Macroeconomics and Microeconomics Notes: A Physicist's Perspective

Kyle J. Bunkers

September 13, 2020

Contents

Sources

The notes are based off of what I have learned from MRU (marginal revolution "university"[1]). I am using their introduction to microeconomics and some of their introduction to macroeconomics sequences. In addition, a series of Introduction to Advanced Macroeconomics lectures by Professor Burda from Humboldt-Universität zu Berlin available on YouTube provided a more advanced (as you might expect) analysis of macroeconomics for these notes. They are mostly an overview of what I took away and thought about and follow the sources pretty closely.

My last real look at this was 2020-08-21.

[1] I put university in quotes because it is not an actual, accredited university, but a series of lectures made available on [and for] the internet, by economists.

Chapter 1

Microeconomic Notes

Appropriately, we start with demand and supply. Demand is conventionally represented as a curve (often a straight line for simplicity of learning the concept) in a price vs quantity graph. Demand "increases" as price decreases, meaning that the quantity demanded increases as price decreases along this curve. However, a change in demand usually means shifting the curve, and a change in quantity demanded would mean moving along the demand curve itself, rather than shifting the entire demand curve. This is illustrated in Figure 1.1. Two other things should be noted. There can be a difference between an individual's demand curve and an aggregate demand curve (of a labor force). We expect that for both that the lower the price more demanded, hence a negative slope in the price vs quantity plane (hence we expect demand to be a monotonically decreasing function of quantity), but for aggregate demand there is a question of whether this is realistic since changes in demand in a country's entire economy affect the supply and economic resources of the country itself. This will be further investigated when we look at macroeconomic concepts. So-called Giffen goods and Veblen goods are possible objects that have portions of positive slope for demand, but these are so rare that one should not worry about them in a basic economics course.

1.1 Demand and Supply Curve Basics

The demand curve can be read as either the quantity demanded given a price, or as the lowest cost willing to be paid at a certain quantity given. The consumers on the upper left of the demand curve are willing to pay a lot for the good while those on the lower right are the ones barely willing to pay for the good. So the uses of the good for those on the upper left are considered "better" in the sense that there is no easy substitute and so people are getting the highest value from the good.

The supply curve has similar interpretations, but there are some caveats. We again have an individual versus an aggregate supply curve. Here an individual supply curve often does have a negative slope. If an individual is paid more, they may supply less since they only care about a certain amount of money. As we will discuss more in macroeconomics, the aggregate supply curve usually does not do such a thing. If people are paid more (higher price), more people will start doing the thing (more quantity). Figure 1.2 shows a typical (linear for simplicity) supply curve. Here an increase in supply moves the curve rightward and downward while a decrease in supply

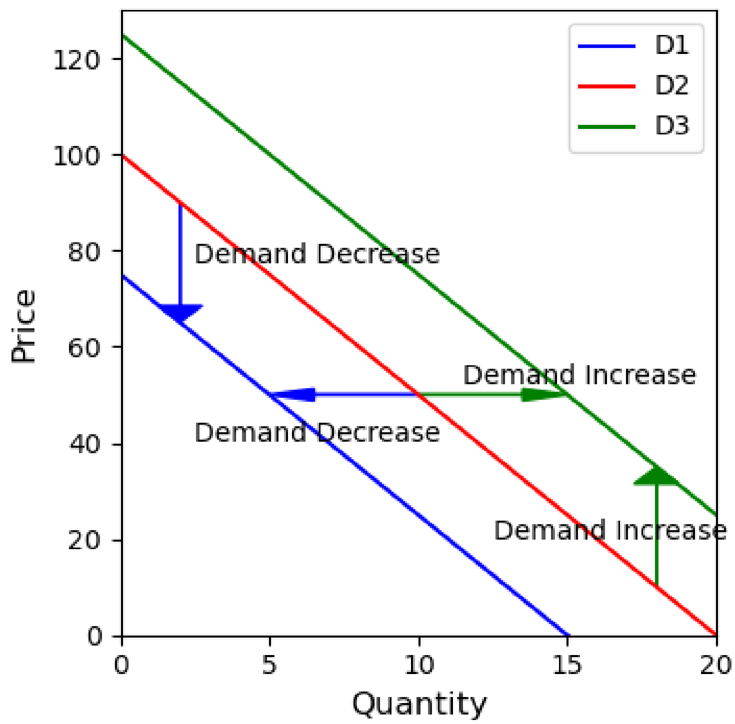

Figure 1.1: This shows a change in demand, which is a shift of the curve itself. We start with curve D2 and an increase in demand to D3 can be viewed as shifting the curve upward or to the right. A decrease in demand from D2 to D1 can be viewed as a shift downward or to the left.

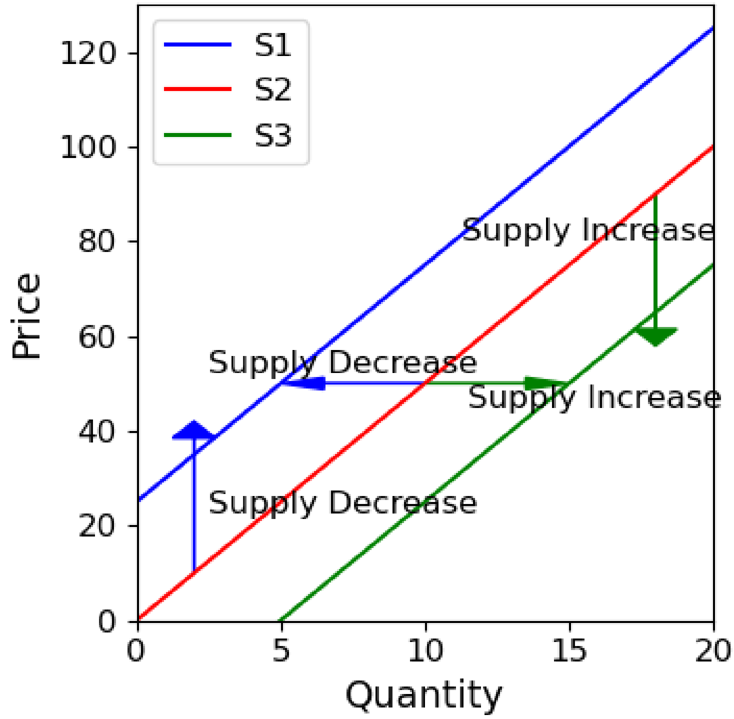

Figure 1.2: This shows a change in supply, which is a shift of the curve itself. We start with curve S2 and an increase in supply to S3 can be viewed as shifting the curve downward or to the right. A decrease in supply from S2 to S1 can be viewed as a shift upward or to the left.

moves the curve leftward and upward.

This of course leads to the "Law of Supply and Demand" which is essentially just saying that the actual quantity and supply of a good is determined by the intersection of the supply and demand curves. If we used $S2$ and $D2$ from Figures 1.1 and 1.2, we could find the actual price at the intersection as shown in Figure 1.3. This is usually called the equilibrium price and quantity since a market may not achieve this value immediately, but will after a period of time assuming that people will exploit any deviations from this equilibrium to make a profit (and hence shifting the quantity and demand to equilibrium values).

One thing to note is that price and quantity are somewhat arbitrary and so who is the "demander" and who is the "supplier" is a somewhat arbitrary designation. We typically think of price in a unit of currency, such as dollars, and quantity as some number or fraction of a good. But, of course, we could consider the good the "price" and the quantity, the amount of currency, in which case the roles of "demander" and "supplier" switch roles. This is just to say when there is a trade, there is equality in roles under this model.

We should talk about the equilibrium state. The idea of why it is a stable equilibrium is simple enough. We have the supply price for quantity supplied $S(Q_S)$ and demand price for quantity demanded $D(Q_D)$. Economic equilibrium occurs when $S = D$ and $Q_S = Q_D$. We can find an equilibrium point of the system, by finding $E(Q_S, Q_D) = S - D$. If $E > 0$ then people will not buy, and so S is forced to decrease. If $E < 0$ then sellers can earn more by increasing the price to match the buyer price. This means the only stable situation is when $E = 0$, so we restrict our

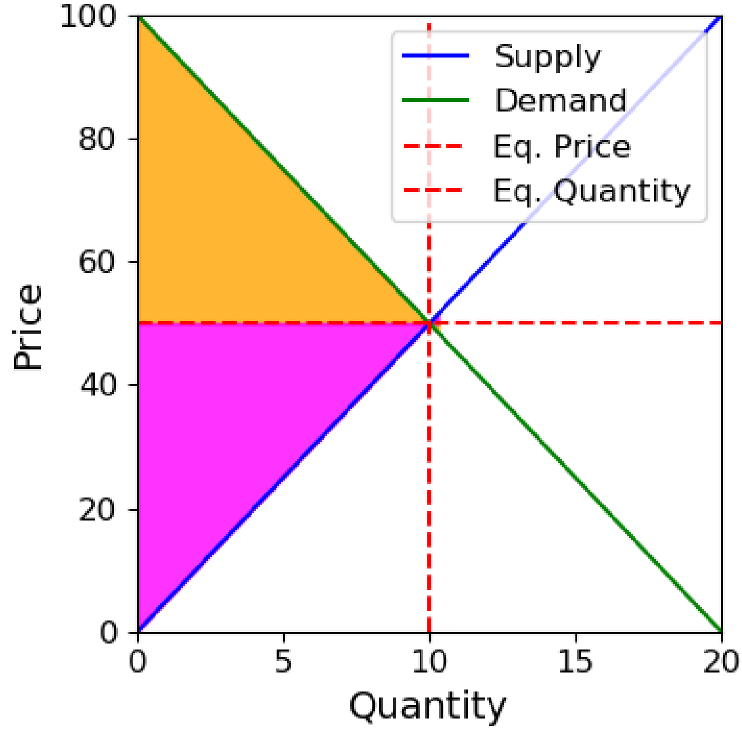

Figure 1.3: Here we find the equilibrium price and quantity at the intersection of the demand and supply curves. The consumer surplus is the orange filled in region and the producer surplus is the magenta filled region.

search along $E = 0 \Rightarrow S(Q_S) = D(Q_D)$. We then can use that if $Q_S > Q_D$ that there will be too much of the good supplied and suppliers will be unable to sell all of their stock, resulting in Q_S decreasing towards Q_D as buyers can either stock up or ask for lower prices. If $Q_D > Q_S$, then suppliers will recognize they can sell even more (at higher prices) and so will start producing more as buyers compete for more of the supply. Only when $Q_D = Q_S$ will things stay stable. Thus, we have a stable equilibrium state

Note that this stable equilibrium only exists under the conditions we used above. First, the market will recognize when prices are out of whack and adjust prices and quantity as described above, and second, the market will detect when quantity is out of whack and adjust. The reactions above are reasonable, but they are not always true, just generally. The idea is simply that buyers will buy more at a lower price, and sellers will sell more at a higher price. If market follows these principles, then the equilibrium will exist and all of the consequences we explore below will follow.

This also brings us to the concepts of consumer surplus and producer surplus. Consumer surplus is the difference between what the consumer is willing to pay and what they actually pay (for all the consumers engaging in trade). The producer surplus is the difference between the price the seller actually sells the good for and the price the seller would be willing to sell the good at (for all producers engaging in trade). In the case I have shown, the consumer and producer surpluses are equal, but this is not necessarily the case. The total of both surpluses is usually called economic surplus, total welfare or Marshallian surplus.

We can now come to the idea of elasticity. In my opinion, elasticity is the inverse of what would be the logical choice, since we use a price (P) vs quantity (Q) diagram rather than a quantity vs price diagram. Elasticity is defined as ϵ and given in a discrete case by

$$\epsilon = \frac{\Delta Q}{\Delta P}\frac{\langle P\rangle}{\langle Q\rangle} = \frac{Q_2 - Q_1}{P_2 - P_1}\frac{P_2 + P_1}{Q_2 + Q + 1} \tag{1.1.1}$$

$$\Delta q \equiv q_2 - q_1 \tag{1.1.2}$$

$$\langle q\rangle \equiv \frac{q_2 + q_1}{2} \tag{1.1.3}$$

where the subscripts indicate a particular slice in time or some dependent variable (so that Q_i and P_i refer to the same slice). In the continuum case, these clearly become derivatives and we get

$$\epsilon = \frac{\mathrm{d}\ln Q}{\mathrm{d}\ln P} = \frac{\mathrm{d}Q}{\mathrm{d}P}\frac{P}{Q} \tag{1.1.4}$$

The reason I say this is not the logical choice, is that typically we think of the derivative as the slope of a curve, and we think of $y = f(x)$ so that $f'(x) = \frac{\mathrm{d}f}{\mathrm{d}x}$ is the slope. But in economic, we have $y = P$ and $x = Q$, so the derivative would be $Q'(P) = \frac{\mathrm{d}Q}{\mathrm{d}P}$, but that is not what happened historically, and so we left with elasticity as it is now defined. If you are wondering about the extra P/Q in the definition, this is to normalize the elasticity. This way curves can be compared even when they have very different magnitudes for Q and P. (One can think of this as the percent change in Q over the percent change in P.)

Often the elasticity is referred to without regard to sign since we know the typical slopes of demand and supply curves. Given a demand or supply curve, the elasticity is easily found by the above formulas (1.1.1) and (1.1.4). When a curve has $|\epsilon| > 1$ we say that it is elastic. When a curve has $|\epsilon| = 1$ we say it is unit elastic, and when a curve has $|\epsilon| < 1$ we say that it is inelastic. On a price vs quantity graph, the more elastic the curve, the closer the curve is to a horizontal line and the more inelastic a curve, the closer it is to a vertical line. Elasticity is a somewhat strange term, but it comes from the idea that the more elastic a curve, the more easily substitutable the good is. So if one has elastic demand, a small price increase can lead to a large quantity demanded/supplied decrease. An inelastic curve has small changes in quantity demanded/supplied leading to large changes in price.

These are shown in Figure 1.4.

Inelastic goods are generally thought of as goods that have few good substitutions, while elastic goods have plenty of substitute goods. A specific good can be elastic in the long run, but inelastic in the short run. This is an economist's way of saying that for a short period of time you may have good a substitute but not for a long period of time. (For example, a car may be elastic short term [you could take the bus, walk, etc.], but in the long term quite inelastic [you need a car to get to work and have time for everything else]). Indeed, the opposite of a short-term inelastic and long-term elastic good is possible, too (If you are held captive in an airport, the food's prices may not matter to you all that much because you have few good substitutes, but in the long run [you get out of the airport] you have plenty of substitute food sources).

This also brings us to inferior goods, normal goods, substitute goods, and complement goods. An inferior good is a good whose demand increases when consumer income decreases. An example could be ramen noodles. Typically as a consumer's income rises they eat fewer noodles (and if their income decreases, they eat more noodles). A normal good is a good whose demand increases

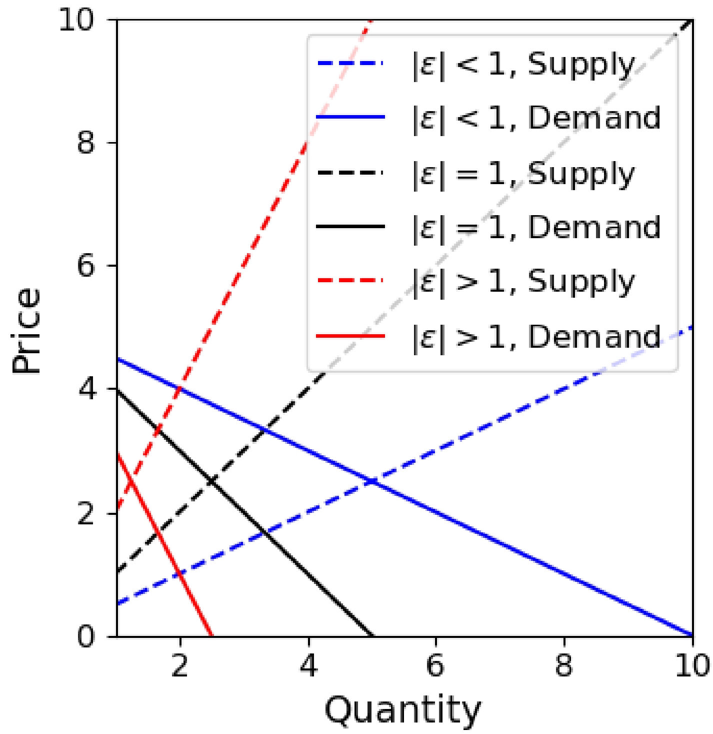

Figure 1.4: This shows the various supply/demand curves for various elasticities.

as a consumer's income increases. Finally, there are substitute goods and complement goods. A substitute good B (always relative to some other good, call good A) is called a substitute to A when the demand for B increases when A's price increases (or when A's price decreases, the demand for B decreases). A complement good C (relative to good A) is a good whose demand increases when the price of A decreases (or when A's price increases, the demand for B decreases). A substitute good for Coca-Cola is Pepsi. A complement good for paper is pencils (toothbrushes are complements of toothpaste conventionally, as well).

We can now consider taxes and subsidies. First let's look at taxes. In reality, a tax is simply one thing that changes the supply and demand curves, but because it is usually a simple one, we can easily compare it to the counterfactual case where there was no tax. Many people think that a tax will fall on whoever is taxed, but remember that who is a "buyer" and a "seller" is fairly arbitrary, and that a person can easily pass off the cost onto the other party either in quantity or price. That is, who is directly taxed does not determine who will bear most of the price.[1]

So what is a tax? Let's first look at a tax on the consumer (say a sales tax of some sort). Say it amounts to adding T to the price P of the good. We will use Figure 1.5 to explain. We start with our regular equilibrium where the dotted black lines cross. Because the consumer now has a price including tax the demand curve will shift downward. It will shift downward in price by the tax T. This is because if before the maximum price willing to pay for a certain quantity Q_P was P. But now the price for Q_P is $P + T$, so you won't pay that. Instead you'll pay for Q_T at price P. This Q_T was price $P - T$ before and so we see that the curve simply shifts downward by T. The

[1]Note that in real life, things are complicated, but if we let ourselves go to equilibrium (a long term in some sense), then these things become true.

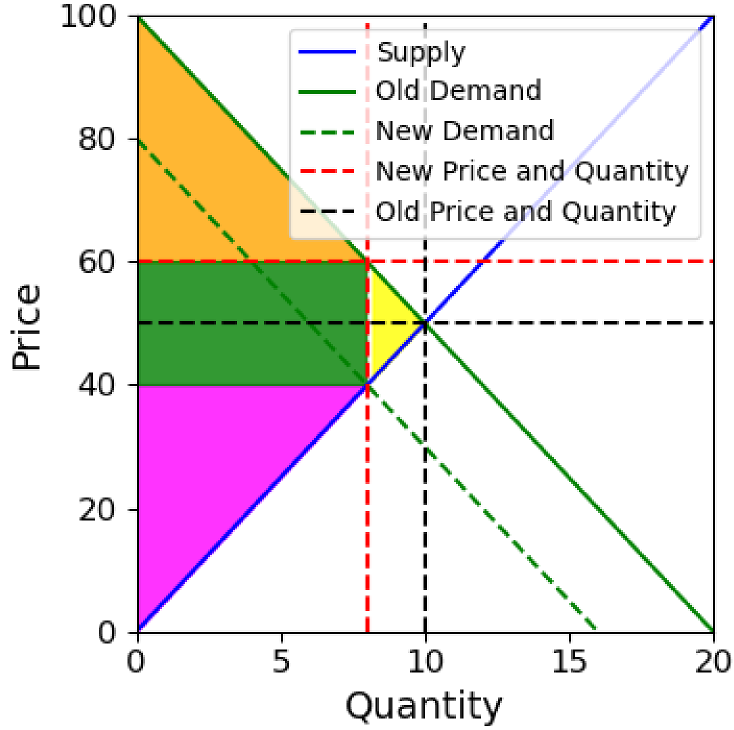

Figure 1.5: This shows what happens when there is a tax on the consumer. The orange region is consumer surplus, the green region goes to whoever collects the tax, the magenta region is producer surplus, and the yellow region is the "deadweight loss". The deadweight loss are trades that would have happened if there were no tax, but now simply do not occur.

suppliers will then only supply up to that amount. The orange region is again consumer surplus and the magenta region is producer surplus. The green region goes to the tax collector, but the yellow region is deadweight loss. That is, trades that no longer occur because of the tax. This is considered bad, since there could be sellers selling and buyers buying and getting value from the trades if not for the tax.

Now, if we instead tax the supplier, they will view it differently and we can see this in Figure 1.6. Here, the supply curve will simply go up by the tax T. This is because the supply curve is the minimum price that a supplier will sell for at some quantity. Thus, the supplier simply adds it on at each quantity shifting the curve upward. The buyers will then adjust to demand less at the new price. The orange region is again consumer surplus and the magenta region is producer surplus. The green region goes to the tax collector, but the yellow region is deadweight loss. That is, trades that no longer occur because of the tax. This is considered bad, since there could be sellers selling and buyers buying and getting value from the trades if not for the tax.

Note that the shaded regions are the exact same in Figures 1.5 and 1.6, showing that it does not matter if we tax the consumer or producer.

What we can notice, however, is that in this case, the consumer and the producer equally provide for the tax in our graphs above. What happens if we make the supply curve more inelastic, however? What about the demand curve the more inelastic one? We get Figure 1.7. We see very clearly that the party with the more inelastic curve is the one who pays more of the tax.

