Price Discrimination

Glenn Ellison

Price Discrimination

A classic definition is selling different units of the same good at different prices.

- Regular vs. student tickets at the theater
- This also applies to nonlinear prices: buy one get one 50%, cell phone data, Disneyland rides(?)

The label is also used when firms sell similar goods at different markups

- Stata/IC vs. Stata/SE, coach vs. business-class seat, iPhone 14 128GB vs. 512 GB
- How costs should matter is not clear, $\frac{p_i}{c_i} \neq \frac{p_j}{c_j}$(?) or $p_i - c_i \neq p_j - c_j$(?), but often differences are obvious.
- Can be considered discrimination to not discriminate when costs differ, e.g. free delivery.

Price discrimination requires market power. Otherwise $p_i = c_i$ for all *i*.

Senior Discount
EBT Discount
Students & Under 25
Military Discount
Harvard Discount
JetBlue Discount

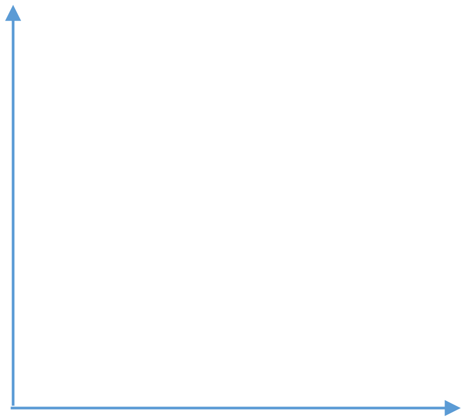
Price Discrimination

Pigou (1920) distinguished between three types of price discrimination:

- **First degree:** Perfect price discrimination. The firm knows consumer preferences completely and can price separately to each consumer.
- **Second degree:** Pricing based on self-selection – consumers with different preferences choose to buy different goods at different prices (e.g., cell phone plans, health insurance contracts, types of gasoline,…)
- **Third degree:** Pricing based on limited observed characteristics, e.g. age, student status, gender, health status,…)

The names are not helpful. And they aren't even in the right order. But somehow they remain the standard.

First-Degree Discrimination



Suppose each consumer *i's* preferences are completely known by the firm, summarized by inverse demand $P_i(x_i)$.

Suppose the seller can set customer-specific nonlinear prices and prevent resale.

The maximum possible profit can be achieved by making each consumer a single take-it-or-leave it offer x_i, T_i.

Consumer i will accept this offer if and only if $T_i \leq \int_0^{x_i} P_i(s)ds$. Clearly setting $T_i = \int_0^{x_i} P_i(s)ds$ is optimal given x_i.

The monopolist solves: $max_{x_i} \int_0^{x_i} P_i(s)ds - cx_i$

<u>FOC</u>: $P_i(x_i^*) = c$

<u>Notes</u>:

1. Quantities are the same as with perfect competition.
2. First-degree discrimination is socially optimal. (As before we ignore distributional issues.)
3. The monopolist can achieve the same outcome with a *customer-specific* two part tariff:

$$T_i(x) = A_i + cx, \text{ with } A_i = \int_0^{x_i} P_i(s)ds - cx$$

Factors Limiting Discrimination

- In reality, a number of factors limit the ability to discriminate:
  - Non-observability of consumer preferences
  - Arbitrage (resale)
  - Inability to monitor consumer purchases
  - Administrative costs
- In some cases these factors can *completely eliminate* the ability to discriminate
  - **Example #1:**
    - Consumers have unit demands
    - Competitive resale market exists $\Longrightarrow p^* = P(\sum_i x_i)$

      Each consumer buys from monopolist only if $T_i \leq p^* x_i$ so might as well charge p^*
  - **Example #2:**
    - Consumers indistinguishable
    - Competitive resale market

      Indistinguishable restricts to T(q). Resale $\Longrightarrow$ $T(q) = A + pq$. Resale also $\Longrightarrow$ $A = 0$.
  - **Example #3:**
    - Unit demands for a single good with non-observable reservation values
    - Risk-neutral consumers

Factors Limiting Discrimination

FROM RICHARD OWEN (London Times)
IN ROME

AN ITALIAN supermarket that offered 20 per cent discounts to the over-60s has discovered the flaw in the scheme: resourceful and unscrupulous grandmothers have been hiring themselves out to shoppers who are under 60 but still like the idea of a discount on their grocery bill.

The management of the supermarket in Udine, near Venice, was surprised and delighted by the "unimaginable success" of its scheme. The "senior citizens discount" produced a 40 per cent surge in takings, making Udine, a medieval town hitherto best known for its Tiepolo frescoes, a mecca for elderly shoppers.

It took a year for the penny to drop. "We seem to have been providing a social service," the manager said ruefully yesterday. "The ordinary shoppers made a saving, and the old people made a fortune in tips."

He knew there had been cases of middle-aged people pretending to be older, but he had no idea that teams of grandmothers had been operating a systematic scam, loading up trolleys of groceries and unloading them into the cars of strangers. The more enterprising had hovered near the entrance, whispering "Need any help with your shopping?" as customers arrived.

Third-Degree Discrimination

Now suppose monopolist can distinguish classes of consumers, but is limited to simple linear pricing within each group.

