Robust Multidimensional Screening

Rahul Deb

BOSTON COLLEGE

Northwestern Summer School
August 26, 2026

Multidimensional screening: What is it?

A monopoly seller of multiple goods faces a single buyer with a private type.

The buyer's type determines her value for different bundles of goods.

- The type could be the value for each good (additive case).
- More generally, goods could be complements or substitutes.

How should the seller optimally sell these goods?

- Should she charge a separate price for each good and allow the buyer to pick his own bundle?

Multidimensional screening: What is it?

A monopoly seller of multiple goods faces a single buyer with a private type.

The buyer's type determines her value for different bundles of goods.

- The type could be the value for each good (additive case).
- More generally, goods could be complements or substitutes.

How should the seller optimally sell these goods?

- Should she charge a separate price for each good and allow the buyer to pick his own bundle?
- Should she sell bundles of goods?

Multidimensional screening: What is it?

A monopoly seller of multiple goods faces a single buyer with a private type.

The buyer's type determines her value for different bundles of goods.

- The type could be the value for each good (additive case).
- More generally, goods could be complements or substitutes.

How should the seller optimally sell these goods?

- Should she charge a separate price for each good and allow the buyer to pick his own bundle?
- Should she sell bundles of goods?
- Or both?

Multidimensional screening: What is it?

A monopoly seller of multiple goods faces a single buyer with a private type.

The buyer's type determines her value for different bundles of goods.

- The type could be the value for each good (additive case).
- More generally, goods could be complements or substitutes.

How should the seller optimally sell these goods?

- Should she charge a separate price for each good and allow the buyer to pick his own bundle?
- Should she sell bundles of goods?
- Or both?
- What about lotteries?

Multidimensional screening: Why is it important?

Many firms offer multiple goods.

Traditional examples are large retailers.

- Amazon, Costco, Walmart etc.
- Typically, the seller just chooses prices for each good but there is some bundling.

Multidimensional screening: Why is it important?

Many firms offer multiple goods.

Traditional examples are large retailers.

- Amazon, Costco, Walmart etc.
- Typically, the seller just chooses prices for each good but there is some bundling.

Digital streaming services are also good examples.

- Netflix, Spotify etc.
- These firms choose a pure bundling approach.

Multidimensional screening: Why is it important?

Many firms offer multiple goods.

Traditional examples are large retailers.

- Amazon, Costco, Walmart etc.
- Typically, the seller just chooses prices for each good but there is some bundling.

Digital streaming services are also good examples.

- Netflix, Spotify etc.
- These firms choose a pure bundling approach.

Why do we see this difference?

How would we advise a retailer or a streaming service?

Without understanding simple benchmark cases like monopoly, it is hard to study competition.

Point of departure: The simplest case

Think of the simplest case:

- Two goods.
- Values are additive.
- The buyer's value for each good is iid $\mathbb{U}[0,1]$.

Questions

1. What is the optimal mechanism?

Point of departure: The simplest case

Think of the simplest case:

- Two goods.
- Values are additive.
- The buyer's value for each good is iid $\mathbb{U}[0, 1]$.

Questions

1. What is the optimal mechanism?
2. Does the form of the optimal mechanism change if I shift the support of the distribution to $[\underline{\theta}, \underline{\theta} + 1]$ where $\underline{\theta} > 0$?

Lecture aim and plan

We have already seen robustness being motivated as a desirable modeling property.

Today I will emphasize additional benefits of robust optimality.

- It can sometimes help make hard problems tractable.
- It can yield more natural solutions.

Lecture aim and plan

We have already seen robustness being motivated as a desirable modeling property.

Today I will emphasize additional benefits of robust optimality.

- It can sometimes help make hard problems tractable.
- It can yield more natural solutions.

I will demonstrate these benefits in the context of multidimensional screening.

- I will start by explaining why the problem is hard.
- I will then discuss different ways to model robustness.
- This highlights how different modeling choices can lead to qualitatively different solutions.
- The central point I will emphasize is that general instances of a problem may be hard to solve, but worst case instances may be easier.

Model: Types

Seller has one unit of each of $N = \{1, \ldots, n\}$ goods with $n \geq 2$.

The seller has a *constant marginal cost* $c \geq 0$ of producing the goods.

One buyer with *types* $\theta = (\theta_1, \ldots, \theta_n) \in \Theta = [\underline{\theta}, \overline{\theta}]^n$ with $\overline{\theta} > \underline{\theta} \geq 0$.

θ is jointly distributed with distribution $F \in \Delta(\Theta)$.

The buyer has utility $u(\theta, b)$ for a *bundle* $b \subseteq N$.

- We will mostly focus on the additive case $u(\theta, b) = \sum_{i \in b} \theta_i$.
- We refer to the bundle $b = N$ of all goods as the *grand bundle*.

The central issue is how to model what the seller knows about F.

- And, of course, how this affects the optimal mechanism.

Model: Mechanism

Throughout, we will employ the revelation principle.

- If we were to consider auctions, we might worry about the issues Ben raised.

A (direct) *mechanism* (q, t) consists of

- an allocation $q : \Theta \to \Delta(2^N)$, with $q(\theta, b) \in [0, 1]$ representing the probability of receiving bundle $b \subseteq N$, and
- a transfer $t : \Theta \to \mathbb{R}$.

We use $q_i(\theta) = \sum_{b \supseteq \{i\}} q(\theta, b)$ to denote the allocation probability of good i.

- Note that we don't use this as the primitive because values may not be additive.
- Getting each good independently with probability $\frac{1}{2}$ may yield a different utility than getting the grand bundle with probability $\frac{1}{2}$ (and nothing with probability $\frac{1}{2}$).

Model: Incentive compatibility (IC) and individual rationality (IR)

We require that (q, t) satisfies the standard IC and IR constraints:

$$\sum_{b \subseteq N} u(\theta, b) q(\theta, b) - t(\theta) \geq \sum_{b \subseteq N} u(\theta, b) q(\theta', b) - t(\theta') \quad \forall \theta, \theta' \in \Theta, \tag{IC}$$

$$\sum_{b \subseteq N} u(\theta, b) q(\theta, b) - t(\theta) \geq 0 \quad \forall \theta \in \Theta. \tag{IR}$$

We denote the set of all IC+IR mechanisms by $\mathcal{M}$.

Model: Mechanism

I will repeatedly mention two special classes of mechanisms.

- **Pure bundling**: Only the grand bundle is for sale at a given price.
- **Separate sales**: There is a price for each good, the buyer can choose any bundle.

Formally, a mechanism is *pure bundling* at price $\overline{p}$ if for all $b \neq \emptyset$,

$$q(s,b) = \begin{cases} 1 & \text{if } u(s,N) \geq \overline{p} \text{ and } b = N, \\ 0 & \text{otherwise,} \end{cases} \quad \text{and} \quad t(s) = \begin{cases} \overline{p} & \text{if } u(s,N) \geq \overline{p}, \\ 0 & \text{otherwise.} \end{cases}$$

Formally, a *separate sales mechanism* at prices $p = (p_1, \ldots, p_N)$ is

$$q(s,b) = \begin{cases} 1 & \text{if } b = \hat{b}(s,p), \\ 0 & \text{otherwise,} \end{cases} \quad \text{and} \quad t(s) = \begin{cases} \sum_{i \in b} p_i & \text{if } b = \hat{b}(s,p), \\ 0 & \text{otherwise,} \end{cases}$$

where $\hat{b}(s,p) \in \arg\max_{b' \subseteq N} \{u(s,b') - \sum_{i \in b'} p_i\}$.

Returning to the simplest case

Let $\theta = (\theta_1, \theta_2) \sim \mathbb{U}[0,1]^2$. Assume additive utility and cost $c = 0$.

Separate sales

The monopoly price of each good $i \in \{1,2\}$ is $p_i = 1/2$, so

$$\Pi_{\text{ss}} = 2\left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = \frac{1}{2}.$$

Returning to the simplest case

Let $\theta = (\theta_1, \theta_2) \sim \mathbb{U}[0,1]^2$. Assume additive utility and cost $c = 0$.

Separate sales

The monopoly price of each good $i \in \{1,2\}$ is $p_i = 1/2$, so

$$\Pi_{\text{ss}} = 2\left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = \frac{1}{2}.$$

Pure bundling

Since $\theta_1 + \theta_2$ has a triangular distribution, the optimal price $\overline{p}$ for the grand bundle is

$$\overline{p} = \sqrt{\frac{2}{3}}, \qquad \Pi_{\text{pb}} = \frac{2}{3}\sqrt{\frac{2}{3}} \approx .544.$$

Returning to the simplest case

Let $\theta = (\theta_1, \theta_2) \sim \mathbb{U}[0,1]^2$. Assume additive utility and cost $c = 0$.

