

Unit 5.4: Monopoly

Michael Malcolm

June 18, 2011

1 Price Making

A firm has a *monopoly* if it is the only seller of some good or service with no close substitutes. The key is that this firm has the power to set the price. In perfect competition, there is a downward sloping market demand, but each firm takes the market price as given. By contrast, in monopoly, the firm can choose any price/quantity combination on its demand curve.

2 Profit Maximization

Consider a monopoly facing market demand curve $P = 300 - 2q$ and with total cost function $TC = 100 + q^2$. The monopoly's objective is to maximize its profit. Recall that profit is:

$$\begin{aligned}\Pi &= TR - TC \\ &= P * q - TC\end{aligned}$$

The key difference is that, for a perfectly competitive firm, the price P is fixed. For a monopoly firm, the price P depends on the level of output that it sells. Specifically, selling more output requires lowering the price. Substituting in the demand and cost functions, profit is:

$$\begin{aligned}\Pi &= (300 - 2q)q - (100 + q^2) \\ &= 300q - 2q^2 - 100 - q^2\end{aligned}$$

The firm maximizes its profit by setting the derivative of the profit function equal to zero.

$$\begin{aligned}\frac{d\Pi}{dq} &= 300 - 4q - 2q = 0 \\ 300 &= 6q \\ q &= 50\end{aligned}$$

Substitute this back into the demand curve to find the price:

$$P = 300 - 2q = 300 - 2(50) = 200$$

So this monopoly maximizes profit by selling 50 units of output, each for \$200.

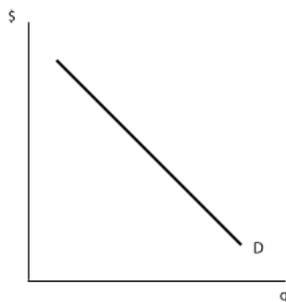


Figure 1: A monopoly faces a downward sloping demand curve

3 Marginal Revenue

To delve into this decision a bit deeper, first note that a monopolist faces a downward sloping demand curve, as shown in figure 1. Unlike a perfect competitor, who can sell as much as he wants at a price that is fixed from the perspective of a single firm, a monopolist controls the demand curve and must cut the price in order to sell more output.

Unlike in a competitive market, where marginal revenue is equal to price, marginal revenue for a firm with market power is *lower* than the price. To see why, think about the following simple demand function.

P	Q	TR	MR
12	0	0	–
11	1	11	11
10	2	20	9
9	3	27	7

When the firm sells the first unit at \$11, it earns \$11 of additional revenue. However, when the firm sells the second unit at \$10, the marginal revenue is only \$9. To see why, notice that the firm cut the price to \$10 for *both* units, not just the second unit. So, while the firm gets \$10 of revenue from the second customer, it loses \$1 because of selling to the first customer at \$10 instead of \$11.

The increase of \$10 is known as the *quantity effect* – the increase in revenue because more units are sold. The decline of \$1 is known as the *price effect* – the decline in revenue from cutting price for units that were previously sold for a higher price. The two combine for the net change in revenue of \$9. Notice that, in perfect competition, there is no price effect because the firm faces a fixed price no matter how much output it sells.

Because the monopolist charges a single price to all customers, selling more output by cutting the price requires cutting price for *all* units, not just for the marginal units sold. Thus, the additional revenue that the firm earns is less than the price for which the marginal unit is sold.

To see this graphically, consider figure 2. Initially, the firm sells q_0 units of output, each at P_0 . The firm decides to sell more units, raising output to q_1 by lowering the price to P_1 . Total revenue is initially the rectangle around P_0 and q_0 , changing to the rectangle around P_1 and q_1 . The quantity effect is the increase in revenue that results from the new units that are sold at P_1 . The price effect is the decline in revenue that results from lowering the price from P_0 to P_1 for the units that were previously sold at P_0 .

It is possible for the price effect to offset the quantity effect. That is, the firm might actually *lose* revenue by selling additional units. The demand curve shows the price that consumers are willing to pay, so graphically the marginal revenue curve is lower than the demand curve since $MR < P$, as in figure 3. The point where marginal revenue falls below zero is exactly the point where the price effect offsets the quantity effect, so that revenue declines by selling additional units.

