \|\theta - \theta'\|.$$

4. P_θ is continuous in θ with respect to Total-variation metric.
5. $\sup_\theta \mathbb{E} \left[\left(\sup_{\theta' \in \Theta} \log f(x, \theta') \right)^2 \mid \theta \right] < \infty$.
6. **Fisher information** matrix $I(\theta)$ as defined by

$$I(\theta) := - \left(\mathbb{E} \left[\frac{\partial^2}{\partial \theta_g \theta_{g'}} \log f(x, \theta) \mid \theta \right] \right)_{g, g' \in G}$$

is positive definite for each θ.

back