- Two groups $i = 1,2$
- Demands $X_i(p_i)$
- Constant marginal cost c

With <u>discrimination</u>

$max_{p_i}(p_i - c)X_i(p_i) \;\; \Rightarrow$

$$X_i(p_i^*) + (p_i^* - c)X_i'(p_i^*) = 0 \text{ for all } i$$

$\Leftrightarrow$

$$\frac{p_i^* - c}{p_i^*} = -\frac{1}{\epsilon_i(p_i^*)}$$

Label the markets so that $p_1^* < p_2^*$.

Third-Degree Discrimination

- Two groups $i = 1,2$. Demands $X_i(p_i)$. Constant marginal cost c.

If we ban discrimination:

$$max_p (p-c)X_1(p) + (p-c)X_2(p) \Rightarrow$$

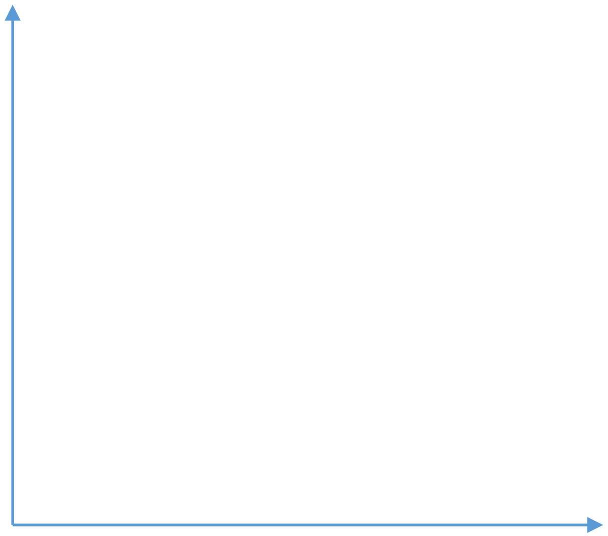
$$\sum_i X_i(p^*) + (p^*-c)X_i'(p^*) = 0$$

<u>Notes</u>:

1. If $\pi_i(p) \equiv (p-c)X_i(p)$ is single-peaked and concave in p for $p \leq p_2^*$, then $p^* \in (p_1^*, p_2^*)$.
   Here, banning discrimination helps one group and hurts the other.

   An elasticity description is $\frac{p^*-c}{p^*} = -\frac{1}{\epsilon(p^*)}$ with

$$\epsilon(p) = \frac{\sum_i X_i'(p)}{p\sum_i X_i(p)} = \sum_i \underbrace{\left(\frac{X_i(p)}{\sum_j X_j(p)}\right)}_{group\ i\ share} \epsilon_i(p)$$



2. Otherwise, the monopolist may choose to serve only one market and set $p^* = p_2^*$.
   No one is better off when discrimination is banned.

Third-Degree Discrimination

Welfare Effects

To think about welfare one should keep in mind two sources of inefficiency:

- Deadweight loss
- Misallocation of goods sold

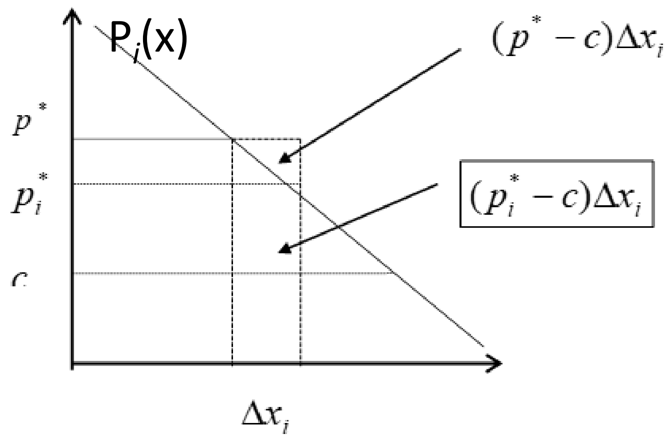
Discriminatory and nonlinear pricing schemes introduce the second. A basic result illustrating this consideration is:

Proposition. *If 3rd degree discrimination does not increase total output relative to uniform pricing, then it reduces social welfare.*

Proof

Write Δx_i and ΔW_i for the change in output and welfare in market i when we shift from uniform to discriminatory pricing. Note that

$$(p^* - c)\Delta x_i \geq \Delta W_i \geq (p_i^* - c)\,\Delta x_i$$



Note: The inequalities are strict if $\Delta x_i \neq 0$.

Third-Degree Discrimination

Welfare Effects

Proposition. *If 3rd degree discrimination does not increase total output relative to uniform pricing, then it reduces social welfare.*

Proof

Write Δx_i and ΔW_i for the changes when we shift from uniform to discriminatory pricing. Note that

$$(p^* - c)\Delta x_i \geq \Delta W_i \geq (p_i^* - c)\,\Delta x_i$$

Summing over i gives

$$(p^* - c)\sum_i \Delta x_i \geq \Delta W \geq \sum_i (p_i^* - c)\,\Delta x_i$$

If discrimination doesn't increase output, the left inequality shows welfare weakly decreases. It strictly decreases if some Δx_i is nonzero.

