## 1 Behind the demand and supply curves

## 2 Monopoly

Main ideas so far:
Supply and Demand Model
Analyzing the model - positive - e.g. equilibrium, shocks
Welfare/Efficiency- normative
Policies/intervention - taxes, price controls, minimum wages
Now we are familiar with the demand/supply model we want to deepen our understanding by looking underneath the demand and supply curves to the underlying theories of firms and consumers.

Behind Supply and Demand

## Demand

For consumers utility theory is what we use to model consumer behavior. The most important implication of this model is that it generates a downward sloping demand curve

Our theory of the consumer is UTILITY THEORY. We assume consumers allocate their income to maximize their utility. Most important implication of this theory is the EQUIMARGINAL PRINCIPLE. For any goods, good 1 and good 2, the following equation is satisfied.

$$
\mathrm{MU}_{1} / \mathrm{p}_{1}=\mathrm{MU}_{2} / \mathrm{p}_{2}
$$

Can think of this in terms of BANG FOR THE BUCK - the consumer allocates each dollar in such a way that at the margin it is generating the same "bang" - in terms of utility which is what the consumer cares about.

If the consumer did not do so it would be possible to reallocated spending an improve utility - say $p_{1}=p_{2}=1$ and $M U_{1}$ is larger than $M U_{2}$ then shifting spending from good 2 to good one will generate more utility $\left(\mathrm{MU}_{1}\right)$ than it gives up $\left(\mathrm{MU}_{2}\right)$ - the consumer will therefore want to do this, and indeed will keep doing so until the Equimarginal condition holds. Shifting consumption brings the bang-per-buckinto line because of DIMINISHING MARGINAL UTILITY - As you consume more of something you get less and less satisfaction out of the marginal unit.

Think about the effect of price on demand:
Increasing $p_{1}$ implies $M U_{1}$ also has to increase to keep the equimarginal condition satisfied. But by diminishing marginal utility $\mathrm{MU}_{1}$ INCREASES when consumption of good 1 DECREASES - this is another way of saying demand falls as price increases, so we have a downward sloping demand curve.

## Supply

Now let's think about the behavior of firms that underlies the supply curve. For firms we also have a theory that underlies the supply curve and explains why it is upward sloping.

The primitive concept when dealing with firms is their COST CURVE. We take the cost curve as a given;

There are various different concepts of costs, all inter-related. All of them we think of relating some notion of costs to the level of production Q - so implicitly the decision variable of the firm is always the quantity of production

## DEFINITIONS:

TOTAL COST: $\mathrm{C}(\mathrm{Q})-$ This is the dollar cost of producing total output Q
In some sense this is the primitive cost concept.

MARGINAL COST MC(Q)- This is the dollar cost of producing the marginal unit at output Q - i.e when output is already Q what is the cost of the next unit (either positive or negative)

NOTE - THERE IS SOME SUBTLETY ABOUT DEFINING MARGINAL COST SINCE OFTEN THE COST OF THE NEXT UNIT UP IS LESS THAN THE COST OF THE NEXT UNIT DOWN - IN THIS CASE WE CONSIDER BOTH CHANGES AND DEFINE THE MARGINAL COST AS THE MEAN OF THE 2 MEASURES.

AVERAGE COST $\mathrm{AC}(\mathrm{Q})$ - This is the dollar cost per unit, or unit cost. Defined as $\mathrm{C}(\mathrm{Q}) / \mathrm{Q}$

2 other terms:
FIXED COST - component of total cost that is independent of Q
VARIABLE COST - component of total cost that varies with $Q$

## Example 1:

The following table is generated from the cost function $C(Q)=5+Q^{A} 2$
One way to represent this information is to set out the fuction in a table and calculate the various cost concepts.

| Q | $\mathrm{C}(\mathrm{Q})$ | $\mathrm{MC}(\mathrm{Q})$ | $\mathrm{AC}(\mathrm{Q})$ |
| :--- | :--- | :--- | :--- |
| 0 | 6 | $1-$ Here we only <br> have upward <br> adjustment margin | NA - Q =O |
| 1 | 7 | 2 | 7 |
| 2 | 10 | 4 | 5 |
| 3 | 15 | 6 | 3 |
| 4 | 20 | 8 | 5 |
| 5 | 31 | 10 | 6.2 |

To get the MC we just look at the differences (up and down) of costs for the production of the next unit.

