543: Game Theory **Multidimensional Screening**	**Week 5**

Updated July 26, 2021.

Preliminary and incomplete.

Thanks to 534 Fall 2020 students for proofreading! (Remaining errors are mine.)

Source(s): [McAfee et al., 1989], [Daskalakis et al., 2014], [Haghpanah and Hartline, 2021].

We now study multidimensional screening. Our main application will be selling multiple products. We will first see that optimal mechanisms can be complicated: they may offer products as bundles, may involve randomization, and may offer infinitely many options to buyers. Then we discuss when simple mechanisms are optimal.

Examples

Setting.

- Single seller, single buyer, two products (seller cost zero)
- Values are "additive": The value of consuming products together is $\theta_1 + \theta_2$, where θ_i is the value of consuming only product i.

What do optimal mechanism look like (we will define what a "mechanism" is later)?

Example: Buyer type $(\theta_1, \theta_2) \in \{1, 2\}^2$, uniformly distributed.

- Sell separately? Revenue = 2
- Only as a bundle? price = 3, revenue = $9/4 > 2$. Optimal among <u>all</u> mechanisms!
- So the seller benefits from bundling even though products are "doubly independent".

What are all mechanisms? As before, a mechanism is any extensive form game played by the single player whose terminal nodes specify distributions over products and transfers. By the revelation principle, we focus on direct IC and IR mechanisms. A direct mechanism is a pair of functions $q : \Theta \to [0,1]^2$ and $t : \Theta \to R$. (Notice that $q(\theta)$ only specifies the probability of getting each product, and not the joint distribution over products. Why?)

In the example above, we represent a mechanism as a menu, and not as a direct mechanism. The corresponding direct mechanism is obtained by mapping each type to the choice she makes given the options. E.g., when the only option is to buy the grand bundle at price 3, the corresponding direct mechanism is $q_1(\theta) = q_2(\theta) = 1$ and $t(\theta) = 3$ if $\theta_1 + \theta_2 \geq 3$, and $q_1(\theta) = q_2(\theta) = t(\theta) = 0$ otherwise. This observation is general, and is called the "taxation principle": for any mechanism (not necessarily even direct), there exists a "menu representation" of that mechanism, such that the outcomes of the two mechanisms are the same when the player chooses an optimal strategy in each mechanism.

Example: $(\theta_1, \theta_2) \in [0,1]^2$, uniformly distributed.

- Sell separately? Revenue = 1/2.
- Sell as bundle? Sell at price $\sqrt{2/3}$, Revenue $\simeq 0.54$
- Optimal: choose 1 of 3 options
  - Get both product. Price = $\frac{4-\sqrt{2}}{3}$
  - Get product 1 only. Price = 2/3
  - Get product 2 only. Price = 2/3
  - Get nothing, pay nothing.

Example: Buyer type $(\theta_1, \theta_2) \in \{1,2\} \times \{1,3\}$, uniformly distributed.

- Optimal mechanism: choose 1 of 3 options
  - Get both products. Price = 4
  - Get product 1 w.p. 1 and product 2 w.p. 1/2. Price = 2.5
  - Get nothing, pay nothing.

Example: Two "Beta" distributions. Optimal mechanism has infinitely many choices!

So optimal mechanisms can be quite complex. When are "simple" mechanisms optimal?

When is selling separately optimal?

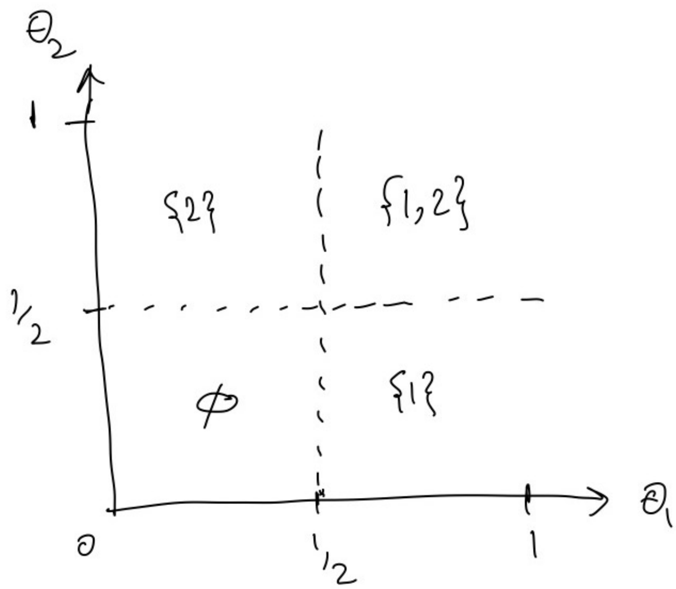
Based on [McAfee et al., 1989].