Notes:

1. Welfare can go either way. A clear example of increasing welfare is if the monopolist will only serve high types under uniform pricing.
2. Equity concerns could go either way: financial aid vs. health insurance.

Third-Degree Discrimination

Welfare Effects

Proposition. *If 3rd degree discrimination does not increase total output relative to uniform pricing, then it reduces social welfare.*

Corollary. *If demands are linear and all markets are served under uniform pricing, then 3rd degree discrimination reduces social welfare.*

Proof:

Suppose $X_i(p) = A_i - B_i p$. Solve to find $p_i^* = \frac{c}{2} + \frac{A_i}{2B_i}$ and $x_i^* = \frac{1}{2}A_i - \frac{1}{2}B_i c$.
Summing gives $\sum_i x_i^* = \frac{1}{2}\sum_i A_i - \frac{1}{2}\left(\sum_i B_i\right)c$.

This is exactly the same as the total output under uniform pricing with demand $X(p) = \sum_i A_i - \sum_i B_i p$.

Notes:

1. This is a special property of linear demands.
2. It highlights a general concern. When one is specifying structural models that will be estimated to make a counterfactual inference, it is important to make sure one is estimating and not assuming the result.

Third-Degree Discrimination

Welfare Effects – Bergemann, Brooks, and Morris *AER* (2015)

The effects of allowing price discrimination will depend on the monopolist's information.

Consider our standard example where $v \sim U[0,1]$ and c=0.

- If the monopolist has no information we have standard monopoly pricing: $p^m = \frac{1}{2}, \pi = \frac{1}{4}, CS = \frac{1}{8}.$ (Labeled A in the picture)
- If the monopolist had full information we get 1st degree discrimination: $\mathrm{p(v) = v}, \pi = \frac{1}{2}, CS = 0.$ (Labeled B in the picture)
- If the monopolist has partial information we can get higher CS. For example, if we pool all consumers with $\mathrm{v} \sim U$[0.01, 0.02] with an appropriately smaller mass with $\mathrm{v} \sim U$[0.02, 1] both p=0.01 and p=0.5 are profit maximizing to this segment. Having the monopolist will set p=0.01 in this segment increases CS and leaves profit unaffected

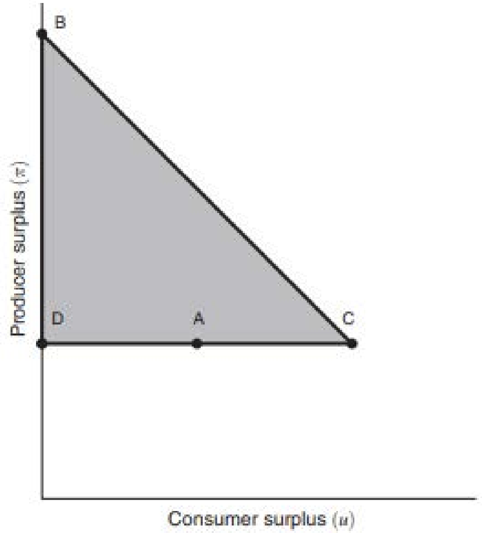


Figure 1. The Surplus Triangle

BBM show the set of profit-CS outcomes that are possible is the right triangle with the 1st degree and CS maximizing points bounding the hypotenuse.

We can robustly say that allowing price discrimination is weakly good for the monopolist – the monopolist can always ignore any information it has.

Whether price discrimination helps or hurts consumers and social welfare depends on the monopolist's information

Second-Degree Discrimination

Two Type Model

Suppose monopolist cannot observe consumer preferences, but sells a good of variable quality/quantity and can prevent resale between consumers.

- Consumers of type θ get utility $v(q, \theta) - T$ if they buy quality/quantity q at total price T and utility 0 if they do not purchase.
- Utilities satisfy $\frac{\partial v}{\partial q} > 0$, $\frac{\partial v}{\partial \theta} > 0$, $\frac{\partial^2 v}{\partial \theta \partial q} > 0$, and $\frac{\partial^2 v}{\partial q^2} < 0$.
- Assume that the cost of producing a quality q good/producing quantity q is *cq*.
- For now assume there are just two types with $\theta_2 > \theta_1$ and write $v_i(q)$ for $v(q, \theta_i)$.

If θ were observable, this would be a first-degree discrimination model. The monopolist would offer consumers of each type θ_i one option: they could purchase quality/quantity q_i^* at price T_i^*, where $\frac{\partial v}{\partial q}\left(q_i^*, \theta_i\right) = c$ and $T_i^* = v(q_i^*, \theta_i)$.

Second-Degree Discrimination

Two Type Model

- Consumers of type θ get utility $v(q,\theta) - T$ if they buy quality/quantity q at total price T and utility 0 if they do not purchase.
- Utilities satisfy $\frac{\partial v}{\partial q} > 0$, $\frac{\partial v}{\partial \theta} > 0$, $\frac{\partial^2 v}{\partial \theta \partial q} < 0$, and $\frac{\partial^2 v}{\partial q^2} < 0$.
- Assume that the cost of producing a quality q good/producing quantity q is *cq*.
- For now assume there are just two types with $\theta_2 > \theta_1$ and write $v_i(q)$ for $v(q,\theta_i)$.