## Example 2:

Another way to handle the information is to work with the functions themselves
Suppose the cost function is $C(Q)=10+Q+Q^{A} 2$

Here what can we say:
i) $\quad \mathrm{AC}=10 / \mathrm{Q}+1+\mathrm{Q}$

Calculating the marginal cost - those of you that know calculus bear with me-you'll see the answer we get is just the derivative of the cost function

Think of 1 more unit:
$\mathrm{C}(\mathrm{Q}+1)-\mathrm{C}(\mathrm{Q})$
$=10+Q+1+(Q+1)^{\wedge} 2-10-Q-Q^{\wedge} 2$

```
=10+Q+1+Q^2+2Q+1-10-Q-Q^2
=1+2Q+1=2Q+2
C(Q)-C(Q-1)
=10+Q+Q^2-10-(Q-1)-(Q-1)^2
=10+Q+\mp@subsup{Q}{}{\wedge}2-10-Q+1-\mp@subsup{Q}{}{\wedge}2+2Q-1
=2Q
```

Thus taking the average of these two measures we find that:
$\mathrm{MC}(\mathrm{Q})=1 / 2(2 \mathrm{Q}+2+2 \mathrm{Q})=2 \mathrm{Q}+1$
The MC is increasing which means that each unit is getting more and more expensive
Many of you find the idea of increasing marginal costs counter-intuitive as you have in mind the idea of scale economies - this shows up in the AC


We see that AC is first declining and then increasing. Why is this?
When AC is declining we say we have INCREASING RETURNS TO SCALE (IRTS) which means that as you scale the process up it becomes more efficient. In this case the increasing returns to scale are related to the FIXED COST. With higher levels of production this fixed cost gets spread over more units and so becomes a smaller fraction of the unit price.

If this were the only thing going on then we would have unlimited scale economies bigger always better. In practice we think scale economies are limited and this is due to the other component of the cost function. The variable costs are increasing in Q.- this introduces a component of DECREASING RETURNS TO SCALE (DRTS)- as you scale up the process it becomes less efficient. Why might this be?
i) Might be working with a fixed plant capacity - even if you employ more workers you would really need to invest more capital to scale up the process effectively.
ii) Other factors such as managerial input might also not scale up effectively larger organizations are harder to manage.
iii) For some industries increasing input prices might be a problem - as firm expands these prices rise raising the costs of production.

Thus in general when thinking about AC we have a tradeoff between IRTS and DRTS - at first costs decline due to the fixed cost effect and then start to rise due to the rising MC

Another important point is that MC INTERSECTS AC AT THE MINIMUM AC. Why?

## Because:

i) $\quad \mathrm{MC}>\mathrm{AC}$ implies that if you add one more unit it costs above the average so it is going to pull up the AC for all units - thus AC rising
ii) $\mathrm{MC}<\mathrm{AC}$ implies that if you add a unit its cost is below average so AC is falling.

If AC is falling to the left and rising to the right, it must be exactly at its minimum when $\mathrm{MC}=\mathrm{AC}$ - we will come back to this when we study equilibrium

## Competitive Supply - Theory of the Supply Curve

Armed with the cost curve we can go and think about what a competitive firm is going to do:

The basic assumption we make in dealing with firms is that there goal is to MAXIMIZE ECONOMIC PROFIT. This is not always the best assumption, but it is a good place to start.

It's important to bear in mind that even when we study different types of behavior by firms this is a consequence of putting them in different economic environments - the motive for that behavior is always the same - maximize profits - it is just that it manifests itself in different ways depending on the competitive environment.

DEFINITION: Profit=Revenue-Costs
This is our basic definition.
We think of the decision as a choice of quantity of output Q - This is why we defined the cost function of the firm in terms of Q .

At the same time we can think of revenues in terms of Q :

## Revenue $=P Q$

How do we maximize profits. As always we think about this in terms of decisions AT THE MARGIN.

Think about the firm gradually increasing Q . For the first unit if the revenue it brings in exceeds its cost of production, that makes a profit for the firm and so the firm produces that unit.

In jargon $\mathrm{MR}>\mathrm{MC}$ so the marginal profit is positive.
Think about the next unit - if MR>MC for this unit the firm produces this as well - note that both MR and MC can be changing as we increase Q - recall DRTS implies that marginal costs are usually increasing. And marginal revenue will be falling if the price is falling (more on that later). Thus typically as you increase output the marginal profit on each unit falls.

When does the firm reach its maximium? Precisely at the point when...

## PROFIT MAXIMIZING CONDITION: MR=MC

This is the fundamental profit max condition for the firm. Basically it says that if the profit on the marginal unit is zero, that's when to stop - all the other units made a profit, and any further units would make a loss.

We can think of this graphically if you find it useful.
For a competitive firm what is the MR curve? Competitive firm faces a given price, so $\mathrm{MR}=\mathrm{p}$. Each unit is sold at the same price p

What about MC - we discussed this above - this is an increasing curve
Sketch the two curves one under the other


If we sketch the marginal profit we see that it is increasing up to $M C=p$ and then starts declining

Thinking about this for a moment what have we derived?

For a competitive firm this is precisely the SUPPLY CURVE. We have figured out for every price p how to determine the quantity that the firm wants to supply. Thus the MC curve is just the supply curve - now we can see the precise way in which the supply curve is related to the costs of the firm.

## Profits and Entry

Discuss how the AC determines the entry decision and in long run eqm AC is minimized this is why it is so important that MC cuts AC at minimum since at this point both $\mathrm{p}=\mathrm{mc}$ and minAC are possible together.


