M

make or buy decisions

Robert E. McAuliffe

A firm's decision to make an input in production rather than buy that input in the market is an important one that firms frequently fail to reconsider. What a firm chooses to produce is one of the most fundamental decisions it can make. But if markets are so efficient in providing resources, why should firms even exist? Coase (1937) answered that there were costs from using the market to obtain resources just as there are costs of producing them within the firm. The TRANSACTIONS COSTS of using the market must then be weighed against the coordination costs of producing the input internally.

Williamson (1985) argued that internal production requires additional bureaucracy to manage the new activity along with higher production costs. But if the firm obtained the input in the market, it would be costly (in terms of time) to locate suitable suppliers, monitor quality levels, and insure that supplies will be delivered on time. When the part is unique and very specialized, it is costly to use the market because there may only be one or two suppliers. In this case, the firm may decide to produce the part internally.

Porter (1980) suggests that VERTICAL INTEGRATION may also achieve strategic goals for the firm. For example, a firm may want to improve the quality or lower the production costs of its product but would need closer coordination with its suppliers to accomplish this goal. Another possibility is that internal production will improve the firm's flexibility to produce new products.

Bibliography

Coase, R. H. (1937). The nature of the firm. *Economica*, **4**, 386–405.

Keat, P. G. and Young, P. K. Y. (2002). *Managerial Economics: Economic Tools for Today's Decision Makers*, 4th edn. Upper Saddle River, NJ: Prentice-Hall.

Perloff, J. M. (2004). *Microeconomics*, 3rd edn. Boston: Pearson Addison-Wesley.

Porter, M. E. (1980). *Competitive Strategy: Techniques for Analyzing Industries and Competitors*. New York: Free Press.

Shughart, W. F. (1990). *The Organization of Industry*. Boston: Richard D. Irwin.

Williamson, O. E. (1985). *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting*. New York: Free Press.

marginal cost

Robert E. McAuliffe

The marginal cost is the change in TOTAL COSTS due to a unit (or incremental) change in output. That is, for discrete changes in output, marginal cost is given by:

$$\mathrm{MC} = \frac{\Delta \mathrm{TC}}{\Delta \mathrm{q}} \tag{1}$$

where MC is the marginal cost, ΔTC is the change in total costs, and Δq is the change in output. In the SHORT RUN, FIXED COSTS do not vary with output, so marginal costs can also be written as:

$$\mathrm{MC} = \frac{\Delta \mathrm{TVC}}{\Delta \mathrm{q}} \tag{2}$$

where ΔTVC is the change in TOTAL VARIABLE COST. If the total cost function is known or has been estimated, marginal cost would be the derivative of total costs with respect to output. Therefore, if total costs are:

$$\mathrm{TC} = 1{,}000 + 5 \times q + 8 \times q^2 \qquad (3)$$

marginal costs would be:

$$\mathrm{MC} = \frac{d\mathrm{TC}}{d\mathrm{q}} = 5 + 16 \times q \qquad (4)$$

Understanding marginal costs is crucial for optimal decision-making because, unless there are resource constraints, any activity should be continued as long as the additional (or marginal) benefits exceed the marginal costs. If two or more activities have marginal benefits that exceed their marginal costs but a manager cannot fully fund all of them, then each activity should be funded until each provides the same marginal benefits per dollar (*see* PROFIT MAXIMIZATION; UTILITY MAXIMIZATION). For any *change* in an activity (production, pricing, ADVERTISING, etc.), a manager must compare the marginal benefits of the change against the marginal costs to make the best decision.

For example, consider a manager whose only variable input is labor. The change in total costs will then be the change in labor input (ΔL) multiplied by the wage rate (w), and since marginal costs are the change in total variable costs divided by the change in quantity (Δq), from equation (2) we have:

$$\mathrm{MC} = \frac{\Delta\mathrm{TVC}}{\Delta q} = \frac{w \times \Delta L}{\Delta q} = w \times \left\{\frac{\Delta L}{\Delta q}\right\} \qquad (5)$$

The term in braces is the inverse of MARGINAL PRODUCT of labor, so equation (5) shows the relationship between marginal costs, variable input costs (the wage rate, w), and productivity (the marginal product of labor). An increase in the cost of variable inputs, w, will increase the marginal cost of production, while an increase in productivity (a rise in the marginal product of labor) will reduce marginal costs, as would be expected (see Shughart, Chappell, and Cottle, 1994; Friedman, 1976).

As Baumol (1977) notes, managers frequently use average values rather than marginal values when making decisions, and this can lead to incorrect conclusions. This occurs because much of the accounting information provided to managers is in the form of average of total figures. In addition, marginal calculations may require information a company does not yet have, such as the marginal benefits and costs of an increase in advertising. The important issue is not whether past advertising expenditures have generated net revenues that exceeded the costs, but whether any increase (or decrease) in advertising is justified. Finally, because marginal costs reflect the change in total costs for the last unit(s) produced, they reflect changes more quickly than average total costs, which are based on the costs for all units produced to that point (*see* AVERAGE TOTAL COST). Thus even if the average increase in net revenues from advertising fell by a small amount, the *marginal* change could be substantial and profits could be increased by reducing advertising expenditures.

Bibliography

Baumol, W. J. (1977). *Economic Theory and Operations Analysis*. Englewood Cliffs, NJ: Prentice-Hall.

Keat, P. G. and Young, P. K. Y. (2002). *Managerial Economics: Economic Tools for Today's Decision Makers*, 4th edn. Upper Saddle River, NJ: Prentice-Hall.

Friedman, M. (1976). *Price Theory*, 2nd edn. Chicago: Aldine.

Salvatore, D. (2004). *Managerial Economics in a Global Economy*, 5th edn. Mason, OH: South-Western Publishing.

Shughart, W. F., Chappell, W. F., and Cottle, R. L. (1994). *Modern Managerial Economics*. Cincinnati, OH: South-Western Publishing.

marginal product

Lidija Polutnik

The marginal product of any input, such as labor or capital, measures the change in TOTAL PRODUCT as a result of a small change in the usage of that input, holding all other inputs constant. For example, the marginal product of labor or capital is the additional output from using one more labor or machine hour.

The marginal product of an input that is infinitely divisible can be obtained by taking the

partial derivative of the production function with respect to that input (*see* PRODUCTION FUNCTIONS). In the relevant range of production, marginal product is positive. An increase in the marginal product (or an increase in productivity) of an input means that through better technology, more efficient management, or an increase in work effort, the company is able to obtain more output from the same amount of inputs. Alternatively, marginal product can be improved through the restructuring of the company by decreasing the number of work hours while holding the addition to total output constant.

The total product and marginal product curves demonstrate the LAW OF VARIABLE PROPORTIONS. This law states that as the quantity of a variable input increases, holding the amount of other productive factors constant, beyond some point the marginal product of the variable input begins to decrease; total product continues to increase, though, at a decreasing rate. DIMINISHING RETURNS arise because the fixed amount of plant and equipment is gradually spread among an even greater number of workers, leaving a smaller amount of CAPITAL to work with for each variable input.

By enabling us to measure the change in total product as a result of a change in a variable input, marginal product permits comparisons with the short run cost of production (*see* SHORT RUN COST CURVES). If the costs of labor and machine hours are known, we can calculate the MARGINAL COST of production by dividing an input's cost by the marginal product of the input. Therefore, marginal costs are inversely related to the marginal product of the variable inputs. For example, when the marginal product of labor is falling, the marginal cost of output is increasing, and vice versa. When the marginal product of labor is constant, additional units can be produced with constant marginal cost.

Knowledge of the marginal products of inputs enables managers to determine the short run optimal level of employment of capital and labor in the production process. In order to maximize total profit (*see* PROFIT MAXIMIZATION), inputs should be hired up to the point where the marginal input cost equals the MARGINAL REVENUE product of labor. Marginal input cost is defined as the amount that an additional unit of the variable input adds to TOTAL COSTS. Marginal revenue product is defined as the amount that an additional unit of the variable unit adds to TOTAL REVENUE and equals the marginal product of the variable input times the marginal revenue resulting from the increase in output produced.

Bibliography

Douglas, E. J. (1992). *Managerial Economics*, 4th edn. Englewood Cliffs, NJ: Prentice-Hall.

Perloff, J. M. (2004). *Microeconomics*, 3rd edn. Boston: Pearson Addison-Wesley.

Salvatore, D. (2004). *Managerial Economics in a Global Economy*, 5th edn. Mason, OH: South-Western Publishing.

marginal rate of substitution

Gilbert Becker

The marginal rate of substitution of good X for good Y ($\text{MRS}_{x,y}$) is defined as the amount of good Y which the consumer is willing to give up in exchange for one additional unit of good X, in order to maintain the same level of TOTAL UTILITY. Since the same total utility is achieved from the two choices, $\text{MRS}_{x,y}$ concerns the movement along a consumer's indifference curve (*see* INDIFFERENCE CURVES). Specifically, it measures the slope of the indifference curve between any two points. This identifies the trade-off of one good for another that the consumer is *willing* to make, independent of the prices of the two goods which indicate the trade-off conditions *required* by the market.

The marginal rate of substitution between two goods describes the relative importance of the two goods to the consumer and it is defined as the ratio of the marginal utilities of the two goods. For example, a $\text{MRS}_{x,y}$ with the value of 3 to 1 indicates that the consumer is currently willing to give up 3 units of good Y in order to receive 1 extra unit of good X. Here, the MARGINAL UTILITY of good X(MU_x) is three times that of good Y. This can be formally expressed as:

$$\text{MRS}_{x,y} = -\frac{\Delta Y}{\Delta X} = \frac{\text{MU}_x}{\text{MU}_y} = 3$$

where Δ indicates the change in the quantity of the good in question.

Since the marginal utility of either good will typically decrease with greater consumption of that good, the marginal rate of substitution and therefore the slope of the indifference curve is typically not constant. One exception is the case where the two goods in question are perfect SUBSTITUTES. Here the marginal rate of substitution will be constant, e.g., 2 aspirin may be substituted for every 1 acetaminophen tablet.

The rate at which consumers are willing to trade off one product for another can be altered by changing their tastes. Here, ADVERTISING and other forms of promotion can be useful.

See also *diminishing marginal utility; utility maximization*

Bibliography

Maurice, S. and Thomas, C. (2002). *Managerial Economics*, 7th edn. Boston: McGraw-Hill Irwin.

Nicholson, W. (2000). *Intermediate Microeconomics and its Application*, 8th edn. New York: Harcourt.

marginal rate of technical substitution

Gilbert Becker

The production of virtually every good requires the use of two or more inputs such as labor, capital, and raw materials. Typically, these inputs can be used in varying proportions to generate the same level of output. For example, in agriculture varying amounts of labor and capital can be used to plant or harvest the same number of acres of land per day. Modern farming techniques in the US and other developed nations use large harvesting combines requiring relatively few workers, while less sophisticated farming operations often use a highly labor-intensive approach. Since reducing the amount of any one input generally tends to reduce total output, a greater quantity of a second input is needed to compensate for the lost output. The marginal rate of technical substitution measures the rate at which any one input must be substituted for one unit of another input in order to maintain a constant level of output.

Formally, the production of varying levels of output using different combinations of inputs is described using an isoquant map (*see* ISOQUANT-ISOCOST CURVES). The marginal rate of technical substitution is measured by the slope of a given isoquant curve between any two points on that curve. Since increased use of any one input tends to reduce that input's marginal productivity, the marginal rate of technical substitution tends to decrease with greater substitution toward an input. That is, the isoquant curve tends to be convex.

The marginal rate of technical substitution between inputs is important information for managers selecting the optimal combination of inputs. Just as the consumer trade-off between two goods is determined by the ratio of the MARGINAL UTILITY of the two *goods* (*see* MARGINAL RATE OF SUBSTITUTION), the marginal rate of technical substitution is determined by the ratio of the MARGINAL PRODUCT of the two *inputs*. Cost minimization requires that the marginal rate of technical substitution between inputs be equal to the ratio of the prices of those inputs. This is the production requirement equivalent to the consumer optimization problem (*see* UTILITY MAXIMIZATION).

Bibliography

Browning, E. and Zupan, M. (2004). *Microeconomics: Theory and Applications*, 8th edn. New York: John Wiley.

Maurice, S. and Thomas, C. (2002). *Managerial Economics*, 7th edn. Boston: McGraw-Hill Irwin.

Perloff, J. M. (2004). *Microeconomics*, 3rd edn. Boston: Pearson Addison-Wesley.

marginal revenue

Gilbert Becker

Marginal revenue is the change in the TOTAL REVENUE resulting from a one-unit increase in the sales of a product. It is important to a firm because PROFIT MAXIMIZATION requires the firm's managers to find the output level where this extra revenue equals the MARGINAL COST. While greater sales resulting from price reductions generally cause marginal revenue to fall, marginal cost eventually rises with greater

output. Thus, any output below the level where the two are equal leaves the possibility for additional profits to be made, but output beyond that level costs more to produce than the gain in revenues from the additional sale.

The nature of the relation between marginal revenue and the level of output sold varies with the basic type of market under consideration (*see* BASIC MARKET STRUCTURES). Under PERFECT COMPETITION the firm, being too small to influence market outcomes, can sell as much output as it desires without changing the market price. As such, for any additional sales, the marginal revenue is constant at the level of the market-generated price. In other words, the firm faces a demand which is perfectly elastic for its own product (*see* ELASTICITY).

Under OLIGOPOLY and other forms of imperfect competition, additional sales may occur as a result of a change in the quality of the product or a change in ADVERTISING (*see* DORFMAN–STEINER CONDITION), or they may be brought about by a change in the product's price. The LAW OF DEMAND indicates that, all else being equal, an increase in the number of units sold requires a reduction in price. In the absence of PRICE DISCRIMINATION, the marginal revenue from an additional unit sold declines as the lower price needed to generate a larger number of unit sales will offset the gains from the increased quantity sold. Moreover, the estimation of marginal revenue becomes more difficult in oligopoly since the amount of additional sales resulting from a price reduction depends in part on the reaction of rivals, who may or may not respond with price cuts of their own.

Although it is not always consistent with profit maximization, managers may at times wish to maximize sales revenues (*see* OBJECTIVE OF THE FIRM). Here, marginal revenue and its relation to price elasticity of demand are important. Where demand is elastic, a decline in price results in a large increase in unit sales, and total revenue is growing. In this case, the marginal revenue from the additional units sold is positive. For revenue maximization, production should continue to increase until marginal revenue is zero, i.e., where demand is unit elastic.

See also perfect competition; profit maximization

Bibliography

Maurice, S. and Thomas, C. (2002). *Managerial Economics*, 7th edn. Boston: McGraw-Hill Irwin.

Perloff, J. M. (2004). *Microeconomics*, 3rd edn. Boston: Pearson Addison-Wesley.