If θ were observable, this would be a first-degree discrimination model. The monopolist would offer consumers of each type θ_i one option: they could purchase quality/quantity q_i^* at price T_i^*, where $\frac{\partial v}{\partial q}\left(q_i^*, \theta_i\right) = c$ and $T_i^* = v(q_i^*, \theta_i)$.

<u>Note</u> that with θ unobserved this does not work. The high type θ_2 consumer gets zero utility from buying q_2^* at price T_2^*, but would get positive utility from instead buying q_1^* at price T_1^* (because the θ_1 types get nonnegative utility and $\frac{\partial v}{\partial \theta} > 0$.)

Second-Degree Discrimination

Two Type Model

With θ unobservable the monopolist will want to allow consumers to choose (q, T) from a menu of offers.

In the two type case it suffices to offer a two item menu $\{(q_1, T_1), (q_2, T_2)\}$. The monopolist's profit-maximization problem is

$$\max_{q_1, T_1, q_2, T_2} T_1 + T_2 - c(q_1 + q_2)$$

st

(IR1) $v_1(q_1) - T_1 \geq 0$

(IR2) $v_2(q_2) - T_2 \geq 0$

(IC1) $v_1(q_1) - T_1 \geq v_1(q_2) - T_2$

(IC2) $v_2(q_2) - T_2 \geq v_2(q_1) - T_1$

Second-Degree Discrimination

Two Type Model

$$\max_{q_1, T_1, q_2, T_2} T_1 + T_2 - c(q_1 + q_2)$$

st

(IR1) $v_1(q_1) - T_1 \geq 0$

(IR2) $v_2(q_2) - T_2 \geq 0$

(IC1) $v_1(q_1) - T_1 \geq v_1(q_2) - T_2$

(IC2) $v_2(q_2) - T_2 \geq v_2(q_1) - T_1$

The first step in solving problems like this is typically to figure out which constraints are and are not binding. Here note:

1. (IC2) + (IR1) $\Longrightarrow$ (IR2).

2. (IC1) seems unlikely to bind. Students rarely consider buying first-class tickets.

Ignoring (IR2) and (IC1) the simplified problem is:

Second-Degree Discrimination

Two Type Model

$$\max_{q_1,T_1,q_2,T_2} T_1 + T_2 - c(q_1 + q_2)$$

st

(IR1) $v_1(q_1) - T_1 \geq 0$

(IC2) $v_2(q_2) - T_2 \geq v_2(q_1) - T_1$

Clearly, one wants to increase T_1 if (IR1) doesn't bind. This implies

$$T_1 = v_1(q_1)$$

One will also want to increase T_2 if (IC2) doesn't bind. This implies

$$T_2 = T_1 + (v_2(q_2) - v_2(q_1)) = v_1(q_1) + v_2(q_2) - v_2(q_1)$$

As in the durable good problem, we can only have the high type pay what the low-types pay plus their incremental value for the higher quality product.

With these T_1 and T_2 both constraints hold (with equality) and can now be ignored.

Second-Degree Discrimination

Two Type Model

$$\max_{q_1,q_2} v_1(q_1) + v_1(q_1) + v_2(q_2) - v_2(q_1) - c(q_1 + q_2)$$

Taking the FOCs for this maximization we find:

$$\frac{\partial v_2}{\partial q}(q_2^{2D}) = c \Longrightarrow q_2^{2D} = q_2^*$$

$$\frac{\partial v_1}{\partial q}(q_1^{2D}) = c + \left(\frac{\partial v_2}{\partial q}(q_1^{2D}) - \frac{\partial v_1}{\partial q}(q_1^{2D})\right) \Longrightarrow q_1^{2D} < q_1^*$$

Observations:

1. The quality/quantity received by the low type is distorted downward from what is efficient. This distortion in product characteristics is a key welfare loss from 2nd degree discrimination.
2. The high type consumers receive surplus ("information rent") while low types do not.
3. Welfare is worse than with 1st degree discrimination. The comparison with offering just a single good is ambiguous.

Second-Degree Discrimination

Continuum of Types

Modern IO theory papers will work in a continuum-of-types model. It's covered pretty well in 14.124 so I'll have Roi cover it in recitation instead of doing it here.

- Continuum of consumers with types θ with density $f(\theta)$ on $[\underline{\theta}, \overline{\theta}]$.
- Type θ consumer's gross utility from quality/quantity x is $v(x,\theta)$ where $v(0,\theta) = v(x,0) = 0, v_x > 0, v_\theta > 0$, $v_{x\theta} > 0$, and $v_{xx} < 0$,

The model gives generalizations of the insights from the two-type model:

- There is again no reason to distort the product provided to the very highest type $\overline{\theta}$.
- When you provide higher quality to some type θ it increases the incremental utility $v(x, \theta + d\theta) - v(x,\theta)$ that you have to provide to slightly higher types. This makes it optimal to provide all types below $\overline{\theta}$ with suboptimal quality.
- Information rents accumulate, so you especially distort quality to the lowest type. Giving them a better product means many others get information rents.
- You would also distort quality a lot if there are few consumers of some type θ. Any gains from providing better quality are limited by the number of consumers of that type. But information rent losses come just from having the offer available.