Essentially never!

Proposition 0.1 *Suppose that θ_1, θ_2 are drawn independently from distribution with pdf $f(\theta) > 0$ over $[\underline{\theta}, \bar{\theta}]$. Then selling separately is not optimal.*

Proof: We show this only for the uniform distribution on $[0,1]^2$. We already claimed in example 2 that selling separately is not optimal. Here we provide a "marginal improvement" argument that generalizes to any distribution.

Optimal prices for selling separately? 1/2. The choices of different types are shown below.



Improvement: buy a single product at price 1/2. Buy bundle at price $1-\epsilon$.

Gain:

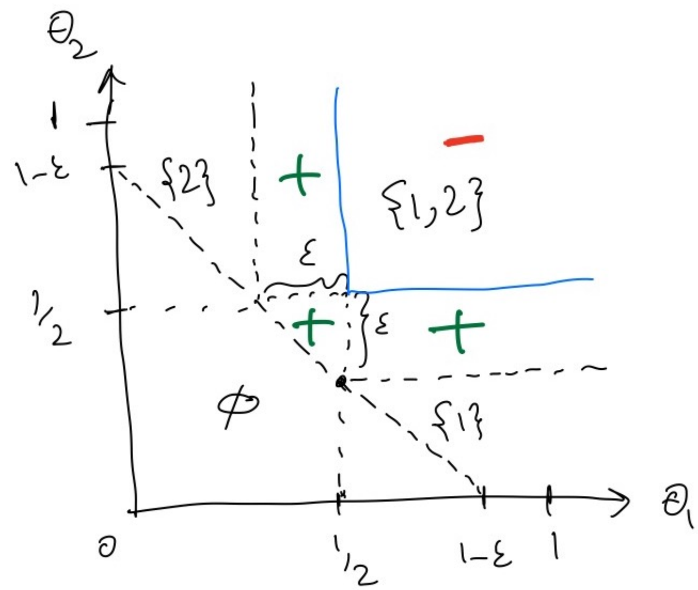
$$\epsilon^2/2(1-\epsilon) + 2\epsilon/2(1/2-\epsilon).$$

Ignoring terms with ϵ^2, ϵ^3, the gain is $\epsilon/2$.

Loss:

$$\epsilon/4.$$

So gain > loss for small ϵ. **QED**



When is pure bundling optimal?

Based on [Haghpanah and Hartline, 2021].

One answer: when values are "perfectly negatively correlated".

- Suppose $\theta_1 + \theta_2 = c$ for all types.
- Then selling only the grand bundle at price c extracts the full surplus.

Based on this observation: pure bundling is optimal with "many" doubly independent products.

- Suppose that $\theta_i \sim F_i$ independently for each product i, with mean μ_i
- As the number of products grow, $\sum_i \theta_i$ concentrates around $\sum_i \mu_i$ by the law of large numbers
- So pure bundling (at a price below $\sum_i \mu_i$) becomes optimal.

Can pure bundling be optimal even if it doesn't extract the full surplus? Yes!