Schiller, B. (2003). *The Micro Economy Today*, 9th edn. New York: Irwin McGraw-Hill.

marginal utility

Gilbert Becker

Marginal utility is defined as the change in the TOTAL UTILITY resulting from the consumption of one additional unit of a good by a consumer. For any individual good, consumer behavior typically follows the law of DIMINISHING MARGINAL UTILITY, which states that in any given time period, as the rate of consumption of a good rises, the additional utility or satisfaction acquired by the consumer eventually declines.

It can be shown that for an individual who is purchasing several goods simultaneously, the maximization of total utility requires equating the marginal utility per dollar spent on each good. More formally, this ideal combination of goods occurs when

$$MU_1/P_1 = MU_2/P_2 = \ldots = MU_i/P_i \quad (1)$$

where MU and P indicate the marginal utility and price of the ith good. An imbalance between any two of these ratios, for example

$$MU_2/P_2 > MU_1/P_1 \quad (2)$$

would result in a situation where a reallocation of the consumer's income toward purchasing more of good 2 would cause a net increase in total utility. This is because the additional satisfaction gained from consumption of an extra unit of good 2 would more than offset the loss of utility from the decreased consumption of good 1. An inequality such as in equation (2) would occur if, for example, the price of the second good, P_2, was decreased. The equation demonstrates why a decrease in the price of this good would lead to a shifting away from good 1 and an increase in the purchase of good 2 (*see* SUBSTITUTION EFFECT). This helps to generate the negative

relationship between the price of good 2 and its quantity demanded, which is the foundation of the LAW OF DEMAND. Similarly, equation (2) can be used to demonstrate how a manager may create a more favorable position for a product by using increased ADVERTISING. This strategy, when successful, creates imbalances between the ratios by altering the marginal utility of one good relative to the others.

The marginal utility of income to a gambler, or profit to an investor, can be used to identify the three categories of risk behavior which individuals may portray. An individual is identified as a risk preferrer if his marginal utility of income grows at an increased rate as his level of income rises. An individual is said to show RISK AVERSION if the marginal utility of an extra dollar of income grows at a decreasing rate as income rises, and RISK NEUTRALITY if the marginal utility of an extra dollar of income is constant (see Varian, 1992). A risk averter typically sees risk as being undesirable, and is willing to tolerate greater uncertainty in an outcome only if rewarded with a greater expected return. A risk lover is a gambler who prefers the increased uncertainty (*see* EXPECTED PRESENT VALUE).

See also *marginal rate of substitution*

Bibliography

Maurice, S. and Thomas, C. (2002). *Managerial Economics*, 7th edn. Boston: McGraw-Hill Irwin.
Nicholson, W. (2000). *Intermediate Microeconomics and its Application*, 8th edn. New York: Harcourt.
Schiller, B. (2003). *The Micro Economy Today*, 9th edn. New York: Irwin McGraw-Hill.
Varian, H. R. (1992). *Microeconomic Analysis*, 3rd edn. New York: Norton.

market definition

Gilbert Becker

An economic definition of the term "market" is the group of all firms willing and able to sell a similar product or service to the same potential buyers. As such, the market definition for a given product is the identification of the relevant group of sellers who may be seen as being in COMPETITION with one another for the sale of that product. Two boundaries are used to define a market:

1 the product boundary;
2 the geographic boundary.

Properly identifying these boundaries for a particular market will result in a more accurate count of the number and size distribution of firms in that market. This information is useful for a firm in developing its competitive strategy. For markets under PERFECT COMPETITION the product boundary is relatively easy to establish since the product is homogeneous across firms. Given the complexity of most markets though, and especially those where PRODUCT DIFFERENTIATION exists, this boundary becomes blurred. Similar difficulties arise in identifying correct geographic boundaries. Thus, managers should be aware that the term market may be a source of significant confusion. Sales representatives and other subordinates may erroneously use the term to describe their personal selling area or limit their definition to the specific product line associated with their selling effort. These approaches may underestimate the true scope of the economic market in question.

Market definition is also essential for the government and the courts in developing antitrust policy and in deliberating antitrust cases. For example, the CLAYTON ACT proscribes mergers which substantially lessen competition "in any line of commerce" (i.e., product market) and "in any section of the country" (i.e., GEOGRAPHIC MARKET). Thus the proper identification of market boundaries will be critical in indicating the degree to which competition may have declined. An expanded market definition, for example, which includes a greater number of rivals, will naturally tend to diminish the market share of the merging firms and thus make the merger less potentially onerous in the eyes of the antitrust enforcers. Moreover, in cases of monopolization, proper identification of the market is essential to determine the existence of MARKET POWER, which is the first necessary requirement for a finding of an antitrust violation (*see* MICROSOFT ANTITRUST CASE; SHERMAN ACT).

The central issues in market definition are where and how these boundaries are to be drawn. Numerous authors including Church and Ware (2000) have recognized that market definition and measurement, when applied to antitrust policy, may differ from that of the simple economic definition given above. As such, several tools have been introduced and are categorized below as either theoretical or practical measures, each having its own relative strength in terms of its use in purely economic versus public policy applications.

Theoretical Measures

Cross elasticities are one set of market measurement tools which have been used for decades. The cross elasticity of demand between two goods measures the percentage change in unit sales of one firm's product given a 1 percent change in the price of a rival's product. As such it measures the degree to which the two goods are substitutes in the eyes of consumers. A high cross-elasticity measure indicates close substitutability and thus the firms in question should be counted as being in the same market. Similarly, cross elasticities of supply may be considered. High values for this measure indicate a firm's willingness and ability to produce more of their own good in reaction to a rival's price increase. It may also indicate a similarity in production technological processes between firms and the potential to switch product lines and readily and easily enter a product market where price has been increased above the competitive level. As a result, both measures of cross elasticities are needed to properly define the relevant product market. These measures attempt to determine an economic definition of a market by examining relevant rivals who establish a market's price.

While these measures are conceptually appealing, Scherer and Ross (1990), Shepherd and Shepherd (2004), and others have identified several difficulties that arise in their application. First, what constitutes a "high" value for cross elasticity is necessarily an arbitrary decision. Moreover, possible values for these measures fall along a continuum. Distinct gaps between high and low values which could indicate a break in the substitutability continuum for these goods may not arise in the actual estimates examined. Third, two different estimates for the cross elasticity of demand between two goods A and B will be generated for the relatively similar cases of a 1 percent price increase in product A, and a 1 percent price decrease in product B.

Finally, the correct estimation of cross elasticities may require using price changes originating from the competitive price level, which may not be readily available. Where market power already exists and a firm's price is set significantly above that of its rivals, its cross price elasticity of demand may be higher since further price increases may induce considerably more consumer reaction than before. A small price increase for product A may induce a larger change in product B's sales when A's price is 50 percent above B's than when it is only 5 percent above that of B. Thus, if a monopolist already using its market power has set its price above the competitive level, the cross elasticity measured at that price level may be high, indicating the existence of many rivals and a broad market. The high cross elasticity measure here may lead to the erroneous conclusion that market power is limited, when in fact that power is already in use.

As a result, more recent antitrust analysis has focused on defining a market by examining the existence of sustainable market power. Perhaps the first such measure is the Lerner index, which examines the extent by which a firm's price exceeds its marginal cost. Since the Lerner index can be shown to be inversely related to a product's own price elasticity of demand, the latter may also be used to help identify an antitrust market.

Boyer (1979) developed the basis of a theoretically more advanced measure which has been used by the federal antitrust enforcement agencies. This method can be used for both product market and geographic market definitions. The method involves examining a hypothetical monopolist in the geographic area in question, and its ability to introduce a small but significant nontransitory increase in prices (SSNIP) above the current level. Boyer argues that a market can be identified, from the point of view of a single firm, as the smallest group of rivals necessary to organize in order to successfully act as a cartel by imposing such a SSNIP (*see* cartels). If this group of firms (or hypothetical monopolist)

could raise prices without bringing about reaction by other firms that was sufficient to force a retreat in the cartel price hike, then these other firms can correctly be seen as being outside of the market. This may be due to their producing a sufficiently different (nonsubstitutable) product, or because their business takes place in a geographically distinct location, or even simply because their reaction was slow or limited. If, on the other hand, the initial group of firms in question could not successfully collude for a significant period of time without feeling the disciplinary effects of the other rivals' increases in supply, then the relevant market should be expanded to include these additional firms. In revising its guidelines for mergers in 1982 and 1992, the Department of Justice (DOJ) introduced a variation of Boyer's method, using a 5 percent price rise as the SSNIP to be examined.

Several shortcomings of this new method have been cited. Boyer's work recognized that problems similar to those involved with using cross elasticities still exist with the new approach. Moreover, Scherer and Ross argue that the use of the 5 percent rule in a situation in which market power already exists will tend to yield biased market definitions which could increase the likelihood of allowing the mergers that are being examined. Shepherd and Shepherd (2004) argue that estimates concerning future responses to a hypothetical monopolist's actions are speculative, and that what constitutes a sufficient reaction to a "small" and "significant" price rise is arbitrary. Finally, they maintain that both this measure and that of cross elasticity of supply improperly treat the question of BARRIERS TO ENTRY by combining it with the definition of market, thereby blurring the distinction between potential rivals and those actually in existence.

Practical Measures

Perhaps because of the weaknesses cited above, SSNIP has not yet been widely accepted by the courts. Instead, courts continue to use more pragmatic indicators. Given the complexity of markets, it remains unlikely that any one measure can accurately and completely determine the boundaries of markets. Perhaps realizing this, the US Supreme Court has offered its own guides. (For an expansive historical examination of the US courts' development of the concept of market definition, see Werden, 1992.) In an important antitrust case involving the Du Pont Company's monopolization of the cellophane market, the court established the standard of "reasonable interchangeability" of products as the measure for market delineation (see Breit and Elzinga, 1996: 196–9 for details). It cited a product's price, qualities, and uses as being significant in examining this demand-side substitutability. Shortly thereafter, in 1962, Chief Justice Warren, in a merger case involving the Brown Shoe Company, identified several additional market boundary determinants, including public recognition of a market, unique production facilities, differing prices among rivals, and even specialized vendors (Breit and Elzinga, 1996: 133–41). Moreover, the court here recognized that the Clayton Act allowed for examining submarkets for antitrust purposes.

Shepherd and Shepherd (2004) include, among other items, transportation costs relative to the product's value and shipping distances. These often limit competition to a certain geographic area. In many antitrust cases, the type of actual shipping data needed is readily available as evidence. Moreover, Breit and Elzinga's 20 percent rule gives a simple tool for discerning those who influence the price of a product and should be considered as market participants. Finally, numerous other forms of qualitative data, from surveys of consumer behavior to information regarding SEARCH COSTS, have been considered.

This list of market boundary variables is almost certainly overinclusive for any one case, but some items such as shipping information clearly warrant examination in many cases. Given the different techniques available for measuring market boundaries, and the different economic philosophies regarding government's role in regulating market activity (*see* CAPITALISM), debate over market definition methodology will likely continue.

See also *antitrust policy (US); geographic market; horizontal merger guidelines; merger guidelines, 1992–7*

Bibliography

Boyer, K. (1979). Industry boundaries. In T. Calvani and J. Siegfried (eds.), *Economic Analysis and Antitrust Law*. Boston: Little, Brown.

Breit, W. and Elzinga, K. (1996). *The Antitrust Casebook: Milestones in Economic Regulation*, 3rd edn. Fort Worth, TX: Dryden Press.

Church, J. and Ware, R. (2000). *Industrial Organization: A Strategic Approach*. Boston: McGraw-Hill Irwin.

Scherer, F. and Ross, D. (1990). *Industrial Market Structure and Economic Performance*, 3rd edn. Boston: Houghton-Mifflin.

Shepherd, W. and Shepherd, J. (2004). *The Economics of Industrial Organization*, 5th edn. Long Grove: Waveland Press.

Shepherd, G., Shepherd, H., and Shepherd, W. (2000). Sharper focus: Market shares in the merger guidelines. *Antitrust Bulletin*, **45**, 835–85.

US Department of Justice and Federal Trade Commission (April 2, 1992; revised April 8, 1997). *Horizontal Merger Guidelines*. Washington, DC: US Government Printing Office.

Waldman, D. and Jensen, E. (2001). *Industrial Organization: Theory and Practice*, 2nd edn. Boston: Addison-Wesley.

Werden, G. (1992). The history of antitrust market delineation. *Marquette Law Review*, **76**, 123–215.

market microstructure

Robert E. McAuliffe

This is an element of market transactions that economists tend to overlook: the actual processes and outcomes of exchange. How do specific trading mechanisms affect the distribution of prices? Market microstructure deals with this question by investigating how different trading mechanisms affect the prices observed when exchange occurs. Researchers investigate how price-setting rules evolve in markets and basically how prices are determined in markets.

Economists have not focused on the determination of the EQUILIBRIUM price because they argue that the market determines the price. But the question remains: how do markets determine prices? Would the price change if the pricing mechanism were changed? Market microstructure examines these issues, normally in the context of asset markets for stocks and bonds.

In his work on intermediaries, Daniel Spulber (1999) found that intermediary firms in markets lower TRANSACTIONS COSTS. He finds that intermediaries can improve the gains from trade because they offer price commitments to both sellers and buyers of the good. Thus intermediaries can evolve and operate in a market even when direct exchange possibilities between buyer and seller exist. He also found that as transactions costs fell to zero, the intermediary price approached the market price but intermediaries set prices otherwise. When uncertainties in exchange are allowed or inventories, the importance of intermediaries is increased as they provide market-making services to the economy.

Bibliography

Frankel, J. A., Galli, G., and Giovanni, A. (eds.) (1996). *Microstructure of Foreign Exchange Markets*. Chicago: University of Chicago Press.

Harris, L. (2002). *Trading and Exchanges: Market Microstructure for Practitioners*. Oxford: Oxford University Press.

O'Hara, M. (1997). *Market Microstructure Theory*. Oxford: Blackwell.

Spencer, P. D. (2000). *The Structure and Regulation of Financial Markets*. Oxford: Oxford University Press.

Spulber, D. F. (1999). *Market Microstructure: Intermediaries and the Theory of the Firm*. Cambridge: Cambridge University Press.

market performance

Gilbert Becker

Market performance measures the degree to which firms in a market achieve certain socially desirable goals, including EFFICIENCY and equity. Several types of efficiency have been identified. The first, **production efficiency**, occurs when suppliers minimize the AVERAGE TOTAL COST of producing a good, in both the SHORT RUN and the long run. By choosing the optimal combination of labor and other variable inputs (*see* ISOQUANT-ISOCOST CURVES), managers can minimize the short run per unit cost of production. By achieving MINIMUM EFFICIENT SCALE, long run costs of production can also be minimized. As a result,

production efficiency enhances firms' profitability. In addition, lower production costs and COMPETITION lead to lower prices, which benefit consumers. Moreover, this type of efficiency benefits society on the whole by minimizing the amount of its scarce resources needed to produce the product.

