7.5 Profits and producer surplus

Let us return to our long-run supply function in Figure 30. Suppose that indeed the price is p". In that case, the firm produces q", as we have argued. Also, in this case the firm makes positive profits (remember, positive economic profits –i.e., extraordinary profits–, not accounting profits). Make sure you know how to graphically identify these profits as the area of the rectangle with vertices in the vertical axis at p" and $AC(q")$ and vertices in the horizontal axes 0 and q". That is the simplest way to identify these profits: $q(p - AC(q))$.

There is another way to identifying these profits that is (less direct but) more illuminating: the profit is the area above the marginal cost curve and below the horizontal line at p", and between 0 and q". Why is this so? Note that the revenues of the firm are simply $p" \times q"$, that is, the rectangle below the horizontal line at q"and the horizontal axis, and between 0 and q". The profit is the difference between this and the cost. But the cost is the area below the marginal cost curve and the horizontal axis, and between 0 and q"! Why?

Think of the marginal cost at some q as the (extra) cost of producing that unit q. So, if we begin with $MC(1)$, that is the extra cost of producing the first unit... beginning with producing 0 units (which costs 0 in the long run). Then, if we produce the second unit, we add extra cost $MC(2)$ to the cost, so that the cost of the first two units is $MC(1) + MC(2)$. If we continue this process until unit q", we have that the cost of producing q" units is ($q" \times AC(q)$ or $C(q")$ or) $MC(1) + MC(2) + ... + MC(q")$. That, if you make this computation adding tiny bits of output at a time instead of one unit at a time is... yes, the area below the marginal cost curve and the horizontal axis, and between 0 and q".

Now consider the short run, and look at Figure 31. At price p' the firm produces q' and makes... negative profits, since the price is below the average cost of producing the q' units.

However, if we compute the area of the rectangle between the price and the average *variable* cost, between 0 and q', we see a positive value. That value is what we call **producer surplus**.

What do we obtain if, as we did for the long run, we compute the revenues (again, the area of the rectangle $q' \times p'$) and then substract the area below the short-run marginal cost, between 0 and q'? What is $MC_{SR}(1) + MC_{SR}(2) + ... + MC_{SR}(q')$? As before, it is the (short-run) *extra* cost of producing the first q' units... except that this time this is not the same as $C(q)$! Why? Well, $MC(1)$ is the *extra* cost of producing the first unit, so the cost of producing one unit is the cost of producing 0 units plus $MC(1)$, as before. This time, however, the cost of producing 0 units is not 0, but the fixed cost! That is, $C_{SR}(q') = MC_{SR}(1) + MC_{SR}(2) + ... + MC_{SR}(q') + FC$. Or, in other words, $MC_{SR}(1) + MC_{SR}(2) + ... + MC_{SR}(q') = C_{SR}(q') - FC$. That is, $VC(q')$!

Therefore, the producer surplus is simply the revenues minus the variable costs $p \times q - VC(q)$. Or, put it differently, profits *plus* fixed costs, since

$$\pi = p \times q - C(q) = p \times q - (FC + VC(q)).$$

In the long run, since there is no fixed cost, producer surplus and profit coincide. In any case, both are (to all effects) equivalent ways of measuring the value for a firm of participating in the market. That is, the welfare measure that, for firms, corresponds to the welfare measure that for consumers we defined as the consumer surplus.

Remember that the consumer surplus was the area between the (inverse) demand function and horizontal line at the price, and between 0 and the quantity traded. Similarly, the producer surplus is the area between the horizontal line at the price and the (inverse) supply function (i.e., the marginal cost function), and also beteween 0 and the quantity traded.