Separate sales

The monopoly price of each good $i \in \{1,2\}$ is $p_i = 1/2$, so

$$\Pi_{\text{ss}} = 2\left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = \frac{1}{2}.$$

Pure bundling

Since $\theta_1 + \theta_2$ has a triangular distribution, the optimal price $\overline{p}$ for the grand bundle is

$$\overline{p} = \sqrt{\frac{2}{3}}, \qquad \Pi_{\text{pb}} = \frac{2}{3}\sqrt{\frac{2}{3}} \approx .544.$$

Why is pure bundling more profitable than separate sales?

- Bundling reduces the dispersion in total value: a high value for one good offsets a low value for the other.
  - The classic example is negatively correlated values.
- But the gain must outweigh the loss from charging one price to buyers with high values for both goods.

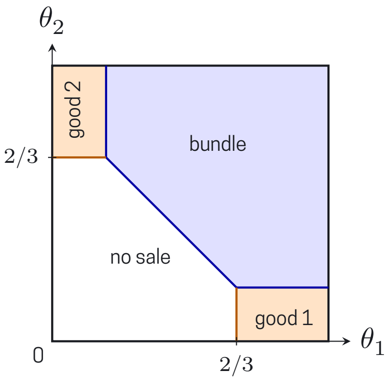
Mixed bundling uses both dimensions

The optimal menu (Pavlov, 2011) is the following:

- Each good is priced separately at $\frac{2}{3} > \frac{1}{2}$.
- The grand bundle is priced at $\frac{4-\sqrt{2}}{3} > \sqrt{\frac{2}{3}}$.
- $\Pi_{\text{mb}} = \frac{12+2\sqrt{2}}{27} \approx .549 > \Pi_{\text{pb}} > \Pi_{\text{ss}}$.

Why it works:

- The discount captures the aggregation gain from balanced types with high total value.
- Stand-alone options retain asymmetric types who value only one good highly.
- Self-selection serves both groups without imposing one common cutoff.



Shifting the support changes the optimal menu (Pavlov, 2011)

Now let $\theta_i \overset{\text{iid}}{\sim} \mathbb{U}[\underline{\theta}, \underline{\theta}+1]$.

Mixed bundling: $\underline{\theta} = 0$

Price $p_i = \frac{2}{3}$ for good i,

Price $\overline{p} = \frac{4-\sqrt{2}}{3}$ for the grand bundle.

Stochastic mixed bundling: $0 < \underline{\theta} < .077$

Price $\overline{p}_\beta$: Good i and good $j \neq i$ with probability $\beta \in (0,1)$.

Price $\overline{p} > \overline{p}_\beta$: The grand bundle.

β rises with $\underline{\theta}$.

Pure bundling: $\underline{\theta} \geq .077$

Price

$$\overline{p} = \frac{4\underline{\theta}}{3} + \frac{\sqrt{4\underline{\theta}^2 + 6}}{3}.$$

for the grand bundle.

Raising the lower bound makes exclusion more costly. For a small shift, partial bundles trade off rent extraction against participation; for a large enough shift, both forces favor giving both goods, so the partial-bundle options collapse into pure bundling.

Why multidimensional screening is hard

In one dimension, IC is equivalent to an increasing allocation $q(\theta)$ with transfers pinned down by the envelope formula

$$t(\theta) = t(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} q(x)dx.$$

One can use this to write the seller's problem pure in terms of the allocation.

- IC is simply the monotonicity of the allocation.

Optimality can be established either pointwise (by distributional arguments) or by ironing.

With multiple goods, types are *not ordered*: IC is characterized by global cyclic-monotonicity (Rochet, 1987).

More generally

To the best of my knowledge, the optimal mechanism with iid $\mathbb{U}[0,1]^n$ is not known for $n > 6$.

More generally

To the best of my knowledge, the optimal mechanism with iid $\mathbb{U}[0,1]^n$ is not known for $n > 6$.

Daskalakis+Deckelbaum+Tzamos (2017) show the optimal mechanism with two goods and iid Beta distributions consists of an infinite menu of lotteries.

More generally

To the best of my knowledge, the optimal mechanism with iid $\mathbb{U}[0,1]^n$ is not known for $n > 6$.

Daskalakis+Deckelbaum+Tzamos (2017) show the optimal mechanism with two goods and iid Beta distributions consists of an infinite menu of lotteries.

Hart+Reny (2015) show that the seller's maximal revenue may decrease if the buyer's value increases (in a first-order stochastic dominance sense).

More generally

To the best of my knowledge, the optimal mechanism with iid $\mathbb{U}[0,1]^n$ is not known for $n > 6$.

Daskalakis+Deckelbaum+Tzamos (2017) show the optimal mechanism with two goods and iid Beta distributions consists of an infinite menu of lotteries.

Hart+Reny (2015) show that the seller's maximal revenue may decrease if the buyer's value increases (in a first-order stochastic dominance sense).

Hart+Nisan (2019) show that simple mechanisms—such as separate or pure bundling—may yield only a negligible fraction of the optimal revenue.

More generally

To the best of my knowledge, the optimal mechanism with iid $\mathbb{U}[0,1]^n$ is not known for $n > 6$.

Daskalakis+Deckelbaum+Tzamos (2017) show the optimal mechanism with two goods and iid Beta distributions consists of an infinite menu of lotteries.

Hart+Reny (2015) show that the seller's maximal revenue may decrease if the buyer's value increases (in a first-order stochastic dominance sense).

Hart+Nisan (2019) show that simple mechanisms—such as separate or pure bundling—may yield only a negligible fraction of the optimal revenue.

Summary:

- Solving for the optimal mechanism is hard.
- The optimal mechanism may be weird. Weird mechanisms are not practical.

Moving forward

In practice, firms do not know the precise joint distribution F of the buyer's values.

- This is one starting point for taking a robust approach to mechanism design.

But even if they did, mechanisms in practice are not super complicated.

- Most sellers either choose separate sales, pure bundling or offer bundle discounts.
- Randomization is not observed in practice.

So, to reiterate, the standard model is both intractable and perhaps not very realistic.

How should we model a robust seller?

There are many possible ways to model robustness in multidimensional screening.

- What are natural choices?

Why can a max-min approach be more tractable?

- While it is hard to derive the optimal mechanism for arbitrary F, it might be possible for certain distributions F.
- If the worst case distributions take this tractable form, we are back in business.

Robustness and separation in multidimensional screening

Known marginals but not the joint distribution (Carroll, 2017)

Suppose the seller knows the marginal distribution F_i of the buyer's value for good i.

But the seller does not know the joint distribution F.

Thus, the set of possible joint distributions is

$$\mathcal{F} = \{F \in \Delta(\Theta) : \operatorname{marg}_i F = F_i \text{ for every } i \in N\}.$$

The seller solves

$$\Pi^\star = \sup_{(q,t)\in\mathcal{M}} \inf_{F\in\mathcal{F}} \mathbb{E}_F\left[t(\theta) - c\sum_{i=1}^{n} q_i(\theta)\right].$$

Profit guarantee

A mechanism $(q, t) \in \mathcal{M}$ *guarantees* a profit of Π if

$$\inf_{F \in \mathcal{F}} \mathbb{E}_F \left[t(\theta) - c \sum_{i=1}^{n} q_i(\theta) \right] \geq \Pi.$$

$\Pi^\star$ is the *highest profit guarantee* for the seller.

- Throughout ties will be broken in favor of the seller.
- Unlike Ben's lecture yesterday, this is not an issue in this one agent setting.

Separate sales give a correlation-free guarantee

The seller can always choose a separate sales mechanism $(q^\star, t^\star)$ with the per good monopoly prices and per good profits:

$$p_i^\star \in \operatorname*{arg\,max}_{p\in[\underline{\theta},\overline{\theta}]}(p-c)(1-F_i(p)) \quad \text{and} \quad \Pi_i^\star := (p_i^\star - c)(1-F_i(p_i^\star)).$$

For every $F \in \mathcal{F}$, this separate sales mechanism guarantees the seller a profit of

$$\mathbb{E}_F\left[t^\star(\theta) - c\sum_{i=1}^n q_i^\star(\theta)\right] = \sum_{i=1}^n \Pi_i^\star.$$

Separate sales are robustly optimal

Theorem

No IC+IR mechanism $(q, t) \in \mathcal{M}$ can provide a higher profit guarantee than $\sum_{i=1}^{n} \Pi_i^\star$.

