Market Structure: Oligopoly (Imperfect Competition)

I. Characteristics of Imperfectly Competitive Industries

A. Monopolistic Competition

- large number of potential buyers and sellers
- differentiated product (every firm produces a different product)
- buyers and sellers are small relative to the market
- no barriers to entry or exit

B. Oligopoly

- large number of potential buyers but only a few sellers
- homogenous or differentiated product
- buyers are small relative to the market but sellers are large
- barriers to entry

C. What is product differentiation?

- Product differentiation occurs when firms sell similar but not identical products.
- Examples?
- Two kinds of product differentiation.

    - Real product differentiation = actual differences exist between the goods produced by different firms.
    - Imaginary product differentiation = no actual differences but consumer believe there are and act as if there were differences between the goods produced by different firms.

- Impact of product differentiation on firm demand.

    Recall that a perfectly competitive firm is a price taker with demand that is perfectly elastic. A price taker cannot raise its price without losing all of its quantity demanded. If that firm can differentiate its product then it will no longer be a price taker. Rather, it can now raise its price and not lose all of its quantity demanded, although it will still lose some. Thus, product differentiation causes a firm's demand to become <u>larger</u> and <u>less elastic.</u>

- Relationship between firm demand and market demand.

    Firms have more competitors than does the entire market because they have both the competitors from other goods that the market has plus the competition from other firms within the same market. Hence, market demand is split into firm demand. As the number of firms in the market increase then firm demand will get smaller.

    The increased competition also leads to more substitutes for firms and, hence, firm demand is more elastic than is market demand. As the number of firms in the market increase then firm demand will become more elastic.

- Market Power

    Market power is the ability of a firm to raise price and not lose all of its quantity demanded. That is, firms with market power have downward sloping demand curves. Perfectly competitive firms have no market power. Monopolies have market

power because of lack of competition. However, product differentiation, as discussed above, also by itself causes demand to be downward sloping. That is, product differentiation is a second source of market power.

II. Monopolistic Competition

A. The short-run in monopolistically competitive industries.

Monopolistically competitive industries look like monopolies in the short-run, as is shown in the graph below. The firm has a downward sloping demand curve because of product differentiation. Profit can be positive (as shown below), negative or equal to zero dependent upon market conditions. The firm produces where marginal revenue equals marginal cost. Price is given by the demand curve at profit maximizing output and profit equals (p – ATC)Q.
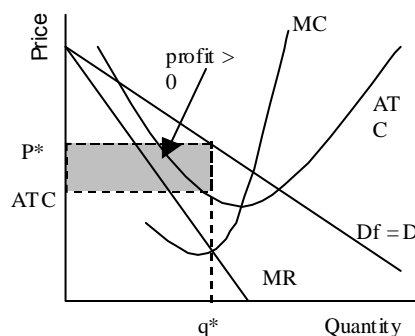
The only difference between monopolistic completion and monopoly in the short-run is that discussed in the previous section – firm demand is smaller and more elastic than market demand for monopolistic competition whereas for monopoly firm demand equals market demand.

Similar to both monopoly and perfect completion, firms in monopolistic competition may decide to shut down. The decision is the same for all firms in the short-run:

- o If P > ATC => profit > 0 => produce
- o If P = ATC => profit = 0 => produce
- o If P < ATC => profit < 0 => decision to produce or shutdown depends on:
  - ▪ If P < AVC => shutdown
  - ▪ If P ≥ AVC => produce

B. The long-run in monopolistically competitive industries.

Monopolistically competitive industries act like perfectly competitive industries in the long-run. This is because, like perfect competition, firms can freely enter and exit the industry. That is, no entry barriers exist to keep out competition. As a result, similar to perfect competition, profit serves as a signal to firms to either enter or exit the industry in the long-run.

- o If profit > 0 => entry occurs driving down prices and profit.

2

With entry and more competition market demand is split between more competing firms. Hence, market demand falls and becomes more elastic.

- o   If profit < 0 => exit occurs driving up prices and profit.

    With exit and less competition market demand is split between fewer competing firms. Hence, market demand rises and becomes less elastic.

- o   Therefore, profit moves to profit = 0 where there exists no entry or exit => profit = 0 is the long-run equilibrium in the market, just as it is in perfect completion.