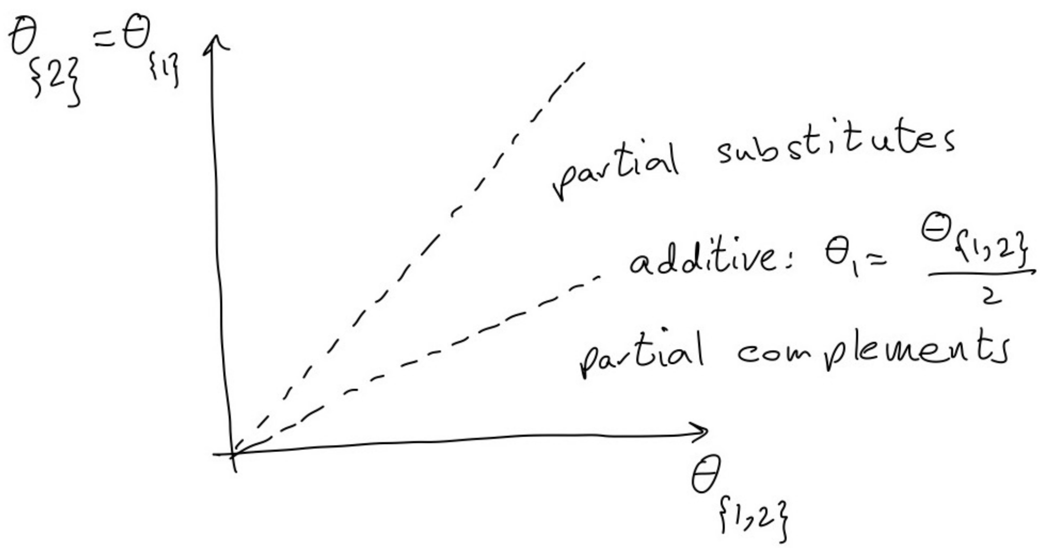
Consider a more general setting where additivity is relaxed.

- Products $1, 2$

- There is a finite set of types $\theta \in \{\theta^1, \ldots, \theta^n\}$
- Let $f^1, \ldots, f^n$ be probabilities.
- A type specifies a value for each bundle $\theta = (\theta_1, \theta_2, \theta_{\{1,2\}}) \in R^3$
- Value for bundle $b \subseteq \{1, 2\}$ is θ_b (value for the empty bundle is zero)
  - Additive $\theta_{\{1,2\}} = \theta_1 + \theta_2$.
  - Partial substitutes $\theta_{\{1,2\}} \leq \theta_1 + \theta_2$.
  - Partial complements $\theta_{\{1,2\}} \geq \theta_1 + \theta_2$.

It is a bit annoying to draw types in 3 dimensions. An advantage of additivity was that we could draw types in two dimensions. But we don't want to assume additivity. So let's instead make another assumption. Namely that products are "identical": $\theta_1 = \theta_2$ for all θ. So we will think of a type as a two-dimensional vector $(\theta_1, \theta_{\{1,2\}})$. We can still distinguish between additive, partial substitutes, and partial complements.

- Additive $\theta_{\{1,2\}} = 2\theta_1$.
- Partial substitutes $\theta_{\{1,2\}} \leq 2\theta_1$.
- Partial complements $\theta_{\{1,2\}} \geq 2\theta_1$.



So let's consider $r(\theta) = \frac{\theta_1}{\theta_{\{1,2\}}}$ as a measure of complementarity between products. Higher r means products are less complementary. Notice that this measure is type-dependent.

Proposition 0.2 *Suppose that different types have different values for the grand bundle, $\theta^i_{\{1,2\}} \neq \theta^j_{\{1,2\}}$ for all $i \neq j$. Then pure bundling is optimal for all distributions if and only if for any pair of types θ^i, θ^j where $\theta^i_{\{1,2\}} < \theta^j_{\{1,2\}}$, we have*

$$r(\theta^i) \leq r(\theta^j).$$

Proof: We Prove given a regularity assumption that we define later. The proof follows problem 1 of PS 1 to obtain an upper bound on revenue, as well as conditions under which the upper bound is tight. When then use a new argument to see how the upper bound can be optimized.

Suppose without loss of generality that $\theta^1_{\{1,2\}} < \theta^2_{\{1,2\}} < \ldots, < \theta^n_{\{1,2\}}$.

Since the two products are identical, we can assume without loss of generality that a mechanism never gives a buyer only product 2 (if it does, we can give product 1 instead and no type's utility changes). So a direct mechanism is (q, t) where $q : \Theta \to [0, 1]^2$ such that $q_{\{1\}}(\theta) + q_{\{1,2\}}(\theta) \leq 1$ for all θ and $t : \Theta \to R$.

The utility of type θ^i from reporting θ^j is $(\theta^i_{\{1\}} q_{\{1\}}(\theta^j) + \theta^i_{\{1,2\}} q_{\{1,2\}}(\theta^j)) - t(\theta^j)$. We can write this in dot product form as $\theta^i \cdot q(\theta^j) - t(\theta^j)$. So now this looks very similar to the screening problem of selling a single product, except that scalar product $\theta^i q(\theta^j)$ is replaced by dot product $\theta^i \cdot q(\theta^j)$.