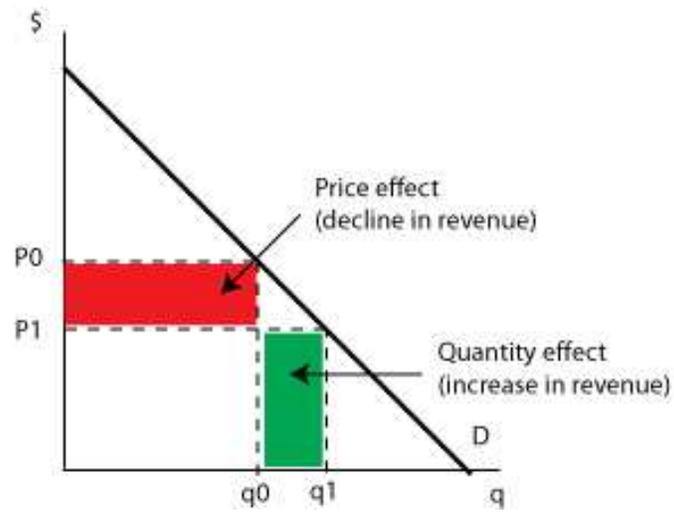


Figure 2: Price effect and quantity effect

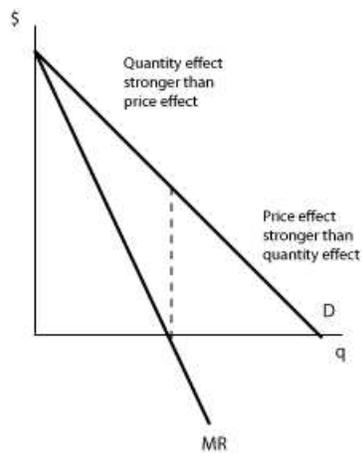


Figure 3: MR is lower than demand, and can even be negative

For a linear demand curve, there is a simple derivation of the marginal revenue curve. Consider any demand curve with the form $P = a - bq$. Total revenue is:

$$\begin{aligned} TR &= Pq \\ &= (a - bq)q = aq - bq^2 \end{aligned}$$

Marginal revenue from selling one more unit is the first derivative with respect to q .

$$\begin{aligned} MR &= \frac{dTR}{dq} \\ &= a - 2bq \end{aligned}$$

When the demand curve is $P = a - bq$, then the corresponding marginal revenue curve is $MR = a - 2bq$. That is, for a linear demand curve, the marginal revenue curve has the same intercept as the demand curve, with twice the slope.

Notice that, in our problem earlier with $P = 300 - 2q$, the corresponding marginal revenue curve is $MR = 300 - 4q$. Since the total cost function is $TC = 100 + q^2$, marginal cost is $MC = 2q$.

Recall our profit maximization problem. Profit was

$$\begin{aligned} \Pi &= (300 - 2q)q - (100 + q^2) \\ &= 300q - 2q^2 - 100 - q^2 \end{aligned}$$

The first-order condition for profit maximization was

$$\begin{aligned} \frac{d\Pi}{dq} &= 300 - 4q - 2q = 0 \\ 300 - 4q &= 2q \end{aligned}$$

In other words, the $MR = MC$ condition familiar from principles courses is just the first-order condition for profit maximization. Graphically, figure 4 shows monopoly profit maximization. The firm sets its output where $MR = MC$ and finds the price on the demand curve corresponding to this level of output.

4 Supply Curve

A monopoly (or any firm with market power) has no supply curve. The monopolist can choose any price/quantity combination on the demand curve that it wants to choose (in fact, it picks the price/quantity combination that maximizes profit). It is not constrained by any supply curve.

In essence, a supply curve is an answer to the question: "If the price is P , how much output q does the firm want to sell?" The question makes no sense for a monopolist since the monopolists actually *sets* the price.

5 Shut Down

Like any other firm, the monopolist shuts down in the short run if it cannot earn enough revenue to cover its variable cost. Stated in per-unit terms, the monopolist shuts down when there is no price (read from the demand curve) that covers its AVC . Figure 5 shows a firm that should shut down.