7.6 Long run and free entry

Consider what happens if technology is not proprietary, but is in the public domain, and all the inputs are readily available for everyone that is willing to pay their prices (as we are assuming, since we are assuming that the input markets are competitive). Suppose that the price is p' in Figure 30. That is, a firm makes positive (and so extraordinary, above the market return of capital) profits. In that case, we would expect that more firms would consider using the technology, buying the necessary inputs, and enter the market by replicating what the firm in Figure 30 is doing. If entry is indeed free, then in the long run we would have firms entering as long as the price was like p'. That is, as long as the price was above the minimum value of the average cost function.

In that case, in the long run, the supply would be infinity (OK, very large, too large) for any price above that minimum. Or, put in other words, the supply would be flat at a price equal to that minimum of the average cost.

(The level of output at which that minimum is attained is called **minimum efficient scale**.) That is, with long run competition and free entry we would expect firms to produce at the output level that minimizes the per-unit cost of production.

7.7 Equilibrium in a competitive market

In a competitive market, every participant, buyer or seller, takes the price as given. So, who sets that price? Economics does not have a good answer for this question. Instead, we import the concept of **equilibrium** from physics, and postulate that, "somehow" we expect the system (the market) to find a resting place. That is, to find a state of the system –a price– at which, absent external forces, the moving parts of the system are at rest –do not change–. What we call an **equilibrium price**: one that no internal forces would disturb.

Consider what this means for our case: each consumer makes decisions as to how much to buy given the price, in total $D(p)$. Each firm makes decisions as to how much to produce, in total $S(p)$. All of then go to the market and... well, if the price happens to be p such that

$$D(p) = S(p) \tag{17}$$

all is good: everyone can fulfill their plan. But if $S(p) > D(p)$, then some firms are not going to be able to sell as much as they planned: a force inside the system will somehow change the state. Probably (but this is only a conjecture which we have not modelled!) the price will "somehow" go down, so that sellers want to sell less (supply is upward slopping, in general) and buyers want to buy more (demand is downward slopping, in general). On the opposite, if $S(p) < D(p)$ we expect the price to go up.

We use this idea of equilibrium to close our model with a prediction of what is going to happen in the (competitive) market: the price, p, will be one such that, when consumers and firms behave as we have postulated (which is summarized by $D(p)$ and $S(p)$ respectively), (17) is satisfied.

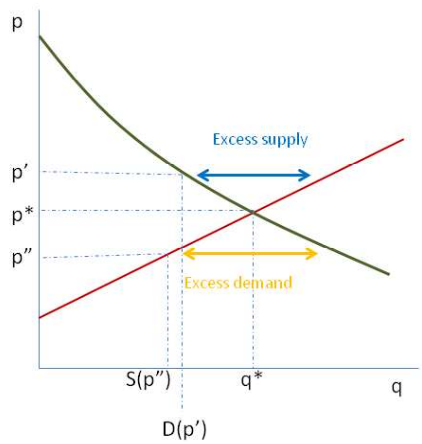

Figure 32

That price is illustrated in Figure 32. At price p^*, the decisions of buyers and sellers are compatible, and so feasible. Thus, the model predicts that this will be the price and that $q^* = D(p^*) = S(p^*)$ units of the good will be traded in that market. If the price was, say, p', then sellers would plan to sell $S(p')$ but would be able to sell only $D(p') < S(p')$, because only those many units would be purchased by consumers. Likewise, if the price was, say, p", then consumers would plan to buy $D(p")$ but they would be able to buy only $S(p")$ units because this is how much would be supplied by sellers. Thus, in the first case, we can talk of **excess supply**, and in the case of p" we can talk of **excess demand**.

7.8 Competitive equilibrium and efficiency.

Pause now and consider what the **market** *inverse* demand and supply functions measure. Recall that each of them is the horizontal sum of the corresponding individual demand and supply functions. Also, each individual inverse demand function represents how much the consumer is willing to pay for the corresponding unit, and each individual inverse supply function represents how much *extra* it costs the firm to produce the corresponding unit.

