MERGER ANTITRUST LAW

LAWJ/G-1469-05	Tuesdays and Thursdays, 11:10 am – 1:10 pm
Georgetown University Law Center	Dale Collins
Fall 2023	wdc30@georgetown.edu
	www.appliedantitrust.com

READING GUIDANCE

The supplemental materials contain links to YouTube videos that explain the concepts we will examine in this class. If you are having trouble grasping one of the concepts, I encourage you to look at the related YouTube video or search the Internet for other explanations.

Class 10 (September 28): Competition Economics (Unit 8)

In this second of two classes on basic competition economics, we will examine four topics: substitutes and complements, elasticities and cross-elasticities, diversion ratios, and models for competitive, monopolistic, and imperfectly competitive markets. The concepts in this class are used on a daily basis in merger antitrust analysis and you will see them in expert reports, briefs, and court opinions. We will use them throughout the rest of the course.

Substitutes and complements. *Substitutes* are products that consumers substitute for one another so that their demands by consumers are negatively correlated. For example, if you need a printer and decide to buy a Brother, you will have little or no need for an HP printer. Conversely, if you buy an HP printer, you will have no remaining demand for a Brother printer. You will either buy one or the other (or some other printer model), but you do not need both. While substitutability may be an all-or-nothing proposition for some products (such as printers), it is a matter of degree for other products. For example, if you buy more fish, you will demand less meat, so fish and meat are substitutes even for consumers who purchase both products.

The essence of competition is to induce the consumer to buy the firm's product instead of the substitute products sold by the firm's competitors. A firm competes by selecting product attributes and prices that will encourage consumers to choose its products over its competitor's substitutes. By definition, horizontal mergers involve firms that sell substitute products.

The characteristic of substitutes that an increase in demand for product A will decrease demand for the substitute product B also has implications for the relationship between the demand for A (holding its price constant) and the price of B. If you lower the price of B, the demand for B will increase, and hence the demand for A will decrease. While the quantities demanded move in opposite directions for substitutes, the quantity demanded for a product and the price of its substitutes move in the same direction.

Complements are products whose demands are positively correlated. If demand for printers increases, then the demand for printer ink and toner will also increase, so printers and ink/toner are complements. Also, when the price of a product is increased, the demand for its complements will fall. If the price of printers increases significantly, fewer printers will be sold, and, as a result, customers will demand less ink/toner. Vertical mergers involve firms in the chain of manufacture or distribution of a product (such as an LCD manufacturer and a TV manufacturer or a car manufacturer and a car dealership) and are a common example of mergers involving complements. But the analytical principles and tools used in vertical mergers generally apply to all mergers involving complements.

Elasticities and cross-elasticities. Substitutability and complementarity can be matters of degree, so it is helpful to have a metric that will tell us whether, for example, two products are highly substitutable or weakly substitutable for one another. We could look at the numerical relationship between the quantities demanded, but this relationship will change if we change the unit of measure for one of the products. Say if you increase your demand for fish by one pound and, as a result, your demand for meat decreases by one pound, then the ratio of changes in demand ($\Delta q_{fish}/\Delta q_{meat}$ = 1/-1) is -1. But if we change the unit for fish to ounces while leaving meat in pounds, the ratio becomes -16. This problem of unit dimensions also affects the sensitivity of the demand for a product to changes in the product's price (i.e., the slope of the demand curve): the slope will be different if the price is measured in dollars as opposed to yen.

To deal with this problem, economists have created a metric known as *elasticity* that is independent of the units of measurement. Rather than measure changes in absolute units, elasticity measures changes in percentages. We begin by defining elasticity of demand (usually denoted by the Greek letter epsilon ε), which measures the sensitivity of a change in the quantity demanded to the change in the product's price:

$$\varepsilon = \frac{\%\text{change in quantity}}{\%\text{change in price}} = \frac{\frac{\Delta q}{q}}{\frac{\Delta p}{p}}$$

This is often called the *own-elasticity of demand* since the changes in quantity demanded and price relate to the same product. Rearranging the terms of the right-hand side of the above equation, we can also see that:

$$\varepsilon = \frac{\Delta q}{\Delta p}\frac{p}{q}.$$

The term $\Delta q/\Delta p$ is the slope of the demand curve, which is constant and negative when the demand curve is linear and downward sloping. Since the ratio of p and q changes along the demand curve, this means—somewhat counterintuitively—that the elasticity of demand continuously changes along a linear demand curve and is not a constant.