Two Part Tariffs

Suppose x represents a quantity and the firm can identify and charge consumers who consumer a positive amount, but cannot prevent resale of incremental units across consumers.

The firm is then restricted to "two part tariffs."

- $T(x) = F + px$
- In two-type model:
  - If only type H is served, *p* = *c* and *F* = Type H's consumer surplus.
  - If both types are served, $p \in (c, p^m)$ and *F* = Type L's consumer surplus.
  - Aggregate surplus rises relative to simple linear pricing **if** both types are served (with and without two-part tariffs).

Two Part Tariffs

Proof that $p \in (c, p^m)$:
Let $CS_i(p)$ be type *i* consumer surplus at price *p.*

If selling to both types, firm sets $F = CS_L(p)$

Profit will be $\Pi(p) = \underbrace{(p - c)X(p)}_{\pi(p)} + 2CS_L(p)$

Differentiating, $\Pi'(p) = \underbrace{(p - c)X'(p) + X(p)}_{\pi\prime(p)} + 2CS_L{}'(p)$, so:

- $\Pi'(c) = X(c) + CS_L'(c) = X_H(c)X_L(c) - 2X_L(c) > 0$
- $\Pi'(p^m) = CS_L'(p^m) < 0$

Tying

One theory of tied goods is that by tying a base good to a product that meters its usage the firm can implement a two-part tariff version of second-degree discrimination.

Bundling

- How should a multi-product firm sell its products?
  - One option: sell each good using that good's optimal selling mechanism.

- Often, some form of "bundled" pricing is optimal.
  - With two goods and perfect negative correlation, can get first-best profit (Stigler)
  - With two goods and independent valuations, mixed bundling always better (McAfee, McMillan, Whinston)
  - With *N* products with independent valuations, bundling can get approximate first-best profit (Bakos-Brynjolfsson; Armstrong)

Bundling

Negative Correlation

Stigler provided a very simple example illustrating how negatively correlated values can make bundling profitable.

Suppose types θ are uniformly distributed on [0, 1].

- Type θ consumers get utility θ from watching a superhero movie.
- Type θ consumers get utility $1 - \theta$ from watching a romantic comedy.

If the monopolist prices separately it sets $p^* = \frac{1}{2}$ for each movie and earns profit $^1/_4$ on each.

If the monopolist follows a "pure bundling" strategy of offering only the choice of getting both movies for $P_B = 1$, then it will sell to all consumers and earn a profit of 1.

This feels like price discrimination. You know that consumers don't like the second product as much, so you are willing to sell it at a low price.

Bundling

Independent valuations

Suppose v_1 and v_2 are independent.

Let P_1^* and P_2^* be the optimal separate prices. Mixed bundling is obviously weakly optimal. Set $P_B = P_1^* + P_2^*$.

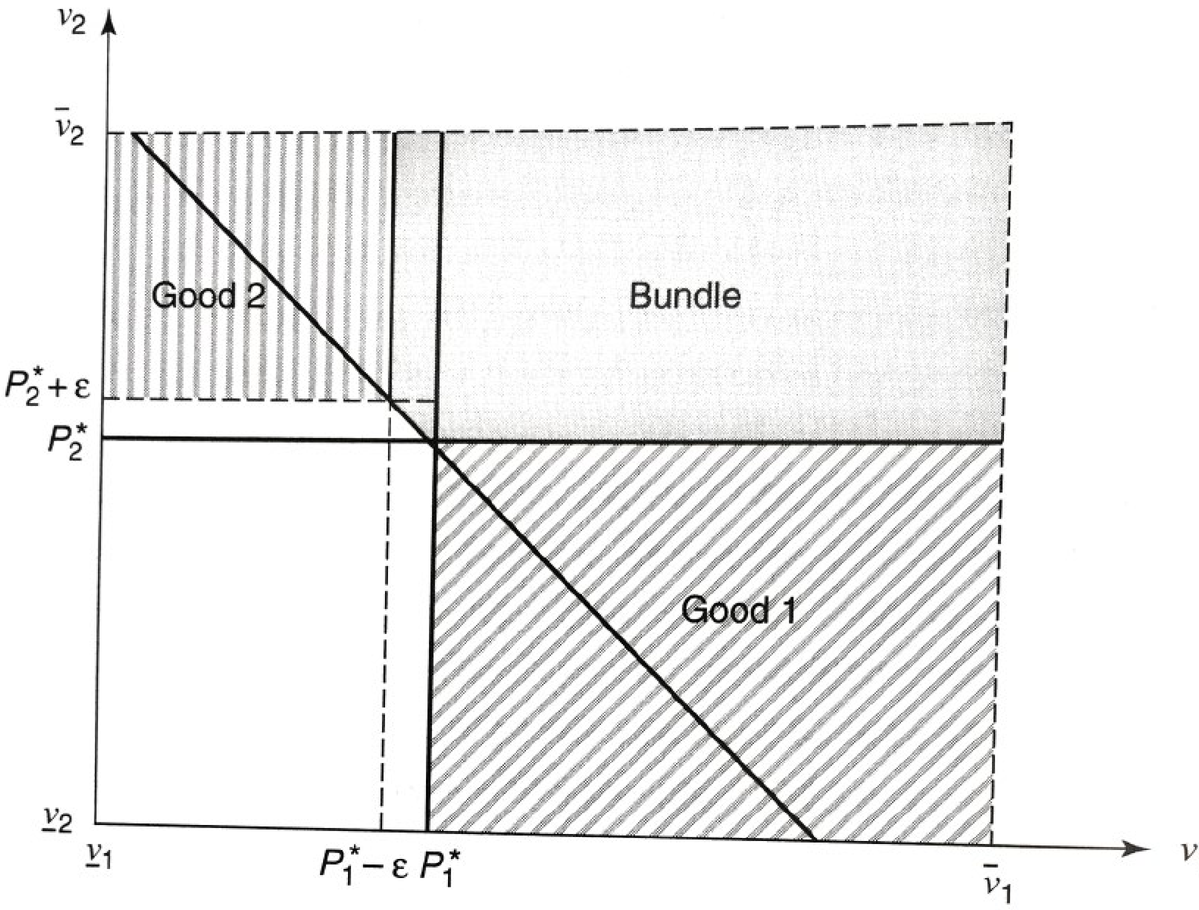
Suppose, we do this and then raise P_2 to $P_2^* + \varepsilon$,