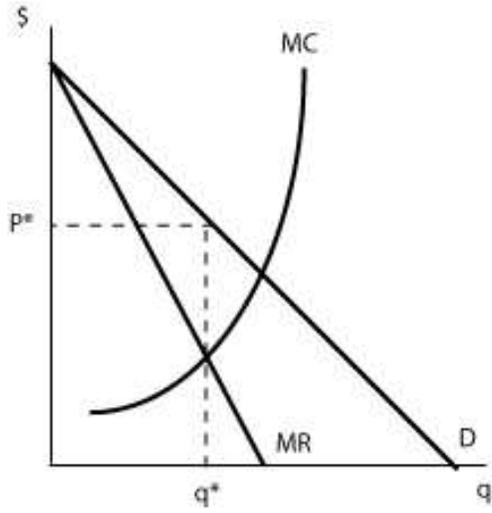


Figure 4: Profit-maximization for a single price monopoly

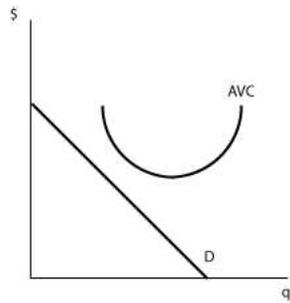


Figure 5: A monopoly that should shut down

6 Characterizing the Profit-Maximizing Choice

Total revenue is $TR = P \cdot q$. The important thing to notice is that, for a monopoly, P is a function of q . We calculate marginal revenue by taking the first derivative of the total revenue function. Using the product rule.

$$MR = \frac{dTR}{dq} = \frac{dP}{dq} \cdot q + 1 \cdot P$$

For a perfect competitor $\frac{dP}{dq} = 0$ since a single seller cannot influence the market price, so $MR = P$. However, continuing with the calculation for the monopoly case, multiply the first term by $\frac{P}{P}$ and recall that the price elasticity of demand is defined as $\varepsilon = \frac{dq}{dP} \frac{P}{q}$.

$$\begin{aligned} MR &= \frac{dP}{dq} \cdot q \cdot \frac{P}{P} + 1 \cdot P \\ &= P \left(\frac{dP}{dq} \cdot \frac{q}{P} + 1 \right) \\ &= P \left(\frac{1}{\varepsilon} + 1 \right) \end{aligned}$$

Notice that as $\varepsilon \rightarrow -\infty$ that $MR \rightarrow P$. That is, as the demand becomes more and more elastic, the market starts to approximate perfect competition.

Now, the monopoly maximizes profit where $MR = MC$:

$$\begin{aligned} MR &= MC \\ P \left(\frac{1}{\varepsilon} + 1 \right) &= MC \\ P &= \left(\frac{\varepsilon}{1 + \varepsilon} \right) MC \end{aligned}$$

This characterizes the monopoly's profit-maximizing price. For example, when $\varepsilon = -2$, the profit-maximizing price is $P = 2MC$, but when $\varepsilon = -4$ then the price is $P = \frac{4}{3}MC$. As demand becomes more inelastic (meaning that the firm has more market power), the price gets distorted further and further above marginal cost. Again, as $\varepsilon \rightarrow -\infty$, notice that $P \rightarrow MC$, so the closer a firm is to having no market power (i.e. no ability to raise the price), the outcome gets closer and closer to the competitive outcome where $P = MC$.

Notice that the formula doesn't make sense if $\varepsilon < -1$. The reason for this is just that a monopoly never sets its price on the inelastic part of the demand curve. If demand is inelastic, then raising price causes total revenue to rise since price rises by more than quantity falls. At the same time, since output falls (even only a little), costs fall too. If demand is inelastic, revenue rises and costs fall when price is increased, meaning that profit rises. So, any time a firm faces inelastic demand, it should continue to raise its price – doing so will raise profit.