Demand is called inelastic when moderate percentage changes in price result in only small percentage changes in demand. In other words, inelastic demand is not very sensitive to price changes. Demand has unit elasticity when a given percentage change in price results in the same percentage change in demand. Demand is called *elastic* when a small percentage price change results in a large change in demand. Elastic demand is sensitive to price changes. We can summarize these definitions with the following formulae:

$$\varepsilon = \frac{\%\text{change in quantity}}{\%\text{change in price}} \begin{cases} > -1 & \text{Inelastic demand (low price sensitivity)} \\ = -1 & \text{Unit elasticity} \\ < -1 & \text{Elastic demand (high price sensitivity)} \end{cases}$$

These terms are used all the time in merger antitrust analysis, so it is important to get a feel for what they mean.

Note that a product's elasticity is bounded from above by zero. A zero elasticity means that the demand for the product is completely insensitive to changes in the product's price. In this case, demand is called *perfectly inelastic*. On the other hand, when demand for a product falls to zero for any increase in price, no matter how small, elasticity approaches minus infinity (-∞), in which case demand is called *perfectly elastic*.

As a quick aside, we should note that this is another area of economic sloppiness. Since the demand curve is downward sloping, the percentage change in quantity will always have the opposite sign of the percentage change in price, so ε will be a negative number. But since some people have problems thinking about relationships in negative numbers (e.g., -100 is smaller than -10), as a matter of convention, economists usually speak in terms of the *absolute value* of own-elasticity and so treat it as a positive number.[1] Now the definitions become:

$$|\varepsilon| = \left|\frac{\%\text{change in quantity}}{\%\text{change in price}}\right| \begin{cases} <1 & \text{Inelastic demand (low price sensitivity)} \\ =1 & \text{Unit elasticity} \\ >1 & \text{Elastic demand (high price sensitivity)} \end{cases}$$

Many people find this a much more natural way of thinking about elasticity since elasticity increases as the absolute values increase. When expressed in absolute numbers, inelastic demand is demand with an elasticity less than one and that approaches zero in the extreme. Conversely, elastic demand is demand with an elasticity greater than one and that approaches infinity as demand becomes more sensitive to price. Since economists rarely say whether they are speaking of true own-elasticity or its absolute value, you will need to figure that out from the context.

While the formalities of elasticity can be confusing, the basic concept (and the common usage) is very straightforward: Inelastic demand is not very sensitive to price changes, while elastic demand is sensitive to price changes. There are several reasons why demand may be inelastic, including:

1. *No substitutes*. If you must use a car to drive to work and to the store, there is no alternative but to buy gasoline to power the car. Moderate percentage changes are unlikely to affect the demand for gasoline. The U.S. Energy Administration estimates that gasoline's short-run elasticity is between -0.02 and -0.04, meaning it takes a 25% to 50% increase in the price of gasoline to decrease automobile travel by around 1%.[2]
2. *Little competition*. This is a special case of few substitutes. Beer prices in sports stadiums are higher than at grocery stores because stadiums usually allow only one vendor to sell

[1] The absolute value of a number is simply the value of the number without the sign and is denoted "| |." So |10| = |-10| = 10.