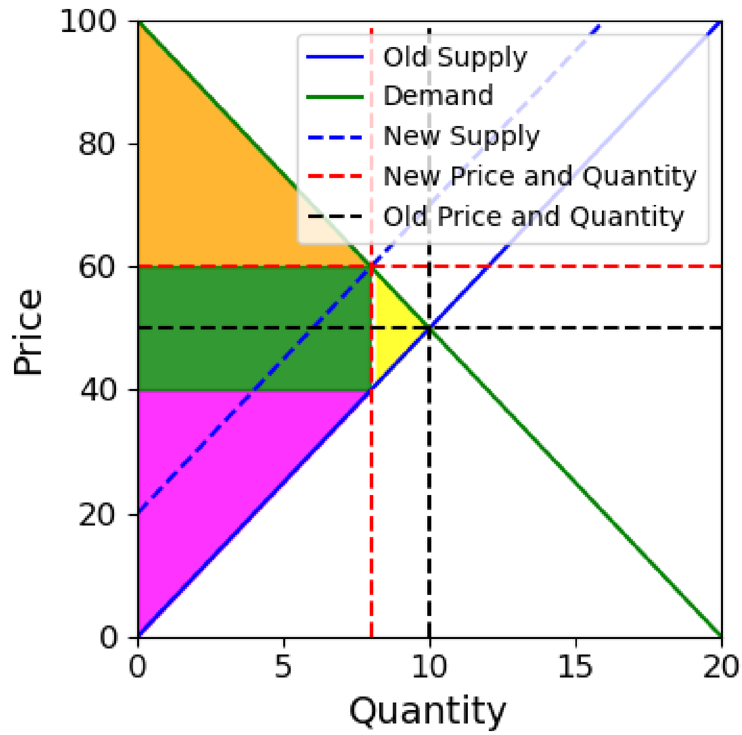

Figure 1.6: This shows what happens when there is a tax on the producer/supplier/seller. The orange region is consumer surplus, the green region goes to whoever collects the tax, the magenta region is producer surplus, and the yellow region is the "deadweight loss". The deadweight loss are trades that would have happened if there were no tax, but now simply do not occur.

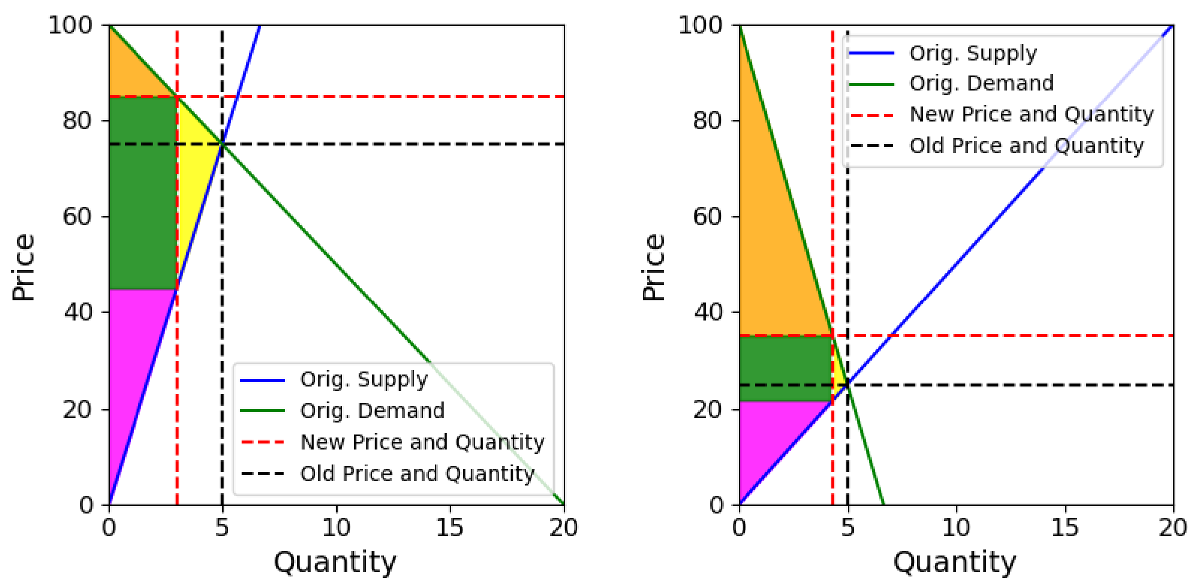

Figure 1.7: This shows what happens either the supply curve is more inelastic (left) of the demand curve is more inelastic (right). We see that the party with the more elastic curve pays less of the tax, or, equivalently, the more inelastic party pays more of the tax. So the supplier pays more of the tax on the left figure, and the consumer pays more of the tax on the right figure.

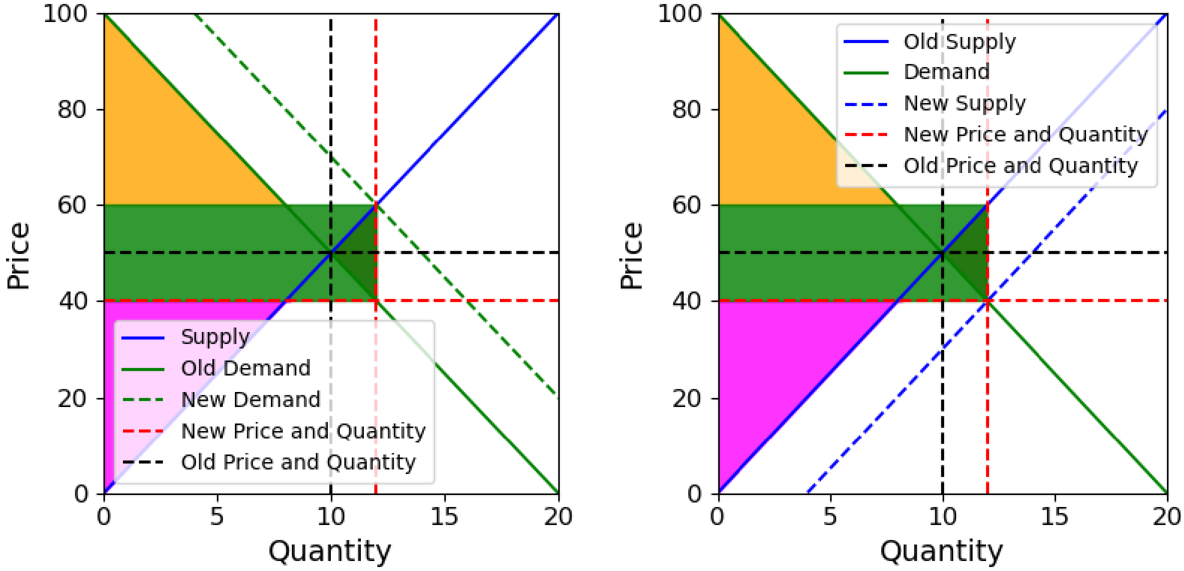

Figure 1.8: The effects of subsidies on either consumers or producers is shown. The orange region plus the part of the lighter green region above the equilibrium price line is the consumer surplus, the purple region plus the lighter green region below the equilibrium price line is the producer surplus, and the green (both lighter and darker) region is the cost of the subsidy (which is distributed among consumers and producers). The darker green triangle is considered a deadweight loss because without the subsidy these trades would not happen and so are not "properly" valuing the good.

Now we can get to subsidies. These are essentially the opposite of a tax, and again, it doesn't matter if we subsidize by giving to the consumer or the producer. In this case, a subsidy to the consumer makes them willing to buy more because the price is effectively lower to them, so the curve now shifts upward instead of downward. When the producer/supplier gets a subsidy, they can lower their price at the same quantity and so the supply curve moves downward instead of upward. Either scenario is shown in Figure 1.8. We see that the part with the more inelastic curve benefits from the subsidy more in Figure 1.9.

I should now add that while taxes and subsidies are often considered "bad" because of these properties, this is not exactly true, as we will learn when we get to externalities. It only really makes sense if the people in the demand and the people in the supply space have to pay for all the costs from the trade/transaction. If some bystander or society in general has to pay part of the cost, then it may actually be optimal to tax or subsidize to let consumers and producers know the "true" cost. But, in general, it is best to think of things in the simplistic way unless you have specific information that tells you it is wrong. Most of the time markets value things close to properly, and it will be hard to create a subsidy or tax that actually gives the market the "proper" value evaluations.

We can then consider what happens if we impose a price ceiling or a price floor. This is different than a tax or subsidy, but we can use some of our same tricks with our price vs quantity graphs. We find Figure 1.10. The effects are bad for both cases since we get deadweight losses. For a price ceiling we find that we get shortages of the desired good, and people then queue or search for the resource by using some other resource than price to get their good. In addition, some

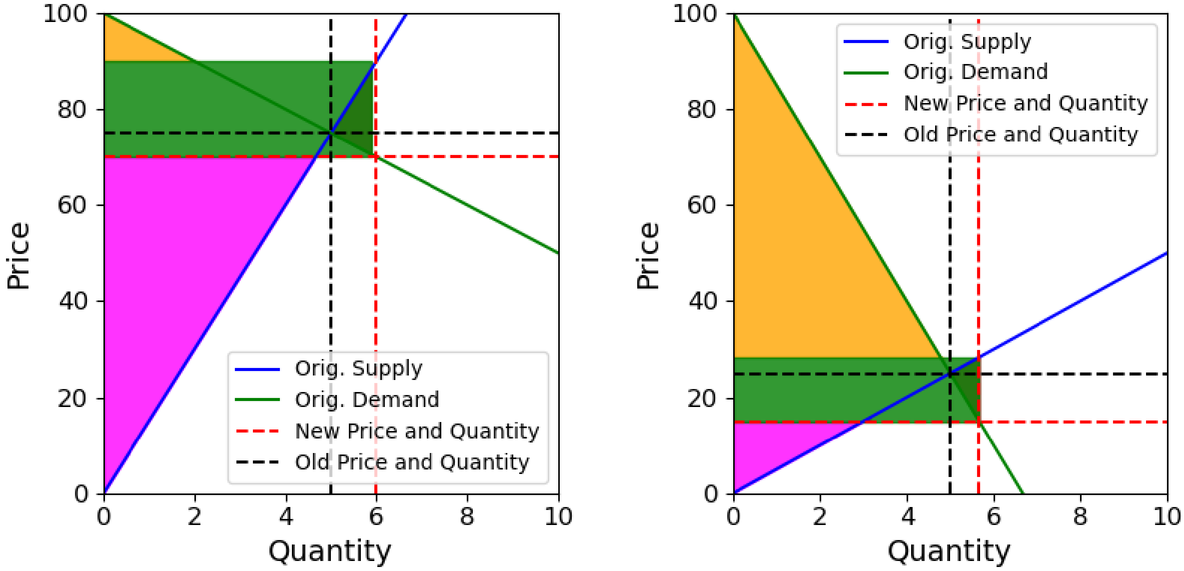

Figure 1.9: The effects of subsidies on either consumers or producers is shown when there are different elasticities. We see that the party with the more inelastic curve is the one that benefits most from a subsidy (just as they are penalized more by a tax). So the left figure shows the consumer getting more of a benefit from a subsidy and the right figure shows the producer getting more of a benefit from a subsidy.

producers decide not to produce because of the restriction leading to even fewer goods available. For a price floor, there is a surplus of the good because there is less demand than there are goods available. This also leads to a deadweight loss since some consumers are unwilling to buy at the price floor. A common example of a price floor is a minimum wage with the supply being labor and the producers being employers.

We can see that taxes or subsidies are generally a better mechanism since while we have deadweight losses, we do not have shortages or surpluses in general.

Externalities are the next thing to focus on. An externality can be either a benefit or a cost. A externality benefit is something like for a person taking a vaccine. They benefit the community but are not (usually) paid for this benefit they provide. An externality cost is something like pollution, where the polluter and the people buying the product (which causes the pollution) don't have to pay for the pollution, and it harms bystanders, as they either have to pay to clean it up or have worse health outcomes. An external cost is analogous to a subsidy and an external benefit is analogous to a tax. The deadweight loss regions are the same, but now we interpret the original equilibrium price and quantity as the inefficient one (because the market is not properly taking into account costs and benefits).

To make a market properly take into account externalities, the government can step in. Usually it is better for there to be a tax (for greenhouse gases, for example, one can trade permits that allow the release of so much gas, with the number of permits determining the overall amount of greenhouse gas permitted) than for the government to demand a specific solution. The government specific solution method can be superior if a solution that works is well-known and weak compliance will not solve the problem (eliminating small pox with vaccination is a commonly cited example).

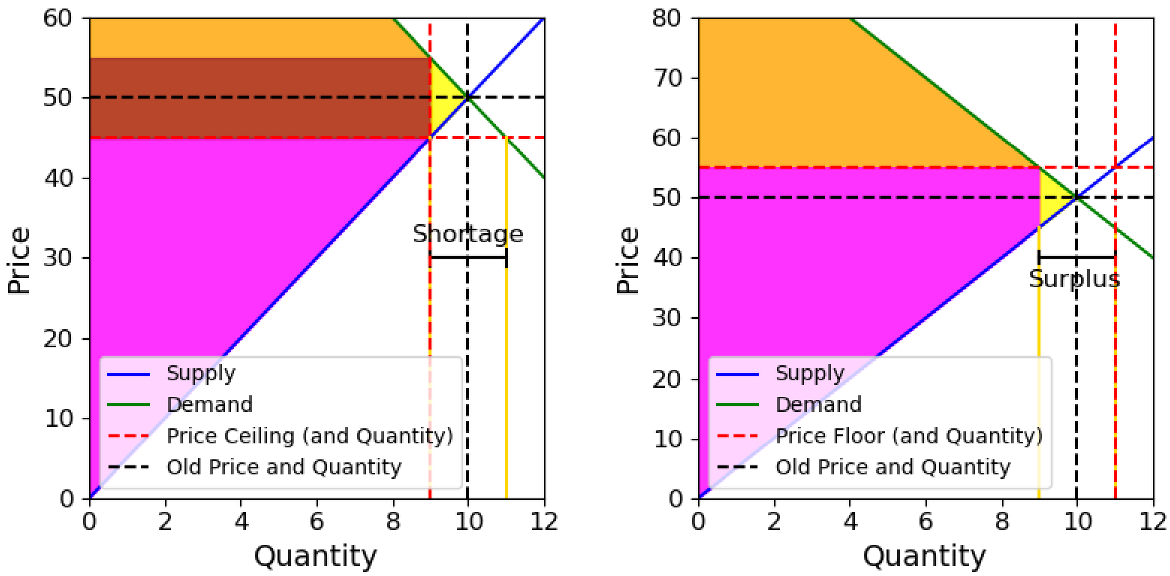

Figure 1.10: The effect of a price ceiling is shown on the left figure. For a price ceiling we see that there is a shortage, with the darker rectangle showing the lost time (the cost in time or something else used to get the goods for those buyers). There is also a deadweight loss in the yellow region. The effect of a price floor is shown on the right figure. Here we see there is a surplus of the good, and there are deadweight losses in the yellow region.

See Figure 1.11 for the graphical interpretations of these two possibilities.

Finally, besides creating taxes, one can give some entity property rights over the externality and if it satisfies Coase's theorem, then the market will properly assess the value of the externality in transactions. Coase's theorem says that if 4 things are satisfied, then assigning property rights will properly assess the value of the externality. These four things are

- Property rights must be clearly defined.
- There must be little to no transaction costs.
- There must be few affected parties (transaction costs are high if one must navigate many parties)
- There are no wealth effects. The efficient solution will be the same regardless of which party gets the property rights.

Note that there are behavioral economic critiques of Coase's theorem, but in some cases it can apply without controversy.

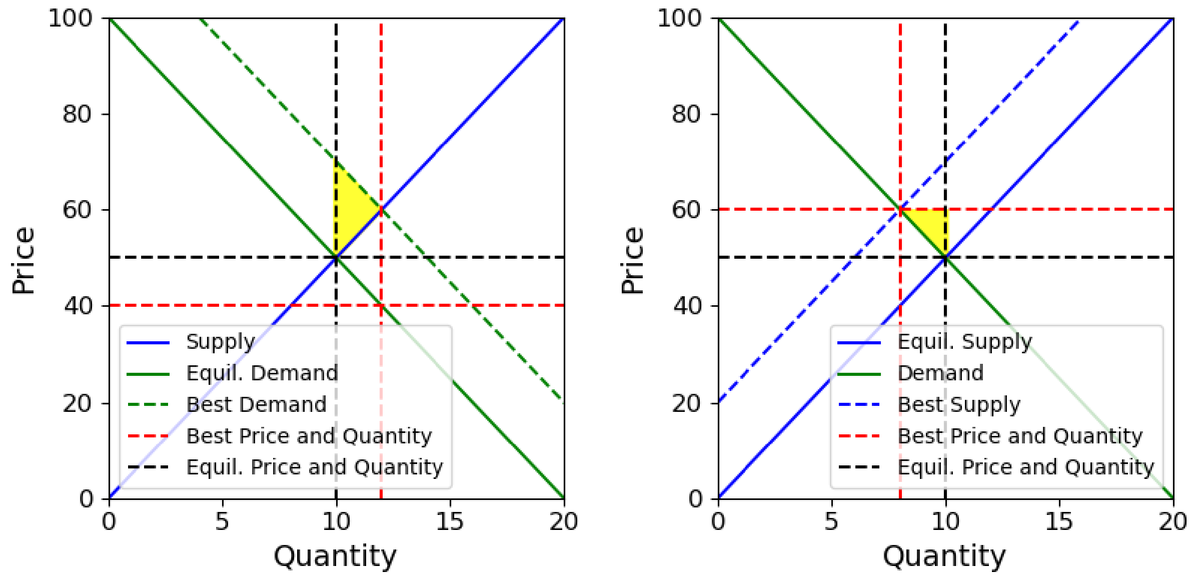

Figure 1.11: These show the regions of deadweight loss in yellow. An external benefit is shown on the left and an external cost is shown on the right.

1.2 Competitive Firms

A competitive firm is a firm in a market where there are many small buyers and sellers, with each seller selling a virtually identical product. They are small in the sense that the quantity that any one firm produces of a product does not affect the world price. In this sense, the price is set for them. The demand curve for a competitive firm's product is completely flat because of this. The price is set, and so at any quantity they will get the same price for their product, (perfectly horizontal line means perfectly elastic on a price vs quantity graph).

We can then think about how such a competitive firm maximizes its profits. For an economist, profits include a calculation of opportunity costs[2], so that an accountant's profit would be calculated slightly differently. Given that opportunity costs are often impossible to accurately calculate[3], accounting costs are generally better to use outside of theoretical models. There are other costs, for example there are fixed costs and variable costs. A fixed cost is something that does not change in cost as you produce more of your product. A variable cost is one that changes as one produces more of your product (generally variable costs increase as you produce more of a product).

Economists use profit (Π) is the same as total revenue (TR) minus total costs (TC) which is often written

$$\Pi = TR - TC \tag{1.2.1}$$

with each variable considered dependent on the quantity produced (since a competitive firm can't affect the world price).

[2]Costs incurred for not doing something else with your time which would have gotten you more value, usually money.

[3]How does one know what the optimal use of resources would have been without omniscience? It comes down to some subjective decisions on what is more valuable.

Some other terms to get used to are marginal revenue,[4] marginal cost, and average cost. These are usually denoted MR, MC, and AC, respectively. They come from the derivatives of (1.2.1) with respect to quantity. We have

$$\frac{\mathrm{d}\Pi}{\mathrm{d}Q} = \overbrace{\frac{\mathrm{d}TR}{\mathrm{d}Q}}^{=MR} - \overbrace{\frac{\mathrm{d}TC}{\mathrm{d}Q}}^{=MC} \tag{1.2.2}$$

We maximize profit by setting the above equal to zero which means profit is maximized at $MR = MC$. As we stated above, the demand for a competitive firm's product is perfectly elastic, and so each extra unit of product sold will go at the world price, P. Thus the marginal revenue is simply the world price, $MR = P$. Also, for an economist, maximizing profit does not mean that profit is greater than zero, simply that this is the greatest profit possible given the world price and the firm's capabilities. We can easily see that $MC = P$ will maximize profit from our previous reasoning. We'd like to know if our profit is positive, though and we can use average cost AC to figure this out. Remember that $AC = TC/Q$ by definition. This means if we look at

$$\frac{\Pi}{Q} = \frac{TR - TC}{Q} = \frac{TR}{Q} - AC \tag{1.2.3}$$

If we use $TR = MR \times Q$ then

$$\Pi = Q(MR - AC) = Q(P - AC) \tag{1.2.4}$$

Thus, if $P - AC > 0$ or $P > AC$ then profits will be positive. This is fairly intuitive, as well. If your average cost is more than the price you sell for, you are losing money. If it is much greater than the price you sell for, you're making a lot of profit. It is also worth noting that $AC = (FC + VC)/Q$ where FC is fixed costs and VC is variable costs. Generally, FC is a constant and VC increases as a function of the quantity produced Q, so that $AC = FC/Q + VC/Q$ with VC/Q a growing function of Q so that AC is somewhat parabolic shaped.

One last thing, zero profit in economics means "usual" or "normal" profit. It means that the company is willing to keep going like it has and so a zero profit company economically may be a huge profit maker from an accounting point of view (the economist considers the accountant's profits as money being used for something and so perhaps not as a "profit").

Now one can determine if an industry is an increasing costs industry, constant costs industry, or decreasing costs industry. An increasing costs industry means that the supply curve for the entire industry slopes upward and to the right on a price vs quantity graph (as we typically assume). A constant costs industry means that the supply curve is flat (perfectly elastic). A decreasing costs industry is somewhat rare, but means that the supply curve slopes downward and to the right, like a demand curve. This leads to an interesting phenomenon where there is a locus of the industry at a single spot, often called a cluster. This is because either the technique or the machines needed to make the product are especially valuable when they are near each other and push costs way down. Silicon Valley was considered a decreasing costs industry for making silicon chips when it originally sprang up.

An increasing cost industry is marked by having inputs that are difficult to duplicate (such as natural resources like oil or gold), so that it is more difficult to get more of the natural resource so

[4] Remember that revenue is not profit. Revenue is simply the money taken in, not considering the costs.

that increasing prices actually does lead to new suppliers coming online. A constant cost industry is marked by the fact that an increase in price leads to new suppliers which make up for the new demand, and so all individual suppliers end up with the same quantity produced and the price normalized back to its original.

1.3 Monopolies

Let's now consider the effect of a monopoly. The monopoly may be due to a patent or some sort of regulation scheme, but we assume that one company has control over all of the supply of a product. A monopoly has market power, the power to raise prices above the marginal cost without competitors coming in to drive the price down.

Profits are still determined by (1.2.1), and so (1.2.2) still applies meaning profit maximization occurs at $MR = MC$. The difference is that $MR \neq P$ because of the monopoly. Therefore the $P > MR$ (if it were not, then the monopoly would be losing money, which makes little sense as the monopolist could just not produce whatever it is they are making). Thus a monopoly has a demand curve for their product that slopes down and to the right on a price vs quantity graph. This means there is a point at which the amount of revenue gets smaller as one increases the quantity sold.