Similarities:
i) still describe firm by its cost function
ii) monopolist still wants to maximize profits by setting $\mathrm{MR}=\mathrm{MC}$

Differences:
i) no longer have a supply curve - this is the relationship between price and quantity, but the monopolist chooses quantity taking into account that the price depends on his decision
ii) it is no longer true that $\mathrm{MR}=\mathrm{p}$. Now when the monopolist thinks about the extra unit the price declines as he moves down the demand curve.

Let's see how this works with an example:

What are the primitives of the economic environment?
We take as given a demand curve which describes the consumer side of the market and a marginal cost curve which describes the production side of the market.

The monopolist has to choose a quantity of production, or equivalently since every Q implies a P on the demand curve a price.

Suppose that the demand curve is
$\mathrm{P}=15-2 \mathrm{Q}$ and the MC curve is Q


What is the optimal quantity for the monopolist. Just like in competition profits are maximized at the point where $\mathrm{MR}=\mathrm{MC}$

Here the main difficulty is calculating the MR. At each quantity Q what happens to R if we increase Q by 1 unit.

Here we have 2 counteracting effects:
i) extra unit brings in more revenue
ii) price falls so that revenue falls on all the other units

This implies that $\mathrm{MR}<\mathrm{D}$ - if price did not fall then MR at every point on the demand curve would just be the current price. Since the price falls MR is always strictly below.

Let's work it out explicitly:
$\mathrm{R}=\mathrm{PQ}=(15-2 \mathrm{Q}) \mathrm{Q}=15 \mathrm{Q}-2 \mathrm{Q}^{\wedge} 2$
We can calculate MR just as we did for MC
Consider adding a unit:
$\mathrm{R}(\mathrm{Q}+1)-\mathrm{R}(\mathrm{Q})=15(\mathrm{Q}+1)-2(\mathrm{Q}+1)^{\wedge} 2-15 \mathrm{Q}-2 \mathrm{Q}^{\wedge} 2$
$=15-4 \mathrm{Q}-2$
$\mathrm{R}(\mathrm{Q})-\mathrm{R}(\mathrm{Q}-1)=15 \mathrm{Q}-2(\mathrm{Q})^{\wedge} 2-15(\mathrm{Q}-1)+2(\mathrm{Q}-1)^{\wedge} 2$
$=15-4 \mathrm{Q}+2$
Taking the average of these we get:
$M R=15-4 Q$
This is an example of the TWICE AS STEEP RULE which works for linear demand curves - to find the MR curve just multiply the slope of the curve by 2.

With the MR curve we can now solve the monopoly problem:
$\mathrm{MR}=\mathrm{MC}$
$15-4 \mathrm{Q}=\mathrm{Q}$
Q=3 $=5 \mathrm{Q}$
Pres given $G_{y}$ demand $\Rightarrow r=15-4 Q_{y}=9$
Once we have the quantity we can find the price, which is an equivalent way of thinking about the question - we can think that the monopolist is just setting the price.

Can illustrate all the economics in the diagram below.

$$
\begin{aligned}
& \frac{\text { Monopoly }}{C S \quad P S} D W L \\
& A+B \quad C+D \quad E \\
& \text { Competition } \\
& \text { Set Supply }(M C)= \\
& \text { Demand } 15-2 Q=Q \Rightarrow Q=5 \\
& \Rightarrow \text { DWL disappears }
\end{aligned}
$$



What are the important properties of the solution:

## i) MONOPOLIST RESTRICTS QUANTITY - EXPLOITS MARKET POWER BY CREATING AN ARTIFICIAL SCARCITY WHICH DRIVES UP THE PRICE AND RAISES PROFITS <br> ii) MONOPOLIZED MARKET IS NOT EFFICIENT - THE ARTIFICIAL SCARCITY CREATES A LOSS RELATIVE TO COMPETITION

It's important to understand that the loss is fi-om the point of view of society as a whole monopoly is good for the producer, but there are some transactions that would create value but which are not carried out.

Note: Once we get away from perfect competition and allow firms to set prices there are many outcomes possible. Monopoly is just the beginning. Even thinking about the monopolist we can see that there are maybe other pricing strategies.

Example.

## Perfect Price Discrimination

Can see that the problem from the point of view of the monopolist is that has to charge a single price to consumers. If he could charge different prices to different consumers there would be no need to restrict quantity.

In the limit if you can charge every consumer the price they are willing to pay can carry out every efficient transaction. However monopolist gets all surplus!

In practice such PRICE DISCRIMINATION might be difficult for two reasons
i) don't always know which consumer is which, so don't know who to charge high prices to
ii) low value consumers can buy cheap and sell on to other consumers - this ARBITRAGE prevents price discrimination and is counterproductive for the firm.

Nevertheless many pricing strategies can be understood in this way:
i) tuition scholarships at university
ii) student discounts at movie theater, magazine subscriptions etc
iii) airline tickets depend on how far in advance they are purchased