[2] *See* U.S. Energy Admin., Gasoline Prices Tend To Have Little Effect on Demand for Car Travel (Dec. 15, 2014); *but see* Lutz Kilian & Xiaoqing Zhou, *Federal Res. Bank of Dallas, Gasoline Demand More Responsive to Price Changes than Economists Once Thought* (June 16, 2020) (noting studies estimating the short-run U.S. price elasticity of gasoline demand to be between -0.27 to -0.37).

beer and prohibit fans from bringing their own beer, whereas grocery stores have to compete for beer sales with many other vendors.[3]

3. *Income level*. Elasticity of demand for any product is generally less for higher income level groups than for lower income groups. This is because consumers tend to become less price sensitive as the purchase price of a good becomes a smaller share of income. This factor can be generalized to the share of the consumer's consumption budget.
4. *Time period of measurement*. As a general rule, the longer the period of time over which elasticity is measured, the more elastic the demand. The idea is that consumers have little time to adjust during short time periods, while over longer time periods, more customers can adjust. For example, if gasoline prices increase sharply for a sustained time period, in the short run consumers must simply pay the higher price, but in the long run they can adjust by purchasing more fuel-efficient cars and thereby decrease their demand for gasoline.

The more inelastic a firm's residual demand curve, the more the firm can charge for its product. Much of anticompetitive behavior can be analyzed in terms of the conduct a firm pursues to decrease the elasticity of its residual demand curve by decreasing the relative availability or attractiveness of substitutes. For example, a horizontal merger of two significant competitors in a concentrated market will reduce the attractiveness of the merger partners' products as substitutes by coordinating the prices of the merged firm's products.

Cross-elasticity of demand measures the sensitivity of demand for one product to changes in the price of a different product:

$$\varepsilon_{ij} = \frac{\%\text{change in quantity of product } i}{\%\text{change in price of product } j} = \frac{\dfrac{\Delta q_i}{q_i}}{\dfrac{\Delta p_j}{p_j}}$$

If ε_{ij} is greater than zero—that is, if the demand for product *i* increases with an increase in the price of product *j*—then the two products are *substitutes*. If ε_{ij} is less than zero (i.e., a negative number), the two products are *complements*. If ε_{ij} is equal to zero, then the two products are unrelated. So with cross-elasticities of demand, the sign of ε_{ij} matters.

Note that the greater the value of cross-elasticity, the more sensitive the demand for product *i* is to price changes in product *j*. Fortunately, this is the natural way to think about cross-elasticities, so there is no need to employ absolute values. Cross-elasticities are often used to indicate how substitutable or competitive two products are. The higher the cross-elasticity, the more competitive the products. While own-elasticity ultimately will determine the firm's profit-maximizing price, it is the cross-elasticities that the firm will want to manipulate to reduce the availability or attractiveness of substitutes to make the firm's residual demand curve more inelastic and enable the firm to increase its prices.[4]

[3] *Query*: Why do stadiums only allow one vendor to sell in the stadium? (Remember, the answer to this question is "To make money." The second question is "How does this permit the firm to make money?").

[4] A slightly more rigorous explanation is that when the two merging partners were separate, they competed on price, but once they merge they will coordinate on price. So premerger, if firm 1 increased its price while firm 2 held its

Diversion ratios. Diversion ratios are another measure of the degree of substitutability or competitiveness between two products that have become as, if not more, important than cross-elasticity in modern merger antitrust analysis. Consider two products, *i* and *j*, that are substitutes. If the price of product *i* increases, then two things happen: the quantity demanded of product *i* will decrease and the quantity demanded of product *j* will increase. The diversion ratio D_{ij} is the increase in the demand in product *j* divided by the total loss of demand in product *i*:

$$D_{ij} = \frac{\text{Units captured by Firm } j \text{ as a result of Firm } i\text{'s price increase}}{\text{Total units lost by Firm } i \text{ as a result of Firm } i\text{'s price increase}} = \left| \frac{\Delta q_j}{\Delta q_i} \right|$$

When the two products substitute one-for-one for one another (e.g., you either buy one red car or one blue car), you can think of the diversion ratio as the percentage of the lost demand for product *i* that is *diverted* to product *j*. By convention, diversion ratios are positive numbers, which is why we look at the absolute value of the right-hand side of the equation. As with cross-elasticities of demand, the higher the diversion ratio between two products, the more substitutable and hence more competitive the two products are with one another.[5]

Perfectly competitive markets. Antitrust law is about preserving competition, so it is helpful to have a model for markets that exhibit the maximum degree of competition. Economists call these *perfectly competitive markets*. These markets are characterized by firms that are *price-takers*, that is, firms that act as if their individual choice of production level will not affect the market-clearing price. The usual way to envision this is to think of a market with many firms, each producing only an infinitesimal fraction of total market supply. If each firm only produces a small fraction of total supply, then no firm would perceive that a small change in its production level would have any influence on the market-clearing price even if the aggregate demand curve is downward sloping.