A second type of efficiency is **dynamic efficiency**, which focuses on technological progress. Here efficiency is achieved if market conditions induce a socially optimal amount of research and development, leading to a maximum level of innovation and invention. Schumpeter (1950) first proposed that large firm size and high industry concentration enhanced innovation. Many explanations have been advanced in favor of this hypothesis. These include the argument that ECONOMIES OF SCALE exist in research and development, and that there exists a disproportionate availability of funds for large firms in imperfect CAPITAL MARKETS. Also, the existence of economic profits in concentrated markets (*see* ECONOMIC PROFIT) has been argued as a factor enabling greater research. Other arguments conflict with Schumpeter, maintaining that the incentive to innovate is diminished in markets where competition is imperfect. Cohen and Levin (1989) survey this literature and find that although there is a wealth of empirical studies of these hypotheses, the results are inconclusive. They suggest that other variables affecting dynamic efficiency, such as the demand for innovation and the technological opportunity for innovation, need additional investigation.

A third type of efficiency is that of **allocative efficiency**. This is achieved when the amount of resources allocated to (and thus the number of goods produced in) each market maximizes society's welfare by reducing the DEADWEIGHT LOSS associated with MARKET POWER. This type of efficiency occurs in a market when every consumer who is willing and able to pay the MARGINAL COST of producing that good receives it. PERFECT COMPETITION is one MARKET STRUCTURE that reaches this result since the competitive market price is determined by the marginal cost of the last unit produced. MONOPOLY may also achieve this result if the firm is able to practice first-degree PRICE DISCRIMINATION.

The final performance criterion, **equity**, involves, in part, a question of the fairness of the levels of prices and profits established in markets. The topic is quite controversial. By one definition, zero long run economic profit constitutes a fair result, since the buyers pay only for the economic costs of production, and the sellers have all of their costs covered, including the owners' OPPORTUNITY COSTS. An alternative definition of fairness establishes that any level of economic profit is acceptable as long as there is a voluntary exchange between buyers and sellers in a market which is reasonably competitive. This divergence of opinion as to an appropriate definition of fairness has resulted in differences in state and federal government policies concerning antitrust enforcement (*see* ANTITRUST POLICY (US)). As one example, Folsom (1990) demonstrates that during the 1980s the National Association of Attorneys General, fearing that market power would increase the likelihood of unfair transfers of wealth, established more stringent standards for mergers and for vertical restraints of trade than those introduced by the federal government.

Measurement

Most empirical studies investigating issues concerning market performance have used one of two measures. Perhaps most prominent are studies using measures of firm and industry **economic profit**. If long run rates of return on assets or owner's equity are derived in part from the existence of market power, such profits may indicate a lack of market fairness. Moreover, these profits may be indicative of inefficiency resulting from this power if BARRIERS TO ENTRY or other factors foster an environment wherein maximum efficiency is not forced by competition.

Unfortunately, the use of profit-rate data suffers from a number of measurement problems. Fisher and McGowan (1983) point out several of these flaws. For example, the correct profit levels will not be measured if the value of the firm's capital, or the DEPRECIATION thereof, is measured using accounting as opposed to economic definitions (*see* ECONOMIC DEPRECIATION). Similarly, the value of ADVERTISING and research and development

cannot be correctly associated with only the year in which the expenditure takes place. As a result, the measure of ACCOUNTING PROFIT reported for that year may diverge from that of the true economic profit. In addition, profit rates need to be adjusted for differences in risk across industries if a correct evaluation of the extent of excessive profit is to be made. Finally, profit rates suffer conceptually as a measure of performance since firms with high profits may be more efficient and thus demonstrate good performance.

A second tool used to measure performance is the **price–cost margin**. One theoretically appealing example is the LERNER INDEX, which measures the difference between price and marginal cost as a percentage of price. While also indicating profitability, the extent of allocative inefficiency is directly measured in this index since the deadweight loss to society depends on the size of the gap between the marginal value that consumers give to the last unit consumed (price) and its marginal cost of production. Unfortunately, data on marginal cost are also difficult to obtain. As a result, AVERAGE VARIABLE COSTS are often used as a proxy. In addition, recent studies of market performance have attempted to estimate cost markups by examining pricing data in specific industries over time, or by introducing models of GAME THEORY. These methods suffer from the need to make numerous assumptions concerning demand and cost conditions, or about oligopolistic behavior. These and other issues continue to make the evaluation of market performance difficult.

See also *structure-conduct-performance paradigm; market power*

Bibliography

Carlton, D. W. and Perloff, J. M. (2000). *Modern Industrial Organization*, 3rd edn. Reading, MA: Addison-Wesley.

Cohen, W. and Levin, R. (1989). Empirical studies of innovation and market structure. In R. Schmalensee and R. D. Willig (eds.), *Handbook of Industrial Organization*. New York: North-Holland.

Fisher, F. M. and McGowan, J. J. (1983). On the misuse of accounting rates of return to infer monopoly profits. *American Economic Review*, **73**, 82–97.

Folsom, R. (1990). State antitrust remedies: Lessons from the laboratories. *Antitrust Bulletin*, **35**, 941–84.

Schumpeter, J. (1950). *Capitalism, Socialism, and Democracy*, 3rd edn. New York: Harper and Row.

market power

Gilbert Becker

Market power is generally defined as the ability of one or more firms to influence price. Greer (1992) broadens the definition somewhat to include "the ability to subdue rivals," including suppliers and potential entrants as well as customers. This power may arise from firms involved in either single or multiple markets. Market power stems from a variety of possible sources and has a number of potential consequences. It has been examined theoretically and estimated empirically by an ever-growing number of authors.

The principal source of market power is size – that of the single firm alone, or a group of firms – relative to the overall size of the market. This size, in theory, generates power through its effect on market supply. For the individual firm in the case of MONOPOLY, the firm's size is equal to that of the market. As such the firm enjoys the ability to influence the market price since it supplies the entire market and can choose the point along the market demand curve where it wishes to operate. In contrast, in PERFECT COMPETITION the individual firm's size is so small that its output has no noticeable impact on the market supply and it becomes a *price-taker*.

In antitrust cases concerning monopoly, once the market has been properly defined the next step traditionally has been to examine the market share of the alleged monopolist in order to estimate its market power. For example, Microsoft Corporation, in its recent antitrust trial, was found to control nearly 90 percent of its market, and thereby was seen as holding substantial market power (*see* MICROSOFT ANTITRUST CASE). Similarly, for groups of firms, as the collective market share of a small group rises, the potential for (and perhaps likelihood of) COLLUSION increases, again raising the ability to influence price (*see* OLIGOPOLY).

Current antitrust guidelines have broadened their focus beyond simple market share analysis since market share alone may not indicate sustainable market power. Recently, guidelines in the US and European Union (EU) have moved toward defining a market as the smallest area in which market power exists and a small but significant nontransitory increase in prices (SSNIP) can be sustained (*see* ANTITRUST POLICY (US); EU COMPETITION POLICY, 2004; MARKET DEFINITION).

The acquisition of market power can be achieved through any of several methods. Perhaps the most obvious approach is through the development of a product that is both unique in its characteristics and desired by consumers. Building a better mousetrap, or personal computer operating system for that matter, will ordinarily lead consumers to a firm's door, and increase its market share and thus its ability to influence price. A second means for expanding a firm's size and power is through a horizontal merger with a rival already within the market. Third, in order to sustain, if not create, market power, it is essential that BARRIERS TO ENTRY exist. These barriers, which are sometimes natural (i.e., extraordinary start-up costs), sometimes government created (i.e., occupational licensing requirements), and in other instances are the result of deliberate STRATEGIC BEHAVIOR, all tend to retard the growth in the number of firms in a market, and thereby protect the above-normal prices and profits of existing firms. Finally, while firm size relative to the market is the critical source of market power, a firm's absolute size may also play a role. Sheer size in terms of assets or number of employees may generate political power, which could in some instances help create or enhance market power. The evidence here is sporadic at best. Moreover, it is important for managers to be aware that for antitrust purposes, the courts have consistently held that large size alone is not onerous.

Several measures have been constructed that examine market power. One of the first theoretical models developed was the **Rothschild index**. This index is defined as the ratio of the slope of the individual firm's demand curve to that of the market demand. Following the analysis presented above, the ratio for a monopolist will have a value of 1 as the demand for its product is equal to the market demand. Overall the value of the index ranges between zero (for perfect competition) and 1, with higher values being associated with greater market power. The practical application of this index has been limited by the difficulty of estimating demand curves (*see* ESTIMATING DEMAND).

Three other measures have been more widely used in measuring market power. Each goes beyond the ability of (a) firm(s) to influence the market and instead focuses on the outcome of using that power which typically manifests itself in some measure of increased profit. The first of these three is the LERNER INDEX, measuring the extent of the divergence between price (P) and MARGINAL COST (MC). It is calculated as

$$L = \frac{(\mathrm{P} - \mathrm{MC})}{\mathrm{P}}$$

and, as was true of the Rothschild index, may range in value from zero for perfectly competitive firms to positive values with a limit approaching 1 in cases where price and marginal cost diverge. Once again, greater values for L indicate the presence of more market power. One of the difficulties in using this measure lies in the measurement of marginal cost, the data for which are not readily available. As a result, early studies used AVERAGE VARIABLE COSTS as an estimate for marginal cost. More recent studies estimate the value of MC econometrically (see Bresnahan, 1989; Church and Ware, 2000).

A third measure of market power, presented by Bain (1941), involves a direct measure of **excess profits**. Under perfect competition economic profits should equal zero in the long run (*see* ECONOMIC PROFIT), while the existence of market power in some imperfectly competitive markets should provide positive profits. Although ACCOUNTING PROFIT data are readily available and have been used in numerous studies following Bain's work, several difficulties arise in properly adjusting these data to reach a value for economic profits (*see* MARKET PERFORMANCE for details on measurement problems of this and other variables). In addition, positive economic profits may occur in the short run even in the absence of market power.

For example, exogenous shocks to market demand may temporarily increase price and result in short-term profit even in the case of perfect competition. Consequently, in order to correctly use this measure of market power, these profits should not be examined with a single static estimate. Instead, a long-term trend of sustained economic profits is necessary to establish the existence of market power.

A final measure of market power that has been used in some studies is **Tobin's q**. This measure, calculated as the ratio of the current market value of a firm's assets relative to the replacement cost of those assets, is an alternative method for measuring excess profits. Values of this ratio that exceed 1 indicate the existence of such profits. This measure has the advantage of not requiring an estimate of economic profits, as in Bain's measure, but does require the estimation of replacement cost, which in some instances may be quite difficult.

A great number of studies investigating the existence of market power have been undertaken over the past half-century. In a survey of this literature Bresnahan (1989) identifies two classes of studies. The early studies followed Bain (1951) and the structuralist school. Tests of this school of thought commonly involved CROSS-SECTION ANALYSIS, across several industries, of the effects of MARKET STRUCTURE on market performance. Here market power, arising from various structural factors, was expected to manifest itself in terms of firm and industry profitability, which was used as an indicator of performance. As noted above, the primary source of market power is firm size. As such, these tests examined the effects of single-firm market shares and/or industry concentration (sources of power) on the value of the Lerner index or profit rates (the outcome of the use of power).

These studies found a consistently positive but weak link between size (market power) and profitability. Their conclusion that size caused profitability was hampered by the argument that the efficiency of large size may also be driving this profitability. Because of this, Bresnahan argues, a new class of studies identifiable as industry-specific econometric case studies has arisen. In surveying the results of these studies he concludes that:

1 some concentrated industries possess significant amounts of market power; and
2 anticompetitive conduct is a significant cause of market power.

Numerous consequences may arise from the existence of market power. First, positive economic profits which result from the use of this power indicate a transfer of income from consumers to producers. Second, poor market performance, ranging from allocative inefficiency to a reduced rate of product innovation, are alleged to result from this power. Third, following Greer's definition, market structure can be changed if suppliers and rivals can be "subdued." The use of PRICE DISCRIMINATION, exclusive dealing, PREDATORY PRICING, and other forms of conduct often requires the existence of some power, but also may enable the further extension of that power across markets and over time.

See also *structure-conduct-performance paradigm*

Bibliography

Bain, J. S. (1941). The profit rate as a measure of monopoly power. *Quarterly Journal of Economics*, **55**, 271–93.
Bain, J. S. (1951). Relation of profit rate to industry concentration: American manufacturing, 1936–1940. *Quarterly Journal of Economics*, **65**, 293–324.
Bresnahan, T. (1989). Industries with market power. In R. Schmalensee and R. D. Willig (eds.), *Handbook of Industrial Organization*. New York: North-Holland.
Church, J. and Ware, R. (2000). *Industrial Organization: A Strategic Approach*. Boston: McGraw-Hill Irwin.
Greer, D. F. (1992). *Industrial Organization and Public Policy*, 3rd edn. New York: Macmillan.
Martin, S. (1994). *Industrial Economics*, 2nd edn. New York: Macmillan.
Maurice, S. and Thomas, C. (2002). *Managerial Economics*, 7th edn. Boston: McGraw-Hill Irwin.

market structure

Gilbert Becker

Market structure consists of those relatively fixed features of a firm's environment which identify the competitive nature of the industry. As such it is related to MARKET POWER and

MARKET PERFORMANCE and may be influenced by antitrust policy (*see* ANTITRUST POLICY (US)). Its first element is the number and size distribution of the sellers in the market. As the number of sellers increases, the market moves toward PERFECT COMPETITION and market power tends to diminish. Several measures of the size distribution of firms, including various CONCENTRATION INDICES and the HERFINDAHL-HIRSCHMAN INDEX, emphasize the collective market shares of an industry's leading firms. High values in these indices indicate a potential for market power. The market share of a single firm may also help to explain the market's structure and competitive conditions (*see* PRICE LEADERSHIP). The second element of market structure is the number and size distribution of buyers, which is important as it offers an indication of the extent of their countervailing power. A proper count of the number of sellers and buyers in any market requires careful identification of the degree of product substitutability between firms (*see* SUBSTITUTES) and the relevant GEOGRAPHIC MARKET in which the competition takes place (*see* MARKET DEFINITION).

A third vital element of market structure is the condition of ENTRY. High BARRIERS TO ENTRY are a central feature of MONOPOLY and are common to OLIGOPOLY. The condition of entry is important in understanding the competitive process in two ways. First, it helps to explain the number and size distribution of firms currently in the market. Second, it helps to evaluate the potential for new competitors.

Some researchers cite this element of structure as being uniquely important. Baumol, Panzar, and Willig (1982) maintain that in the absence of any barriers to entry or exit (*see* SUNK COSTS) markets become contestable and the number of rivals, their size, and other structural variables become irrelevant in determining the outcome of market performance (*see* CONTESTABLE MARKETS). Their research cites examples such as the airline industry where high resource mobility and low barriers to entry into new geographic markets assure market EFFICIENCY. Despite this, Shepherd (1984) and others have argued that the number of truly contestable markets is extremely limited, and that the level of COMPETITION already in existence in a market is of far greater importance than the degree of potential competition.