When we add two demand functions horizontally what we are doing is, in a sense, ordering (from highest to lowest) the willingness to pay for additional units of the good by *the two* consumers. It is as if you sit the two consumers together and ask each of them how much s/he is willing to pay for their first unit. You take the highest answer and that is the value $P^d(1)$. Then you ask the consumer who answered with the highest number to imagine that s/he had one unit, and repeat the question: now that this consumer has one unit and the other has none, how much is each of them willing to pay for one (additional) unit. The highest answer is the value of $P^d(2)$, etc.

Make sure you understand this with the following example. Table 6 displays the inverse demand function (willingness to pay) of two consumers. We are going to compute the aggregate demand function.

q	Ms A	Mr B
1	15	14
2	12	13
3	9	11
4	6	10
5	3	8
6	1	5

Table 6

That is, we are going to compute the quantity demanded by the market for each price. You can see that at a price above \$15 the demand is 0. At a price of \$15 (or below but above \$14) the demand is 1. At a price of \$14 (or below but above \$13), the demand is 2, etc. So we can represent (part of) this (market) demand as:

p	q
15	1
14	2
13	3
12	4
11	5
10	6
9	7

The inverse of this function, that is, the **market inverse demand** is this same table with only writing the q column on the left and the p column on the right. (For any price between two consecutive prices in the p column, the demand is the demand at the minimum of them, of course.)

What does it mean then that $P^d(5) = 11$, in this case? As you can see, it means that, if the consumers already have 4 units *distributed in the priority order given by their willingness to pay* (the first goes to Ms A, the second and third to Mr B, the fourth to Ms A), then the (highest) willingness to pay for the fifth unit is \$11 (in this case, by Mr B).

Try to convince you that something similar can be said for the horizontal sum of the supply functions: the inverse supply function is simply the "marginal cost of the industry", understanding by that the cost of increasing total output by one unit when q units are being produced in the cheapest way possible: assigning the production of each of these units according to the firm that needs to incur the lowest extra (marginal) cost to produce it.

Do we agree? If so, you realize that the market inverse demand function measures what we can term, in a very precise way, the market's (or society's) willingness to pay for additional units, much as the individual inverse demand function measures the individual's willingness to pay for additional units. That is, $P^d(q)$ is the value for society of an additional unit when q units are consumed. (Again, to be more precise when the good is divisible, the per-unit value of a tiny extra bit of the good.) Likewise, the market inverse supply function $P^s(q)$ measures the industry's cost of producing one more unit when q units are being produced.

Consequently, and just as in the case of one consumer, the (market) consumer surplus (when q units are consumed and the price is p) is the value of the area below the inverse demand function and above the price p, between 0 and q. Also, the (market) producer surplus (i.e., the profits plus fixed costs, if there are any) is simply the area below the price p and above the inverse supply curve, between 0 and q. And so, the **total surplus**, i.e., the sum of consumer and producer surplus, is simply the value of the area below the inverse demand function and above the inverse supply function, between 0 and q.

Under several conditions that include:

1.-absent any fairness criterion (i.e., if the distribution of income is not judged);

2.-the input prices –price of labor included– measure the real cost for society of producing those inputs;

3.-no third party (i.e., other than the consumers and producer involved) is affected by the consumption or production of this good (no **externalities**);

4.-the good is **private**, i.e., one unit of the good consumed by one consumer cannot be consumed by another;

the total surplus (which is measured in dollars: try to see this) can then be seen as a measure of the value that trading in this market creates for society. That is, as a measure of the **gains from trade** in this marke.

Note that total surplus only depends on how many units of the good are produced. In particular, it does not depend on the price at which they are traded. Indeed, an extra unit traded at a price p when q are being produced generates an extra consumer surplus of $P^d(q) - p$ and an extra producer surplus of $p - P^s(q)$. Thus, it generates a total surplus of

$$P^d(q) - p + p - P^s(q) = P^d(q) - P^s(q),$$

which is independent of the price p.(The price is money outlay for one member of society, the consumer, and money incoming for another, the firm.)