The graph below shows a monopolistically competitive firm in long-run equilibrium with zero profit.
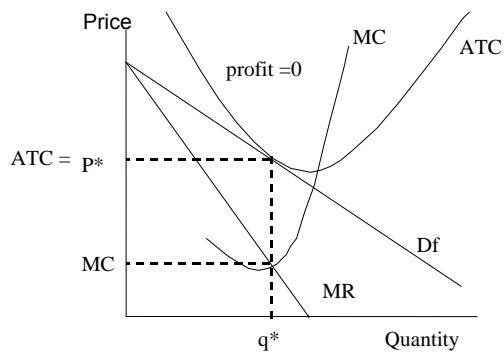
Use the graph above and compare to long-run equilibriums in perfect competition and monopoly. The graph will also be used to evaluate monopolistic competition with respect to technological and allocative efficiency. From the graph we can see that the following is true:

1. P=ATC.
2. MC = MR.
3. P > MC.

Efficiency requires:

- o   Technological Efficiency

    ▪   Firm Technological Efficiency

    Recall that for firm technological efficiency we ask the question: Does the firm produce on its cost curves?

    Clearly, as is true with perfect competition, the firm <u>must</u> produce on its cost curves. Competition requires it do so because if it does not profit becomes negative and the firm is driven out of business. Hence, monopolistically competitive firms will be efficient in this manner.

    ▪   Industry Technological Efficiency

    Recall that for industry technological efficiency we ask the question: Does the firm produce at the minimum point of the average total cost curve in the long-

3

run? In other words, we are asking whether the firm takes advantage of all economies of scale.

Clearly, the answer here is no. Zero profit requires a tangency between the downward sloping demand curve and the long-run average total cost curve. But that tangency ensures that the firm will be producing with long-run costs too high, inefficiency results.

- o Allocative Efficiency

Recall that for allocative efficiency we ask the question: Is price (or marginal social benefit) equal to marginal (social) cost? Clearly, from the graph above price exceeds marginal cost resulting in allocative inefficiency. Similar to a monopoly, a monopolistically competitive firm produces too little at too high a price.

The above discussion seems to imply that monopolistically competitive firms are just as inefficient as monopoly firms but this is not correct. In monopolistic competition entry and exit ensure that price falls so that profit equals zero, which is lower than for monopolies. The lower price is monopolistic completion means that they are more efficient than monopolies but less efficient than perfectly competitive firms.

- o How could monopolistically competitive firms be made more efficient?

The source of the inefficiency in monopolistic competition is product differentiation. Thus, wouldn't getting rid of product differentiation make the market more efficient? The answer to this question turns out to be not as clear as it might seem.

For example, no product differentiation means that all consumers would only be consuming exactly the same product. For example, everyone would drink only Fresca, drive only a dodge Dakota, eat only beef, etc. These are my preferences for soft drinks, a car, and food but may not be your preferences. Consumers actually get value from having choices because preferences vary between different consumers. That is, product differentiation is valuable in and of itself. Of course, this is only true for actual and not imaginary product differentiation.

III. Oligopoly

Recall that the characteristics of an oligopoly are:

- large number of potential buyers but only a few sellers
- homogenous or differentiated product
- buyers are small relative to the market but sellers are large
- barriers to entry

The above characteristics imply that there are two kinds of oligopolies:

- Pure oligopoly – have a homogenous product. Pure because the only source of market power is lack of competition.

  An example of a pure oligopoly would be the steel industry, which has only a few producers but who produce exactly the same product.

- Impure oligopoly – have a differentiated product. Impure because have both lack of competition and product differentiation as sources of market power.

An example of an impure oligopoly is the automobile industry, which has only a few producers who produce a differentiated product.

A. Measuring market or monopoly power via Concentration Ratios

A concentration ratio measures <u>only</u> the first source of market power, lack of competition. A concentration ratio takes the ratio of total sales of the ___ largest firms in the industry divided by the total industry sales. For example:

- A four firm concentration ratio = (total sales of the 4 largest firms)/(total industry sales)

- A one firm concentration ratio = (total sales of the largest firm)/(total industry sales)

- A eight firm concentration ratio = (total sales of the 8 largest firms)/(total industry sales)

A concentration ratio can be defined for any number of firms not just 1, 4, or 8.

