SYLLABUS FOR INDUSTRIAL ORGANIZATION EC 328

Fall, 2007

Prof. Ken Hendricks
Economics, Room 3.102B,
Email: hendrick@eco.utexas.edu
Phone: (512) 475-8532

Office Hours: 12:00 - 2:00 Thursday

TA: Rao Fu, raofu@mail.utexas.edu

Main Texts

Industrial Organization: Contemporary Theory and Practise, Lynne Pepall, Daniel Richards, and George Norman, Thomson, 3rd Edition, 2005.
The Antitrust Casebook: Milestones in Economic Regulation, William Breit and Kenneth Elzinga, Dryden Press, 3rd Edition, 1996.
"Lectures Notes for an Undergraduate Course in Market Structure and Regulation", Ken Hendricks and Asher Wolinsky

Other Texts:

Industrial Organaization: A Strategic Approach, Jeffrey Church and Roger Ware, McGraw-Hill, 2001
The Antitrust Revolution, by J. Kwoka and L. White, Oxford University Press, 1999.

Course Summary

This is an advanced undergraduate course in industrial organization. We study the behavior of firms and the structure of markets. The first goal is to apply microeconomic theory to understand when and how firms exercise market power and its impact on market efficiency and consumer welfare. The second goal is to apply the theoretical insights to anti-trust cases.

The lectures are based on my lecture notes, which will be posted prior to each class. The focus is on models. The textbook by Pepall, Richards and Norman covers most of the material in more detail. The case studies (and a couple of empirical studies) are used to illustrate how the models can and have been applied. The case studies are drawn primarily from the casebook by William Breit and Kenneth Elzinga although I have included a couple of HBS cases.

Assessment

Assessment will be based on class participation (15%), assignments (25%), a midterm (20%), and a comprehensive final exam (40%).

Course Outline

1. Review (Chapters 2, 4).

   - Review of cost concepts, profit maximization, welfare measurement
   - Antitrust Statutes: Sherman Act, Clayton Act, FTC Act.

2. Monopoly (Chapter 5)

   - Theory/Regulation
   - Durable good
   - Applications: U.S.vs Alcoa (1945), U.S. vs United Shoe Machinery, U.S. vs E.I. duPont de Nemours & Co.

3. Price Discrimination (Chapter 5, 6, 8)

   - First, second, and third degree price discrimination
   - Two part tariffs
   - Bundling and tie-in sales
   - Applications: Utah Pie Co. vs Continental Baking Co. (1967), FTC vs Morton Salt (1948), Eastman Kodak vs Image Technical Services;

4. Oligopoly in Homogenous Good Markets (Chapters 9, 10a)

   - Capacity competition: the Cournot model
   - Price-markups, measures of market concentration.
   - Price competition: the Bertrand model.
   - Application: Genesove, D. and W. Mullin, "Testing Static Oligopoly Models: Conduct and Cost in the Sugar Industry, 1890-1914," RAND Journal, 1998.

5. Oligopoly in Differentiated Good Markets (Chapter 10b)

   - Differentiated products
   - Location models
   - Application: HBS: "Breakfast Cereal Case"

6. Cartels and Price Fixing (Chapter 11, 14, 15)

   - Dynamic games
   - Creating and enforcing cartels
   - Cooperation and repeated play
   - Applications: U.S. vs Addyston Pipe and Steel Co., Interstate Circuit, Inc. vs U.S. (1939), E.I. duPont and Ethyl Corporation v FTC (1984). Porter, R. "A Study of Cartel Stability: The Joint Executive Committee, 1880-1886", Bell Journal of Economics, autumn, 1983.

MIDTERM

7. Horizontal Mergers (Chapter 16)

- Theory
- Merger guidelines for horizontal mergers
- Application: Standard Oil of New Jersey vs U.S., FTC vs Coca-Cola Bottling Company.

8. Vertical Mergers and Restraints (Chapters 17, 18, 19)

- Theory
- Application: Brown Shoe vs U.S; Ford Motor Co. vs U.S.; U.S. vs Colgate (1919), Continental TV vs GTE Sylvania.

9. Entry (Chapter 12, 13)

- Limit pricing
- Predatory pricing
- Raising rival's costs
- Application: Northeastern Telephone Co. vs AT&T (1981); Matsushita Electric vs. Zenith Radio Corp (1986).

10. Advertising (Chapter 20, 21)

- Theories of advertising
- Advertising and competition
- Application: Ackerberg, D. "Empirically Distinguishing Informative and Prestige Effects of Advertising" RAND Journal of Economics, 32 (2), 316-333.

11. Innovation, Research and Development (Chapters 22, 23)

- Patents
- Joint ventures
- Application: OCS leasing

12. Auctions (Chapter 25)

- First price and second price auctions: revenue equivalence.
- Common value auctions and the "winner's curse"
- Application: Hendricks, K. and R. Porter (1988), "An Empirical Study of an Auction with Asymmetric Information", American Economic Review, December 1988, 865-83.

Lecture 1: Preliminaries

January 30, 2007

The purpose of this lecture is to review some basic concepts of microeconomics that we will be using throughout the course.

1 Review of Cost Concepts

In the following discussions the firms' technologies as well as the conditions they face in the input markets will be summarized by cost functions. The cost function $C(y)$ specifies the minimal expenditure needed to produce output level y.

We sometimes distinguish between the *fixed* and the *variable* components of cost.

$$C(y) = F + V(y),\ \ y > 0$$

The fixed cost does not depend upon the level of production, but it should not be confused with sunk cost.

$$C(y) = \begin{cases} F_0, & y = 0 \\ F_0 + F_1 + V(y), & y > 0 \end{cases}$$

Here F_0 is sunk; F_1 is fixed but not sunk. The important behavioral distinction is that short-run supply behavior is independent of sunk costs but does depend upon fixed costs since these can be avoided by not producing. An example of a fixed costs that are sunk are the costs of cell phone towers. The cell phone company has to invest and maintain its network of cell-phone towers. These costs do not depend on call usage. Usage can be zero and the company will still have to pay these costs. An example of fixed costs that are not sunk is the allocation of a plane to fly from one city to another in a city-pair market. The costs of the fuel and crew are independent of how many seats are booked as long at least one seat is booked. However, if no seats are booked, then the airline can allocate the plane to another city-pair market. Thus, the fixed costs are avoidable.

Marginal cost is denoted $MC(y) = C'(y)$. Roughly speaking, it is the cost increment resulting from a unit increase in output. Average cost is denoted $AC(y) = C(y)/y$. Other cost per unit concepts such as $AVC(y) = V(y)/y$ and $AFC(y) = F/y$ look at other cost components per unit of output. Note that $C(y)$ is measured in dollars, while $MC(y)$ and $AC(y)$ are measured in dollars per unit output, that is, MC and AC are measured by the same units as prices.

Example 1 $C(y) = (1 + y^2)$ *for all nonnegative* y.

$$\begin{aligned} F &= 1, V(y) = y^2; \\ MC(y) &= 2y; AC(y) = y + 1/y, AVC(y) = y, AFC(y) = 1/y \end{aligned}$$

Here F *is sunk.* $MinAC$ *is achieved at* $y = 1$.

2 Review of Profit Maximization

In most our discussions, we shall assume that the firms' objective is to maximize profit. Let $P(y)$ denote the inverse demand function faced by the firm. The firm's revenue:

$$R(y) = yP(y);\ MR(y) = R'(y).$$

The firm's profit:

$$\pi(y) = R(y) - C(y)$$

Let y^* be the profit maximizing output. If $y^* > 0$, then $\pi'(y) = 0$. That is,

$$MR(y) = MC(y).$$

For firms in competitive markets, $P(y)$ is constant, so $MR(y) = P$ and $P = MC(y^*)$.

Example 2

$$C(y) = \begin{cases} 1 + y^2, & y > 0 \\ 0 & y = 0 \end{cases}$$

Here $F = 1$ *and is not sunk.* $MC(y) = 2y$. *Suppose* $P(y) = 4$. *Then, if the firm operates, the output that equates marginal cost to price is* $y = 2$. *But should the firm operate? Need to check that firm profits are positive.*

$$\pi(2) = 4 \times 2 - (1 + 4) = 3 > 0.$$

Profits are positive, so $y^* = 2$. *But suppose* $P(y) = 1.8$. *In this case, the output level that equates marginal cost to price is* $y = 0.9$ *and profits are*

$$\pi(0.9) = (1.8)(0.9) - 1 - 0.81 < 0.$$

Here profits are negative so the solution to the profit maximization problem lies on the boundary: $y^* = 0$.

The notion of a supply function refers to the special case in which demand function faced by the firm is a constant price P. The supply function describes the profit maximizing output for each constant P that the firm might face. To keep with the above notation, let us denote the supply function by $y^*(P)$. Thus, $y^*(3)$ gives the profit maximizing output if the firm faces the constant price 3 and so on. Note that $\pi(0)$, and hence the supply function, depends on the specification of the sunk costs.

For the cost function of the example given above, the supply function is

$$y(P) = \begin{cases} 0 & \text{if } P < 2 \\ P/2 & \text{if } P \geq 2 \end{cases}$$

At $P = 2$, the output that equates marginal cost to price is $y = 1$, and profits at this output level are zero. Hence, the firm is indifferent between producing $y = 2$ and shutting down.

If the fixed cost were sunk, then the supply function of a firm already in the industry will be $P/2$ for all $P \geq 0$.

If $P(y)$ is not constant, as in the case of a monopolist, the optimal output of the firm does not depend only on the price and it is therefore meaningless to speak of a supply function in that case.

2.1 A Numerical Example

Why do we care about cost functions? The reason is that the cost function of the firm, combined with the market demand function, completely determine prices and outputs in the short-run equilibrium and prices, outputs, and number of firms in the long-run equilibrium in competitive markets. The short-run is defined by fixed costs (i.e., cell phone towers) that cannot be avoided and a fixed number of firms. In the long-run, all inputs are avoidable (e.g., the cell phone towers have to be replaced eventually) and the number of firms is endogenous.

Consider an industry in which, for each plant, the cost of producing output y is

$$C(y) = \begin{cases} 0, & y = 0 \\ 32 + (.5)y^2, & y > 0 \end{cases}$$

Here the fixed cost is avoidable in the short run so the short run is defined by a fixed number of firms. Market demand for the product, as a function of the price P, is $D(P) = 136 - P$ for $P \leq 136$.

First, we derive the various cost functions for a plant.

$$MC(y) = y; AC(y) = (32/y) + (y/2)$$

Note that the average cost is U-shaped. Differentiating $AC(y)$ with respect to y and setting the derivative equal to zero yields the output level that minimizes average cost. solves (differentiate

$$-32/y^2 + (1/2) = 0 \implies y = 8.$$

The value of average costs at $y = 8$ is

$$AC(8) = 32/8 + 8/2 = 8.$$

Next we derive the firm's supply function. Assume that each firm owns one plant. If the firms shuts down the plant, it gets zero. Hence, in order to operate, it has to earn positive profits. The firm cannot make positive profits if price is less than average costs. Since the lowest average costs can be is 8, this means that the firm will not operate if price is less than 8.At higher prices, its profit-maximizing output equates marginal cost to price. Hence,

$$y(P) = \begin{cases} P \text{ if } P \geq 8 \\ 0 \text{ if } P < 8. \end{cases}.$$

Now suppose there are n firms in this industry. Then the market supply function is simply n times the individual firm supply function.

$$Y(P) = ny(P) = \begin{cases} nP \text{ if } P \geq 8 \\ 0 \text{ if } P < 8. \end{cases}.$$

Equating supply to demand yields the equilibrium price in the short-run.

$$nP = 136 - P \Longrightarrow P^* = \frac{136}{(n+1)}.$$

In particular, suppose that $n = 7$. Then $P^* = 17 = y^*$, and $Y^* = 119$. Note that firm profits are positive since the equilibrium price exceeds $AC(17) = (32/17) + (17/2) < 17$.

Since firms are earning positive profits in the short-run, this industry will attract entrants. Thus, over time, more plants are built, more output is supplied to the market, and price will fall. What will be the long-run equilibrium price, output, and number of firms? Recall that long-run equilibrium is characterized by three conditions:

1. Each firm is making zero economic profits $\Longrightarrow P^{**} = AC(y^{**})$.
2. Each firm is maximizing profits $\Longrightarrow P^{**} = MC(y^{**})$.
3. Supply is equal to demand $\Longrightarrow D(P^{**}) = Y(P^{**})$.

Here the ** denotes the long-run equilibrium value. Applying the first two conditions implies that each firm's long-run output minimizes its average cost :

$$y^{**} = 8.$$

Substituting this value into either condition (1) or (2) yields the long-run price:

$$P^{**} = 8.$$

Finally, we can solve for the number of firms using condition (3):

$$n^{**}8 = 136 - 8 \Longrightarrow n^{**} = 16.$$

3 Review of Welfare Measurement

The maximum sum that consumer k is willing to pay for y units, when faced with the choice of y or none at all, is captured by the area under his inverse demand curve between 0 and y. Let $T_k(y)$ denote this area. Thus, $T_k(y)$ is the value (in terms of money) that consumer k attributes to having y units. Here $T'_k(y)$ is consumer $k's$ marginal willingness to pay at y - it measures consumer $k's$ willingness to pay for another unit when she already has y units. (Actually, T'_k is her inverse demand function.) The (inverse) market demand curve is obtained by summing the consumer (inverse) demand curves. Therefore, the maximum sum that the market is willing to pay for y units is given by the area under the inverse market demand curve from 0 to y. Let $T(y)$ denote the market's willingness to pay for y units. Here $T'(y)$ is the market's marginal willingness to pay at y units and is represented by the inverse market demand function.

Individual k's consumer surplus from purchasing y units at price P is simply the difference between what she is willling to pay and the amount that she actually has to pay,

$$CS_k(y) = T_k(y) - Py.$$

Clearly, the individual should only purchase units up to the point where her willingness to pay for an additional unit is equal to the price, that is,

$$T'_k(y) = P.$$

Let $y_k(P)$ represent the amount that consumer k is willing to buy when the price is P (i.e., the consumer's demand function). Then total consumer surplus at price P can be defined as

$$CS(Y(P)) = \sum_{k=1}^{K} CS_k(y_k(P)) = \sum_{k=1}^{K} T_k(y_k(P)) - PY(P)$$

where

$$Y(P) = \sum_k y_k(P).$$

Producer surplus from selling y units is the difference between what the producer gets from selling the units and the amount that she has to pay for producing them. For firm i, we can define it as

$$PS_i(y) = Py - C_i(y).$$

Clearly, each firm i should produce only up to the point where marginal cost is equal to the price, that is,

$$P = C'_i(y).$$

Let $y_i(P)$ denote the amount that firm i is willing to supply at price P - it is the firm's supply function. The total producer surplus at price P is simply the sum of the firms' producer surpluses.

$$PS(Y(P)) = \sum_{i=1}^{n} PS_i(y_i(P)) = PY(P) - \sum_{i=1}^{n} C_i(y_i(P))$$

where

$$Y(P) = \sum_{i=1}^{n} y_i(P)$$

Note: producer surplus is not always equal to profits. Producer surplus ignores fixed costs that are sunk.

Welfare at price P is simply the sum of consumer and producer surpluses. More precisely,

$$\begin{aligned} W(Y(P)) &= \sum_{k=1}^{K} CS_k(y_k(P)) + \sum_{i=1}^{n} PS_i(y_i(P)) \\ &= \sum_{k=1}^{K} T_k(y_k(P)) - \sum_{i=1}^{n} C_i(y_i(P)). \end{aligned}$$

The amount that consumers pay is the amount that firms get so these terms cancel out. The final expression says that welfare at price P is equal to the sum of the consumers' willingness to pay for the amounts that they demand at price P less the sum of the costs that firms incur from supplying the market at this price.

In the competitive markets, the amount produced by each firm satisfies the condition that price is equal to marginal cost. Since consumes buy amounts where the willingness to pay of the last unit is equal to price, this means that the marginal willingness to pay is equal to marginal cost. In words, the amount that each consumer is willing to pay for the last unit that she purchases is equal to the cost of producing that unit. Hence, there are no further gains from trade. The cost of producing an additional unit exceeds everyone's willingness to pay for that unit. As we shall see in the next lecture, this is not true of monopoly markets.

Example 3 *Consider a market with ten individuals, $k = 1, .., 10$, with discrete demands. Each individual wants only one unit. Individual k values the first unit at r_k (known as the consumer's reservation price) and any additional units at zero. Order the individuals by their valuations from highers to lowest and assume that $r_k = 11 - k$.. Marginal costs of supply is zero. Thus, his demand function is*

$$y_k(P) = \begin{cases} 1 \text{ if } P \leq r_k \\ 0 \text{ if } P > r_k \end{cases}.$$

His total willingness to pay at price P is

$$T_k(y_k(P))) = \begin{cases} r_k \text{ if } P \leq 1 \\ 0 \text{ if } P > r_k \end{cases}.$$

The market demand at price P is

$$Y(P) = \sum_{k=1}^{10} y_k(P)$$

and total willingness to pay is defined similarly. For example, suppose the price $P = 5$. Then market demand, $Y(P) = 6$, since consumers 1 through 6 are willing to buy a unit. Their total willingness to pay, $T(P = 5) = 45$. The amount that they have to pay is 30. Hence, consumer surplus is 15. Producer surplus is revenues minus costs, which is equal to 30. Total welfare is 45. The competitive solution: P=0, CS = 55, PS = 0, W=55.

More generally, individuals will have continuous demands.

Example 4 *Consider a market with ten identical individuals. Each individual k's demand function may be given by*

$$y_k(P) = 1 - P,$$

Her total willingness to pay at price P is

$$T_k(y_k(P)) = \frac{1}{2}(1 - p^2),$$

the area under her inverse demand curve. Similarly, the market demand is given by

$$Y(P) = 10(1 - P)$$

The market's total willingness to pay for Y units is the area under this curve, which is equal to

$$T(Y(P)) = \frac{10}{2}(1 - p^2).$$

For example, suppose $P = 1/2$. Then each consumer's willingness to pay is 3/8, and her consumer surplus is 3/8-1/4=1/8. The market willingness to pay is 30/8 and total consumer surplus is 30/8-5/2=10/8. Producer surplus is 5/2. Thus total welfare is 5/2 + 10/8 = 15/4. The competitive solution: P=0, CS = 5, PS = 0, W = 5.

Conclusion 1 *Total surplus is maximized at the perfectly competitive solution.*

Remark 1 *Partial equilibrium - maximizing total surplus in one market may not be desirable if surplus is not maximized in all other markets.*

Lecture 2: Regulation of a Monopoly

February 5, 2007

1 Monopoly

The objectives of this lecture are to describe: (1) the simple of model of monopoly pricing for a perishable good and rationale for regulation, (2) methods of regulation when the regulator knows market demand and the monopolist's cost function and (3) methods of regulation when the regulator is not fully informed.

Definition 1 *A firm is a monopoly if it is the only supplier of a product in a market.*

One reason why a firm may be a monopolist is because it holds a patent. The patent prevents rival firms from entering the market and competing away the monopoly rents. These rents are the prize that firms receive for creating a new product or developing a new technology. Without these rents, a firm would not be able to recover its research and development costs and, as a result, it would have no incentive to invest in creating new products or technologies. The pharmaceutical industry is an example of an industry which is largely driven by patents. Production costs of a drug are typically measured in cents whereas the R&D costs that went into creating the drug are often measured in the hundreds of millions of dollars. A key issue is the optimal duration of a patent. In the United States, patent duration was 17 years.but it has recently been lengthened to 20 years.

A second reason why a firm may be a monopolist is because the industry is a natural monopoly. Assume the technology of the industry under consideration is described by the cost function $C(Y)$ and that it is accessible for all.

Definition 2 *An industry is Natural Monopoly (NM) with respect to quantity Y, if the least expensive way to produce Y is by one firm. That is, for any quantities*

$$y_1 + .. + y_n = Y,$$

$$C(Y) \leq C(y_1) + ... + C(y_n).$$

An industry is natural monopoly in some range of quantities if it is a NM for every quantity in that range.

Example 1 $C(Y) = F + cy$. *To verify that this is a NM, observe that the variable cost of producing quantity Y is cY regardless of how many firms produce it. But if Y is produced by more than one firm, the total cost is higher, since the fixed cost is duplicated.*

If $C(\cdot)$ displays decreasing AC, the industry is NM for all quantities. But the industry might be NM in the relevant range, even if AC is not strictly decreasing. The range of quantities for which a U-shaped AC is a NM extends to some point to the right of minimum AC.

Common examples of NM industries are those industries which require large fixed cost in networks of some sort: telephone network, electricity, and water supplies.

How will a monopolist behave? More precisely, what price and output will it choose? The key difference between a monopolist and a competitive firm is that the monopolist does not take price as fixed, independent of the quantity that it supplies to the market. The monopolist faces a tradeoff when it tries to increase output: more output can often only be sold by lowering price. To illustrate this point, suppose that the total demand for product y is as follows:

$$y(p) = \begin{cases} 0 \text{ if } \; p > 4 \\ 10 \text{ if } 1 < p \leq 4 \\ 20 \text{ if } 0 \leq p \leq 1 \end{cases}$$

(Graph the demand function.) One interpretation of this demand curve: consumers have unit demands, ten consumers are willing to pay \$4 per unit, and ten consumers are willing to pay only \$1 per unit. Assume that production costs are zero and that the monopolist can set only one price. Then the profit-maximizing price is \$4 since profits at this price are \$40 and profits at \$1, the only other price worth considering, are \$20. But, at this price, there are ten consumers who are willing to pay at least \$1 for the good but are not served even though the cost of producing these units is zero. These are trades that are not made even though both sides could benefit. Of course, the reason why the monopolist does not supply these units is because it would have to lower its price on the infra-marginal units, the units it sells to the ten consumers who are willing to pay \$4.per unit. The gains from trades that are not made at the monopoly solution are known as *deadweight loss*. Here the value of the deadweight loss is \$10.

More generally, if demand is continuous (i.e., smooth), then monopolist solves the following maximization problem:

$$\max_{y} \; P(y)y - C(y)$$

where $P(y)$ is the inverse demand function. Then, differentiating with respect to y and setting the equation equal to zero yields the familiar $MR = MC$ rule:

$$\frac{\partial P}{\partial y} y - P(y) - C'(y) = 0.$$

In elasticity form, it can be expressed as

$$\frac{P - MC}{P} = \frac{1}{\varepsilon(P)}.$$

This is known as the inverse elasticity rule. The key point is that price exceeds marginal cost in monopoly markets. Thus, at the monopoly allocation, consumers' willingness to pay for one more unit exceeds the costs of supplying that unit. Hence, in monopoly markets, the allocation is not efficient: gains from trade are not exhausted.

Example 2 *Consider a market with ten identical individuals. Each individual k's demand function may be given by*

$$y_k(P) = 1 - P,$$

so the market demand is given by

$$Y(P) = 10(1 - P).$$

Soving for the inverse demand yields

$$P = 1 - \frac{Y}{10}.$$

Assume marginal costs are zero. Then the monopolist chooses Y to maximize profits:

$$\max_Y (1 - \frac{Y}{10})Y$$

Differentiating profits and setting the derivative equal to zero

$$1 - \frac{2}{10}Y = 0 \Longrightarrow Y^M = 5, P^M = \frac{1}{2}.$$

Hence, the monopoly solution is $CS = 5/4, PS = 5/2, W = 15/4$. Deadweight loss is $5 - 15/4 = 5/4$. The competitive solution is $P = 0, CS = 5, PS = 0, W = 5$, and maximizes welfare.

This conclusion contrasts sharply with the outcome of competitive markets. In these markets, the amount produced by each firm satisfies the condition that price is equal to marginal cost. Thus, if firm i offered to pay firm j to produce some of its output, it could not offer a deal that firm j would be willing to accept. Consumers buy amounts such that their willingness to pay for the last unit purchased is equal to price. Hence, consumers who have purchased output cannot gain from reselling to each other or to consumers with lower willingness to pay. Finally, since the amount that each consumer is willing to pay for the last unit purchased is equal to the cost of producing that unit, there are no further gains from trade between firms and consumers.

2 Regulation

The basic motivation for regulating a monopoly is the existence of deadweight loss. The purpose is to eliminate or reduce this loss. The possible courses of action are (1) encourage competition by breaking up the monopoly or (2) Alter the firm behavior by public ownership or by dictating to it some of its actions. An example of the latter is maximum price regulation.

The promotion of competition requires regulation of industry structure - preventing monopolization; the other approach requires supervision of the firm's production and pricing decisions. The promotion of competition may seem more attractive since it may require less information and less activity on the part of the regulator. This is particularly true if the source of the original monopoly was entry barrier such as licensing that can be removed once and for all. However, when the industry is a natural monopoly, the promotion of competition may be undesirable. There are two difficulties. First, production is most efficient with one firm. Second,.even if more firms are let into the industry, the result will not be perfect competition. For these reasons, the presence of NM may call for direct regulation.