We can find the marginal revenue curve from the demand curve. For a linear demand curve, it turns out the marginal revenue curve is simply the curve with twice the (negative) slope of the demand curve. That is given

$$D = -m_D Q + D_0 \tag{1.3.1}$$

for the demand curve, then the marginal revenue curve is

$$MR = -2m_D Q + D_0 \tag{1.3.2}$$

We can figure this out from the fact that revenue R is given by

$$R = P(Q)Q \tag{1.3.3}$$

where $P(Q)$ is the price at quantity produced Q. Then

$$\frac{\mathrm{d}R}{\mathrm{d}Q} \equiv MR = \frac{\mathrm{d}P}{\mathrm{d}Q}Q + P = P\left(1 + \frac{\mathrm{d}P}{\mathrm{d}Q}\frac{Q}{P}\right) = P\left(1 + \frac{1}{\epsilon}\right) \tag{1.3.4}$$

(one can see again why the elasticity is essentially the inverse of what would normally be useful since we use price vs quantity relationships). We then see that for $P = -m_D Q + D_0$ as our demand curve that

$$MR = -m_D Q + (-m_D Q + D_0) = -2m_D Q + D_0 \tag{1.3.5}$$

In general, one can use the more general formula (1.3.4) to find a result.

We find that the monopoly markup very much is determined by the elasticity of demand. If the demand is highly inelastic, then the markup is large, and if the demand is elastic, then the markup is small. This makes some sense, as if the monopoly has control and the demand is inelastic, this

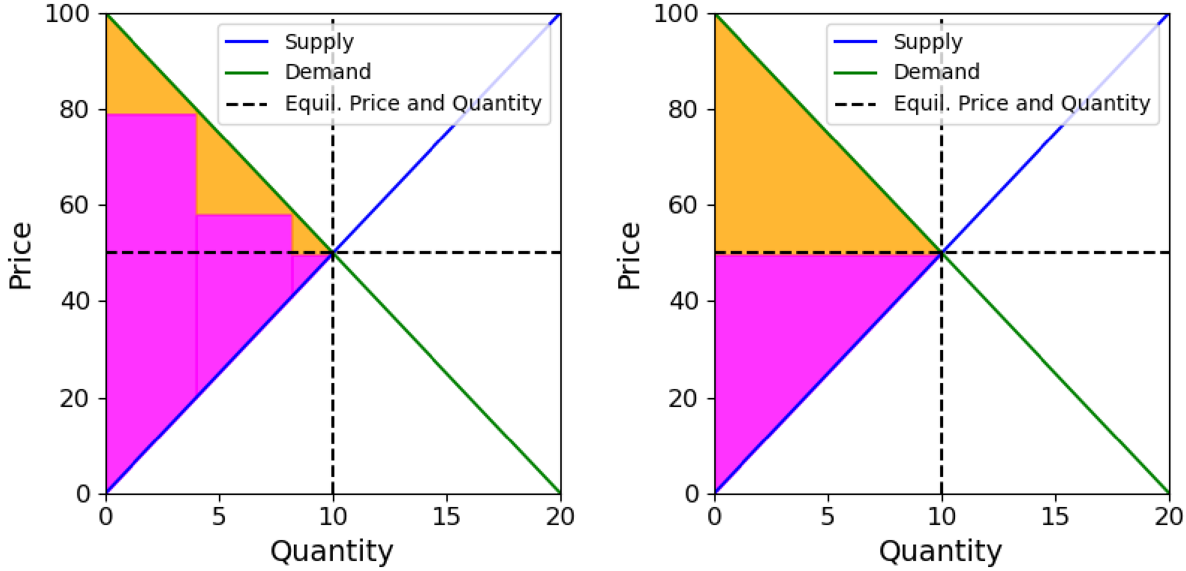

Figure 1.12: How price discrimination increases monopoly revenues (producer surplus in magenta) via taking away consumer surplus (the orange regions). The left shows what happens under price discrimination, while the figure on the right shows what usually happens without price discrimination.

means there are few substitutable goods and so the monopoly can take advantage of those who have to pay for the good.

Now, we can look at price discrimination. Price discrimination is when a firm charges a different price to different consumers (essentially dividing the consumer market up and charging each section the highest price it can). This is difficult to do in a competitive market, since competition drives the price down to the marginal cost, but with a monopoly the monopolist can control the price.

If we think about our normal supply and demand curves, we are basically supplying a product at a price along the demand curve. In a competitive market this usually wouldn't work because the consumers would see that they can get a better price and so buy from someone else. Thus, price discrimination is a way for a monopolist to increase its revenues via changing consumer surplus to producer surplus. We can see how three price levels for a monopoly could work in Figure 1.12

You might think that price discrimination is always bad from this, but that is not necessarily so. Price discrimination may mean that more people can get the good, because if only one price existed, the producer would produce less of the good (they are not going to go away from profit maximizing). That is, the right image in Figure 1.12 assumes that the same quantity would be produced without price discrimination. This is often not true, and the equilibrium quantity would actually be smaller without price discrimination (thus, there are some deadweight loss trades here).

How does price discrimination work in practice, though? Well, there are a couple of mechanisms. Tying works by having one good require another good to function properly. For example, a printer will require ink or toner. Those who use more ink or toner will have to buy more and so they will spend more than those who only print a little bit, a form of price discrimination.

Another type of price discrimination is bundling. This occurs when a firm puts two or more products together. This is because people do not value the products in the bundle equally. Some find a lot of value in product A, other in product B, etc. By putting the products in a bundle, the firm can ensure a sale on all of their products by pricing just right. This is because people will probably find value in most of the products in the bundle, and so they will be willing to pay for the bundle, even if they weren't willing to pay for each product on its own. What I am saying is that you can reduce the variance. This works best for goods with very low to zero marginal cost, since increasing sales costs essentially nothing (thus software and cable channels are often bundled). This type may or may not be harming the customer surplus depending on the specifics.

1.4 Types of Goods

It is worth stopping and considering the different types of goods that are typically considered. For this we need the concept of a rivalrous good and an excludable good. A rivalrous good is one where the good is used up after someone pays for it, so no one else can consume it. An apple is a rivalrous good since it is literally consumed. An idea is a non-rivalrous good since when one person uses it, it does not prevent another person from using the same idea. An excludable good is a good where it is easy to prevent people from using the good if they don't pay for it. The apple is excludable since by eating it or keeping it, no one else can use the apple. A non-excludable good would be something where someone can gain the benefit of the good without paying for it. National defense is a common example, but so is the timber in an unowned forest. If you pay for the national defense, but your neighbor does not, they still get the benefit from you paying.

This allows us to create a matrix of goods: private goods, open access common goods, club goods, and public goods, as shown in Table 1.1.

Markets work extremely well with private goods, often work with club goods, but have problems with open access common goods and public goods. We have the tragedy of the commons, where people exploit the resource using it all up to the detriment of all (if you fish all the fish in the ocean, there is no longer a fishing industry), or no one pitches in even though it would be better if everyone did just a little bit (national defense, where if no one pays, you can all lose everything).

	Excludable	Non-Excludable
Rivalrous	Private Goods	Open Access Common Goods
Non-Rivalrous	Club Goods	Public Goods

Table 1.1: This shows the matrix of some common names for goods with different rivalrous and excludability properties.

1.5 Asymmetric Information

My final section on introductory microeconomics will be on some of the terminology and basic ideas around the fact that sometimes consumers and producers have very differing expertise or asymmetric information. Asymmetric information can benefit either party. For example, a health insurer has asymmetric information in that they only know a very limited amount about any person they are insuring, whereas the person knows a lot about their personal situation. This means that people have a reason to only get insurance when they know they will need it. This is called adverse selection because the health insurer will only insure people who are the most expensive to maintain, and so will be unable to make money.[5] In the naive health insurance market, the willingness to get health insurance signals that one should not give health insurance to that person, thus an offer conveys negative information.

Another phenomenon associated with information asymmetry is moral hazard. Moral hazard is the temptation to exploit an information asymmetry. For example, a computer repair person may know a lot more about your broken laptop than you. If they sell you a bunch of new things you don't need, knowing you won't be able to tell these things are unnecessary. This is a moral hazard. It is an entering into an agreement with some misleading way. Additionally, moral hazard may occur if one does not guard against a risk because one is shielded from the consequences (if you bought home insurance but then skimp on protections for your house from damage).[6]

Finally, one can make up for information asymmetry sometimes by signalling. For example, if you want to prove that your product is durable, you can offer a very generous warranty. This will send the signal that your product is of high quality since otherwise you would be losing a lot of money by offering such a warranty.

[5]In fact, this is not what happens in real life because people who are risk-averse are more likely to purchase insurance in the first place, and so there is not such a strong adverse selection effect.

[6]Note how this specific use is essentially covered by the "misleading" expression in the previous explanation.

Chapter 2

Macroeconomics

Now we get to macroeconomics, which differs from microeconomics by not (directly) considering individual decisions leading to an economy, but the combination of all those decisions at once. This is a bit like in physics going from statistical physics to thermodynamics. Statistical physics determines thermodynamics and microeconomics determines macroeconomics. But you can miss a forest for the trees, and thermodynamics often allows you to ask and answer questions easily that would be subtle to nearly impossible with statistical physics only. In the same way, macroeconomics lets us reason about large scale economic phenomena without knowing all of the microeconomic data and decisions.

2.1 Macroeconomics Basics

To start with, we need to talk about the variables and quantities we will be considering. First let's discuss GDP, nominal and real GDP, per capita GDP, and GNP/GNI. Gross domestic product (GDP) is the market value of all finished goods and services produced within a geographic area (typically a country's borders) in one year. Market value is what would these goods/services be sold at, and finished goods are goods that will not be sold again as a part of another good. Thus, if iron is sold to make nails, then we don't count the iron being sold to the nailmaker and the nail being sold to the consumer. We only count the nail being sold to the consumer so that we do not overcount. The iron is an intermediate good, which will become a finished good in some final product. GDP also only considers goods produced in the geographic area. So if country A produces a finished good and sells it in country B, the good counts towards GDP in country A only. If you live in country A and buy a finished good produced in country B, then that counts toward country B's GDP.

Gross national product (GNP)[1] is the market value of all goods and services produced in one year by labor and property supplied by the citizens of a country. Note that this means if the owner of a factory owned in Germany is American, then one will count the material sold from that factory as contributing to American GNP, and to French GDP.

There are some limitations even to the above definitions. Some goods/services are not available on

[1]This is more commonly called GNI or gross national income (GNI) today. GNI is now used because it is typically calculated from GDP rather than from the GNP definition, even though GNP and GNI actually denote the same idea/number.

a market and so are not counted towards GDP or GNP. Generally speaking, though, GDP/GNP are correlated with each other and a higher GDP means a bigger economy.

Next we consider the difference between nominal GDP (nGDP) and real GDP (rGDP).[2] When we simply report the numbers we see each year, we are using nGDP. If there were no inflation, then nGDP would be the same as rGDP. We usually want to use rGDP since that tells how much more the economy is actually able to produce rather than what numbers were associated that year with the production of goods/services.

Finally, an important correlate to well-being is the GDP per capita (pcGDP), which is simply the total GDP divided by the population size. Note that there is again nominal and real versions which I will denote npcGDP for nominal per capita GDP and rpcGDP for real per capita GDP.[3] It may not super surprising but the higher rpcGDP is, the better off people tend to be in a country.

Finally, when comparing different country GDPs, one has to make adjustments for the fact that they usually use different currencies. There are multiple ways of doing this, with purchasing power parity (PPP) being one of them. This uses a "handbasket" of goods and compares prices in each currency for this handbasket to determine the corrections. Another prominent adjustment is using market exchange rates or the current currency exchange rate (how much it would cost to convert a currency from one to another on the international market). PPP is generally considered the better method for within country comparison, and a GDP using PPP is pretty much automatically converted into a real GDP. Market exchange rates tell us how much international purchasing power a country has since they will probably have to convert their currency into another.

When looking at GDP, one often considers two different approaches. The national spending approach and the factor income approach. The national spending apprach looks at by splitting between consumption (goods and services consumed), investment (in capital stock, usually by businesses) and government purchases; this approach then adds exports and subtracts imports to get the GDP number. Government purchases should not be confused with government spending. Government purchases mean that the government bought something directly. Government spending is something like sending a check to every person, which is just a transfer of wealth (we don't want to double count, as this money will be used by the people with the check on consumption or investments). These divisions are used because we know that the different divisions can act fairly independently of each other and so we can have useful analyses of what caused what with these divisions.

The other approach, the factor income approach, looks at the other side of the coin. It adds up employee compensation, rent, interest payments, and profit to find the total GDP. Note that essentially by definition, the money spent must equal the money earned, hence this is an equivalent approach to the national spending approach. We may get different values in practice because of the difficulty of proper accounting, but if these are similar then we know we have probably calculated the GDP correctly. This is essentially just saying that every transaction requires two people so income for one person is spending for the other. This income approach is often called gross domestic income or GDI (for the same reason there is a GNI).

[2]As a physicist, I find it annoying that many measures don't have a standardized unit that differentiates nominal and real GDP and so I have adopted nGDP and rGDP. It is like having two different measures of angles, radians and degrees, but never labeling which is which. I will also employ $^{n}\$$ for nominal dollars and $^{r}\$$ for real dollars so that we don't confuse the two.

[3]Since per capita is Latin (for "by heads" or "for each head"), it seems we can put it before or after GDP and people understand what is meant.

At this point, economists really like to show that a small but positive growth rate over a long time leads to enormous growth. People are poor at understanding exponential growth, but it really is just the fact that continual increases make a big difference. If you can sustain a small percentage increase in growth, then over the long haul you will reap enormous benefits.[4]

The importance of institutions is then usually explained. Institutions are laws, regulations, and cultural norms. Cultural norms can be things like how honest people are with each other, how much they value innovation and innovators, etc. Property rights, political stability, a good legal system, honest government and competitive and open markets are examples of good institutions. These are usually hard to change, but can often explain the differences in outcomes between countries.

2.2 Solow-Swan Model

Now we get to some of the more interesting ideas. This is a simple model that explains the difference between catching up growth (like that done by China, Japan, Germany, and South Korea post World War II, small rpcGDP countries) and cutting edge growth (USA, most of Western Europe, and modern Japan and modern South Korea, high rpcGDP countries).

There is a production function $Y(t)$ that gives us our production output given a time t. In fact, we write $Y(t) = Y(K(t), A(t), L(t))$ where $K(t)$ is capital stock (machines, factories, etc.), $A(t)$ is ideas/innovation, and $L(t)$ is labor (for example a more educated work force has a larger L; if one adds more workers of the same effectiveness then L also gets larger). A represents knowledge that makes labor more effective (so education can be a part of A). This is usually rendered into a model via

$$Y(t) = K(t)^{\alpha}[A(t)L(t)]^{1-\alpha} \tag{2.2.1}$$

with $0 < \alpha < 1$ called the elasticity of output with respect to capital (since α is the exponent of capital K). As a physicist, I like to think about the units of everything. We remember that GDP is actually a flow (because it is per year) and so $Y(t)$ should be able to be measured in ${}^{r}\$/\text{yr}$. If we measure $K(t)$ and $L(t)$ in these units as well, with $A(t)$ dimensionless, then everything will work out fine. We can then reason as follows. $Y(K, A, L)$ holding A and L constant ($Y(K, A, L)|_{A,L} = Y(K)$) should increase as K increases, but should increase less and less as K increases. This is saying $\frac{\partial Y}{\partial K} > 0$ but $\frac{\partial^2 Y}{\partial K^2} < 0$ in calculus language. Thus we have an increasing concave down function. This is simply saying that we have diminishing returns and explains why α is forced to remain between 0 and 1. This also helps explain why catching up growth is faster. If you're at a low level of K, adding one more unit of K leads to a greater increase than when you add one more unit of K when K is high. One other consideration is that K also depreciates (because the capital stock rusts, breaks down, needs fixing, etc.).

We can say that depreciation means that $\frac{\mathrm{d}K(t)}{\mathrm{d}t} = -\delta K(t)$ so that it is constantly depreciating for some constant δ. In addition, to make up for this, we can invest part of the output into paying for more K and countering the depreciation. Thus

$$\frac{\mathrm{d}K(t)}{\mathrm{d}t} = \overbrace{-\delta K(t)}^{\text{depreciation}} + \overbrace{\gamma Y(t)}^{\text{investment}} \tag{2.2.2}$$

[4]Since we usually do not know what types of interventions are sustainable that lead to growth, this becomes a tautological statement of little relevance to policy except for big interventions (such as communism [as it has existed] vs capitalism [as it has existed] where history provides a guide).

If we assume that AL is a constant in time so $[AL]^{1-\alpha} \equiv \nu$, then we can find the steady state value by setting $\frac{\mathrm{d}}{\mathrm{d}t} \to 0$ and we find

$$\delta \overbrace{K(t \to \infty)}^{\equiv K_s} = \gamma\nu K(t \to \infty)^{\alpha} \tag{2.2.3}$$
$$K_s^{1-\alpha} = \frac{\gamma\nu}{\delta} \tag{2.2.4}$$
$$K_s = \left(\frac{\gamma\nu}{\delta}\right)^{\frac{1}{1-\alpha}} \tag{2.2.5}$$

We can determine if this is a stable or unstable equilibrium by plugging in $K_s + h$ where h is a perturbation to K_s. Then

$$\frac{\mathrm{d}(K_s + h)}{\mathrm{d}t} = -\delta(K_s + h) + \gamma\nu(K_s + h)^{\alpha} \tag{2.2.6}$$
$$\frac{\mathrm{d}h}{\mathrm{d}t} = -\delta K_s - \delta h + \gamma\nu K_s^{\alpha}\left(1 + \frac{h}{K_s}\right)^{\alpha} \approx \cancel{-\delta K_s} - \delta h + \gamma\nu K_s^{\alpha}\left(\cancel{1} + \alpha\frac{h}{K_s}\right) \tag{2.2.7}$$
$$\approx h\left[-\delta + \alpha\gamma\nu K_s^{\alpha-1}\right] \tag{2.2.8}$$

Thus we are stable if

$$-\delta + \alpha\gamma\nu K_s^{\alpha-1} < 0 \tag{2.2.9}$$
$$\alpha\gamma\nu\left(\frac{\gamma\nu}{\delta}\right)^{\frac{\alpha-1}{1-\alpha}} < \delta \tag{2.2.10}$$
$$\alpha\gamma\nu\frac{\delta}{\gamma\nu} < \delta \tag{2.2.11}$$
$$\alpha < 1 \tag{2.2.12}$$

assuming δ, γ, ν are not zero.[5]

This shows that if we hold AL fixed, we will eventually hit a point of no economic growth, because we will hit the steady state value. Increasing γ can get us to a larger final Y value, but $0 \leq \gamma \leq 1$.

Note that if we hold A and K fixed, then we can say the same things about L that we did about K. It has $\frac{\partial Y}{\partial L} > 0$ and $\frac{\partial^2 Y}{\partial L^2} < 0$. More L increases Y but with diminishing returns. In addition, labor also has costs of depreciation (people get older and drop out of the workforce).

Thus, this model says poorer (less productive) countries should catch up to richer (more productive) countries as time goes on if A does not differ between countries. This is only somewhat true, of course. Essentially, what we are saying is that two countries with similar institutions will converge to the same production, while those with different institutions will most likely diverge (and we do see this prediction mostly borne out in economic data). However, we don't see that economies get to zero growth. Here is where A plays the biggest role.

If A is a growing function of t, then essentially we find that ν is a function of t and if it grows without bound then the K goes to infinity, and we have no steady state. Thus, a larger A is the key to cutting edge growth and economic improvement over the long run.

[5] Note that this is sometimes used as an argument why $\alpha < 1$. Otherwise, there would be no steady state because the production function could grow without bound, and hence K and L could, too.

Some institutions that enhance A are an appreciation for creating new knowledge, patents, prizes for innovation, etc. These are important for economic growth and thus to people's well-being (generally speaking).

Now, the more advanced treatment uses a production function with the properties $Y(K, AL)$ (since we only have A and L in the combo AL) with the idea of constant returns to scale (in normal mathematical/physical language, this implies that Y is a homogeneous function and so Euler's homogeneous function theorem will apply) which is just a statement that

$$Y(tK, tAL) = tY(K, AL) \tag{2.2.13}$$
$$\tag{2.2.14}$$

for $t \geq 0$ [also $Y(0,0) = 0$]. Then Euler's homogeneous function theorem which says for function satisfying ($\mathbf{V}$ is a vector with components V_i)

$$f(\alpha\mathbf{V}) = \alpha^n f(\mathbf{V}) \tag{2.2.15}$$

then

$$\mathbf{V}\cdot\frac{\partial f(\mathbf{V})}{\partial\mathbf{V}} \equiv \sum_i V_i \frac{\partial f(\mathbf{V})}{\partial V_i} = nf(\mathbf{V}) \tag{2.2.16}$$

for n an integer. So for us, we find

$$K\frac{\partial Y}{\partial K} + AL\frac{\partial Y}{\partial(AL)} = Y \tag{2.2.17}$$

We then divide by AL to normalize our function and write

$$y \equiv \frac{Y(K, AL)}{AL} = \overbrace{\frac{K}{AL}}^{\equiv k}\frac{\partial Y}{\partial K} + \frac{AL}{AL}\frac{\partial Y}{\partial(AL)} = k\frac{\partial Y}{\partial K} + \frac{\partial Y}{\partial(AL)} \tag{2.2.18}$$

What we'd like to show is that

$$y(k) = y(K, AL) = Y(K, AL)/(AL) \tag{2.2.19}$$

for simplicity define $\ell = AL$

$$\frac{\partial y}{\partial K} = \frac{\mathrm{d}y}{\mathrm{d}k}\frac{\partial k}{\partial K} = \frac{\mathrm{d}y}{\mathrm{d}k}\frac{1}{\ell} \tag{2.2.20}$$
$$\frac{\partial y}{\partial \ell} = \frac{\mathrm{d}y}{\mathrm{d}k}\frac{\partial k}{\partial \ell} = \frac{\mathrm{d}y}{\mathrm{d}k}\frac{-K}{\ell^2} \tag{2.2.21}$$

We also must have

$$\frac{\partial(Y/\ell)}{\partial K} = \frac{1}{\ell}\frac{\partial Y}{\partial K} \tag{2.2.22}$$
$$\frac{\partial(Y/\ell)}{\partial \ell} = \frac{1}{\ell}\frac{\partial Y}{\partial \ell} - \frac{Y}{\ell^2} \tag{2.2.23}$$

which means

$$\frac{\mathrm{d}y}{\mathrm{d}k}\frac{1}{\ell} = \frac{1}{\ell}\frac{\partial Y}{\partial K} \tag{2.2.24}$$

$$\frac{\mathrm{d}y}{\mathrm{d}k}\frac{-K}{\ell^2} = \frac{1}{\ell}\frac{\partial Y}{\partial \ell} - \frac{Y}{\ell^2} \tag{2.2.25}$$

or

$$\frac{\mathrm{d}y}{\mathrm{d}k} = \frac{\partial Y}{\partial K} \tag{2.2.26}$$

$$\frac{\mathrm{d}y}{\mathrm{d}k} = -\frac{\ell}{K}\frac{\partial Y}{\partial \ell} + \frac{Y}{K} \tag{2.2.27}$$

$$\tag{2.2.28}$$

which means

$$\frac{\partial Y}{\partial K} = -\frac{\ell}{K}\frac{\partial Y}{\partial \ell} + \frac{Y}{K} \tag{2.2.29}$$

$$Y = K\frac{\partial Y}{\partial K} + Y\frac{\partial Y}{\partial \ell} \tag{2.2.30}$$

which is our definition of a homogeneous function. This means we can simply use our steps in reverse to define k and show that it implies that $y(k) = Y(K, AL)/(AL)$.