Note the statement of the theorem. It is not claiming uniqueness.

- More on this at the end. Recall Alessandro's discussion yesterday.

Separate sales are robustly optimal

Theorem

No IC+IR mechanism $(q, t) \in \mathcal{M}$ can provide a higher profit guarantee than $\sum_{i=1}^{n} \Pi_i^\star$.

Note the statement of the theorem. It is not claiming uniqueness.

- More on this at the end. Recall Alessandro's discussion yesterday.

Implication:

- The complex multidimensional screening reduces to n trivial one-good problems.
- Provides one explanation for separate sales in practice.

One proof strategy:

- Find one distribution $F \in \mathcal{F}$ for which the optimal mechanism yields profit $\Pi^\star$.
- Then the profit guarantee cannot be higher.

Candidate distribution: Comonotone types

Suppose every F_i has a continuous, positive density f_i.

Let z be a quantile and recall that its distribution is $z \sim \mathbb{U}[0, 1]$.

Candidate distribution: Comonotone types

Suppose every F_i has a continuous, positive density f_i.

Let z be a quantile and recall that its distribution is $z \sim \mathbb{U}[0,1]$.

Set

$$\theta_i(z) = F_i^{-1}(z), \qquad i \in N.$$

The induced joint distribution F^{co} of $\theta(z) = (\theta_1(z), \dots, \theta_n(z))$ is *comonotonic*.

- Each $\theta_i(z)$ has marginal CDF F_i, so $F^{\text{co}} \in \mathcal{F}$.
- The type vector $z \mapsto (\theta_1(z), \dots, \theta_n(z))$ is effectively one-dimensional.
- A high value for one good never offsets a low value for another: comonotonicity removes the diversification benefit of bundling.

The mechanism is still unrestricted; q and t may depend on the entire vector θ.

- But it suffices to define it $(q(z), t(z))$ only on the support of F^{co}.

Monotone virtual values

Define the one-good virtual value

$$\phi_i(\theta_i) = \theta_i - \frac{1 - F_i(\theta_i)}{f_i(\theta_i)}.$$

Suppose the F_is are such that each virtual value is increasing (*monotone hazard rate*).

Note that in this case, we have

$$\phi_i(p_i^\star) = p_i^\star - \frac{1 - F_i(p_i^\star)}{f_i(p_i^\star)} = c.$$

We argue that no mechanism $(q, t) \in \mathcal{M}$ yields a higher profit than $\Pi^\star$ for distribution F^{co}.

This, in turn, yields the robust optimality of separate sales.

Profits in terms of virtual values

The standard envelope formula and integration by parts imply that

$$\mathbb{E}_{F^{\text{co}}}\left[t(\theta(z)) - c\sum_{i=1}^{n} q_i(\theta(z))\right] \leq \sum_{i\in N}\int_0^1 \left[\theta_i(z) - c - \frac{1-z}{f_i(\theta_i(z))}\right] q_i(z)\,dz$$
$$= \sum_{i\in N}\int_0^1 (\phi_i(\theta_i(z)) - c) q_i(z)\,dz.$$

where the inequality follows from the IR constraint.

Critically, this standard trick does *not* work in general for multiple goods.

- This is because IC is not characterized in this simple way.
- So this is where the comonotonicity is really helpful.

Profits in terms of virtual values

The pointwise maximizer satisfies IC+IR and is given by

$$q_i^\star(z) = \begin{cases} 0, & \phi_i(\theta_i(z)) < c \iff \theta_i(z) < p_i^\star, \\ 1, & \phi_i(\theta_i(z)) \geq c \iff \theta_i(z) \geq p_i^\star. \end{cases}$$

This yields exactly the profit $\Pi^\star = \sum_{i=1}^n \Pi_i^\star$ and we are done!

Effectively, the argument shows that (when virtual values are monotone) one worst case distribution is comonotonic.

- When the virtual values are not monotone, this simple approach does not work.
- Separate sales may not be the profit maximizing mechanism for a comonotonic distribution whose marginals do not satisfy the monotone hazard rate assumption.
- A lot more work is required but the message stays the same.

Separate sales versus pure bundling

Separate sales versus pure bundling

Now suppose that the marginals $F_1 = \cdots = F_n$ are the same.

- The monopoly price $p_1^\star = \cdots = p_n^\star = p^\star$ for each good is the same.

For the comonotonic joint distribution, the optimal pure bundling yields the same profit as the optimal separate sales.

- Choose the grand bundle price $\overline{p} = \sum_{i=1}^n p^\star$ to be the sum of the monopoly prices.
- In this case, $\sum_{i=1}^n \theta_i(z) \geq \overline{p} \iff \theta_i(z) \geq p^\star$ for any $i \in N$.

Separate sales versus pure bundling

Now suppose that the marginals $F_1 = \cdots = F_n$ are the same.

- The monopoly price $p_1^\star = \cdots = p_n^\star = p^\star$ for each good is the same.

For the comonotonic joint distribution, the optimal pure bundling yields the same profit as the optimal separate sales.

- Choose the grand bundle price $\overline{p} = \sum_{i=1}^n p^\star$ to be the sum of the monopoly prices.
- In this case, $\sum_{i=1}^n \theta_i(z) \geq \overline{p} \iff \theta_i(z) \geq p^\star$ for any $i \in N$.

For such marginals, is pure bundling also robustly optimal?

Pure bundling is not robustly optimal

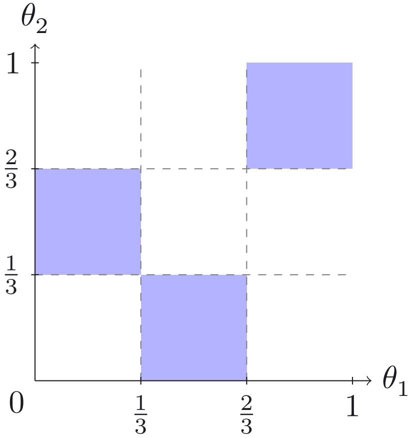
Let the marginals be $F_1 = F_2 = \mathbb{U}[0, 1]$; this satisfies the monotone hazard rate assumption.

Thus, $p_1^* = p_2^* = 1/2$, and the separate sales profit is $\Pi_1^\star + \Pi_2^\star = 1/2$.

Pure bundling is not robustly optimal

Let the marginals be $F_1 = F_2 = \mathbb{U}[0,1]$; this satisfies the monotone hazard rate assumption.

Thus, $p_1^* = p_2^* = 1/2$, and the separate sales profit is $\Pi_1^\star + \Pi_2^\star = 1/2$.



Take the joint distribution to be uniform on the shaded region.

Both marginals are $\mathbb{U}[0,1]$.

If the grand-bundle is priced $p_1^* + p_2^* = 1$, then it is only purchased in the top right square.

This yields profit $\frac{1}{3} < \frac{1}{2} = \Pi_1^\star + \Pi_2^\star$.

Pure bundling is not robustly optimal.

Why is pure bundling not robustly optimal?

The comonotonic distribution is *one* worst case distribution for the separate sales mechanism.

- But it is not the only worst case distribution.
- Indeed all distributions in $\mathcal{F}$ yield the same profit.

But all distributions in $\mathcal{F}$ do not yield the same profit for pure bundling.

- The comonotonic distribution is not a worst case distribution for pure bundling.
- Tthis mechanism cannot provide the same profit guarantee.

To establish the uniqueness of separate sales, we have to identify one distribution for every other mechanism to show that they provide a lower profit guarantee.

Discussion

Why did robustness make the problem easier?

- We did not have to solve the problem for arbitrary F.
- We simply needed to guess a candidate optimum (which is guided by the modeling assumptions).
- And then find *one distribution* at which the highest profit is $\Pi^\star$.

Discussion

Why did robustness make the problem easier?

- We did not have to solve the problem for arbitrary F.
- We simply needed to guess a candidate optimum (which is guided by the modeling assumptions).
- And then find *one distribution* at which the highest profit is $\Pi^\star$.

What kind of data would allow the seller to estimate the marginals but not the joint distribution?

Discussion

Why did robustness make the problem easier?

- We did not have to solve the problem for arbitrary F.
- We simply needed to guess a candidate optimum (which is guided by the modeling assumptions).
- And then find *one distribution* at which the highest profit is $\Pi^\star$.