Methods of regulation:
1. Welfare maximization - MC pricing, with subsidy if necessary (first best). (Illustrate graphically.)
2. Welfare maximization subject to a zero profitability constraint - AC pricing (second best). (Illustrate graphically.)

In theory, first best regulation does not create deadweight loss. But, since the required subsidy has to be raised somewhat, it must cause distortions elsewhere in the economy.

2.1 Regulation Under Incomplete Information

The methods of direct regulation described above assume that the regulator has substantial amount of information about cost and demand. Often this is not the case. Such information may be unavailable to the regulator or costly to obtain. In such a situation, the regulator would like to impose a scheme that gives the monopoly incentives to behave properly, and that does not require as much information to implement.

2.1.1 The Loeb-Magat Proposal

Suppose that the monopoly knows all the cost and demand information, while the regulator knows only the demand curve. Consider the following scheme. The monopoly is free to choose any price p it wishes. The regulator will pay the monopoly subsidy $S(p) =$ consumer surplus at p. Since the regulator knows the demand curve, this scheme can be implemented.

Under the scheme, the monopoly's revenue is

$$S(p) + px(p).$$

Observe that this is equal to the total willingness to pay $T(x(p))$. Thus, the monopoly solves the problem

$$\max_p \pi(p) = T(x(p)) - C(x(p)).$$

In other words, the monopoly maximizes welfare.

One problem with this method is that it requires a large subsidy, which awards all the social surplus to the monopoly. However, some of this can be collected back through a franchise fee. Another problem is that this method still requires the regulator to have substantial information.

2.1.2 Cost Plus Regulation

Suppose that the regulator can observe total output, total cost, and total revenue, but not the demand or the cost function. Consider the effect of regulation which restricts the monopoly profit to $\$s$ per unit of output. That is, if the monopoly sells y units it is allowed to keep only $\$sy$ as profit after paying out its costs. Under this regulation, the monopoly's problem becomes:

$$Maximize \;\; y \;\; subject.to. \; sy \leq yP(y) - C(y)$$

It is easy to see that the solution to this problem is the largest y such that

$$P(y) - AC(y) = s.$$

Notice that if s is sufficiently small, the monopoly will choose to produce more than an unregulated monopoly.

Example 3 *Costs:* $C(y) = 4y$; *Demand:* $P(y) = 20 - y$. *Monopoly output is* $y_M = 8$ *and profit is 64. For* $s = 1$*, the regulated monopolist's output is 15.*

This method has two important drawbacks. First, it may induce inefficient excessive production, as illustrated by the example. Second, it may induce cost padding. Suppose that, in the above example, the monopoly can claim additional \$2 per unit as costs which the owners can appropriate in the form of money, reduced effort, or perks (travel, entertainment, etc.). Then under the $s = 1$ regulation the output will be 13. The official profit is just \$13 but there are an extra \$26 in cost padding.

2.1.3 Rate of Return Regulation

A wide spread form of regulation is called rate of return regulation under which the restriction on the profit that the monopoly is allowed to keep is tied to capital stock rather than to the output. Suppose that the production technology is: $y = F(K, L)$ where K denotes the value of the capital stock employed by the firm and L denotes the firm's labor input. Let r denote the cost of capital and w denote the wage. Thus, the profit of an unregulated firm that uses input quantities K and L is:

$$p(F(K, L))F(K, L) - rK - wL.$$

Let s denote the regulator's mandated restriction on the rate of return on capital. Thus, the monopoly is not allowed to keep more than sK as profit. Under this regulation, the monopoly's problem becomes: choose K and L to maximize

$$sK \quad \textit{subject to } (i)\ y = F(K, L) \textit{ and} (ii)\ sK \leq yP(y) - rK - wL.$$

Example 4 *A monopoly produces with a technology that requires one unit of capital and 3 units of labor for each unit of output produced per time period. The prices of capital and labor are $r = w = \$1$ per time period so the cost function is $C(y) = 4y$. The demand function is given by $P = 20 - y$. The monopoly solution is $y_M = 8$ and the monopoly profit is* 64. *Now suppose that the regulator can observe the amount of capital employed and allows the firm to keep as profit only* \$0.2 *for each \$1 of capital. Thus, if the monopoly produces the original quantity of* 8, *it will employ* 8 *units of capital and can keep only* \$1.6. *To maximize profit that it can take under this restriction, the monopoly has to expand its capital stock and hence production. Here, maximum profit will be achieved at* $p = \$4.2$ *and* $y = 15.8$, *yielding a profit of* \$3.16

Note that, in this example, the rate of return regulation works like the method that ties profits to output. To increase its profit, the monopoly has to increase its capital stock and hence production. Thus, by tying profit to capital stock, the firm is induced to accumulate more capital than required for efficient production. Some of the deadweight loss is eliminated, but some productive inefficiency is created in return.

When the monopoly faces a choice of technology, this method suffers from an additional drawback. Suppose now that there is also an alternative technology with which each unit of output requires 3 units of capital and 2 units of labor. This technology is less efficient since it requires \$5 per unit of output. Under the rate of return restriction to \$0.2 per \$1 of capital, the firm will adopt this technology, since by charging $p = 5.6$ and producing $y = 14.4$, it will keep profit of \$8.64, which is larger than what it makes with the more efficient technology.

3 Regulation of Multiproduct Firms (Optional)

Consider a firm which produces two outputs, x and y. Its costs depends upon the quantities of each output and is given by $C(x, y)$. Marginal cost of x is denoted by $MC_x(x, y)$. Similarly, the maginal cost of y is given by $MC_y(x, y)$.

Example 5 $C(x, y) = F + cx + dy$. $MC_x(x, y) = c$ *and* $MC_y(x, y) = d$.

Note that while it is possible to define a notion of average cost, such definition may not be very meaningful.

Definition 3 *A multiproduct industry is Natural Monopoly with respect to quantities X, Y, if production of X and Y by one firm is less expensive than their production by any combination of firms. The industry is a NM with respect to some ranges of X and Y if the above is true for any pair of quantities in the range.*

Definition 4 *A NM displays Economies of Scope. These are present if* $C(X,Y) < C(X,0) + C(0,Y)$.

Notice that the technology in the above example is a NM for all $x \geq 0, y \geq 0$.

In discussing welfare we shall assume, for simplicity, that the demands for the two products are independent. This imples that the total willingness to pay for the combination (x,y) can be measured by the sum $T_x(x)+T_y(y)$, i.e. the sum of the area under the demand curve of X and the area under the demand curve for Y. Note that, if the demands were not independent, the measurement of the willingness to pay would not be simply the sum of these areas.

Regulation:

1. First best: MC pricing, $P_i = MC_i(x,y), i = x,y$, coupled with a subsidy if necessary.
2. Second best: pricing subject to nonnegative profit constraint.

Note that in the multiproduct case there are often many possible combinations of outputs in both markets, which cover costs. The problem is to choose the best such combination.

Definition 5 *Ramsey prices are prices* p_R *and* q_R *which maximize welfare subject to the contraint that profit is nonnegative. That is,* p_R *and* q_R *are the prices which maximize*

$$T_x(x(p)) + T_y(y(q)) - C(x(p), y(q)) \quad s.t. \;\; px(p) + qy(q) \geq C(x(p), y(q)).$$

Example 6 *Costs:*

$$C(x,y) = 5 + x + y; C(x,0) = 4 + x; C(0,y) = 2 + y.$$

Demands:

$$p = 4 - x; \quad y(q) = \begin{cases} 3 \; for \; q \leq 4 \\ 0 \; for \; q > 4 \end{cases}$$

Solution: $p_R = 1$ *and* $q_R = 8/3$.

Note that in this case Ramsy prices achieve the first best. Since the dmeand for y is inelastic, prices below 4 involve no distortion. Actually, Ramsy price here are not unique - any q between 8/3 and 4 will do.

Now modify the above example as follows:

$$y(q) = \begin{cases} 3 \; if \; q \leq 2 \\ 0 \; if \; q > 2 \end{cases}.$$

Ramsey prices for this example are $p_R = 2$ and $q_R = 2$. Note that now Ramsy prices involve deadweight loss in the market for x.

In both of the above cases Ramsey prices put as much as possible of the burden on product y, the demand for which is inelastic. This is a general principle: the deadweight loss is minimized when products with inelastic demands are "taxed" more heavily. Finally, note that finding Ramsy prices is the same as finding an optimal commodity tax.

3.0.4 Cross Subsidy

The second best prices may be such that the production of one product by a stand alone firm is cheaper than this product's regulated price.

Definition 6 *Assume that the regulated monopoly sells quantities x and y at prices p and q such that $px + qy = C(x, y)$. If $px > C(x, 0)$ or $qy > C(0, y)$, then the regulated prices p and q involve cross subsidy.*

Example 7 *Costs:*

$$C(x, y) = 5 + x + y; C(x, 0) = 4 + x; C(0, y) = 2 + y$$

Demands:

$$p = 4 - x; \quad y(q) = \begin{cases} 3 \ for \ q \leq 4 \\ 0 \ for \ q > 4 \end{cases}$$

Ramsey prices are $p_R = 1$ and $q_R = 8/3$. Since $y(8/3)(8/3) = 8 > 2 + 3 = C(0, 3)$, at the Ramsy prices, product y subsidizes x.

Cross subsidy poses problems to the regulated monopoly:
1. An entrant may enter the subsidizing market, although efficiency dictates joint production.
2. Consumers of the subsidizing good may object, since they essentially are paying a "tax" which subsidizes consumer of the other product.
Examples: long distance vs. local telephone service; urban vs rural transportation; large vs. small cities in airline networks.

Conclusion 1 *The possibility of cross subsidy may add an additional constraint to the regulator's problem. Sometimes the regulator may want to impose the best prices possible which do not involve cross subsidy.*

In the above example, the requirement that prices be subsidy free implies that $q \leq 5/3$. Therefore, the prices which maximize welfare subject to both the nonnegative profit constraint and the subsidy free constraint are $p = 2$ and $q = 5/3$. Note that these prices involve deadweight loss, since $p > 1$,while the Ramsey prices in this case do not.

The problem in general is that such prices do not always exist.

Example 8 *Three products, x, y, and z. The demand for each product is perfectly inelastic. Costs:*

$$\begin{aligned} C(x, 0, 0) &= C(0, y, 0) = C(0, 0, z) = 300; \\ C(x, y, 0) &= C(x, 0, z) = C(0, y, z) = 400; \\ C(x, y, z) &= 650. \end{aligned}$$

This is a NM!. Subsidy free prices p,q, and r have to satisfy:

$$\begin{aligned} px &\leq 300; \quad qy \leq 300; \quad rz \leq 400; \\ px + qy &\leq 400; \; px + rz \leq 400; \; qy + rz \leq 400 \end{aligned}$$

But the above inequalities imply

$$2(px + qy + rz) \leq 1200,$$

so that

$$(px + qy + rz) \leq 600.$$

But cost coverage implies that

$$(px + qy + rz) \geq 650.$$

Therefore, there are no subsidy free prices in this case.

Conclusion 2 *It is possible to have a NM which cannot "defend" its markets against entry in some of them.*

This lends some theoretical justification to a policy conclusion that, to enjoy the benefits of NM, sometimes the regulator has to regulate entry to some of the monopoly's markets. Of course, one has to distinguish between a theoretical justification and a practical policy. There are many drawbacks and substantial scope for abuse in such policy.
Examples: entry restriction into long distance telephone service; entry restriction into city-pair markets in air travel.

Lecture 3: Durable Good Monopoly

February 5, 2007

1 DURABLE GOOD MONOPOLY

A durable consumer good is a good which provides a stream of consumption services over time. The issue of durability introduces two complications for the monopolist interested in profit maximization. The first is that consumers may be able to wait and buy tomorrow. That is, the monopolist today has to compete against itself in the future. The second is that the monopolist creates competition for itself in the future by selling today. The existence of a second-hand market suggests that the market power of the monopolist supplier in the future is determined in part by the production of the monopolist today.

1.1 Pacman Versus Coase

The issue is whether the monopolist can engage in intertemporal price discrimination. That is, can the monopolist use time to sort consumers by their valuations and make them pay their reservation values? The answer (roughly) is that it depends upon which side of the market can credibly threaten to outwait the other side. This in turn depends upon how quickly the monopolist can change prices and the distribution of consumer valuations.

Suppose the monopolist has two units of the durable good to sell. The good lasts forever, storage costs are zero. There are two consumers, each demanding one unit of the good, but with different valuations. One consumer has a reservation value of 10 for the good, and the other has a reservation value of x, where $20 > x > 10$. (These values should be interpreted as the present value of the stream of services provided by a unit of the good.) The monopolist, of course, does not know which consumer has the high valuation. The common discount rate is δ.

Pricing Strategies:
1. *Pacman strategy:* set price in period t equal to the highest reservation value of any consumer that has not yet purchased a unit. This is called the Pacman strategy since it specifies that the monopolist will move down the demand curve selling to consumers sequentially in order of their reservation prices.
2. *Competitive strategy:* sell all of the units immediately at a price equal to the lowest consumer valuation. This is the pooling strategy in which the monopolist is unable to price

discriminate.

1.1.1 Two Periods

The credibility of the Pacman strategy depends critically upon an infinite number of selling periods. To see this, suppose the monopolist has only two periods to sell its supply of the durable good and it attempts to use the Pacman pricing strategy. If the firm has sold a unit in period 1, then it will set $p_2 = 10$ and sell the last unit to the low valuation buyer in period 2. It it has not sold any units in period 1, then the Pacman strategy calls for a price $p_2 = x$, which yields a payoff of x. But, a more profitable strategy at this point, given that the monoplist cannot wait until the following period, is to set $p_2 = 10$, sell both units and get a payoff of 20. Hence, the threat to set $p_2 = x$ if no one buys in period 1 is not credible, and the high valuation consumer should anticipate this fact.

Suppose then that $p_2 = 10$. Then the high valuation buyer can choose to either buy today at p_1 or wait until period 2 and buy at a price of 10. The ability to arbitrage any price differential between the two periods implies that, for the high valuation buyer to purchase in period 1, p_1 must satisfy

$$(x - p_1) \leq \delta(x - 10) \Longrightarrow p_1 \leq x(1 - \delta) + \delta 10.$$

Alternatively, the monopolist could just use the competitive strategy and set $p_1 = 10$ and sell both units in period 1 for a profit of 20. It is easy to see that this strategy is more profitable that trying to engage in intertemporal price discrimination, that is,

$$x(1 - \delta) + \delta 20 < 20.$$

Conclusion 1 *The monopolist is unable to engage in intertemporal price discrimination and is forced to sell both units at the competitive price.*

1.1.2 Infinite Number of Periods

With an infinite number of periods, the monopolist's threat to use the Pacman strategy in every period t is credible. Given this strategy, the high valuation buyer has no incentive to delay beyond period 1 since the monopolist will not reduce its price until a unit is sold at a price of x. Consequently, the Pacman strategy is implementable. However, it may not be an equilibrium.

Suppose the monopolist orginally had N units and sold N-2 using the Pacman strategy. It now has two units left to sell which, according to the Pacman strategy, involves setting price equal to x in the current period and 10 in the following period. What is the best alternative for the monopolist to the Pacman strategy? There is no point lowering the price in the current period unless the monopolist can sell both units. Consequently, the best alternative is the competitive strategy in which it sells both units now at a price of 10.

Comparing payoffs to the two strategies, we obtain the result that the Pacman strategy is more profitable if and only if

$$x + 10\delta > 20 \Longrightarrow \delta > (20 - x)/10.$$

If the monopolist is not too impatient, and valuations between the two buyers are not too large, then the Pacman strategy is preferred. But if the prices cannot be changed too quickly, which means δ is bounded away from 1, then the competitive strategy is preferred when valuations of buyers are quite close (i.e., x approaches 10).

Conclusion 2 *The basic tradeoff is between the speed with which the monopolist can change prices (i.e., committment to period length) and the discreteness in buyer valuations. If buyer valuations are "lumpy" and known, then the monopolist can use the Pacman strategy to price discriminate assuming it is not too impatient (i.e., price changes can be made frequently). However, if buyer valuations are distributed more or less continuously (i.e., demand curve is continuous), then the result is the elimination of market power and competitive pricing. Markets where there is a large number of buyers and small differences in willingness to pay favor Coase-like outcomes, but small markets with large differences in willingness to pay favor Pacman discriminatory outcomes.*

1.2 Strategies to Mitigate Loss of Monopoly Power

Leasing

If the monopolist leases the units each period, then the good is returned to the monopolist at the end of each period which in turn affects its optimal pricing decision in the following period. Retaining ownership gives the monopolist a way to commit not to increase supply and lower prices in the future.

Examples of firms that have used this option: United shoe Machinery Corporation prior to 1953 has a lease only policy. The U.S. government brought a civil suit against United Shoe seeking an injunction against its lease policies and it was forced to eliminate the lease only policy. IBM had a similar policy for its main frame computers prior to 1956; the same is true of Xerox for photocopiers prior to 1975. One problem with leasing contracts is that the renters may abuse the equipment, so it is not always possible for a durable good monopolist to practise leasing. In this regard, it is interesting to note that Xerox required the lessees to sign maintenance contracts which stipulated that only Xerox personnel were permitted to service the machines.

Reputation

A monopolist may be able to establish a reputation for not lowering prices in subsequent periods. Two prominant examples are De Beers and Disney.

De Beers has a policy of never lowering prices and indeed, its price increases for the first fifty years or so of its existence exceeded the rate of inflation. In order to maintain its reuptuation for price support of uncut gems, de Beers manages the price of gems by regulating the supply of gems and withholding excess supplies - diamonds that are then stockpiled. It regulates supply through its exclusive supply contracts with cartel members and by buying diamonds on the open market from independent sources and leaked diamonds from cheating cartel members. At times, this policy has been expensive. The USSR dumped large volumes of diamonds in 1984 to finance its war in Afghanistan, and again in 1995.

With the videocassetter recorder revolution, Disney has a remarkable opportunity to release its classic animated films on videocassette. But how to get consumers to buy today at high prices? Disney's strategy was to make the videos available for a limited time only and accompany the release with an advertising campaign in which Disney claims that the video will only be available for a limited time period and then never again. For example, *Bambi* was available for only two months in the spring of 1997. The "buy now because otherwise you cannot get it" strategy appears to have had some success. The film, *Snow White and the Seven Dwarfs*, sold 20 million copies in the first three months of a six month release.

Contractual Arrangements

Monopolist can sign contracts with buyers promising to buy back the good at the original selling price or it adopt best-price clauses which commit the monopolist to retroactively reduce the price of customers who purchase today in the event prices are decreased in the future. General Electric and Westinghouse used best-price clauses to stabilize prices in the market for electric turbogenerators in 1963.

Planned Obsolescence

Idea here is to increase future demand by decreasing the stock available from past sales. Examples: textbooks?

1.3 Recycling

This has been a much studied problem. Not too surprisingly, the theoretical research finds that the constraint which a competitive recycling sector exerts on the primary product monopolist depends on how efficient the recycling sector is at recovering scrap and turning it into a secondary product. If there is no shrinkage or depreciation, then eventually the market power of the monopolist is eliminated.

The classic case is Alcoa who was a monopolist in the market for aluminium in the US in the 1930s. Alcoa's share of the primary or virgin aluminum was 90%. However, its share dropped to about 65% if secondary production is included in the same market. The judge ruled that the appropriate market share was 90% and on that basis found that Alcoa was a monopolist. He argued that Alcoa controlled the secondary market via its control of the primary market. Was he right? The combined rate of shrinkage and depreciation in the production of aluminum is on the order of 25%. The theory suggests that this is enough for Alcoa to have market power. Subsequent empirical studies provide evidence that this was the case.

Lecture on Price Discrimination

February 12, 2007

1 PRICE DISCRIMINATION

The objectives of this lecture are (i) describe various kinds of pricing schemes that firms employ to reduce the deadweight loss and (ii) to consider the welfare implications of these schemes. Price discrimination takes place when a firm sells different units of the same product at different prices.

Definition 1 *First degree price discrimination is when the firm gets for each unit the maximum willingness to pay for that unit.*

This form of discrimination is also known as perfect price discrimination. It is primarily a theoretical benchmark concept and it is hard or impossible to think of practical examples where this form was applied.

Definition 2 *Third degree price discrimination is when the firm sells to different and distinguishable groups of consumers at different prices.*

Among the obvious examples are discounts to members of identifiable groups such as senior citizens or students. The discrimination is typically based on age or location of the consumer.

Definition 3 *Second degree price discrimination is when the firm sells different units at different prices but it cannot exclude consumers from any offer that it makes.*

The pricing and product decisions of the firm induce consumers to sort themselves volountarily into groups that pay different prices. Quantity discounts are an example of this method: consumers who buy larger quantities pay less per unit, but nobody is excluded from buying large quantities. Discounted airfares for people who buy their tickets sufficiently early is another example.

1.1 First-degree Price Discrimination

At first glance, it may appear that first-price discrimination is impossible. However, if consumers are identical, and the monopolist knows their demand curves, then it can extract all of the consumer surplus from consumers using a non-linear tariff. A good is sold at non-linear prices when the cost of purchase is not proportional to quantity, i.e., cost is not equal to a constant common price times the quantity. One simple form of non-linear prices is *two-part tariff* (TPT). To purchase quanitity y under a TPT arrangement, a buyer has to pay the sum $E + PQ$ (i.e., a fixed fee, E, plus a per unit price of P). Telephone and electricity supply services are often priced in this manner.

Consider a simple version of the example given in the textbook (Chapter 6), page 115). The example involves a jazz club, which offers music and drinks. The inverse demand of the typical patron for the club's services is given by

$$P = V - Q$$

where P is the price of drinks and Q is the quantity of drinks. The jazz club owner incurs a cost of c dollars per drink served plus a fixed cost F of operating the club each night.

Now suppose the jazz club is a traditional monopolist. Then it chooses Q to maximize

$$\pi(Q) = N[(V - Q)Q - cQ] - F$$

where N is the number of consumers. The first-order condition for this optimization problem is

$$V - 2Q - c = 0$$

(i.e., $MR(Q) = MC$). Here the solution is

$$Q^M = \frac{V - c}{2}, P^M = \frac{V + c}{2}, \pi^M = \frac{1}{4}(V - c)^2$$

where π^M represents profit per patron. To obtain total club profits, multiply π^M by N and deduct fixed costs.

Example 1 *Suppose* $V = 16$, $c = 4$. *Then* $P = 10$, $Q = 6$, *and* $\pi^M = 36$.

The club owner's manager has taken a course in economics and tells the owner that she should adopt a different pricing scheme: in addition to charging a price for each drink, the club should also levy a cover charge, E. What are the profit-maximizing values for E and P? At any P, the amount that a consumer is willing to pay to enter the club is given by her consumer surplus.

$$E(P) = CS(P) = \frac{1}{2}(V - P)^2.$$

Thus, the club's maximization problem consists of choosing P to

$$\max_P \pi(P) = (P - c)(V - P) + \frac{1}{2}(V - P)^2.$$

The first-order condition for this problem is

$$V + c - 2P - (V - P) = 0 \Longrightarrow P^* = c, E^* = \frac{1}{2}(V - c)^2.$$

The monopolist prices drinks at marginal cost and extracts the consumer's willingness to pay through the cover charge. There is no deadweight loss so the outcome is efficient!

Example 2 *Suppose* $V = 16$, $c = 4$. *Then* $P = 4$, $Q = 12$, $E = 72$ *and* $\pi^M = 72$.

1.2 Third Degree Price Discrimination

What is the impact of price discrimination on profits, consumer surplus, and welfare? We will study this question using demand functions that are step functions.

Suppose that the total demand for product y is as follows:

$$y(p) = \begin{cases} 0 \text{ if } \ p > 4 \\ 10 \text{ if } 3 < p \leq 4 \\ 50 \text{ if } 0 \leq p \leq 3 \end{cases}$$

(Graph the demand function.)

Assume that production costs are zero. The monopoly's price is 3 and profit is 150. Clearly, if the monopoly can divide this market into two, such that in one market the demand is inelastic at 10 up to a price of 4, and in the other market the demand is inelastic at 40 up to a price of 3, then it can increase its profit by charging in these markets 4 and 3 respectively.

In the above case, price discrimination does not affect the total quantity and surplus. It only transfers the consumers' surplus to the firm. It turns out that the effect of price discrimination on total quantity and surplus is not always in the same direction. Depending on the circumstances, it may increase or decrease anyone of these magnitudes, relative to the corresponding regular monopoly (which serves all markets at the same price). The following examples demonstrate the different possible effects of discrimination on quantity and surplus. (In all of them the production costs are assumed zero.)

Example 3 *Discrimination results in lower quantity and surplus*

$$y_1(p_1) = \begin{cases} 0 \ \ if \ \ p_1 > 4 \\ 10 \ \ if \ \ 3 < p_1 \leq 4 \\ 12 \ \ if \ \ 0 \leq p_1 \leq 3 \end{cases} ; \qquad y_2(p_2) = \begin{cases} 0 \ \ if \ \ p_2 > 3 \\ 48 \ \ if \ \ 0 \leq p_2 \leq 3 \end{cases}$$

In the absense of discrimination the monoply price is $p = p_1 = p_2 = 3$. *Price discrimination between these two markets will result in* $p_1 = 4, p_2 = 3$.

