I. Concepts of Competition

Whether a firm can be regarded as competitive depends on several factors, the most important of which are:

- The number of firms in the industry. As the number of firms increases, the effect of any one firm on the price and quantity in the market declines.
- The degree of rivalry. If firms are vigorously trying to claim market share at each other’s expense, each firm will be more constrained in its choices.
- The degree of homogeneity of the product. Differences in quality or other properties mean that the products of different firms are not perfect substitutes for each other, and that means customers will absorb some price differences among firms.
- The ease of entry and exit. If new producers can easily enter and exit the market, existing firms may behave as though there are more firms than there appear to be, because there are more potential competitors.

Entry into a market can be deterred by barriers to entry. Important barriers to entry include:

(a) High start-up costs. These may cause it to take a very long time for new firms to enter the market, during which time the market is less competitive than it otherwise would be.
(b) Brand loyalty. If people are reluctant to consider new alternatives, the established firms in an industry face less of a threat from new competitors.
(c) Government restrictions. These restrict the ability of competitors to contest the market. Examples include licensing of electricians, plumbers, doctors, etc.; patents; protectionist tariffs and quotas; and absolute limits on the number of suppliers, such as NYC’s taxicab medallion system.

We ask whether a market satisfies the above conditions, and to what degree, in order to decide what market structure best describes it.

II. Firm Demand Under Perfect Competition

When a market has a large number of firms, free entry and exit, and a relatively homogeneous product, it can generally be modeled as perfectly competitive, or PC for short.

The key condition for a competitive market is as follows: every firm (and every consumer) must take the price of the good, determined on the market, as given. That is, each firm must regard the price as being unaffected by its own choice of quantity. Increasing its output will not drive down the price, nor will cutting back its quantity raise the price. We call this the assumption of price-taking.
The assumption of price-taking behavior can be represented graphically by having the firm face a horizontal demand curve. It will be horizontal at the market price, established by supply and demand on the market as a whole. Note that with a horizontal demand curve, all levels of output from this one firm correspond to the same price, p*.

You can think of the firm’s demand curve as having a very, very slight downward slope to see that if the firm raises its price even slightly, all of its customers immediately leave and go to other firms. This indicates the very high degree of competition in the market.

The practical effect of this assumption is that a firm can always sell another unit at the market price p* without affecting the price of its own previous units. As a result, when the firm decides to raise its quantity, there is only one effect: the “people buy more” effect. There is no “people pay less” effect as in the general model of profit-maximization. When the firm sells one more unit of the good, the added (marginal) revenue is just the price of the good, p*. As a result, the PC firm’s MR is exactly the same as its firm demand curve: MR = p*.

Note that this is just a special case of the general model of profit-maximization. Recall that we can find MR by taking the firm demand curve and doubling its slope. But in this case, the slope of the demand is zero, so the slope of the MR is also zero.

### III. Profit Maximization Under PC

As usual, the firm sets MC = MR. But now, MR = p*, so the firm sets MC = p*. In other words, as long as the price of a unit exceeds the cost of producing it, produce the unit.

\[
\begin{align*}
4 &= 2 + .25q \\
2 &= .25q \\
q^* &= 8
\end{align*}
\]

Example: Suppose MC = q^2 – 8q + 17, and p* = 10. Then
\[
\begin{align*}
10 &= q^2 – 8q + 17 \\
0 &= q^2 – 8q + 7 \\
0 &= (q – 7)(q – 1)
\end{align*}
\]
The correct answer for this example is \( q^* = 7 \), not \( q = 1 \). Why did we get two answers, and how did we choose? It’s because this MC is a parabola, and it crosses the firm demand curve twice – once from above, and once from below. But it turns out that the lower quantity doesn’t maximize profits; in fact, it minimizes profits. Notice that increasing quantity above that point means getting higher MR than MC, so it makes sense to do so.

Just as in the general theory of profit-maximization, we need to use ATC to find out the firm's actual profits or losses.

IV. The Whole Picture (in the Short Run)

The complete representation of a PC market has two side-by-side graphs, market on the left and a typical firm on the right.

In this SR equilibrium, we have the firm making a profit. We could also have drawn it with the firm making a loss. In addition, there is no particular reason why all PC firms
must be identical in the SR, because while they face the same technology, they may have made different capital choices and therefore be on different SRATC curves.

If demand increases or decreases, this will affect the choices and profits of all the firms in the market. Note that the increase in price tends to increase profits or decrease losses. Obviously, the reverse would be true of a decrease in price.

V. Perfect Competition in the Long Run

In the long run, entry and exit become possible. Why? Because potential firms can buy fixed inputs and become actual firms. And existing firms can sell off or stop renting their fixed inputs and go out of business.

Firms will choose to enter the industry if the existing firms in the industry are making economic profits. The profits are an incentive to enter. Now, remember that the market supply curve is just the summation of all the individual firms’ supply curves. If we add more firms, therefore, the market supply must shift to the right. But what effect does a rightshift of supply have? It drives down the price on the market, thereby reducing the profits of each firm.