What kind of data would allow the seller to estimate the marginals but not the joint distribution?

What are examples of multiple good sellers employing separate sales?

- What data do these sellers have access to?

What are other alternative ways of modeling robustness?

Discussion

Why did robustness make the problem easier?

- We did not have to solve the problem for arbitrary F.
- We simply needed to guess a candidate optimum (which is guided by the modeling assumptions).
- And then find *one distribution* at which the highest profit is $\Pi^\star$.

What kind of data would allow the seller to estimate the marginals but not the joint distribution?

What are examples of multiple good sellers employing separate sales?

- What data do these sellers have access to?

What are other alternative ways of modeling robustness?

How can we explain practices like pure bundling?

Model: Informational robustness

Seller does not know how the buyer learns (Deb+Roesler, 2024)

Consider a framework that is (a little) more general than additive values.

Given type θ, the buyer's value for a bundle $b \subseteq N$ is

$$u(\theta, b) = \kappa_b \sum_{i \in b} \theta_i,$$

where $\kappa_b \geq 0$ and we normalize $\kappa_N = 1$, and $u(\theta, \emptyset) = 0$.

Assumption (**weak free-disposal**): The greatest surplus is generated by trading the grand-bundle N. That is,

$$u(\theta, N) \geq u(\theta, b) \quad \text{for all } b \subseteq N, \text{ and all } \theta \in \Theta.$$

Model: Informational robustness

Seller does not know how the buyer learns (Deb+Roesler, 2024)

Consider a framework that is (a little) more general than additive values.

Given type θ, the buyer's value for a bundle $b \subseteq N$ is

$$u(\theta, b) = \kappa_b \sum_{i \in b} \theta_i,$$

where $\kappa_b \geq 0$ and we normalize $\kappa_N = 1$, and $u(\theta, \emptyset) = 0$.

Assumption (**weak free-disposal**): The greatest surplus is generated by trading the grand-bundle N. That is,

$$u(\theta, N) \geq u(\theta, b) \quad \text{for all } b \subseteq N, \text{ and all } \theta \in \Theta.$$

- This does not require $\kappa_{b'} \geq \kappa_b$ for proper subsets $b \subseteq b' \subset N$.
- When $\underline{\theta} > 0$ we can have $\kappa_b > \kappa_N = 1$ but we need $|b|\kappa_b \leq n$.

Model: Prior distribution

Assumption *(exchangeable prior):* The prior distribution F is exchangeable.

- This means that F is symmetric.
- Permuting the goods does not change the distribution.

This is stronger than required (Che+Zhong, 2024) but it simplifies the proof.

The buyer lears about her type.

- The seller does not know how the buyer learns.
- She is robust with respect to all possible signals.

Model: Buyer learning

The buyer learns about his type θ by observing a signal $s \in S$.

Signals are distributed by $G_{S\times\Theta}(s \mid \theta)$.

It is wlog to restrict attention to *unbiased experiments* where $S = \Theta$ and

$$s = \mathbb{E}_{G_{S\times\Theta}}[\theta \mid s].$$

In words, s itself is the *posterior estimate* of θ.

We use G to denote the distribution over posterior estimates induced by $G_{S\times\Theta}$.

We denote the set of these posterior estimate distributions by $\mathcal{G}$.

- The seller is robust with respect to this set.

Since $S = \Theta$, IC+IR mechanisms are defined identically.

Optimal informationally robust mechanism: Definition

Once again, a mechanism (q, t) provides a *profit guarantee* of Π if

$$\min_{G \in \mathcal{G}} \left\{ \int_S \left[t(s) - c \sum_{i=1}^{n} q_i(s) \right] dG(s) \right\} \geq \Pi.$$

An *optimal informationally robust mechanism* $(q^\star, t^\star)$ is the mechanism that provides the highest revenue guarantee.

Formally,

$$(q^\star, t^\star) \in \arg\max_{(q,t)} \min_{G \in \mathcal{G}} \left\{ \int_S \left[t(s) - c \sum_{i=1}^{n} q_i(s) \right] dG(s) \right\}.$$

Optimal informationally robust mechanism: Characterization

A *random pure bundling mechanism* is a mechanism in which the buyer is only ever allocated the grand bundle but the allocation could be random.

We use $\mathcal{M}^{rPB} \subset \mathcal{M}$ to denote the set of random pure bundling mechanisms.

> **Theorem**
>
> There is a random pure bundling mechanism that is an optimal informationally robust mechanism.

Optimal informationally robust mechanism: Discussion

Why does pure bundling arise?

- Selecting a one-dimensional mechanism effectively shrinks the set of signals the seller has to guard against.
- This is because only the distribution of the grand bundle value matters. Given this, the marginal distributions of each dimension do not matter, nor do the correlations across dimensions.

Implications:

- Simple one-dimensional mechanisms are effective when the seller does know the type distribution precisely.
- Perhaps one reason why pure bundling is commonly used for digital goods.

Once again, uniqueness for the robust optimum is not established.

Perfectly correlated signals

Every signal G induces a distribution $\overline{G}$ over grand bundle estimates

$$\bar{s} = \mathbb{E}[\theta_1 + \cdots + \theta_n | \bar{s}].$$

Multiple signals can induce the same distribution over grand bundle estimates.

- For example, the distribution $\bar{s}$ conditional on $(\theta_1, \ldots, \theta_n)$ can be "permuted."

In particular, one of those signals is comonotone (perfectly correlated).

- This reduction is one dimension is once again useful.

Lemma

For every signal $G \in \mathcal{G}$, there exists a signal $G' \in \mathcal{G}$ that is comonotone (perfectly correlated) such that G and G' induce the same distribution $\overline{G} = \overline{G}'$ over grand bundle estimates.

Comonotone signals: Intuition

Take a distribution $\overline{G}$ over grand bundle estimates induced by some signal G.

Consider a signal where the buyer only learns the grand bundle value but nothing further.

- Let the distribution of grand bundle estimates from this signal be $\overline{G}$.
- This is a valid signal.

Comonotone signals: Intuition

Take a distribution $\overline{G}$ over grand bundle estimates induced by some signal G.

Consider a signal where the buyer only learns the grand bundle value but nothing further.

- Let the distribution of grand bundle estimates from this signal be $\overline{G}$.
- This is a valid signal.

Note that, because the prior F is exchangeable, the posterior will satisfy

$$\mathbb{E}[\theta_1|\overline{s}] = \cdots = \mathbb{E}[\theta_n|\overline{s}]$$

for any grand bundle estimate $\overline{s}$.

Comonotone signals: Intuition

Take a distribution $\overline{G}$ over grand bundle estimates induced by some signal G.

Consider a signal where the buyer only learns the grand bundle value but nothing further.

- Let the distribution of grand bundle estimates from this signal be $\overline{G}$.
- This is a valid signal.

Note that, because the prior F is exchangeable, the posterior will satisfy

$$\mathbb{E}[\theta_1|\overline{s}] = \cdots = \mathbb{E}[\theta_n|\overline{s}]$$

for any grand bundle estimate $\overline{s}$.

Thus, the buyer's posterior about $(\theta_1, \ldots, \theta_n)$ from this signal is comonotone.

Optimal informationally robust mechanism: Proof overview

$$\max_{M\in\mathcal{M}} \min_{G\in\mathcal{G}} \{\Pi_M(G)\} \geq \max_{M\in\mathcal{M}^{rPB}} \min_{G\in\mathcal{G}} \{\Pi_M(G)\}$$
$$= \max_{M\in\mathcal{M}^{rPB}} \min_{G\in\mathcal{G}^{PC}} \{\Pi_M(G)\} = \Pi^\star_{Du}$$

$$\max_{M\in\mathcal{M}} \min_{G\in\mathcal{G}} \{\Pi_M(G)\} \leq \max_{M\in\mathcal{M}} \min_{G\in\mathcal{G}^{PC}} \{\Pi_M(G)\}$$
$$= \max_{M\in\mathcal{M}^{rPB}} \min_{G\in\mathcal{G}^{PC}} \{\Pi_M(G)\} = \Pi^\star_{Du}$$

Notation:

$\mathcal{G}^{PC} =$ set of perfectly correlated signals.

$\mathcal{M} =$ set of IC mechanisms.

$\mathcal{M}^{rPB} =$ set of random pure bundling mechanisms.

$\Pi_M(G) =$ seller profit from mechanism M given signal G.

$\Pi^\star_{Du} =$ max-min profit with 1 good.