Example 4 *Discrimination results in higher quantity and surplus.*

$$y_1(p_1) = \begin{cases} 0 \ \ if \ \ p_1 > 4 \\ 100 \ \ if \ \ 0 < p_1 \leq 4 \end{cases} ; \qquad y_2(p_2) = \begin{cases} 0 \ \ if \ \ p_2 > 3 \\ 20 \ \ if \ \ 0 \leq p_2 \leq 3 \end{cases}$$

In the absense of discrimination the monoply price is $p = p_1 = p_2 = 4$. Price discriminatoin between these two markets will result in $p_1 = 4, p_2 = 3$.

Example 5 *Discrimination results in higher quantity but lower surplus.*

$$y_1(p_1) = \begin{cases} 0 \ \ \textit{if} \ \ p_1 > 4 \\ 10 \ \ \textit{if} \ \ 2 < p_1 \leq 4 \\ 19 \ \ \textit{if} \ \ 0 \leq p_1 \leq 2 \end{cases} ; \qquad y_2(p_2) = \begin{cases} 0 \ \ \textit{if} \ \ p_2 > 2 \\ 2 \ \ \textit{if} \ \ .02 < p_2 \leq 2 \\ 201 \ \ \textit{if} \ \ 0 \leq p_2 \leq .02 \end{cases}$$

In the absense of discrimination the monoply price is $p = p_1 = p_2 = 2$,the quantity is 21 and the surplus is 62. Under discriminatoin, $p_1 = 4, p_2 = .02$,total quantity is 211, but total surplus is 47.98.

This last example highlights the fact that price discrimination may result in inefficient allocations in the quanitity produced. That is, some units are allocated to people who do not value them the most while others who value them more remain without the product.

So far it has been simply assumed that the monopoly can discriminate, but it should be emphasized that this may not always be possible. For price discrimination to be exercised, it should be that: (i) the seller can divide its market into separate sub-markets with different patterns of demand; and (ii) these markets can indeed be separated in the sense that *resale* between them is not a viable possibility.

The latter condition may hold for a service such as a lawyer's advice, or in the case of a product which is difficult to store or transport between the markets. Another example is airline tickets which are not permitted to be resold.

The former condition can be satisfied when the markets are geographically separated cities with sufficiently different populations and hence demands. It can also be satisfied if a clearly identified group of customers, say students who can be identified by a university I.D., have a different pattern of demand from the rest of the population.

1.2.1 Continuous Demand

When demand curves are downward sloping, then the optimal rule for a price-discriminating monopolist is to set quantities in each market so that marginal revenue in each market is equated to marginal cost. We use the Harry Potter example from the textbook to illustrate the basic logic underlying this rule.

The commodity is a hardcover copy of *Harry Potter and the Order of Phoenix.* The inverse demand for the book in the United States is given by

$$P_U = 36 - 4Q_U;$$

in Europe, the inverse demand is given by

$$P_E = 24 - 4Q_E.$$

Here quantity is measured in millions of books. Solving for quantity as function of price yields the demand functions

$$Q_U = 9 - \frac{P_U}{4}$$

for $P_U \leq 36$ and

$$Q_E = 6 - \frac{P_E}{4}$$

for $P_E \leq 24$. Marginal costs of production are $4 per book. We will assume that profits exceed fixed costs.

Let us first determine what the monopolist would charge if it not able to discriminate between American and European consumers (i.e., $P_E = P_U$). As in the analysis above, we first add the demand curves.

$$Q(P) = Q_E(P) + Q_U(P) = \begin{cases} 0 \text{ if } P > 36 \\ 9 - \frac{P}{4} \text{ if } 36 \leq P \leq 24 \\ 15 - \frac{P}{2} \text{ if } P < 24. \end{cases}.$$

For P higher than 36, demand in each market is zero; for P between 24 and 36, demand in US is positive but zero in Europe, and for P below 24, demand is positive in each market. The monopolist's optimization problem is to choose P to

$$\max_P \pi(P) = PQ(P) - 4Q(P)$$

or, equivalently, choose Q to

$$\max_Q \pi(Q) = (30 - 2Q)Q - 4Q.$$

Here we have assumed that the optimal price is less than 30. Differentiating π with respect to Q and setting the derivative equal to zero yields

$$30 - 4Q - 4 = 0 \Longrightarrow Q^M = 6.5, P^M = 17, \pi^M = 84.5$$

Now suppose that the monopolist can discriminate between the two markets. Its optimization problem consists of choosing Q_U and Q_E to maximize total profits

$$\max_{Q_U,Q_E} \pi(Q_U, Q_E) = (36 - 4Q_U)Q_U - (24 - 4Q_E)Q_E - 4(Q_E + Q_U).$$

Differentiating with respect to each variable yields the following two first-order conditions:

$$\begin{aligned} 36 - 8Q_U - 4 &= 0 \\ 24 - 8Q_E - 4 &= 0. \end{aligned}$$

Thus, in each market, the monopolist chooses quantity to equate marginal revenue to marginal cost. The solution is

$$\begin{aligned} Q_U^* &= 4, P_U^* = 20, \pi_U^* = 64 \\ Q_E^* &= 2.5, P_E^* = 14, \pi_E^* = 25. \end{aligned}$$

Price goes up in the American (high demand) market and down in the European (low demand) market. Total profits are 89 million, which is 4.5 million higher than the monoplist's profits when it cannot price discriminate.

1.2.2 Two-Part Tariffs

Consider once again the example of the jazz club. If the owner can identify different groups of consumers, say by their age, then he will want to charge different cover charges to different classes of customers but still price drinks at marginal cost! The idea is that marginal cost pricing maximizes the consumer surplus of each type of consumer, which can be obtained by the monopolist by setting cover charges equal to the consumer surplus for each type of consumer. For example, suppose the inverse demand for drinks by old consumers is

$$P_o = V_o - Q_o$$

and by young consumers is

$$P_y = V_y - Q_y$$

where $V_o > V_y$. If the monopolist uses price to discriminate between consumers, then it solves

$$\max_{Q_o, Q_y} (V_o - Q_o - c)Q_o + (V_y - Q_y - c)Q_y.$$

Equating marginal revenue to marginal cost in each market and solving for quantities yields

$$Q_o^M = \frac{V_o - c}{2}, Q_y^M = \frac{V_y - c}{2}.$$

Equilibrium prices are

$$P_o^M = \frac{V_o + c}{2}, P_y^M = \frac{V_o + c}{2}, \pi^M = \frac{1}{4}[(V_o - c)^2 + (V_y - c)^2].$$

However, if the monopolist instead sets $P^* = c$, it can charge

$$E_o^* = \frac{1}{2}(V_o - c)^2, E_y^* = \frac{1}{2}(V_y - c)^2,$$

and earn twice as much profit. Thus, the club owner should use the cover charge, and not the price of drinks, to discriminate among different classes of consumers. Similarly, if the club owner books a more popular band, raising the willingness to pay of all consumers, then it should increase the cover charge but not the price of drinks.

Example 6 *Suppose* $V_o = 16$, $V_y = 12$, *and* $c = 4$. *Then* $P_o^* = P_y^* = 4$, $Q_o^* = 12, Q_y^* = 8, E_o^* = 72$,$E_y^* = 32$,*and* $\pi^M = 104$.

1.3 Second Degree Price Discrimination

Often the buyers are not separated to distinct markets so that a seller who would like to price discriminate has to think of a clever scheme to divide its market into sub-markets in a way that will allow profitable discrimination. For example, it is common for cell phone companies to offer multiple calling plans. Each plan consists of a pair (E, p), where E is the fixed fee and p is the price per call. A flat-rate plan is one in which E is high but $p = 0$ (i.e., unlimited calling); in a metered plan, E is lower but p is positive. The idea behind offering the two plans is that the flat-rate plan is designed to appeal to consumers who are high users, and the metered plan to consumer who are low users. But, in order to discriminate in this way, the monopolist has to offer plans that induce consumers to sort themselves. The high demand users must prefer the flat-rate plan and the low demand users must prefer the metered plan. These sorting constraints will prevent the monopolist from extracting all of the consumer surplus.

Consider the above jazz club example under the assumption that the differences in willingness to pay does not depend upon age but some other characteristic that the club owner cannot observe (e.g., fans versus toursist). However, the jazz club owner knows the two demand curves. Suppose it offers the following menus:

$$\begin{aligned} A &: \quad \text{entry plus 12 drinks for \$120} \\ B &: \quad \text{entry plus 8 drinks for \$64.} \end{aligned}$$

The \$108 and \$64 represent the amounts that the high and low demand types spend respectively when the monopolist can observe their types and price discriminate perfectly using two-part tariffs (see above example). Will the high willingness to pay customers choose A and low willingness to pay custormers choose B? If the answer is yes, then the fact that consumers know their type but the jazz club owner does not is not an obstacle. To check, let us compute the consumer surplus of the two types when they choose either of the two plans.

$$CS_o(A) = 0; CS_o(B) = 32; CS_y(A) = -48, CS_y(B) = 0.$$

Clearly, the high demand types will choose B rather than A so the monopolist will not be able to capture all of the consumer surplus. Private information protects the consumer.

What is the optimal menu? Note that the problem here is with the high demand consumers. Menu B extracts all of the consumer surplus from the low demand types. But the monopolist cannot extract all of the consumer surplus from the high demand types. It needs to offer them a menu that they would prefer to B. The following menu is the one that does the trick:

$$\begin{aligned} A &: \quad \text{entry plus 12 drinks for \$88.} \\ B &: \quad \text{entry plus 8 drinks for \$64} \end{aligned}$$

The consumer surplus that the high demand types obtain from A is exactly 32 so they are indifferent between the two plans. The two types of consumers sort themselves by choosing different menus. Notice that A can be interpreted as offering a quantity discount. The price per drink is \$7.33 rather than \$8.

Example 7 ***Sorting through product choice - bundling with a "bad".*** *Consider the demand*

$$y(p) = \begin{cases} 0 & \text{if } p > 4 \\ 20 & \text{if } 2 < p \leq 4 \\ 50 & \text{if } p \leq 2 \end{cases}$$

Suppose that the valuation 4 consumers value their time at \$3 per hour while the others value time at \$1 per hour. Suppose also that the firm cannot distinguish among its customers. It can, however, offer two packages: (i) \$3.99 *per unit and no waiting time and (ii) \$1 per unit and an hour waiting time. The valuation consumers will choose (i), while the rest will choose (ii). The firm's profit will be* \$110 *as apoosed to* \$100 *that it can get as a regular monopoly.*

Note that the waiting time is a cost to society. It is incurred by consumers but not captured by the monopoly. It is however profitable for the monopoly since it uses time to sort its customers for price discrimination.

Example 8 ***Sorting Using "Damaged" Goods***

When Intel made the 486 chip widely available, they also produced a budget chip, the 486SX, that was manufactured by taking a fully functioning 486 chip and disabling the integrated math coprocessor. The processor greatly speeds up computations. Since disabling the coprocessor was costly, the 486SX chip cost more to produce than the original 486. Nevertheless, it sold for less \$333 as compared to \$588 for the 486.

Another example of this phenomenon is the IBM LaserPrinter E. In 1990 IBM introduced the LaserPrinter E as an alternative to the well estalbished Laser Printer. The E was identical to the standard model in every respect except that it contained extra chips whose function was to slow down the printing speed. The E sold for about \$1000 less than the original model.

Essentially the strategy here is to differentiate among consumers on the basis of their valuation of the computational or printing speed in much the same way the airlines try to differentiate between travelers who have different valuations on time.

1.3.1 Two-Part Tariffs (Optional)

Suppose that a monopoly with cost function

$$C(y) = cy$$

faces buyers of two types denoted by 1 and 2. Their respective demands for the product are $y_1(p)$ and $y_2(p)$, and their respective numbers are N_1 and N_2. Assume that buyer type 1 has a higher willingness to pay for each quantity. Let us derive now the profit maximizing

TPT. First, notice that, given a TPT scheme (E,p), the product will be bought by buyers of type i only if $E \leq s_i(p)$. Thus it will be bought by all buyers if $E \leq s_2(p)$, it will be bought only by type 1 buyers if

$$s_1(p) \geq E > s_2(p^m),$$

and it will not be bought by anyone if $E > s_1(p^m)$. Thus, the profit of the firm from the scheme (E,p) is

$$\pi(p,E) = \begin{cases} [N_1 y_1(p) + N_2 y_2(p)](p-c) + (N_1+N_2)E & \text{if } E \leq s_2(p), \\ N_1[y_1(p)(p-c)+E] & \text{if } E > s_2(p), \\ 0 & \text{if } E > s_1(p), \end{cases}$$

Notice that for any p, the firm's profit at that p will maximized by choosing

$$E = s_1(p) \textit{ or } E = s_2(p).$$

The former excludes type 2 customers but extracts the maximum possible from type 1's, while the latter will extract the maximum possible from type 2's. Thus, the monopoly's profit maximization problem can be divided into two steps. First, find the price p_1 that maximizes $\pi(p, s_1(p))$ and price p_2 that maximizes $\pi(p, s_2(p))$. Second compare $\pi(s_1(p_1), p_1)$ to $\pi(s_2(p_2, p_2)$. The profit maximizing TPT will be $(E^T, p^T) = (s_1(p^1), p^1)$ or $(s_2(p^2), p^2)$, according to whether $\pi(s_1(p_1), p_1)$ is smaller or larger than $\pi(s_2(p_2), p_2)$.

Example 9 *Let*

$$c = 0, \; N_1 = N_2 = N, \; y_1(p) = 1-p, \; y_2(p) = a(1-p), a < 1.$$

It follows that

$$s_1(p) = (1-p)^2/2; s_2(p) = a(1-p)^2/2.$$

To find p_2 we have to maximize over p the expression

$$\pi(s_2(p), p) = 2Ns_2(p) + Np(1-p) + Np(a(1-p)) = N[a(1-p)^2 + p(1-p)(1+a)]$$

Taking the derivative and equating to 0 yields the solution

$$p_2 = (1-a)/2.$$

Substituting back, we get

$$\begin{aligned} s_2(p_2) &= a(1-a)^2/8, \\ \pi(s_2(p_2), p_2) &= N(1+a)^2/4. \end{aligned}$$

An immediate calculation yields $p_1 = 0$ and $s_1(0) = 1/2$. Hence, $\pi(s_1(p_1), p_1) = N/2$. Therefore, the overall profit maximizing scheme is:

$$\begin{gathered} (1/2, 0) \;\; \textit{if} \;\; N/2 > N(1+a)^2/4. \\ [a(1-a)^2/8, (1-a)/2] \;\; \textit{if} \;\; N/2 < N(1+a)^2/4. \end{gathered}$$

Now,

$$N/2 > N(1+a)^2/4 \ if \ a < \sqrt{2} - 1.$$

Thus if a is sufficiently small, the profit maximizing scheme is (1/2,0) and type 2 buyers are excluded. If a is larger, the maximizing scheme is

$$(E^T, p^T) = [a(1-a)^2/8, (1-a)/2].$$

Remark: TPT is not always the best non-linear scheme that the monopoly can adopt. When the consumers served under the TPT have different demand curves, there are non-linear schemes which are more profitable. To see this suppose that there are two types as above, let (E, p) be the TPT scheme and let q_1 and q_2 be the quantities bought by consumers of types 1 and 2 under this scheme. Consider a non-linear scheme that offers only two deals. Quantity q_2 is offered for the payment $E + pq_2$, and the quantity q_1 is offered for slightly more than $E+pq_1$. Clearly, type 2 consumers will choose the first option while type 1 consumers will choose the second. This scheme is more profitable than the TPT since type 2 consumers get exactly the same quantity and pay exactly the same as under TPT, while type 1's get the same quantity as under the TPT but pay slightly more.

1.4 The Welfare Consequences of PD

There are three sources of inefficiency under price discrimination.
1. Like regular monopoly, the PD monopoly has an incentive to produce too little in any of its markets.
2. Inefficient allocation across markets in the sense that customers who value the product more highly are left without it while others get it
3. Sorting costs.

The first type of inefficiency may be more or less pronounced than under the corresponding regular monopoly. The other two are special to the PD monopoly and do not appear under the regular monopoly.

Is there room for regulation that prohibits discrimination? The answer of course depends on the circumstances. As shown above, sometimes PD enhances welfare and sometimes it is detrimental. It may be unwise to prohibit PD when it is in fact beneficial.

Bundling and Tying

In the previous lecture, we examined various strategies that monopolists employ in which different consumers end up paying different prices for the same commodity. In this lecture, we study how multi-product monopolists use bundling and tying strategies to achieve a similar outcome.

Bundling

Definition 1: A pure bundling strategy is when the monopolist offers to sell units of two commodities, 1 and 2, only as a bundle at a price P_{12}. A mixed bundling strategy is when the monopolist offers to sell a unit of product 1 alone at price P_1, a unit of product 2 alone at price P_2, and units of both products at price P_{12}.

Stigler Example: One distributor, two films, 1 and 2, and two television stations, A and B. The demand information is as follows:

	Maximum Willingness to Pay for Film 1	Maximum Willingness to Pay for Film 2
Station A	\$8,000	\$2,500
Station B	\$7,000	\$3,000

Optimal uniform pricing policy: $P_1 = 7{,}000$, $P_2 = 2{,}500$ and sell both films to both stations. Revenues = 14,000 + 5,000 = 19,000.

Optimal Pure Bundling Strategy: offer the two films in a package deal at P_{12} = \$10,000. Revenues = 20,000.

This example clearly demonstrates how bundling can improve profits.

Comments:

1. The distributor has to know the stations' willingness to pay.

2. The gains from bundling arise from differences in the buyers' relative valuations. Station A values film 1 more highly than station B, and station B values film 2 more than station A. For example, suppose that station A valued film 2 at \$3,000 and station B valued it at \$2,500. Then the optimal bundle price is \$9,500, which is simply the sum of the optimal uniform prices.

In the Stigler example, there is no reason for the distributor to consider a mixed bundling strategy. But this is not always true. In some cases, a mixed bundling strategy will be more profitable. It depends upon how the willingness to pay varies across consumers and, to some extent, on difference in marginal costs.

A More General Model

Unit demands, two goods, marginal costs of goods 1 and 2 are zero respectively.

R_1 = a consumer's reservation price for good 1
R_2 = a consumer's reservation price for good 2
R_{12} = consumer's reservation price for the bundle of one unit of each good.

Consumer surplus: $CS_i = R_i - P_i$, i = 1,2.

Assumptions:

1. Additive values: $R_{12} = R_1 + R_2$.

2. R_1 and R_2 are each distributed independently and uniformly on the interval [0, 1].

Assumption 2 implies that the probability that a randomly selected consumer is willing to buy good i at price P is equal to the probability that its reservation price exceeds P. This probability is equal to

$$\int_p dr = 1 - P.$$

Now assume that N, the number of consumers, is quite large. Then, applying the law of large numbers,

$$D_i(P) \approx N(1 - P).$$

I. Uniform Pricing

The monopolist sets P_1 and P_2 to equate marginal revenue to marginal cost in each market. In our example: $P_i^M = ½$, $\pi^M = ¼ + ¼ = ½$.

These prices partition the space of consumers into four rectangles:

$$\Gamma(0,0) = \{(R_1,R_2)|\ R_1 < 1/2, R_2 < 1/2\}$$

$$\Gamma(1,0) = \{(R_1,R_2)|\ R_1 \geq 1/2, R_2 < 1/2\}$$

$$\Gamma(0,1) = \{(R_1,R_2)|\ R_1 < 1/2, R_2 \geq 1/2\}$$

$$\Gamma(1,1) = \{(R_1,R_2)|\ R_1 \geq 1/2, R_2 \geq 1/2\}.$$

The four sets correspond to the sets of consumers who do not buy either good, buy good 1 only, buy good 2 only and buy both goods. (Illustrate the partition.)

II. Pure Bundling

The monopolist offers a bundle consisting of one unit of each good at price P_{12} and does not allow consumers to buy the goods separately.

The profit-maximizing price of the bundle cannot exceed the sum of the monopoly prices: $P_{12} \leq P_1^M + P_2^M$.

In our example, we can compute the optimal bundle price as follows. For any P_{12}, demand is given by 1 – area of the triangle in the unit square under the line $R_2 = P_{12} - R_1$.

$$D(P) = N(1 - P^2/2).$$

The profit-maximizing bundle price is $P_{12} = (2/3)^{1/2} \approx .82$ Therefore, demand is 2/3 and profits are $\pi_{12} \approx 0.55$.

The pure bundling strategy partitions the space of consumers into two sets:

$$\Gamma(0,0) = \{(R_1,R_2)|\ R_1 + R_2 < .82\}$$

$$\Gamma(1,1) = \{(R_1,R_2)|\ R_1 + R_2 \geq .82\}.$$

(Illustrate.) The two sets correspond to the set of consumers who do not buy either good and the set who buy both. The partition is no longer rectangles.

Now compare the two partitions.

1. The consumers who buy both goods under the uniform pricing will continue to do so under pure bundling since $P_{12} < 1$.

2. Some consumers who previously did not buy either good now buy both.

3. Some consumers who previously bought only one good now buy both (high reservation value for the one good is enough to subsidize the purchase of the second) and some now buy neither good (relatively low reservation values for both goods).

Thus, the monopolist faces a tradeoff in choosing between pure bundling and uniform pricing: it loses sales from some customers and gains sales from others. Which pricing policy is better depends upon the distribution of reservation prices in the population. In the uniform case, bundling is more profitable.

Remark: Suppose unit costs are positive. Under pure bundling, the allocation is inefficient for two reasons: (i) some consumers who do not purchase units have a willingness to pay that exceeds costs and (ii) some consumers who purchase units have a willingness to pay that is less than cost.

III. Mixed Bundling

Note that mixed bundling in which the stand-alone prices P_1 and P_2 exceed the bundle price P_{12} is effectively a pure bundling policy. No one will want to buy one good alone when the bundle is cheaper (assuming free disposal).

Therefore, mixed bundling is always at least as profitable as pure bundling and it can be more profitable.

Determining the optimal prices in our example is not trivial. It is not necessary to do so, however, to illustrate the tradeoffs. Let us assume that the monopolist uses the optimal uniform prices, $P_1 = ½$ and $P_2 = ½$, and the optimal bundle price, $P_{12} = .82$.

Then mixed bundling partitions the set of consumers into four sets defined as follows:

$$\Gamma(0,0) = \{(R_1,R_2)|\ R_1 < ½\ ,\ R_2 < ½\ ,\ R_1 + R_2 < .82\}$$

$$\Gamma(1,0) = \{(R_1,R_2)|\ R_1 - ½ \geq \max\{0, R_1 + R_2 - .82\}$$

$$\Gamma(0,1) = \{(R_1,R_2)|\ R_2 - ½ \geq \max\{0, R_1 + R_2 - .82\}\}$$

$$\Gamma(1,1) = \{(R_1,R_2)|\ R_1 + R_2 - .82 > \max\{0, R_1 - ½\ ,\ R_2 - ½\}\}.$$

In words, Γ(0,0) represents the set of consumers whose consumer surplus from buying one good only or buying the bundle is negative; Γ(1,0) is the set of consumers for whom buying good 1 only yields positive consumer surplus and which exceeds the consumer surplus from buying the bundle. The other sets are defined similarly.

(Graph the partition).

The graph illustrates the tradeoffs. When the monopolist offers unit prices that are less than the bundle price, some consumers who were not willing to buy the bundle will buy a unit of one of the goods. On the other hand, some consumers who were previously buying the bundle will switch to buying a unit of only one of the goods. This is the tradeoff that the monopolist must consider in deciding whether or not to set non-trivial unit prices.

Remark: The choice between mixed bundling and the other two schemes is further complicated when unit costs are positive and differ across the two goods. Profits are determined by markups, not by sales.

Examples: albums, airlines, cable TV.

Tying

Definition: Tying exists when a seller of a product requires as a condition of sale that the customer also purchase a second product (the tied product).

Tying is closely related to bundling. The main difference is that in tie-in sales the seller does not pre-specify the amounts of each good to be purchased, only that at least one unit of the first product must be purchased in order for the buyer to purchase units of the second product.

Example: good 1 = camera; good 2 = film.

Two types of consumers: High and Low

$$H: Q = 16 - P$$
$$L: Q = 12 - P$$

where P is the price of developing a picture. Film is produced in a competitive market at a marginal cost of \$2 per photo. The monopoly is in camera market.

Suppose the consumer can purchase film at \$2.

- H type consumer will take 14 pictures and is willing to pay up to \$98 (i.e., total surplus) to lease the camera.
- L type consumer will take 10 pictures and is willing to pay \$50 to lease the camera.

Thus, the best the monopolist can do is lease the camera at \$50 and make profits of \$100.

Now suppose the monopolist enters the film market and redesigns the camera so that it can only be used with film that the monopolist supplies. The monopolist's costs are also \$2 per photo. It prices the film at \$4 per photo.