The k is called the capital per effective unit of labor (it would be dimensionless if we are measuring K, L, and Y in the same units with A dimensionless). Our new $y(k)$ has the simple property that $\frac{\mathrm{d}y}{\mathrm{d}k} > 0$ and $\frac{\mathrm{d}^2y}{\mathrm{d}k^2} < 0$. (Note that this basically follows from $\frac{\mathrm{d}y}{\mathrm{d}k} = \frac{\partial Y}{\partial K}$ and the properties we noted for Y).

Typically, one then stipulates $L(t) = L_0\exp(nt)$ and so our previous equation (2.2.2) says

$$\frac{\mathrm{d}K}{\mathrm{d}t} = \gamma Y - \delta K \tag{2.2.31}$$

and finally that $A = A_0\exp(at)$ and so is also growing exponentially. We note that the $\mathrm{d}K/\mathrm{d}t$ equation is sometimes called the Goldsmith equation.

Economists like to say that A and L are exogenous in the model (specified from the outside because we gave them a specific form) and the capital stock K is endogenous because it is determined by things within the model (we did not dictate a specific form from the outset). Remember that γY is the savings, and that we can switch into the language of y. Then we must have $\gamma y = \gamma Y/(AL)$ and

$$\frac{\mathrm{d}K}{\mathrm{d}t}\frac{1}{AL} = \gamma y - \delta k \tag{2.2.32}$$

$$\frac{\mathrm{d}k}{\mathrm{d}t} - K\frac{\mathrm{d}(AL)^{-1}}{\mathrm{d}t} = \gamma y - \delta k \tag{2.2.33}$$

$$\frac{\mathrm{d}k}{\mathrm{d}t} + \frac{K}{(AL)^2}\left[\frac{\mathrm{d}A}{\mathrm{d}t}L + \frac{\mathrm{d}L}{\mathrm{d}t}A\right] = \gamma y - \delta k \tag{2.2.34}$$

$$\frac{\mathrm{d}k}{\mathrm{d}t} = -\frac{k}{AL}\left[aAL + nAL\right] + \gamma y - \delta k \tag{2.2.35}$$

$$\frac{\mathrm{d}k}{\mathrm{d}t} = -k\left[a + n\right] + \gamma y - \delta k \tag{2.2.36}$$

$$\frac{\mathrm{d}k}{\mathrm{d}t} = \gamma y - [\delta + a + n]k \tag{2.2.37}$$

This means the steady state $y = y_s$, $k = k_s$ yields

$$y_s = \frac{\delta + a + n}{\gamma} k_s \tag{2.2.38}$$

Remember that Y and K are essentially defined by $Y = yAL$ and $K = kAL$. By construction Y and K are growing, but we see that $Y/K = y_s/k_s$ is a constant (note that y_s and k_s reach a steady state value because they normalize to labor, whereas Y and K increase without bound).

If you wanted to choose a savings rate γ to maximize $(1-\gamma)y$ [this is the consumption, the non-invested part], what would be the optimal savings rate to choose? Essentially, we are asking we want the most consumption for our savings rate possible. That is given δ, a, and n, what γ gives us the largest $E = (1-\gamma)y_s$. We find

$$\frac{\partial E}{\partial \gamma} = -y_s - \gamma \frac{\partial y_s}{\partial \gamma} = -y_s + \frac{\delta + a + n}{\gamma^2} k_s = 0 \tag{2.2.39}$$

$$\frac{y_s}{k_s} = \frac{\delta + a + n}{\gamma^2} \tag{2.2.40}$$

This is only possible when $\gamma = 1$, which is clearly a minimum, and not the solution we'd like. Thus to find the maximum, we need to think about this differently. Basically what we'd like is that the $y'(k)$ to be the same as $\frac{\gamma y_s}{k_s}$. If $y'(k) < \frac{\gamma y_s}{k_s}$ then increasing k will lead to a smaller portion of consumption to savings because savings is growing faster. Conversely, by decreasing k we would be increasing consumption. If $y'(k) > \gamma y_s/k_s$ then increasing k will increase consumption relative to savings because the savings rate is smaller than the increase in consumption. Thus, only when they are equal do we get optimal consumption to savings. A graphical way to think of this is to plot $y(k)$ and $(\delta + a + n)k$ on the y axis with k on the x axis. Then only when $y(k)$ is parallel to $(\delta + a + n)k$ will we be the farthest away, since otherwise the $y(k) - (\delta + a + n)k$ will be increasing or decreasing in distance.

The same limitations as before apply. This model has weak microeconomic pinnings, with many exogenous assumptions driving the behavior. Also, the AY is only similar for countries with similar institutions, otherwise you don't get convergence.

Remember that the Solow model tells us the production function can be changed via three different routes. Via growth in labor input (i.e., more workers or more efficient workers), technical change (i.e., better use of workers and capital stock/products, new products, new methods), and growth in capital stock (more labor-saving products, etc.).

We can see most everything involved in Figure 2.1.

Last, it is useful to consider the behavior linearized around the k_s, the steady state value. Here we can write $k = k_s + \Delta k$. We use $\frac{\mathrm{d}k}{\mathrm{d}t} = \frac{\mathrm{d}k}{\mathrm{d}y}\frac{\mathrm{d}y}{\mathrm{d}t}$ or equivalently $\frac{\mathrm{d}y}{\mathrm{d}t} = \frac{\mathrm{d}k}{\mathrm{d}t}\frac{\mathrm{d}y}{\mathrm{d}k}$ we then find

$$\frac{\mathrm{d}k}{\mathrm{d}t} = \gamma y(k) - [\delta + a + n]k \tag{2.2.41}$$

$$\frac{\mathrm{d}y}{\mathrm{d}t} = \frac{\mathrm{d}y}{\mathrm{d}k}\left(\gamma y(k) - [\delta + a + n]k\right) \tag{2.2.42}$$

$$\frac{\mathrm{d}\ln y}{\mathrm{d}t} = \frac{\mathrm{d}y}{\mathrm{d}k}\left(\gamma - [\delta + a + n]\frac{k}{y}\right) \tag{2.2.43}$$

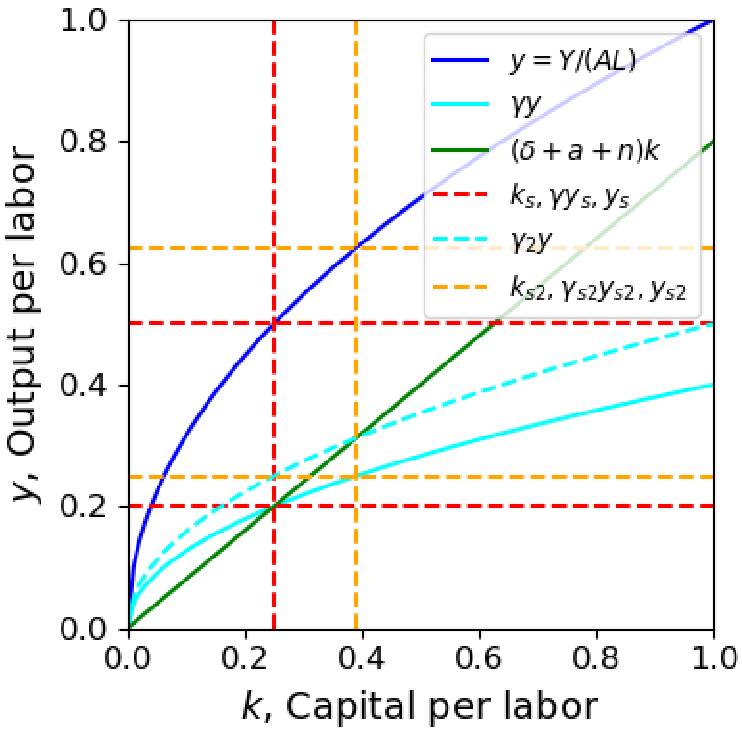

Figure 2.1: We see the Solow-Swan model. Here γ is the savings rate, γ_2 is the optimal savings rate for the most consumption, and subscript s quantities are the steady-state values. We see with our first savings rate γ that the consumption is not optimizied, whereas the k_{s2}, y_{s2} values for γ_{s2} yield the largest possible consumption.

$$\frac{\mathrm{d}\ln y}{\mathrm{d}t} = \left(\left[\frac{\mathrm{d}y}{\mathrm{d}k}\right]_{k=k_s} + \mathcal{O}((\Delta k^2))\right)\left(\gamma - [\delta + a + n]\frac{k_s + \Delta k}{y(k_s) + \Delta k\, y'(k_s)}\right) \tag{2.2.44}$$

$$\frac{\mathrm{d}\ln y}{\mathrm{d}t} = \left(\left[\frac{\mathrm{d}y}{\mathrm{d}k}\right]_{k=k_s} + \mathcal{O}((\Delta k^2))\right)\left(\gamma - [\delta + a + n]\frac{k_s + \Delta k}{y(k_s)}\left[1 - \Delta k\,\frac{y'(k_s)}{y_s} + \mathcal{O}((\Delta k)^2)\right]\right) \tag{2.2.45}$$

$$\frac{\mathrm{d}\ln y}{\mathrm{d}t} = \left[\frac{\mathrm{d}y}{\mathrm{d}k}\right]_{k=k_s}\left(\cancel{\gamma - [\delta + a + n]\frac{k_s}{y_s}} - \frac{[\delta + a + n]}{y(k_s)}\left[-\frac{k_s y'(k_s)}{y_s} + 1\right]\Delta k\right) \tag{2.2.46}$$

$$\frac{\mathrm{d}\ln y}{\mathrm{d}t} \approx -\frac{\Delta y}{\Delta k}\Delta k\,\frac{[\delta + a + n]}{y(k_s)}\left[1 - \frac{k_s y'(k_s)}{y_s}\right] \tag{2.2.47}$$

$$\frac{\mathrm{d}\ln y}{\mathrm{d}t} \approx -\lambda\frac{y - y(k_s)}{y(k_s)} \tag{2.2.48}$$

$$\lambda \approx [\delta + a + n]\left[1 - \frac{k_s}{y_s}y'(k_s)\right] \tag{2.2.49}$$

If we define $\alpha = y'(k_s)k_s/y_s$ as the elasticity of output to capital at the steady state with $y_s = y(k_s)$ then this becomes simply

$$\frac{\mathrm{d}\ln y}{\mathrm{d}t} \approx -\lambda\frac{y - y(k_s)}{y(k_s)} \tag{2.2.50}$$

$$\lambda \approx [\delta + a + n]\,[1 - \alpha] \tag{2.2.51}$$

Empirically, economists think that λ is fairly small, on the order of 0.01 to 0.05 when expressing things in terms of annual changes (so 1% to 5% in percent changes) meaning that it takes about 14 to 70 years to "close half the gap" or converge halfway given this model. That is, the Solow-Swan model claims that if all countries have the same production model, then there should be convergence within a couple of centureis of all country economies to the same per capita values. Clearly, this isn't really true as some are much faster than this (South Korea, Japan, China, etc.). Thus, as stated before, maybe other factors are important giving a different production function (perhaps other than K or AL, we need law system, property rights, and other institutions).

2.3 Factor Production

A reasonable assumption is that in real terms, that a competitive firm will pay workers a wage (that is hire enough workers with a wage) such that the marginal cost of another worker is equal to the marginal revenue (that is they will hire until a worker costs more than what they are paid to produce). That is the factor payment is the same as the marginal product. The real capital rental rate is given by $r = \frac{\partial Y}{\partial K} = y'(k)$ and the real (annual) wage is given by $\omega = \frac{\partial Y}{\partial L} = y(k) - ky'(k)$. One interesting effect is that this is true regardless of the size of K and L, only k really matters. That is, competitive firms and constant returns to scale (implies we have a homogeneous function) then we find factor payment must be equal to the marginal product. We get the same marginal product if K and L are scaled up together.

We can then think of the factor price frontier, which is all of the ways an economy could distribute outputs. A plot can be made of the "frontier" with ω on the x axis and r on the y axis. We saw that the neoclassical theory implies that factor payments equal marginal product, but we could imagine other payment options given that we are not going to be in such an equilibrium. Some may not work (not paying workers anything), but we recognize that any units given to labor is not profit and vice versa so there is a trade-off between investments in capital/savings and paying labor. That is given ω and r, we know that there is some sort of trade-off between these two so we get a curve that is convex in the sense that all points above the curve form a convex polygon.

We use that (let $A = 1$ for simplicity)

$$Y(K, L) = L\frac{\partial Y}{\partial L} + K\frac{\partial Y}{\partial K} \tag{2.3.1}$$

$$y(k) = \frac{\partial Y}{\partial L} + k\frac{\partial Y}{\partial K} = \omega + kr \tag{2.3.2}$$

We can of course write this as

$$\omega = y(k) - kr \tag{2.3.3}$$

We can then view this as function $\omega = \omega(r)$. Thus

$$\frac{\mathrm{d}\omega}{\mathrm{d}r} = \frac{\partial y}{\partial r} - \frac{\partial k}{\partial r}r - k \tag{2.3.4}$$

$$\frac{\mathrm{d}\omega}{\mathrm{d}r} = \frac{\partial y}{\partial k}\frac{\partial k}{\partial r} - \frac{\partial k}{\partial r}r - k = \frac{\partial k}{\partial r}\left(\frac{\partial y}{\partial k} - r\right) - k \tag{2.3.5}$$

we can then use that $\frac{\partial y}{\partial k} = r$ so that we find

$$\frac{\mathrm{d}\omega}{\mathrm{d}r} = -k \tag{2.3.6}$$

which is a fairly interesting and simple result.

We might think about how depreciation would possibly change this derivation. In that case, we know that k will be decreasing in time. What this then means is that there is no change in our equations. We simply have a different k and also a different r. However, it means that there are three different places output can be placed into. Thus output is divided between wages, capital, and depreciation so that we have

$$y = \omega + kr + \delta k = \omega + k(r + \delta) \tag{2.3.7}$$

Note that all we really have done is split the original r into two components, that for capital depreciation δ and that for no capital depreciation r. Hence, our equations do not change, just the interpretation of r. We simply recognize that $r_{\text{no }\delta} + \delta = r$. If δ is a constant, then it is especially simple and $\mathrm{d}r_{\text{no }\delta} = \mathrm{d}r$.

We can then plot this. An increase in technology means that the factor price frontier shifts upward and to the right since we can have a higher annual wage given a rental rate for capital (and vice versa). What this is actually saying is that if you choose a high wage, then the rental rate for capital is small (you are using your money on labor rather than on the capital). We also see that high wages are achieved via larger k. This means high wages are achieved when the capital stock to labor ratio is high. Essentially, you have low wages if you don't save and invest enough in capital stock or it is easier to employ labor rather than capital to do a lot of the work.

2.4 Overlapping Generations Model

This was developed by Allais (1947) and also by Samuelson (1958) and was originally a way of understanding fiat money value in equilibrium. That is, why do things get accepted as money that don't have intrinsic value. Diamond (1965) then applied it to saving, investment, and capital.

The model uses discrete time, with any period having two overlapping generations. The two generations in each period are a laborer and a capitalist. "Young people" are the laborers and "old people" are the capitalists. There is perfect knowledge maximizing utility. It has time-separable utility, so that the utility today doesn't affect consumption tomorrow. Essentially, this means that no one regrets their decisions even if things change in time. We use capital depreciation, a closed economy, no government, and full employment to endogenously determine factor prices for this simple model.

We use c_{1t} for consumption of young at time period t and c_{2t} for consumption of the old at time period t. Then s_t is the savings of the young at time period t, w_t is the wage and income of the young at time period t, δ is the capital depreciation, R_{t+1} is the gross rate of return in $t+1$ paid on savings in t ($R_{t+1} = 1 - \delta + y'(k)$ at market equilibrium), $r_{t+1} = 1 - R_{t+1}$ is the capital rate of return in period $t+1$, $\beta \leq 1$ is the private discount rate, θ is the central planner's discount rate, n is the growth rate of the population N or employed labor force L, and ξ_t is Benassy's time-varying social weighting factor. That is ξ_t weights different generations differently through time periods.

Firms optimize as before, hiring until factor price is the same as marginal product. Household saving is simply save enough to live off of. Thus, savings should be determined by

$$\max_{c_{1t}, c_{2,t+1}} \left(U(c_{1t}) + \beta U(c_{2,t+1})\right) \tag{2.4.1}$$

subject to a budget constraint of

$$c_{2,t+1} = (\omega_t - c_{1t})R_{t+1} \tag{2.4.2}$$

where U is a concave function of consumption. Larger β weights old age utility more.

One then takes ω_t and R_{t+1} as given with $U'(c) > 0$ and $U''(c) < 0$.

Economists have come up with the term comparative statics to describe the process of holding all variables but one constant and seeing how the system evolves. This concept is essentially looking at partial derivatives, allowing us to see the dynamics when changing only a single independent variable.

Using a Lagrange multiplier method simply shows that it results in the same answer as simple substitution (and that it more rigorously gets our answer without worrying about what the derivatives are with respect to). We start with

$$\max_{c_{1t}, c_{2,t+1}} \left[U(c_{1t}) + \beta U(c_{2,t+1}) + \lambda(c_{2,t+1} - [(\omega_t - c_{1t})R_{t+1}])\right] \tag{2.4.3}$$

We then take derivatives of the expression to be maximized with respect to $c_{1t}, c_{2,t+1}, \lambda$ and find

$$\frac{\partial U}{\partial c_{1t}} = \lambda R_{t+1} \tag{2.4.4}$$

$$\beta \frac{\partial U}{\partial c_{2,t+1}} = \lambda \tag{2.4.5}$$

$$(\omega_t - c_{1t})R_{t+1} = c_{2t} \tag{2.4.6}$$

Thus

$$\frac{1}{R_{t+1}}\frac{\partial U}{\partial c_{1t}} = \beta\frac{\partial U}{\partial c_{2,t+1}} \tag{2.4.7}$$

or equivalently

$$\frac{\partial U}{\partial c_{1t}} = \beta R_{t+1}\frac{\partial U}{\partial c_{2,t+1}} \tag{2.4.8}$$

We'd then like to know how things change as we change s_t, R_{t+1} and ω_t. To do so, we write a consumption/savings function as a function of ω_t and R_{t+1}. That is, we assert $s_t = S(\omega_t, R_{t+1})$ as a relationship. We then note that $\partial S/\partial\omega_t > 0$ but that $\frac{\partial S}{\partial R_{t+1}}$ is of either sign. That is, people save more if the wage increases but a better interest rate could induce more savings or more consumption (depending on how much savings you have already).

Note again that this is the simplest possible model. Economists have considered most of the more realistic complications and incorporated them into the model, so that this "model" should really be thought of as one variant of the OLG models. For example, uncertainties, regard for children, and environmental concerns have been considered as ways of extending and making the model more realistic.