Notice that concentration ratios must vary between 0 and 1. If the industry is a monopoly then the concentration ratio is always equal to 1. If the industry is perfectly competitive then the one, four, or eight largest firms in the industry are small relative to the industry and the concentration ratio is very close to 0. Thus, the larger the concentration ratio then the more market power exists because of lack of competition. Likewise, the smaller is the concentration ratio then the more competitive is the industry.

Concentration ratios have some limitations. For example, consider the following two true/false statements.

1. T or F: A monopoly will always have a concentration ratio that equals 1.

2. T or F: An industry with a concentration ratio of 1 must be a monopoly.

Statement 1 is true while statement 2 is false. Statement 1 is true because in a monopoly the single firm is the market. Hence, total sales of 1, 2, 3, or any number of firms will always equal total industry sales. Statement 2 is false because, for example, in a four firm concentration ratio there are <u>four</u> ways to get a concentration ratio equal to 1: (1) the industry is a monopoly, (2) the industry has 2 firms, (3) the industry has 3 firms, (4) the industry has 4 firms. Only one of these four ways is a monopoly. Statement 2 can be changed to be true in the following manner:

3. Tor F: An industry with a one firm concentration ratio of 1 must be a monopoly.

B. Why does oligopoly exist? Barriers to entry in oligopoly industries.

In general, the discussion of barriers to entry in the chapter on monopolies still applies here for oligopoly with a few exceptions. These differences are discussed here. For the full list see the monopoly chapter.

- Natural Oligopoly – Natural Barriers to Entry

Recall that a natural monopoly exists when only one firm can produce at the lowest cost or when LRAC is declining over the entire range of demand. For a natural oligopoly there must again be substantial economies of scale but enough to support more than just one firm. In other words a natural oligopoly would have a LRAC curve and a demand curve that looks like:

Price

Quantity
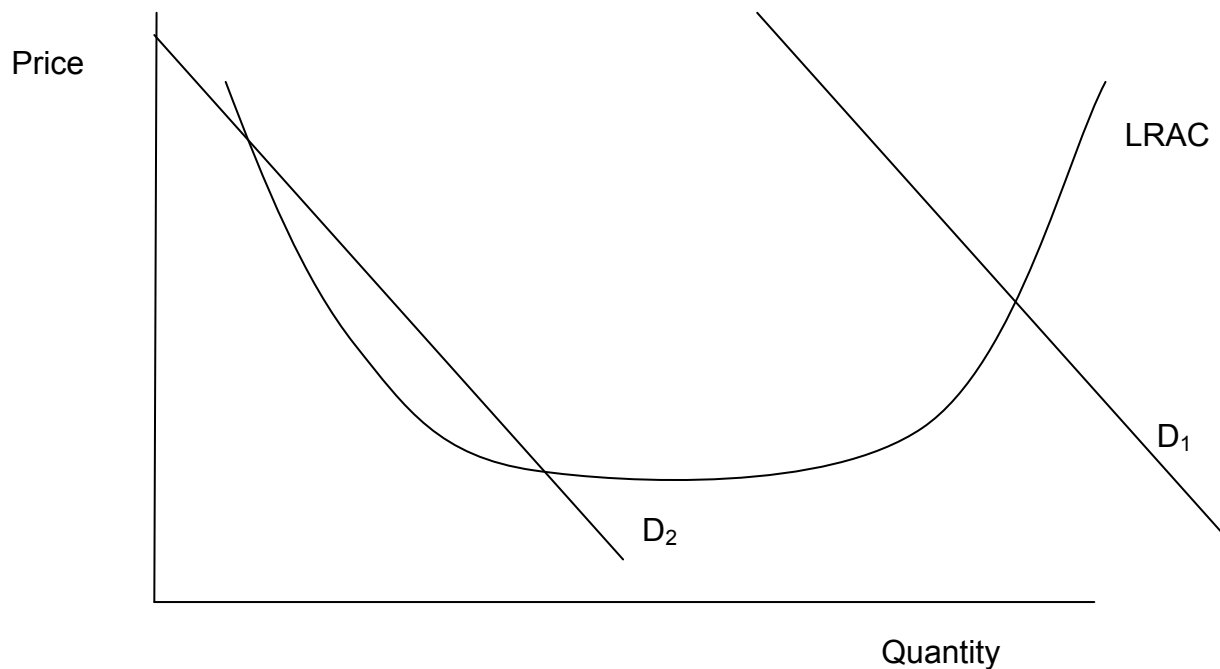
In the graph above, a demand equal to D_2 would result in a natural monopoly while a demand equal to D_1 would result in a natural oligopoly. The natural monopoly results because only one large firm can always produce at a lower cost while at D_1 demand is large enough that one firm would actually have <u>higher</u> costs of production given the diseconomies of scale (upward sloping LRAC).