Now the firms are making profits, but smaller profits than before. But if there are still economic profits being made, more firms will enter. This must continue until there are no economic profits. What has to be true when profits equal zero?

\[ TR = TC \]
\[ p^*q = q \times ATC \]
\[ p^* = ATC \]

So entry finally stops when firms are producing at their lowest average total cost. Here is a diagram of the final, long-run equilibrium under perfect competition:
(The supply curve is rather far to the right in the picture above, because there were a lot of profits that had to be eliminated. That’s not the important thing about the picture. What’s important is that in the end there are zero profits.)

What if typical firm is making losses? Then the reverse process will take place. Firms will exit the market, causing a leftshift of market supply, causing a rise in market price, causing a reduction of losses. This continues until losses are zero.

Thus, LR competitive equilibrium consists of two conditions:

- $p^* = MC$
- $p^* = \text{minimum ATC}$

The first condition is caused purely by profit maximization, and it’s true in both the SR and the LR. The second condition, however, is caused by entry and exit in the LR. It won’t necessarily be true in the SR.

These two conditions have important efficiency implications. Marginal-cost pricing ($p^* = MC$) means that consumers who buy the product face the true opportunity cost of their choices. They will only buy the good if the value to them is greater than the price, which represents the value of the resources that went into making the product. Minimum-average-cost pricing ($p^* = \text{minimum ATC}$) means that the product is being made at the lowest average cost possible, so that no resources are being wasted in its production.

The conclusion that firms make zero profit in the LR may seem odd, given the profits that many firms earn in reality. What could explain the difference between theory and reality? (1) Reality may differ from the perfectly competitive model, and to that extent economic profits can be made. But also, (2) the profits we generally hear about are accounting profits, not economic profits. To find out whether these “profitable” firms are really making economic profits, we’d need more information about their implicit costs. Finally, (3) we may be observing short-run profits, not long-run profits.
VI. Perfect Competition and Plant Size

It turns out that the perfectly competitive firm produces not just at the minimum of its SRATC, but also its LRATC. Why? Because any PC firm not at its minimum LRATC will, in the LR, change its input combination to take advantage of lower average costs. If firms are able to make positive profits by moving outward on the LRATC curve, those profits will attract entrants into the industry in the usual fashion. So by the same arguments as before, profits will eventually dissipate to zero. The price must be at the bottom of the LRATC, not just the SRATC.

VII. Perfect Competition and the LR Supply Curve

As we have seen, changes in demand in a PC market create profits and losses for firms. In the SR, this has no effect on the supply curve; but in the LR, firms enter for profits and leave to escape losses, leading to supply curve shifts. We want to use this information to derive a LR supply curve. A LR supply curve, just like a SR supply curve, shows the total quantity that will be supplied in a market at different prices; but unlike the SR supply curve, it shows the quantity supplied after all long-term changes, including entry and exit of firms, have been taken into account.

Aside: In the basic supply-and-demand framework, notice that we can use demand curves and equilibrium points to “trace out” the supply curve. If you look at three different demand curves, and then mark the equilibrium point on each one, you can connect the equilibrium points to find where the supply curve must be.

Now we’re going to use the same basic technique to trace out the LR supply curve. We can do this by changing demand, and then finding the equilibrium points after allowing LR adjustments, including entry and exit. Start with an initial (short-run) supply and demand. If we are in long-run equilibrium, profits are zero. Now, let demand shift to the right. In the short-run, price rises a lot. But the higher price creates profits, and profits attract entry in the long run. So eventually supply shifts to the right as well, pushing price back down (though possibly not as low as it was before). Once profits are back to zero again, you’re in a new long-run equilibrium. Do this all again to find a third long-run equilibrium, and then connect the dots to get the long-run supply curve.
The interpretation of the LR supply curve is pretty much the same as the SR supply curve: it shows the willingness of producers to sell at each price. But the LR supply curve measures this willingness in the broadest sense, including all firms that might potentially supply this product.

Notice that the LR supply curve is flatter than the SR supply curve. This must be so, since the LR supply curve takes into account the quantity responses of all firms, not just the ones currently in the market, but potential firms as well. It is even possible that the LR supply curve can be downward-sloping. Why?

Consider what must happen if entry and exit do not affect the cost curves of individual firms. Then after all adjustment to a change in demand has taken place, the market price must have returned to the lowest point on the LRATC, which is exactly where it was before. So in this case, the LR supply curve must be horizontal (instead of upward-sloping as in the picture above). We call this a constant-cost industry. This is most likely to be the case when the industry in question constitutes only a small portion of the demand for its inputs.

If the industry in question has a large impact on the markets for its inputs, then the LR supply curve may slope upward or downward. If the effect of entry into the industry is to bid up the price of inputs, so that a firm’s cost curves rise as a result of the entry of new firms, then the market price after adjustment will be higher than it was before. In this case, the LR supply curve must be upward-sloping as in the picture above; this is called an increasing-cost industry, which results from external diseconomies. On the other hand, if entry into the industry creates a greater demand for inputs that allows those inputs to be produced through mass production techniques (i.e., at lower average cost), then the industry can benefit from lower costs of production. In this case, the LR supply curve is downward-sloping. This is called a decreasing-cost industry, which results from external economies.