Suppose we use $U(c) = \ln c$, then we note that at optimum we require

$$\frac{1}{c_{1t}} = \frac{\beta R_{t+1}}{c_{2,t+1}} \tag{2.4.9}$$

$$\frac{1}{\omega_t - s_t} = \frac{\beta R_{t+1}}{R_{t+1}s_t} \tag{2.4.10}$$

$$\omega_t - s_t = \frac{s_t}{\beta} \tag{2.4.11}$$

$$s_t\left(\frac{1}{\beta}+1\right) = \omega_t \tag{2.4.12}$$

$$s_t \equiv S(\omega_t, R_{t+1}) = \frac{\omega_t}{\frac{1}{\beta}+1} = \frac{\omega_t\beta}{1+\beta} \tag{2.4.13}$$

We can also consider (N_t is the population at time t)

$$K_{t+1} = N_t s_t = N_t S(\omega_t, R_{t+1}) \tag{2.4.14}$$

$$N_{t+1} = (1+n)N_t \tag{2.4.15}$$

$$k_{t+1} \equiv \frac{K_{t+1}}{N_{t+1}} = \frac{S(\omega_t, R_{t+1})}{1+n} \tag{2.4.16}$$

This means the return on capital is given by

$$R_{t+1} = 1 + y'(k_{t+1}) - \delta \tag{2.4.17}$$

So that we get a change from k and a deprecation of capital through δ. That is, the rate of return is the marginal return minus depreciation. Then

$$r_{t+1} = R_{t+1} - 1 = y'(k_{t+1}) - \delta \tag{2.4.18}$$

We can find the slope of a curve in $\omega(r)$ space and so find (use $\mathrm{d}R_{t+1} = \mathrm{d}r_{t+1}$)

$$\mathrm{d}r_{t+1} = y''(k_{t+1})\,\mathrm{d}k_{t+1} = y''(k_{t+1})\frac{\frac{\partial S}{\partial \omega_t}\,\mathrm{d}\omega_t + \frac{\partial S}{\partial R_{t+1}}\,\mathrm{d}R_{t+1}}{1+n} \tag{2.4.19}$$

$$\mathrm{d}r_{t+1} = y''(k_{t+1})\frac{\frac{\partial S}{\partial \omega_t}\,\mathrm{d}\omega_t + \frac{\partial S}{\partial R_{t+1}}\,\mathrm{d}r_{t+1}}{1+n} \tag{2.4.20}$$

or

$$\mathrm{d}r_{t+1}\left(1 - \frac{y''(k_{t+1})}{1+n}\frac{\partial S}{\partial R_{t+1}}\right) = \mathrm{d}\omega_t y''(k_{t+1})\frac{\partial S}{\partial \omega_t} \tag{2.4.21}$$

$$\frac{\mathrm{d}\omega_t}{\mathrm{d}r_{t+1}} = \frac{1 - \frac{y''(k_{t+1})}{1+n}\frac{\partial S}{\partial R_{t+1}}}{\frac{y''(k_{t+1})}{1+n}\frac{\partial S}{\partial \omega_t}} = \frac{1+n-y''(k_{t+1})\frac{\partial S}{\partial R_{t+1}}}{y''(k_{t+1})\frac{\partial S}{\partial \omega_t}} \tag{2.4.22}$$

We then see that $\mathrm{d}\omega_t/\mathrm{d}r_{t+1} < 0$ if $\partial S/\partial R_{t+1} > 0$ and $\partial S/\partial \omega_t > 0$ since $y''(k) < 0$ if y is a proper production function (for in this case the numerator is positive and the denominator is clearly negative). For $\partial S/\partial R_{t+1} < 0$ and $\partial S/\partial \omega_t > 0$ then we need $1+n > y''(k_{t+1})\frac{\partial S}{\partial R_{t+1}}$, and similarly for other cases. The other cases are not worth looking at since we require $\frac{\partial S}{\partial \omega_t} > 0$ for a sensible savings function. Simple models usually just assume $\partial S/\partial R_{t+1} > 0$ for simplicity.

We can then let values go to steady-state values to find the behavior over time. We can conceptualize this as

$$k_{t+1} = \frac{S(\omega_t, R_{t+1})}{1+n} = \frac{S\left(y(k_t) - k_t y'(k_t), 1 + y'(k_{t+1}) - \delta\right)}{1+n} \tag{2.4.23}$$

Then $t \to \infty$ implies $q_t \to q_{t+1} \to q_s$ for any quantity and so

$$k_s = \frac{S(y_s - k_s y'_s, 1 + y'_s - \delta)}{1+n} \tag{2.4.24}$$

Note that we can find ($y' = y'(k_t)$ for brevity)

$$\frac{\mathrm{d}k_{t+1}}{\mathrm{d}k_t} = \frac{\mathrm{d}S}{\mathrm{d}k_t}\frac{1}{1+n} = \frac{\frac{\partial S}{\partial \omega_t}\frac{\partial \omega_t}{\partial k_t} + \frac{\partial S}{\partial R_{t+1}}\frac{\partial R_{t+1}}{\partial k_t}}{1+n} \tag{2.4.25}$$

$$= \frac{\frac{\partial S}{\partial \omega_t}(\cancel{y' - y'} - k_t y'') + \frac{\partial S}{\partial R_{t+1}}\left(y''(k_{t+1})\frac{\mathrm{d}k_{t+1}}{\mathrm{d}k_t}\right)}{1+n} \tag{2.4.26}$$

$$\tag{2.4.27}$$

so that

$$\frac{\mathrm{d}k_{t+1}}{\mathrm{d}k_t}\left(1 - \frac{\frac{\partial S}{\partial R_{t+1}}y''(k_{t+1})}{1+n}\right) = \frac{\partial S}{\partial \omega_t}\frac{-k_t y''}{1+n} \tag{2.4.28}$$

$$\frac{\mathrm{d}k_{t+1}}{\mathrm{d}k_t} = \frac{-\frac{\partial S}{\partial \omega_t}k_t y''}{1+n-\frac{\partial S}{\partial R_{t+1}}y''(k_{t+1})} \tag{2.4.29}$$

which at steady-state then approaches

$$\left(\frac{\mathrm{d}k_{t+1}}{\mathrm{d}k_t}\right)_s = \frac{-\left(\frac{\partial S}{\partial \omega_t}\right)_s k_s y''_s}{1+n-\left(\frac{\partial S}{\partial R_{t+1}}\right)_s y''_s} \tag{2.4.30}$$

We can use that $y'' < 0$, with all partial derivatives of S being positive to see that the expression is positive. This can be read as

$$\left(\frac{\mathrm{d}k_{t+1}}{\mathrm{d}k_t}\right)_s = \frac{-k_s}{\left(\frac{1+n-\left(\frac{\partial S}{\partial R_{t+1}}\right)_s y''_s}{\left(\frac{\partial S}{\partial \omega_t}\right)_s y''_s}\right)} = \frac{(\mathrm{d}\omega/\mathrm{d}r)_s}{(\mathrm{d}\omega_t/\mathrm{d}r_{t+1})_s} \tag{2.4.31}$$

where $\frac{\mathrm{d}\omega}{\mathrm{d}r}$ is the slope of the FPF (factor price frontier) curve from (2.3.6), and $\frac{\mathrm{d}\omega_t}{\mathrm{d}r_{t+1}}$ is the market clearing condition from (2.4.22). For there to be such a steady state away from infinities, then $\mathrm{d}k_{t+1}/\mathrm{d}k_t < 1$ so that the FPF curve is flatter than the market clearing curve. This is simply saying that k as a function of t is not growing or (decreasing) exponentially. One has to check that the stability occurs.

Economists often prefer to look at logarithms of the sides (essentially they are normalizing the variables and so assuming that when we adjust for the different sizes of economies, the same dynamics are at work). Thuse we use

$$k_{t+1} = \frac{S(\omega_t, R_{t+1})}{1+n} \tag{2.4.32}$$

$$\ln k_{t+1} = \ln S(\omega_t, R_{t+1}) - \ln(1+n) \tag{2.4.33}$$

$$\ln k_s = \ln S_s - \ln(1+n) \tag{2.4.34}$$

If we assume we are near the steady state, we can find the dynamics near there with a Taylor series

$$\ln(k_s + \Delta k_{t+1}) = \ln S(\Delta k_t, \Delta k_{t+1}) - \ln(1+n) \tag{2.4.35}$$

$$\ln(k_s + \Delta k_{t+1}) = \ln(1 + \frac{\Delta k_{t+1}}{k_s}) + \ln(k_s) \tag{2.4.36}$$

$$\begin{aligned}\ln S(k_s + \Delta k_t, k_s + \Delta k_{t+1}) &= \ln\left(S(k_s, k_s) + \left(\frac{\partial S}{\partial \omega_t}\right)_s \left(\frac{\partial \omega_t}{\partial k_t}\right)_s \Delta k_t + \left(\frac{\partial S}{\partial R_{t+1}}\right)_s \left(\frac{\partial R_{t+1}}{\partial k_{t+1}}\right)_s \Delta k_{t+1}\right) \\ &= \ln\left(1 + \frac{1}{S_s}\left(\frac{\partial S}{\partial \omega_t}\right)_s (-k_s y''_s)\Delta k_t + \frac{1}{S_s}\left(\frac{\partial S}{\partial R_{t+1}}\right)_s y''_s \Delta k_{t+1}\right) + \ln S_s\end{aligned} \tag{2.4.37}$$

Note that this "Taylor Series" has slightly more stringent conditions. We cannot have $k \to 0$, and we require $\left(\frac{\partial S}{\partial \omega_t}\right)_s / S_s \Delta k$ (and its R_{t+1} analogue) to be well-defined and small. Let's define $\left(\frac{\partial \ln S}{\partial q}\right)_s \equiv \frac{1}{S_s}\left(\frac{\partial S}{\partial q}\right)_s$ for a quantity q for convenience.[6] We have then found (assuming the Δ quantities are small enough)

$$\ln(k_s + \Delta k_{t+1}) = \ln S(\omega_t(\Delta k_t), R_{t+1}(\Delta k_{t+1})) - \ln(1+n) \tag{2.4.38}$$

$$\frac{\Delta k_{t+1}}{k_s} + \cancel{\ln k_s} = \left(\frac{\partial - \ln S}{\partial \omega_t}\right)_s k_s y''(k_s)\Delta k_t + \left(\frac{\partial \ln S}{\partial R_{t+1}}\right)_s y''(k_{t+1})\Delta k_{t+1} + \cancel{\ln S_s - \ln(1+n)} \tag{2.4.39}$$

$$\frac{\Delta k_{t+1}}{k_s}\left(1 - k_s\left(\frac{\partial \ln S}{\partial R_{t+1}}\right)_s y''_s\right) = \left(\frac{\partial \ln S}{\partial \omega_t}\right)_s k_s y''_s \Delta k_t \tag{2.4.40}$$

[6]These are really the same expressions in general, but there could be ambiguity if one is unfamiliar with logarithmic derivatives.

which means

$$\frac{\Delta k_{t+1}}{k_s} = \frac{-\left(\frac{\partial \ln S}{\partial \omega_t}\right)_s k_s y_s''}{\left(1 - k_s \left(\frac{\partial \ln S}{\partial R_{t+1}}\right)_s y_s''\right)} \Delta k_t \tag{2.4.41}$$

$$\frac{\Delta k_{t+1}}{k_s} = \frac{-\left(\frac{\partial S}{\partial \omega_t}\right)_s k_s y_s''}{\left(S_s - k_s \left(\frac{\partial S}{\partial R_{t+1}}\right)_s y_s''\right)} \Delta k_t \tag{2.4.42}$$

$$\frac{\Delta k_{t+1}}{k_s} = \frac{-\left(\frac{\partial S}{\partial \omega_t}\right)_s k_s y_s''}{\left(1 + n - \left(\frac{\partial S}{\partial R_{t+1}}\right)_s y_s''\right)} \frac{\Delta k_t}{k_s} \tag{2.4.43}$$

We can now ask the question of whether the market clearing curve is the best that could be done. That is, could a clever central planner do a better job of optimizing utility? Or, equivalently, how could a market fail? The central planner can think about total utility across all generations (whereas markets may only consider the current generation). This requires controversial ideas, such as being able to add utility across generations and what should the weightings across generations should be, but it is worth considering.

Prof. Burda uses a model with discounting across time given by $1/(1+\theta)^t$ through ξ_t. The social planner then has the problem

$$\max_{\forall c_{1t}, c_{2t}, k_t} \beta U(c_{2,0}) + \sum_{t=0}^{T-1} \left[\xi_t \left(U(c_{1,t}) + \beta U(c_{2,t+1})\right)\right] \tag{2.4.44}$$

subject to the restraints $t = 0, \ldots, T$

$$c_{1,t} + \frac{c_{2,t}}{1+n} + (1+n)k_{t+1} = (1-\delta)k_t + y(k_t) \tag{2.4.45}$$

for initial k_0 and terminal condition $k_{T+1} = 0$ or $k_{T+1} = k$ for some k.

The initial $\beta U(c_{2,0})$ are the lucky "first" generation of old people in the model who didn't save initially. The use of β is a conventional factor put in front. The constraints are from the social planner requiring that consumption (properly weighted) is equal to the resources produced during that time. We use $\xi_t = 1/(1+\theta)^t$ so that $\theta > 0$ means the planner cares less about future generations and $\theta = 0$ means all generations are equally weighted. Note that this yields a geometric discount. Also, people are not weighted equally in this case (remember there are more people in future generations). To weight people equally, we require $\xi_t = (1+n)^t$ so $(1+n)^t = 1/(1+\theta)^t$ or $(1+n)(1+\theta) = 1^{-t} = 1$.

It is claimed by Professor Burda that substitution is easier to do to solve this equation, but I think Lagrange multipliers are usually more straightforward.

Our problem with Lagrange multipliers is given by

$$\begin{aligned} F(c_{1t}, c_{2t}, k_t, \lambda_t) = \beta U(c_{20}) + \sum_{t=0}^{T-1} \xi_t \left[U(c_{1t}) + \beta U(c_{2,t+1})\right] \\ - \sum_{t=0}^{T} \lambda_t \left[c_{1t} + \frac{c_{2t}}{1+n} + (1+n)k_{t+1} - (1-\delta)k_t - y(k_t)\right] \end{aligned} \tag{2.4.46}$$

with t indexing the set. Thus we need

$$\frac{\partial F}{\partial c_{1t}} = 0 \tag{2.4.47}$$
$$\frac{\partial F}{\partial c_{2t}} = 0 \tag{2.4.48}$$
$$\frac{\partial F}{\partial k_t} = 0 \tag{2.4.49}$$
$$\frac{\partial F}{\partial \lambda_t} = 0 \tag{2.4.50}$$

for our extremum. Thus

$$\xi_t U'(c_{1t}) - \lambda_t = 0 \tag{2.4.51}$$
$$\xi_{t-1}\beta U'(c_{2,t}) - \frac{\lambda_t}{1+n} = 0 \tag{2.4.52}$$
$$\lambda_{t-1}(1+n) - \lambda_t\left[(1-\delta) + y'(k_t)\right] = 0 \tag{2.4.53}$$

This means that we get

$$\xi_t U'(c_{1t}) = (1+n)\xi_{t-1}\beta U'(c_{2t}) \tag{2.4.54}$$
$$U'(c_{1t}) = (1+n)(1+\theta)\beta U'(c_{2t}) \tag{2.4.55}$$

and

$$\xi_{t-1}U'(c_{1,t-1})(1+n) = \xi_t U'(c_{1t})[1-\delta+y'(k_t)] \tag{2.4.56}$$
$$U'(c_{1,t-1})(1+n)(1+\theta) = U'(c_{1t})[1-\delta+y'(k_t)] \tag{2.4.57}$$
$$U'(c_{1t})(1+n)(1+\theta) = U'(c_{1,t+1})[1-\delta+y'(k_{t+1})] \tag{2.4.58}$$

reproducing the conditions given.

We can then substitute one into the other

$$U'(c_{1,t+1}) = (1+n)(1+\theta)\beta U'(c_{2,t+1}) \tag{2.4.59}$$
$$U'(c_{1,t+1}) = \frac{U'(c_{1t})(1+n)(1+\theta)}{1-\delta+y'(k_{t+1})} \tag{2.4.60}$$
$$\frac{U'(c_{1t})(1+n)(1+\theta)}{1-\delta+y'(k_{t+1})} = (1+n)(1+\theta)\beta U'(c_{2,t+1}) \tag{2.4.61}$$
$$U'(c_{1t}) = \beta[1-\delta+y'(k_{t+1})]U'(c_{2,t+1}) \tag{2.4.62}$$

this looks like our previous market conditions. Thus, we reproduce part of the solution. Thus, with $\beta R_{t+1} = 1-\delta+y'(k_{t+1})$, we have the same as for our market conditions before, we get an optimal solution given the current levels of stuff. We are not guaranteed however to get the best solution in aggregate terms. That is, the path to get to the optimum is different than the market way. Essentially the central planner can force the first generations to save to the correct level and give successive generations the optimal value of utility for the entire economy (the markets only do the optimum given the savings and capital levels given initially).

That is, the market will hit a steady state, but that steady state is not necessarily the optimal steady state possible, it just includes that possibility. That is the steady state market outcome is given by

$$\frac{U'(c_{1s})}{U'(c_{2s})} = \beta[1 - \delta + y'(k_s)] \tag{2.4.63}$$

at steady state we'd expect $c_{1s} = c_{2s}$ so that

$$y'(k_s) = \frac{1}{\beta} - 1 + \delta \tag{2.4.64}$$

There is no guarantee that k_s is the k_s satisfying the Solow-Swan like conditions that the best steady state $y'(k_s^*) = \theta + n + \delta$. This is simply saying that if people don't care about the future, then they will not take actions that would improve the next generations' lives if it goes against the current generation's interests. (For example, global warming would be a problem for future generations, but if one doesn't care about future generations, then one won't take actions against global warming. Then we aren't in an optimum situation for all generations.)

One way of connecting the model to the real world is to consider pensions systems. In the first type of pension system we will oconsider, it is fully-funded. We tax the younger generation by τ_t, invest it, and then the older generation lives off of these returns $R_{t+1}\tau_t$. The younger generation can still save so we are essentially just making $c_{1t} = \omega_t - s_t - \tau_t$ and $c_{2,t+1} = R_{t+1}(s_t + \tau_t)$. This clearly just introduces a new variable $q_t = s_t + \tau_t$ which reproduces the same equilibrium trajectories. The government does what the household (should?) would do anyway. This is tax the young, use the proceeds for the old in some way.

Another pension system idea is pay-as-you-go. Here the tax goes to the current older generation. It is not invested. The benefits are simply given out to the older generation. Then c_{1t} remains the same, and $c_{2,t+1} = R_{t+1}s_t + b_{t+1}$ but b_{t+1} is not just $R_{t+1}\tau_t$ as it was before. This causes aggregate savings to decline.[7] In the pay-as-you-go k_s decreases which may or may not be closer to k_s^*, the ideal k_s.

To reiterate, the optimal social planner and the decentralized market don't get the same solution in general. The decententralized market has at steady-state that $c_{1s} = c_{2s}$ so that

$$\frac{U'(c_{1s})}{U'(c_{2s})} = \beta[1 - \delta + y'(k_s)] \tag{2.4.65}$$

$$y'(k_s) = \frac{1}{\beta} + \delta - 1 \tag{2.4.66}$$

whereas the optimal planner would have chosen a different k, the best one k_b which is given when all generations' utility are important:

$$y'(k_b) = \theta + n + \delta \tag{2.4.67}$$

In general the decentralized $k_s \neq k_b$.

If we consider what will actually happen when compared against the optimal, there are really only two cases. When $r > \theta + n$ (equivalently, $k_b < k_s$), it is called dynamic efficiency. Future

[7] Wages are being reduced to give to the old, and the old have more to spend.

generations could be better off at the cost of reducing current consumption (instead of being redirected into capital accumulation). This requires a sacrifice from current generations. The other case is $r < \theta + n$ (equivalently, $k_b > k_s$), which is dynamic inefficiency. This is when there is too much saving/captial accumulation. So the current generation could improve further generations by actually consuming more![8] This may have happened in the USSR, where there was too much steel production and it wasn't being all used in consumption.

Thus, we see that introducing a pay-as-you-go system in a dynamically efficient situation leads to an even worse outcome, but in a dynamically inefficient situation leads to an improvement for succeeding generations. There are still risks, though, because population growth n and technological change a change the optimal k value, and so the pay-as-you-go may not be great for all time. Thus, it is better to use a diversity of complementary savings mechanisms so that one isn't protected against failures from any single system.

Another possibility is using government debt. A government can go in debt to spur more saving or more consumption. Say that they tax the old to fund debt for the young. This makes no change in the private intertemporal budget constraint, as it is simply moving money around a bit. But this is only good if the government doesn't waste tax, roll debt to unborn generations, etc. This also doesn't consider that goverments usually have better interest rates than that for private debt.

Ramsey (we will consider him next, but he was an economist and mathematician) asked how should we treat future unborn generations to find a social optimum. We can introduce something like this in our market model by having agents that care about their "children", the future generation. That is, we can use an agent that also cares about future generations' utilities. Specifically, we will have agents care about their "children", the next generation. So

$$V_t = U(c_{1t}) + \beta U(c_{2t}) + \frac{1}{1+\theta} V_{t+1} \tag{2.4.68}$$

is the new agent, which values the next generations' utility V_{t+1}. We keep the same constraints on the c_{it}.

$$c_{1t} + s_t = \omega_t + b_t \tag{2.4.69}$$

$$c_{2,t+1} + b_{t+1} = (\omega_t + b_t - c_{1t}) R_{t+1} \tag{2.4.70}$$

Here b_t is a bequest, and b_{t+1} is a decision for the agent (how much money to leave for the next generation).

If we let this go on recursively forever, we see that we get

$$V_t = \sum_{i=0}^{\infty} \left(\frac{1}{1+\theta} \right)^i \left[U(c_{1,t+i}) + \beta U(c_{2,t+1+i}) \right] \tag{2.4.71}$$

We can still think of maximizing V_t, and can use the recursive form with the budget constraints (choosing b_{t+1}) to find an optimum. We can think of it as

$$\widetilde{V}(b_t) = \max_{c_{1t}, c_{2,t+1}, b_{t+1}} \left[U(c_{1t}) + \beta U(c_{2t}) + \frac{1}{1+\theta} \widetilde{V}(b_{t+1}) \right] \tag{2.4.72}$$

[8]This is a sort of have your cake and eat it, too situation! You can go out and buy more, and it actually helps future generations.

with the constraints and $b_{t+1} \geq 0$. The key insight is that $\widetilde{V}_t$ is not really a function of time directly, but of b_t the bequests. We assume that agents at different times will act the same if they have the same bequest b_t. It can then be shown that $\widetilde{V}_t$ is an increasing concave function of b_t. Here $\beta^{-1} - 1$ does not necessarily need to equal θ. We can then show that this leads to the socially optimum of the planner from before.

The above can be considered a Lagrange multiplier problem with the constraints as The constraint is simply the $c_{2,t+1}$ equation.

$$\widetilde{V}(b_t) = U(c_{1t}) + \beta U\left(c_{2,t+1}\right) + \frac{1}{1+\theta}\widetilde{V}(b_{t+1}) + \lambda(c_{2,t+1} + b_{t+1} + [-\omega_t - b_t + c_\tau]R_{t+1}) \tag{2.4.73}$$

and so

$$\frac{\partial \widetilde{V}(b_t)}{\partial c_{1t}} = U'(c_{1t}) + \lambda R_{t+1} = 0 \tag{2.4.74}$$

$$\frac{\partial \widetilde{V}(b_t)}{\partial c_{2t}} = \beta U'(c_{2,t+1}) + \lambda = 0 \tag{2.4.75}$$

$$\frac{\partial \widetilde{V}(b_t)}{\partial b_{t+1}} = \frac{\widetilde{V}'(b_{t+1})}{1+\theta} + \lambda = 0 \tag{2.4.76}$$

This implies

$$\frac{U'(c_{1t})}{R_{t+1}} = \frac{\widetilde{V}'(b_{t+1})}{1+\theta} \tag{2.4.77}$$

$$\frac{U'(c_{1t})}{R_{t+1}} = \beta U'(c_{2,t+1}) \tag{2.4.78}$$

Note that if we consider b_t a separate variable, we also get

$$\frac{\partial \widetilde{V}(b_t)}{\partial b_t} = \widetilde{V}'(b_t) = -R_{t+1}\lambda = U'(c_{1t}) \tag{2.4.79}$$

Thus

$$U'(c_{1t}) = \frac{R_{t+1}\widetilde{U}'(c_{1,t+1})}{1+\theta} \tag{2.4.80}$$

$$U'(c_{1t}) = R_{t+1}\beta U'(c_{2,t+1}) \tag{2.4.81}$$

which is our familiar condition for the social planner's optimum! A market with people caring about future generations can find the optimal solution!