- H type takes 12 pictures and is willing to pay \$72 (i.e., its consumer surplus) to lease the camera.
- L type takes 8 pictures and is willing to pay \$32 (its consumer's surplus) to lease the camera.

Thus, the monopolist can lease the cameras for \$32 and make \$40 of profit on film for total profits of \$104.

Notice that this example is basically the same as the jazz club example. Thus, the best strategy for the monopolist is to design two cameras, one that can take 14 pictures and one that can take 10 pictures (since marginal costs is \$2). It can then lease them for \$86 and \$70 respectively.

Check:

Willingness to pay of L types for 10 pictures is (10)(2) + (10)(10)(1/2) = 70.
Willingness to pay of H types for 10 pictures is (10)(6) + (10)(10)(1/2) = 110.
Willingness to pay of H types for 14 pictures is (14)(2) + (14)(14)(1/2) = 126.

Therefore, if monopolist charges $70 for camera with 10 pictures, the most that he can charge the H types for the camera with 14 pictures is

$$110 - 70 = 126 - R \rightarrow R = \$86.$$

Total profit: $106 > $104.

Note: book states $88 but I think this is a mistake.

Lecture on Oligopoly Pricing in Homogenous Good Markets

February 19, 2007

1 Introduction

An industry with a small number of competing firms (greater than 1), where "small" means that the decisions of each firm have nonnegligible effect on the others so that firms cannot ignore this effect in their considerations. In this lecture, we consider the following question: how are prices and output determined when there are a small number of firms producing a homogenous - identical - product?

The monopoly models feature a simple decision problem, namely, optimization in environments that react in a very predictable way. The modeling of oligopoly is conceptually more complicated, since when an oligopolist considers an action, such as announcing a price or introducing a new product, it has to consider how its competitors might react. It is not reasonable to suppose that they will not react or that their reactions are fully predictable.

Although the situations in which we are interested are inherently dynamic (i.e., repeated play), we initially look at a much simpler static model. The hope is that it will give us some insights into this complicated problem.

1.1 Cournot Duopoly Model

We present the Cournot duopoly model formally as a game in normal form. A game in normal form consists of three elements: player set, strategies, and payoffs.

- Player Set: $i = 1, 2$
- Strategy for firm i: $y_i \in [0, \infty)$;
- Payoffs for firm i:
$$\pi_i(y_i, y_j) = P(y_i + y_j)y_i - C(y_i)$$

Interpretation: two firms decide simultaneously what quantity to produce and supply to the market. Given the amounts supplied, price adjusts so as to clear the market. This model of price formation makes some sense for agricultural commodities where farmers commit to production before prices are determined or resources such as oil. In Cournot's original example, the commodity was spring water.

Assume demand is given by

$$\begin{aligned} P(Y) &= a - bY, \\ Y &= y_1 + y_2 \end{aligned}$$

and technology by

$$C(y_i) = cy_i, i = 1, 2.$$

Now consider the individual firm's optimization problem. When firm 1 contemplates what to supply, it has to think about what firm 2 might supply. Suppose it believes that 2 is going to produce y_2. Then its problem, anticipating the price formation process, is to find y_1 that maximizes its profit

$$\pi_1(y_1, y_2) = [a - b(y_1 + y_2) - c]y_1.$$

Notice that its demand curve is essentially the market demand with the intercept equal to $a - by_2$ instead of a. (Illustrate.) Thus, firm 1's marginal revenue also depends on the output of firm 2. Differentiating and setting marginal revenue equal to marginal cost implies that

$$y_1 = (a - c - by_2)/2b.$$

The above relationship is known as firm 1's *best reply.* Similarly, firm 2's best reply is

$$y_2 = (a - c - by_1)/2b.$$

Clearly, each firm should be optimizing given its belief about the other firm's action. But to say anything about what is likely to happen in this market we need to posit beliefs for each firm. Can we say anything about these beliefs? In general, no. However, we can suggest a criterion that under certain circumstances makes some sense.

Definition 1 *A* Nash Equilibrium *is a pair of quantities* (y_1^*, y_2^*) *satisfying*

$$\pi_i(y_i^*, y_j^*) \geq \pi_i(y_i, y_j^*)$$

for any $y_i \in [0, \infty), i \neq j, i, j = 1, 2.$

Interpretation: each firm is behaving optimally given its conjecture about its rival's choice of quantity and, in equilibrium, their conjectures are correct.

Mathematically, (y_1^*, y_2^*) lies on the best reply curves of both firms i.e., it is an intersection point. (Illustrate.) It can be obtained by solving the pair of best replies which, in this example, yields the solution

$$y_1^* = y_2^* = (a - c)/3b.$$

Equilibrium total output is

$$Y^* = \frac{2(a - c)}{3b}$$

and the equilibrium price is

$$P^* = \frac{a + 2c}{3},$$

and each firm's profit is

$$\pi^* = \frac{(a - c)}{9b}.$$

How might the firms arrive at this solution? If firms reach a non-binding agreement prior to production that these are the quantities they will produce, then the agreement will be self-enforcing (i.e., each firm will want to honor it if it believes that the other will). Such a situation may also be reached through repeated interaction where firms come to expect each other to produce these quantities because of past experience. It may also be possible that each firm arrives at its conjecture by solving its own and its rival's problems and recognizing that the above solution is the only pair of quantities in which conjectures are mutually consistent.

Example 1 *Suppose $a = 120, b = 20, c = \$20$. Then each firm's best reply is given by*

$$y_i = \frac{100}{40} - \frac{y_j}{2}.$$

Imposing symmetry and solving yields

$$y^* = \frac{5}{3}.$$

Equilibrium market output is $\frac{10}{3}$ and market price is $\frac{160}{3}$.

1.1.1 Symmetric Cournot Oligopoly with N Firms

The analysis of this case is an immediate extension of the above. Let

$$Y_{-i} = \sum_{j \neq i}^{n} y_j$$

denote the combined output of all firms other than firm i. Then firm i's best reply to the belief that its rivals will produce Y_{-i} is given by

$$y_i = (a - bY_{-i} - c)/2b.$$

Therefore, in a symmetric equilibrium in which each firm produces y^*

$$y^* = [a - b(N - 1)y^* - c]/2b$$

which yields the solution

$$y^* = (a - c)/[b(N + 1)].$$

Total output is

$$Y^* = Ny^* = \frac{N(a-c)}{b(N+1)}$$

and market price is

$$P^* = \frac{a}{N+1} + \frac{N}{N+1}c$$

Now consider what happens as N varies from 1 (monopoly) to infinity (perfect comptition. Each firm's output decreases and gets negligible in the limit; total output goes from the monopoly output of $\frac{a-c}{2b}$ to $\frac{a-c}{b}$, and price converges to c. Thus, the Cournot solution always lies between monopoly and perfect competition, and converges to perfect competition as the number of firms gets large.

1.1.2 An Asymmetric Cournot Duopoly

Suppose $C_i(y_i) = c_i y_i$ and $i = 1, 2$. The pair of best reply curves are given by

$$y_1 = \frac{(a - c_1 - by_2)}{2b}$$

$$y_2 = \frac{(a - c_2 - by_1)}{2b}.$$

(Illustrate.) Solving the pair of equations yields

$$y_1^* = \frac{(a + c_2 - 2c_1)}{3b}, \; y_2^* = \frac{(a + c_1 - 2c_2)}{3b}.$$

Thus, the more efficient firm (i.e., the one with the lower marginal cost) produces more and has higher market share but the relatively inefficient firm also produces. Unlike perfect competition, marginal costs are not equated.

1.1.3 Properties of the Cournot solution

Recall that each firm i chooses output to solve

$$\max_{y_i} [a - bY_{-i} - by_i]y_i - c_i y_i.$$

The first-order condition determining firm i's best reply to Y_{-i}, evaluated at the equilibrium, is

$$[a - bY_{-i}^* - 2by_i^*] - c = 0.$$

The term in brackets is firm $i's$ marginal revenue. We can rewrite this equation as

$$P^* - c = by_i^*.$$

Dividing both sides of this equation by the price yields

$$\frac{P^* - c}{P^*} = s_i^* \frac{bY^*}{P^*}$$

where $s_i^* = \frac{y_i^*}{Y^*}$ is firm i's market share. Now recall the definition of demand elasticity:

$$\eta(P) = \frac{\partial Y(P)}{\partial P} \frac{P}{Y(P)}.$$

Inverting the inverse demand to obtain the demand function,

$$Y(P) = \frac{a}{b} - \frac{P}{b}.$$

Therefore, for linear demand,

$$\eta(P) = \frac{-P}{bY(P)}.$$

Substituting, we obtain the markup equation:

$$\frac{P^* - c}{P^*} = \frac{s_i^*}{\eta}.$$

Several observations follows immediately from the above equation:
1. Oligopolists in the Cournot solution exercise market power since price is above marginal cost.
2. Market power is limited by the market elasticity. The more elastic demand, the less the mark-up.
3. Firms with lower marginal costs will have greater market shares.
4. The greater the number of competitors, the smaller is each firm's market share and the less its market power.
5. In the symmetric case, the solution lies between the competitive equilibrium and monopoly in terms of quantities, prices, profits and surplus.

1.2 The Bertrand Duopoly Model

Let $D(P)$ denote demand in the market when price is P. In the Bertrand model, the normal form of the game is given as follows:

- Player set: two firms, indexed by $i = 1, 2$.
- Strategy for firm i: $p_i \in [0, \infty)$;
- Payoffs for firm i:

$$\pi_i(p_i, p_j) = \begin{cases} (p_i - c_i)D(p_i) & \text{if } p_i < p_j \\ (p_i - c_i)D(p_i)/2 & \text{if } p_i = p_j \\ 0 & \text{if } p_i > p_j \end{cases}$$

Interpretation: two firms decide simultaneously what price to charge and produce to demand. Consumers know the prices charged and buy from the firm that charges the lowest price. Each firm is assumed to have sufficient capacity to service all of the market demand at its quoted price.

The Bertrand model captures an important feature of many markets, namely, that firms set prices. However, it yields a very stark result.

Definition 2 *A* Nash Equilibrium *is a pair of prices* (p_1^*, p_2^*) *satisfying*

$$\pi_i(p_i^*, p_j^*) \geq \pi_i(p_i, p_j^*)$$

for all $p_i \in [0, \infty), i \neq j, i, j = 1, 2.$.

The only pair of prices satisfying the two conditions given above is

$$p_1 = p_2 = c$$

[Show graphically why marginal cost pricing is the only equilibrium.] This result is known as the Bertrand paradox: "One is monopoly, two is perfect competition". Example: pricing of air travel in a city-pair market.

1.2.1 An Asymmetric Duopoly

Suppose the unit costs of the firms are c_1 and c_2 with $c_1 < c_2 < p^M(c_1)$. Assuming prices are denominated in pennies, the undercutting logic yields a Nash equilibrium in which $p_1 = c_2 - \$.01$, $p_2 = c_2$. [Prove this claim]. That is, the lower cost firm prices just below the unit cost of the higher cost firm and captures all of the market.

1.3 Remarks

For many markets, price competition seems more natural than quantity competition. However, the Bertrand model yields predictions that appear inconsistent with observation. What is wrong? Need to focus on three crucial assumptions.

1. Unlimited capacity

In practise, firms typically do not have the capacity to service the entire market demand. That is, in the short run, capacity is fixed. Customers are rationed, which means that the higher price firm will face a residual demand curve against which it can optimize. This typically leads to equilibrium prices above marginal cost. But, capacity itself should be viewed as a choice, one which firms will make anticipating how the outcome of the price competition will vary depending upon the capacity choices taken. If they anticipate correctly, how much capacity will duopolists choose? In a remarkable result, Kreps and Scheinkman have shown that, given certain regularity conditions and efficient rationing, duopolists will choose capacities equal to the Cournot quantities. In other words, they will commit not to engage in ruthless price competition. In fact, the equilibrium of the pricing game will be that each firm sets the same price and that price is equal to the Cournot price.

2. Homogenous good

Consumers care only about price and respond en mass to the slightest difference in price by buying from the lowest price firm. In reality, firms often have preferences over suppliers and are willing to accept some differential before responding. Costs of adjustment or differentiated products will allow the more "continuous" adjustment to price cuts. It will not be true that a penny difference in price can induce a huge change in demand.

3. Static game

The price game is played only once. In practise, firms compete against each other repeatedly.

1.4 Bertrand with Capacity Constraints

We consider relaxing each of these restrictions. To illustrate how capacity constraints undermine the Bertrand paradox, consider the following example. The world famous Whistler Ski Resort in Vancouver, BC. offers skiing on two adjacent mountains: Blackcomb and Whistler. Gondolas take skiers from Whistler village to their respective mountains. The quality of skiing offered by both mountains is essentially the same so skiers will choose the mountain that posts the lowest price of a daily lift ticket. Blackcomb has fewer runs than Whistler and can accommodate only 1000 skiers per day. Whistler can accommodate 1400 skiers per day. Total daily demand for skiing is given by

$$Q = 6,000 - 60P$$

where P is the price of a daily lift ticket. Marginal costs of providing lift services is \$10 per skier.

Clearly, marginal cost pricing is not an equilibrium. If the price of a ticket is \$10, then total demand is 5,400. Demand far exceeds the combined capacity of the two mountains, which is 2400. Each mountain could raise the price of a ticket above \$10 and still sell out their capacity.

What is the equilibrium? The highest price that Blackcomb and Whistler could charge that would equate demand and supply is to charge a price of \$60. At this price, total demand is 2,400, which is exactly equal to the combined capacity of the two mountains. Will either mountain have an incentive to change its price? Obviously, neither mountain wants to lower its price; they are already selling to capacity, so a lower price will not give them more customers. Will either mountain want to raise price? To answer this question, we need to make an assumption about which skiers will get to buy at the lower price. The assumption that we make is that the skiers with the highest willingness to pay will buy at the lower price. Hence, if Blackcomb charges \$60, and Whistler raises its price above \$60, the residual demand curve that it faces is

$$Q_W = 5,000 - 60P_W.$$

This is a parallel shift in the demand curve: 1000 skiers with the highest willingness to pay buy their ticket from Blackcomb and the rest have to buy from Whistler. Blackcomb's profits are given by

$$\pi(P_W) = (5000 - 60P_w)(P_W - 10).$$

Differentiating with respect to P_W, we get that

$$\frac{\partial \pi(P_W)}{\partial P_W} = 5000 - 60P_W - 60(P_w - 10) = 4400 - 120P_W.$$

Evaluating this derivative at $P_W = \$60$, we find that

$$\frac{\partial \pi(60)}{\partial P_W} = -2800.$$

Thus, Whistler's profits decrease if it tries to raise price. A similar calculation holds for Blackcomb. Hence, the Nash equilibrium price is \$60.

Horizontal Mergers

November 29, 2005

Horizontal mergers occur when the firms joining together are formerly competitors in the same product market. An example is the merger between AT&T and Cingular, who were previously competing against each other in the market for cellular services. Mergers and other forms of asset or stock acquisitions are covered under Section 7 of the Clayton Act. The criteria for illegality is whether the effect of the merger would be to "substantially lessen competition". The problem is how to codify this criteria. There are basicly two measurement issues: the definition of a market and a measure of concentration in that market.

1 Market Definition

1. In 1993, the U.S. Department of Justice brough a suit to prevent the merger of two makers of premium fountain pens, Gillette, through its brand subsidiary, Waterman, and Parker Pen Holdings. The key issue was the definition of market. The government argued that Waterman and Parker pens were very close substitutes in the product space, whereas other fountain pens, and other products, were much less close substitutes. For example, their expert stated "I do not know of any fountain pens with suggested retail prices under $50 that have solid gold nibs. Compared to Parker and Waterman premium fountain pens with gold-plated nibs, fountain pens with suggested retail prices under $50 may not available with more than one nib width, and they may be perceived to be of poor quality or to lack of brand image or brand reputation comparable to Parker and Waterman." The defendant's expert argued that the $50 lower bound was spurious, and that many pens priced below $50 competed effectively with those above. He argued for a broader definition of market, the market of "highline pens" - all pens with suggested retail prices over $10. The court decided for the defendant.

2. In 1996, Staples, an office super store (OSS), entered into an agreement to acquire through merger its major rival, Office Depot, for $4 billion. The FTC filed an injunction to prevent the merger. The FTC argued that office super stores had created da well defined product market of their own, specializing in high volume, convenient distribution, and low prices. This market is distinct from the market in which many of the same products are sold by other retailers, smaller stores, and non-specialist stores. The FTC also argues that the geographic boundary consists of individual city markets. Thus, the market structure is

either monopoly, duopoly, or triopoly, depending on which of the three OSSs (Staples, Office Depot, and OfficeMax) are active in a given market. The merger would create monopolies and duopolies, and significant price increases would result, at least in the 5-10% range. The defendents argued that the relevant market is office products in which the merged entity would have only a 6% market share. They also claim that there is easy entry or product repositioning from either specialized office supply stores or major retailing chains like Wal Mart and K Mart.

2 The Guidelines

In defining a market, we need to find a gap in the chain of demand substitution across different products. For example, Teva Sport Sandals are to some extent substitutes for Reeboks and Birkenstocks, which are themselves substitutes for Rockport Casual shoes, which are in turn substitutes for Church's English dress shoes. (Recall the Hotelling model.) If all of these shoes are very close substitutes for each other, then none of these suppliers have very much market power. More precisely, a small increase in price will lead to a large reduction in quantity demanded, which is simply another way of saying that the own price elasticity is high (or equivalently, the cross-price elasticities between shoes is high). For many years, market definition in antitrust cases was defined in terms of substitutability and quantified using price correlations and elasticity estimates. This approach gave rise to inconsistent judgments. Markets were defined too broadly in some cases and too narrowly in others. Elasticity measurements taken before the merger are not always a good proxy for elasticity measurements taken after the merger (e.g., du Pont cellophane fallacy). In any case, the main point is an antitrust case is not market shares per se but whether the merged firm can increase prices.

The recent 1982 guidelines for mergers distinguish between the economic concept of a market, which is based upon substitutability, and an antitrust market, which is based upon the power to raise prices. An antitrust market is defined by applying the "Hypothetical Monopolist" test:

"a market is defined as a product or group of products and a geographic area in which it is sold such that a hypothetica, profit-maximizing firm, not subject to price regulation, that was the only present and future seller of those products in that area would impose a small but significant and nontransitory increase in price above prevailing or likely future levels."

The basic idea here is to take the merged entity and find out what would be the smallest group of products, containing that entity, which if owned by a single monopolist, could profitably increase price by a significant amount.

1. If a set satisfies the hypothetical monopolist test, then adding products to the set generates sets that also satisfy the test. The rule here is to choose the smallest set. However, even after applying this rule, there may be more than one minimal set.

2. The process of market delineation involves only the demand side. The possibility of entry into the market by other firms in response to the merger is a consideration that needs to

be argued separately.

3 Concentration Measures

Having defined the market, the next step is to measure concentration before and after the merger is permitted and ask whether the change in concentration will allow the merged entity to raise price significantly. Two measures have been adopted by the courts.
1. Order the market (output) share of firms from largest to smallest. The m firm concentration ratio is the sum of the market shares of the largest m firms:

$$CR_m = \sum_{i-1}^{m} s_i$$

The most frequently used are four firm (CR_4)and eight firm (CR_8) ratios. Here $s_i = y_i/Y$, so the range of CR_m is 0 to 1.
2. Hefindahl-Hirschman Index is the sum of the squares of market shares of all firms in the industry.

$$HHI = \sum_{i=1}^{N} (100 s_i)^2$$

Multiplying the share by 100 converts the share into a percent. The range of the index is 10,000/N (equal sharing) to 10,000 (monopoly). It has several desirable properties.
(a) HHI decreases with the number of firms. For example, if market shares are equal, HHI = (1/N) which decreases with the number of firms.
(b) HHI increases with the variance in the distribution of firm sizes.
(c) Recall from the Cournot model, that in equilibrium,

$$(P^* - c_i)/P^* = s_i/\eta(Y^*).$$

Multiplying through by $10000 s_i$ and summing over all i yields

$$\sum_{i=1}^{N} 10000 s_i (P^* - c_i)/P^* = (HHI)/\eta(Y^*).$$

Thus, HHI is proportional to a weighted average of each firm's percentage markup of price over cost. It is a summary statistic of market power. Holding demand elasticity constant, an industry with a higher HHI value has a higher average percent markup. The reason why the above equation is useful is because price-cost margins are not observable, but market shares and demand elasticities are observable.

3.1 The Challenge Criteria

In 1982, the guidelines suggested that a merger should not be challenged if post-merger HHI is:
1. Less than 1000
2. Between 1000 and 1800, and $\triangle HHI < 100$;
3. $HHI > 1800$ and $\triangle HHI < 50$.

These suggestions were qualified marginally in 1984.

Remark 1 *The measure assumes that the shares of the merging firms are equal to the sum of the pre-merger shares. Theory suggests that this is not reasonable.*

4 Horizontal Mergers in Homogenous Good Markets

If firms have identical marginal costs and compete in quantities, then a merger between two firms increases price and reduces welfare.

Example 1

Suppose demand is given by $P(y) = 1 - Y$. There are three firms, marginal costs are equal to zero, and they compete in quantities. We first derive the Cournot equilibrium. Firm 1 chooses its output y_1 to

$$\max_{y_1} \pi(y_1) = (1 - y_1 - y_2 - y_3)y_1.$$

Differentiating with respect to y_1 and solving for its best reply yields

$$y_1 = \frac{1}{2}(1 - y_2 - y_3).$$

Similarly, the best replies for firms 2 and 3 are given by

$$\begin{aligned} y_2 &= \frac{1}{2}(1 - y_1 - y_3). \\ y_3 &= \frac{1}{2}(1 - y_1 - y_2). \end{aligned}$$

Solving these three equations yields the Cournot triopoly equilibrium. Imposing symmetry and solving, we obtain

$$y_i^* = 1/4, p^* = 1/4, \pi_i^* = 1/16, S^* = 9/32, T^* = 15/3$$

where S is consumer surplus and T is total surplus (total profits + consumer surplus).

Now suppose firms 1 and 2 merge into firm 12 so the industry is a duopoly. In that case, the merged firm chooses its output y_{12} to

$$\max_{y_{12}} \pi(y_{12}) = (1 - y_{12} - y_3)y_{12}.$$

Differentiating and solving for its best reply to y_3, we obtain

$$y_{12} = \frac{1}{2}(1 - y_3).$$

Firm 3's best reply to y_{12} is given by

$$y = \frac{1}{2}(1 - y_3).$$

Solving the two equations yields the Cournot duopoly equilibrium. The solution is

$$\widetilde{y}_{12} = y_3 = 1/3, \widetilde{p} = 1/3, \widetilde{\pi}_{12} = \pi_3 = 1/9, \widetilde{S} = 2/9, \widetilde{T} = 4/9$$

Conclusion 1 *Industry output falls from 3/4 to 2/3; price increases from 1/4 to 1/3; profits to the merging firms falls - their combined pre-merger profits are 1/8, but their post merger profits are 1/9. The merger is unprofitable! This is known as the merger paradox. Notice that merger confers a positive externality to firm 3, the outside firm: its profits go from 1/16 to 1/9. What has happened is that it expands output in response to the merging firm's reduction in output. Thus, its market share goes up, from 1/3 to 1/2, and, in addition, the price is higher. Hence, its benefits big time from the merger. Consumer surplus falls, profits rise, but inefficiency costs increase as social welfare falls by 5.2%.*

Example 2 *Practise Problem 16.1*

Here demand is given by

$$P = 130 - Y,$$

marginal costs is 30, and the industry has 20 firms. To derive the Cournot equilibrium, consider first firm 1's profit maximization problem under the assumption that everybody else is producing $\overline{y}$:

$$\max_{y_1} \pi(y_1) = (130 - y_1 - 19\overline{y})y_1 - 30y_1.$$

Differentiating, we obtain the first order condition

$$130 - 2y_1 - 19\overline{y} - 30 = 0$$

which implies that firm 1's best reply is

$$y_1 = \frac{1}{2}(100 - 19\overline{y}).$$

In the symmetric equilibrium, each firm produces

$$y^* = \frac{100}{21}$$

and the price is

$$P^* = 130 - (20)\frac{100}{21} = 34.76.$$

Substituting these values into each firm's profit function yields

$$\pi^* = 22.67.$$

Now suppose six firms, firms 15-20, merge. Repeating the above steps for 15 firms, firm 1's maximization problem becomes

$$\max_{y_1} \pi(y_1) = (130 - y_1 - 14\overline{y})y_1 - 30y_1.$$

Differentiating, we obtain the first order condition

$$130 - 2y_1 - 14\overline{y} - 30 = 0$$

which implies that firm 1's best reply is

$$y_1 = \frac{1}{2}(100 - 14\overline{y}).$$

In the symmetric equilibrium, each firm produces

$$y^* = \frac{25}{4}$$

and the price is

$$P^* = 130 - (15)\frac{25}{4} = 36.25.$$

Substituting these values into each firm's profit function yields

$$\pi^* = 39.06.$$

Conclusion 2 *Prior to the merger, the six firms earned a combined profit of 136.02; after the merger, their profit is 39.06!. Price has gone up but most of the benefits of the merger has gone to firms 1 through 14. In fact, one can show that for the merger to be profitable, over 80% (or 16 firms) have to merge in this industry for the merger to be profitable.*

How can we resolve the merger paradox? One approach is to assume that the merger generates cost efficiencies. One form of cost efficiencies is lower fixed costs. For example, in the AT&T and Cingular merger, the fixed costs that each firm incurs are the costs of installing and maintaining the network of cell-towers. With the merger, only one set of cell towers is needed. This kind of efficiency is often mentioned as the reason for a merger.