Optimal informationally robust mechanism: Proof overview

$$\max_{M\in\mathcal{M}} \min_{G\in\mathcal{G}} \{\Pi_M(G)\} \geq \max_{M\in\mathcal{M}^{rPB}} \min_{G\in\mathcal{G}} \{\Pi_M(G)\}$$

$$= \max_{M\in\mathcal{M}^{rPB}} \min_{G\in\mathcal{G}^{PC}} \{\Pi_M(G)\} = \Pi^\star_{Du}$$

$$\max_{M\in\mathcal{M}} \min_{G\in\mathcal{G}} \{\Pi_M(G)\} \leq \max_{M\in\mathcal{M}} \min_{G\in\mathcal{G}^{PC}} \{\Pi_M(G)\}$$

$$= \max_{M\in\mathcal{M}^{rPB}} \min_{G\in\mathcal{G}^{PC}} \{\Pi_M(G)\} = \Pi^\star_{Du}$$

Notation:

$\mathcal{G}^{PC} =$ set of perfectly correlated signals.

$\mathcal{M} =$ set of IC mechanisms.

$\mathcal{M}^{rPB} =$ set of random pure bundling mechanisms.

$\Pi_M(G) =$ seller profit from mechanism M given signal G.

$\Pi^\star_{Du} =$ max-min profit with 1 good.

Max taken over smaller set.

Optimal informationally robust mechanism: Proof overview

$$\max_{M\in\mathcal{M}} \min_{G\in\mathcal{G}} \{\Pi_M(G)\} \geq \max_{M\in\mathcal{M}^{rPB}} \min_{G\in\mathcal{G}} \{\Pi_M(G)\}$$

$$= \max_{M\in\mathcal{M}^{rPB}} \min_{G\in\mathcal{G}^{PC}} \{\Pi_M(G)\} = \Pi^\star_{Du}$$

$$\max_{M\in\mathcal{M}} \min_{G\in\mathcal{G}} \{\Pi_M(G)\} \leq \max_{M\in\mathcal{M}} \min_{G\in\mathcal{G}^{PC}} \{\Pi_M(G)\}$$

$$= \max_{M\in\mathcal{M}^{rPB}} \min_{G\in\mathcal{G}^{PC}} \{\Pi_M(G)\} = \Pi^\star_{Du}$$

Notation:

$\mathcal{G}^{PC} =$ set of perfectly correlated signals.

$\mathcal{M} =$ set of IC mechanisms.

$\mathcal{M}^{rPB} =$ set of random pure bundling mechanisms.

$\Pi_M(G) =$ seller profit from mechanism M given signal G.

$\Pi^\star_{Du} =$ max-min profit with 1 good.

Consequence of the lemma.

Optimal informationally robust mechanism: Proof overview

$$\max_{M\in\mathcal{M}}\min_{G\in\mathcal{G}}\{\Pi_M(G)\} \geq \max_{M\in\mathcal{M}^{rPB}}\min_{G\in\mathcal{G}}\{\Pi_M(G)\}$$

$$= \max_{M\in\mathcal{M}^{rPB}}\min_{G\in\mathcal{G}^{PC}}\{\Pi_M(G)\} = \Pi^\star_{Du}$$

$$\max_{M\in\mathcal{M}}\min_{G\in\mathcal{G}}\{\Pi_M(G)\} \leq \max_{M\in\mathcal{M}}\min_{G\in\mathcal{G}^{PC}}\{\Pi_M(G)\}$$

$$= \max_{M\in\mathcal{M}^{rPB}}\min_{G\in\mathcal{G}^{PC}}\{\Pi_M(G)\} = \Pi^\star_{Du}$$

Notation:

$\mathcal{G}^{PC} =$ set of perfectly correlated signals.

$\mathcal{M} =$ set of IC mechanisms.

$\mathcal{M}^{rPB} =$ set of random pure bundling mechanisms.

$\Pi_M(G) =$ seller profit from mechanism M given signal G.

$\Pi^\star_{Du} =$ max-min profit with 1 good.

One-dimensional max-min problem (Du 2018).

Optimal informationally robust mechanism: Proof overview

$$\max_{M\in\mathcal{M}} \min_{G\in\mathcal{G}}\{\Pi_M(G)\} \geq \max_{M\in\mathcal{M}^{rPB}} \min_{G\in\mathcal{G}}\{\Pi_M(G)\}$$
$$= \max_{M\in\mathcal{M}^{rPB}} \min_{G\in\mathcal{G}^{PC}}\{\Pi_M(G)\} = \Pi^\star_{Du}$$

$$\max_{M\in\mathcal{M}} \min_{G\in\mathcal{G}}\{\Pi_M(G)\} \leq \max_{M\in\mathcal{M}} \min_{G\in\mathcal{G}^{PC}}\{\Pi_M(G)\}$$
$$= \max_{M\in\mathcal{M}^{rPB}} \min_{G\in\mathcal{G}^{PC}}\{\Pi_M(G)\} = \Pi^\star_{Du}$$

Min taken over smaller set.

Notation:

$\mathcal{G}^{PC} =$ set of perfectly correlated signals.

$\mathcal{M} =$ set of IC mechanisms.

$\mathcal{M}^{rPB} =$ set of random pure bundling mechanisms.

$\Pi_M(G) =$ seller profit from mechanism M given signal G.

$\Pi^\star_{Du} =$ max-min profit with 1 good.

Optimal informationally robust mechanism: Proof overview

Notation:

$\mathcal{G}^{PC} =$ set of perfectly correlated signals.

$\mathcal{M} =$ set of IC mechanisms.

$\mathcal{M}^{rPB} =$ set of random pure bundling mechanisms.

$\Pi_M(G) =$ seller profit from mechanism M given signal G.

$\Pi^\star_{Du} =$ max-min profit with 1 good.

$$\max_{M\in\mathcal{M}} \min_{G\in\mathcal{G}} \{\Pi_M(G)\} \geq \max_{M\in\mathcal{M}^{rPB}} \min_{G\in\mathcal{G}} \{\Pi_M(G)\}$$

$$= \max_{M\in\mathcal{M}^{rPB}} \min_{G\in\mathcal{G}^{PC}} \{\Pi_M(G)\} = \Pi^\star_{Du}$$

$$\max_{M\in\mathcal{M}} \min_{G\in\mathcal{G}} \{\Pi_M(G)\} \leq \max_{M\in\mathcal{M}} \min_{G\in\mathcal{G}^{PC}} \{\Pi_M(G)\}$$

$$= \max_{M\in\mathcal{M}^{rPB}} \min_{G\in\mathcal{G}^{PC}} \{\Pi_M(G)\} = \Pi^\star_{Du}$$

No loss in restricting attention to $\mathcal{M}^{rPB}$ when $G \in \mathcal{G}^{PC}$.

Optimal informationally robust mechanism: Proof overview

Notation:

$\mathcal{G}^{PC} =$ set of perfectly correlated signals.

$\mathcal{M} =$ set of IC mechanisms.

$\mathcal{M}^{rPB} =$ set of random pure bundling mechanisms.

$\Pi_M(G) =$ seller profit from mechanism M given signal G.

$\Pi^\star_{Du} =$ max-min profit with 1 good.

$$\max_{M\in\mathcal{M}} \min_{G\in\mathcal{G}}\{\Pi_M(G)\} \geq \max_{M\in\mathcal{M}^{rPB}} \min_{G\in\mathcal{G}}\{\Pi_M(G)\}$$

$$= \max_{M\in\mathcal{M}^{rPB}} \min_{G\in\mathcal{G}^{PC}}\{\Pi_M(G)\} = \Pi^\star_{Du}$$

$$\max_{M\in\mathcal{M}} \min_{G\in\mathcal{G}}\{\Pi_M(G)\} \leq \max_{M\in\mathcal{M}} \min_{G\in\mathcal{G}^{PC}}\{\Pi_M(G)\}$$

$$= \max_{M\in\mathcal{M}^{rPB}} \min_{G\in\mathcal{G}^{PC}}\{\Pi_M(G)\} = \Pi^\star_{Du}$$

Consider $s = (\hat{s}, \dots, \hat{s})$ in the support of $G \in \mathcal{G}^{PC}$.

Optimal informationally robust mechanism: Proof overview

Notation:

$\mathcal{G}^{PC} =$ set of perfectly correlated signals.

$\mathcal{M} =$ set of IC mechanisms.

$\mathcal{M}^{rPB} =$ set of random pure bundling mechanisms.

$\Pi_M(G) =$ seller profit from mechanism M given signal G.