One key result in competitive markets is that each firm chooses its production level so that its marginal cost equals the market price. Like all profit-maximizing firms, the price-taking firm sets its production level so that its marginal revenue equals its marginal cost. But in a perfectly competitive market, each firm perceives that its choice of production level will not change the market-clearing price in any way. Consequently, the firm perceives that no downward adjustment of the firm's price is necessary to sell out its products when it increases its production. As a result, the firm perceives its marginal revenue as the prevailing market price. A second key result is that a perfectly competitive market produces the maximum consumer surplus of any market structure. Accordingly, a perfectly competitive market is the ideal market structure for antitrust under a consumer welfare standard.

Perfectly monopolized markets. The polar opposite of a perfectly competitive market (think a multitude of firms) is a perfectly monopolized market (think one firm). In this case, the

---

price constant, there will be some substitution by firm 2's product for firm 1's product. But once the merger takes, if the price of firm 1's product is increased and the merged firm management also instructs firm 2 to increase its price as well, firm 2' product will not be as attractive a substitute as it was premerger. This decreases the cross-elasticity between the products of the merged firm, decrease the own-elasticity of firm 1's products, and hence enables firm 1 to increase the price of its product postmerger.

[5] You probably noticed that there are two diversion ratios for any pair or products: one for the diversion of product *i* to product *j* (D_{ij}) and one for the diversion of product *j* to product *i* (D_{ij}). As a general matter, the two diversion ratios need not be the same and indeed may differ considerably.

monopolist's residual demand curve is the same as the aggregate demand curve for the market as a whole. The monopolist, as a profit-maximizing firm, sets its production level so that its marginal revenue equals its marginal cost, but now the monopolist is very cognizant of the effect its choice of its production level has on the market-clearing price. The upshot is that the monopolist chooses a production level lower than the aggregate production level in a perfectly competitive market facing the same aggregate demand curve and charges a higher price. Indeed, in a homogenous product market with a single price, the perfectly monopolized market shifts the maximum amount of wealth from consumers to producers and creates the maximum profits for producers.[6] It is the worst market structure from a consumer welfare perspective for the market.

Imperfectly competitive markets. In between a perfectly competitive market structure and a perfectly monopolized one, there exist imperfectly competitive markets. These are markets where at least some firms recognize that their choice of production level will affect the market price, although this effect will depend not only on any given firm's choice of production level but also on the choices of all firms taken together.

One measure of the degree of market power a firm possesses is the *Lerner index* (L). The Lerner index is the difference between the market price and the firm's marginal cost and dividing it by the market price:

$$L = \frac{p - mc}{p}$$

As you can see, in a perfectly competitive market (where $p = mc$), each firm's Lerner index is zero. The Lerner index reaches its maximum when the market-clearing price is the monopoly price.

We can generalize the Lerner index to the market as a whole. Let L_i be the Lerner index for firm i given the firm's price and marginal costs. Summing these indices across all n firms in the market weighted by each firm's market share s_i gives the market' Lerner index:

$$L = \sum_{i=1}^{n} L_i s_i = \sum_{i=1}^{n} \frac{p_i - mc_i}{p_i} s_i .$$

Think of the market Lerner index as the market-share weighted average of the percentage price-cost margins of all firms in the market.

The class notes address three models of imperfect competition. Each of these models is used in merger antitrust analysis. The choice of the model to use depends on which model most closely approximates the conditions of the actual market being analyzed.