What is then the quantity q that maximizes the total surplus, this measure of value for society (under a few conditions, as we have seen)? This answer is straightforward, isn't it? We just have to go one unit at a time: is q the optimal one? Let us consider what additional surplus would one more unit generate: $P^d(q) - P^s(q)$. If this is positive, it means that producing and trading that extra unit increases the total surplus. That is, q is not the one that makes total surplus maximum. If this is negative, it means that reducing production and consumption by one unit the total surplus is reduced by a negative amount. Reducing by a negative amount is, ... wait, yes: increasing! Thus, again, q is not the quantity that maximizes total surplus.

Thus, the q that maximizes total surplus is the one at which $P^d(q) = P^s(q)$! That is, the one that is traded in a competitive market, in equilibrium!

Suppose that, for whatever reason, the output traded in the market is not that q that maximizes total surplus, call it q^c (c for competitive outcome). This may be due to several reasons. For instance, market power (sellers are not price-takers). We may still compute the total surplus for that output (and remember that total surplus only depends on output, in particular not in prices). Figure 33 represents this output and the corresponding total surplus. Note that the maximum possible surplus is larger than the total surplus at output q' in the figure. In fact, the difference, that is, the total surplus that could be obtained if q^* units were traded but is not obtained because $q^* - q'$ of theese units are not, is the area of that sort of triangle that is shaded.

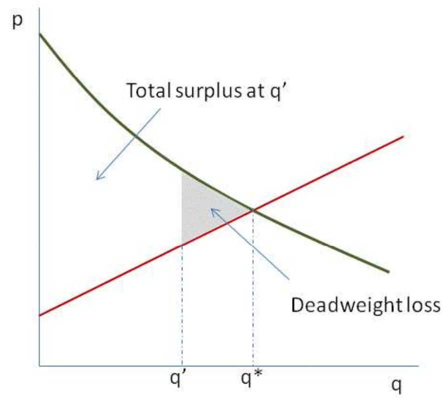

Figure 33

We call that "loss" of surplus **deadweight loss**. Note that surplus, and so the deadweight loss, is measured in dollars (or whatever currency prices are measured in). Thus, it may be compared with anything that is also measured in dollars.

A little thought exercise: why is item 3 among the conditions for the competitive equilibrium to be surplus-maximizing? That is, why if consuming or producing the good has effects on non-trading agents then the competitive output q^* may not be surplus-maximizing? As an example, imagine that the production of the good generates pollution that increases the cost that the city incurs in keeping the streets clean. Say, it raises this costs t dollars per unit produced. The city is not involved in trading the good, though. Can you conjecture (with a graph, perhaps) what would be the level of output that would maximize the value for the society (city) generated by trade in this market?

8 Monopoly

When we discussed perfect competition, we mentioned that such (behavioral) assumption is typically justified when there are many, similarly small firms in a market. This is not the most common market structure, of course. Often, only a few firms operate in a market, and so they will take into account that their output decisions (for example) affect the price. That is, the price is not independent of their actions. Or, equivalently, firms are not price-takers: firms have market power.

In this course, we are going to consider only the most extreme of such cases: **monopoly**. A monopoly is a market structure where only one seller operates. The reasons for this, that is, the reasons why no other firm operates in the market, may be several. Increasing returns to scale (for scales comparable to the "size" of market demand), patents,... are some of these reasons. But, given this, how would this sole producer behave?

There are two things that competitive and monopolistic firms have in common, as long as the monopolist is also competitive in the markets for inputs. (That is, as long as the monopolist is only one of may buyers of the inputs it uses.) Firstly, in both cases the objective for the firm is to make as high a profit as possible. Secondly, whatever output the firm decides to produce, maximizing profits requires that the output be obtained with the input combination that costs the least among the ones that allow the firm to produce that output. That is, monopolists and competitive firms alike will solve problem (11), which will determine their cost function, $C(q)$.