$\Pi^\star_{Du} =$ max-min profit with 1 good.

$$\max_{M\in\mathcal{M}} \min_{G\in\mathcal{G}} \{\Pi_M(G)\} \geq \max_{M\in\mathcal{M}^{rPB}} \min_{G\in\mathcal{G}} \{\Pi_M(G)\}$$

$$= \max_{M\in\mathcal{M}^{rPB}} \min_{G\in\mathcal{G}^{PC}} \{\Pi_M(G)\} = \Pi^\star_{Du}$$

$$\max_{M\in\mathcal{M}} \min_{G\in\mathcal{G}} \{\Pi_M(G)\} \leq \max_{M\in\mathcal{M}} \min_{G\in\mathcal{G}^{PC}} \{\Pi_M(G)\}$$

$$= \max_{M\in\mathcal{M}^{rPB}} \min_{G\in\mathcal{G}^{PC}} \{\Pi_M(G)\} = \Pi^\star_{Du}$$

$q^{rPB}(s,N) = \sum_b \frac{\kappa_b|b|}{n} q(s,b),$

$q^{rPB}(s,b) = 0$ if $b \neq N,$

Optimal informationally robust mechanism: Proof overview

Notation:

$\mathcal{G}^{PC} =$ set of perfectly correlated signals.

$\mathcal{M} =$ set of IC mechanisms.

$\mathcal{M}^{rPB} =$ set of random pure bundling mechanisms.

$\Pi_M(G) =$ seller profit from mechanism M given signal G.

$\Pi^\star_{Du} =$ max-min profit with 1 good.

$$\max_{M\in\mathcal{M}} \min_{G\in\mathcal{G}}\{\Pi_M(G)\} \geq \max_{M\in\mathcal{M}^{rPB}} \min_{G\in\mathcal{G}}\{\Pi_M(G)\}$$

$$= \max_{M\in\mathcal{M}^{rPB}} \min_{G\in\mathcal{G}^{PC}}\{\Pi_M(G)\} = \Pi^\star_{Du}$$

$$\max_{M\in\mathcal{M}} \min_{G\in\mathcal{G}}\{\Pi_M(G)\} \leq \max_{M\in\mathcal{M}} \min_{G\in\mathcal{G}^{PC}}\{\Pi_M(G)\}$$

$$= \max_{M\in\mathcal{M}^{rPB}} \min_{G\in\mathcal{G}^{PC}}\{\Pi_M(G)\} = \Pi^\star_{Du} \quad t^{rPB}(s) = t(s).$$

Optimal informationally robust mechanism: Proof overview

Notation:

$\mathcal{G}^{PC} =$ set of perfectly correlated signals.

$\mathcal{M} =$ set of IC mechanisms.

$\mathcal{M}^{rPB} =$ set of random pure bundling mechanisms.

$\Pi_M(G) =$ seller profit from mechanism M given signal G.

$\Pi^\star_{Du} =$ max-min profit with 1 good.

$$\max_{M\in\mathcal{M}} \min_{G\in\mathcal{G}} \{\Pi_M(G)\} \geq \max_{M\in\mathcal{M}^{rPB}} \min_{G\in\mathcal{G}} \{\Pi_M(G)\}$$

$$= \max_{M\in\mathcal{M}^{rPB}} \min_{G\in\mathcal{G}^{PC}} \{\Pi_M(G)\} = \Pi^\star_{Du}$$

$$\max_{M\in\mathcal{M}} \min_{G\in\mathcal{G}} \{\Pi_M(G)\} \leq \max_{M\in\mathcal{M}} \min_{G\in\mathcal{G}^{PC}} \{\Pi_M(G)\}$$

$$= \max_{M\in\mathcal{M}^{rPB}} \min_{G\in\mathcal{G}^{PC}} \{\Pi_M(G)\} = \Pi^\star_{Du}$$

$u(s,N)q^{rPB}(s',N) = n\hat{s}\sum_b \frac{\kappa_b|b|}{n} q(s',b),$

Optimal informationally robust mechanism: Proof overview

Notation:

$\mathcal{G}^{PC} =$ set of perfectly correlated signals.

$\mathcal{M} =$ set of IC mechanisms.

$\mathcal{M}^{rPB} =$ set of random pure bundling mechanisms.

$\Pi_M(G) =$ seller profit from mechanism M given signal G.

$\Pi^\star_{Du} =$ max-min profit with 1 good.

$$\max_{M\in\mathcal{M}} \min_{G\in\mathcal{G}} \{\Pi_M(G)\} \geq \max_{M\in\mathcal{M}^{rPB}} \min_{G\in\mathcal{G}} \{\Pi_M(G)\}$$

$$= \max_{M\in\mathcal{M}^{rPB}} \min_{G\in\mathcal{G}^{PC}} \{\Pi_M(G)\} = \Pi^\star_{Du}$$

$$\max_{M\in\mathcal{M}} \min_{G\in\mathcal{G}} \{\Pi_M(G)\} \leq \max_{M\in\mathcal{M}} \min_{G\in\mathcal{G}^{PC}} \{\Pi_M(G)\}$$

$$= \max_{M\in\mathcal{M}^{rPB}} \min_{G\in\mathcal{G}^{PC}} \{\Pi_M(G)\} = \Pi^\star_{Du}$$

$u(s,N)q^{rPB}(s',N) = \sum_b \kappa_b |b| \hat{s}\, q(s',b),$

Optimal informationally robust mechanism: Proof overview

Notation:

$\mathcal{G}^{PC} =$ set of perfectly correlated signals.

$\mathcal{M} =$ set of IC mechanisms.

$\mathcal{M}^{rPB} =$ set of random pure bundling mechanisms.

$\Pi_M(G) =$ seller profit from mechanism M given signal G.

$\Pi^\star_{Du} =$ max-min profit with 1 good.

$$\max_{M\in\mathcal{M}} \min_{G\in\mathcal{G}} \{\Pi_M(G)\} \geq \max_{M\in\mathcal{M}^{rPB}} \min_{G\in\mathcal{G}} \{\Pi_M(G)\}$$
$$= \max_{M\in\mathcal{M}^{rPB}} \min_{G\in\mathcal{G}^{PC}} \{\Pi_M(G)\} = \Pi^\star_{Du}$$

$$\max_{M\in\mathcal{M}} \min_{G\in\mathcal{G}} \{\Pi_M(G)\} \leq \max_{M\in\mathcal{M}} \min_{G\in\mathcal{G}^{PC}} \{\Pi_M(G)\}$$
$$= \max_{M\in\mathcal{M}^{rPB}} \min_{G\in\mathcal{G}^{PC}} \{\Pi_M(G)\} = \Pi^\star_{Du}$$

$u(s,N)q^{rPB}(s',N) = \sum_b u(s,b)q(s',b).$

The same argument works for other ambiguity sets

The key feature of the argument was that for every distribution of grand bundle estimates, there exists a comonotone distribution that induces the same distribution of grand bundle estimates.

This would also be true if the ambiguity set were

$$\{G \in \Delta(\Theta) : \mathbb{E}_G[\theta_1 + \cdots + \theta_n] = \lambda\}$$

or

$$\left\{G \in \Delta(\Theta) : \mathbb{E}_G[\theta_1] = \cdots = \mathbb{E}_G[\theta_n] = \frac{\lambda}{n}\right\}$$

for some known constant $\lambda > 0$.

The same argument can be applied to this case to show the robust optimality of random pure bundling.

Discussion

Observe the similarities between the arguments for the last two papers.

- In both cases, the worst case involves comonotone distributions.
- This makes the problem one-dimensional (and hence tractable).

Discussion

Observe the similarities between the arguments for the last two papers.

- In both cases, the worst case involves comonotone distributions.
- This makes the problem one-dimensional (and hence tractable).

In (the symmetric case of) Carroll:

- There is a single worst case distribution.
- Both separate sales and pure bundling yield the same profit for this distribution.
- But pure bundling cannot provide the same profit guarantee.

Discussion

Observe the similarities between the arguments for the last two papers.

- In both cases, the worst case involves comonotone distributions.
- This makes the problem one-dimensional (and hence tractable).

In (the symmetric case of) Carroll:

- There is a single worst case distribution.
- Both separate sales and pure bundling yield the same profit for this distribution.
- But pure bundling cannot provide the same profit guarantee.

In Deb+Roesler:

- The proof does not identify a single worst case distribution; hence randomization is necessary.
- But randomizing the price of each good independently is not the same as randomizing the price of the grand bundle.
- Indeed if the prior is independent across dimensions, no random separate sales mechanism can achieve the same profit guarantee.