- Artificial Barriers to Entry

  For the most part, the artificial barriers to entry discussed under monopoly are still the same for oligopoly. Both legal and illegal business practices do differ between monopoly and oligopoly industries. For example, model changes, advertising, proliferation of brands and so forth are legal business practices that tend to be used as entry barriers in oligopoly industries.

C. Mutual Interdependence

One of the problems with modeling oligopoly industries is to answer the question: How do firms respond to their competitors' behavior?

In the other market structures discussed to date we did not address this question because:

- For a monopoly no competitors existed.

- In a competitive market, either perfect or monopolistic, individual competitors were so small relative to the market that they could be safely ignored.

In oligopoly markets, however, other competitors both exist and are large enough that a firm must respond to their actions or be driven out of business. Therefore, the crucial question is how firms respond. Oligopoly firms are considered to be <u>mutually interdependent:</u> what one firm does affect their competitors.

Oligopoly markets are difficult to model because there are many ways that oligopoly firms may and do interact. We will focus on only a few of these models and will mostly use <u>game theory</u> to build the models.

D. Game Theory

6

Game theory is a mathematical method of analyzing the outcomes of choices made by people or organizations that are interdependent.

Definitions:

- Players – those who make choices in the game.

- Strategies – the possible choices the players can make to achieve their goals.

- Payoffs – the returns or profits to different choices.

- Payoff matrix – a matrix that shows how different choices affect the payoffs.

For example, suppose that A owns a house that he values at $60,000. B values the same house at $80,000 and has $70,000 in cash. Clearly, the possibility of a mutually beneficial exchange exists. There are two possible outcomes:

- Cooperative Solution = the two parties reach an agreement over price and the mutually beneficial exchange occurs.

  Assume that the price = $70,000. In this case A gets $70,000 and B gets the house, worth $80,000. Total value is $150,000.

- Non-cooperative Solution = the two parties negotiate but cannot reach an agreement over the price so no exchange occurs. How likely is this to happen or do people ever not have a mutually beneficial exchange because they cannot agree on the terms of the exchange?

  A keeps the house worth $60,000 while B keeps $70,000. Total value is $130,000.

- Cooperative surplus = the extra value gained by cooperating and exchanging the house = $20,000.

What does the payoff matrix look like in this example?

You must first decide what are the choices or decisions each of the parties can make. The issue here is bargaining to get more of the cooperative surplus. Both parties want more of the surplus. Therefore, suppose the decision is to bargain hard or soft. Further, suppose that:

- If one party bargains hard and the other soft then the one bargaining hard gets all of the cooperative surplus.

- If both parties bargain soft then they equally split the cooperative surplus.

- If both parties bargain hard then no exchange occurs.

The payoff matrix looks as follows:

		A's choice	
		Hard	Soft
B's choice	Hard	70K, 60K	90K, 60K
	Soft	70K, 80K	80K, 70K

The first number in each cell is person B's payoff while the second number is person A's payoff. Note that an exchange occurs and the cooperative solution is reached in three of the four cells. The only cell where the exchange does not occur is when both play hard.

What will the likely outcome be in the game above? To answer this question look at each individual's choice and assume that they will each act rationally.

- Individual A's choice

  o If B chooses Soft then A would be better off with hard.

  o If B chooses Hard then A does not care between hard and soft (gets $60K in either case.)

  o As a result, A is likely to choose hard.

- Individual B's choice

  o If A chooses Soft then B would be better off with hard.

  o If A chooses Hard then B does not care between hard and soft (gets $60K in either case.)

  o As a result, B is likely to choose hard.

Notice that even though a cooperative surplus exists the nature of the game itself, the fact that both parties are bargaining for more of the surplus, results in neither getting any of the surplus. Does this ever happen in real life?

This is also called a <u>dominant strategy game</u>. A dominant strategy game is one where each party's choice is not dependent upon what the other party decides.