The *Cournot model* is used for markets that closely resemble a homogeneous product market. In a true homogeneous product market, all the products the various firms produce are identical, so consumers buy only from firms charging the lowest price. This has four important implications:

[6] Homogeneous products are products that consumers perceive are completely identical to one another, so consumers pick the product to purchase based entirely on the product's price.

1. There is only a single price in the market.[7] Any firm offering its product at a higher price will sell nothing.
2. Since an individual firm cannot meaningfully set its own price, the firm's *control variable* is its production level.
3. The market-clearing price will depend on the aggregate production Q by all of the firms in the market (i.e., the sum of each firm's production level):

$$Q = \sum_{i=1}^{n} q_i.$$

   where there at n firms in the market.
4. The Lerner index for the market in a Cournot model is a function of market concentration and the elasticity of aggregate demand (in absolute value):

$$L = \frac{HHI}{|\varepsilon|}$$

   where ε is the elasticity of aggregate demand and HHI is the *Herfindahl-Hirschman Index* (HHI). The HHI is equal to the sum of the squares of the market shares s_i for every firm in the market:

$$HHI = s_1^2 + s_2^2 + \cdots + s_N^2 = \sum_{i=1}^{n} s_i^2.$$

   The HHI today is the primary measure of market concentration in merger antitrust analysis. Note that as the market concentration (as measured by the HHI) becomes larger, the Lerner index increases, indicating that more market power is being exercised in the market.[8] The HHI ranges from zero (where there are so many firms that the market share for each firm is close to zero) to 10,000 (where there is only one firm).[9]

Firms in a Cournot model compete in their choice of outputs. A *Nash equilibrium* in a Cournot model is defined to be an output level q_i for each firm i such that q_i maximizes the firm's profit where the market price is $p(Q)$ (see #3 above) and all of the other firms are assumed to hold their production levels constant at their respective original profit-maximizing levels. Under this assumption, no firm in a Nash equilibrium has a profit-maximizing incentive to change its production level. You do not have to be able to do the math to find a Nash equilibrium, but it will be helpful if you have an idea of the basic outline of the technique. Here is an illustration in a two-firm Cournot duopoly:[10]

[7] Remember, we are assuming that there are no search costs for consumers. If there are search costs, then homogeneous markets can sustain multiple prices in equilibrium.

[8] This is one of the theoretical justifications for relating increases in concentration to a reduction of competition, which, as we will see, is a central (rebuttable) presumption in the merger antitrust analysis of horizontal mergers.

[9] This is the case where market shares are measured from 0 to 100. The Merger Guidelines use this measure. When market shares are measured fractionally from 0 to 1, the HHI ranges from 0 to 1.

[10] A duolopy is a market with only two firms.

1. Start with the first-order condition $mr(q_1, q_2) = mc(q_1)$. In this case, marginal revenue is a function of both q_1 and q_2 (which together add to Q, which in turn determines the market-clearing price).
2. Solve the first-order condition for q_1, which will be a function of marginal cost and q_2. This yields a *reaction function* for firm 1 since it tells firm 1 its profit-maximizing output level given any level of firm 2's output.
3. Do the same thing for firm 2 to derive firm 2's reaction function, which will depend on firm 2's marginal cost and q_1.
4. This yields two reaction functions with two unknown variables. As a general rule, you can solve for n unknown variables if you have n independent equations. Solve this system of two equations for the unknown profit-maximizing quantities q_1* and q_2*.

If you want to see an example of the Nash equilibrium in a two-firm Cournot model with linear demand, take a look at How to Solve Cournot Problem: Algebra-Based Solution, a short YouTube video.

While the control variable in a Cournot model is quantity, the control variable in a *Bertrand model* is price. While in a homogenous product market, the aggregate production of the individual firms determines the market price, in a Bertrand model, each firm has some control over the price of its product. This makes the Bertrand model useful in thinking about differentiated product markets, where firms produce products that are not identical but still compete with one another to a greater or lesser extent. Because different consumers can have different preferences over these various differentiated products, the market supports multiple prices. Each firm confronts its own distinct residual demand curve that allows it to set its own prices in its profit-maximizing interest.