That is the good news: we can use all we have learnt in the section on costs when we analyze the monopoly. Thus, we only need to reconsider the choice of q. Here, though, things are different. As we mentioned, the monopolist affects the price with its output decisions. Indeed, if it decides to (produce and) sell q units, the highest price it can (and will) charge is given by the inverse demand function, $P^d(q)$. That is, with its decisions on output, the monopolist also **determines** the price. The monopolist is not a price-taker, but a price-setter!

That means, in turn, that the revenue that the monopolist obtains if it decides to produce q units is $P^d(q) \times q$ instead of some given number p times q.

Note the big difference between $P^d(q) \times q$ and $p \times q$. In principle, any firm, competitive or not, decides how much to produce and what price to ask. However, selling that output at that price is another issue: if consumers don't buy so much at that price then the firm cannot fulfill its plans. If the firm is competitive, it can sell NOTHING, unless it sets a price no higher than what the rest of the market is asking, p. And at that price p, it can sell anything it wants. However, if the firm is a monopoly, it can sell a quantity q if it sets a price no higher than $P^d(q)$, the market's "willingness to pay" for that last unit q. (If the monopolist wants to q units, it will not charge a lower price that this maximum price that allows it to do so, of course.)

Thus, since the monopolist is still interested in maximizing the difference between its revenue and its cost (the latter being just $C(q)$), the optimal decision for the monopolist will be the solution to

$$\max_q \left\{ P^d(q) \times q - C(q) \right\}. \tag{18}$$

Compare this problem with the problem that the competitive firm solves. When producing an additional unit of output, the firm increases its costs in $MC(q)$ in both cases, and the additional unit is sold at the market price in both cases. However, in order to sell that extra unit, the monopolist will need to reduce the price it charges. Thus, the q units that it was selling before increasing output will fetch a lower price. Indeed, obtaining the first order condition for (18), we see these three effects:

$$P^d(q) - MC(q) + \frac{dP^d(q)}{dq} \times q = 0. \tag{19}$$

The first two terms are the same as in the competitive firm´s problem (when we evaluate p at the market price $P^d(q)$). However, the third term is new: it measures the change in the price, $\frac{dP^d(q)}{dq}$ (negative, since the demand is downward slopping), times the number of units sold. Remember: the competitive firm can increase input if it wants without the need to reduce the price.

Note that this third effect, again, absent for a competitive firm, is a **disincentive** to increase output. That is, we expect the output of the monopoly to be too low as compared to the output if the industry was competitive. Indeed, we can write (19) as

$$P^d(q) - MC(q) = -\frac{dP^d(q)}{dq} \times q, \tag{20}$$

that is, the equilibrium price will not be equal to the marginal cost, but higher than the marginal cost! (Note again that $\frac{dP^d(q)}{dq} < 0$, and so the right hand side of the above inequality is positive.)

The monopolist, as the competitive firm, produces up to the point at which the marginal cost equals the **marginal revenue**. However, in this case the marginal revenue is not the price, but $P^d(q) + \frac{dP^d(q)}{dq} \times q$. The sencond term being negative, and $P^d(q)$ being the price resulting from the monopolist's decisions, this is lower than the price. In Figure 34 we have illustrated the solution to (19) and the marginal revenue for the monopolist.

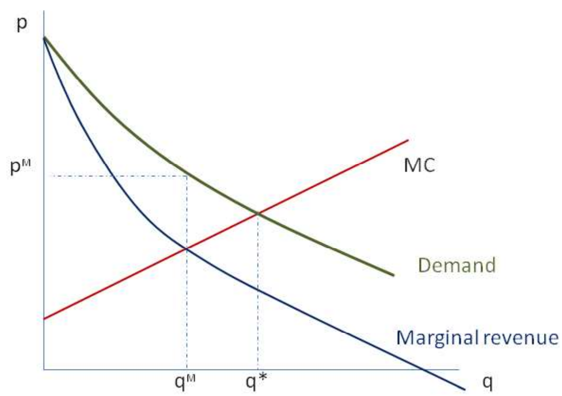

Figure 34

8.1 Demand elasticity at the monopoly solution; the Lerner index of market power

Consider (20), and divide both sides of the equation by $P^d(q)$. Then we can write that equation (remember: it is simply the f.o.c. of (18)) as

$$\frac{P^d(q) - MC(q)}{P^d(q)} = -\frac{dP^d(q)}{dq} \times \frac{q}{P^d(q)}. \tag{21}$$