Two other elements of market structure are the degree of PRODUCT DIFFERENTIATION and the extent of VERTICAL INTEGRATION existing in the market. Both influence the nature of industry costs and the STRATEGIC BEHAVIOR among the rivals within the market. Both may also play a role in determining the condition of entry by increasing the costs and risk of entry.

The structure of a market is dependent on several factors, the two most basic of which are the underlying consumer demand and the industry's production cost conditions. Industry technology which offers substantial ECONOMIES OF SCALE relative to market demand may require large firm size and greater market concentration and thus limit the room for, and number of, existing rivals. In addition, these circumstances may limit the number of potential new entrants, which may face cost disadvantages stemming from an inability to produce at MINIMUM EFFICIENT SCALE.

Market structure may also be influenced by government policy ranging from patent laws and licensing requirements, which influence entry, to antitrust policy such as restrictions on mergers (*see* HORIZONTAL MERGER GUIDELINES), which may influence the number and size distribution of existing rivals. Finally, LIMIT PRICING and other forms of strategic behavior can be used by rival firms to alter the structure of a market.

The work by Porter (1980) on competitive strategy emphasizes the importance of market structure to successful business management. Five basic forces (including potential entrants, substitute goods industries, and the rivalry of materials suppliers) that exist in every market are identified. Porter demonstrates that from these forces evolve the competitive strategies that firms must adopt in order to be profitable. He argues that a sound evaluation of market structure and market definition is essential for managers in order to properly develop offensive and defensive strategies, assess the company's strengths and weaknesses, and examine its ability to cause changes in market structure.

Analysis of market structure is also important in the development of industrial policy. Shepherd (1982) uses an analysis of market structure to investigate the extent to which goods in the US economy were produced in competitive markets. He uses structural elements to create categories such as "effective competition" (wherein industries have low concentration ratios, unstable market shares, and low entry barriers) and "tight oligopoly" (where concentration ratios exceed 60 percent, and entry barriers are medium or high). He finds that more than 75 percent of the economy's national income in 1980 was generated in markets that were competitive. Moreover, this percentage had increased sharply (from somewhat more than 50 percent) since the 1950s. Shepherd cites active government antitrust policy as the primary explanatory variable, and imports as a secondary variable causing this shift. The policy implications here, and surrounding market structure in general, are controversial (*see* STRUCTURE–CONDUCT–PERFORMANCE PARADIGM).

Bibliography

Adams, W. and Brock, J. (2001). *The Structure of American Industry*, 10th edn. Upper Saddle River, NJ: Prentice-Hall.

Bain, J. S. (1956). *Barriers to New Competition*. Cambridge, MA: Harvard University Press.

Baumol, W. J., Panzar, J. C., and Willig, R. D. (1982). *Contestable Markets and the Theory of Industry Structure*. San Diego: Harcourt Brace Jovanovich.

Carlton, D. W. and Perloff, J. M. (2000). *Modern Industrial Organization*, 3rd edn. Reading, MA: Addison-Wesley.

Deutsch, L. (1998). *Industry Studies*, 2nd edn. Armonk, NY: M. E. Sharpe.

Porter, M. E. (1980). *Competitive Strategy: Techniques for Analyzing Industries and Competitors*. New York: Free Press.

Shepherd, W. (1982). Causes of increased competition in the U.S. economy, 1939–1980. *Review of Economics and Statistics*, **64**, 613–26.

Shepherd, W. G. (1984). "Contestability" vs. competition. *American Economic Review*, **74**, 572–87.

Shepherd, W. and Shepherd, J. (2004). *The Economics of Industrial Organization*, 5th edn. Long Grove: Waveland Press.

markup pricing

Kostas Axarloglou

This can be an optimal pricing policy where prices are set to cover all direct costs plus a percentage markup for profit contribution. It applies only for market structures where companies have some MARKET POWER, and therefore are able to set the prices of their products (*see* MARKET STRUCTURE). In other words, it is used by *price-setting* companies as opposed to *price-taking* ones in a perfectly competitive market (*see* PERFECT COMPETITION).

Markup pricing is one of several different pricing policies companies use to price their products. For that, they calculate the cost per unit of output, to which they add a profit margin, or a price markup, to determine the final price of their product. The cost per unit is the variable production and marketing cost per unit of output plus the average overhead cost; the final price is then determined by adding a certain percentage margin to the unit cost.

The price markup is usually expressed as the difference between the final price and the marginal cost, as a percentage of either the price or the marginal cost. Here, the MARGINAL COST represents the per unit cost of production. The former method gives the markup on price, while the latter gives the markup on cost. If P represents the final price of the product, and MC its marginal cost, then the markup on price equals to $\frac{P - \text{MC}}{P}$, and the markup on cost is $\frac{P - \text{MC}}{\text{MC}}$.

Although the markup pricing policy is relatively easy to apply, there are some problems associated with its use. First, the overhead cost per unit is calculated based on the expected overhead cost divided by the "normal" amount of the product the company expects to produce. However, in many cases, the actual production deviates from the expected volume of production, and this becomes a source for cost miscalculations. Also, in calculating the per unit cost of production, companies usually use actual expenses only, and they ignore the OPPORTUNITY COSTS of their resources. Consequently, they may underestimate their cost of production. Finally, since companies do not change their prices very frequently, they need to determine

the optimal price that will cover the period of time until the next price adjustment.

Economic theory suggests that the pricing rule, which allows companies to maximize their profits (*see* PROFIT MAXIMIZATION), is $P = \frac{1}{1+\frac{1}{e_p}}$ MC, where e_p is the price ELASTICITY of the firm's demand for its product (when there are no strategic interactions between firms; *see* GAME THEORY and OLIGOPOLY for a discussion of interaction between rival firms). Obviously, the profit-maximizing price depends on the marginal cost, and it is inversely related to e_p, the price elasticity of demand (notice that e_p is always a negative value). In other words, companies with very elastic market demand (large absolute value of e_p), and therefore a market demand very sensitive to changes in prices, will charge a price close to their marginal cost. Their price markup will be low, since high prices might result in a significant decline in their sales. In a perfectly competitive market, where e_p is a very large number (in extreme cases it approaches ∞), P is very close to MC, or the price markup is close to zero. From the optimal pricing equation above, the price to cost markup can be derived as $\frac{P - \text{MC}}{P} = -\frac{1}{e_p}$, which is called the LERNER INDEX, introduced by Abba Lerner in 1934. Obviously, the price markup is inversely related to the price elasticity of market demand. High e_p makes the company very aggressive in pricing its products and willing to accept a low price markup as a way to stay competitive and preserve its market share. On the other hand, low e_p (demand is relatively price inelastic) implies large profit-maximizing price markups. In such a case, the company has some market power, and for that reason it can charge a rather large price markup over its marginal cost. (For the factors that influence e_p, see the discussion in the entry on elasticity).

The above discussion suggests that the price markup is closely related to the degree of market power a company can exercise. More competitive market structures are usually characterized by lower price markups. Domowitz, Hubbard, and Petersen (1986), by using data from US industries between 1958 and 1981, found that those industries with high concentration (with a four-firm concentration ratio in the range of 80 to 100) show a markup on cost ratio of 0.32, while those industries with low concentration (with four-firm concentration ratio in the range of 0 to 20) show a lower markup on cost ratio of about 0.23, as expected. Also, Hall (1986), using data from 48 US industries in manufacturing for both durable and non-durable products, shows that in most of these industries price markups are significantly different from zero, or that in US manufacturing overall, companies seem to have significant market power.

Rotemberg and Saloner (1986) and Rotemberg and Woodford (1991, 1992) present evidence that price markups are correlated with the changes of economic activity during the different phases of the business cycle. They usually increase during economic recessions and decline during economic booms. Finally, Axarloglou (1994) finds some evidence of the same behavior in the publishing industry.

Bibliography

Axarloglou, K. (1994). Product variety and countercyclical price markups: An empirical assessment. Working paper, Fuqua School of Business, Duke University.

Domowitz, I., Hubbard, R. G., and Petersen, B. C. (1986). The intertemporal stability of the concentration–margins relationship. *Journal of Industrial Economics*, **17**, 1–17.

Hall, R. (1986). Market structure and macroeconomic fluctuations. *Brookings Papers on Economic Activity*, **2**, 285–322.

Greer, D. F. (1992). *Industrial Organization and Public Policy*, 3rd edn. New York: Macmillan.

Keat, P. G. and Young, P. K. Y. (2002). *Managerial Economics: Economic Tools for Today's Decision Makers*, 4th edn. Upper Saddle River, NJ: Prentice-Hall.

Lerner, A. (1934). The concept of monopoly and the measurement of monopoly power. *Review of Economic Studies*, **1**, 157–75.

Rotemberg, J. and Saloner, G. (1986). A supergame-theoretic model of price wars during booms. *American Economic Review*, **76**, 390–407.

Rotemberg, J. and Woodford, M. (1991). Markups and the business cycle. *NBER Macroeconomic Annual*, 63–129.

Rotemberg, J. and Woodford, M. (1992). Oligopolistic pricing and the effects of aggregate demand on economic activity. *Journal of Political Economy*, **100**, 115–207.

maximin criterion

Eduardo Ley

Suppose that, when a manager is faced with a decision, the payoffs from choosing any of the different strategies available will depend on circumstances beyond her control (e.g., stochastic events, other people's strategies, etc.), which we shall refer to as "states of the world." If we rank the strategies available by each strategy's payoff in the worst-case scenario, then the highest ranked is a maximin strategy. A maximin strategy has the property that the worst that can happen with it is no worse than the worst that could happen with any of the other strategies available. In other words, this is a conservative criterion for choosing a strategy which is the best in the worst possible scenario. As an example, suppose that there are two available decisions, d_j, and three possible states of the world, ω_i, with the probabilities and payoffs shown in the following table:

	d_1	d_2	Prob(ω_i)
ω_1	25	1	0.50
ω_2	10	15	0.35
ω_3	−5	5	0.15

The worst scenario for decision 1 has payoff −5, while the worst payoff with decision 2 is 1. Since $1 > -5$, decision 2 would be chosen by the maximin criterion despite the fact that decision 1 has a higher expected payoff (*see* EXPECTED VALUE). Note that the maximin criterion does not take into account the probabilities of the different states of the world. (In public economics the maximin criterion refers to the welfare criterion, which mandates maximizing the utility of the individual who is worst off.)

Bibliography

Raiffa, H. (1968). *Decision Analysis: Introductory Lectures on Choices under Uncertainty*. Reading, MA: Addison-Wesley.

merger guidelines, 1992–7

Gilbert Becker

These guidelines are the third set of guidelines offered to the US business community in the past quarter-century. They are an effort to clarify the circumstances under which the current administration will challenge a merger which it believes to be in violation of the CLAYTON ACT. The guidelines offer a five-step approach to the analysis of a merger:

1 the definition of the market and measure of industry concentration (*see* CONCENTRATION INDICES);
2 the evaluation of the potential adverse effects of the proposed merger;
3 the analysis of the condition of ENTRY into the market (*see* BARRIERS TO ENTRY);
4 the existence of potential efficiencies from the merger (*see* EFFICIENCY); and
5 the examination of the circumstances in the case where one of the merging firms is failing.

The geographic and product market boundaries are defined using the method known as the 5 percent rule, or SSNIP – a small but significant nontransitory increase in prices (*see* GEOGRAPHIC MARKET; MARKET DEFINITION). The market is then classified as being highly concentrated (if the post-merger HERFINDAHL-HIRSCHMAN INDEX, HHI, has a value greater than 1,800), moderately concentrated (if, after the merger, $1{,}000 < \text{HHI} < 1{,}800$), or unconcentrated (if the post-merger $\text{HHI} < 1{,}000$). While a challenge in the latter case is unlikely, in general the likelihood of a challenge increases, all else constant, for higher levels of concentration and for larger changes in the level of concentration. The numerical ranges used in the 1982 guidelines remain in force here, but new detail is presented (i.e., concerning market definitions where PRICE DISCRIMINATION exists), and a broader description of economic evidence being used by the antitrust agencies is offered.

The second step involves numerous tests which evaluate factors inherent in the industry that facilitate or limit the potential for anticompetitive effects to arise as a result of the proposed merger. Here the potential abuse of MARKET POWER is seen as having two possible sources:

1 coordinated efforts among the sellers in the market;
2 unilateral actions by the newly merged, and now larger, firm.

Evidence indicating a diminished capacity for the abuse of market power will tend to lessen the likelihood of a challenge. When the reverse is true the likelihood of a challenge is enhanced. For example, product standardization in an industry and free-flowing information about rivals' competitive actions are two factors cited as being conducive to COLLUSION. Similarly, the closeness of product attributes of the merging firms' products is cited as facilitating a successful unilateral post-merger increase in prices in some circumstances (*see* PRODUCT ATTRIBUTES MODEL).

The 1992–7 guidelines recognize that any adverse effects of a merger may be offset by the entry of new rivals. Thus the third step of the new guidelines examines the condition of entry into the market and now recognizes the particular importance of three factors:

1 the timeliness of the new entry;
2 the likelihood that it will occur; and
3 the sufficiency of that entry in controlling abuses of market power.

Fourth, the extent to which a proposed merger may offer a method for achieving ECONOMIES OF SCALE or other cost savings is considered as a possible defense in cases which might otherwise be challenged. The antitrust agencies will examine potential efficiencies stemming from the merger, and the new guidelines outline several possible sources of cost savings that may be considered. The 1997 revision of the guidelines is focused on the section concerning efficiencies and emphasizes that only *merger-specific efficiencies* will be considered. These are efficiencies which are likely to be achieved by the merger and unlikely to be achieved in its absence. If these efficiencies are *cognizable* – verifiable and not arising from anticompetitive reductions in output – then they will be considered. Moreover, the greater the potential adverse effect of the merger, i.e., as measured by increased concentration, the greater must be the efficiencies if the merger is to be deemed acceptable.

Finally, the guidelines recognize that the threat of increased market power is minimized in cases where one of the merging firms is in imminent danger of failure. Here the guidelines consider four situations wherein a challenge to the proposed merger is unlikely, including:

1 where the likelihood of the failing firm to successfully reorganize under current BANKRUPTCY laws is poor;
2 where there is an absence of alternative merging partners with whom there exists a decreased potential danger to COMPETITION arising from a merger.

These guidelines are largely an extension of those presented in 1982–4. It is generally accepted that these guidelines have brought more careful economic analysis into the process (see Scheffman, 1993, for a symposium of views). Both sets of guidelines are also widely accepted as being more lenient toward firms that wish to merge than those introduced in 1968. In an examination of recent antitrust activity, Lande (1994) finds that of 61 mergers challenged by the Federal Trade Commission between 1987 and 1992, only one involved a merger with a post-merger HHI value less than 2,000. In addition, only four challenges involved mergers where the change in the HHI was less than 400 points. He cites this and other evidence as indicative of the attenuation of aggressiveness in antitrust enforcement.