Consumers buy the bundle if:

$$v_1 + v_2 \geq P_B$$
$$v_1 + v_2 - P_B \geq v_2 - P_2$$
$$v_1 + v_2 - P_B \geq v_1 - P_1$$

$\Leftrightarrow$

$$v_1 + v_2 \geq P_1^* + P_2^*$$
$$v_1 \geq (P_1^* + P_2^*) - (P_2^* + \varepsilon) = P_1^* - \varepsilon$$
$$v_2 \geq P_2^*$$



There is a trapezoid of values where consumers who used to buy just Good 2 now buy the bundle. We gain $P_1^* - c_1$ on each of these consumers. The increase in profit is first-order in ε.

For consumers with $v_1 \leq P_1^* - \varepsilon$ we get $\pi_2(P_2^* + \varepsilon)$ instead of $\pi_2(P_2^*)$. This loss is second-order in ε.

So mixed bundling is strictly better than separate pricing.

Bundling

Large number of products

N products with each value drawn iid from distribution *F*. Unit demands.

- *Mean* μ*;* variance σ^2

For simplicity assume no costs

One possible strategy: Sell a pure bundle with price $N\mu(1-\varepsilon)$ for small $\varepsilon > 0$

Profit is at least

$$N\mu(1-\varepsilon)\cdot \Pr\left\{\frac{\sum_i v_i}{N} \geq \mu(1-\varepsilon)\right\}$$

$$= N\mu(1-\varepsilon)\cdot [1-\Pr\{\frac{\sum_i v_i}{N} - \mu \leq -\mu\varepsilon)\}]$$

$$\geq N\mu(1-\varepsilon)\cdot [1-\Pr\{\left|\frac{\sum_i v_i}{N} - \mu\right| \geq \mu\varepsilon)\}])$$

$$\geq N\mu(1-\varepsilon)\cdot \left[1-\frac{\frac{\sigma^2}{N}}{(\mu\varepsilon)^2}\right] \quad \text{(by Chebyshev's Inequality)}$$

$$\rightarrow N\mu(1-\varepsilon) \quad \text{as } N\rightarrow\infty$$

Damaged Goods

Deneckere-and McAfee give several interesting examples of firms selling "damaged goods" that were not cheaper to produce.

- The Intel 486SX microprocessor was the 486DX microprocessor with the math coprocessor disabled.
- The 5 page per minute IBM Laser Printer E was an IBM Laser Printer 10 with an extra chip installed that added a wait state in between each instruction.
- *Stata/IC* is presumably *Stata/SE with an added limit on variables, etc.*

In some models you would not want to do this:

- Suppose $\theta \sim F$ on [0, 1]. Consumers get utility 2θ from a high quality good and θ from a low quality good. Suppose both goods cost c.

Deneckere and McAfee note that damaged goods can be useful with other utility specifications.

On Monday I'll discuss a few empirical papers on price discrimination:

- Scott-Morton, Zettelmeyer, and Silva-Risso
- Shiller
- Dubé and Misra

See you then!

Second-Degree Discrimination

Continuum of Types

Suppose monopolist cannot observe consumer preferences, but sells a good of variable quality/quantity and can prevent resale between consumers.

- Continuum of consumers with types θ with density $f(\theta)$ on $[\underline{\theta}, \overline{\theta}]$.
- Type θ consumer's gross utility from quality/quantity *x* is $v(x, \theta)$ where $v(0,\theta) = v(x,0) = 0, v_x > 0, v_\theta > 0$, $v_{x\theta} > 0$, and $v_{xx} < 0$,

We can think of the monopolist's choice of a sales mechanism in two ways:

- Monopolist chooses a nonlinear price schedule T(x) and consumer chooses x.
- Direct mechanism: Consumers announce θ and are given $x(\theta)$ at total price $T(\theta)$.