E. A non-collusive game theory oligopoly model

Collusion occurs when the two firms communicate together and collectively decide their strategies. In most cases in the U.S., collusion between firms is illegal. First, examine non-collusive game theory models. For example, here is an example of a non-collusive, dominant strategy game:

<div align="center">Firm A's choice</div>

		Low Output	High Output
	Low Output	10, 20	9, 30
Firm B's choice	High Output	20, 17	18, 25

In this game we have two firms that are deciding how much output they are going to produce, either high or low output. Recall that the first number in each cell is Firm B's payoff or profit while the second number is Firm A's profit. For example, if both produce low output firm B gets $10 in profit while Firm A gets $20 in profit.

What is the dominant strategy? Both firms are better off, regardless of what the other firm does in choosing to produce high output given the payoffs above. However, let's change just one number in the payoff matrix and we get a different result:

<div align="center">Firm A's choice</div>

		Low Output	High Output
	Low Output	**22**, 20	9, 30
Firm B's choice	High Output	20, 17	18, 25

The payoff that was changed is in bold and is in the low, low cell where B's payoff goes from 10 to 20. What are the choices now?

- Individual A's choice

  o If B chooses Low then A would be better off with High.

  o If B chooses High then A would be better off with High.

  o As a result, A still has a dominant strategy which equals choosing High.

- Individual B's choice

  o If A chooses Low then B would be better off with Low.

  o If A chooses High then B would be better off with High.

  o Notice that B no longer has a dominant strategy. Firm B's strategy depends upon what A chooses.

  o But also notice that as long as B assumes A will be rational then B already knows what A will choose (High). Thus, B will also choose High.

- The equilibrium above is called a "Nash Equlibrium" = each player makes the optimal choice <u>assuming</u> that the other party will also make their optimal choice.

F. The Prisoner's Dilemma

Now look at a game called the Prisoner's Dilemma which is a classic game that is widely applicable to many other decisions including firm decisions. In the classic prisoner's dilemma, two people are arrested for committing the same crime. Suppose that the two people are split up in jail so they can't talk and each has the decision of either confessing or not confessing with the following payoff matrix:

		Individual A's choice	
		Confess	Don't Confess
Individual B's choice	Confess	10 years, 10 years	1, 15
	Don't Confess	15, 1	2, 2

In this case the payoffs are negative and equal the length of the sentence that each person will receive. There are two points to be made here:

1. Both parties have a <u>dominant strategy</u>, which is to confess. Thus, the equilibrium is for both to confess. This is also the <u>Nash equilibrium</u> (why?).

2. A cooperative surplus exists. If both confess, they are not cooperating. Thus, the equilibrium is the non-cooperative solution.

   If they both refused to confess then they would both get 2 years in prison and not 10 years. Why don't they cooperate?

   The answer is <u>not</u> just because they can't trust the other party to not confess. It is because it is <u>always</u> better to confess that not to confess regardless of what the other party chooses. This is what a dominant strategy means.

How can you change the results in the prisoner's dilemma? A crime organization does it by changing the payoff matrix to increase the penalty if one confesses. For example:

	Individual A's choice	
	Confess	Don't Confess
Confess	Death, Death	Death, 15
Don't Confess	15, Death	2, 2

Individual B's choice — Confess / Don't Confess (row labels)

In this case, as long as both parties are sure that to confess is to die then the cooperative solution will also be the Nash equilibrium. This is also a dominant strategy.

G. A collusive game theory oligopoly model

Now apply game theory to the problem of firms who collude, either legally or illegally. Firms often band together in organizations called cartels in order to reduce competition and gain monopoly/market power. Acting like a monopoly will raise profits collectively for all cartel members. However, cartels tend to be unstable because of the large number of problems successful cartels must solve. A successful cartel must solve all of the following problems:

1. Setting total output.

   This is the same problem that a monopolist solves by setting output where marginal revenue equals marginal cost. Although not a trivial problem for a single firm, this problem is relatively easily solved. For multiple firms, there exists more problems as each firm has their own viewpoint on its solution.

2. Setting each firm's share of total output.

   Dividing up total output is equivalent to dividing up profit. Each cartel member would prefer to be a monopoly and, hence, wants all of the profit. Clearly, then, deciding who gets what is a difficult problem. One way to solve the problem is to have exclusive territories. Another method is for the cartel to set production quotas for each member.