Example 3

Suppose industry demand is given by

$$P = 150 - Y,$$

each firm i's cost function is given by

$$C(y_i) = f + 30y_i,$$

and there are three firms. Firm 1 chooses y_1 to

$$\max_{y_1} \pi(y_1) = (150 - y_1 - y_2 - y_3)y_1 - 30y_1 - f.$$

Differentiating profits with respect to y_1 and setting the derivative equal to zero, we obtain

$$150 - 2y_1 - y_2 - y_3 - 30 = 0,$$

which implies that

$$y_1 = \frac{1}{2}(120 - y_2 - y_3).$$

In the symmetric equilibrium, $y_1 = y_2 = y_3 = y^*$. Thus,

$$y^* = \frac{1}{2}(120 - 2y^*) \Longrightarrow y^* = 30, P^* = 60, \pi^* = 900 - f.$$

Now suppose firms 1 and 2 merge into firm 12. In this case, repeating the above steps, we find that firm 12's best reply is

$$y_{12} = \frac{1}{2}(120 - y_3).$$

In the symmetric equilibrium,

$$\widetilde{y}_{12} = \widetilde{y}_3 = 40, \widetilde{P} = 70, \widetilde{\pi} = 1600 - f.$$

Note that the fixed costs of the merged firm is f, not $2f$. Hence, the merger is profitable if

$$1600 - f > 1800 - 2f \Longrightarrow f > 200.$$

Conclusion 3 *Mergers can be profitable if they reduce fixed costs. Note, however, that fixed costs have no impact on price. Thus, for these kinds of mergers, price always rises, and consumer surplus always falls.*

Another kind of cost efficiency arises when one of the firms in the merger has higher marginal costs that the other firms, and the merged entity operates at the lower marginal cost. These kinds of mergers eliminates inefficient firms. There is a tradeoff between market power and enhanced efficiency which may make the merger socially beneficial.

Example 4

Suppose industry demand is given by

$$P = 150 - Y,$$

Firms 1 and 2 have cost functions that are given by

$$C_1(y_1) = 30y_1, C_2(y_2) = 30y_2$$

and firm 3's cost function is

$$C_3(y_3) = 60y_3.$$

In this case, the system of best replies are given by

$$\begin{aligned} y_1 &= \frac{1}{2}(120 - y_2 - y_3) \\ y_2 &= \frac{1}{2}(120 - y_1 - y_3) \\ y_3 &= \frac{1}{2}(90 - y_1 - y_2). \end{aligned}$$

Solving this system for the equilibrium values of output yields

$$y_1^* = y_2^* = \frac{150}{4}; y_3^* = \frac{30}{4}.$$

Equilibrium price and profits are

$$P^* = \frac{270}{4}, \pi_1^* = \pi_2^* = \frac{150^2}{16}, \pi_3^* = \frac{900}{16}.$$

Now suppose firms 2 and 3 merge and operate at marginal costs of 30. The industry becomes a duopoly, and the equilibrium is computed as above:

$$\widetilde{y}_1 = \widetilde{y}_{23} = 40, \widetilde{P} = 70, \widetilde{\pi}_1 = \widetilde{\pi}_{23} = 1600.$$

Thus, price rises but the merger is profitable since

$$1600 > 1406.25 + 56.25 = 1462.25.$$

Conclusion 4 *Mergers can be profitable if they eliminate a higher marginal cost firm. The assumption here is that the merged firm operates at the marginal cost of the most efficient participation firm.*

5 Horizontal Mergers in Differentiated Good Markets

Many mergers are in differentiated product markets. For example, when Universal bought Polygram Records, the artists of Universal and the artists of Polygram continued to compete against each other in the market for music. However, their CD's were produced by one company rather than two separate companies. Similarly, when Pepsi bought 7-UP, it continued to sell both brands. When retailers like Staples and Office Depot merge, they will typically continue to operate the two sets of stores since they are located in different parts of the city. The only difference is that the two sets of stores are owned by one firm rather than two separate firms. What are the gains to merging in these kinds of industries?

Example 5

Consider an industry with three firms selling three differentiated products at zero marginal costs. The demands for the firms are as follows:

$$\begin{aligned} y_1 &= 1 - p_1 + s(p_2 + p_3) \\ y_2 &= 1 - p_2 + s(p_1 + p_3) \\ y_3 &= 1 - p_3 + s(p_1 + p_2). \end{aligned}$$

We will assume that goods are substitutes, that is, $s > 0$. Recall that own-price effects has to exceed cross-price, which means that $s < 1$. An increase in p_2 increases demands for good 1 and for good 3 as marginal buyers shift their demands from good 2 to the other two goods. To compute the Bertrand equilibrium, consider first the maximization problem of firm 1:

$$\max_{p_1} \pi(p_1) = p_1(1 - p_1 + s(p_2 + p_3)).$$

Differentiating profits with respect to p_1 and setting the derivative equal to zero yields

$$1 - 2p_1 + s(p_2 + p_3) = 0.$$

Solving for firm 1's best reply,

$$p_1 = \frac{1}{2}[1 + s(p_2 + p_3)].$$

Similarly, best replies of firms 2 and 3 are given by

$$\begin{aligned} p_2 &= \frac{1}{2}[1 + s(p_1 + p_3)] \\ p_3 &= \frac{1}{2}[1 + s(p_1 + p_2)]. \end{aligned}$$

In the symmetric equilibrium, $p_1^* = p_2^* = p_3^* = p^*$. Substituting into one of the above equations yields

$$p^* = \frac{1}{2}[1 + 2sp^*] \Longrightarrow p^* = \frac{1}{2(1-s)}, y^* = \frac{1}{2(1-s)}, \pi^* = \frac{1}{4(1-s)^2}.$$

Now suppose firms 1 and 2 merge into one firm, firm 12. This firm will continue to produce product 1 and product 2. However, it will now coordinate the prices of these two goods. More precisely, it chooses p_1 and p_2 to maximize profits from product 1 and from product 2:

$$\max_{p_1,p_2} \pi_{12}(p_1,p_2) = \pi_1 + \pi_2 = p_1(1 - p_1 + s(p_2 + p_3)) + p_2(1 - p_2 + s(p_1 + p_3)).$$

To find the optimal prices given p_3, the merged firm differentiates profits with respect to each price and sets the derivatives equal to zero.

$$\begin{aligned} \frac{\partial \pi_{12}}{\partial p_1} &= 0 \Longrightarrow 1 - 2p_1 + s(p_2 + p_3) + sp_2 = 0 \\ \frac{\partial \pi_{12}}{\partial p_2} &= 0 \Longrightarrow 1 - 2p_2 + s(p_1 + p_3) + sp_1 = 0. \end{aligned}$$

The best reply for firm 3 is the same as above. The difference between the first order conditions of firm 12 and the first order conditions of firms 1 and 2 before the merger are the terms sp_2 and sp_1 that are added to the equations for p_1 and p_2 respectively. The term sp_2 in the first equation represents the impact of an increase in p_1 on the profits of firm 12 in market 2; some buyers leave market 1 and some of them go to market 2 where their purchases increases profits in that market. Similarly, the term sp_1 in the second equation measures the impact of an increase in the price of good 2 on firm 12's profits in market 1; marginal buyers leave market 2 and some will purchase good 1, thereby increasing firm 12's profits in that market.

Now it is clear that evaluating firm 12's first order conditions at p^* yields a positive value for the derivatives. That is,

$$1 - 2p^* + s(p^* + p^*) + sp^* > 0.$$

Hence, in order for the first order conditions to be satisfied, firm 12 has to raise prices of goods 1 and 2. Since $s < 1$, an increase in p_1 and p_2 lowers the value of the derivatives. Holding $p_3 = p^*$, we see that the profits of firm 12 exceed the sum of the profits of firm 1 and 2 prior to the merger. The gain arises from the fact that the merged firm internalizes the impact of p_1 on the profits earned in market 2 and impact of p_2 on the profits earned in market 1.

How will firm 3, the outside firm, respond? If p_1 and p_2 increase, then we know from firm 3's best reply that it will increase its price above p^*. But an increase in p_3 means that some buyers will stop buying good 3 and buy goods 1 or 2 instead, which raises profits of firm 12 even further. Thus, a merger of two differentiated firms is highly profitable, and it increases prices in all three markets!

Oligopoly Pricing in Differentiated Markets

March 23, 2007

1 Introduction

Products produced by different firms are often differentiated by their characteristics or attributes. A useful classification is to distinguish between *horizontal* and *vertical* differentiation.

1. Horizontal: attributes are location or type. Examples: ice cream vendors on a beach are differentiated by their location; dark or light beer; different cereals.

2. Vertical: attributes make one of the products better than another - i.e., quality, durabity. Examples: batteries - Duracell vs Energizer; diamonds - number of carats.

The key modeling difference lies in the specification of consumer preferences. If the product space is vertically differentiated, consumers have the same ordering over products - they unanimously agree on which product or brand is better (i.e., there is a "quality" ladder). On the other hand, if the product space is horizontally differentiated, consumer have diverse preferences over products - some will prefer A to B and others will prefer B to A. There is no agreement on which product or brand is preferred. For example, bathers on a beach will prefer ice cream vendor A to B if they are closer to A than B and B to A if they are closer to B. In this course, we focus primarily on horizontal differentiation.

1.1 A Conceptual Framework

In a model with differentiated products, demand functions vary continuously with respect to own and rivals' prices. For example, with two products, the demand functions can be expressed as

$$\begin{aligned} q_1(p_1,p_2) &= a - bp_1 + cp_2; \\ q_2(p_1,p_2) &= a - bp_2 + cp_1; \quad b > 0, b > |c|. \end{aligned}$$

Here, the parameter c measures the change in demand for good i when the price of good j increases. A very small change in price of good j always leads to a small change in demand for good i. The two goods are substitutes if $c > 0$ and complements if $c < 0$. The assumption that $b > |c|$ implies that own price effects dominate cross price effects. It guarantees invertibility of the system. Thus, an equivalent representation of the system of demands is given by the inverse demands

$$\begin{aligned} p_1(q_1,q_2) &= \alpha - \beta q_1 - \gamma q_2; \\ p_2(q_1,q_2) &= \alpha - \beta q_2 - \gamma q_1, \end{aligned}$$

where

$$a = \frac{\alpha(\beta - \gamma)}{\beta^2 - \gamma^2},\ b = \frac{\beta}{\beta^2 - \gamma^2},\ c = \frac{\gamma}{\beta^2 - \gamma^2}.$$

Remark: The restriction $b > |c|$ rules out perfect substitutes and complements. As we shall see below, this condition is also needed to ensure that best replies intersect - i.e., an equilibrium exists. Note also that γ is postive if goods are subsitutes, negative if goods are complements.

As in the case of homogenous good markets, the outcome depends upon whether firms compete in prices or in quantities (capacities). Let us first derive the best reply functions in the game where firms choose prices. To keep the calculations simple, we set marginal costs to zero. Then firm 1 chooses its price to solve the following problem:

$$\pi_1(p_1, p_2) = (a - bp_1 + cp_2)p_1.$$

Differentiating we obtain

$$\frac{d\pi_1}{dp_1} = (a - bp_1 + cp_2) - bp_1 = q_1(p_1, p_2) - bp_1.$$

The tradeoff here is the same as in the monopoly case: the first term measures the gain in revenues from selling q_1 at a higher price and the second term measures the loss in revenue from lower sales. The optimal price equates these two effects, which yields firm 1's best reply function,

$$p_1 = (a + cp_2)/2b.$$

The slope of firm 1's best reply is $c/2b$. If goods are substitutes, then firm 1's best reply slopes upward: an increase in p_2 raises demand for good 1 and firm 1 will want to increase p_1 to offset the increase in q_1. The optimal increase in p_1 is less than the increase in p_2 since $c/2b < 1$ so the net effect on firm 1's sales is positive. If goods are complements, then firm 1's best reply slopes downward: an increase in p_2 lowers demand for good 1 and firm 1 will want to reduce p_1.The net effect on firm 1 sales in this case is negative. Symmetry implies that firm 2's best reply to p_1 is given by

$$p_2 = (a + cp_1)/2b.$$

Solving this pair of equations gives the Nash equilibrium prices:

$$p_1^* = p_2^* = \frac{a}{2b - c}.$$

In the quantity (capacity) game, firm 1 chooses its quantity to solve the following problem:

$$\pi_1(q_1, q_2) = (\alpha - \beta q_1 - \gamma q_2) q_1.$$

Differentiating yields

$$\frac{d\pi_1}{dq_1} = (\alpha - \beta q_1 - \gamma q_2) - \beta q_1 = p_1(q_1, q_2) - \beta q_1.$$

Here the first term measures the gain in revenues from selling more output and the second term measures the loss in revenue from selling output at a lower price. The optimal quantity equates these two effects, which yields firm 1's best reply function,

$$q_1 = (a - \gamma q_2)/2\beta.$$

The slope of firm 1's best reply is $\gamma/2\beta$. If goods are substitutes, then firm 1's best reply slopes down: an increase in q_2 reduces the price at which firm 1 can sell its output and it will respond by lowering its output. The optimal decrease in q_1 is less than the increase in q_2 since $\gamma/2\beta < 1$ so the net effect on firm 1's price is negative. If goods are complements, then firm 1's best reply slopes up: an increase in q_2 increases the price at which firm 1 can sell its output and firm 1 will respond by reducing output and increasing price..The net effect on firm 1's price in this case is positive. Similarly, firm 2's best reply to q_1 is

$$q_2 = (\alpha - \gamma q_1)/2\beta.$$

Solving this pair of equations for the Nash equilibrium quantities yields

$$q_1^* = q_2^* = \frac{\alpha}{2\beta + \gamma}.$$

As we shall see in later lectures, the effect of market power on prices and quantities in differentiated product markets will depend criticially upon whether best reply functions slope upward or downward. If best replies slope upward, then game is one of *strategic complements*; if best replies slope down, then the game is one of *strategic substitutes*. Hence, prices are strategic complements when goods are substitutes and they are strategic substitutes when goods are complements. Quantities (capacities) are strategic substitutes when goods are substitutes and they are strategic complements when goods are complements.

	Substitutes	Complements
quantity	negative slope	positive slope
price	positive slope	negative slope

Which game yields higher prices? Using the relationship between the two sets of parameters, some tedious calculations yields the following outcome matrix:

	price	quantity	profit
Cournot	$\alpha\beta/(2\beta+\gamma)$	$\alpha/(2\beta+\gamma)$	$\alpha^2\beta/(2\beta+\gamma)^2$
Bertrand	$\alpha(\beta-\gamma)/(2\beta-\gamma)$	$\alpha\beta/[(2\beta-\gamma)(\beta+\gamma)]$	$\alpha^2\beta(\beta-\gamma)/[(2\beta-\gamma)^2(\beta+\gamma)]$

Note that the outcomes are identical if $\gamma = 0$. If $\gamma \neq 0$, then differencing the Bertrand price from the Cournot price reveals that the latter is higher. Thus, the pricing game always yields lower prices and higher quantities. However, the ranking on profits depends upon whether goods are substitutes or complements. Profits are higher in the quantity game if products are substitutes but lower if products are complements. It is in this sense that firms play *soft* when goods are strategic substitutes and *hard* when they are strategic complements.

1.2 An Example: The Turnpike Model

The following example illustrates how antitrust policy in markets with complements differs dramatically from antitrust policy in markets with substitutes. Consider a "private" road from point A to point B of length 1. Ownership of the road is divided among N individuals who each own 1/Nth of the road. Let p_i denote the toll that individual i charges travelers for driving on its section of the road. Operating costs are zero. Total demand for traveling on the road is given by

$$D(P) = a - bP$$

where

$$P = \sum_{i=1}^{N} p_i.$$

Here the n products are perfect complements: the consumer buys all n goods or none. Given the tolls charged by the other owners, $p_{-i} = (p_1, .., p_{i-1}, p_{i+1}, ..p_N)$, individual i's profits are

$$\pi(p_i, p_{-i}) = p_i(a - bp_i - b\sum_{i \neq j} p_j).$$

Differentiating with respect to p_i and setting the equation equal to zero yields the best reply function for individual i:

$$p_i = \frac{a - b\sum_{i \neq j} p_j}{2b}.$$

Assuming symmetry, the equilibrium price for each individual is

$$p^* = \frac{a}{(N+1)b},$$

which yields equilibrium market price and quantity

$$P^* = \frac{aN}{(N+1)b}, \quad Q^* = \frac{a}{N+1}.$$

Equilibrium profit of each individual is

$$\pi^* = \frac{a^2}{b(N+1)^2}.$$

In this equilibrium, price is above the monopoly price, since P^* increases in N. In fact, in the limit, as $N \longrightarrow \infty$, $P^* \longrightarrow a/b$ and $Q^* \to 0$. Here consumers are better off with a monopoly than with competition!

To understand the intuition behind this result, you should think of $Q(P)$ as the size of the pie and $p_i Q(P)$ as individual i's slice of the pie. At the monopoly price, each individual has an incentive to increase its toll because the first-order effect of a very small increase in p_i on Q, the size of the pie, is zero but the gain in its share of the pie is positive. If there is a lot of individuals trying to increase their share of the pie, they end up destroying the pie.

Applications of this model: connecting flights between domestic carriers on international flights, game console + games, hardware + software.

1.3 The Hotelling Model

The above model illustrates how product differentiation allow firms to escape the Bertrand trap. We now develop a model in which product differentiation is the outcome of a product location game. Consider a beach or street of unit length (say a mile) with a vendor located at each end. Vendor 1 is located at the left endpoint, which we denote as 0, and vendor 2 is located at the right endpoint, which we denote as 1. Each vendor i posts a price p_i at which it is willing to sell. For simplicity, marginal costs of production are zero. Location costs are fixed and sunk.

M potential consumers are distributed continuously and uniformily along the beach. The uniform distribution means that the number of customers in any section of the beach of length z (measured as fraction of a mile) is equal to $M * z$. (Often, M is normalized to be equal to 1). Each consumer is indexed by her location on the beach relative to 0. Thus, consumer x is the person who is located x miles from 0, where x lies between 0 and 1. Each consumer wants only one unit. The utility of consumer x is given by

$$u(x) = \begin{cases} s - p_1 - tx^2 & \text{if x purchase from vendor 1} \\ s - p_2 - t(1-x)^2 & \text{if x purchases from vendor 2} \\ 0 & \text{otherwise} \end{cases}$$

Here transport costs are assumed to be quadratic in distance. In most textbooks, transport costs are modeled as proportional to distance but this model has problems when you try to vary the locations of the vendors.

Let $\widetilde{x}$ denote the marginal consumer. Given the locations of the vendors, at any pair of prices (p_1, p_2), $\widetilde{x}$ is <u>either</u> someone who is indifferent between buying from vendor 1 <u>or</u> between buying from the lowest cost vendor and not buying at all.

Case 1: Low demand, Local monopoly

Suppose there is no overlap in the market coverage of the two vendors when they set price equal to the monopoly price. In that case, there is an interval of consumers who do not purchase from either vendor and the marginal consumer for each vendor is someone who is indifferent between buying and not buying from them. Let $\widetilde{x}$ denote the marginal consumer for vendor 1 (by symmetry, it will be $1 - \widetilde{x}$ for vendor 2). By definition, $\widetilde{x}$ is determined by the following equation:

$$s - p_1 - t\widetilde{x}^2 = 0 \implies \widetilde{x} = \sqrt{(s - p_1)/t}$$

Vendor 1's problem is to choose price to maximize

$$\pi(p_1) = p_1(\sqrt{(s - p_1)/t})M$$

The solution is $p_1^M = 2s/3$. Consequently, if $s < 3t/4$, then $\widetilde{x} < 1/2$ and each vendor enjoys a local monopoly.

Case 2: High Demand, Duopoly

Suppose $s > 5t/4$. If the two vendors try to set price equal to the monopoly price, everyone buys from either vendor 1 or vendor 2 - there is no one who does not buy. In this case, the market is *covered* and the marginal consumer is someone who is indifferent between buying from vendor 1 or buying from vendor 2. Let $\widehat{x}$ denote her location. It solves the following equation:

$$s - p_1 - t\widehat{x}^2 = s - p_2 - t(1 - \widehat{x})^2 \implies \widehat{x} = (p_2 - p_1 + t)/2t$$

Notice that the location of the marginal consumer depends upon the prices of the two vendors. They compete for this person. Demand for good 1 is $q_1 = \widehat{x}(p_1, p_2)$. Note that the cross-price effect $c = 1/2t$. Vendor 1's problem is to choose price to maximize

$$\pi_1(p_1) = p_1\widehat{x}(p_1, p_2)M = p_1[(p_2 - p_1 + t)/2t]M$$

His best reply function is given by

$$p_1 = (p_2 + t)/2.$$

Vendor 2's problem is to choose her price to maximize

$$\pi_2(p_1) = p_2(1 - \widehat{x})M = p_2[(p_1 - p_2 + t)/2t]M$$

with best reply function

$$p_2 = (p_1 + t)/2$$

The Nash equilibrium is $p_1 = p_2 = t$ and profits are $\pi_1 = \pi_2 = t/2$. Prices exceed marginal cost by an amount equal to the transport cost, a measure of product differentiation, and firms escape the Bertrand trap.

1.3.1 Extensions

Different Locations: Suppose the two vendors are located at 1/4 and 3/4. How do the closer locations affect prices and profits? Assume that s is high enough that the market is covered and that transport costs are proportional to distance traveled. Then the location of the marginal consumer solves

$$s - p_1 - t(\widehat{x} - 1/4)^2 = s - p_2 - t(3/4 - \widehat{x})^2 \implies \widehat{x} = \frac{1}{2} + \frac{p_2 - p_1}{t}.$$

Vendor 1's problem is to choose p_1 to maximize its profits which are

$$\pi_1(p_1) = p_1 \widehat{x}(p_1, p_2) M = p_1 \left(\frac{1}{2} + \frac{p_2 - p_1}{t} \right) M.$$

Differentiating and solving for its best reply yields

$$p_1 = \frac{p_2}{2} + \frac{t}{4}.$$

Vendero 2's problem is to choose p_2 to maximize its profits which are

$$\pi_2(p_2) = p_2(1 - \widehat{x}(p_1, p_2)) M = p_2 \left(\frac{1}{2} + \frac{p_1 - p_2}{t} \right) M.$$

This is the same as Vendor 1's problem so its best reply will also be

$$p_2 = \frac{p_1}{2} + \frac{t}{4}.$$

The Nash equilibrium is $p_1 = p_2 = t/2$ and profits are $\pi_1 = \pi_2 = t/4$. As products locate closer to each other in the product space, they become closer substitutes, the cross-price effects increase, and equilibrium prices fall. In the polar case where both firms locate at 1/2, prices and profits are zero.

More Vendors: Suppose the beach has three vendors: vender 1 is located at 0, vender 2 at 1/2, and vendor 3 at 1. What is the impact of more vendors on prices and profits? In this case, there are two marginal consumers: $\widehat{x}$, who is indifferent between buying from vendor 1 or vendor 2 and $\widehat{y}$, who is indifferent between buying from vendor 2 and vendor 3. The location of these two marginal consumers are determined by these two indifference conditions:

$$s - p_1 - t\widehat{x}^2 = s - p_2 - t(1/2 - \widehat{x})^2 \implies \widehat{x} = \frac{1}{4} + \frac{p_2 - p_1}{t}$$
$$s - p_2 - t(\widehat{y} - 1/2)^2 = s - p_3 - t(1 - \widehat{y})^2 \implies \widehat{y} = \frac{3}{4} + \frac{p_3 - p_2}{t}.$$

Vendor 1 chooses p_1 to maximize

$$\pi_1(p_1) = p_1 \widehat{x}(p_1, p_2) M = p_1 \left(\frac{1}{4} + \frac{p_2 - p_1}{t} \right) M.$$

Vendor 2 chooses p_2 to maximize

$$\pi_2(p_2) = p_2[\widehat{y}(p_2, p_3) - \widehat{x}(p_1, p_2)]M = p_2 \left(\frac{1}{2} + \frac{(p_3 - p_2)}{t} - \frac{(p_2 - p_1)}{t} \right) M.$$

Vendor 3 chooses p_3 to maximize

$$\pi_1(p_1) = p_1(1 - \widehat{y}(p_1, p_2))M = p_1 \left(\frac{1}{4} + \frac{p_3 - p_2}{t} \right) M.$$

The best replies are easily computed to be as follows:

$$\begin{aligned} p_1 &= \frac{p_2}{2} + \frac{t}{8} \\ p_2 &= \frac{p_3 + p_1}{4} + \frac{t}{8} \\ p_3 &= \frac{p_2}{2} + \frac{t}{8}. \end{aligned}$$

The Nash equilibrium prices are $p_1 = p_2 = p_3 = t/4$. Here $\widehat{x} = 1/4$ and $\widehat{y} = 3/4$. Profits are $\pi_1 = \pi_3 = t/16$ and $\pi_2 = t/8$. Note that with more than two vendors, each vendor competes directly only with its immediate neighbours but not with vendors further along the street. In that sense, competition is local.