Heightened concerns by business over increasing foreign competition during the past decade may have fueled the discussion of a need for greater leniency in the guidelines. While the two most recent sets of guidelines specifically recognize the need to include foreign imports in the definition of the market, they also clearly recognize that trade restrictions may limit the ability of foreign firms to respond to domestic price increases.

See also *antitrust policy (US); EU merger guidelines, 2004; EU merger policy, 2004; geographic market; horizontal merger guidelines*

Bibliography

Lande, R. (1994). Beyond Chicago: Will activist antitrust arise again? *Antitrust Bulletin*, **39**, 1–25.

Ordover, J. and Willig, R. (1993). Economics and the 1992 merger guidelines: A brief survey. *Review of Industrial Organization*, **8**, 139–50.

Scheffman, D. (1993). Symposium on new 1992 merger guidelines. *Antitrust Bulletin*, **38** (3).

Shepherd, G., Shepherd, H., and Shepherd, W. (2000). Sharper focus: Market shares in the merger guidelines. *Antitrust Bulletin*, **45**, 835–85.

US Department of Justice and Federal Trade Commission (April 2, 1992; revised April 8, 1997). *Horizontal Merger Guidelines*. Washington, DC: US Government Printing Office.

US Federal Trade Commission Bureau of Competition (1992). *How 1992 Guidelines Differ from Prior Agency Standards*. Washington, DC: US Federal Trade Commission.

Microsoft antitrust case

Gilbert Becker

One of the champions of the computer software revolution over the past two decades, the Microsoft Corporation has also been the center of controversy regarding many of its business strategies. Since the mid-1990s the company has been the subject of a number of antitrust actions, brought by the federal and state governments in the US, by groups of consumers, by rival firms, and by the European Union (EU). At issue are a number of alleged anticompetitive business practices, the injury which these practices may have brought on the market and its various parties, and appropriate public policy toward these acts. Given the magnitude of the impact the new technology has had on numerous markets and the complexity of the issues in question, these cases are seen as an important opportunity to reexamine major antitrust statutes such as the SHERMAN ACT, perhaps in light of new economic thinking about new forms of COMPETITION evolving in the twenty-first century.

Background and Current Status of the Case

Microsoft (MS) is perhaps best known for its development and sale of the product called Windows, which is the operating system currently running on approximately 90 percent of all personal computers in the US. The company has also developed numerous other computer software applications programs (e.g., Microsoft Word for word processing, Excel for spreadsheet analysis, etc.) which run on the Windows system, some of which are at the center of the debate. Other programs, such as Internet Explorer for accessing the Internet, have been written into or have become part of upgraded versions of the Windows operating system.

In 1995, as the result of an investigation by the US Department of Justice (DOJ), Microsoft entered into a consent decree in which it agreed not to tie the sale of other products to that of Windows. In 1997, the DOJ filed suit against Microsoft, claiming that the company had violated the decree by BUNDLING its web browser, Internet Explorer (IE), with Windows. In June of 1998, shortly before a US Court of Appeals ruled in favor of Microsoft, the DOJ in conjunction with 20 US states filed a broadened antitrust suit against the company, claiming illegal monopolistic practices.

That case was initially decided in US District Court in April of 2000 (see *US* v. *Microsoft Corp.*, 1999, 2000). Judge Jackson's three central findings were that Microsoft had violated Sections 1 and 2 of the Sherman Act by:

1 illegally monopolizing the operating system market;
2 illegally attempting to monopolize the web browser market; and
3 illegally tying Windows and IE.

Subsequently, a US Circuit Court of Appeals overturned the last two findings of Judge Jackson (see *US* v. *Microsoft Corp.*, 2001).

Market Definition and Market Power

Numerous economic and legal issues have arisen over which the two parties (Microsoft and the government) and the two courts (District and Circuit) have often disagreed. The first issue to

arise is the proper definition of the market (*see* MARKET DEFINITION). This issue is critical in that illegal monopolization and illegal tying require the demonstration of the existence of MARKET POWER, and the abuse of that power in the relevant market(s). Thus, proper market boundaries must be established before any further economic analysis can take place.

Franklin Fisher, the principal economic witness for the DOJ during the District Court trial, used the new theoretical approach to market definition provided in the government's HORIZONTAL MERGER GUIDELINES. (Much of what follows is based on the testimony of Fisher; see Fisher and Rubinfeld, 2001, for a detailed economic analysis.) He argued that goods which serve to constrain a hypothetical monopolist's market power (*see* SUBSTITUTES) should be included in a proper definition of the market. The DOJ held that two separate economic markets exist: (1) Intel-based PC operating systems and (2) web browsers. In the *Intel-based PC operating system* (OS) market, Windows held approximately 90 percent market share, well above the Supreme Court's previously established standard for demonstrating the existence of MONOPOLY. Here, hand-held substitutes (e.g., Palm Pilots) were seen as outside the market, due to their low CROSS ELASTICITIES, and non-Intel PCs (e.g., Apple) were considered to be a limited damper on Microsoft's power. Moreover, some original equipment manufacturers (OEMs) of PCs testified to the lack of alternatives to purchasing Windows for their machines. Fisher also argued that browsers were a complementary good (*see* COMPLEMENTS) to operating systems, since they initially held a separate price and market of their own.

As for Microsoft's market power, the government's case relies heavily on the theory of NETWORK EXTERNALITIES and the existence of an applications programming barrier to entry (*see* BARRIERS TO ENTRY). Independent programmers have written tens of thousands of software programs that will run on the Windows operating system by relying on Microsoft-written applications programming interfaces. These are sets of commands by which applications connect with the personal computer. Fisher argues that the operating system market is a "winner-take-most" market, since consumers will gravitate to that system with the most programs, given that the more consumers using a program (e.g., Word, Instant Messenger), the more valuable it becomes to other potential users (to "join the network"). In addition, programmers will gravitate to writing almost entirely for that operating system with the most customers, since that is likely to be their most profitable strategy. Reinforcing this focus by programmers are two additional factors:

1 most programming costs are FIXED COSTS leading to dramatic programming ECONOMIES OF SCALE; and
2 costs for rewriting the program to run on another system are often high.

Thus, unless a new operating system can attract a huge volume of applications, a network effects applications barrier yields market power for the dominant operating system.

Microsoft's position, as presented by its principal economics witness Richard Schmalensee, was that the proper market to be examined was that of *software platforms*, which are sets of code upon which applications can be used. (See Evans, Nichols, and Schmalensee, 2001, for a detailed description of Microsoft's economic defense in the case.) While platforms also result in winner-take-all network effects, Schmalensee argued that Microsoft's Windows market position was tenuous, since platforms can be developed from other sources, such as Sun Microsystem's Java or by using a browser (e.g., Netscape's Navigator) to create a web-based platform. Both of these so-called middleware applications threatened the Windows monopoly, albeit in different ways. With respect to Netscape, programmers could write software that would run on its Navigator, which runs on other operating systems and threatened to become a platform substitute replacing Windows. Java, on the other hand, was an open-architecture programming language that was designed to run on any operating system. This meant that (potentially many) new operating systems, each running programs written in the Java language, could be developed as substitutes for Windows and thereby threaten to "commoditize" the once essential product.

Moreover, one "killer" application for any platform could lead to widespread demand by consumers, and a flood of new applications being written for that platform. Thus, while Microsoft was not competing head-to-head with these firms in the Intel-based operating system market, Schmalensee argued that Microsoft's market power is more limited when measured in a framework of dynamic competition in the platform market. In this setting, innovation by "new and unforeseen" competitors poses a continuing threat to the existence of even a firm with a dominant market share. Using the DOJ's theory of a barrier-protected monopolist, he demonstrated Microsoft's limited power by estimating Microsoft's correct profit-maximizing monopoly price to be hundreds of dollars greater than the price that was actually charged for Windows (*see* PROFIT MAXIMIZATION). Finally, Microsoft's position with respect to its IE was that the browser is an important innovation which is a natural addition to an operating system since it allows consumers access to web-based information and can be used by Windows to provide upgrades, and therefore is not a separate product.

In addition, Schmalensee argued that, for several reasons, there does not exist an applications barrier to entry in the market. First, widespread research and development and investment in intellectual property exist across the industry. Second, thousands of software writers exist in the market and barriers to entry in software programming are low. Third, many writers do offer programs for less-used platforms, especially where buyers (e.g., businesses) may be willing to pay more. Finally, since applications are a desirable consumer product, identifying the existence of many applications as a barrier to entry is comparable to identifying improvement of product quality as a barrier.

Anticompetitive Behavior

The government argued that Microsoft was involved in numerous predatory and exclusionary acts designed to maintain its operating system monopoly by protecting its applications barrier to entry. First, Fisher argued that Microsoft tied its browser to the sale of Windows in order to monopolize the browser market. This was not done to ultimately charge a monopoly price for IE. Instead Microsoft's intent, according to Fisher, was to protect its operating system monopoly, since it feared that Netscape's Navigator, a complementary good to the operating system, could develop into a platform for users and programmers which could ultimately run on any operating system.

In addition, the government alleged that Microsoft attempted to foreclose competition in the browser market in several ways, including:

1 by attempting to illegally divide the browser market by soliciting Netscape to not produce a browser that would run on Windows 95;
2 by annual expenditures of $100 million for development of IE, combined with giving the browser away for free, demonstrating obvious predatory behavior;
3 by restricting OEMs from removing IE or its desktop icon, which, combined with PC manufacturers' desire to load only one browser, excluded Netscape Navigator from most new shipments;
4 by using its Windows' market power to induce Internet Service Providers (ISPs) to limit their marketing and distribution of non-Microsoft browsers; and
5 through agreements with other software developers desiring Windows support (e.g., Intuit) to exclusively deal with Microsoft.

Fisher argued that each action, taken on its own merit, was contrary to Microsoft's short-term profit and could thus only be explained as anticompetitive efforts to sustain its monopoly.

The government also identified efforts by which Microsoft attempted to directly protect its operating system monopoly. These included:

1 making threats against Intel Corp. and Apple Computer Co. to withhold Microsoft's support of their products. These threats forced each firm to agree to end work on new technology (e.g., native signal processing, audiovisual streaming) which potentially threatened the Windows monopoly, and to limit support of Netscape Navigator and Sun Microsystem's Java.
2 Microsoft licensed Java from Sun, then created a "polluted Java" of their own which would cause some independent

software developers' programs to run properly on Windows, but not on non-Windows operating systems. This threatened to eliminate a main selling point of Java, which was that it was designed to run on any operating system.

The government argued that these acts harmed markets in several ways. First, consumers were harmed by the suppression of new technology and innovations not reaching the market. Moreover, even if Microsoft later chose to incorporate a version of the new technology in upgrades of Windows, the decision as to which innovations should be allowed access into the market should not be determined solely by Microsoft's interests. Second, consumers and PC manufacturers were unable to choose which browser would be preinstalled, nor were they able to purchase a browserless PC. Third, consumers faced significant time costs and other TRANSACTIONS COSTS if they wished to download Netscape Navigator. Fourth, browser market rivals of Microsoft were being foreclosed from a fair market test of their products based on the quality and features offered to consumers. Finally, and perhaps most importantly, Judge Jackson, in his Findings of Fact, cited harm from Microsoft's signal that it "will use its market power and immense profits to harm any firm that pursues initiatives that could intensify competition against one of Microsoft's core products" (*US* v. *Microsoft Corp*., 1999: 412). In sum, the DOJ's position was that whatever benefits were brought to market by Microsoft's development of IE, they did not exonerate Microsoft from its illegal acts designed to maintain its monopoly. Antitrust action is appropriate as it is designed to protect consumers by *protecting competition*.

Microsoft's position with respect to charges of predation and exclusion hinges on the fact that the platform market is a winner-take-all environment. In this setting Schmalensee argues that each rival may set its price below variable cost, violating the standard test for predation (*see* PREDATORY PRICING), expecting to win even though only one firm will succeed. He further argues that while IE's pricing brought no revenue from IE itself, the strategy was essential for the survival of the Windows platform and of the firm. Microsoft's expenditures of hundreds of millions of dollars on IE, and its subsequent zero pricing for its product, along with its other exclusionary strategies, were acts of competition. As such these acts should not be considered illegal. The inevitable victory of only one firm means that the traditional measure for predation is inappropriate, in large part since it condemns the winner, or as Schmalensee stated: "success, monopolization and exclusion are one and the same" (Evans et al., 2001: 212). As an alternative he offers that a new antitrust standard for predation in winner-take-all markets be the direct examination of the benefits versus costs to consumers of the business actions taken. Here Schmalensee argues that Microsoft's considerable expenditure on IE clearly resulted in a desirable product benefiting consumers, as did the innovation race itself.

In addition, Microsoft argued that many of its actions did not significantly impact the market in question. Its CONTRACTS did not preclude PC manufacturers from installing more than one browser, and Netscape and other rivals offered free downloads over the worldwide web. Moreover, Netscape was able to reach over 100 million potential customers with a CD mailing which only marginally raised its overall costs. Also, other business pressures applied by Microsoft to members of the industry (e.g., ISPs, Apple) did not succeed. In sum, Microsoft held that competition in the market continued to be healthy. Moreover, it maintained that the antitrust case was inappropriate in that it would stifle future innovation by Microsoft and only served to *protect competitors* who were unsuccessful in the market.

The Courts concluded that Microsoft did indeed violate the antitrust laws. The final remedies in this case have yet to be determined and remedies in other antitrust lawsuits against Microsoft are pending. Their implementation may prove to be a force in shaping the industry, and also future public policy. Any legislative changes (or judicial reinterpretation of the current laws) which arise will undoubtedly continue to be debated.

See also *antitrust policy (US); capitalism; EU competition policy, 2004; Microsoft antitrust case: remedies*

Bibliography

Evans, D., Nichols, A., and Schmalensee, R. (2001). An analysis of the government's case in *US* v. *Microsoft*. In W. Commanor (ed.), Symposium: Economists examine Microsoft – its concepts, findings, and remedies. *Antitrust Bulletin*, **46**, 163–252.

Fisher, F. and Rubinfeld, D. (2001). *US* v. *Microsoft*: An economic analysis. In W. Commanor (ed.), Symposium: Economists examine Microsoft – its concepts, findings, and remedies. *Antitrust Bulletin*, **46**, 1–72.

US Circuit Court, District of Columbia (2001). *US* v. *Microsoft Corp*., 253 F.3d 34 (D.C.Cir. 2001). Washington, DC: US Government Printing Office.

US District Court, District of Columbia (1999). *US* v. *Microsoft Corp*., 98 CIV.1232 Court's Findings of Fact. Washington, DC: US Government Printing Office.

US District Court, District of Columbia (2000). *US* v. *Microsoft Corp*., 98 CIV.1232 Court's Conclusions of Law. Washington, DC: US Government Printing Office.

Waldman, D. and Jensen, E. (2001). *Industrial Organization: Theory and Practice*, 2nd edn. Boston: Addison-Wesley.