Revelation Principle: In many circumstances it is sufficient to focus on direct mechanisms that induce consumers to announce truthfully.

Second-Degree Discrimination

The firm's objective function is $\int_{\underline{\theta}}^{\overline{\theta}}[T(\theta) - c(x(\theta))]f(\theta)d\theta$

Write $U(\theta) \equiv v(x(\theta), \theta) - T(\theta)$ for the consumer's equilibrium utility.

The monopolist's maximization problem is:

$$Max_{x(\cdot),U(\cdot)} \int_{\underline{\theta}}^{\overline{\theta}}[v(x(\theta), \theta) - c(x(\theta)) - U(\theta)]f(\theta)d\theta$$

s.t.

IR: $U(\theta) \geq 0$ for all θ

IC: $\theta \in Argmax_{\hat{\theta}} v(x(\hat{\theta}), \theta) - T(\hat{\theta})$ for all θ,

where $T(\theta) \equiv v(x(\theta), \theta) - U(\theta)$.

Second-Degree Discrimination

Local Incentive Constraints

$$Max_{x(\cdot),U(\cdot)} \int_{\underline{\theta}}^{\overline{\theta}} [v(x(\theta),\theta) - c(x(\theta)) - U(\theta)]f(\theta)d\theta$$

s.t.

IR: $U(\theta) \geq 0$ for all θ

IC: $\theta \in Argmax_{\hat{\theta}} v(x(\hat{\theta}),\theta) - T(\hat{\theta})$ for all θ,

The IC constraint in this problem is equivalent to a combination of:

- Monotonicity: $x(\cdot)$ is non-decreasing
- Local IC: $U(\theta) = U(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} v_\theta(x(s),s)ds$

Intuition: Local deviations being unprofitable for the consumer gives

$$U(\theta + \varepsilon) \geq U(\theta) + v(x(\theta), \theta + \varepsilon) - v(x(\theta), \theta)$$

$$U(\theta) \geq U(\theta + \varepsilon) - (v(x(\theta + \varepsilon), \theta + \varepsilon) - v(x(\theta + \varepsilon), \theta))$$

$$\Rightarrow \frac{v(x(\theta+\varepsilon), \theta+\varepsilon) - v(x(\theta+\varepsilon), \theta)}{\varepsilon} \geq \frac{U(\theta+\varepsilon) - U(\theta)}{\varepsilon} \geq \frac{v(x(\theta), \theta+\varepsilon) - v(x(\theta), \theta)}{\varepsilon} \Rightarrow U'(\theta) = v_\theta(x(\theta),\theta)$$

Example: Suppose $v(x,\theta) = \theta\hat{v}(x)$. Then, $v_\theta(x(s),s) = \hat{v}(x(s))$ and $U(\theta) = U(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} \hat{v}(x(s))ds$.

Second-Degree Discrimination
Local Incentive Constraints

$$Max_{x(\cdot),U(\cdot)} \int_{\underline{\theta}}^{\overline{\theta}} [v(x(\theta),\theta) - c(x(\theta)) - U(\theta)] f(\theta) d\theta$$

s.t.

IR: $U(\theta) \geq 0$ for all θ

IC: $\theta \in Argmax_{\hat{\theta}} v(x(\hat{\theta}),\theta) - T(\hat{\theta})$ for all θ,

The IC constraint in this problem is equivalent to a combination of:

- Monotonicity: $x(\cdot)$ is non-decreasing
- Local IC: $U(\theta) = U(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} v_\theta(x(s),s) ds$

Example:

Suppose $v(x,\theta) = \theta\hat{v}(x)$. Then, $v_\theta(x(s),s) = \hat{v}(x(s))$ and $U(\theta) = U(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} \hat{v}(x(s)) ds$

Intuitive tradeoff:

- The first part of the objective function illustrates an incentive to maximize social surplus.
- The second part of the objective indicates we also want to reduce buyer utility.
- The local IC formula indicates that quality we provide to low types increases the utility of all higher type buyers.

Second-Degree Discrimination

Initially, we ignore the Monotonicity constraint. Only the lowest types IR binds.

$$Max_{x(\cdot),U(\cdot)} \int_{\underline{\theta}}^{\overline{\theta}} \left[v(x(\theta),\theta) - c\big(x(\theta)\big) - U(\theta)\right] f(\theta)d\theta$$

s.t.