Smartphones are a good example of a differentiated product market. Apple iPhones and various models of Android phones all compete with one another, but given the differences in consumer preferences, each smartphone manufacturer can set its own price in its profit-maximizing interest, taking into account the prices that its competitors set for their phones and the attributes (think functionality or quality differences) of those competing phones.

A Nash equilibrium in a Bertrand model is defined analogously to one in a Cournot model. A Nash equilibrium in a Bertrand model is a price p_i for each firm i such that p_i maximizes the firm's profit assuming that the price levels of all of the firm's competitors are fixed. The method for solving for a Nash equilibrium in a Bertrand model is also analogous to finding a Nash equilibrium in a Cournot model, except that the first-order condition (FOC) must be derived from the firm's residual demand function and is expressed in terms of prices rather than outputs. When we derived the first-order conditions in the last class, we implicitly used a Cournot assumption that firms set production levels rather than prices. Therefore, we used inverse demand functions as our point of departure for deriving the firm's first-order conditions, which were expressed in terms of outputs. By contrast, the control variable in a Bertrand model is price, so we use the firm's residual demand function to derive the first-order conditions for a profit maximum. If you want to see a Bertrand oligopoly solved for its Nash equilibrium, see Bertrand Oligopoly on YouTube.

The choice of using a Cournot model instead of a Bertrand model in practice depends on whether the firms in the market use production levels or price as their control variables. Since truly homogenous product markets are rare, most modern applied empirical analyses use Bertrand models.

One interesting (and empirically relevant) property of Bertrand models is that the more the differentiated products compete with one another (i.e., the higher the cross-elasticity of demand among the products in the market), the more a Nash equilibrium in a Bertrand model looks like a perfectly competitive equilibrium. Indeed, when the differentiation disappears and the products become identical, the Nash equilibrium is the same as a perfectly competitive market. Conversely, as the differentiation increases and the products compete less with one another, the Nash equilibrium approaches the perfectly monopolized equilibrium where each firm is in its own separate market. An important feature of a Bertrand model with homogenous products is that it only takes two firms to achieve the perfectly competitive equilibrium!

The last model we will examine is a *dominant firm with a competitive fringe*. The math here can be a bit complex and the diagrams are almost unintelligible, but the idea is very simple. Consider a market with a dominant firm and some number of capacity-constrained small firms. Each small firm perceives that its production level is so small that changes in output will not affect the market-clearing price and hence the firm acts as a price-taker (i.e., a perfectly competitive firm). As a result, if the market-clearing price is at or above the small firm's marginal cost, it will produce at capacity. Conversely, if the market-clearing price is below the small firm's marginal cost, it will shut down production. Say that each small firm has a constant marginal cost c and collectively their maximum production given their capacity constraints is Q_f (the subscript f is for "fringe"). Assume that the aggregate demand curve is $Q(p)$. The residual demand curve $q_d(p)$ for the dominant firm is then:

$$\begin{aligned} q_d(p) &= Q(p) - Q_f \quad && \text{if } p \geq c \\ &= Q(p) && \text{if } p < c \end{aligned}$$

The dominant firm then maximizes its profits subject to its (discontinuous) residual demand curve, taking into account the production of the fringe firms if it sets a production level that yields a market-clearing price above the fringe firms' marginal cost.

This model can easily be modified to account for increasing marginal costs or marginal costs that differ among the fringe firms. It can also be modified to apply to a group of large firms that are collectively dominant by applying a Cournot model to the dominant group subject to a group residual demand that accounts for fringe firm production.

If you want to see a dominant firm model solved, look at Economics of the Price Leadership (Dominant Firm) Model on YouTube.[11]

If you have any questions or comments, send me an e-mail. See you in class.

[11] The model in the YouTube video assumes that the fringe firms have increasing marginal costs, which implies that the aggregate supply curve of the fringe firms is upward sloping (i.e., the higher the market-clearing price, the more output the fringe firms collectively provide).