Congestion is an issue that affects many cities and towns in first world countries. A prime example, and the one I will focus on today, is London. London, like many other major cities, is a sprawling, densely populated area which is still constantly attracting more people due to the benefits of agglomeration (I will discuss agglomeration in another article in the future). This led to increased pressure on the transport network in London and eventually led to the traffic volume (cars per mile) exceeding the maximum capacity. What can be done to resolve the issue? Build more roads? Introduce a congestion tax? Limit the extent to which cities are allowed to expand and thereby limit the number of people in the area?

First, I want to discuss congestion itself and define it. In an area, congestion is when the traffic volume is higher than the maximum capacity. This essentially means the network is overloaded and traffic moves slower than it usually would, and this leads to more volume and so the traffic slows further.

Flow is another important term, which is speed multiplied by the travel time. From this we can see there are two measures of congestion: the speed cars are travelling at the travel time in comparison to the speed of cars travel at free-flow (when the network isn't overwhelmed) or the time it takes at free-flow. Both are essentially the same thing but expressed in two different ways. If the speed at free-flow is 30mph but cars are travelling at 28mph, then congestion is likely not the issue, but if they're travelling at 15mph then that would be congestion. And similarly, if a journey usually takes 40 minutes at free-flow but you take 45 minutes then it is likely not congestion, but if it takes you an hour then it is likely congestion. Now that we've established how we determine whether congestion is an issue, we can move onto discussing what causes congestion.

Congestion is caused by too many people trying to travel on the same network at the same time. In economic terms, we could rephrase this as the demand for travel exceeding the supply for travel. But what is the demand for travel? Do people take a bus to school because they enjoy the bus ride? Do people drive to work because they enjoy being stuck in traffic? No, of course not. People travel because it provides access to opportunities, whether that be education or employment or something else is irrelevant, people travel to access opportunity. This means that the demand for travel is derived from the demand for the opportunities offered by travel.

Moving onto looking at the issue from the supply side, we can view congestion as being the result of a market equilibrating process, and so increasing the supply of transport changes the equilibrium but it doesn't decrease congestion. This is because with increased supply, say perhaps a new high-speed train line, all that happens is more people travel and people travel further. But then why do policymakers continue to make new roads? It's not to reduce congestion, it's instead to change the equilibrium. Say a new road allows an extra 1000 people to reach a very productive area that allows more people to experience the benefits of agglomeration, that's still beneficial for the economy but it isn't necessarily the most effective way to reduce congestion. This is the law of peak hour expressway congestion (Downs 1962) which says that the peak hour congestion in urban areas rises to meet maximum capacity. Furthermore, by increasing supply you allow more people to access opportunities in an area, and so the demand increases, therefore the demand is induced as well.

But if increasing the supply of transport doesn't help with congestion, then why do governments continue to invest so much money into infrastructure? It’s because it leads to more people being able to access the benefits from agglomeration in big cities. Thus we see a repeating loop, as land becomes more accessible due to better transport links, it becomes more developed, more people live there, more transport is developed there to serve their needs and this repeats. We could perhaps use it to explain the North/South divide in England; what started as a small gap due to a declining manufacturing sector and developing tertiary sector is now a massive gap, caused by the cumulative causation in terms of infrastructure development.

So what can be done to resolve the issue of congestion if not increasing the supply of transport? One approach is to limit the extent to which cities can expand. This would essentially limit the number of people crammed into a small area and therefore limit the number of people willing and able to use the transport, thus limiting demand and thereby reducing congestion. However, the drawback of this is that it limits the number of people that can access the agglomeration benefits of the city. This can hinder economic growth in that region and for the economy as a whole, as well as reducing the earning potential of the people unable to access the city and thus leading to a fall in their standard of living.

Another possible approach is to introduce a congestion tax, as was done in London. This increases the cost of travelling via chargeable modes of transport and so at every given time and price, without accounting for the congestion charge, fewer people are willing and able to travel due to the congestion charge. This leads to a reduction in congestion. In London where there is a congestion charge, this has been very effective. However, a major flaw of it is that it is a regressive tax, it is a higher proportion of income for those earning less than those earning more since it is the same for all people. This leads to a rise in inequality, which is undesirable.