This means that a profit-maximizing firm will never price on the inelastic part of its demand curve since it could continue to raise its price, and in doing so raise its profit. When demand is elastic and the firm raises price, total revenue falls and total costs fall too. The firm has to calibrate whether to raise price or not on this segment by balancing out the decline in revenue with the decline in costs. But, in any case, *a profit-maximizing firm always sets its price on the elastic part of its demand curve.*

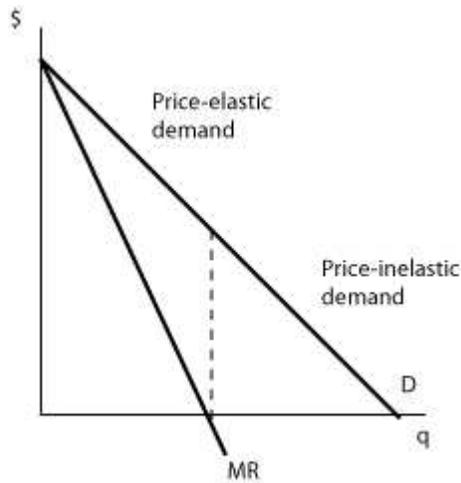


Figure 6: $MR=MC$ on the elastic part of the demand curve

Graphically, this is easy to see. Consider figure 6. The demand curve is elastic to the left of the midpoint of the demand curve, and this midpoint is precisely where the marginal revenue curve hits the axis since it falls twice as fast as the demand curve. Since the firm sets output where $MR = MC$, any time marginal costs are positive this profit-maximizing output will always occur to the left of the midpoint of the demand curve, where demand is elastic.

7 Distortion

Rearranging the characterization of the profit-maximizing price from above:

$$\begin{aligned}
 MR &= MC \\
 P \left(\frac{1}{\varepsilon} + 1 \right) &= MC \\
 P \left(\frac{1}{\varepsilon} \right) + P &= MC \\
 P - MC &= -P \left(\frac{1}{\varepsilon} \right) \\
 \frac{P - MC}{P} &= -\frac{1}{\varepsilon} \\
 \frac{P - MC}{P} &= \frac{1}{|\varepsilon|}
 \end{aligned}$$

The left side of this equation, $\frac{P-MC}{P}$, is sometimes called the *Lerner index*. It is the percentage of the firm's price that is a markup over marginal cost. Economists interpret the Lerner index as a measure of distortion in a market. Notice that, as $|\varepsilon|$ rises and demand becomes more elastic, the Lerner index falls. That is, the more elastic the firm's demand is, the less the distortion of price over marginal cost. A firm with perfectly elastic demand is a competitive price taker, and in this case there is *no* distortion – price is exactly equal to marginal cost.

A firm's market power – its ability to mark up price over marginal cost – is determined by how inelastic its demand is. This makes good economic sense. Firms with more inelastic demand can charge larger markups over marginal cost.

There is an important point here to be careful about. As we emphasized in the last section, market power (price marked up above marginal cost) is *not* the same thing as being able to earn a long run profit. Regardless of how much market power a firm enjoys, if there is no barrier to entry, then this profit will not continue in the long run – new firms will enter and eat up economic profits. The ability to earn long run profit relates to entry conditions, *not* to market power.

8 Efficiency and Price Discrimination

The reason that we compare price to marginal cost as a measure of distortion is that $P = MC$ is the gold standard for allocative efficiency in a market. *Allocative efficiency* just means that maximum gains from trade are realized. If the price that a consumer is willing to pay is higher than the marginal cost of production, then there is surplus to be gained from the trade. As a result, any time the price is set *higher* than marginal cost, there are efficient trades that are lost.

Firms in perfectly competitive markets set price equal to marginal cost, so perfectly competitive markets are allocatively efficient. Consumers earn consumer surplus since some consumers would have been willing to pay more than marginal cost. In fact, notice that *all* of the surplus accrues to consumers in perfectly competitive markets; firms earn no profits in long run equilibrium. See figure 7.

A monopoly firm is not interested in total surplus, but only in the part of the surplus that it can appropriate as profit to itself. Consequently, the monopoly raises price above marginal cost in order to generate a profit (maximizing profit where $MR = MC$). However, in doing so, the monopoly excludes some buyers who were willing to pay more than marginal cost. The lost surplus from these buyers is deadweight loss. Monopoly is allocatively inefficient – price is set higher than marginal cost, generating deadweight loss. Again, see figure 7.