Microsoft antitrust case: remedies

Gilbert Becker

In the monopoly case against Microsoft Corporation (*see* MICROSOFT ANTITRUST CASE), Judge Jackson's decision concerning remedies for the illegal MONOPOLY called for both structural and behavioral relief (see *US* v. *Microsoft Corp*., 2000). It first called for splitting Microsoft into two firms, with one firm responsible for the Windows operating system and the other owning the Microsoft applications software (i.e., Microsoft Office, Internet Explorer, etc.). The rationale behind this split was the expectation that the first firm would continue to innovate in the operating system market, while the second firm would have incentive to make the applications functional for other operating systems, thus allowing new competitors to enter and compete in the operating system market where Microsoft currently holds a monopoly position. His decision also restricted Microsoft's conduct in several ways, including placing restrictions on Microsoft from making exclusive deals with Internet Service Providers (ISPs), and ending its anticompetitive restrictions on PC manufacturers.

On appeal, the US Circuit Court concurred with the US Department of Justice's (DOJ) MARKET DEFINITION and affirmed the lower court's findings of illegal exclusionary and predatory monopolistic practices (see *US* v. *Microsoft Corp*., 2001). It overturned the proposed structural remedy, in part due to the severity of the remedy, the uncertainty of its success, and the brevity of the hearings held on the matter. It remanded the case to the lower court for more complete hearings into an appropriate remedy.

Past antitrust remedies have been designed to end anticompetitive activity, promote a more competitive environment, and punish antitrust offenders. The existence of NETWORK EXTERNALITIES in industries such as computer software may call into question the validity of applying traditional remedies because of possible conflicts arising in achieving these goals. Debate surrounding the question of structural relief in the Microsoft case centers in large part on the economics of these network effects in the industry.

Critics of the government's position on breaking up Microsoft argue that these effects make the District Court's (and other) structural remedies harmful to consumers, software developers, and the economy as a whole. The core of the argument is that increasing the number of competitors in the operating system market will result in a loss of some network benefits to consumers. This happens in part because new software applications ordinarily will not run on all operating systems, and as a result individuals will not be as able to share work or otherwise communicate with one another. Moreover, competition among firms supplying operating systems may take the form of PRODUCT DIFFERENTIATION and lead to increasing incompatibilities across systems. Liebowitz (1999) estimates that software producers would face $30 billion in additional costs for development and marketing if they chose to adapt their existing software to conform to as few as two new operating systems. He argues that these costs would lead to higher software prices for consumers, or a reduced number of software firms as some may not be able to survive the

competitive problems of higher costs and a more fragmented set of consumers.

Microsoft proponents also argue that the firm's monopolizing conduct itself should not be subject to antitrust punishment because the natural result of rivalry in a market having network effects is a winner-take-all outcome, i.e., monopoly. The traditional standards for SHERMAN ACT violations, which make exclusionary and predatory behavior illegal, are seen as protecting only unsuccessful competitors, and not protecting COMPETITION or consumers. Thus such standards are inappropriate in markets having significant network effects (see Rule, 1999).

A central question surrounding the validity of these arguments is whether achieving network effects inevitably requires dominance by a single firm. Proponents of structural relief argue that there is no compelling theoretical proof of this. Commanor (2001) points out that after the breakup of AT&T in the 1980s telephone service compatibility was maintained among the rivals in this market, which also yielded substantial network effects. In addition he argues that while some economic models of OLIGOPOLY can support the claim that product differentiation may create incompatibilities in a restructured market, other models predict convergence toward product standardization. Levinson, Romaine, and Salop (2001) expand this argument, showing that firms producing rival operating systems would each have strong incentives to maintain compatibility across their basic systems, to maximize network benefits for their customers, and thereby maximize their own sales and profits. While this core compatibility is being maintained, each firm simultaneously has incentives to add new features to its own operating system. Moreover, exclusionary and predatory behavior harms consumers and the competitive process if it unnecessarily results in one firm deciding which operating system innovations will be offered as choices to consumers.

In November 2001, the DOJ and nine states reached an out of court agreement as to remedies (see *US* v. *Microsoft Corp.*, 2001). These remedies abandoned structural relief and focused entirely on behavioral solutions. The District Court, as required by law, examined the proposed settlement and simultaneously held hearings concerning an alternative proposal offered by nine states which had refused to sign the agreement drafted by the DOJ. Debate between the DOJ and these states concerned the types and strength of conduct remedies to be imposed. The DOJ, now under the Bush administration, had required what some considered to be only mild concessions from Microsoft, including:

1 allowing greater freedom for PC manufacturers to install non-Microsoft software and remove Microsoft desktop icons, and prohibiting retaliation against those who do;
2 prohibiting Microsoft from entering CONTRACTS which require exclusive support of its products;
3 requiring increased disclosure of technical data aiding programmers to design programs that will fully function on Windows.

The remaining states voiced concerns that the new administration's settlement was too lenient, difficult to enforce, and carried too many exemptions for it to effectively constrain Microsoft's monopoly power. Moreover, they cited examples of Microsoft continuing some of its recently proscribed business practices in product lines that have emerged since the conclusion of the initial trial. These states requested numerous additional remedies, including:

1 requiring Microsoft to sell an unbundled version of Windows (*see* BUNDLING), leaving a more openly competitive environment for rivals selling various software applications such as browsers, media software, and the like;
2 requiring Microsoft to license the source code of Internet Explorer (IE) in order to promote its use by rivals as a platform for their products;
3 forcing Microsoft to write its Office software to run on competing operating systems such as Linux, while restoring its support of the Java programming language; and
4 appointing a special master to monitor Microsoft's business behavior.

Finally, they asked for the courts' findings and remedies to apply to new and future markets in which Microsoft is involved.

Microsoft remained opposed to each of these changes, arguing that the unbundling of Windows would be difficult at best, inefficient, and harmful to consumers. Moreover, it would restrict the company's right to innovate. It further argued that the states' proposals constituted a confiscation of intellectual property and would force Microsoft out of the market.

In November of 2002, the District Court accepted the DOJ settlement proposal as being in the public interest, and largely rejected the remaining states' case. Currently, one state, Massachusetts, remains actively opposed to the settlement agreement, and has filed an appeal with the Circuit Court. The full Circuit Court of Appeals was due to meet in late 2003 to hear this appeal and a separate appeal filed by a consortium of Microsoft's rivals in the markets for server software and other new products. This group has convinced the court to examine the question of whether the District Court erred by not considering alleged antitrust violations by Microsoft in these new markets when handing down its decision regarding the settlement agreement concerning earlier antitrust violations.

Antitrust law also allows for monetary damages to be collected by individuals or firms who initiate private lawsuits. The court's conclusion in *US* v. *Microsoft Corp*. (2001), that Microsoft was guilty of illegal activity, dramatically simplifies plaintiffs' cases for damages. Injured parties need only to demonstrate that the injury occurred as a result of Microsoft's actions and convince the court of a specific dollar amount of damages. The court may then impose treble damage awards.

Consequently, Microsoft faces numerous damage suits. One important case is a set of over 100 class-action lawsuits brought on behalf of consumers who allegedly paid excessive amounts for Windows. An initial out-of-court settlement of this case, whereby Microsoft offered several hundred million dollars in cash, free software, computers, and training to public schools, was rejected by the courts as potentially improving Microsoft's power in the education market (see Buckman and Kulish, 2002). In two additional cases, America on Line (AOL), which now owns Netscape, and Sun Microsystems, which owns Java, both of whom were cited by the Circuit Court as having been injured, filed lawsuits to collect for damages. In July of 2003, AOL and Microsoft reached a settlement in which Microsoft would pay AOL $750 million in damages.

Finally, the European Union (EU) is also investigating Microsoft for violations of its antitrust laws concerning "abuse of a dominant position" (see EU COMPETITION POLICY, 2004). It is also examining Microsoft's efforts to use its Windows monopoly to dominate the market for corporate server software, and its bundling of Windows with its media software. The European Commission, which in recent years is widely considered to have taken a more aggressive antitrust stance than its US counterparts, is considering requiring that Microsoft provide an unbundled version of Windows. It is armed with considerable investigative powers, and also holds, as perhaps its most powerful remedial weapon, the ability to impose a penalty of up to 10 percent of a firm's annual worldwide revenues for violations of the law. In addition it maintains other antitrust remedies similar to those used in the US such as cease and desist orders. The final outcomes of the EU investigation and the various cases in the US concerning Microsoft may have a significant impact on antitrust policy in the US and worldwide in the foreseeable future.

See also *antitrust policy (US); antitrust remedies (US); Microsoft antitrust case; Sherman Act*

Bibliography

Buckman, R. and Kulish, N. (2002). Judge tosses Microsoft consumer class-action deal. *Wall Street Journal*, **244**, A3.

Commanor, W. (2001). The problem of remedy in monopolization cases: The Microsoft case as an example. *Antitrust Bulletin*, **46**, 115–33.

Levinson, R., Romaine, R., and Salop, S. (2001). The flawed fragmentation critique of structural remedies in the Microsoft case. *Antitrust Bulletin*, **46**, 135–62.

Liebowitz, S. (April 30, 1999). Breaking Windows: Estimating the cost of breaking up Microsoft Windows. Working paper, Association for Competitive Technology and the ASCII Group Inc.

Rule, C. (1999). The proper standard for identifying monopolistic practices: Why Microsoft hasn't violated the antitrust laws. *Connecticut Law Review*, **31**, 1387.

US Circuit Court, District of Columbia (2001). *US* v. *Microsoft Corp*., 253 F.3d 34 (D.C.Cir. 2001). Washington, DC: US Government Printing Office.

US District Court, District of Columbia (2000). *US* v. *Microsoft Corp*., Civil Action No. 98-1233 Court's Final Judgement. Washington, DC: US Government Printing Office.

US District Court, District of Columbia (2001). *US* v. *Microsoft Corp*., Civil Action No. 98-1232 (CKK) Proposed Final Judgement. Washington, DC: US Government Printing Office.

minimum efficient scale

Robert E. McAuliffe

The minimum efficient scale (MES) represents the smallest output level for a firm at which long run average costs are at a minimum (*see* LONG RUN COST CURVES). If the long run average cost curve were U-shaped and continuous, then the MES firm size would be unique. However, statistical estimates of cost curves for various industries suggest that long run average cost curves are L-shaped where there are significant ECONOMIES OF SCALE at low levels of output which are exhausted relatively quickly, then average costs remain constant (see Johnston, 1960; Scherer, Beckenstein, and Kaufer, 1975). This means that the MES represents a lower bound on firm size but not necessarily an upper bound.

Estimates of the MES have also been obtained using engineering surveys and survivor studies. With engineering surveys, industrial engineers and other experts provide information concerning the expected changes in costs as the scale of operations increases, and from this the MES is determined. Recognizing the difficulties in estimating the MES, Stigler (1958) suggested that those firms which survived in the competitive environment should be the most efficient (*see* EFFICIENCY). Survivor studies examine the changes in the number of firms in different size classes over time to determine the optimum size plant (or, as Stigler suggested, optimum range of sizes). In the SHORT RUN firms may not be operating at the optimal scale, so reliable survivor estimates must be based on industries in long run EQUILIBRIUM. Unlike engineering studies these studies use data from operating firms, but those firm sizes that survive could have done so through anticompetitive behavior or because of BARRIERS TO ENTRY, which would not reflect efficiency. Nevertheless, survivor studies tend to confirm the results obtained from the engineering and statistical cost studies: there appears to be a wide range of optimum plant sizes, suggesting a range of output levels where long run average costs are constant.

Knowledge of the plant-level MES in an industry is important to understand the feasible number of firms that could operate in the industry. Scherer et al. (1975) concluded from their estimates of MES that actual concentration ratios in US industries are much higher than required by the estimated minimum efficient scale. If correct, these estimates indicate that antitrust policies which would break up large firms might not cause inefficiency. However, ECONOMIES OF SCOPE are also important in industry and estimates of economies of scale and the MES may fail to detect these additional causes of larger firm size, so any policy actions must be carefully considered. See Gold (1981) for a critical survey of the theoretical issues regarding firm size.

The behavior of costs for plants operating at less than MES is also important. If costs increase significantly when plants are smaller than the MES, then the disadvantages of small size are much greater and this could deter ENTRY into the industry.

Bibliography

Carlton, D. W. and Perloff, J. M. (2000). *Modern Industrial Organization*, 3rd edn. Reading, MA: Addison-Wesley.

Gold, B. (1981). Changing perspectives on size, scale, and returns: An interpretive survey. *Journal of Economic Literature*, **19**, 5–33.

Johnston, J. (1960). *Statistical Cost Analysis*. New York: McGraw-Hill.

Scherer, F. M., Beckenstein, A., and Kaufer, E. (1975). *The Economics of Multi-Plant Operation: An International Comparisons Study*. Cambridge, MA: Harvard University Press.

Stigler, G. J. (1958). The economies of scale. *Journal of Law and Economics*, **1**, 54–71.

monopolistic competition

Kostas Axarloglu

Monopolistic competition is a MARKET STRUCTURE characterized by a large number of firms where every firm has some MARKET POWER with respect to its products. The concept was introduced by Edward Chamberlin (1933) to study deviations in terms of prices and the number of firms in the market from the perfectly competitive market structure (*see* PERFECT COMPETITION).

The basic market characteristics of a monopolistically competitive industry are the following:

- there is a large number of firms, and every firm has a small market share;
- every firm sells a differentiated product (*see* PRODUCT DIFFERENTIATION) from its competitors, and therefore it has some market power (downward-sloping market demand schedule);
- there is free ENTRY and EXIT of companies into and from the industry in the long run, which usually leads to lower market share and zero ECONOMIC PROFIT;
- because of the large number of firms in the industry, each firm does not consider the reactions of its rivals in its own decisions (*see* OLIGOPOLY).

The market structure of monopolistic competition is suitable to study several different questions which cannot be easily addressed in the context of other market structures. First, since every firm in the industry sells a different variety, the level of PRODUCT VARIETY in the market can be assessed along with its social welfare implications (*see* OPTIMAL VARIETY). Second, this market structure allows economists to explore the reasoning behind companies' decisions on the type, design, and selection of the varieties (brands) they offer in the market. Finally, the implications of brand selection on companies' pricing policies can also be assessed.

In analyzing the monopolistically competitive market structure, two major families of models have been developed. **Chamberlinian models**, sometimes also called the *representative consumer* models, and **Hotelling-type models**, also known as *location* or *address* models.

Chamberlinian Models

Introduced by Chamberlin (1933) and extended by Dixit and Stiglitz (1977), the Chamberlinian model is a representative consumer model with a large number of firms, each one of which offers a distinct product variety in the market. In the model, consumers have an insatiable desire for product variety and perceive every brand as an equally good substitute for every other brand available in the market. The ELASTICITY of substitution among different brands is exogenously given and constant, and is not related to market entry or exit. In other words, the representative consumer does not have an "ideal" variety, i.e., a variety he prefers the most. Firms produce one variety each and compete only in terms of prices, so they do not compete in the design and the particular characteristics of their products.