Multi-Product Vendors: Suppose vendor 1 is located at 0 and vendor 2 has ice cream carts at locations 1/2 and 1. How would the equilibrium prices change, if at all? (See problem set.)

Endogenous Location: Suppose the vendors play a two-stage game in which they first choose locations (i.e., pay the fixed cost associated with location) and then compete in prices. Clearly, in making their location choices, the vendors will want to anticipate the kind of prices that will occur in the second stage. Two forces affect location decision: (1) locate where the demand is $\Longrightarrow$ move to the center and (2) stay away from the competition $\Longrightarrow$go to the ends. Which force dominates depends upon the transport cost function. In quadratic case, the second force dominates $\Longrightarrow$ maximal differentiation. In the linear case, the first dominates $\Longrightarrow$ minimal differentiation. However, the latter leads to existence problems (Illustrate).

Sequential and Repeated Games

March 6, 2007

1 Introduction

In this lecture, I want to discuss the impact of repeated interaction on the behavior of firms. In order to do so, we need to develop and master some more tools from game theory. I begin by discussing how to formulate a game in which players move sequentially and how to solve it.

1.1 Sequential Move Games

Consider first an entry game. Firm 1 is a potential entrant. Firm 2 is an incumbent firm. If the entrant stays out, firm 1 gets its normal return of \$1 million from an alternative investment, the incumbent is a monopolist and can earn \$5 million. If firm 1 enters, they compete for consumers. If firm 2's response to entry is to fight, each firm gets 0; if firm 2's response is to accommodate entry, then each firm gets \$2 million. The game in normal form can be described as follows:

$$\begin{array}{ccc} & Fight & Acquiese \\ Enter & 0,0 & 2,2 \\ Stay\ out & 1,5 & 1,5 \end{array} \tag{1}$$

This game in normal form has two Nash equilibria: (Stay Out, Fight) and (Enter, Acquiese). But the first equilibrium makes no sense. Firm 1 stays out because firm 2 threatens to fight if it enters. But, if firm 1 does enter, it is not in firm 2's best interest to carry out the threat since acommodating entry yields a payoff of 2 and fighting generates a payoff of 0. The threat is not credible: firm 2 will not carry out the threat if called upon to do so. Hence firm 1 should not believe the threat but enter.

This example illustrates a flaw in Nash equilibrium. In multi-stage games, Nash equilibria does not rule out play based on incredible threats. This issue has led economists to study games in extensive form rather than normal form. (Present the tree diagram and solve the game backwards.) The solution concept that is used for the game in extensive form is known as subgame perfection: play in every subgame must be a Nash equilibrium. The subgame perfect equilibrium for the above game is unique: it is (Enter, Acquiese). Every subgame perfect equilibrium is a Nash equilibrium but, as we have seen above, not every Nash equilibrium is subgame perfect.

1.1.1 Stackelberg in Quantities

Consider next the Cournot model in which two firms compete in a homogenous good market. However, instead of choosing quantities (or capacities) simultaneously, let us assume that the two firms choose their quantities sequentially with firm 1 chooses its capacity first and firm 2 chooses its capacity after observing firm 1's capacity choice. The game in normal can be represented as follows:

- Player Set: $i = 1, 2$
- Strategies: $y_1 \in [0, \infty)$; $y_2 = y_2(y_1)$.
- Payoffs for firm 1 and firm 2:

$$\pi_1(y_1, y_2(y_1)) = P(y_1 + y_2(y_1))y_1 - C(y_1)$$

$$\pi_1(y_1, y_2(y_1)) = P(y_1 + y_2)y_2 - C(y_2)$$

Interpretation: firm 1 decides what quantity to produce and supply to the market, then firm 2 decides what quantity to produce and supply the market. Given the amounts supplied, price adjusts so as to clear the market.

Assume demand is given by

$$\begin{aligned} P(Y) &= a - bY, \\ Y &= y_1 + y_2 \end{aligned}$$

and technology by

$$C(y_i) = cy_i, i = 1, 2.$$

Now consider firm 2's optimization problem. When firm 2 contemplates what to supply, it knows what firm 1 is going to supply. Hence, anticipating the price formation process, firm 2 chooses y_2 that maximizes its profit

$$\pi_2(y_1, y_2) = [a - b(y_1 + y_2) - c]y_2.$$

Notice that its demand curve is essentially the market demand with the intercept equal to $a - by_2$ instead of a. (Illustrate.) Differentiating profits, setting the derivative equal to zero, and solving for y_2 implies that

$$y_2 = (a - c - by_1)/2b.$$

The above relationship is known as firm 2's *best reply.*

Now consider firm 1's problem. In choosing its output, it has to anticipate how firm 2 will respond to its choice. Thus, it chooses y_1 to maximize

$$\pi_2(y_1, y_2(y_1)) = [a - by_1 - ((a - c - by_1)/2b) - c]y_1$$

Differentiating and solving for y_1 yields

$$y_1^* = \frac{a-c}{2b}.$$

Substituting this solution back into firm 2's best reply yields

$$y_2^* = \frac{a-c}{4b}.$$

The equilibrium price is

$$p^* = \frac{1}{4}(a+3c).$$

Note that total output is higher in the sequential move game than in the simultaneous move game. Here it is $\frac{3(a-c)}{4b}$; in the Cournot equilibrium, it was $\frac{2(a-c)}{3b}$. Hence, price is lower in the sequential move game than in the Cournot equilibrium. The leader has a larger market share and greater profits. This finding illustrates a more general result: in a game with strategic substitutes, the first-mover has an advantage.

Example 1 *Suppose $a = 120, b = 20, c = \$20$. Then firm 2's best reply is given by*

$$y_2 = \frac{100}{40} - \frac{y_1}{2}.$$

Solving for the Nash equilibrium,

$$y_1^* = \frac{5}{2}, y_2^* = \frac{5}{4}.$$

Equilibrium market output is $\frac{15}{4}$ and market price is 45.

1.1.2 Stackelberg in Prices

Consider next the Hotelling model in which firm 1 sets its price first and then firm 2 sets its price. We assume that s is large enough that the market is covered. Recall that, in this case, the marginal consumer is someone who is indifferent between buying from vendor 1 or buying from vendor 2. Let $\widehat{x}$ denote her location. It solves the following equation:

$$p_1 + t\widehat{x}^2 = p_2 + t(1-\widehat{x})^2 \implies \widehat{x}(p_1, p_2) = (p_2 - p_1 + t)/2t$$

Notice that the location of the marginal consumer depends upon the prices of the two vendors. Now consider vendor 2's problem. It knows the price that vendor 1 has set. Therefore, given p_1, it chooses its price to maximize

$$\pi(p_2) = p_2(1 - \widehat{x}(p_1, p_2))M = p_2[(p_1 - p_2 + t)/2t]M$$

Differentiating with respect to p_2, setting the derivative to zero, and solving for p_2 yields

$$p_2(p_1) = (p_1 + t)/2.$$

How about vendor 1's problem? She has to anticipate vendor 2's response. Hence, she chooses her price to maximize

$$\pi(p_1) = p_1 \widehat{x}(p_1, p_2(p_1))M = p_1[(\frac{p_1 + t}{2} - p_1 + t)/2t]M = p_1[\frac{3}{4} - \frac{p}{4t}]M.$$

Differentiating and solving yields

$$p_1^* = \frac{3t}{2}.$$

Substituting into vender 2's best reply yields

$$p_2^* = \frac{5t}{4}.$$

The location of the marginal consumer is

$$\widehat{x}^* = \frac{3}{8}.$$

Thus, prices in this game are higher than in the simultaneous move game. The price leader has a higher price, lower market share, and lower profits. This finding illustrates a more general result: in a game of strategic complements, the second-mover has an advantage.

1.2 Repeated Play

A group of firms who have agreed to coordinate pricing and output in order to exercise market power are termed a *cartel* (e.g., OPEC). Antitrust action usually requires evidence of such explicit pricing agreements. The main goal of this lecture is to show that such agreements are not required for firms to collude and obtain monopoly profits. They can recognize their mutual interdependence and the advantages of raising prices and engage in tacit collusion for which there is no "smoking gun" evidence.

Suppose two firms play the following Cournot game twice. The inverse demand curve is given by

$$P(y) = 70 - y$$

and firm i's cost function is given by

$$C(y_i) = 10y_i, i = 1, 2.$$

In the symmetric Cournot equilibrium to the one-shot game, outputs are

$$y_1^* = y_2^* = 20, p^* = 30$$

and each firm's profit is 400. The monopoly output is 30, which yields a total profit of 900.

A strategy for a firm in this two period game is a contingent plan which specifies (i) a level of output in period 1 and (ii) output in period 2 as a function of the outputs chosen by both firms in period 1. An example is the following:

"in period 1, choose 15 units; in period 2, choose 20 units if both firms produced 15 units in period 1 and 60 units otherwise."

Call the strategy given above the "C" strategy for cooperation. If both players play this strategy, then each will earn (assuming no discounting) profits of $450 + 400 = 850$. But is (C,C) a Nash equilibrium? To answer this question, we need to evaluate the best possible deviation and compare its profitability to that of C.

Best deviation to C:

Note first that if both firms choose 15 in period 1, neither firm has any incentive to deviate from 20 in period 2 since this is a Nash equilibrium in the period 2 stage game. Therefore, if a firm is going to deviate, it should do so in period 1. Fixing firm j's strategy at "C", it is easily shown that firm i's optimal deviation in period 1 is to choose 22.50 units, which earns a profit of 506.25. In period 2, firm i's best reply to $y_j = 60$ is to produce 0 Call this strategy the "D" strategy. Aggregate profits from "D" is 506.25.

Conclusion 1 *(C,C) is a Nash equilibrium.*

Is this equilibrium plausible? Firm i is deterred from deviating in period 1 by firm j's threat to flood the market in period 2. But is this threat credible? That is, will firm j carry out its threat if called upon to do so? The answer has to be no. Flooding the market and driving price down to marginal cost following a deviation by firm i in period 1 cannot be rationalized under any conjecture about firm i's choice of output in period 2. In fact, if firm j truely expects firm i to produce zero, firm j should deviate from C and produce the monopoly output. But if firm j has no incentive to follow through on its threat, why should firm i believe it? Clearly, it should not.

We need to find a solution for the repeated game which rules out strategies that involve incredible threats. The way we do this is to require that the strategies represent a Nash equilibrium for every subgame. We call such a solution the subgame perfect equilibrium. In the context of a repeated game, each period defines a new subgame. Thus, the number of subgames is equal to the number of periods.

Applying this concept to the two period repeated game given above implies that firms have to choose an output of 20 in period 2, *regardless of what happened in period 1.* The output pair (20,20) is the unique Nash equilibrium for the second period subgame. But, if both firms choose 20 in period 2 independently of what is played in period 1, then equilibrium play in period 1 is also 20.

Conclusion 2 *The unique subgame perfect equilibrium in a finitely repeated Cournot stage game is to play the Cournot quantity in every period.*

1.3 The Infinitely Repeated Game

Suppose the two firms live forever so the Cournot stage game is repeated an infinite number of times. For simplicity, let us restrict the set of outputs that each firm can choose in any period to $\{15, 20, 22.5\}$. The payoffs to the stage game are:

	$y_2 = 15$	$y_2 = 20$	$y_2 = 22\frac{1}{2}$
$y_1 = 15$	$450, 450$	$375, 500$	$337\frac{1}{2}, 506\frac{1}{4}$
$y_1 = 20$	$500, 375$	$400, 400$	$350, 393\frac{3}{4}$
$y_1 = 22\frac{1}{2}$	$506\frac{1}{4}, 337\frac{1}{2}$	$393\frac{3}{4}, 350$	$337\frac{1}{2}, 337\frac{1}{2}$

A strategy for the stage game is simply a choice of one of these three outputs. The unique Nash equilibrium is, of course, (20,20).

A strategy for the infinitely repeated game is a plan which specifies for each period t, a choice of one of the three outputs as a function of the history of outputs chosen in prior periods. A pair of strategies generates a pair of outputs for each period $\{(y_{1,t}, y_{2,t})\}_{t=0}^{\infty}$. Firm i's evaluation of this stream of outputs is given by the discounted sum

$$\prod_i = \sum_{t=0}^{\infty} \rho^t \pi_i(y_{1,t}, y_{2,t}).$$

The Trigger Strategy

Consider the following plan for firm i, which we will call the "T" strategy:

> "In each period t, play 15 if both firms have played 15 in every prior period; otherwise play 20."

That is, firm i cooperates as long as both firms have cooperated in the past. However, if anyone has deviated, then firm i plays the Cournot quantity forever.

Now suppose both firms play the T strategy. Is it an equilibrium? The profits from the T strategy is

$$\prod_i^C = \sum_{t=0}^{\infty} \rho^t (450) = 450/[1-\rho].$$

Now suppose firm i deviates in period t and chooses $y_{i,t} = 22.5$. Then its payoff from that period forward is

$$\prod_i^D = 506\frac{1}{4} + \sum_{t=1}^{\infty} \rho^t (400) = 506\frac{1}{4} + 400\rho/[1-\rho].$$

Clearly, firm i has no incentive to deviate if $\prod_i^D < \prod_i^C$. This inequality holds true if firm i is not too impatient. More precisely, $\rho > .529$.

Conclusion 3 *If firms are not too impatient and they cannot envision a last period, then collusion is a noncooperative equilibrium outcome.*

Vertical Integration and Restraints

April 4, 2007

A firm is vertically integrated if it engages in more than one successive stage of the production and distribution process. For example, a fast food chain that produces the products and later distributes them in its outlets is vertically integrated in comparison to a fast food restaurant that buys its food products from independent suppliers.

1 Reasons for Vertical Integration

1.1 Elimination of double monopoly markup

Consider a monopoly that produces its product y using a single input x. The technology requires one unit of input for each unit of output (i.e., $y = x$). Thus, if the input's price is c per unit, the monopoly's production cost will be $C(y) = cy$. Suppose that the input is produced at a cost of w per unit, and is sold by a competitive industry. Does the monopoly have an incentive to vertically integrate (backward) with the input selling industry? No, because in a competitive industry, $c = w$ and this will be its cost if it produces the input for itself.

Suppose the input industry is competitive but operates under conditions of rising marginal costs. Does the monopoly have an incentive to vertically integrate in this case? Again, the answer is no. After such integration, a profit maximizing integrated firm will produce exactly the same quantity that the monopoly would have bought from the competitive industry. The integrated firm will own the profit of the competitive industry, but this is presumably the price it would have had to pay for this industry.

Suppose next that the upstream industry is also a monopoly and that it determines the input's price unilaterally. The input is produced at a constant average cost w. The final (inverse) demand is given by

$$P(y) = a - by.$$

The downstream monopolist chooses its output to

$$\max \pi_D(y) = (a - by)y - cy.$$

Differentiating with respect to y and setting the derivative equal to zero yields

$$a - 2by - c = 0$$

or

$$y = \frac{a-c}{2b}.$$

This equation describes how the downstream monpolist supply varies with the input price c. Since $x = y$, this equation is also the downstream monopolist's input demand function. Inverting this equation yields the inverse demand function for x,

$$c(x) = a - 2bx.$$

The upstream monopolist chooses its output to

$$\max \pi_U(x) = (a - 2bx)x - wx.$$

Differentiating with respect to x, setting the derivative equal to zero, and solving for x yields

$$x^* = \frac{a-w}{4b}.$$

Substituting this value of x into the inverse demand for x yields the equilibrium input price,

$$c^* = (a + w)/2$$

and the final product's price,

$$P^* = (3a + w)/4.$$

The profits of the upstream and downstream monopolists are

$$\pi_U^* = \frac{(a-w)^2}{8b}, \pi_D^* = \frac{(a-w)^2}{16b}.$$

The combined profit of both upstream and downstream industries is $3(a - w)^2/16b$.

Now suppose the two firms merge. In the case, the vertically integrated firms simply transfers from the upstream division to its downstream division at $c = w$ and chooses y to

$$\max \pi_{VI}(y) = (a - by)y - wy.$$

Differentiating, setting the derivative equal to zero, and solving for y yields

$$\widetilde{y} = \frac{a-w}{2b}.$$

The equilibrium price is

$$\widetilde{P} = \frac{a+w}{2},$$

and profits are

$$\pi_{VI} = \frac{(a-w)^2}{4b}.$$

Thus, output is higher, price is lower and profits are higher after the merger. This difference is not surprising since by vertically integrating, the industry eliminates the deadweight loss that accompanies the upstream monopoly pricing. (Illustrate with graphs.)

Conclusion 1 *Upstream monopoly power creates an incentive for vertical integration that will eliminate the double monopoly markup.*

Notice that this model does not really distinguish between integration and other contractual arrangements. For example, the elimination of the double marginalization can be achieved through a two part tariff rather than integration. A two-part tariff is like a franchisee fee. In the above example, the upstream monopolist should charge $c = w$ per unit of input and a franchise fee

$$F = \frac{(a-w)^2}{4b}.$$

It could also be eliminated using a retail price maintenance agreement. In the above example, the upstream monopolist could impose a price ceiling on the retailer that it cannot charge a price higher than $\widetilde{P}$. Since the retailer has no reason to charge a lower price, it will set $P = \widetilde{P}$ and sell $\widetilde{y}$ units if the wholesale price $w \leq \widetilde{P}$. The upstream monopolist can then set $w = \widetilde{P}$ and earn the same profits that it would if it were vertically integrated.

1.2 Transaction Costs

Another motive for vertical integration is the elimination of transaction costs that may arise in market relationships. For example, market transactions may require complicated contracts and some of the accompanying legal costs may be avoided in a vertically integrated firm. Of course, the operation of the integrated firm also involves transaction costs, but in some cases these may be lower.

Suppose for example that the quality of the input is difficult to monitor. The downstream firm may then want to produce this input in house. There is of course the problem of quality control inside the firm, which may involve monitoring the workers' efforts. However, if the quality of the input largely depends on the quality of the raw material which is easy to control, the firm may find it easier to monitor the input quality by vertically integrating and controlling the raw material supplies.

One of the problem which interferes with non-integrated structure is the hold-up problem. Suppose that a firm needs a special input which can be supplied by a supplier who is not part of the firm. The value of this input depends on whether the supplier makes some specific investment prior to production. The situation is described schematically as follows:

Investment	Value	Cost	Supplier Revenue	Supplier Profit
0	3	2	2.5	0.5
1	4	1.5	2.75	0.25

At date 0, the supplier has to make a decision of whether to invest (1) or not (0). At date 1, the input is supplied at a cost. If the supplier has not made any investment at date 0, then the value of the input is 3 and cost is 2, so the surplus is $(3-2) = 1$. Assume that the supplier and the buyer bargain and reach an agreement that split equally the surplus at date 1. Then the supplier's revenue is 2.5 and its profit is 0.5. Now suppose the supplier makes the prior investment. Then the value of the input is enhanced, increasing from 3 to 4.

In addition, costs fall from 2 to 1.5. Thus the surplus is greater and given by $4 - 1.5 = 2.5$. Once again, assume that the supplier and the firm split equally the existing surplus at date 1. Then each party gets 1.25, since at date 1 the specific investment is already sunk and does not affect the bargaining. The supplier's total revenue is 2.75 and its net profit is 0.25. In that case, the supplier gets a profit of 0.5 if it does not make the investment and 0.25 if it does make the investment. Hence it is better off not making the investment even though from the point of view of their combined surplus, the investment is desirable. The problem here is that, because the investment is sunk, the supplier cannot recover it in the subsequent bargaining session with the firm. This problem could be avoided if the two firms can write a contract which conditions the price on the level of investment or the value of the input. But these magnitudes might be impossible or difficult to verify to a third party, such as a court, thus rendering the contract unenforceable. If this is the case, the firm may want to overcome the hold-up problem through integration.

1.3 Monopolization

Sometimes vertical integration monopolizes an industry and hence increases profits for the same reasons that horizontal integration does.

Suppose the downstream industry is competitive and the upstream industry is Cournot duopoly. For concreteness, assume as above that the final product's demand is

$$P(y) = a - by,$$

that one unit input makes a unit output, and that the input is produced at a constant per unit cost w. The downstream industry's (inverse) demand for the input is

$$c = a - bz,$$

where c and z are the input's price and quantity respectively. Each of the upstream firms earns a duopolistic profit which in this case is $(a - w)^2/9b$. Now, if one of the upstream firms vertically integrates forward and acquires the entire downstream industry, it becomes a monopoly, and its profit will be $(a-w)^2/4b$. Notice that in this case the increase in profit is not due to the vertical integration per se but rather due to the horizontal integration of the downstream industry that took place in the process.

1.4 Price discrimination

Suppose that an upstream monopoly sells to two downstream industries with different demands. The demands are such that the monopoly would like to price discriminate between these two industries, but there is no easy way to exercise such discrimination due to inability to prevent resale. By vertically integrating into one of these industries, the resale problem is resolved.

1.5 Internalizing Externalities

Consider a manufacturer who distributes its product through a network of independent distributors. Suppose that the distributors are responsible for the sales promotion. If there is spillover so that one distributor can free ride on the advertising done by others, these distributors may invest less than optimal effort in promotion. The manufacturer then has an incentive to vertically integrate and take this responsibility upon itself. This is of course true not only for advertising but other forms of sales effort or reputation maintenance.

2 Vertical Restrictions

An alternative to vertical integration is for manufacturers impose on distributors restrictions in an attempt to capture some of the benefits. These restrictions include exclusive territories, resale price maintenance (RPM), exclusive dealing sales quotas.

2.1 RPM

We have already seen how RPM can replicate the vertically integrated outcome. However, we observe RPM arrangements even if the downstream market is competitive. In fact, in many cases, RPM sets a price floor (i.e., $P \geq \widetilde{P}$) rather than a price ceiling. The following simple model illustrates why this may be the case and why RPM constracts can actually benefit consumers. An upstream firm U produces product at zero cost. It sells the product at price c to downstream competing duopolists 1 and 2. The final demand for the product is inelastic at quantity M up to price 1. Bertrand competition between the duopolists results in price $p_1 = p_2 = c$. The manufacturer's optimal strategy is to set $c = 1$ and earn profits of M.

Now suppose that each duopolist can advertise and thus increase demand. If firm i chooses advertising volumeA_i, its cost will be A_i^2 and demand increases from M to $M + A_i$. Therefore, if firms 1 and 2 choose advertising volumes A_1 and A_2, demand for firm i will be

$$D_i(p_i, p_j) = \begin{cases} M + A_1 + A_2 & \text{if } p_i < p_j \text{ and } p_i \leq 1 \\ (M + A_1 + A_2)/2 & \text{if } p_i = p_j \leq 1 \\ 0 & \text{otherwise.} \end{cases}$$

In the absence of RPM, the equilibrium price in the downstream industry is $p_1 = p_2 = c$ and $A_1 = A_2 = 0$.

Now suppose U makes an agreement with its dealers to fix the downstream price p at 1. Then i's problem is to

$$\max_{A_i} (1 - c)(M + A_i + A_j)/2 - (A_i)^2$$

The solution is

$$A_i = (1 - c)/4.$$

U's profit with RPM is thus

$$\pi(c) = c[M + (1 - c)/2].$$

Maximizing $\pi(c)$ with respect to c yields the solution

$$c = \begin{cases} M + 1/2 & \text{if } \; M + 1/2 < 1 \\ 1 & \text{otherwise} \end{cases}$$

The main point here is that RPM may be profitable for the upstream monopoly, who might want to choose $c < 1$ and employ RPM to create the right incentives for the downstream firms to promote the product.

Of course, if U were vertically integrated with the dealers, then it could increase its profits even further by choosing $A_1 = A_2 = 1/2$. Thus, the relevance of RPM contracts assumes that vertical integration is not possible for legal or organizational reasons.

A similar issue arises when A_i is interpreted as service (i.e., providing information about the products) provided by retailer i. Even if consumers go to the retailer that has a more informative sales staff, they buy from the retailer that offers the lowest price. In the absence of a RPM, the equilibrium consists of both retailers offering zero service, which is not good for either consumers or the manufacturer.