Notice that, for the monopolist, the next unit beyond Q_M would have been efficient – the consumer is willing to pay more than the marginal cost of production. If the monopolist could sell to *only* this consumer for the lower price, it would do so. The problem is that the monopolist does not want to cut the price for *all* consumers. So, in some sense, the source of the allocative inefficiency problem with monopolists is the assumption that a single price is charged to all customers.

In contrast, suppose that the monopoly can *perfectly price discriminate* – meaning that each customer is charged a price exactly equal to his maximum willingness to pay. In this case, the firm will sell to any customer willing to pay more than the marginal cost of production, charging each customer a different price. The firm will even sell to customers willing to pay only slightly above marginal cost. It certainly wouldn't want to lower the price for *everyone* down to this level, but in the case of perfect price discrimination, it need not do so.

As shown in figure 7, a perfect price discriminator sells the same level of output as a perfect competitor. However, since each customer is charged a price equal to his willingness to pay, all the surplus goes to the producer. The point is that perfectly price discriminating monopoly is allocatively efficient, *but* all of the surplus accrues to the producer, in contrast to the competitive case where all of the surplus accrues to the consumer.

Perfect price discrimination resolves the allocative inefficiency problem that arises with monopoly, but it's important to be aware that all of the gains from trade go to the producers in this case.

Perfect price discrimination is impossible in practice since a seller would have to be psychic to determine exactly what each consumer is willing to pay; also he has to be able to prevent resale from low-price to high-price customers. However, there are all kinds of ways that firms can imperfectly price discriminate among different customers and charge different prices to different groups of customers. Although perfect price discrimination is infeasible in practice, it is important to understand as the limit case. The better firms get at price discriminating among different customers (i.e. as price discrimination gets closer to perfect),

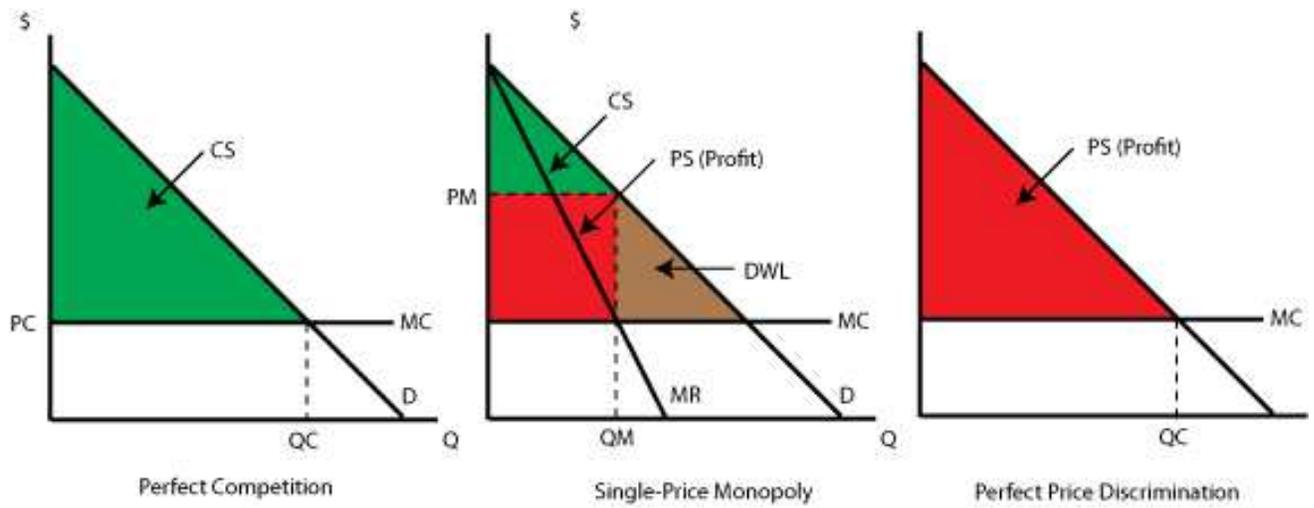


Figure 7: Surplus in perfect competition, single-price monopoly and perfectly price discriminating monopoly

allocative efficiency improves, but more of the surplus goes to producers rather than to consumers.