Using historical data to improve screening

In both previous settings, we started off with the assumption that the seller knew something about the set of buyer's type distributions.

But where does the seller get this information from?

Using historical data to improve screening

In both previous settings, we started off with the assumption that the seller knew something about the set of buyer's type distributions.

But where does the seller get this information from?

- One natural source is data on consumption from previously offered mechanisms.

Suppose the seller has offered several mechanisms in the past.

- She observes what groups of buyers chose from each of those mechanisms.
- These mechanisms are not assumed to be optimal.

Can the seller use this historical data to offer a new mechanism with a strictly higher profit guarantee?

- The past choices reveal some information about the buyers' type distribution.

Using historical data to improve screening

In both previous settings, we started off with the assumption that the seller knew something about the set of buyer's type distributions.

But where does the seller get this information from?

- One natural source is data on consumption from previously offered mechanisms.

Suppose the seller has offered several mechanisms in the past.

- She observes what groups of buyers chose from each of those mechanisms.
- These mechanisms are not assumed to be optimal.

Can the seller use this historical data to offer a new mechanism with a strictly higher profit guarantee?

- The past choices reveal some information about the buyers' type distribution.

This is the question posed by Rosenthal (2022).

Screening with historical data: A possible strategy

The seller can just calculate the profits she obtained from each past mechanism.

And she can choose the one that gave her the highest profit.

Question

1. Would this yield the optimal robust mechanism?

Screening with historical data: A possible strategy

The seller can just calculate the profits she obtained from each past mechanism.

And she can choose the one that gave her the highest profit.

> **Question**
>
> 1. Would this yield the optimal robust mechanism?
> 2. If not, does the seller just choose from previously offered prices or menu items?

Example: Choosing from past mechanisms?

Assume two goods $n = 2$, the seller has cost $c = 0$ and the buyer has additive utility.

Suppose the seller offered two separate sales mechanisms in the past:

$$M^{t_1} = (p_1^{t_1}, p_2^{t_1}) = (1, 2) \text{ and } M^{t_2} = (p_1^{t_2}, p_2^{t_2}) = (4, 5).$$

Example: Choosing from past mechanisms?

Assume two goods $n = 2$, the seller has cost $c = 0$ and the buyer has additive utility.

Suppose the seller offered two separate sales mechanisms in the past:

$$M^{t_1} = (p_1^{t_1}, p_2^{t_1}) = (1, 2) \text{ and } M^{t_2} = (p_1^{t_2}, p_2^{t_2}) = (4, 5).$$

She observes the purchase behavior of three groups (A, B, C) of equal mass:

group	$(p_1^{t_1}, p_2^{t_1}) = (1, 2)$	$(p_1^{t_2}, p_2^{t_2}) = (4, 5)$
A	both	good 1 only
B	good 2 only	good 2 only
C	good 1 only	good 1 only

Example: Choosing from past mechanisms?

Assume two goods $n = 2$, the seller has cost $c = 0$ and the buyer has additive utility.

Suppose the seller offered two separate sales mechanisms in the past:

$$M^{t_1} = (p_1^{t_1}, p_2^{t_1}) = (1, 2) \text{ and } M^{t_2} = (p_1^{t_2}, p_2^{t_2}) = (4, 5).$$

She observes the purchase behavior of three groups (A, B, C) of equal mass:

group	$(p_1^{t_1}, p_2^{t_1}) = (1, 2)$	$(p_1^{t_2}, p_2^{t_2}) = (4, 5)$
A	both	good 1 only
B	good 2 only	good 2 only
C	good 1 only	good 1 only

These mechanisms yielded profit:

$$\Pi_{M^{t_1}} = \frac{1}{3}(p_1^{t_1} + p_2^{t_1}) + \frac{1}{3}p_2^{t_1} + \frac{1}{3}p_1^{t_1} = 2, \qquad \Pi_{M^{t_2}} = \frac{1}{3}p_1^{t_2} + \frac{1}{3}p_2^{t_2} + \frac{1}{3}p_1^{t_2} = \frac{13}{3}.$$

What does the panel reveal?

The data reveals that the type distributions of each group must be supported in

group	Good 1	Good 2
A	$\theta_1 \geq 4$	$2 \leq \theta_2 \leq 5$
B	$\theta_1 \leq 1$	$\theta_2 \geq 5$
C	$\theta_1 \geq 4$	$\theta_2 \leq 2$

What does the panel reveal?

The data reveals that the type distributions of each group must be supported in

group	Good 1	Good 2
A	$\theta_1 \geq 4$	$2 \leq \theta_2 \leq 5$
B	$\theta_1 \leq 1$	$\theta_2 \geq 5$
C	$\theta_1 \geq 4$	$\theta_2 \leq 2$

I now argue that the seller can receive a strictly higher profit guarantee by mixed bundling.

So consider the mechanism M^{mb} given by the prices

$$\underbrace{p_1 = 4,\ p_2 = 5,}_{\text{individual good prices}}\ \underbrace{p_{12} = 6}_{\text{bundle price}}\ .$$

Observe that the grand bundle was not historically offered in the data.

Buyer demand

group	Good 1	Good 2
A	$\theta_1 \geq 4$	$2 \leq \theta_2 \leq 5$
B	$\theta_1 \leq 1$	$\theta_2 \geq 5$
C	$\theta_1 \geq 4$	$\theta_2 \leq 2$

What is the profit guarantee of this mixed bundling menu?

- Group A: buys the grand bundle since $\theta_2 \leq p_2 = 5$ and $\theta_1 + \theta_2 - p_{12} = (\theta_1 - 4) + (\theta_2 - 2) \geq \theta_1 - 4$;
- Group B: buys good 2 since $\theta_1 \leq p_1 = 4$ and $\theta_2 \geq p_2 = 5$;
- Group C: buys good 1 since $\theta_1 \geq p_1 = 4$ and $\theta_2 \leq p_2 = 5$.

Hence mixed bundling guarantees profit

$$\Pi_{M^{mb}} = \frac{1}{3}p_{12} + \frac{1}{3}p_2 + \frac{1}{3}p_1 = \frac{6+5+4}{3} = 5 > \max\{\Pi_{M^{t_1}}, \Pi_{M^{t_2}}\} = \frac{13}{3}.$$

Mixed-bundling is robustly optimal

group	Good 1	Good 2
A	$\theta_1 \geq 4$	$2 \leq \theta_2 \leq 5$
B	$\theta_1 \leq 1$	$\theta_2 \geq 5$
C	$\theta_1 \geq 4$	$\theta_2 \leq 2$

I now argue that the seller cannot receive a strictly higher profit guarantee.

Mixed-bundling is robustly optimal

group	Good 1	Good 2
A	$\theta_1 \geq 4$	$2 \leq \theta_2 \leq 5$
B	$\theta_1 \leq 1$	$\theta_2 \geq 5$
C	$\theta_1 \geq 4$	$\theta_2 \leq 2$

I now argue that the seller cannot receive a strictly higher profit guarantee.

Consider the distribution $\underline{F}$ that places all the weight on the lower bound of the supports.

$$\begin{aligned}
\text{Group } A &: \text{Probability 1 on } \theta^A = (4, 2), \\
\text{Group } B &: \text{Probability 1 on } \theta^B = (0, 5), \\
\text{Group } C &: \text{Probability 1 on } \theta^C = (4, 0).
\end{aligned}$$

Mixed-bundling is robustly optimal

group	Good 1	Good 2
A	$\theta_1 \geq 4$	$2 \leq \theta_2 \leq 5$
B	$\theta_1 \leq 1$	$\theta_2 \geq 5$
C	$\theta_1 \geq 4$	$\theta_2 \leq 2$

I now argue that the seller cannot receive a strictly higher profit guarantee.

Consider the distribution $\underline{F}$ that places all the weight on the lower bound of the supports.

$$\begin{aligned} &\text{Group } A: \text{Probability 1 on } \theta^A = (4,2), \\ &\text{Group } B: \text{Probability 1 on } \theta^B = (0,5), \\ &\text{Group } C: \text{Probability 1 on } \theta^C = (4,0). \end{aligned}$$

The maximal surplus from trade for this distribution is

$$\frac{1}{3}(4+2) + \frac{1}{3}(0+5) + \frac{1}{3}(4+0) = 5 = \Pi_{M^{mb}}.$$

But this is the profit guarantee of the mixed bundle ($p_1 = 4,\ p_2 = 5,\ p_{12} = 6$).

- Every other type in the support spends weakly more.