2.5 Ramsey Model

The last example was a Ramsey-like Model. This will extend thinking about the time and growth of economies. We are going to use continuous time. We will be choosing an optimal function rather than a series of optimal values (like the discrete time analogue before). This will explicitly use physics concepts like a Hamiltonian and a phase space.

When we do this, we ask the question how much should a nation save, and our discrete to continuous means we cans switch from discount θ to $\exp(-\theta t)$. We still use no government, closed economy (full altruism), and no overlapping generations (that is, no disconnected generations for a closed economy with no new migrants).

We still maximize an objective function with future discounting, but over a set of functions. This is the calculus of variations or dynamic optimization. They call it Pontryagin's Maximum Principle. Economists consider state variables and costate variable, because they use a different form of the Hamiltonian that is analogous but not exactly the same as that used in physics. That is, their Hamiltonian is of the form $H(\mathbf{q}, \mathbf{p}, \mathbf{u}, t)$ instead of $H(\mathbf{q}, \mathbf{p}, t)$, but the ideas are the same. H is a Legendre transform of a Lagrangian that is different than ones that come up in physics, is all. State variables are not under control, but costate variables are and so are often called control variables. The costate variables can be thought of as affecting state variables. Initial conditions are, of course, very important for the dynamics, as well.

We have a population L [economists like using Q_t to indicate the function $Q(t)$, and also use Newton dot notation for time derivatives], and so have

$$L_t = L_0 \exp(nt) \dot{L}_t = \frac{\mathrm{d}L_t}{\mathrm{d}t} = N L_t \tag{2.5.1}$$

The capital K_t is given by

$$\dot{K}_t = Y_t - C_t - \delta K_t \tag{2.5.2}$$

where $Y_t = F(K_t, A_t L_t)$ is a production function with constant returns in K and L and normalize $A = 1$ (no technological progress) so we can go back to $y_t(k)$ with $k = K_t/L_t$ and $y_t(k) > 0$, $y_t'(k) > 0$, and $y_t''(k) < 0$.

We let $u(c_t)$ be the utility given consumption c_t requiring $u' > 0$ and $u'' < 0$. Then the total utility over time is

$$\int_0^\infty \mathrm{d}t\ \exp(-\theta t) u(c_t) \tag{2.5.3}$$

If we were to view this as weighting people with ρ with $N = L$ we'd use

$$\int_0^\infty \mathrm{d}t\ N \exp(-\rho t) u(c_t) = \int_0^\infty \mathrm{d}t\ N_0 \overbrace{\exp(nt)\exp(-\rho t)}^{\exp(-\theta t)} u(c_t) = N_0 \int_0^\infty \mathrm{d}t\ \exp(-\theta t) u(c_t) \tag{2.5.4}$$

Ramsey originally thought $\theta = \rho - n > 0$ was an immoral choice, but mathematically $\theta \leq 0$ is problematic as the integral generally diverges.

This is the equivalent of the action integral in Hamiltonian classical mechanics. We want to make this functional maximum with respect to c_t. That is we desire

$$\max_{c_t} \int_0^\infty \mathrm{d}t\ \exp(-\theta t) u(c_t) \tag{2.5.5}$$

with the constraints

$$\dot{k}_t = \frac{\dot{K}_t}{L_t} - \dot{L}_t \frac{K_t}{L_t^2} = \frac{\dot{K}_t}{L_t} - n L_t \frac{k_t}{L_t} = \frac{Y_t - C_t - \delta K_t}{L_t} - n k_t = y(k_t) - c_t - (n+\delta) k_t \tag{2.5.6}$$

and an initial prescribed k_0.

This problem is such that we can form a Hamiltonian function. The way to see that is that we have an extremal function, so that our integral expression is essentially Hamilton's principle of least action, and so we can form the Hamiltonian function in the usual way. This Hamiltonian is given by

$$H(c_t, k_t, \mu_t) = \exp(-\theta t)u(c_t) + \mu_t \left[f(k_t) - c_t - (n+\delta)k_t\right] \tag{2.5.7}$$

where μ_t is a costate variable. This is similar to forming a Lagrange multiplier. It is useful to write $\mu_t = \exp(-\theta t)\lambda_t$ so that we have

$$H(c_t, k_t, \lambda_t) = \exp(-\theta t)\left\{u(c_t) + \lambda_t \left[f(k_t) - c_t - (n+\delta)k_t\right]\right\} \tag{2.5.8}$$

The λ_t is the current value costate variable.

A "recipe" is then given for maximization. It involves $\frac{\partial H}{\partial p} = 0$ for each p costate (or control) variable and $\frac{\partial H}{\partial q} = -p$ for each q state variable, impose a transversality condition at $t = \infty$ which means that you don't save for infinite time, you actually consume what you are saving in a finite time horizon.

The way of getting here, is actually fairly interesting, so I will delve into it. Our original problem can be rewritten as

$$I = \int_0^\infty \mathrm{d}t \ \left[f(\mathbf{q}, \mathbf{u}, t) - \boldsymbol{\lambda}(t) \cdot \left[\mathbf{g}(\mathbf{q}, \mathbf{u}, t) - \dot{\mathbf{q}}\right]\right] \tag{2.5.9}$$

where $\mathbf{g}(\mathbf{q}, \mathbf{u}, t) = \dot{\mathbf{q}}$ is the constraint equation. We note that we can then integrate the last $\boldsymbol{\lambda}(t) \cdot \dot{\mathbf{q}}$ term by parts to find

$$I = \int_0^\infty \mathrm{d}t \ \left[f(\mathbf{q}, \mathbf{u}, t) - \boldsymbol{\lambda}(t) \cdot \mathbf{g}(\mathbf{q}, \mathbf{u}, t) + \dot{\boldsymbol{\lambda}} \cdot \mathbf{q}\right] - \left[\boldsymbol{\lambda}(t) \cdot \mathbf{q}(t)\right]_{t=0}^\infty \tag{2.5.10}$$

We must also have

$$0 = \delta I = \int_0^\infty \mathrm{d}t \ \left[\frac{\partial f}{\partial \mathbf{q}} \cdot \mathrm{d}\mathbf{q} + \frac{\partial f}{\partial \mathbf{u}} \cdot \mathrm{d}\mathbf{u} - \left[\frac{\partial \mathbf{g}}{\partial \mathbf{q}} \cdot \boldsymbol{\lambda} \cdot \mathrm{d}\mathbf{q} + \frac{\partial \mathbf{g}}{\partial \mathbf{u}} \cdot \boldsymbol{\lambda} \cdot \mathrm{d}\mathbf{u}\right] + \dot{\boldsymbol{\lambda}} \cdot \mathrm{d}\mathbf{q}\right] - \left[\boldsymbol{\lambda}(t) \cdot \mathrm{d}\mathbf{q}\right]_{t=0}^\infty \tag{2.5.11}$$

Each coefficient of $\mathrm{d}\mathbf{q}$ or $\mathrm{d}\mathbf{u}$ must be zero, thus

$$\frac{\partial f}{\partial \mathbf{q}} + \frac{\partial \mathbf{g}}{\partial \mathbf{q}} \cdot \boldsymbol{\lambda} + \dot{\boldsymbol{\lambda}} = 0 \Rightarrow \frac{\partial f}{\partial q_i} + \sum_j \frac{\partial g_j}{\partial q_i}\lambda_j + \dot{\lambda}_j = 0 \tag{2.5.12}$$

$$\frac{\partial f}{\partial \mathbf{u}} + \frac{\partial \mathbf{g}}{\partial \mathbf{u}} \cdot \boldsymbol{\lambda} = 0 \Rightarrow \frac{\partial f}{\partial u_i} + \sum_j \frac{\partial g_j}{\partial u_i}\lambda_j = 0 \tag{2.5.13}$$

and if $\mathbf{q}(\infty)$ and $\mathbf{q}(0)$ is fixed, we are done, but if they are not then we must also require $\boldsymbol{\lambda}(\infty) = 0$ to enforce the optimality condition. This is the so-called transversality condition (this may be called transversality because $\boldsymbol{\lambda} \cdot \mathrm{d}\mathbf{q}$ must be zero and so $\boldsymbol{\lambda}$ is "orthogonal" or transverse to $\mathrm{d}\mathbf{q}$).

With

$$H \equiv f(\mathbf{q}, \mathbf{u}, t) + \boldsymbol{\lambda} \cdot \left[\mathbf{g}(\mathbf{q}, \mathbf{u}, t) - \dot{\mathbf{q}}\right] \tag{2.5.14}$$

we see that the conditions

$$\frac{\partial H}{\partial \mathbf{q}} = -\dot{\boldsymbol{\lambda}} \tag{2.5.15}$$

$$\frac{\partial H}{\partial \mathbf{u}} = \mathbf{0} \tag{2.5.16}$$

reproduces these optimality conditions.

Let's do this for our $\mathbf{q} = [k_t]$ and $\mathbf{u} = [c_t]$ with

$$H(c_t, k_t, \mu_t) = \exp(-\theta t)\left\{u(c_t) + \lambda_t\left[y(k_t) - c_t - (n+\delta)k_t\right]\right\} \tag{2.5.17}$$

Then

$$\frac{\partial H}{\partial c_t} = u'(c_t) - \lambda_t = 0 \tag{2.5.18}$$

$$\frac{\partial H}{\partial k_t} = \exp(-\theta t)\lambda_t[y'(k_t) - (n+\delta)] = -\frac{\mathrm{d}\mu_t}{\mathrm{d}t} = -\exp(-\theta t)\left[-\theta\lambda_t + \dot{\lambda}_t\right] \tag{2.5.19}$$

and transversality implies $\lim_{t\to\infty}\mu_t = 0$. This can be rewritten as

$$u'(c_t) = \lambda_t \tag{2.5.20}$$

$$\lambda_t[y'(k_t) - (n+\delta)] = -\dot{\lambda}_t + \theta \tag{2.5.21}$$

$$\lim_{t\to\infty}\exp(-\theta t)\lambda_t = 0 \tag{2.5.22}$$

In fact, the economics transversality condition enforces $\lim_{t\to\infty} k_t\mu_t = 0$ rather than what I used in the general case. This shows us that λ_t is the marginal utility of consumption in the first equation, then how the marginal product of capital changes given growth factors, and finally the transversality condition that eliminates solutions where we get utility from the far future. Dr. Bund calls the transversality condition the "cake-eating" condition because it means that you have to eat the cake at some time to get the utility.

We can then rewrite this as

$$\dot{c}_t = -\frac{u'(c_t)}{u''(c_t)}\left[y'(k_t) - (n+\theta+\delta)\right] \tag{2.5.23}$$

$$\dot{k}_t = y(k_t) - c_t - (n+\delta)k_t \tag{2.5.24}$$

The $-u'/u''$ is almost the elasticity of (intertemporal) substitution, how ready we are to trade off consumption today and consumption of the future. The actual quantity that is the elasticity of substitution is $\sigma = -cu''/u'$ and $\sigma = \infty$ means completely elastic (willing to move consumption between time completely) whereas $\sigma = 0$ means completely inelastic (not willing at all to move consumption between time).

The Keynes-Ramsey rule is simply that first equation, normalized and with σ in it

$$\frac{\dot{c}_t}{c_t} = \frac{1}{\sigma}\left[y'(k_t) - (n+\theta+\delta)\right] \tag{2.5.25}$$

In words, it says that consumption per capita increases, remains constant, or decreases depending upon whether the marginal product of capital, net of population growth, and depreciation exceeds, is equal to, or is less than the rate of time preference. The rate of increase of consumption depends on σ.

(A quick note on economic terminology, they call the conditions we derive FONCs, or first order necessary conditions.)

We can easily find the steady state values

$$y_s'(k_s) = n + \theta + \delta \tag{2.5.26}$$
$$y_s(k_s) = c_s + (n + \delta)k_s \tag{2.5.27}$$

The first equation has θ which is "Ramsey's impatience". Note one can then analyze the direction in a phase-space diagram (alternatively, simply solve the differential equations above). This will give us two quadrants that are impossible based on actual economics in the phase space dynamics. These are rising consumption and falling capital and falling consumption and rising capital, the "transversality" conditions. This is essentially only allowing one to go along a saddle point in one way. Thus it is called a saddle path or stable manifold.

The plots make comparative statics/dynamics much easier. Statics means we only care about the steady state, whereas dynamics cares about the path (only allowing saddle stable paths). We can control θ pretty clearly, but we can also think about δ and n or even A.

A decline in θ implies a change in c_s (it shifts the c down, and so it gets carried by the phase space dynamics to the right). When θ is smaller, means we are more patient and hence consume less.

Suppose δ increases, so more depreciation. Then both c_s and k_s are affected. Larger δ should lead to a drop in consumption, as we have to save more to try to keep the same amount of stuff. It also should harm the k_s capital stock so that we are deflected down and to the left. Then we have to redraw the curves and see where the phase space dynamics leads us.

Finally, consider A increases, so there's been a technological development that improves productivity. Clearly we can now consume more. This leads to a deflection upwards and the phase space dynamics will pull us towards more capital stock k, to the right. Note that we replace $y'(k_t)$ with $Ay'(k_t)$ to get this result.

2.5.1 Market Ramsey Model

Now we can try decentralizing the Ramsey model to see what markets might get to. We use a representative consumer. This will be similar to the OLG model. It turns out that if we use a private discount rate ρ and $\rho = \theta$, we get the same solution as the social planner's optimum.

We now include ω_t the wage paid to each person who supplies labor inelastically. We will put θ as the private and central planner's discount rate. We also include R_t, the gross return paid to owners in t for capital services deriving from capital stock before depreciation, and $r_t = R_t - \delta$ the net rate of return on capital in t or real interest rate in t.

The new problem is maximize

$$\int_0^\infty \mathrm{d}t \ \exp(-\theta t)u(c_t) \tag{2.5.28}$$

where wealth evolves under

$$\dot{k}_t = \frac{\dot{K}_t}{L_t} - nk_t = \frac{\omega_t L_t + R_t K_t - C_t - \delta K_t}{L_t} - nk_t = \omega_t + R_t k_t - c_t - (n+\delta)k_t \tag{2.5.29}$$

which can be rewritten as

$$\dot{k}_t = \omega_t + (r_t - n)k_t - c_t \tag{2.5.30}$$

with an initial k_0.

We form the Hamiltonian

$$H = \exp(-\theta t)u(c_t) + \mu_t(\omega_t + (r_t - n)k_t - c_t) \tag{2.5.31}$$

Then

$$\frac{\partial H}{\partial k_t} = \dot{\mu}_t = (r_t - n)\mu_t = -\exp(-\theta t)\left[-\theta\lambda_t + \dot{\lambda}_t\right] \tag{2.5.32}$$

$$\frac{\partial H}{\partial c_t} = \exp(-\theta t)u'(c_t) - \mu = 0 \tag{2.5.33}$$

We can rewrite these as

$$\begin{aligned}(r_t - n)\lambda_t &= \theta\lambda_t - \dot{\lambda}_t \\ \frac{\dot{\lambda}_t}{\lambda_t} &= \theta - (r_t - n) = \theta + n - r_t = \theta + n + \delta - R_t\end{aligned} \tag{2.5.34}$$

$$u'(c_t) = \lambda_t \tag{2.5.35}$$

The transversality condition applies again $\lim_{t\to\infty} k_t\mu_t = 0$.

Thus, we find

$$\frac{\dot{c}_t}{c_t} = -\frac{u'(c_t)}{c_t u''(c_t)}\left[R_t - (n + \delta + \theta)\right] \tag{2.5.36}$$

$$\dot{k}_t = \omega_t - c_t + (R_t - n - \delta)k_t \tag{2.5.37}$$

Now we have exchanged ω_t for $y(k_t)$ and R_t for $y'(k_t)$ from the Ramsey model. We then have a private Ramsey-Keynes rule

$$\frac{\dot{c}_t}{c_t} = \frac{1}{\sigma}\left[R_t - (n + \delta + \theta)\right] \tag{2.5.38}$$

with σ the elasticity of intertemporal substitution.

This means that the (well-)planned economy and the market economy have similar dynamics. The First and Second Fundamental Welfare theorems say a market will tend toward a competitive equilibrium that is weakly Pareto optimal and out of all possible Pareto optimal outcomes one can achieve any particular one by enacting lump-sum wealth redistribution and then letting the market take over. This will fail if there are no constant returns to production, if there is not perfect competition, if there are no externalities, and if there is not full information for all agents.

This means that low real interest rates can have a clear interpretation if we assume the planner's interest rate is related to the private interest rate. We must have

$$r_t = y'(k_t) - \delta = \theta + n + a \tag{2.5.39}$$

where a is technological change. Then low r_t implies $\theta+n+a$ must all be low. It is not, necessarily, the central bank's fault for having low real interest rates.

Ricardian equivalence then says so long as taxation doesn't affect agents' incentives at the margin with government and private households at the same discount rate, the temporal path of taxes is irrelevant for agents' decisions. Only the present value of the government's expenditures affects the general equilibrium. Even with a more complicated discount factor, then the budget constraint is shifted back by the present value of the government's total resource claim.

Essentially, this is saying that agents take into account the government's tax ideas.

Last, notice that the modern Ramsey model is now called the Ramsey-Cass-Koopmans model.

2.6 Money

We start with David Hume's idea, that money is the oil that greases the wheels of trade. So "it is evident that the greater or less plenty of money is of no consequence." William Stanley Jevons said money has the properties of being

1. Medium of Exchange
2. Standard of Payment
3. Store of Value
4. Means of Deferred Payment

Then Groucho Marx's "While money can't buy happiness, it certainly lets you choose your own form of misery."

So why does money matter? The competition between money being a medium of exchange and a store of value. So money is a dominated asset that is generally accepted as a means of payment. So it has a liquidity premium, which is a fragile thing; it requires people to be willing to accept the money. Five prominent money models

CIA Cash-in-Advance models

OLG Overlapping-Generations models

MUF Money in the Utility Function

VC Money Velocity Costs in the budget constraint

LP Limited Participation

The CIA approach or Cambridge equation can be described as the following. Money is required to conduct transactions. If you want Y goods in each period at price P, then the demand for money is PY. We then use that GDP is the final purchase of goods and services as one way of measuring this (note that GDP is only final goods, so doesn't cover all the other transactions that don't produce GDP but require money) and use the Cambridge equation given by

$$MV = PY \tag{2.6.1}$$

with M the money supply and V is the velocity of money: how many times the money is spent on GDP in a period. When V is about constant, this gives us the interesting insight that $M = PY/V$ and so the price level can be inferred from the amount of money. That is $P = MV/Y$. We can then take logarithms, and take a time derivative to find

$$\ln M + \ln V = \ln P + \ln Y \tag{2.6.2}$$

$$\frac{\mathrm{d}\ln M}{\mathrm{d}t} + \frac{\mathrm{d}\ln V}{\mathrm{d}t} = \frac{\mathrm{d}\ln P}{\mathrm{d}t} + \frac{\mathrm{d}\ln Y}{\mathrm{d}t} \tag{2.6.3}$$

$$\mu + 0 = \pi + g \tag{2.6.4}$$

where π is the inflation rate in a period, g is the real economic growth in a period, and μ is the money growth in a period. We assumed $\frac{\mathrm{d}V}{\mathrm{d}t} = 0$, so there is no corresponding quantity. This is called monetary neutrality, where the velocity of money is stable. In the long run, it is well attested.

We return now to OLG, with Allais and Samuelson. Here we talk about fiat money, an asset that is an asset because others believe it or trust it to have real value in the future. Now the young use the money as a vehicle for saving. The old use money to trade for consumption when their income is low or zero. We again use the utility of the N_t households of the young generation in t is

$$U(c_{1,t}) + \beta U(c_{2,t+1}) \tag{2.6.5}$$

with utility U satisfying $U' > 0$, $U'' < 0$ and population growth at rate n. One then assumes there is no productive capital, so can only store value with money. Note that if there was no store, then old people would simply starve as they would have nothing. An improvement is sharing resources across generations. Thus, money is a good way of doing so. Even if the goods have some storage across generations, it is not optimal to do only that; some money is necessary. We now have a budget constraint involving money for households. We have $P_t(1 - c_{1,t}) = M_t^D$ and $P_{t+1}(c_{2,t+1}) = M_t^D$ where P_t is the price and M_t^D is the demand for money at time t. So long as $P_t < \infty$ then money will also have value to the young (in the initial money given generation). We imagine that the old are given H money. The FONC becomes

$$-\frac{U'(c_{1,t})}{P_t} + \frac{U'(c_{1,t+1})}{P_{t+1}} = 0 \tag{2.6.6}$$

Thus there is a demand for money depending on prices. So

$$\frac{M_t}{P_t} = L(P_t/P_{t+1}) \tag{2.6.7}$$

$$L' \gtreqless 0 \tag{2.6.8}$$

Then P_t/P_{t+1} is the gross deflation rate. Here L is a savings function. The sign is indeterminate. If there is inflation, then money is a "hot potato" because it buys less in the future.