IC:

$$\begin{aligned} t(\theta^i) - t(\theta^{i-1}) &\leq \theta^i \cdot q(\theta^i) - \theta^i \cdot q(\theta^{i-1}) \\ &= \theta^i \cdot (q(\theta^j) - q(\theta^{i-1})) \end{aligned}$$

where we define $t(\theta^0) = 0$ and $q_b(\theta^0) = 0$ for all b. So the above constraint for $i = 1$ is the IR constraint for type θ^1, $t(\theta^1) \leq \theta^1 \cdot q(\theta^1)$.

Sum over all $j \leq i$,

$$\begin{aligned} t(\theta_i) &= \sum_{j \leq i} t(\theta_j) - t(\theta_{j-1}) \\ &\leq \sum_{j \leq i} \theta^j \cdot (q(\theta^j) - q(\theta^{j-1})), \end{aligned}$$

with equality if all IC constraints from every type θ^j to θ^{j-1} hold with equality.

So expected revenue of any IC-IR mechanism is at most

$$\begin{aligned} &\sum_i t(\theta_i) f(i) \\ &\leq \sum_i \big(\sum_{j \leq i} \theta^j \cdot (q(\theta^j) - q(\theta^{j-1})) \big) f(i) \\ &= \sum_i \big(\theta^i - \frac{(\theta^{i+1} - \theta^i) \sum_{j>i} f(j)}{f(i)} \big) \cdot q(\theta^i) f(i) \end{aligned}$$

with equality if the IC constraint of every type θ^i to θ^{i-1} holds with equality.

Define $\phi(\theta^i) = \theta^i - \frac{(\theta^{i+1}-\theta^i)\sum_{j>i} f(j)}{f(i)}$. This is a vector generalization of the virtual value from the single product case. Now ϕ is a vector $(\phi_{\{1\}}(\theta^i), \phi_{\{1,2\}}(\theta^i))$.

So upper bound on revenue is

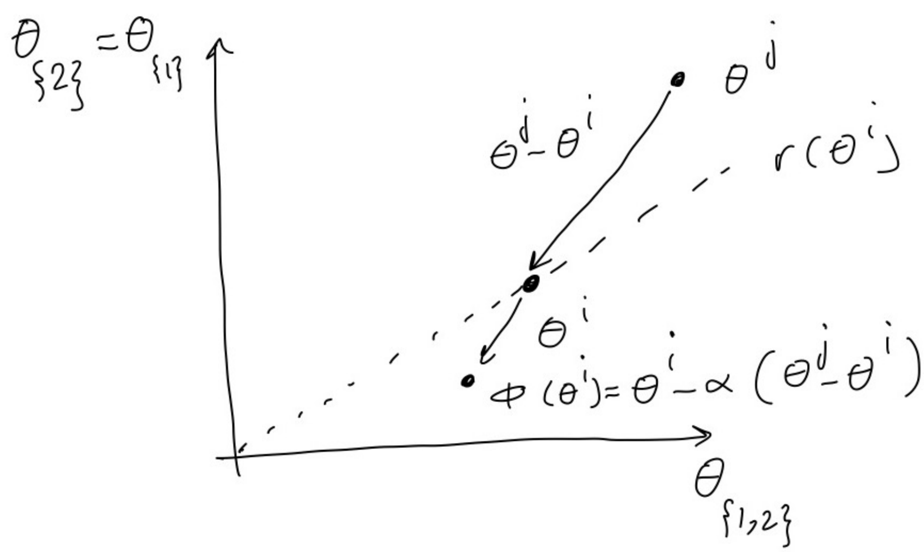
$$\max_q \sum_i \phi(\theta^i) \cdot q(\theta^i) f(i)$$
$$\text{s.t.} \sum_b q_b(\theta^i) \leq 1, \forall \theta^i$$

What is the solution? If $\max_b \phi_b(\theta^i) < 0$ then $q_b(\theta^i) = 0$. Otherwise $q_b(\theta^i) = 1$ for b that maximizes $\phi_b(\theta^i)$.