2.2 Exclusive Territories

If the retailers are geographically differentiated and manufacturers use exclusive territories agreements to prevent competition between retailers, the same ambiguity arises as in the case of RPM. For example, consider a Hotelling model in which buyers either live in Town R, which is located at the right end point of the line, or at town L, which located at the other end of the segment. The towns are the same size. Travel costs are zero and demand is the same as in the RPM model given above where A is service. The manufacturer's costs are also the same. Thus, the Bertrand equilibrium is for each retailer to provide no service and set its price equal to c and sell only to consumers living in its town. However, if the manufacturer does not allow the retailers to sell to buyers in the other town, then each retailer can increase demand by providing service and will do so if $c < 1$..If the increase in demand is large enough, the optimal wholesale price for the manufacturer is to lower its price below 1 and for the retailers to provide service

3 Conclusion

The conclusion of this lecture is that vertical integration or vertical restraints need not be detrimental to welfare. This is in contrast to horizontal integration which always results in increased monopoly power. Vertical integration may decrease welfare when it increases monopoly power because it is accompanied by some sort of horizontal integration with or without disguise.

Thus, vertical integration and restraints need to be evaluated on a rule of reason criterion rather than a per se criterion. The main objective of antitrust policy should be to prevent increased monopolization but to abstain from intervention otherwise. This objective has not always been achieved with cases involving vertical integration or restraints.

Static Models of Market Entry

April 11, 2007

1 Theory of Entry Deterrance and Accommodation

The existence of rents from market power will attract entry and, in the absence of effective entry deterrence, new firms will enter and compete away the economic profits of the incumbents. In fact, in a world in which entry is easy (i.e., not costly), the issues of market power and anticompetitive practises disappears. Thus, from an antitrust perspective, the question of entry deterrence is central to a determination of market power. Our focus in this lecture is to identify conditions under which an incumbent firm can expand capacity to deter entry or at least limit the size at which an entrant enters.

The prototypical case of DuPont. In 1972 Dupont was the largest producer of a chemical agent known at titanium dioxide which is used to whiten paper and paint. It had approximately 35% of industry capacity, most of which used its proprietary chloride process. In the early 1970s two exogenous shocks gave DuPont's proprietary process a considerable cost advantage over the other two technologies that were available. One was the adoption of stricter pollution controls which threatened the viability of the sulfate process. The second was that the price of the raw material used by the other chloride process tripled. This gave DuPont a substantial cost advantage over its rivals.

The management of DuPont discussed two strategies for responding to their advantageous position. The "maintain status quo" strategy called for increasing DuPont's market share as the sulfate process exited the industry from 35% to 40% by 1985, investing $192 million in new capacity and no real change in prices. The "growth" strategy was based upon exploiting the strategic advantage. The ideas was to invest in $394 million in new capacity, building a new plant and expanding several existing plants, and expanding capacity to 65%. The management team made estimates about several limit prices that would trigger competitors expansions and imports. They proposed a pricing strategy that was high enough to generate the cash flow needed to finance the expansion but low enough to discourage entry. Dupont estimated the "growth" strategy to be more profitable and pursued it.

By 1979, Dupont had acheived its goal and had 60% of the capacity. The FTC filed a complaint charging DuPont with monopolizing the market by engaging in limit pricing and holding excess capacity. The court dismissed the charges, primarily because the FTC failed to demonstrate that DuPont had invested in excess capacity. Subsequent econometric analysis by Hall found that DuPont's investments reduced its short-run marginal costs and that strategic effect limited the output and capacity expansion of its rivals and prevented entry. However, DuPont's expansion did not involve excess capacity.

There are two issues that we need to address. First, does it make sense for an incumbent to invest in excess capacity, which it then holds in reserve until entry or expansion by rivals? In this case, the excess capacity is a warning to rivals that entry or expansion will be met by an aggressive price war. The second issue is whether an incumbent will overinvest in capacity but use the capacity even if entry does not occur.

1.1 Capacity Deterrence

Consider a homogenous good industry with inverse demand

$$P(y) = 1 - y,$$

where y denotes industry output. Firm 1 has been servicing this industry as a monopolist because of a patent. However, its patent is about to expire and a rival, firm 2, is poised to enter when this happens. Can it use its capacity to deter entry?

To produce one unit of output requires one unit of capacity. The cost of a unit of capacity is $c < 1/2$. Prior to the entry decision of firm 2, firm 1 chooses a level of capital k_1 which is then fixed. Firm 2 observes k_1 and then decides whether to enter or not. Entry costs are denote by F. If firm 2 decides to entry, it chooses a level of capital k_2, which is also fixed. Given their capacities, each firm then chooses output simultaneously subject to the restriction that $y_i \leq k_i$. Production costs are zero. We assume without loss of generality that firm 1 cannot augment its capital stock.

Suppose any indivisibilities in capital are negligible relative to demand so k_i is any real number in the interval $[0, 1]$. In this case, it makes sense to assume that the post-entry game is Cournot. We study the equilibrium to this game by working backwards.

Suppose firm 1 has invested in capacity k_1 and firm 2 has entered (i.e., paid F). Firm 2 has to choose its capacity and output. Since it has no incentive to invest in more capacity than it needs to produce a given level of output, we may assume that $y_2 = k_2$. Firm 2 then chooses y_2 to

$$\max_{y_2} \pi_2(y_1,y_2) = (1 - c - y_1 - y_2)y_1.$$

Solving this optimization problem yields firm 2's best reply function,

$$R_2(y_1) = (1 - c - y_1)/2$$

Firm 1 chooses its output to

$$\max_{y_1} \pi_1(y_1,y_2) = (1 - y_1 - y_2)y_1 \quad \text{s.t.} \quad y_1 \leq k_1$$

This yields firm 1's best reply function,

$$R_1(y_2) = \begin{cases} (1 - y_2)/2 & \text{if } (1 - y_2)/2 < k_1 \\ k_1 & \text{if } (1 - y_2)/2 \geq k_1 \end{cases}$$

The important point to note about these best replies is that firm 1's best reply does not depend upon capacity costs. The reason is that its investment is sunk, so its marginal cost of production in the post-entry game is zero. By contrast, firm 2 has to invest in capacity to produce, so it faces a marginal cost of c per unit of output. As we shall see, firm 1 can use this asymmetry to deter entry.

If firm 1 is not capacity constrained, then the equilibrium to the post-entry game is given by

$$\begin{aligned} y_1^* &= \frac{1+c}{3} \\ y_2^* &= \frac{1-2c}{3}. \end{aligned}$$

If firm 1 is capacity constrained, then the equilibrium is

$$\begin{aligned} y_1^* &= k_1 \\ y_2^* &= \frac{1-c-k_1}{2}. \end{aligned}$$

Idle Threats Figure 1 depicts the best replies when $c = 1/5$ and $k_1 = 1/2$. The equilibrium to the post-entry game is $(y_1, y_2) = (2/5, 1/5)$, which yields a price of 2/5. Profits to firm 2 are 1/25; profits to firm 1 are 2/25. Therefore, as long as entry costs are sufficiently low, that is, $F < 1/25$, firm 2 enters and earns positive profits. Note, however, that firm 1 produces less than capacity. If firm 1 produces to its capacity of 1/2 after firm 2 enters, firm 2's best reply is 3/20. But, if firm 2 produces that much, then firm 1 will want to reduce its output to 17/40. Hence, firm 1's threat to produce up to capacity is not credible. Furthermore, firm 2 should anticipate this fact in making its entry decision.

This argument leads to a very important point. If firm 1 is not constrained in its output choice in the equilibrium of the post-entry game, it should reduce its capacity prior to firm 2's entry decision. The extra capacity is costly and has no benefit since it does not affect the entrant's decision and is not used following entry. In the above example, firm 1's equilibrium output is 2/5 but its capacity is 1/2. If it reduces $k_1$to 2/5, it lowers investment costs with no change in revenues. We therefore obtain the following conclusion:

> *In the absence of any commitment to producing to capacity, it is never optimal for firm 1 to invest in idle capacity.*

Therefore, in equilibrium, $y_1 = k_1$.

It is worth emphasizing that firm 1 may want to commit to producing to capacity if it could. For example, if it could credibly threaten to produce 1/2 in the post-entry game, firm 2's profits are much lower. The outcome to the post-entry game would be $(y_1, y_2) = (1/2, 3/20)$ which yields $p = .35$. Profits to firm 2 are 9/400. Hence, if $F > 9/400$, firm 2 would not enter in response to firm 1's threat and, in turn, firm 1 would earn monopoly profits of 1/4. But, how can firm 1 make its threat credible? There is no way in this model.

Equilibrium In what follows, we denote the post-entry equilibrium outputs as (y_1^*, y_2^*).

Case 1: High Entry Costs

Entry is said to be *blockaded* if the monopoly choice of capacity by firm 1 is sufficient to deter entry. The monopoly choice of output in our model is

$$k^m(c) = (1-c)/2.$$

To determine the value of F for which this choice blockades entry, we need to compute the profits of firm 2 in the post-entry equilibrium and check when they are less than F.

Define

$$\overline{k}(c) = (1+c)/3.$$

Recall that $\overline{k}$ is the maximum amount that firm 1 can credibly threaten to produce in the post-entry equilibrium. Note that $\overline{k}$ can be less than k^m. More precisely,

$$k^m \gtreqless \overline{k} \Leftrightarrow c \lesseqgtr 1/5.$$

Therefore, in evaluating the profitability of entry, we need to distinguish between two cases. When $c < 1/5$, firm 2 computes its profits from entry assuming that firm 1 produces $\overline{k}$, which is less than its capacity of k^m. When $c > 1/5$, firm 2 computes its profits from entry assuming that firm 1 produces to capacity, k^m, .which is less than $\overline{k}$.

The post-entry equilibrium can be characterized as follows:

$$y_1^* = \begin{cases} \overline{k} & \text{if } c < 1/5 \\ k^m & \text{if } c \geq 1/5 \end{cases}, \quad y_2^* = \begin{cases} (1-2c)/3 & \text{if } c < 1/5 \\ (1-c)/4 & \text{if } c \geq 1/5 \end{cases} \quad p^* = \begin{cases} (1+c)/3 & \text{if } c < 1/5 \\ (1+3c)/4 & \text{if } c \geq 1/5 \end{cases}$$

Therefore, entry is blockaded if

$$F > \begin{cases} (1-2c)^2/9 & \text{if } c < 1/5 \\ (1-c)^2/16 & \text{if } c \geq 1/5 \end{cases}$$

The equilibrium outcome when F satisfies the above inequality is that firm 1 invests in k^m, firm 2 does not enter, and firm 1 produces to capacity.

Case 2: Low Cost Advantage

Suppose F does not satisfy the above inequality for $c < 1/5$. Recall that the maximum output that firm 1 can threaten to produce in the post-entry equilibrium is $\overline{k}$. Since firm 2's profits at this output level exceed F, firm 1 cannot deter entry. Knowing this, how much capacity should firm 1 invest? Applying the no idle capacity result, the optimal choice of capacity for firm 1 is determined by the following problem:

$$\max_{k_1} \ \pi_1(k_1) = (1 - k_1 - c - R_2(k_1))k_1 \ \text{ s.t. } \ k_1 \leq \overline{k}.$$

The solution to the unconstrained problem is

$$k_1 = (1-c)/2,$$

which is, of course, the monopoly output, k^m. (This result is due to the linearity of the demand function and is not true in general.) But, for $c < 1/5$, $k^m > \overline{k}$. Hence, in this case, firm 1 should simply invest in capacity $\overline{k}$. Firm 2 enters with a capacity of $(1-2c)/3$, which is less than that of firm 1, and each firm produces to capacity.

Case 3: High Cost Advantage

Suppose F does not satisfy the above inequality and $c \geq 1/5$. Firm 1 cannot deter entry by choosing its preferred capacity of k^m. But, in this case, $k^m < \overline{k}$, so it can increase capacity beyond k^m (but not beyond $\overline{k}$) and credibly threaten to use that capacity in the event that firm 2 enters. Hence, in this range of capital costs, firm 1 may have two possible choices. One choice is to invest in capacity k^m, let entry occur at a capacity of $(1-c)/4$, and then produce to capacity. The other choice is to choose a higher level of capacity, k^d, deter firm 2 from entering, and then produce to capacity as a monopolist. Moreover, it will use that capacity in the event that firm 2 does not enter since $\overline{k} < 1/2$ for $c < 1/2$.

Here k^d is defined as the solution to the following equation:

$$\pi_2(k_1, R_2(k_1)) = [1-c-k_1-(1-c-k_1)/2](1-c-k_1)/2 = F$$

which is given by

$$k^d = 1 - c - 2\sqrt{F}.$$

Firm 1 will utilize this option when c is relatively large for then k^m is much smaller than $\overline{k}$ and an increase in output to $\overline{k}$ substantially lowers firm 2's profits.

Example 1 $c = 2/5$.

The post-entry equilibrium at $k_1 = k^m$ is

$$k^m = 3/10, \quad R_2(k^m) = 3/20, \quad p^* = 11/20, \quad \pi_2^* = 9/400, \quad \pi_1^* = 9/200$$

The post-entry equilibrium at $k_1 = \overline{k}$ is

$$\overline{k} = 7/15, \; R_2(\overline{k}) = 1/15, \; p^* = 7/15, \; \pi_2^* = 1/225, ..\pi_1^* = 7/225$$

Now consider an entry cost such as $F = 4/400$. It is less than $9/400$, the level required to blockade entry at k^m, and exceeds $1/225$, the entry-deterring level at $\overline{k}$. But firm 1 does not have to commit to a capacity of $\overline{k}$ to deter entry, it need only expand to k^d, which is $2/5$. If it does so, then firm 2 does not enter, firm 1 enjoys a monopoly. Output, price and profits are

$$y_1 = 2/5, \; p^* = 3/5, \pi_1^* = 2/25.$$

In this case, k^d yields higher profits for firm 1 than k^m.

1.2 Learning By Doing

In the 1970s, several consulting firms including Boston Consulting Group recommend to their clients that they should sacrifice short-run profits early in the product life cycle in order to gain a strategic advantage over rivals later in the cycle. The argument is that by cutting price and producing a lot early, a firm slides down the learning curve and lowers its costs in subsequent periods. This in turn will allow the firm to enjoy a larger cost advantage vis-a-vis its rivals and perhaps deter entry. The key assumption here is that learning is not transferable (i.e., no spillovers), it can only be achieved throught production. We examine this argument in this lecture.

The phenomenon of learning-by-doing was first discussed by Alchian (1950) in a study of wartime airframe production. Last year, a Yale Ph.D economics student estimated the magnitude of learning-by-doing effects in the airplane production and found that they were huge and were an important determinant of product exit and entry patterns. Lieberman (1982, 1984) estimated learning curves in production of chemical products. More recently, Dick (1991) has studied strategic effects of learning in the semiconductor industry, and Jarmin (1994) has conducted a similar study of the US Rayon industry in the period 1920-1938. He finds that the leading rayon producers overinvested in learning to push down their learning and reduce their rivals' markets shares in subsequent periods.

1.2.1 Model

There are two periods. Demand in each period is given by

$$P(y) = 1 - y.$$

In period 1, firm A is a monopolist. Its unit cost in period 1 is c. Its unit cost in period 2 depends upon its output in period 1, denoted y_1, according to the equation

$$c_2 = c - \theta y_1$$

where $\theta < c$. For simplicity, we assume that firm 1 does not discount profits, so its objective is to choose outputs (y_1, y_2) to maximize profits summed over the two periods. Firm A is worried about firm B entering in period 2.

Monopoly Before studying how the threat of entry affects firm A's choice of output in period 1, let us first establish the monopoly benchmark. Firm A's maximization problem is as follows:

$$\max_{y_1, y_2} \ (1 - y_1 - c)y_1 + (1 - y_2 - c + \theta y_1)y_2$$

Let us first optimize with respect to second period output. The first-order condition for optimality requires

$$y_2 = (1 - c + \theta y_1)/2$$

Susbstituting this relation into the maximization problem given above and simplifying yields

$$\max_{y_1} \; (1-c-y_1)y_1 + (1-c+\theta y_1)^2/4.$$

Differentiating with respect to period output yields

$$1-c-2y_1+\theta(1-c+\theta y_1)/2 = 0.$$

The profit maximizing solution is

$$y_1^M = \frac{(1-c)}{(2-\theta)}, \;\; y_2^M = \frac{(1-c)}{2-\theta}$$

Intrepretation: If the monopolist cared only about first period profits, it would produce $(1-c)/2$. When it takes into account the effect of first-period production of second-period costs, it produces more. The larger is the value of θ, the higher is the monopolist's first period output.

Threat of Entry Suppose now that firm A worries about entry by firm B in period 2. Let x denote output by firm B. It produces at a unit cost of c. Assume that y_1 is observed by firm B. Then, if it enters, then its optimization problem is

$$\max_x (1-c-y_2-x)x$$

Differentiating with respect to x yields its best reply

$$R(y_2) = (1-c-y_2)/2.$$

Similarly, firm A's best reply in period 2 is

$$R(x) = (1-c_2-x)/2.$$

Solving for the intersection of the best replies yields the Cournot solution

$$y_2 = (1+c-2c_2)/3, \quad x = (1+c_2-2c)/3.$$

Equilibrium profits in period 2 are

$$\pi_2^A = (1+c-2c_2)^2/9, \;\; \pi_2^B = (1+c_2-2c)^2/9.$$

Firm A's first period optimization problem, assuming it accommodates the entrant, is

$$\max_{y_1} (1-c-y_1)y_1 + (1-c+2\theta y_1))^2/9.$$

Solving yields

$$y_1^D = \frac{(1-c)}{2}\left[\frac{(9+4\theta)}{(9-4\theta)}\right]$$

It is not difficult to show that y_1^D is larger than y_1^M. Intuitively, firm 1 wants to be more aggressive in period 1 since it then commits itself to higher output in period 2. This is in turn causes the entrant to choose a lower output, increasing profits to firm 1.

If firm 2 incurrs an entry cost of F, and its profits at the above solution exceed F, firm 1 may want to increase output even further to deter entry. Once again, there may be two solutions: the output level which assumes accommodation and a larger output which deters entry.

Example 2 *Suppose* $c = \theta = .25$. *Then*

$$y_1^M = 3/7, \quad y_1^D = 15/32.$$

Since output in period 1 is approximately equal to $1/2$*, firm 1's unit costs in period 2 are approximately* $1/2$ *of its unit cost in period 1. Hence* $y_2 \approx 1/3$ *and* $x \approx 5/24$*. Total output in period 2 is approximately* $13/24$*, which is only slightly less than the amount in period 1. Hence, prices do not decline very much, even though period 1 is monopoly and period 2 is duopoly.*

1.3 Complementarities

Given the freedom to choose their own route structure and prices following deregulation in 1979, most airlines transformed their networks into hub and spoke networks. Interlining traffice (i.e., changing airlines at a connecting point) declined dramatically. Bamberger and Carlton (1993) report that interlining traffic as a share of connecting traffic fell from 38.8% in 1979 to 4.5% in 1989. This reflects the growth of single carrier hub airports. By 1986, Pickrell and Oster report that virtually all of the commuter and regional airlines were tied contractually to one of the major airlines operating hub and spoke networks. We provide an explanation of this phenomenon in this lecture and also why hub and spoke networks are a deterrent to entry by small airlines.

1.3.1 Model

There are $n \geq 4$ cities and individuals living in each city who wish to travel to other cities. Individuals who wish to travel from city g to city h are assumed to have no desire to travel anywhere else (i.e., no substitutability in demand across city-pair markets). Individuals care only about reaching their destination at the lowest price, not how this destination is reached. In particular, they are indifferent to distance traveled, number of stops incurred, or the airline that is flying them. Demand in each city-pair market is the same. If p_{gh} is the price of the cheapest return ticket from city g to city h, then the number of g-h travelers is given by

$$D(p_{gh}) = 1 - p_{gh}$$

Here the g-h market is distinct from the h-g market. For simplicity, we assume that transport cost per passenger per flight are 0. The fixed costs of offering a direct flight between cities g and h is F. The flight services both the g-h and h-g markets. Here we assume that F is not sunk.

Monopoly Hub Operator Suppose airline H is a monopolist and operates a hub and spoke network. Let city 1 denote the hub city. The total number of markets are n(n-1). Of these, 2(n-1) are serviced by a direct flight, and (n-1)(n-2) city-pair markets that are serviced by a one-stop flight. Since length does not matter either to airline or travelers, the monopolist charges the same price in each city-pair market. Define

$$\pi^M \equiv \max_p p(1-p)$$

to be the monopoly profits in a city-pair market. It is easily verified that

$$p^M = 1/2, \quad \pi^M = 1/4.$$

Network profits to the hub operator are

$$\Pi^M = (n-1)[n/4 - F].$$

Threat of Spoke Entry A low-cost regional airline is contemplating entry into one of the spoke markets, say the (1-2) and (2-1) markets. Its marginal costs are also 0 and its fixed cost is F_E. Will the hub-operator concede the spoke market to the lower cost regional airline? To answer this question, we need to compute the equilibrium profits of the hub-operator when it does and does not concede the market.

Suppose it does not concede the market. Then the pricing in the (1-2) and (2-1) markets is essentially Bertrand. Each airline will undercut the other until price is equal to marginal cost. Each airline earns zero profits in these markets. Prices in all other city-pair markets are unaffected. In particular, the hub operator continues to charge monopoly prices in each (g-2) and (2-g) market, $g \neq 1, 2$, and in each of its hub markets. Therefore, its losses are -1/2.

Now, suppose the hub operator concedes the (1-2) and (2-1) markets by withdrawing its flights. By doing so, it saves the fixed cost F. What does it lose?

Let s denote the price charged by the regional carrier for traveling on its (1-2) or (2-1) flight. It can charge only one price on each flight since it cannot discriminate among travelers on the basis of their origin or destination. The hub operator faces a similar problem trying to discriminate between (g-1) and (g-2) travelers (as well as (1-g) and (2-g) travelers). Because the latter fly with the regional traveler, the hub operator is constrained to charge the (g-1) and (g-2) travelers the same price. Let p denote this price. It is determined by the following optimization problem:

$$\pi^s \equiv \max_p p(1-p) + p(1-s-p)$$

which yields the best reply function

$$p = (1-s)/2.$$

Note that π^s consists of profits in the hub market, (g-1), and the connecting market, (g-2). Given the hub operator's price p in these markets, and assuming this price is the same for every $g \geq 3$, the regional carrier solves

$$\max_s s(1-s) + (n-1)s(1-p-s)$$

which yields the best reply

$$s = \frac{(n-1)-(n-2)p}{2(n-1)}.$$

The equilibrium consists of

$$s = \frac{n}{3n-2}, \quad p = \frac{n-1}{3n-2}, \quad \pi^s = \frac{n-1}{3n-2}$$

There are 2(n-2) connecting markets associated with city 2. Therefore, the loss in operating profits from accommodation is

$$2(n-2)(2\pi^M - \pi^s) - F = 2(n-2)\left(\frac{1}{2} - \frac{n-1}{3n-2}\right) - F > 2(n-2)\frac{1}{6} - F.$$

Clearly, given any F such that the regional operator makes a profit, n does not have to be very large for the hub operator to be better off not conceding the market. If it concedes the (1-2) and (2-1) market, the hub operator has to lower prices in its hub markets and share profits in each connecting market involving city 2. Note that prices in the latter markets are higher with entry due to double marginalization.

The complementarities of hub-spoke networks make it difficult for a regional carrier to enter a spoke markets or to survive as an independent operator. In the former case, the hub operator will not want to concede the market and in the latter case, it has a strong incentive to invade its markets. Therefore, it is not surprising that regional carriers quickly became allied with the hub operators. The alliance also allowed the airlines to codeshare and price discriminate.

It is not hard to see how the argument given above for airline networks extends to sets of complementary products such as computer software and hardware. The economic forces suggest that one firm will offer the entire set of products and that it will be difficult for any firm to invade a part of the product space. Furthermore, competition may not be a good thing in these circumstances. Prices tend to be lower when the entire set of products are offered by a single firm.

1.4 Brand Proliferation

In 1972, the FTC filed a complaint against Kellogg, General Mills, General Foods and Quaker Oats - the four largest ready-to-eat (RTE) breakfast cereals in the U.S. They were charged with violating Section I of the Sherman Act. These four firms accounted for 91% of the market in 1970. Despite substantial growth in industry sales in the period from 1947 to 1970, the number of competitors had declined from 55 in 1947 to 30 in 1967. Each of the four firms enjoyed rates of return which were double the rates enjoyed by other firms in the food industry.

The FTC argued that the Big Four colluded on price and engaged in a number of exclusionary practices designed to restrict competition among themselves and exclude entry. The practices included (1) a shelf-space allocation plan designed to stablize market shares (2) intensive advertising to create the impression of product differentiation and to create barriers to entry and (3) brand proliferation. According to the FTC, the Big Four occupied virtually every single profitable position in the product spectrum of the RTE breakfast cereals market. With so many brands available in the product spectrum, it was hard for a new entrant to find a niche to establish itself. Furthermore, when an entrant did manage to find a niche, the four firms immediately brought out new brands to compete in that location
.

This lecture addresses the following question. Is product proliferation on the part of an incumbent monopolist a profit-maximizing strategy?