3. Reducing production by non-cartel members.

   The cartel must get all major firms to join the cartel or competition from outside the cartel will continue to reduce price and profit. However, the problem here is that it is better for the last firm to not join the cartel. A cartel acts like a monopoly by having its members voluntarily reduce their output and increase price. A non-cartel firm would, therefore, have an advantage if it didn't join because its competitors are voluntarily ceding market share to the non-member.

4. Cheating by cartel members.

   The incentive to cheat by cartel members is similar to the incentive not to join the cartel in the first place. If other cartel members don't cheat then they are voluntarily reducing output and ceding market share to the cheating firm. Thus, the cheating firm can increase its profit by increasing output. Of course, all firms have this same incentive and if they all cheat then the cartel fails. There are two main problems to solve in order to solve the cheating incentive:

1. How do cartel members detect cheating?

   The cheating firm(s) have an incentive to hide their activity so that their increased profits can continue. The successful cartel must have mechanisms in place to detect the cheating.

2. How do cartel members deter cheating?

   The most common method of deterring cheating, once detected, is to also cheat. Profits for the cheating cartel member go down as a result. Of course, if everyone cheats then the cartel fails.

Let us examine an example of a cartel game where the choice by each firm is to cheat or not to cheat. Notice that cheating is the equivalent of high output and not cheating is the equivalent of low output.

		Firm A's choice	
		Cheat	Don't Cheat
Firm B's choice	Cheat	100, 100	250, 50
	Don't Cheat	50, 250	200, 200

Notice that this is an example of a prisoner's dilemma game. The cooperative solution, where neither firm cheats, has the highest total profits. These are the monopoly profits, which each firm equally shares. However, for both firms cheating is <u>always</u> the dominant strategy. For example, if firm A doesn't cheat then B can increase its profit to $250 from $200. If firm A does cheat then B can increase its profit to $100 from $50 by also cheating. The same is true for decision making by Firm A.

As a result, the Nash equilibrium is for both parties to cheat even though overall profits decrease from $400 to $200. The solution is the same as is true for any prisoner's dilemma, the payoffs must be changed so that the cooperative solution is the Nash equilibrium. This could be done by, perhaps, fining firms that cheat. For example:

|		Firm A's choice	
		Cheat	Don't Cheat
Firm B's choice	Cheat	-100, -100	50, 50
	Don't Cheat	50, 50	200, 200

In this game, a firm that is caught cheating is fined $200. Notice that with this change the dominant strategy for both firms is to not cheat, which is the cooperative surplus.

H. Are Oligopoly Industries Efficient?

Recall our efficiency/market structure continuum:

Least Market Power/
Most Efficient

Most Market Power/
Least Efficient

Perfect
Competition

Monopoly

What we've learned so far about monopolistic competition seems to indicate that market power does lead to inefficiency. The same is true for oligopoly. More market power in oligopoly industries tends to lead to inefficiencies and for the same reasons as in monopoly and monopolistic competition. In fact, a successful cartel acts like a monopoly and is as inefficient as a monopoly. Hence, we would conclude that filling in the blanks in the continuum leads to the following:

Least Market Power/ Most Efficient   Most Market Power/ Least Efficient

Perfect Competition   Monopolistic Competition   Oligopoly   Monopoly
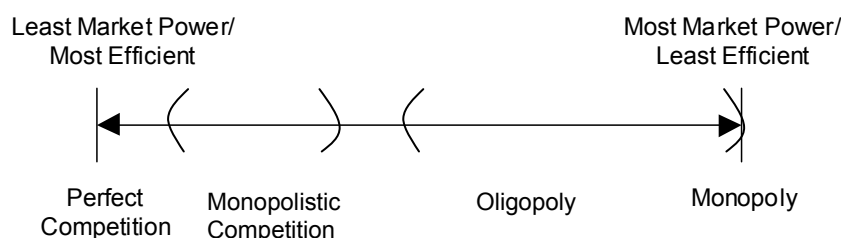
IV. Antitrust and Regulation

If market power leads to inefficiency then what is the correct solution to that market power? U.S. public policy has focused on either anti-monopoly laws that make firms having or attempting to gain monopoly power illegal or regulation of markets to make them legal. Recall our discussion of the exceptions to the idea that monopoly power leads to inefficiency in the chapter titles "Monopoly and Perfect Competition Compared". There we suggested that in the real world the simply conclusion that market power leads to inefficiency was not always correct. Regulation, then, theoretically would be used for those situations.