IR: $U\big(\underline{\theta}\big) \geq 0$

Local IC: $U(\theta) = U\big(\underline{\theta}\big) + \int_{\underline{\theta}}^{\theta} v_\theta(x(s),s)ds$

The Local IC constraint indicates that $U(\cdot)$ is not a separate choice once the monopolist has chosen $x(\cdot)$. Noting that we will optimally set $U\big(\underline{\theta}\big) = 0$ and substituting we have an unconstrained maximization:

$$Max_{x(\cdot)} \int_{\underline{\theta}}^{\overline{\theta}} \left[v(x(\theta),\theta) - c\big(x(\theta)\big) - \int_{\underline{\theta}}^{\theta} v_\theta(x(s),\theta)ds\right] f(\theta)d\theta$$

A trick for simplifying this is to note that integrating by parts gives:

$$\int_{\underline{\theta}}^{\overline{\theta}} \left[\int_{\underline{\theta}}^{\theta} v_\theta(x(s),\theta)ds\right] f(\theta)d\theta = \left\{\left[\int_{\underline{\theta}}^{\theta} v_\theta(x(s),\theta)ds\right] F(\theta)\right\}_{\underline{\theta}}^{\overline{\theta}} - \int_{\underline{\theta}}^{\overline{\theta}} v_\theta(x(\theta),\theta)F(\theta)d\theta$$

$$= \int_{\underline{\theta}}^{\overline{\theta}} v_\theta(x(\theta),\theta)(1 - F(\theta))d\theta$$

Second-Degree Discrimination

Substituting for the separate integral gives a simpler expression for the problem:

$$Max_{x(\cdot)} \int_{\underline{\theta}}^{\overline{\theta}} \underbrace{\left[\overbrace{v(x(\theta),\theta) - v_\theta(x(\theta),\theta)\frac{(1-F(\theta))}{f(\theta)}}^{\text{"virtual valuation" of type }\theta} - c(x(\theta)) \right]}_{\equiv \boldsymbol{\Lambda}(\boldsymbol{x}(\boldsymbol{\theta}),\boldsymbol{\theta})} f(\theta)d\theta$$

An important feature of this expression there are no interactions across θ's. Hence, we maximize the integral by just maximizing $\Lambda(x(\theta),\theta)$ for each θ.

The FOC for this maximization is

$$v_x(x^*(\theta),\theta) - c'(x^*(\theta)) - v_{x\theta}(x^*(\theta),\theta)\left(\frac{1-F(\theta)}{f(\theta)}\right) = 0 \text{ for all } \theta \qquad (*)$$

Note: $x^*(\theta)$ satisfies Monotonicity provided that $\Lambda(x,\theta)$ has increasing differences in (x,θ).

In the $\mathrm{v}(x,\theta) = \theta\hat{v}(\mathrm{x})$ example, a sufficient condition is $\frac{1-F(\theta)}{f(\theta)}$ decreasing. This is known as the monotone hazard rate condition.

Second-Degree Discrimination

Substituting for the separate integral gives a simpler expression for the problem:

$$Max_{x(\cdot)} \int_{\underline{\theta}}^{\overline{\theta}} \underbrace{\left[\overbrace{v(x(\theta),\theta) - v_\theta(x(\theta),\theta)\frac{(1-F(\theta))}{f(\theta)}}^{\text{"virtual valuation" of type } \theta} - c\big(x(\theta)\big) \right]}_{\equiv \boldsymbol{\Lambda}(\boldsymbol{x}(\boldsymbol{\theta}),\boldsymbol{\theta})} f(\theta)d\theta$$

The FOC for this maximization is

$$v_x(x^*(\theta),\theta) - c'\big(x^*(\theta)\big) - v_{x\theta}(x^*(\theta),\theta)\left(\frac{1-F(\theta)}{f(\theta)}\right) = 0 \text{ for all } \theta \qquad (*)$$

In the $v(x,\theta) = \theta\hat{v}(x)$ example, this is

$$v_x(x^*(\theta),\theta) = c'\big(x^*(\theta)\big) + \hat{v}_x\big(x^*(\theta)\big)\left(\frac{1-F(\theta)}{f(\theta)}\right) = 0$$

This illustrates that units of x are sold at a markup over marginal cost. The markup depends on how many higher types there are and the density of types at θ.

Second-Degree Discrimination

The FOC for this maximization is

$$v_x(x^*(\theta),\theta) - c'\big(x^*(\theta)\big) - v_{x\theta}(x^*(\theta),\theta)\left(\frac{1-F(\theta)}{f(\theta)}\right) = 0 \text{ for all } \theta \qquad (*)$$

Observations:

1. The efficient quality/quantity is sold to the highest type. All other types have quality distorted downward.

2. Can implement with nonlinear tariff $\hat{T}(x) = T\left({x^*}^{-1}(x)\right)$ as both $T(\cdot)$ and $x^*(\cdot)$ are monotone. The consumers FOC gives $\hat{T}'\big(x^*(\theta)\big) = v_x(x^*(\theta),\theta)$.

3. Suppose $c(x) = cx$, $v(x,\theta) = \theta\hat{v}(x)$, and $\underline{\theta} = 0$. Then,

- The Lerner index is decreasing in in θ.

$$\frac{v_x(x^*(\theta),\theta) - c}{v_x(x^*(\theta),\theta)} = \left(\frac{v_{x\theta}(x^*(\theta),\theta)}{v_x(x^*(\theta),\theta)}\right)\left(\frac{1-F(\theta)}{f(\theta)}\right)$$

$v(x,\theta) = \theta\hat{v}(x) \Rightarrow \left(\frac{v_{x\theta}}{v_x}\right) = \frac{1}{\theta}$ is constant and $\left(\frac{1-F(\theta)}{f(\theta)}\right)$ is decreasing with a monotone hazard rate.

- The prices involve quantity discounts, i.e. $T(x)/x$ is decreasing in x.