Discussion: Why is mixed-bundling required?

group	Good 1	Good 2
A	$\theta_1 \geq 4$	$2 \leq \theta_2 \leq 5$
B	$\theta_1 \leq 1$	$\theta_2 \geq 5$
C	$\theta_1 \geq 4$	$\theta_2 \leq 2$

In order to extract the full surplus at distribution $\underline{F}$, the seller must

- Sell both goods to Group A at a price of 6.
- Sell good 2 to Group B at a price of 4.
- Sell good 1 to Group C at a price of 5.

To extract the surplus from Groups B and C, the seller must offer both goods separately at prices $p_1' = 4$ and $p_2' = 5$.

But then Group A must be offered a bundle discount in order for them to buy both goods.

Example: Are optimal mechanisms always simple?

Once again, assume two goods $n = 2$, zero costs $c = 0$, and additive utility.

The seller has historically offered two separate-sales mechanisms:

$$M^{t_1} = (p_1^{t_1}, p_2^{t_1}) = (1, 1), \qquad M^{t_2} = (p_1^{t_2}, p_2^{t_2}) = (2, 3).$$

She observes the purchase behavior of two groups of equal mass:

group	M^{t_1}	M^{t_2}
A	both	good 2 only
B	both	good 1 only

The choice histories imply

group	Good 1	Good 2
A	$1 \leq \theta_1 \leq 2$	$\theta_2 \geq 3$
B	$\theta_1 \geq 2$	$1 \leq \theta_2 \leq 3.$

A candidate deterministic optimum

group	Good 1	Good 2
A	$1 \leq \theta_1 \leq 2$	$\theta_2 \geq 3$
B	$\theta_1 \geq 2$	$1 \leq \theta_2 \leq 3.$

Consider the distribution $\underline{F}$ that:

$$\text{places probability 1 on } \theta^A = (1,3) \text{ and } \theta^B = (2,1).$$

Consider the deterministic menu M^d that offers good 2 only as an add on:

$$\underbrace{p_1 = 2,}_{\text{good 1}} \qquad \underbrace{p_{12} = 4}_{\text{grand bundle}} .$$

- Group A buys the bundle: $\theta_1^A + \theta_2^A - 4 \geq 0 \geq \theta_1^A - 2$.
- Group B buys good 1: $\theta_1^B - 2 \geq 0 \geq \theta_1^B + \theta_2^B - 4$.

A candidate deterministic optimum

group	Good 1	Good 2
A	$1 \leq \theta_1 \leq 2$	$\theta_2 \geq 3$
B	$\theta_1 \geq 2$	$1 \leq \theta_2 \leq 3.$

Consider the distribution $\underline{F}$ that:

$$\text{places probability 1 on } \theta^A = (1,3) \text{ and } \theta^B = (2,1).$$

Consider the deterministic menu M^d that offers good 2 only as an add on:

$$\underbrace{p_1 = 2}_{\text{good 1}}, \qquad \underbrace{p_{12} = 4}_{\text{grand bundle}} .$$

- Group A buys the bundle: $\theta_1^A + \theta_2^A - 4 \geq 0 \geq \theta_1^A - 2$.
- Group B buys good 1: $\theta_1^B - 2 \geq 0 \geq \theta_1^B + \theta_2^B - 4$.

Thus M^d guarantees

$$\Pi_{M^d} = \frac{1}{2}(4+2) = 3.$$

A candidate deterministic optimum

No other deterministic mechanism guarantees more

group	Good 1	Good 2
A	$1 \leq \theta_1 \leq 2$	$\theta_2 \geq 3$
B	$\theta_1 \geq 2$	$1 \leq \theta_2 \leq 3.$

For any deterministic IC+IR mechanism, let (q^A, t_A) and (q^B, t_B) denote the chosen allocations and payments for both groups.

- If $q^B \neq (1,1)$, IR implies that $t_B \leq 2$ (allocated one good) and $t_A \leq 4$ (allocated both goods).
- If $q^B = (1,1)$ but $q^A \neq (1,1)$, IR implies that $t_B \leq 3$ (allocated both goods) and $t_A \leq 3$ (allocated one good).
- If $q^A = q^B = (1,1)$, IC gives $t_A = t_B$, while B's IR constraint gives $t_A = t_B \leq 3$.

A candidate deterministic optimum

No other deterministic mechanism guarantees more

group	Good 1	Good 2
A	$1 \leq \theta_1 \leq 2$	$\theta_2 \geq 3$
B	$\theta_1 \geq 2$	$1 \leq \theta_2 \leq 3.$

For any deterministic IC+IR mechanism, let (q^A, t_A) and (q^B, t_B) denote the chosen allocations and payments for both groups.

- If $q^B \neq (1,1)$, IR implies that $t_B \leq 2$ (allocated one good) and $t_A \leq 4$ (allocated both goods).
- If $q^B = (1,1)$ but $q^A \neq (1,1)$, IR implies that $t_B \leq 3$ (allocated both goods) and $t_A \leq 3$ (allocated one good).
- If $q^A = q^B = (1,1)$, IC gives $t_A = t_B$, while B's IR constraint gives $t_A = t_B \leq 3$.

In every case,

$$t_A + t_B \leq 6 \quad \Longrightarrow \quad \mathbb{E}_{\underline{F}}[t] \leq 3.$$

Hence $\Pi_{M^d} = 3$ is the highest deterministic profit guarantee.

Randomization can do better

group	Good 1	Good 2
A	$1 \leq \theta_1 \leq 2$	$\theta_2 \geq 3$
B	$\theta_1 \geq 2$	$1 \leq \theta_2 \leq 3.$

Offer the previously untried menu M:

- $\left(\left(1, \frac{1}{2}\right), \frac{5}{2}\right)$: Good 1 for sure and Good 2 with probability $\frac{1}{2}$ at price $\frac{5}{2}$.
- $((1, 1), 4)$: The grand bundle at price 4.

Group A buys the bundle because

$$(\theta_1 + \theta_2 - 4) - \left(\theta_1 + \frac{1}{2}\theta_2 - \frac{5}{2}\right) = \frac{1}{2}\theta_2 - \frac{3}{2} \geq 0.$$

Group B buys the lottery because the preceding difference is weakly negative, and

$$\theta_1 + \frac{1}{2}\theta_2 \geq 2 + \frac{1}{2} = \frac{5}{2}.$$

Therefore

$$\Pi_M = \frac{1}{2}\left(4 + \frac{5}{2}\right) = \frac{13}{4} > 3.$$

Intuition

group	Good 1	Good 2
A	$1 \leq \theta_1 \leq 2$	$\theta_2 \geq 3$
B	$\theta_1 \geq 2$	$1 \leq \theta_2 \leq 3.$

A deterministic menu has only coarse ways to screen the two groups:

- Group B will pay 2 for good 1.
- Giving B the entire second good would turn its allocation into the bundle and prevent the seller from charging A a higher bundle price.

The lottery fills the gap between good 1 and the grand bundle:

- A 50% chance of good 2 raises B's payment from 2 to $5/2$.
- Group A, which values good 2 more strongly, still selects the certain bundle at 4.

Note that types are not ordered.

Discussion

Rosenthal (2022) does not establish a general result for this environment.

Data driven robustness is a promising direction for future research.

- Instead of a panel, a seller might also observe a repeated cross-section.

We can also think of dynamics in this setting.

- The seller can trade off worst case profits against learning.

Selecting amongst robustly optimal mechanisms

As we discussed, there may be multiple robustly optimal mechanisms.

How should we pick amongst them?

One way is to use the criterion of Mishra+Patel+Pavan (2026).

- The principal has a prior that can be used to pick amongst robustly optimal mechanisms.

Another criterion is that of Börgers (2017).

- Consider two robustly optimal mechanisms M and M'.
- They both have the same worst case profits.
- But M yields weakly (and sometimes strictly) higher profits for all distributions.
- Then we should choose M instead of M'.

The first criterion is sharper, the second is more robust.

Concluding remarks

Robustness can make multidimensional screening problems more tractable.

- This approach can be both more realistic and more useful.
- One can study other environments: for example, non-additive preferences, multi-dimensional sequential screening (Bergemann+Deb 2025), auctions etc.

Different assumptions of the ambiguity set can lead to qualitatively different solutions.

- This is both a feature and a bug.

The right optimality concept?

- Think back to Alessandro's discussion yesterday.
- Should we use plain max-min or max-min+prior or regret etc.?
- Again, a feature and a bug.

Uniqueness is hard to establish.