A. Anti-trust Laws

The U.S. has passed a number of anti-trust (a trust is a synonym for a monopoly) laws over the past century. The general theory underlying anti-trust laws are as discussed in the previous sections: monopoly power leads to market inefficiencies and must be curtailed to make markets more efficient. In this section we will discuss both some of the major antitrust laws and their major effects and how courts have interpreted those laws.

Major Anti-trust Laws

3. The Sherman Act (1890) has two major provisions:

    4. Section 1: Actions taken by firms ("contract, combination in the form of trust or otherwise, or conspiracy") that restrain trade or competition in markets are illegal.

    5. Section 2: forbids individuals from monopolizing, attempting to monopolize or to conspire with others to monopolize trade or commerce. This section, therefore, makes monopolies illegal.

Two major problems arose with the Sherman Act. First, how is a monopoly to be defined? Here, the main issue is how to define a relevant market. For example, which would be the relevant market: (1) gala apples, (2) apples, or (3) fruit? It is often the case in anti-trust cases that defining the market differently changes whether a "monopoly" exists.

The second issue is what market share must a firm attain in order to be considered a monopoly? Our discussion to date seems to imply that a monopoly exists only when a firm has 100 percent of the market. But wouldn't a firm with 99 percent of the market also be cause for concern?

12

6. The Clayton Act (1914), which amends the Sherman Act and has three major provisions.

   7. Section 2 forbids price discrimination that occurs not on the basis of cost or quality.

      Price discrimination is allowed when it doesn't reduce competition, when there are differences in quality or cost, and for intangible services (e.g., legal or physician services).

   8. Section 3 forbids restrictive agreements between firms. That is, it forbids agreements between firms that restrict the actions of firms and reduce competition in the market.

      This section is referring to practices such as agreements to tie sales of one good to sales of another good that reduce competition.

   9. Section 4 forbids mergers where the merge reduces competition.

10. The Federal Trade Commission Act (1914) which created the Federal Trade Commission (FTC) to investigate and challenge in the courts violations of anti-trust law.

11. Robinson-Patman Act (1936), which amends the Clayton Act.

   12. Prohibits discounts and other price concessions. Allows price discrimination only when it is (1) based on differences in cost, (2) a good faith attempt to react to competition in the market, and (3) caused by differences in the marketability of the firm's product.

Courts and Anti-trust Law

The U.S. legal system relies upon courts and ultimately judges to determine the meaning of language in laws passed by Congress. The courts have had two major approaches to anti-trust law:

13. The rule of reason approach to anti-trust.

   The early judicial interpretation by the U.S. Supreme Court found that the meaning of the Sherman Act meant that monopolies could be "good" or "bad". The court found that only the bad monopolies and bad monopoly practices were illegal under the Sherman Act. Whether a given practice was considered bad or good depended upon things like how much of the market was affected and how long the practice would last. The courts spent much of their time over the next few years ruling on the legality of specific practices.

14. The Per Se Doctrine.

   The U.S. Supreme Court can change its mind over time, which happened with its interpretation of anti-trust law in 1945 in the case U.S. v. Aluminum Co. of America (ALCOA). The Court found that the Sherman Act does not allow "good" monopolies and monopoly practices and disallow "bad" ones but, rather, simply disallows all monopolies and monopoly practices *per se* (because they exist).

Exemptions to Anti-trust Law

Some industries are exempt from anti-trust law, either because of an exemption written into the original law or because the exemption has subsequently arisen. The major industries exempt include:

1. Labor Unions

2. Agriculture

3. Sports Organizations

4. Export Industries

5. Regulated Industries

B. Regulation

The main issue here is why industries are regulated. Economists have offered three separate theories.

The Public Interest Theory

The public interest theory suggests that regulation is in the best interest of the public. For example, there are a number of instances where competitive markets fail to act efficiently and where regulation is often used to improve the market's efficiency. The discussion below just gives some examples but does not discuss all possible examples of regulation intended to serve the public interest.

- Regulating monopoly power, the case of the natural monopoly.

    Recall that a natural monopoly is one where the LRAC curve declines over the entire range of demand so that <u>one</u> firm can always produce the output demanded for a lower cost. As a result, a monopoly develops because of the competitive advantage a single firm has over larger numbers of more competitive firms. It is easy to show that the natural monopoly is more efficient than a competitive market but that the monopoly remains inefficient.