The number of different varieties in the market is equal to the number of firms in the industry. Firms will enter the market until all profitable opportunities are exhausted (zero economic profits). Overall, the model predicts that monopolistically competitive industries will have a large number of varieties when there are low FIXED COSTS in production, and/or a low elasticity of substitution among different varieties, and/or a large market demand.

However, the free entry of firms has some significant welfare implications on the economy (see the entry on optimal variety). Overall, these models predict that the product variety provided in the market can either be greater or less than optimal because of two contradicting forces. On the one hand, a firm will enter the market with a new brand if it believes entry will be profitable, but the new brand may take business from other firms already in the market. Since the entrant does not worry about the negative effect it will have on established firms, there tends to be too much entry and too many varieties. Offsetting this force is the fact that society may benefit more from a new brand than it costs to produce, but if the firm cannot practice PRICE DISCRIMINATION, it cannot capture all the benefits to society from introducing a new

brand and this force tends to reduce the number of brands.

The main criticism against these models, aside from the fact that they are based on a representative consumer, is the fact that they cannot explain how companies choose the attributes of their brands. There is no basis for a theory of *product choice* and *product design*. Also, the assumption that the representative consumer does not have an "ideal" variety is rather simplistic and not very appealing. Finally, the fact that the degree of substitution among different varieties is exogenously given, and is independent of the number of varieties in the market, does not allow economists to study the implications of the level of product variety on companies' pricing policies.

Recently, there have been some efforts to improve the basic Chamberlinian model, while preserving some of its fundamental characteristics. Perloff and Salop (1985) introduce a model where every consumer places a different relative value on each available brand. The consumer's preferences, along with the price of a brand, results in CONSUMER SURPLUS, and the consumer buys the brand which gives the highest surplus (the "best buy"). However, in their structure, all brands compete symmetrically with all the others (a Chamberlinian element).

Hotelling-Type Models

In this family of models, product variety is due to the large number of consumers in the market with diverse preferences. In fact, it is usually assumed that consumers have an "ideal" variety for every product they consume. In other words, consumers can be "located" along the product characteristics space, a space where each point represents a different product variety. The consumer's location in this space indicates her ideal variety.

Introduced by Hotelling (1929), the model assumes that consumers are uniformly distributed along the product characteristics space and each one has an ideal product variety. The product characteristics space is linear (a straight line) and bounded, and can be considered equivalent to the market for the product. There are also many firms in the market, and each one produces one brand of the product. Firms decide about the particular attributes of these brands by choosing the location of their brands along the product characteristics space. Brands compete with only their neighboring products on either side of their location in the product characteristics space. Hence, the degree of substitution among different brands is not the same for all brands.

Based on this structure, Hotelling showed that in a case of a *duopoly* (a market with just two firms), when companies do not compete in terms of price, it is optimal for them to locate next to each other in the middle of the product characteristics space. In other words, firms choose varieties which are very close SUBSTITUTES, since that way they can achieve maximum market share. The reason for this is that consumers will purchase from the store that is closest to them in the market. If consumers were located along a one-mile stretch of road, the first firm in this market would want to locate in the middle – one half-mile down the road – because this location allows the firm to attract the greatest number of consumers. Given this, the next firm would want to locate next to the first because it will also attract the greatest number of consumers, and since the firms do not compete on price, only their location matters to customers. This result, known as *minimum product differentiation*, depends on specific assumptions of the model, as will be discussed later. High market demand, low fixed costs (weak ECONOMIES OF SCALE), and weak substitution among different varieties lead to high product variety in the market. Overall, this model predicts a larger number of varieties in the market than the Chamberlinian models do, because there is weaker COMPETITION among brands, since varieties compete only against their neighbors and not against all the other varieties in the market.

As mentioned, Hotelling's result of minimum product differentiation is a special one. It depends on the assumptions that companies do not compete in terms of price and that the market is bounded. Generalizing Hotelling's results, Eaton and Lipsey (1975) show that in a case of a large number of firms, there is an EQUILIBRIUM where firms locate in pairs along the product characteristics space with some space between the pairs. In other words, there is some *clustering* of varieties in the market.

Other researchers derive Hotelling's result of minimum product differentiation by assuming instead that consumers are not uniformly distributed along the product characteristics space but are clustered around certain points (in some cases in the middle of the product characteristics space). In other words, consumers tend to prefer some varieties relative to others. Then, companies have incentives to locate in those areas of high demand (*thick markets*). This is one of the causes of geographic concentration of economic activity in large metropolitan areas (*spatial agglomeration*).

Finally, Lancaster (1979) introduces a model where consumers perceive products as bundles of characteristics and have preferences over different collections of characteristics, not necessarily over individual products. Therefore, by combining different products in their consumption, they can end up with the set of characteristics and qualities they find most desirable. Called the *characteristics approach*, this model allows formalization of different situations such as whether products can or cannot be combined in consumption, or whether such combinations preserve the characteristics of the separate products.

Monopolistic Competition and Price Markups

Recent research has studied how the entry and exit of firms in a monopolistically competitive market influences the pricing policy of companies, and especially their price markups (*see* MARKUP PRICING). In the context of the Chamberlinian models, price markups are constant and not related to entry and exit because the degree of substitution among different varieties is constant and not influenced by the number of brands in the market. In terms of Hotelling-type models though, price markups are inversely related to the entry of firms since market entry implies a larger number of brands, stronger competition, a higher price elasticity of demand, and lower price markups. The conclusion is that periods of high market demand and therefore high market entry (e.g., an economic boom or a seasonal increase in market demand during Christmas) are associated with lower price markups. Weitzman (1982) and others present theoretical justification for this result, while Barsky and Warner (1995) offer some empirical support.

Bibliography

Barsky, R. B. and Warner, E. J. (1995). The timing and magnitude of retail store markdowns: Evidence from weekends and holidays. *Quarterly Journal of Economics*, **110**, 321–52.

Chamberlin, E. H. (1933). *The Theory of Monopolistic Competition*. Cambridge, MA: Harvard University Press.

Dixit, A. K. and Stiglitz, J. E. (1977). Monopolistic competition and optimum product diversity. *American Economic Review*, **67**, 297–308.

Eaton, B. C. and Lipsey, R. G. (1975). The principle of minimum differentiation reconsidered: Some new developments in the theory of spatial competition. *Review of Economic Studies*, **42**, 27–49.

Hotelling, H. (1929). Stability in competition. *Economic Journal*, **39**, 41–57.

Lancaster, K. J. (1979). *Variety, Equity, and Efficiency*. New York: Columbia University Press.

Perloff, J. M. and Salop, S. C. (1985). Equilibrium with product differentiation. *Review of Economic Studies*, **52**, 107–20.

Salop, S. C. (1979). Monopolistic competition with outside goods. *Bell Journal of Economics*, **10**, 141–56.

Weitzman, M. (1982). Increasing returns and the foundation of unemployment theory. *Economic Journal*, **92**, 787–804.

monopoly

Robert E. McAuliffe

A monopoly exists when there is a single seller of a product in the industry. As a consequence, the monopolist's demand curve is the same as the industry demand curve and is thus downward sloping (*see* DEMAND CURVES). A monopolist must be protected by BARRIERS TO ENTRY to remain a single seller while earning ECONOMIC PROFIT. The monopolist may be protected by licensing requirements, patents, its own STRATEGIC BEHAVIOR, or ECONOMIES OF SCALE which prevent ENTRY into the industry.

Since the industry demand curve facing the monopolist is the *average revenue* curve, when it is decreasing the MARGINAL REVENUE curve must lie below it. To maximize profits the monopolist will produce until marginal revenue equals MARGINAL COST, a condition all

profit-maximizing firms must fulfill regardless of MARKET STRUCTURE (*see* PROFIT MAXIMIZATION). If the monopolist sells a single, nondurable product and market demand is

$$P = f(Q) \tag{1}$$

where P is the market price of the product, Q is the quantity sold, and $f(Q)$ is the inverse DEMAND FUNCTION, then the marginal revenue for the monopolist is

$$\text{MR} = \frac{\Delta \text{TR}}{\Delta Q} = \text{P} + Q \times \frac{\Delta P}{\Delta Q} < P \tag{2}$$

where ΔTR is the change in TOTAL REVENUE and ΔQ is the change in the quantity sold (see Martin, 1994). The marginal revenue curve slopes downward because, in the absence of PRICE DISCRIMINATION, the monopolist must lower the price to all consumers of the product in order to sell an additional (marginal) unit, and this is why the third term on the right in equation (2) above is less than P. Multiplying and dividing the next to last term in equation (2) by P yields

$$\text{MR} = P \times \left(1 + \frac{Q \Delta P}{P \Delta Q}\right) = P\left(1 + \frac{1}{e_p}\right) \tag{3}$$

where e_p is the price ELASTICITY of market demand for the product. Since the elasticity of demand is always negative, the marginal revenue in equation (3) above will be negative whenever market demand is inelastic (when $|e_p| < 1$). This means that the monopolist could increase revenues (and profits) by selling fewer units and raising price, so an unregulated, profit-maximizing monopolist should never operate on the inelastic portion of its demand curve. Setting marginal revenue in equation (3) above equal to marginal cost (MC) yields the profit-maximizing markup for the monopolist (*see* MARKUP PRICING):

$$\begin{aligned} \text{MC} - P &= \frac{P}{e_p} \\ \frac{P - \text{MC}}{P} &= -\frac{1}{e_p} \end{aligned} \tag{4}$$

Thus the optimal markup for a monopolist is inversely related to the absolute value of the elasticity of demand. The term on the left-hand side is the LERNER INDEX of monopoly power and shows that the monopolist will charge a higher price markup over marginal cost when consumers have fewer SUBSTITUTES (market demand is less elastic).

The Social Costs of Monopoly

In the preceding discussion it was shown that a monopolist will set its price above marginal cost, while in perfectly competitive markets price equals marginal cost (*see* PERFECT COMPETITION). This means that some consumers who are willing to pay the cost to society of producing this product do not receive it, and this is called DEADWEIGHT LOSS. Consider an example where, for simplicity, there are no FIXED COSTS in production, so that the long run marginal cost (LRMC) is equal to long run average cost (LRAC) in figure 1.

To maximize profits the monopolist will produce Q_m units of output (where MR = MC) and the market price for Q_m units will be P_m. If this industry were perfectly competitive it would produce until $P = \text{MC}$, which would be Q_c units sold at price P_c above. Compared with a perfectly competitive industry, the monopolist produces less output and charges a higher price. In figure 1, the rectangle P_m,M,B,P_c is the monopolist's profit, which is a transfer from consumers (who lose this amount in CONSUMER SURPLUS) to the monopolist. Since this transfer makes consumers worse off and the monopolist better off, the welfare consequences cannot be judged on the grounds of Pareto optimality (*see* PARETO OPTIMAL ALLOCATION). However, the triangle M,B,C is lost consumer surplus that is not gained by the monopolist. This is the deadweight (efficiency) loss to society from a monopoly. Once the monopolist has earned its profits, hypothetically it should be willing to produce additional units for those consumers who are willing to pay the marginal cost of production (that is, those consumers along the demand curve between points M and C). This production would make consumers (and therefore society) better off and leave the monopolist no worse off, so it is a more efficient (Pareto optimal) allocation. Since the monopolist does not produce and sell these additional units, this loss in allocative efficiency is one of the costs to society from monopoly.

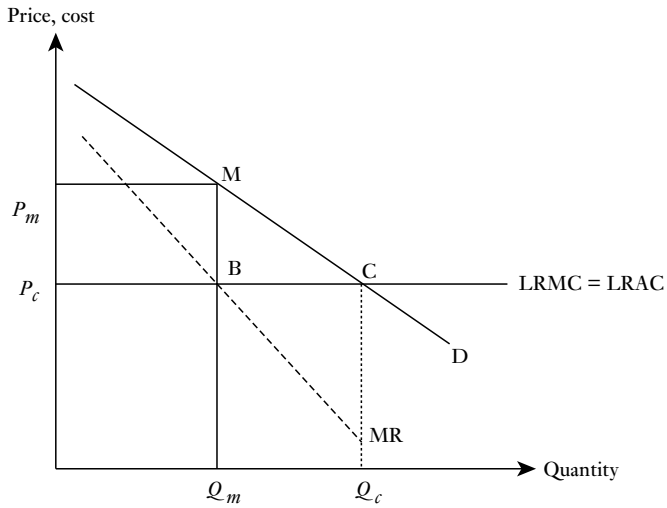

Figure 1 Profit-maximizing price and output under monopoly

Another potential cost of monopoly is what Leibenstein (1966) termed X-inefficiency (*see* X-EFFICIENCY), the failure to minimize production costs. Without the pressure of competition, a monopolist has less incentive to be efficient and may waste resources in production. Therefore the monopolist's costs would not be equal to those achievable under perfect competition, and so the social costs of monopoly would include the monopolist's higher costs of production in addition to the deadweight loss.

Potential monopoly profits may encourage firms or individuals to expend resources attempting to acquire or maintain a monopoly. Such behavior is called RENT SEEKING and if these expenditures do not create benefits to society, then the effort to monopolize is costly to society. Posner (1975) argued that if firms compete for these monopoly profits, they would expend resources until the net expected profit was zero, and this meant that the social cost of monopoly was all of the monopolist's profits. Although this represents an extreme upper bound, some resources are wasted in rent-seeking behavior and this increases the costs to society from a monopoly (see Tirole, 1988).

An interesting issue considered by Arrow (1962) was the effect of market structure on the incentives to innovate. In his model, firms could innovate to reduce their marginal costs of production, and he examined the returns from this innovation under perfect competition and monopoly. Arrow found that the monopolist had *less* incentive to innovate because it was already earning monopoly profits and the innovation would replace its existing position. On the other hand, a competitive firm could become a monopoly and would earn higher profits. Therefore monopolies may also be costly to society in terms of dynamic efficiency: these firms may fail to innovate or introduce new products. However, if the monopolist is threatened by entry, it has a greater strategic incentive to innovate and preempt potential entrants because in doing so it will remain a monopolist, while an entrant will have to share the market with the monopolist (see Tirole, 1988, for a survey).

Monopolies can be beneficial to society under certain conditions. When there are economies of scale, a single firm may be able to serve the market more efficiently than several firms (*see* NATURAL MONOPOLY). Governments grant monopoly licenses to firms and individuals through patents in an effort to encourage innovation. In this case, rent-seeking behavior has socially desirable benefits which society promotes.

Bibliography

Arrow, K. J. (1962). Economic welfare and the allocation of resources for inventions. In R. Nelson (ed.), *The Rate and Direction of Inventive Activity*. Princeton, NJ: Princeton University Press.

Keat, P. G. and Young, P. K. Y. (2002). *Managerial Economics: Economic Tools for Today's Decision Makers*, 4th edn. Upper Saddle River, NJ: Prentice-Hall.