As shown below, $\phi(\theta^i)$ is "below" θ^i. That is

$$\phi_1(\theta^i) \leq \frac{\theta_1^i}{\theta_{1,2}^i} \cdot \phi_{1,2}(\theta^i).$$

(Note: In class I incorrectly wrote the condition as $\frac{\phi_1(\theta^i)}{\phi_{1,2}(\theta^i)} \leq \frac{\theta_1^i}{\theta_{1,2}^i}$. The two conditions are the same if $\phi_{1,2}(\theta^i) \geq 0$ but are different otherwise. The correct interpretation of $\phi(\theta^i)$ is "below" θ^i is that $\phi_1(\theta^i) \leq \frac{\theta_1^i}{\theta_{1,2}^i} \cdot \phi_{1,2}(\theta^i)$. For example, if $\frac{\theta_1^i}{\theta_{1,2}^i} = r = 0.5$, then $(\phi_1, \phi_2) = (-1, -1)$ is below the line with slope 0.5, and satisfies $\phi_1(\theta^i) \leq \frac{\theta_1^i}{\theta_{1,2}^i} \cdot \phi_{1,2}(\theta^i)$, but violates $\frac{\phi_1(\theta^i)}{\phi_{1,2}(\theta^i)} \leq \frac{\theta_1^i}{\theta_{1,2}^i}$.)



Now consider two cases. If $\phi_{1,2}(\theta^i) \geq 0$, then $\phi_1(\theta^i) \leq \frac{\theta_1^i}{\theta_{1,2}^i}\phi_{1,2}(\theta^i) \leq \phi_{1,2}(\theta^i)$. So the solution is to set $q_{\{1,2\}}(\theta^i) = 1$. If $\phi_{1,2}(\theta^i) < 0$, then $\phi_1(\theta^i) \leq \frac{\theta_1^i}{\theta_{1,2}^i} \cdot \phi_{1,2}(\theta^i) < 0$ so we also have $\phi_1(\theta^i) < 0$. So the solution is to set $q_b(\theta^i) = 0$ for both bundles.

Now suppose that $\phi_{\{1,2\}}(\theta^i)$ is increasing in i (this is the regularity assumption). Let i^* be smallest i such that $\phi_{\{1,2\}}(\theta^i) \geq 0$. The solution is

$$q_b(\theta) = \begin{cases} 1 & \text{if } \theta > \theta^i \& b = \{1,2\} \\ 0 & \text{otherwise} \end{cases}$$

We can implement this allocation by selling the grand bundle $\{1,2\}$ at price $\theta^{i*}_{\{1,2\}}$. **QED**

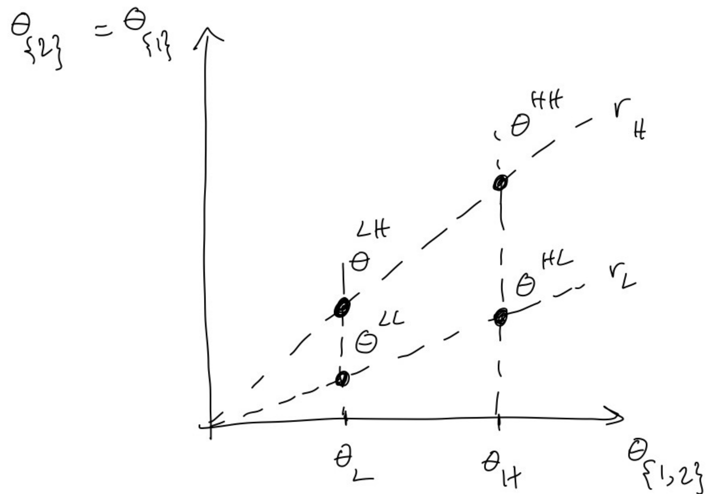
Discussion:

- Pure bundling may be optimal even if $r(\theta) > 0.5$ for both types. That is, if the products are partial substitutes for both types.
- Similarly, pure bundling may not be optimal even if $r(\theta) < 0.5$ for both types.
- So is there a connection between the optimality of pure bundling and complements/substitutes? Yes. What is important is how complementarity changes across types. Pure bundling is optimal if products are "less complementary" for higher types.
- Alternative interpretation: a seller can produce a good in two different qualities. A type θ is (θ_L, θ_H), where θ_L is the value of low quality, and θ_H is the value of high quality. Similarly q_L, q_H are the probabilities of receiving low and high quality, and utility is $\theta \cdot q - t$. This is mathematically equivalent to our problem above. Then "selling only the highest quality is optimal" if θ_L/θ_H is increasing in θ_H.
- For example, in the dynamic allocation application from last time, the ratio has a nice interpretation: it is the discount factor. In the dynamic mechanism design lecture, we assumed that $\theta_L/\theta_H = \delta$ was constant.
- So we have a generalization: selling immediately is optimal if (and to some extent only if) types with a higher value for the product have a higher $\delta(\theta)$, that is, they are more patient.

What if different types don't have different values for the grand bundle?

Consider 4 types $\theta^{LL}, \theta^{LH}, \theta^{HL}, \theta^{HH}$ as drawn below.

Let q be the probability that $\theta_{\{1,2\}} = \theta_H$. Also let $q^L = Pr[r(\theta) = r^H | \theta_{\{1,2\}} = \theta_L] = q^H = Pr[r(\theta) = r^H | \theta_{\{1,2\}} = \theta_H]$ be the probabilities of having a high r conditioned on either having a low or high $\theta_{\{1,2\}}$. Suppose first that $r(\theta)$ and $\theta_{\{1,2\}}$ are independent, that is, $q_L = q_H$.



Notice that we can write the distribution as a convex combination of two distributions,

$$\begin{aligned}(Pr[\theta^{LL}], Pr[\theta^{LH}], Pr[\theta^{HL}], Pr[\theta^{HH}]) &= ((1-q)(1-q_L), (1-q)q_L, q(1-q_L), qq_L) \\ &= q_L(0, 1-q, 0, q) + (1-q_L)(1-q, 0, q, 0).\end{aligned}$$

The interpretation is that we can draw a type from the distribution by first choosing one of two distributions μ^1 or μ^2 with probabilities $q_L, 1-q_L$, and then selecting a type from the chosen distribution.

Suppose now that the seller has a superpower: he can observe whether μ^1 or μ^2 was chosen, and can design his mechanism accordingly. Clearly sellers revenue with this superpower is at least his revenue without. With the superpower, the seller would design the optimal mechanism for the chosen distribution. The optimal revenue for either distribution is $\Pi = \max(\theta_L, q\theta_H)$, which is the revenue of pure bundling for the original distribution. So optimal revenue with superpower is $q^L\Pi + (1-q^L)\Pi = \Pi$. So pure bundling is optimal.

Similarly, if $q^L \leq q^H$, we can write the distribution as a convex combination of three distributions,

$$\begin{aligned}((1-q)(1-q_L), (1-q)q_L, q(1-q_H), qq_H) &= q_L(0, 1-q, 0, q) + (q_H - q_L)(1-q, 0, 0, q) \\ &\quad + (1-q_H)(1-q, 0, q, 0).\end{aligned}$$

Now revenue with superpower is $q_L\Pi + (q_H - q_L)\Pi + (1-q_H)\Pi = \Pi$, and so pure bundling is optimal.

Proposition 0.3 *Pure bundling is optimal for all distributions if "$r(\theta)$ is (first order) stochastically increasing in $\theta_{\{1,2\}}$", that is,*

$$Pr[r(\theta) \leq r | \theta_{\{1,2\}} = \hat{v}]$$

is non increasing in $\hat{v}$ for all r.

The intuition is the same as before. The proof is based on the idea we saw above, and relies on a classical result from statistics which says that if a distribution first order stochastically dominates another, then we can find a monotone mapping from the dominated distribution to the dominating distribution.

References

[Daskalakis et al., 2014] Daskalakis, C., Deckelbaum, A., and Tzamos, C. (2014). The complexity of optimal mechanism design. In SODA. To appear.

[Haghpanah and Hartline, 2021] Haghpanah, N. and Hartline, J. (2021). When is pure bundling optimal? The Review of Economic Studies, 88(3):1127–1156.

[McAfee et al., 1989] McAfee, R. P., McMillan, J., and Whinston, M. D. (1989). Multiproduct monopoly, commodity bundling, and correlation of values. The Quarterly Journal of Economics, 104(2):371–383.