    The graph to the right shows a natural monopoly with average cost of producing 20,000 units of output of $10. Having two firms instead of one producing the total output will increase average costs to $20. Forcing the market to be truly competitive, with a large number of very small firms, will increase costs even more. This will in turn raise prices in the market.


    However, the monopoly firm remains inefficient just as is true of all monopolies, with price too high (albeit lower than in competition) and quantity too low. One way to address this is to regulate the natural monopoly as a method of reducing price.
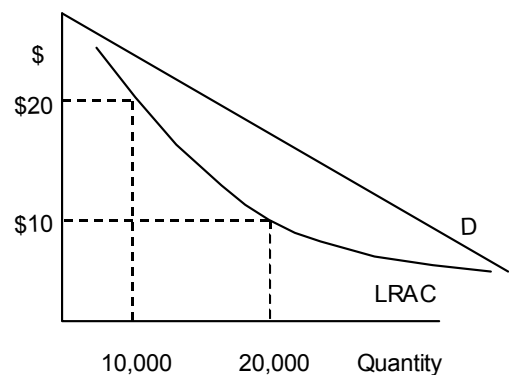
- Externalities

Recall our discussion of externalities in the topic titled "Monopoly and Perfect Competition Compared". We noted there that, as is shown in the graph below, that both positive and negative externalities result in allocative inefficiency.

Also recall that the common solution to these types of market failures is to regulate
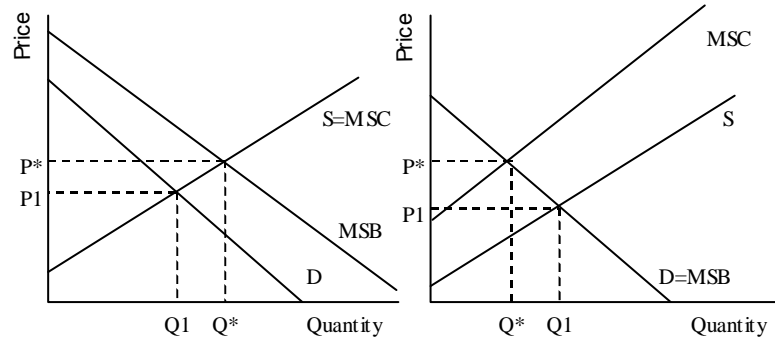
Positive Externality   Negative Externality

the market. In the case of the positive externality, where too little is produced the regulation is to increase output. For negative externalities, too much output is being produced so the regulation, as in the case of pollution, is to reduce the level of output.

The Capture Theory

The capture theory focuses on a second group that may also benefit from the regulation, the firms in the industry being regulated. The capture theory suggests that firms may try and often succeed in "capturing" the regulation and using it as a barrier to entry to keep out competitors and thereby increase their prices and profits.

One example that is often used here is professional licensing. All states in the U.S. require professionals such as medical doctors and lawyers to obtain a license before they can practice. The licensing requirements vary by state but commonly have education and testing requirements. Although it is often argued that such licensing is in the public interest by assuring quality it is certainly the case that the licensing increases the costs of entering the industry and, therefore, serves as an entry barrier.

The Public Choice Theory

The public choice theory is sometimes known as the political economy theory. This theory focuses on a third party beside the public (consumers) or the industry being regulated (firms). Rather, the public choice theory assumes that the regulation serves the best interest of those individuals who are doing the regulation, either as elected officials or as bureaucrats in a regulatory agency. The theory suggests that the regulators use the regulation, essentially "selling" it in an indirect manner, in order to make themselves better off.

Consider, for example, how elected officials might "sell" the regulation. First notice that the goal of all elected officials is to get re-elected. In our political system this is done by raising money from contributors, such as firms lobbying for regulation in their industry. Being re-elected also requires that you satisfy consumers to the point where they are willing to vote for you in the election.

Hence, the political economy theory suggests a tradeoff between regulation serving the interests of consumers and serving the interests of firms. Notice that if regulation never

serves the interest of consumers then consumers are likely to vote out of office those they hold responsible. However, also notice that if regulation never serves the interests of firms then they are likely to give money to other candidates who will eventually win elections against incumbents.

Hence, the political economy theory predicts that regulation will sometimes serve the needs of consumers and sometimes serve the needs of firms. Without going into details, the evidence in the literature generally finds this to be the case, which implies that the political economy theory best explains why regulation exists.