Leibenstein, H. (1966). Allocative efficiency vs. "X-efficiency." *American Economic Review*, **56**, 392–415.

Martin, S. (1994). *Industrial Economics*, 2nd edn. New York: Macmillan.

Posner, R. (1975). The social costs of monopoly and regulation. *Journal of Political Economy*, **83**, 807–27.

Tirole, J. (1988). *The Theory of Industrial Organization*. Cambridge, MA: MIT Press.

monopsony

Lidija Polutnik

Monopsony exists when one firm is the only buyer of an input or product. The classic example of a monopsony is a one-company town where the only employer in the town is a textile mill or a coal mine. In purchases of inputs the monopsonist can derive its power from other factors as well. For example, purchases of "house brand" goods where suppliers supply most of their output to a single retailer provide the buyer with a significant amount of monopsony power (also buyer power).

How many inputs should the monopsonist hire? What should be their pay? In a competitive labor market, each employer takes the wage rate as given, and the MARGINAL COST of hiring one more worker is simply the wage rate. In this case the marginal cost of the input curve is horizontal and coincides with its supply curve. The supply curve facing the monopsonist is far different. A single buyer of labor in a textile mill town faces the entire labor supply of the town. While the competitive firm can hire all the labor it wants at the going wage, the monopsonist cannot. The monopsonist must pay successively higher wage rates to attract additional workers and the higher wage rate applies to all the workers previously hired. As a result the marginal cost of input (labor) curve lies above its supply curve. A profit-maximizing monopsonist hires workers up to the point where the MARGINAL REVENUE product as defined by the derived demand for labor of the last worker hired equals the marginal input cost. The wage rate is determined by the height of the labor supply curve at the level of employment. The monopsonist restricts hiring in order to pay a lower price for labor or other inputs. For example, in a monopsonistic labor market the firm pays a lower wage rate than would prevail in the case of a competitive market, because workers lack alternative sources of employment and specialized inputs in general can be sold to one or, at best, a limited number of users. A gap between the derived demand of an input and the supply curve reflects the loss in EFFICIENCY in a monopsony.

A monopsony, like a monopoly, confronts the long run problem of EXIT and ENTRY. If the single buyer succeeds in depressing the input price below the competitive price, it may not be feasible for a specialized input to exit immediately. However, in the long run suppliers will not replace specialized equipment when it becomes obsolete and the incentive to produce inputs that are limited to one or a few uses will decrease. A producer that is concerned about the long run relation with its supplier may be better off in not exercising its monopsony power in the SHORT RUN (Carlton and Perloff, 2000).

Perry (1978) demonstrates that a powerful incentive for BACKWARD INTEGRATION is provided by higher profits than can be earned if a monopsony integrates backward. If the monopsonist purchases its supplier, then it does not need to pay a higher input price on all the units used if it chooses to expand an input usage. The marginal input cost schedule becomes irrelevant and the integrated firm chooses to hire the input up to where the derived demand of input equals the market supply curve of the input. Complete backward integration in the input markets results in the allocatively efficient outcome of a competitive input market.

Bibliography

Carlton, D. W. and Perloff, J. M. (2000). *Modern Industrial Organization*, 3rd edn. Reading, MA: Addison-Wesley.

Perloff, J. M. (2004). *Microeconomics*, 3rd edn. Boston: Pearson Addison-Wesley.
Perry, M. K. (1978). Vertical integration: The monopsony case. *American Economic Review*, **68**, 562–70.

moral hazard

Alexandra Bernasek

Moral hazard is an incentive problem that arises in cases where the actions of individuals cannot be observed and contracted upon, creating ASYMMETRIC INFORMATION among individual parties to a transaction. Moral hazard arises commonly in insurance markets, financial markets, and labor markets. It arises in situations involving cooperative effort by two or more people, and in the context of principal–agent relationships (*see* PRINCIPAL–AGENT PROBLEM). In some cases of moral hazard the actions of individuals are unobservable, but in many cases the problem is that the costs of observability are prohibitively high. Those costs are described as "monitoring costs." The nature of transactions characterized by moral hazard is such that individuals do not have incentives to behave in ways that lead to Pareto efficient outcomes (*see* PARETO OPTIMAL ALLOCATION). Solving the problem of moral hazard involves designing incentive CONTRACTS that reduce monitoring costs and combine risk sharing with the creation of appropriate incentives.

Insurance markets provide a good illustration of the moral hazard problem; if people can insure themselves against certain risks they are less likely to act with an appropriate level of care. Consider the example of auto theft. If the probability of a theft occurring depends on the actions of the car owner (e.g., where the individual parks, whether or not they lock their car doors, etc.), then the insurance company faces an incentive problem; they want the car owner to take actions that minimize the probability of theft. In this case, full insurance will not be optimal since the insurance company will want the car owner's wealth to depend on his or her actions, thus creating incentives for the car owner to take the proper amount of care (e.g., this can be used to explain deductibles). The amount of insurance an individual can purchase at actuarially fair rates is effectively rationed in the presence of moral hazard. An early article establishing these results is Spence and Zeckhauser (1971).

Holmstrom (1979) has shown that in the presence of moral hazard, Pareto optimal risk sharing in the context of a principal–agent problem is generally not possible because it won't induce the agents to take actions that are desired by the principal. A second-best solution to the problem is possible which trades off some risk-sharing benefits for the provision of incentives for the desired behavior. One solution is to spend resources on monitoring agents' actions and use this information in the contract. Holmstrom shows that any additional information about the agent's actions, no matter how imperfect, can be used to improve the welfare of both the principal and the agent.

Shavell (1979) also explores what Pareto optimal fee schedules would look like in agency relationships characterized by moral hazard. He finds that the characteristics of these fee schedules are related to both the principal's and the agent's attitudes toward risk (*see* RISK AVERSION). He discusses the implications of his results for four examples of principal–agent relationships: strict liability versus negligence standards in the control of stochastic EXTERNALITIES, insurance, the lawyer–client relationship, and the relationship between STOCKHOLDERS and managers of a firm.

The analysis of moral hazard in the principal–agent context extends beyond the principal–single agent to a multi-agent setting. Two features of a multi-agent setting that are not present in a single-agent setting are free-riding and COMPETITION. Holmstrom (1982) explores these as a way of better understanding the organizational design of firms. His focus is on "moral hazard in teams," where a team refers to a group of individuals who are organized in a way that productive inputs into the firm are related. He then examines the organization of production in this context when the agents' inputs are imperfectly observed. He finds that the free-rider problem can be resolved if ownership and labor are to some extent separated, and that relative performance evaluation, such as peer production averages, can reduce moral hazard costs. The latter finding is one

explanation for executive incentive packages that base rewards on comparisons with peer firms (firms within the same industry).

Varian (1990) explores several issues related to group incentives in the context of a real-world credit institution: the Grameen Bank in Bangladesh. One of the issues Varian explores is the use of groups of borrowers to reduce monitoring costs by the bank as a way of overcoming the moral hazard problem inherent in the credit relationship. The Grameen Bank creates incentives for borrowers to monitor one another by requiring them to form groups of five from among their peers and accept joint responsibility for repaying all group members' loans. Because their access to loans is tied to the repayment performance of the group, agents have an incentive to monitor their fellow group members to make sure they are taking actions that are consistent with repaying their loans. Stiglitz (1990) also models this use of "peer monitoring" as a solution to the moral hazard problem in credit markets.

In an early article on the relationship between agency costs and the ownership structure of the firm, Jensen and Meckling (1976) examine the problem of moral hazard in the context of the separation of ownership and control in the modern corporation. There is divergence between the interests of owners of the firm and managers. A moral hazard problem arises because of the difficulty owners have observing the actions of managers and thus writing complete contracts which require them to act in the owners' best interests. Real resources will therefore be allocated to monitoring activities by the owners (shareholders). In their definition of agency costs, Jensen and Meckling include the monitoring expenditures of shareholders: the costs associated with their attempts to measure and observe the actions of managers, as well as the costs of their efforts to control the actions of managers. Examples of these activities are audits, formal control systems, budget restrictions, and incentive compensation schemes.

Grossman and Hart (1980) examine the role of takeover bids in creating incentives for efficient management of the "stock market corporation." They assume significant costs in ensuring that managers and directors of corporations act in the interests of the shareholders. They model the stock market corporation as a common property resource (*see* PUBLIC GOODS PROBLEM) and explain the deviation between potential benefits and actual benefits of collective action in terms of the problem of moral hazard. The extent of this deviation will depend upon the extent of the unpredictability associated with takeover bids. They argue that the higher the threat of a raid, the more efficient current management of a corporation will be. Shareholders will thus weigh the costs and benefits of creating the possible threat of a takeover bid as a way of monitoring the actions of managers.

The role of moral hazard in explaining the organization of production within firms has been an important area of research. Alchian and Demsetz (1972) examine information problems associated with the organization of production to explain how "team production" induces the contractual process of firms. Team production involves a cooperative activity, the essence of which is that the contribution of individual cooperating inputs to output cannot be identified. Their central question is one of incentives: how can members of a team be rewarded and induced to work efficiently? They conclude that monitoring to reduce shirking can be achieved more efficiently in a firm than through market bilateral negotiations among inputs. Mirrlees (1976) also explores issues of optimal payment schedules and organizational structure in cases where the performance of individuals in production can only be imperfectly observed. His conclusion that IMPERFECT INFORMATION binds the organization together is consistent with the conclusions of Alchian and Demsetz.

Attempts to study the effects of moral hazard in the labor market have been important for explaining disequilibrium phenomena such as involuntary unemployment. Firms may pay above market-clearing wages in order to prevent workers from shirking when their effort cannot be directly observed. These "efficiency wages" may increase worker effort for two reasons. One is the threat of unemployment that is created by the above-market wages, and another is improved worker morale. The efficiency wage is the wage at which labor costs are minimized. Efficiency wage theory has also been used to explain the observed downward rigidity of wages, and layoffs. For a survey of efficiency

wage models of the labor market, see Akerlof and Yellen (1986).

See also *asymmetric information; imperfect information; principal–agent problem*

Bibliography

Akerlof, G. A. and Yellen, J. L. (eds.) (1986). *Efficiency Wage Models of the Labor Market*. Cambridge: Cambridge University Press.

Alchian, A. and Demsetz, H. (1972). Production, information costs, and economic organization. *American Economic Review*, **62**, 777–95.

Grossman, S. and Hart, O. (1980). Takeover bids, the free rider problem, and the theory of the corporation. *Bell Journal of Economics*, **11**, 42–64.

Holmstrom, B. (1979). Moral hazard and observability. *Bell Journal of Economics*, **10**, 74–91.

Holmstrom, B. (1982). Moral hazard in teams. *Bell Journal of Economics*, **13**, 324–40.

Jensen, M. and Meckling, W. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, **3**, 305–60.

Keat, P. G. and Young, P. K. Y. (2002). *Managerial Economics: Economic Tools for Today's Decision Makers*, 4th edn. Upper Saddle River, NJ: Prentice-Hall.

Mirrlees, J. A. (1976). The optimal structure of incentives and authority within an organization. *Bell Journal of Economics*, **6**, 105–31.

Shavell, S. (1979). Risk sharing and incentives in the principal and agent relationship. *Bell Journal of Economics*, **10**, 55–73.

Spence, A. M. and Zeckhauser, R. (1971). Insurance, information and individual action. *American Economic Review*, **61**, 380–7.

Stiglitz, J. (1990). Peer monitoring and credit markets. *World Bank Economic Review*, **4**, 351–66.

Varian, H. R. (1990). Monitoring agents with other agents. *Journal of Institutional and Theoretical Economics*, **146**, 153–74.

multicollinearity

Alastair McFarlane

Multicollinearity refers to the situation where the explanatory variables of a LINEAR REGRESSION are highly correlated with each other. This phenomenon is especially common in TIME-SERIES DATA because of the presence of lagged variables and common time trends among explanatory variables. Perfect multicollinearity exists when an explanatory variable is an exact linear combination of other explanatory variables, making it mechanically impossible to calculate regression coefficients. This extreme case is unlikely unless the researcher has constructed a poorly specified model.

It is difficult to disentangle the separate effects of explanatory variables on the explained variable when multicollinearity is present. Coefficients may have the wrong sign or an implausible magnitude and small changes in the data can produce wide swings in the estimated coefficients. When combined with high VARIANCE in the error term and low variance in the explanatory variables, a high degree of multicollinearity will lead to high standard errors of the coefficients and low t-statistics (*see* STANDARD ERRORS OF THE COEFFICIENT; T-STATISTIC). However, multicollinearity does not adversely affect the predictive power of the regression model as a whole.

One method of detecting multicollinearity is to examine the correlation matrix for high correlation among the explanatory variables. A signal that multicollinearity is problematic is a high R^2 and low t-statistics. More formal measures of the degree of multicollinearity are the condition number, the variance-inflation factor, and Theil's measure (see Maddala, 2001).

A remedy is to increase the sample by adding more data which, by providing additional information, will help to lower variances. Since obtaining more data is usually difficult, researchers may resort to ad hoc techniques such as ridge regression, principal components regression, or omitting the offending variables. These ad hoc techniques have been criticized for producing biased or meaningless estimates. In such cases it may be advisable to accept multicollinearity and interpret the regression coefficients with caution.

Bibliography

Greene, W. H. (2002). *Econometric Analysis*, 5th edn. Upper Saddle River, NJ: Prentice-Hall.

Gujarati, D. N. (2002). *Basic Econometrics*, 4th edn. New York: McGraw-Hill Irwin.

Maddala, G. S. (2001). *Introduction to Econometrics*, 3rd edn. New York: John Wiley.

multi-unit auctions

Brett Katzman

These are more elaborate AUCTIONS used to sell multiple units of a single good. Multi-unit auctions have been used for decades by the US Treasury to sell notes, bills, and bonds, are used every day on eBay, and have garnered much attention due to their success in the sale of PCS spectrum rights.

There are four primary ways of auctioning multiple units. A method used for years by the US Treasury is the discriminatory (or pay-your-bid) auction, where bidders submit sealed bids. Those submitting the highest bids for units being sold are awarded the objects and must pay the (perhaps different and therefore discriminatory) amount of their specific bid(s). The US Treasury has recently begun experimenting with uniform price auctions where bidders also submit sealed bids. Those submitting the highest bids are awarded the items, but each winner pays the same price per unit, which is determined by the market-clearing bid. The uniform bid is similar to the single-unit second-price auction, so much so that it is commonly (and falsely) considered the multi-unit analog thereof. The actual analog to a second-price auction is the Vickrey auction (named after Nobel Laureate William Vickrey), which has a much more complex payment rule. Finally, perhaps the simplest method for auctioning multiple units is to sell them sequentially using single-unit auctions.

Bibliography

Cramton, P. and Ausubel, L. (2002). Demand reduction and inefficiency in multi-unit auctions. Working paper, University of Maryland.

McAfee, R. P. (2002). *Competitive Solutions: The Strategist's Toolkit*. Princeton, NJ: Princeton University Press.