We can then define a money market equilibrium so that supply equals demand. The real supply for the old is H/P_t in period t. The demand by the young people at the same period is $N_{t+1}L\left(P_t/P_{t+1}\right)$. So

$$\frac{H}{P_t} = N_{t+1}L(P_t/P_{t+1}) \tag{2.6.9}$$

$$\frac{H}{P_{t+1}} = N_{t+2}L(P_{t+1}/P_{t+2}) \tag{2.6.10}$$

$$\frac{P_{t+1}}{P_t} = \frac{N_{t+1}L(P_t/P_{t+1})}{N_{t+2}L(P_{t+1}/P_{t+2})} \tag{2.6.11}$$

we can then use $N_{t+1}/N_{t+2} = 1/(1+n)$ and so

$$\frac{P_{t+1}}{P_t}(1+n) = \frac{L(P_t/P_{t+1})}{L(P_{t+1}/P_{t+2})} \tag{2.6.12}$$

This means that because there are more people, the price must fall given the constant supply of H. So there's deflation. One equilibrium state would be that the ratio of prices between periods is constant

$$\frac{P_t}{P_{t+1}} = \frac{P_{t+1}}{P_{t+2}} \tag{2.6.13}$$

which means the ratio of L's is a constant. This means

$$\frac{P_t}{P_{t+1}} = 1+n \tag{2.6.14}$$

This says that steady state inflation is given by $-n$. If n is negative (falling population) then one should expect inflation. Constant supply things will generally deflate, but to get inflation in this model, the money supply must increase (assuming positive n).

Remember that for this to work, everyone must expect money to be valuable infinitely into the future. Back to storage, if there is a rate of return r, then so long as $r < n$ there are advantages to using money. When $r > n$ then money isn't useful.

Let's replace H with M_t so the government decides to print money. Say they give new cash in the amount $\mu_{t+1}M_t$ to the old people of each period. Now the price constraints are $P_t(1-c_{1,t}) = M_t^D$ but $P_{t+1}c_{2,t+1} = M_t^D(1+\mu_{t+1})$. Here μ_{t+1} is exogenous, decided by the government. We have the same FONC, but

$$\frac{M_t}{P_t} = N_{t+1}L(P_t/P_{t+1}, \mu_{t+1}/P_{t+1}) \tag{2.6.15}$$

$$\frac{(1+\mu_{t+1})M_t}{P_{t+1}} = N_{t+2}L(P_{t+1}/P_{t+2}, \mu_{t+2}/P_{t+2}) \tag{2.6.16}$$

so money demand depends on yield on money and real income. If we assume that the rates are the same again, so that $P_t/P_{t+1} = P_{t+1}/P_{t+2}$ and $\mu_{t+1}/P_{t+1} = \mu_{t+2}/P_{t+2}$ then

$$\frac{P_{t+1}}{P_t}(1+n) = (1+\mu_{t+1})\frac{L(P_t/P_{t+1}, \mu_{t+1}/P_{t+1})}{L(P_{t+1}/P_{t+2}, \mu_{t+2}/P_{t+2})} \tag{2.6.17}$$

where the ratio of L's is one again and so

$$(1+\pi)(1+n) = (1+\mu) \tag{2.6.18}$$

when $\pi n \ll 1$ then $\pi + n \approx \mu$. So inflation is the rate of growth of the money supply (μ) minus the growth rate of the population.

Money demand is a function of the price of goods in terms of money tomorrow. In the future we can ask why inflation happens in these models. In addition, we will look at how to stop inflation. Expectations about the future are an important aspect to this.

To do so, we will consider the Cagan model. We'll have a discontinuous and continuous model. We pay attention to money because we want to avoid hyperinflation. It is not always easy to defeat hyperinflation, and so credibility and expectations are very important.

We start with logarithm of the money demand

$$\ln M_t - \ln P_t = \ln Y_t - \eta i_t \tag{2.6.19}$$

here η is the semi-elasticity of money demand with the interest rate i_t. Y_t is production (income). Thus money demand depends on income (positively) and the nominal interest rate (negatively). We will use $\ln Q_t = q_t$ as shorthand. We then use the Fisher relation

$$i_t = r_t + \pi_t^e = r + \frac{\Delta P^{\,e}_{\,t+1}}{P_t} \approx r + p^e_{t+1} - p_t \tag{2.6.20}$$

where the superscript e means expected. Thus, the nominal interest rate includes expectations about future inflation. We can then take differentials, which the economists call differences (even though they use calculus rules, and so the Δ would only be correct when the difference is approaching zero). Thus

$$\Delta m_{\,t} - \Delta p_{\,t} = \Delta y_{\,t} - \eta \Delta i_{\,t} \tag{2.6.21}$$

(so assuming η is constant) and with $i_t = r_t + \pi_t^e$ we get

$$\Delta m_{\,t} - \Delta p_{\,t} = \Delta y_{\,t} - \eta(\Delta r_{\,t} + \Delta \pi^{\,e}_{\,t}) \tag{2.6.22}$$

Note that when $\Delta i_{\,t} = 0$ we simply have

$$\Delta m_{\,t} - \Delta y_{\,t} = \Delta p_{\,t} \tag{2.6.23}$$

which is what we had derived before. Essentially, the money growth rate minus the real growth rate gives us the inflation rate.

The first lesson then is that monetary collapse is driven by expectations of future inflation. This creates a larger interest rate, and so people hold on to more money, a self-fulfilling prophecy.

Let's set $r = 0$ for easy analysis and set $y = 0$, as well. Then we have

$$m_t - p_t = -\eta(p^e_{t+1} - p_t) \tag{2.6.24}$$

$$p_t = \frac{m_t + \eta p^e_{t+1}}{1+\eta} \tag{2.6.25}$$

This means p_t is endogenous in this model. We can then use this for multiple times with $p^e_{t+1} \to p_{t+1}$ and find

$$p_{t+1} = \frac{m_{t+1} + \eta p^e_{t+2}}{1+\eta} \tag{2.6.26}$$

$$p_t = \frac{m_t + \eta p^e_{t+1}}{1+\eta} \to \frac{m_t + \eta \frac{m_{t+1}+\eta p^e_{t+2}}{1+\eta}}{1+\eta} = \frac{m_t}{1+\eta} + \frac{\eta m_{t+1} + \eta^2 p^e_{t+2}}{(1+\eta)^2} \tag{2.6.27}$$

Continuing this we'd find

$$p_t = \sum_{j=0}^{N} \frac{\eta^j m_{t+j}}{(1+\eta)^{j+1}} + \frac{\eta^N p^e_{t+N+1}}{(1+\eta)^N} \overset{N\to\infty}{\Longrightarrow} + \sum_{j=0}^{N} \frac{\eta^j m_{t+j}}{(1+\eta)^{j+1}} \tag{2.6.28}$$

which if m_{t+j} is small enough, is a geometric series (or less than one) and so a finite number. The class factors a $(1+\eta)$ out so that

$$(1+\eta)p_t = \sum_{j=0}^{N} \left(\frac{\eta}{1+\eta}\right)^j m_{t+j} + \frac{\eta^{N+1} p^e_{t+N+1}}{(1+\eta)^N} \tag{2.6.29}$$

This simply says that the price level today is a function of expected future money supplies. Now we need to think about how people's expectations change. Where does money supply growth come from? Often from fiscal deficits monetized by central bank, large current account surpluses financed by money creation, wars, etc. Money growth may be driven by higher inflation (the Tanzi effect, inability to collect taxes, can force the government to print more money).[9] So fiscal deficits can cause money growth.

Now we can go to the continuous time model. We then choose

$$M_t/P_t = \exp\left(-\eta \pi^e_t\right) \tag{2.6.30}$$

and in the long run $\pi^e_t = \pi_t$. We use zero output growth $g = 0$ so that $\gamma = \mu$ in steady state (rather than $\gamma < \mu - g$). Then the rate of money growth is given by $M/P = \exp(-\eta\mu)$. We need to explain μ. We take the logarithmic derivatives and find

$$\frac{\mathrm{d}\ln M_t}{\mathrm{d}t} - \frac{\mathrm{d}\ln P_t}{\mathrm{d}t} = -\eta \frac{\mathrm{d}\pi^e_t}{\mathrm{d}t} \tag{2.6.31}$$

$$\mu_t - \pi_t = -\eta \frac{\mathrm{d}\pi^e_t}{\mathrm{d}t} \tag{2.6.32}$$

The monetary growth rate is determined by the real government budget deficit. If G is the real government expenditures and T is the tax income then the real government budget deficit is given by $G - T$, and to get it in nominal terms, we multiply by the price so $P(G - T)$ is the nominal budget deficit/surplus. This means

$$\frac{\mathrm{d}M}{\mathrm{d}t} = P(G-T) \tag{2.6.33}$$

$$\frac{\mathrm{d}\ln M}{\mathrm{d}t} = \frac{P}{M}(G-T) \tag{2.6.34}$$

$$\mu = \frac{(G-T)}{(M/P)} \tag{2.6.35}$$

[9] People know that if they postpone paying their taxes, it will cost them a lot less because of the inflation.

This means μ is the ratio of the real deficit $(G-T)$ to the real money (M/P) in the economy. Thus, if there is inflation, then P increases, so M/P decreases which means that μ increases. So inflation can require more printing as μ will increase, and so it runs away. At a steady state, we need $\pi = \mu$. So in the long run the inflation acts as a tax as the real deficit is fixed by the government (in this model) and so the tax is because real money is being put under pressure because of the need to finance the debt.

We can rewrite

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{M}{P}\right) = \frac{\mathrm{d}M}{\mathrm{d}t}\frac{1}{P} + \frac{-M}{P^2}\frac{\mathrm{d}P}{\mathrm{d}t} = \left(\frac{\mathrm{d}\ln M}{\mathrm{d}t} - \frac{\mathrm{d}\ln P}{\mathrm{d}t}\right)\frac{M}{P} = \frac{M}{P}\left(\mu_t - \pi_t\right) \tag{2.6.36}$$

which with our expression for μ_t gives

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{M}{P}\right) = (G-T) - \pi_t\frac{M_t}{P_t} \tag{2.6.37}$$

So that the real balances increase only if the primary deficit exceeds the inflation tax. Thus, it would be nice to know how π changes in time. We'd like $\mathrm{d}\pi/\mathrm{d}t$ which is still an open question in economics. There are multiple ideas such as myopic, adaptive, or smart responses by people in the economy. Cagan used adaptive expectations so

$$\frac{\mathrm{d}\pi_t^e}{\mathrm{d}t} = \beta(\pi_t - \pi_t^e) \tag{2.6.38}$$

with β a positive number. People change their expectations a lot if they're way off, but only a little if their expectations are only a little off. We use $\mu_t - \pi_t = -\eta\mathrm{d}\pi_t^e/\mathrm{d}t$ and plug in this value to find

$$\pi_t = \mu_t + \eta\pi_t^e = \mu_t + \eta\beta(\pi_t - \pi_t^e) \tag{2.6.39}$$
$$\pi_t(1-\eta\beta) = \mu_t - \eta\beta\pi_t^e \tag{2.6.40}$$
$$\pi_t = \frac{\mu_t - \eta\beta\pi_t^e}{1-\eta\beta} \tag{2.6.41}$$

For stability, we need $\eta\beta < 1$.

Another model is perfect foresight, better than rational expectations (since rational expectations include uncertainty so you can be wrong). Then $\pi_t^e = \pi_t$ always. You can then put together our two equations

$$\frac{M}{P} = \exp(-\eta\pi) \tag{2.6.42}$$
$$\mu = \pi = \frac{G-T}{(M/P)} \tag{2.6.43}$$

and see where they overlap. This leads to inflation growing really quickly if the real government deficit changes.

2.7 Phillips Curve

This curve comes from inductive, empirically-based economic reasoning. It is a relationship between inflation and unemployment. A. W. Phillips was fairly careful about his curve in his paper, and stated it was just a correlation, not necessarily something that must always be obeyed. He looked at the rate of growth of nominal wages and unemployment.

The ideas behind it were an understanding of short to medium term output and input in an economy. High unemployment is associated with low wage inflation while low unemployment is associated with high wage inflation. This is the Phillips curve. Unemployment has an inverse relationship to inflation.

So, there was a relationship between U (unemployment) and π (inflation). All that seemed left to do was find relationships between these and output Y. Okun's law was found for U and Y,[10] and Keynesian theory (aggregate supply curves) provided the link between Y and π.

But this was all done when there was a gold standard with constant inflation of about 1% per year. This meant there were fixed exchange rates, as well.

Milton Friedman then came and said the Phillips curve can't be fundamental. Inflation is a monetary thing and unemployment is a real thing. The relationship must be from anticipation of inflation, so people think that they're getting more, so they work harder, but they actually are not in real terms. Indeed, around 1970, the Phillips curve completely failed in the US. Thus the long run is important, as it is real not monetary items that determine the unemployment rate.

2.8 AS-AD

Thus, we return to the AS-AD picture, where we have aggregate supply (AS) and aggregate demand (AD). The graph has inflation/interest rate on the y axis and real output Y (GDP) on the x axis. The long-run aggregate supply (LAS) is vertical on this plot.

People generally use $\hat{Y}$ for the output gap so that we don't worry about Y increasing year to year.

We can use demand is given by

$$Y = C + I + G + X \tag{2.8.1}$$

with C consumption, I investment, G government spending, and X net exports (foreign contributions). The real supply is simply the $Y = Y(K, L)$ the output. Aggregate demand for goods increases as the inflation rate decreases. An increase in inflation causes either

- (fixed exchange rates) real appreciations, so then X and Y must go down.
- (flexible exchange rages) the central bank raises rates so that consumption and investment decrease as well as X.

This leads to a negative relationship between inflation and aggregate demand.

The aggregate supply depends positively on the inflation rate. This is essentially higher prices make people think that they are getting more if they do more. That is, they think the inflation means they are getting more (as they do in the short run), and so they produce more.

This helps us see if things are happening from aggregate demand or aggregate supply. So if inflation and output both increase/decrease then this is aggregate demand based, if they change opposite each other then it is an aggregate supply effect.

[10] Okun's law says fluctuations of output around its trend are negatively correlated with fluctuations of unemployment around its trend.

The exchange rate determines the slope of the AD curve. In a fixed exchange rate, there is sensitivity of exports and imports to the exchange rate: the more sensitive they are, the flatter the curve. In a flexible exchange rate, investment and consumer spending are sensitive to the exchange rate: the more sensitive they are, the flatter the curve.

Now let's try to mathematically formulate this. This will begin with linear difference equations, and then see how shocks affect our models.

We use aggregate demand and the Taylor rule (flexible exchange rates) without monetary shocks and the average deviation of inflation is $\overline{\pi} = 0$.

$$Y_t = a_1 Y_{t-1} + a_2(i_t - \pi_t) + d_t \tag{2.8.2}$$
$$i_t = c_1 \pi_t + c_2 Y_t \tag{2.8.3}$$

with $0 < a_1 < 1$ and $a_2 < 0$, $c_1 > 1$ and $c_2 > 0$. The aggregate supply/Philips curve

$$\pi_t = \widetilde{\pi}_t + b_1 Y_t + s_t \tag{2.8.4}$$

with $b_1 > 0$. Core inflation/inflationary expectations

$$\widetilde{\pi}_t = \theta \pi_t + (1-\theta)\pi_{t-1} \tag{2.8.5}$$

with $0 < \theta < 1$. Here d_t is a demand shock and s_t is a supply shock.

Then $\overline{Y}_t$ is the potential real GDP in period t.

We can solve this. First for π_t we find

$$\pi_t(1-\theta) = (1-\theta)\pi_{t-1} + b_1 Y_t + s_t \tag{2.8.6}$$
$$\pi_t = \pi_{t-1} + \frac{b_1 Y_t + s_t}{1-\theta} \tag{2.8.7}$$

and

$$Y_t = a_1 Y_{t-1} + a_2([c_1 - 1]\pi_t + c_2 Y_t) + d_t \tag{2.8.8}$$
$$Y_t(1 - a_2 c_2) = a_1 Y_{t-1} + a_2([c_1 - 1]\pi_t) + d_t \tag{2.8.9}$$
$$Y_t = \frac{a_1 Y_{t-1} + a_2([c_1 - 1]\pi_t) + d_t}{1 - a_2 c_2} \tag{2.8.10}$$

We could then subtract off Y_{t-1} to find

$$Y_t - Y_{t-1} = \frac{a_1(Y_{t-1} - Y_{t-2}) + a_2([c_1 - 1])[\pi_t - \pi_{t-1}] + d_t - d_{t-1}}{1 - a_2 c_2} \tag{2.8.11}$$
$$Y_t - Y_{t-1} = \frac{a_1(Y_{t-1} - Y_{t-2}) + a_2([c_1 - 1])\frac{b_1 Y_t + s_t}{1-\theta} + d_t - d_{t-1}}{1 - a_2 c_2} \tag{2.8.12}$$

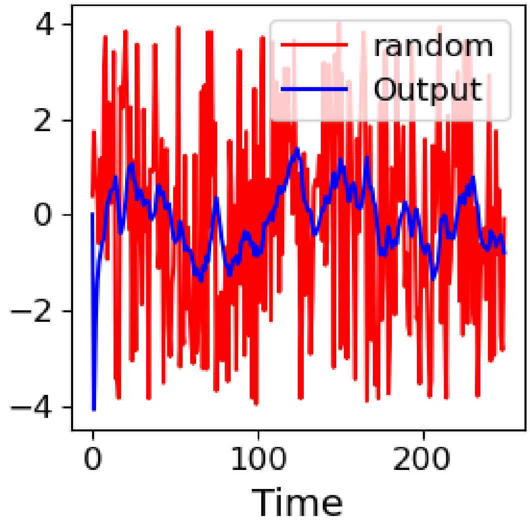

Figure 2.2: This shows the percentage change in the output and random shocks for the model with $a_1 = 0.7$, $a_2 = -0.5$, $b_1 = 0.2$, $\theta = 0.5$, $c_1 = c_2 = 1.9$. We get cycles.

which can be rewritten

$$Y_t\left(1 - \frac{a_2 b_1(c_1-1)}{(1-a_2c_2)(1-\theta)}\right) = \left(\frac{a_1}{1-a_2c_2}+1\right)Y_{t-1} - \frac{a_1}{1-a_2c_2}Y_{t-2} + \frac{d_t - d_{t-1}}{1-a_2c_2} + \frac{a_2(c_1-1)s_t}{(1-a_2c_2)(1-\theta)} \tag{2.8.13}$$

$$Y_t\left(\frac{(1-a_2c_2)(1-\theta) - a_2b_1(c_1-1)}{(1-a_2c_2)(1-\theta)}\right) = \frac{a_1+(1-a_2c_2)}{1-a_2c_2}Y_{t-1} - \frac{a_1}{1-a_2c_2}Y_{t-2} + \frac{d_t-d_{t-1}}{1-a_2c_2} + \frac{a_2(c_1-1)s_t}{(1-a_2c_2)(1-\theta)} \tag{2.8.14}$$

$$\begin{aligned} Y_t = &\frac{(1-\theta)(a_1+1-a_2c_2)}{(1-a_2c_2)(1-\theta)-a_2b_1(c_1-1)}Y_{t-1} - \frac{(1-\theta)a_1}{(1-a_2c_2)(1-\theta)-a_2b_1(c_1-1)}Y_{t-2} \\ &+ \frac{(1-\theta)(d_t-d_{t-1})}{(1-a_2c_2)(1-\theta)-a_2b_1(c_1-1)} + \frac{a_2(c_1-1)}{(1-a_2c_2)(1-\theta)-a_2b_1(c_1-1)}s_t \end{aligned} \tag{2.8.15}$$

$$\begin{aligned} Y_t = &\frac{a_1+1-a_2c_2}{1-a_2c_2-\frac{a_2b_1(c_1-1)}{1-\theta}}Y_{t-1} - \frac{a_1}{1-a_2c_2-\frac{a_2b_1(c_1-1)}{1-\theta}}Y_{t-2} \\ &+ \frac{d_t-d_{t-1}}{1-a_2c_2-\frac{a_2b_1(c_1-1)}{1-\theta}} + \frac{a_2(c_1-1)}{(1-a_2c_2)(1-\theta)-a_2b_1(c_1-1)}s_t \end{aligned} \tag{2.8.16}$$

$$Y_t \equiv \alpha_2 Y_{t-1} + \alpha_2 Y_{t-2} + \epsilon_t \tag{2.8.17}$$

This means it is a second-order difference equation. These new parameters are somewhat ugly looking, but if they change (in time), this is called the Lucas critique, because it means people adapt.

If we put white noise into ϵ_t as input, then we get something that looks like a business cycle.

This can be seen in Figure 2.2. We can also consider what happens after a single shock.

This will involve elementary (to a numerical physicist) examination of difference equations (also known as finite differences). Stochastic difference equations are important for today's economics models.

Evidence is presented for the stochastic second degree model above with economic data.

My own first thoughts are we can look at

$$f_t = \alpha_1 f_{t-1} + \alpha_2 f_{t-2} + \epsilon_t \tag{2.8.18}$$

We first consider the homogeneous version, ignoring ϵ_t at first. We can use an eigenvalue idea that $f_t = f_0 \exp(\lambda t)$ where λ is now an eigenvalue. And so we get

$$f_0 \exp(\lambda t) = \alpha_1 f_0 \exp(\lambda t)\exp(-\lambda) + \alpha_2 f_0 \exp(-2\lambda) \tag{2.8.19}$$
$$1 = \alpha_1 \exp(-\lambda) + \alpha_2 \exp(-2\lambda) \tag{2.8.20}$$

Such an equation defines the eigenvalues λ, but is not very illuminations.

The lecture introduces a linear operator, the lag operator L that takes $L(X_t) = X_{t-1}$. The proof of linearity is

$$L(aX_t + bY_t) = aX_{t-1} + bY_{t-1} = aL(X_t) + bL(Y_t) \tag{2.8.21}$$

Then consider simpler equation

$$Y_t = \alpha Y_{t-1} + \epsilon_t \tag{2.8.22}$$

with ϵ_t an identically, independently distributed (IID) random variable that has $E(\epsilon_t) = 0$ and $E(\epsilon_t^2) = \sigma^2$ with Y_0 the initial value. This can be rewritten

$$Y_t = \alpha L(Y_t) + \epsilon_t \tag{2.8.23}$$
$$(1 - \alpha L)Y_t = \epsilon_t \tag{2.8.24}$$

We can then use that

$$Y_t = \alpha Y_{t-1} + \epsilon_t = \alpha\left(\alpha Y_{t-2} + \epsilon_{t-1}\right) + \epsilon_t = \sum_{j=0}^{s} \alpha^j \epsilon_{t-j} + \alpha^{s+1} Y_{t-s-1} \tag{2.8.25}$$

We say that $Y_{t-s-1} \to 0$ as $s \to \infty$. We require $|\alpha| < 1$ so that the above operations are fine. Note that if we consider $(1 - \alpha L)^{-1}$ as the inverse operator, we have

$$Y_t = (1 - \alpha L)^{-1} \epsilon_t \tag{2.8.26}$$

for a linear operator, we can Taylor expand (with $|\alpha| < 1$, it converges) and so

$$Y_t = \left(1 + \alpha L + \alpha^2 L^2 + \cdots\right)\epsilon_t \tag{2.8.27}$$
$$Y_t = \epsilon_t + \alpha\epsilon_{t-1} + \alpha^2 \epsilon_{t-2} + \cdots \tag{2.8.28}$$
$$Y_t = \sum_{j=0}^{\infty} \alpha^j \epsilon_{t-j} \tag{2.8.29}$$

Dr. Burda approaches this as if it is a division problem, which will work since this is a linear operator. This representation is a moving average representation. It says that the infinite past affects us.