The left hand side is the **mark-up** at which the monopolist sells its output. For the competitive firm, this mark-up was 0. We may equate market power with the ability of sustaining positive mark-ups. Indeed, this mark-up (normalized by the price, that is, as a percentage of the price) is the most popular index of a firms's market power, the **Lerner index**.

Now look at the right hand side. First note that (according to the inverse function theorem)

$$\frac{dP^d(q)}{dq} = \frac{1}{\frac{dD(p)}{dp}}$$

evaluated at $p = P^d(q)$, and so at $q = D(p)$. Therefore the right hand side of (21) is

$$-\frac{1}{\frac{dD(p)}{dp}\frac{p}{D(p)}} = \frac{1}{\varepsilon_D},$$

where ε_D is the price-elasticity of demand (at the monopoly solution). That is, at the monopoly solution,

$$\frac{P^d(q) - MC(q)}{P^d(q)} = \frac{1}{\varepsilon_D}.$$

A monopolist's market power is determined by the elasticity of the demand that if faces. If the elasticity of demand is low, then it will charge a large mark-up, and if the elasticity is high then the optimal mark-up will not be that large. (Note that the competitive firm faces an individual demand that is flat, that is, infinitely elastic, and as a result it charges a mark-up of 0.)

This is intuitive. If the elasticity is high, raising the price (above the marginal cost) induces a relatively large loss of sales, and so will be less attractive than if the elasticity is lower. (For the competitive firm, any such raise means loosing all sales!) Or, equivalently, if the demand is very elastic, in order to make the price raise above the marginal cost the monopolist needs to cut sales by a large amount, and so increasing prices (by reducing output) is less attractive than when the demand is more elastic.

Remark 4 *You can see that a consequence of this observation is that the monopolist will never choose a quantity at which the elasticity of demand is less than* 1. *That is, it will always produce at the elastic portion of the demand function: if* $\varepsilon_D < 1$ *at the monopoly quantity, then the left hand side of the equation above would need to be higher than* 1*, which is impossible since* $MC(q) \geq 0$. *Indeed, try to see that the marginal revenue is negative when* $\varepsilon_D < 1$.

Remark 5 *As you can see in Figure 34 and as follows from our discussion of the first order conditions for the monopoly problem, the monopolist will produce an output which is lower than if it was a price taker. That is, the output will be less than the output at which* $p = MC$*, which as we have argued in the previous section is the output at which the total surplus is maximized. Thus, the monopolist's market power* ***results in a deadweight loss****. Try to identify deadweight loss, consumer surplus, and profits (or producer surplus) in Figure 34.*

Remark 6 *We have assumed that the monopolist chooses* q *taking into account that any such quantity could only be sold at price* $P^d(q)$*. Alternatively, we could postulate that the monopolist chooses the price it will charge,* p*, taking into account that then it will be able to sell* $D(p)$ *units. Convince yourself that both "modelling" choices lead to the same prediction. You can also write our problem (18) if we choose that way of modelling, and convince yourself that indeed, the condition that determines the optimal output is the same.*

Remark 7 *Have you noticed that we have not talked about the monopolist's* ***supply*** *function? Well, that was for a good reason: such thing does not exist. Remember, the supply function measures what quantity a firm/all firms would sell for any* ***given*** *price. The monopolist sets the price (at the same time that it sets quantity). Thus, there is no such mapping simply because nobody* ***gives*** *prices to the monopolist (or to any firm with market power, for that matter).*