1.4.1 Model

Products are differentiated along a single dimension which we model by assuming that tastes of consumers are uniformly differentiated along the unit interval, [0,1]. Each consumer is indexed by her location and each consumer wants only one unit of the good. The utility of consumer x is given by

$$u(x) = s - p - td^2$$

where p is the price of the product purchased and d is the distance to the location of the product. Each product requires a setup cost of f. Marginal costs of production are zero.

There are two firms. Firm 1 is the incumbent firm. We assume that s is sufficiently large relative to t that it is optimal for Firm 1 to service the entire market with a product located at 0. It has been charging customers a per unit price of

$$p^M = s - t.$$

This is the highest possible price that the monopolist can charge and still attract the consumer located at l. The firm will want to do as long as s is sufficiently large relative to t. Its profits were

$$\Pi_0^M = M(s - t) - f$$

where M was the number of consumers. We shall assume that

$$Mt/2 < f < Mt.$$

This assumption implies that no one will enter when the state of demand is M. Recall that in the case of quadratic transport costs, if another firm enters, then it will do so at location 1. Furthermore, if it did so equilibrium prices would be

$$p_1 = p_2 = t$$

where p_i denotes the price charged by firm i. Each firm would serve half of the market. Consequently, if firm 2 entered at location 1, it would earn a profit of

$$\Pi_0^D = \frac{Mt}{2} - f$$

which is negative by the above assumption.

Suppose demand doubles so firm 2 can enter profitably. Firm 1 has the advantage of incumbency and can move first. If it locates a new product at location 1, then its profits are

$$\Pi_1^M = 2M(s - t/2) - 2f.$$

The marginal consumer is now located at 1/2. If firm 1 allows firm 2 to enter at location 1, then the two firms will divide the market at price t and each will earn

$$\Pi_1^D = Mt - f.$$

Under our assumption about f, it is easily checked that $\Pi_1^M > \Pi_1^D$. Thus, firm 1 will try to preempt entry by firm 2 by locating a new product first. Notice that the argument here is exactly the same as in the R&D game. Entry competes away monopoly profits that the incumbent will want to protect by preemptive investment.

Judd (1985) has pointed out that the above argument depends critically upon the immobility of products. If exit costs are negligible, then firm 2 should ignore the fact that firm 1 has located a new product at location 1 and locate its product there as well. Why? Because firm 1 will be hurt by the competition at location 1 and will prefer to withdraw its product.

To see why, let p_{10} and p_{11} denote firm 1's prices of products 0 and 1 and let p_2 denote firm 2's price of its product 1. If firm 1 does not withdraw its product, then Bertrand competition will drive price of good 1 to zero, that is, $p_{11} = p_2 = 0$. This in turn means that the marginal consumer for good 0 is determined by

$$p_{10} + t\widehat{x}^2 = 0 + t(1 - \widehat{x})^2 \Longrightarrow \widehat{x} = (t - p_{10})/2t.$$

Firm 1 seeks to choose p_{10} to maximize

$$\pi(p_{10}) = p_{10}\widehat{x}(p_{10})$$

which yields the solution

$$p_{10}^* = t/2.$$

Total revenues for firm 1 if it insists on offering product 1 are tM. On the other hand, if it withdraws product 1 and leaves the market for firm 2, then the equilibrium price for each product increases to t and firm 1 earns revenues of $2Mt$. Therefore, even if f is sunk, firm 1 is better off withdrawing its product 1 provided exit costs are not too large. If setup costs are not sunk, then it gains even more.

Notice that the argument in this case is exactly the opposite of the one given for hub-spoke networks. The reason is that here products are substitutes, not complements.

Finally, I would like to remark that the proliferation of brands may have nothing to do with entry deterrence. If a cartel fixes prices, then its members may try to compete in other dimension such as advertising and the number of products offered. Rents are dissipated, not by lowering prices, but by offering an excessive number of products.

1.5 Tying

Tying exists when a seller of a product requires as a condition of sale that the customer purchase a second product (the tied product) as well. It is often viewed by antitrust authorities as an exclusionary tactic and a violation of #1 of the Sherman Act. The issue is whether a firm with a monopoly in one market, say market A, can monopolize a second market, say market B, by tying the sale of product B to the sale of product A.

1.5.1 Model

We consider a simple model due to Whinston (1987) and discussed by Tirole in his text. There are two firms, 1 and 2, and two markets, A and B. The potential customers for products A and B are the same and equal to M. Demand in this market is

$$D(q) = \begin{cases} M & \text{if } q \leq v \\ 0 & \text{if } q > v \end{cases}$$

where q is the price of good A. Market B is a differentiated market in which demand for firm i (assuming $q \leq v$) is denoted by $D_i(p_i, p_j)$. Note that, since the set of customers in each market is the same, $D_i(\cdot, \cdot) \leq M$. Let c_A denote the unit production costs of good A and let c_B denote the (common) unit productoin costs of good B.

Tied Sales Suppose firm 1 refuses to sell individual units of goods A and B. It only sells packages consisting of one unit of each good. Let P denote the price of the package. Given firm 2's price for good B, firm 1 chooses P to

$$\pi_1^* \equiv \max_P (P - c_A - c_B) D_1(P - v, p_2).$$

How many customers buy from firm 1? Only those who are willing to pay v for good A **and** $P - v$ for a unit of good B. Customers who are willing to pay v for a unit of good A but whose willingness to pay for good B from firm 1 is less will not buy the package. Let P^* denote the profit-maximizing tying price..

Separate Sales Suppose firm 1 offers to sell the goods separately. It offers good A at price v and good B at price $\widetilde{p}_1$ where $\widetilde{p}_1$ solves the maximization problem

$$\widetilde{\pi}_1 \equiv \max_p (p - c_B) D_1(p, p_2)$$

In this case, it sells good A to all M customers, and good B to $D_1(\widetilde{p}_1, p_2)$. But

$$\widetilde{\pi}_1 \geq (v - c_A)M + (p^* - c_B)D_1(p^*, p_2) \geq (v - c_A + p^* - c_B)D_1(p^*, p_2) = \pi_1^*.$$

The first inequality follows from the fact that $p^* = P^* - v$ does not maximize firm 1's profits in market B. The second follows from the fact that M exceeds the number of customers buying good B from firm 1.

Clearly, firm 1 is better off pricing each product separately and not tying. It can then sell units of good A to customers who prefer to buy from firm 2 or who just want good A at existing prices. This captures the intuition that firm 1 should be able to do better charging two prices than one price.

Conclusion 1 *If firm 2 is in the market, tying is not optimal.*

1.5.2 Tying to Foreclose

The rationale for tying is to foreclose market B to firm 2. The basic idea is that tying allows firm 1 to commit to charging a low price in market B, thereby making this market less profitable to firm 2 and deterring it from entering (assuming production involves fixed costs). The key is to recognize that p^* is lower than $\widetilde{p}$. By definition, p^* maximizes

$$(p^* - [c_B - (v - c_A)])D_1(p^*, p_2).$$

In other words, tying has the effect of reducing firm 1's costs in market B from c_B to $(c_B - (v - c_A))$. Since best reply is increasing in unit cost, it must be the case that $\widetilde{p}$ exceeds p^* at any p_2. Hence, tying causes firm 1's best reply to shifts inward, which implies lower equilibrium prices in market B and less profit for firm B.

Conclusion: tying can be a useful tactic to foreclose competitors from the market.

Research and Development

April 16, 2007

1 RESEARCH AND DEVELOPMENT

R&D is the production of information. The literature often distinguishes between product and process innovation, but this distinction is rather artificial. The commodity information involves some special problems: (i) risks, (ii) indivisibilities, and (iii) appropriability.

The main questions of the economics of R&D are:
1. (Positive) How the allocation of resources is affected by the interaction of these problems with different market structures and with different institutional arrangements.
2. (Normative) The social desirability of institutions and legislation that regulate R&D (e.g., patents).

1.1 Patents

The central and most special problem of R&D is appropriability. In order for there to be an incentive to innovate there must be some possibility to reap its fruits. This is achieved either naturally, due to difficulties to imitate, or through assignment of property rights on inventions. The patent system attempts to do the latter. The main tension is that while patents induce R&D they also confer monopoly power and its resulting inefficiencies. The literature on patents deals with these tradeoffs inquiring about the optimal duration and width of patents and about strategic manipulation of this system. In the following discussion, we shall not discuss these important problems - we shall assume the existence of such a system and inquire about the incentives to innovate in its presence.

1.2 The Incentive to Innovate

Consider a process innovation. Prior to the innovation, the cost of production is c per unit. After the innovation, it is b<c. Let $p^m(c)$ denote the monopoly price when the cost is c.

Definition 1 *An innovation is called* drastic *if* $p^m(b) < c$.

Let V^m denote the value of an innovation under monopoly. Define $\pi(q;c)$ as the profit associated with output q when cost per unit is c. Let $\Pi^m(c)$ denote the maximum profit that can be achieved when the cost per unit is c and let $q^m(c)$ denote the quantity that maximizes the profit in this case. More precisely,

$$\Pi^m(c) = Max_q \pi(q; c) \equiv \pi(q^m(c), c)$$

Then,

$$V^m = \Pi^m(b) - \Pi^m(c) < \Pi^m(b) - \pi(q^m(b); c) = q^m(b)[c - b].$$

The inequality follows from the fact that $q^m(b)$ is not optimal when costs are c.

Let V^c denote the value of an innovation to a firm when the industry is competitive at c. That is, prior to the innovation, $p = c$. After one of the firms has reduced its costs to b, it can undercut its higher cost rivals and service the entire market. Two possibilities can occur. If the innovation is drastic, the innovator becomes a monopolist. If the innovation is not drastic, then the innovator sets $p = c$ (minus a penny) and the innovator produces $q^c(c) = D(c)$. Hence,

$$V^c = \begin{cases} \Pi^m(b) & \text{if drastic} \\ (c-b)q^c(c) & \text{if not} \end{cases}$$

Note that $V^m < V^c$. In case of a drastic innovation this is obviously true. If the innovation is non-drastic, then by definition $P(q^m(b)) > c = P(q^c(c))$ which implies that $q^m(b) < q^c(c)$.

Let V^s denote the social value of the innovation. This is equal to the gain in consumer surplus from reducing price from c to b less the cost of the innovation. It is easy to show that (use graph) $V^s > V^c$. Therefore

$$V^s > V^c > V^m.$$

The reason why V^s exceeds V^c and V^m is clear. Although the patent system awards the innovator with a monopoly right, it still does not enable the innovator to capture the entire surplus generated by the innovation. The value to the innovator accounts only for that part of the surplus that it can capture as profit. The reason that $V^c > V^m$ is that the monopoly enjoys a more profitable position prior to the innovation and therefore the incremental profit resulting from the innovation is lower.

In a competitive industry, a firm will have an incentive to undertake an R&D project if its cost is lower than V^c. Similarly, in a monopoly, the firm will have an incentive if its cost is lower than V^m. Since the R&D project is desirable from a social viewpoint if its cost is less than V^s, the conclusion from above is that both structures will give rise to too little R&D activity, and the monopoly structure will induce even less innovation than a competitive one.

The policy implication from this point is that the patent system itself may not be sufficient to induce the socially optimal level of R&D. Thus, in theory, there is a case to be made for a corrective subsidy to encourage R&D.

The implication of the above analysis is that the monopolist is likely to invest less in R&D than a competitive firm. But this assumes that the monopolist is not threatened. If the monopolist is threatened by an entrant, the monopoly might be willing to pay more in order to protect its market from entry. For example, suppose a monopoly has a patent on technology c and that technology $b < c$ is offered for sale by its inventor. Let $\Pi^d(b, c)$ denote the profit of a duopolist with cost b who faces another duopolist with cost c. It is reasonable to assume that

$$\Pi^m(b) \geq \Pi^d(b, c) + \Pi^d(c, b)$$

and for non-drastic innovation this holds with strict inequality. In that case, the monopoly will be willing to pay more to maintain its monopoly since

$$\Pi^m(b) - \Pi^d(c, b) \geq \Pi^d(c, b)$$

Note that the above inequality implies that, even if the monopoly uses technology b and $c > b$ is offered for sale, it would like to buy it in order to prevent entry, though it would not use it. This is true of course only if b is not drastic with respect to c, since when b is drastic, entry is blocked anyway.

1.3 Licensing of Innovations

Licensing of innovations takes place between an independent inventor and a user firm and also between firms in the industry.

The licensing arrangements which were studied consist of different combinations of fixed fees and royalties.

Consider first a drastic innovation, that is, the new technology $b < c$ is such that $p^m(c) < c$. If the inventor is a member in the industry, it will not license it. If the inventor is not active in the industry itself and if it can charge a fixed fee, then this innovation would be licensed to only one firm. The reason is simple: with a drastic innovation the firm with the patent monopolizes the industry and hence extracts the maximum possible profit. It is therefore obvious why an inside inventor will not license it. By charging fixed fees, an outside inventor can appropriated the entire monopoly profit, and hence it too will not license it to more than one firm.

If however, the payment structure is restricted to royalties, then an outside inventor might want to license it to more than one firm. For example, suppose $c = .9, b = 0$, and $P = 1 - y$. The optimal royalty of licensing to one firm is $1/2$, which will give the inventor revenue of $1/8$. If two firms in this industry will compete as Cournot or Bertrand players, the optimal royalty will again by $1/2$ and it will give the inventor $1/6$ or $1/4$ respectively.

With a non-drastic innovation even a firm in the industry may want to license its innovation. Suppose that the industry is a Cournot duopoly. In this case, after the invention, both firms will continue to serve the market. The firm who owns the patent can always license the innovation for the royalty $c - b$. In this case, the equilibrium behavior remains the same as in the no licensing scenario, but the patent holder's profit is higher by the revenue from the royalty.

As we have noted, restrictions on the structure of payments may affect the licensing decisions. If payments cannot involve royalties, then licensing will not always take place. If, for example, the difference between c and b is large, so that the Cournot equilibrium quantity of the firm without the patent is sufficiently small, the patent holder would not want to license. The reason is that the total industry profit in the post licensing equilibrium is sufficiently smaller, so that even if the patent holder uses a fixed fee to appropriate much of teh other firm's profit, it does not compensate for the decrease in its own profit. If c and b are close, then licensing may be profitable.

Note that licensing here is always socially desirable since it decreases cost by introducing the more efficient technology to the other firm, and does not decrease total quanitity.

If the patent holder knows that a rival can and is likely to innovate around an existing patent, it might be profitable to license the existing patent. In this case, the licensing is socially desirable as well since it avoids costly duplication of the R&D effort.

1.4 Research Joint Ventures

These are arrangements in which the firms share the costs and benefits of an R&D project. Both RJV and licensing have the following effects:
(a) Desirable effects of reducing unnecessary duplicative effort
(b) Mitigate competition in R&D and hence might have some retarding effects on R&D effort
(c) Open possibilities for firms to collude on other aspects of the competition.

Search

An important assumption of the models we have developed in the previous lectures is that consumers know the prices that sellers are charging.

However, in most retail markets, consumers do not know the prices charged by different retailers and have to learn them. This activity is costly, so consumers are unlikely to be fully informed.

- One of the main uses of the Internet is to provide price information cheaply.

Q. How does the consumer's lack of information about prices affect competition among firms?

Varian (1980) developed a simple model to answer this question.

Supply:

- N firms supply a homogenous good.
- Production costs are zero.

Demand:

- Consumers have unit demands and willingness to pay of r.

- Two types of consumers: I informed (shoppers) and M uninformed consumers.

- An informed consumer knows the prices charged by all N firms.

- An uninformed consumer does not know the prices and randomly selects one firm to shop.

The Game: n firms post prices simultaneously.

- Informed consumers buy from the firm(s) setting the lowest price if it does not exceed r.

- Uninformed consumers randomly select a firm and buy if the firm's price does not exceed r.

Let U = M/n. It is the number of uninformed consumers that each firm gets.

Payoffs:

$$\pi_i(p_i, p_{-i}) = \begin{cases} p_i(U + I) & \text{if } \; p_i < p_j, j \neq i \\ p_i(U + \frac{I}{m}) & \text{if } \; p_i = p_j, j = 1, .., m-1 \\ p_i U \text{ if } r \geq p_i > p_j, j \neq i \end{cases}$$

Interpretation:

- Lowest price gets all of the informed consumers. If m firms tie at this price, then I is divided equally among the firms so each gets I/m.

- Each firm can guarantee itself U uninformed consumers if it does not charge more than r.

Claim 1: There is no equilibrium in which all firms charge the same price.

Proof: Without loss of generality, assume that n = 2.

- p = 0 is not an equilibrium because each firm can raise price to r and earn rU.

- If they charge $r \geq p > 0$, then a slight cut in price gets all of the informed consumers, which increases profits. Q.E.D.

Thus, the Law of One Price does not hold!

Claim 2: There is no equilibrium in pure strategies.

Proof: By previous claim, one firm sets a higher price than the other. Let $p^L < p^H$.

- Clearly, $p^H = r$ since firm H's profits are pU which is strictly increasing in p.
- But then firm L should raise price p^L until it is slightly below r.
- But then firm H wants to undercut p^L.

Q.E.D.

We need to look for an equilibrium in which firms cannot forecast the prices of their rivals and undercut them.

In other words, firms are perceived as choosing prices randomly (e.g., hold sales at random times).

Let F(p) denote the probability that a rival posts a price less than p. Then the expected profit to a firm when it charges p is

$\pi(p) = (1 - F(p))^{n-1}p(U+I) + (1 - (1 - F(p))^{n-1})pU$. (*)

- The first term on the RHS is probability that p is the lowest price, in which case firm demand is U+I.

- The second term on the RHS is the probability that p is not the lowest price, in which case firm demand is only U.

In a mixed strategy equilibrium, the firm's profits at every price p has to constant: $\pi(p) = k$.

Solving (*) yields

$$1 - F(p) = \left[\frac{k - pU}{pI}\right]^{\frac{1}{n-1}}.$$

Clearly, the upper bound on prices is r. Therefore, the unknown constant k satisfies:

$$F(r) = 1 \rightarrow k = rU.$$

Substituting, we get

$$1 - F(p) = \left[\frac{(r-p)U}{pI}\right]^{\frac{1}{n-1}}.$$

The lower bound of the set of prices that firms will charge is obtained by setting F(p) = 0:

$$\underline{p} \quad = \quad \frac{rU}{U+I}.$$

This is strictly positive.

- Firms earn positive profits – escape the Bertrand trap.

- The equilibrium density of prices is U-shaped: firms will tend to either price near r or near the lower bound.

Intuition: firm either goes for the informed consumers or is content exploiting the uninformed consumers.

Remark: As U falls, market becomes more competitive, prices fall, and lower prices more likely.

Prior to the Internet, search costs were attributed to dispersion in the geographical location of stores.

- It was costly to visit every store and determine who was offering the lowest price.

However, price search engines like *Pricewatch.com* provide consumers with lots of prices at very low cost. Online travel agents like *Travelocity* and *Expedia* do the same for airline markets.

This led many researchers to predict less price dispersion and lower margins in online markets than in brick and mortar markets.

Furthermore, it is much easier to test this prediction in online markets.

- Even if products are identical (e.g., a CD by Britney Spears, "Da Vinci Code" book) when they are sold in brick and mortar stores, they are differentiated by location.

- Products in online markets are not differentiated by location.

Nevertheless, a long list of papers have found that price dispersion among e-retailers is similar to that of brick and mortar retailers and margins are not extremely low (e.g., Amazon reports average markups of 15%).

One reason may be *obfuscation*. Online retailers try to make it difficult for consumers to determine the true price – shipping cost, taxes, etc.

Advertising

Advertising is ubiquitous and the amounts spent on it are staggering.

- In 2002, U.S. firms spent 236.9 billion dollars (approximately 2.3% of GDP) on advertising.

- Historically, advertising has been around 2% of GDP.

Media (TV, radio, newspapers, magazines, yellow pages, Internet) advertising accounts for roughly 58% of total expenditures.

Non-media include direct mailings, promotions, coupons, catalogs, business publications, sponsorship of special events.

Advertising-to-sales ratio varies widely across industries and products within an industry.

- For example, it is 2% for expensive items like cars, but 14% on less expensive items like soaps and cleaners.

- The variation is not easily explained.

The relationship between advertising and market power is also not clear.

- Old theories argued that advertising was a source of product differentiation and represented a "barrier to entry".

- These studies point to positive correlation between profit rates and advertising rates as evidence.

But correlation $\neq$ causation: advertising could create market power or market power could lead to more advertising.

- Newer theories and evidence suggest that advertising often makes people aware of new goods and hence makes entry easier and increases price competition.

Main problem: we do not understand very well "when and how" advertising affects consumer demand.

Some Questions:

1. What is the interrelationship between price and advertising in particular goods?

- Example: ban on advertising of cigarettes to youths. The lower advertising lead to lower prices (strategic complements). Youths were more price sensitive, so net result was an increase in youth smoking.

2. What is the role of price and advertising in introducing new goods? What is the role of advertising after everyone knows about product?

4. Is advertising a source of market power and, if so, when?

5. What are the welfare implications of advertising?

- Example: in 1997, FDA changed the required content of prescription drug television ads which allowed drug companies to advertise drugs direct to consumers. DTCA on drugs increased dramatically (3.6 billion in 2004) and increased drug purchases. Is this good or bad?

In addressing these questions, economists have generally distinguished between informative advertising and persuasive (image) advertising.

Informative Advertising

- Provides consumers with information about price, characteristics, location of sale, use (e.g., recipes).

This sort of advertising is pro-competitive. When consumers are aware of close substitutes, firms cannot charge high prices. Also reduces product differentiation due to lack of information (e.g., lowers search costs).

- Provides information about product quality that the producers know but the consumer does not.

Basic idea is as follows: producers of high quality products know that consumers who purchase these products are more likely to buy the product again.

Thus, present value of sales from an ad for a high quality product is higher than that of a low quality product.

→ Producers of high quality products are willing to spend more on advertising.

→ Advertising is a signal about product quality.

→ Consumers infer from the advertising the quality of the product and buy products that are heavily advertised.

Remark: The signaling argument is a "money-burning" story. What matters is the amount of money spent on advertising, not the content of the advertising which may be completely uninformative.

Spence offered a similar story to explain role of higher education.

- Advertising that enables more efficient matching between consumers and products.

Much of what marketing is about is figuring out how to direct advertising to consumers who will "find it useful". For example, online retailers frequently respond to purchases by providing buyers with information on "similar" products.

- Advertising may “jog one’s memory” about the existence of a good. Memory loss is one possible explanation for why continual advertising is necessary for a good whose characteristics do not change.

An example is prescription drugs. It appears that advertising increases compliance rates.

- All types of informative advertising are likely to be socially useful to some extent. But there may be too much or too little of it.

Theory does not deliver clear predictions on this issue.

1. Just as in R&D, advertising may generate big spillovers. When one firm advertises its product, it may make consumers aware of a range of products and the benefits are not captured by the advertiser.

- Advertising of drugs often fall to zero after generics enter.
- Pepsi and Coke advertising wars may have been mutually reinforcing.

2. When advertising expands the choice set of individuals, demand becomes more elastic and prices fall. On the other hand, one can also write down models where advertising make consumers less elastic with respect to prices.

- Beham finds prices of eyeglasses fall after firms were allowed to advertise.
- Waldfogel and Mylos find no change in prices after liquor advertising was allowed in Rhode Island.

3. There is a real issue of whether advertising is “mis-informative”. The FTC is charged with the responsibility of monitoring truthfulness of advertising claims but there is a large grey area.

Persuasive Advertising

1. Early view was that image advertising is an attempt by a firm to differentiate its product. Idea is to convince consumers that its product provides more utility than products of its rivals.

- This leads to conclusion that advertising is socially wasteful. Consumers who buy the product may believe that it generates more utility than rival products and hence are willing to pay higher prices but their belief is not correct.

- But need to assume that consumers are consistently fooled!

2. Another possibility is that advertising changes preferences rather than beliefs.

- Problematic for economists who typically take preferences as fixed.

- Example: without advertising, consumers are willing to pay $6 for Brand X detergent. The box costs $2 and firm charges $4. Thus, total surplus is $4. With advertising, consumers are convinced the box is worth $10 and firm sells the box for $6. Consumer surplus has increased to $4 and, if advertising costs are less than $2, the firm's profits have also increased. Total surplus increases.

3. The modern version of this story is that advertising has a psychological or social component that generates utility.

- Advertising makes the product appear more prestigious and desirable because consumers enjoy knowing that the brands of products they buy are widely seen and recognized.

- Example: if I drink the “in” beer and wear the “in” brand of clothing, the group will think I am an “in” person.

In this view, advertising does not dupe consumers nor does it change preferences. Rather advertising is a complement that enhances the value of the product, very much like a nice view is a complement to a hotel room.

- It explains why firms may continue to advertise even after consumers know the quality of the consumption experience.

4. Consumption of some goods like movies and albums have a social component: people want to buy what others buy in order to be part of the social conversation. For these goods, advertising provides a signal that others will buy the product.

- Difficult to distinguish from advertising as a signal of unobserved product quality.