If we look at

$$Y_t = \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \epsilon_t \tag{2.8.30}$$
$$(1 - \alpha_1 L - \alpha_2 L^2) Y_t = \epsilon_t \tag{2.8.31}$$
$$Y_t = (1 - \alpha_1 L - \alpha_2 L^2)^{-1} \epsilon_t \tag{2.8.32}$$

and assume that $|\alpha_i| < 1$ then

$$Y_t = \epsilon_t + (\alpha_1 L + \alpha_2 L^2)\epsilon_t + (\alpha_1 L + \alpha_2 L^2)^2 \epsilon_t + \cdots \quad = \epsilon_t + \alpha_1 \epsilon_{t-1} + (\alpha_2 + \alpha_1^2)\epsilon_{t-2} + \alpha_1^3 + 2\alpha_1\alpha_2\epsilon_{t-3} + \cdots \tag{2.8.33}$$

This is clearly going to lead to combinations (combinatorics).

The preferred method presented is to use $(1 - \alpha_1 L - \alpha_2 L^2) = (1 - \lambda_1 L)(1 - \lambda_2 L)$ with $\alpha_1 = \lambda_1 + \lambda_2$ and $\alpha_2 = -\lambda_1 \lambda_2$. Where we require $|\lambda_i| < 1$ for convergence. We can then simply use

$$Y_t = (1 - \lambda_1 L)^{-1} (1 - \lambda_2 L)^{-1} \epsilon_t \tag{2.8.34}$$
$$Y_t = \sum_{j,k=0}^{\infty} \lambda_1^j \lambda_2^k \epsilon_{t-j-k} \tag{2.8.35}$$

The λ_i are called the characteristic roots of the autoregressive process. We can then consider $\epsilon_t = 0$ for $t < 0$. We can introduce $\nu_i = \lambda_1^j \lambda_2^k$ (with ν_i real valued) and find

$$Y_t = \sum_{i=0}^{\infty} \nu_i \epsilon_{t-i} \tag{2.8.36}$$

We then solve the particular and homogeneous solution parts of

$$Y_t = \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \epsilon_0 \tag{2.8.37}$$

where ϵ_0 is the shock at only time 0. The homogeneous solution is given by

$$Y_t^H = \kappa_1 \lambda_1^t + \kappa_2 \lambda_2^t \tag{2.8.38}$$

and the particular solution by

$$Y_t^P = \frac{\epsilon_0}{1 - \alpha_1 - \alpha_2} \tag{2.8.39}$$

This comes from assuming a steady state and solving.

Thus we see the full solution using $|\lambda_i| = R$ that

$$Y_t = \frac{\epsilon_0}{1 - \alpha_1 - \alpha_2} + R^t \kappa_1 \exp(i\theta t) + R^t \kappa_2 \exp(-i\theta t) \tag{2.8.40}$$

When λ_i are both real, we just get damping. With λ_i complex, we get damped oscillations.

We can now do our previous model as a matrix equation. We use

$$\begin{bmatrix} 1 & \beta_1 & \beta_2 & 0 \\ -\beta_3 & 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} Y_t \\ Y_{t-1} \\ \pi_t \\ \pi_{t-1} \end{bmatrix} = \begin{bmatrix} \beta_4 d_t \\ \beta_5 s_t \end{bmatrix} \tag{2.8.41}$$

Which could be written

$$\begin{bmatrix} 1 & \beta_2 \\ -\beta_3 & 1 \end{bmatrix} \begin{bmatrix} Y_t \\ \pi_t \end{bmatrix} = \begin{bmatrix} \beta_1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} Y_{t-1} \\ \pi_{t-1} \end{bmatrix} + \begin{bmatrix} \beta_4 d_t \\ \beta_5 s_t \end{bmatrix} \tag{2.8.42}$$

One can then invert the matrix on the left hand side and solve in terms of the right hand side. One then gets

$$\mathbf{x}_t = \mathbf{A}\mathbf{x}_{t-1} + \mathbf{A}_0 \boldsymbol{\epsilon}_t \tag{2.8.43}$$

We then repeat the same steps with the linear lag operator. We can write

$$(\mathbb{1} - \mathbf{A}L)\, \mathbf{x}_t = \mathbf{A}_0 \boldsymbol{\epsilon}_t \tag{2.8.44}$$

$$\mathbf{x}_t = (\mathbb{1} - \mathbf{A}L)^{-1} \mathbf{A}_0 \boldsymbol{\epsilon}_t \tag{2.8.45}$$

$$\mathbf{x}_t = \sum_{i=0}^{\infty} (\mathbf{A}L)^i \mathbf{A}_0 \boldsymbol{\epsilon}_t \tag{2.8.46}$$

The last equation is not as useful as simply doing the inversion from the second line. We can then find the eigenvalues of $\mathbf{A}$ and get what we were seeing from the simpler model.

Rational expectations (agents are not consistently making mistakes) is a modeling strategy in modern economics. The next step is then to consider the stochastic growth model (RBC model or real business cycle model).

Of course, agents can still be wrong sometimes which acts like a shock. Previous analyses usually chose $\theta = 0$ so that agents simply used the past as their guide. Rational expectations have different forms

- Strong Form: Agents know model and form mathematical expectation of relevant variables. $p_t^e = E[p_t]$.
- Intermediate Form: Agents form conditional expectations, given a subset of the relevant information $p_t^e = E[p_t|I_t]$.
- Weak Form: agents do not make systemic mistakes.

We will use expectation values $E[\cdot]$, a linear operator. The Law of Iterated Expectations is useful

$$E[E[X|I]|J] = E[X|J] \tag{2.8.47}$$

when J is a subset of I.

The problems with the AS-AD model was that it was not microeconomically founded in either demand or supply and expectations are ad hoc. Shocks are also not internal to the model, but imposed on it.

2.9 Business Cycles

One thing to note is that if you look at the trend across most business cycles, the average finds that there is a sharp decrease at a recession, followed by a slow recovery. Consumption decreases more modestly than investment. Government spending looks basically unaffected. Real money balances go flat before the peak, and so is a leading indicator. Surprisingly, real wages are also basically unaffected. So consumption and investment are procyclical, current accounts are countercyclical, and government spending is acyclical.

Now the RBC model of Kydland and Prescott. They asked can we get business cycles without having a financial sectors. Then shocks will be due to technology.

It is similar to the Ramsey model, but discrete time and stochastic shocks. Just supply shocks. Saddle-path stability is enforced. It will use a decentralized market model, but due to the 1st Welfare theorem, this will give us the social planner optimum as well.

We will then set up preferences of agents and continue with the model. We use additive log utility

$$U(C, 1-L) = \ln C + \frac{[\chi(1-L)-1]^{1-\eta}}{1-\eta} \tag{2.9.1}$$

with the Cobb-Douglas production function $F(K, ZL) = K^{\alpha}(ZL)^{1-\alpha}$.where Z is technological progress. We then have the expected utility of an agent as

$$E_0\left[\sum_{t=0}^{\infty}\beta^t\left(\ln(C_t) + \frac{\chi(1-L_t)^{1-\eta}-1}{1-\eta}\right)\right] \tag{2.9.2}$$

capital accumulation

$$K_{t+1} = (1-\delta)K_t + Y_t - C_t \tag{2.9.3}$$

and an exogenous stochastic process for technology

$$Z_t = \bar{Z}^{1-\rho}Z_{t-1}^{\rho}\exp([1-\rho]t + \epsilon_t) \approx \bar{Z}^{1-\rho}Z_t^{\rho}(1+g)^{(1-\rho)t}\exp(\epsilon_t) \tag{2.9.4}$$

$$\ln Z_t = (1-\eta)\bar{Z} + \rho\ln Z_{t-1} + (1-\rho)gt + \epsilon_t \tag{2.9.5}$$

where $\bar{Z}$ is the steady state value.

We use a single good economy. Households own capital and labor. Firms are owned by households and are run to maximize profits. K is capital, owned by firms in perpetuity for households, and L is labor. No money. The short way of saying this is preferences from households, technology employed by firms, market structure, optimal behavior, equilibrium concept as the assumptions of the model.

We then optimize

$$E_0\left[\sum_{t=0}^{\infty}\beta^t\left(\ln(C_t) + \frac{\chi(1-L_t)^{1-\eta}-1}{1-\eta}\right)\right] \tag{2.9.6}$$

subject to budget constraint

$$K_{t+1} = (1-\delta)K_t + W_tL_t + U_tK_t - C_t + \Pi_t \tag{2.9.7}$$

with wages W_t, prices U_t and profits Π_t given for a household.

We will use Lagrangian multipliers or dynamic programming to find the optimal solution. There are countably infinite Lagrangian multipliers, so this will work. We then have with our multipliers the problem

$$L = E_0\left[\sum_{t=0}^{\infty}\beta^t\left(\ln(C_t) + \frac{\chi(1-L_t)^{1-\eta}-1}{1-\eta} + \lambda_t\left((1-\delta)K_t + W_tL_t + U_tK_t - C_t + \Pi_t - K_{t+1}\right)\right)\right] \tag{2.9.8}$$

The things that are "adjustable" in this model are K_{t+1}, C_t and L_t at each time. Thus the conditions (FONCs) are

$$\frac{\partial L}{\partial K_{t+1}} = 0 = \beta^{t+1}\lambda_{t+1}[(1-\delta)+U_{t+1}] - \beta^t\lambda_t \tag{2.9.9}$$

$$\frac{\partial L}{\partial C_t} = 0 = \frac{\beta^t}{C_t} - \beta^t\lambda_t \tag{2.9.10}$$

$$\frac{\partial L}{\partial L_t} = 0 = \frac{-\beta^t\chi(1-\eta)(1-L_t)^{1-\eta-1}}{1-\eta} + \beta^t\lambda_tW_t \tag{2.9.11}$$

$$\frac{\partial L}{\partial \lambda_t} = 0 = (1-\delta)K_t + W_tL_t + U_tK_t - C_t + \Pi_t - K_{t+1} \tag{2.9.12}$$

which can be rewritten

$$\lambda_t = \beta(1-\delta+U_{t+1})\lambda_{t+1} \tag{2.9.13}$$

$$\lambda_t = \frac{1}{C_t} \tag{2.9.14}$$

$$\lambda_tW_t = \chi(1-L_t)^{-\eta} \tag{2.9.15}$$

$$K_{t+1} = (1-\delta)K_t + W_tL_t + U_tK_t - C_t + \Pi_t \tag{2.9.16}$$

which can be further reduced to

$$\lambda_t = \frac{1}{C_t} \tag{2.9.17}$$

$$\lambda_tW_t = \chi(1-L_t)^{-\eta} \tag{2.9.18}$$

$$\lambda_t = \beta[\lambda_{t+1}(1-\delta+U_{t+1})] \tag{2.9.19}$$

Another transversality condition applies

$$\lim_{t\to\infty} E_0\left[\beta^{t+1}\lambda_{t+1}K_{t+1}\right] = 0 \tag{2.9.20}$$

so people use their capital to increase their utility. We have

$$\Pi_t = \text{sales}_t + \text{costs}_t = Y_t - U_tK_t - W_tL_t \tag{2.9.21}$$

We note that there is a Solow residual S_r, which is the increase in output not due strictly to changes in capital and labor. That is, it is technical change. It is defined by

$$S_r = \left(\frac{\Delta Y}{Y}\right)_t - \alpha\left(\frac{\Delta K}{K}\right)_t - (1-\alpha)\left(\frac{\Delta L}{L}\right)_t \tag{2.9.22}$$

If this were zero, then there is no technological progress. Here α is the share of capital in national income (usually $\sim 1/3$).

In any case, firms want to optimize based off of this, so

$$W_t = (1-\alpha) Z_t^{1+\alpha} K_t^{\alpha} L_t^{-\alpha} \tag{2.9.23}$$

So to the point where marginal product of labor equals the real wage. Similarly, capital will go until the marginal product equals the real user cost

$$U_t = \alpha (Z_t L_t)^{1-\alpha} K_t^{\alpha-1} \tag{2.9.24}$$

These can be used to eliminate W_t and U_t.

We can use Euler's theorem to show that $W_t L_t + U_t K_t = Y_t$ and so there are no net profits in the economy. Remember these are economic profits, not accounting profits. It comes from constant returns to scale and perfect competition. If we plug in our values for Z we find

$$\chi(1-L_t)^{-\eta} = \frac{(1-\alpha) Z_t \left(\frac{K_t}{Z_t L_t}\right)^{\alpha}}{C_t} \tag{2.9.25}$$

$$\frac{1}{C_t} = \beta E_t \left[\frac{1}{C_{t+1}} \left(1 - \delta + \alpha \left(\frac{Z_{t+1} L_{t+1}}{K_{t+1}} \right)^{1-\alpha} \right) \right] \tag{2.9.26}$$

$$K_{t+1} = (1-\delta) K_t + K_t^{\alpha} (Z_t L_t)^{1-\alpha} - C_t \tag{2.9.27}$$

$$Z_t = \bar{Z}^{1-\rho} Z_{t-1}^{\rho} (1+g)^{1-\rho t} \exp(\epsilon_t) \tag{2.9.28}$$

where $E_t[\cdot]$ is the conditional expectation value of the next period's product of marginal utility of resources tomorrow, plus the return of resources today. We have four equations and four unknowns. The first equation is the intratemporal optimality condition (we prefer leisure, but need work to do so). The second is the Euler equation (intertemporal optimality of consumption), then the resource/budget constraint, and the dynamic equation for technology.

We added the expectation value, because we cannot actually know C_{t+1}, so we use what we would expect given the information available. The equilibrium would satisfy all of these for all t.

The next steps are find the steady-state, find the linearized equations around an equilibrium and study the behavior then.

So let's find the steady state. Unfortunately, "steady state" is a misnomer since the actual steady state would require no more growth after some time $t \to \infty$. What economists mean is that $\epsilon_t = 0$ and that the previous equations are satisfied with the additional constraint that

$$\begin{bmatrix} Z \\ Y \\ C \\ K \\ W \end{bmatrix}_t = (1+g)^j \begin{bmatrix} Z \\ Y \\ C \\ K \\ W \end{bmatrix}_{t-j} \tag{2.9.29}$$

$$\begin{bmatrix} U \\ L \end{bmatrix}_t = \begin{bmatrix} U \\ L \end{bmatrix}_{t-j} \tag{2.9.30}$$

for all j. This means that U (rental price of capital) and L (labor supply) are steady state, but other things grow only due to technological change.

This means that Y/K, C/K, WL/Y, U, and L are constant. This is called a balanced growth path.

These assumptions lead to

$$\frac{1}{C} = \frac{\beta}{C(1+g)}\left(1-\delta+\alpha\frac{Y}{K}\right) \Rightarrow (1+g)\beta^{-1} = 1-\delta+\alpha\frac{Y}{K} \tag{2.9.31}$$

$$\chi(1-L)^{-\eta} = \frac{1-\alpha}{C}\frac{Y}{L} \tag{2.9.32}$$

$$K(1+g) = (1-\delta)K + Y - C \Rightarrow g+\delta = \frac{Y-C}{K} \tag{2.9.33}$$

Now we can log linearize, which is the Taylor expansion.

Let's use K_{t+1} as an example. We can write the equation as

$$K_{t+1} = (1-\delta)K_t + Y_t - C_t \tag{2.9.34}$$

and can write these as deviations from the steady state values. We can use a twiddle notation such that $K_{t+1} = \bar{K} + \widetilde{K}$ and so we find

$$\bar{K}_{t+1} + \widetilde{K}_{t+1} = (1-\delta)(\bar{K}_t + \widetilde{K}) + \bar{Y}_t + \widetilde{Y} - \bar{C}_t - \widetilde{C} \tag{2.9.35}$$

and use that $\bar{K}_{t+1} = (1+g)\bar{K}_t$ to write

$$K(1+g) + \widetilde{K}_{t+1} = (1-\delta)(K+\widetilde{K}) + Y + \widetilde{Y} - C - \widetilde{C} \tag{2.9.36}$$

and then we can define $\widehat{Q} = \widetilde{Q}/Q$ so that

$$K(1+g)(1+\widehat{K}_{t+1}) = K(1-\delta)(1+\widehat{K}) + Y(1+\widehat{Y}) - C(1+\widehat{C}) \tag{2.9.37}$$

$$(1+g)(1+\widehat{K}_{t+1}) = (1-\delta)(1+\widehat{K}) + \frac{Y}{K}(1+\widehat{Y}) - \frac{C}{K}(1+\widehat{C}) \tag{2.9.38}$$

Next let's consider the χ equation

$$\chi(1-[\bar{L}_t+\widetilde{L}_t])^{-\eta} = \frac{(1-\alpha)[\bar{Z}_t+\widetilde{Z}_t]\left(\frac{[\bar{K}_t+\widetilde{K}_t]}{[\bar{Z}_t+\widetilde{Z}_t][\bar{L}_t+\widetilde{L}_t]}\right)^{\alpha}}{[\bar{C}_t+\widetilde{C}_t]} \tag{2.9.39}$$

or

$$\chi(1-L[1+\widehat{L}_t])^{-\eta} = \frac{(1-\alpha)Z[1+\widehat{Z}_t]\left(\frac{K[1+\widehat{K}_t]}{ZL[1+\widehat{Z}_t][1+\widehat{L}_t]}\right)^{\alpha}}{C[1+\widetilde{C}_t]} \tag{2.9.40}$$

$$\chi(1-L)^{-\eta}(1-\frac{L\widehat{L}_t}{1-L})^{-\eta} = (1-\alpha)\frac{Z}{C}[1+\widehat{Z}_t]\left(\frac{K}{ZL}\right)^{\alpha}\left([1+\widehat{K}_t][1-\widehat{Z}_t][1-\widehat{L}_t]\right)^{\alpha}[1-\widehat{C}_t] \tag{2.9.41}$$

$$\chi(1-L)^{-\eta}(1+\frac{\eta L\widehat{L}_t}{1-L}) = (1-\alpha)\frac{Z}{C}[1+\widehat{Z}_t]\left(\frac{K}{ZL}\right)^{\alpha}\left(1+\widehat{K}_t-\widehat{Z}_t-\widehat{L}_t\right)^{\alpha}[1-\widehat{C}_t] \tag{2.9.42}$$

$$\chi(1-L)^{-\eta}(1+\frac{\eta L\widehat{L}_t}{1-L}) = (1-\alpha)\frac{Z}{C}[1+\widehat{Z}_t]\left(\frac{K}{ZL}\right)^{\alpha}\left(1+\alpha[\widehat{K}_t-\widehat{Z}_t-\widehat{L}_t]\right)[1-\widehat{C}_t] \tag{2.9.43}$$

$$\left(1+\frac{\eta L\widehat{L}_t}{(1-L)}\right) = \frac{(1-\alpha)Z}{C\chi(1-L)^{-\eta}}\left(\frac{K}{ZL}\right)^{\alpha}\left(\widehat{Z}_t+\alpha[\widehat{K}_t-\widehat{Z}_t-\widehat{L}_t]-\widehat{C}_t\right) \tag{2.9.44}$$

Then from the steady state we know that the 1 on the left and right must cancel and that the Z/C factors must cancel with the $(1-L)^{-\eta}$ for

$$\frac{\eta L\widehat{L}_t}{(1-L)} = \left(\widehat{Z}_t+\alpha[\widehat{K}_t-\widehat{Z}_t-\widehat{L}_t]-\widehat{C}_t\right) \tag{2.9.45}$$

$$\left(\alpha+\frac{\eta L}{(1-L)}\right)\widehat{L}_t = (1-\alpha)\widehat{Z}_t+\alpha\widehat{K}_t-\widehat{C}_t \tag{2.9.46}$$

Then we can consider

$$\frac{1}{\bar{C}_t+\widetilde{C}_t} = E_t\left[\frac{1}{\bar{C}_{t+1}+\widetilde{C}_{t+1}}\left(1-\delta+\alpha\left(\frac{[\bar{Z}_{t+1}+\widetilde{Z}_{t+1}][\bar{L}_{t+1}+\widetilde{L}_{t+1}]}{[\bar{K}_{t+1}+\widetilde{K}_{t+1}]}\right)^{1-\alpha}\right)\right] \tag{2.9.47}$$

$$\frac{1}{C[1+\widehat{C}_t]} = E_t\left[\frac{1}{(1+g)C[1+\widehat{C}_{t+1}]}\left(1-\delta+\alpha\left(\frac{Z(1+g)[1+\widehat{Z}_{t+1}]L(1+g)[1+\widehat{L}_{t+1}]}{K(1+g)[1+\widetilde{K}_{t+1}]}\right)^{1-\alpha}\right)\right] \tag{2.9.48}$$

$$\frac{[1-\widehat{C}_t]}{C} = E_t\left[\frac{[1-\widehat{C}_{t+1}]}{(1+g)C}\left(1-\delta+\alpha\left(\frac{ZL(1+g)}{K}\right)^{\alpha}\left([1+\widehat{Z}_{t+1}][1+\widehat{L}_{t+1}][1-\widetilde{K}_{t+1}]\right)^{1-\alpha}\right)\right] \tag{2.9.49}$$

$$\tag{2.9.50}$$

Steady state cancellations then yield

$$-\frac{\widehat{C}_t}{\cancel{C}} = \cancel{\beta}E_t\left[\cancel{\frac{1}{(1+g)C}\left(1-\delta+\alpha\left(\frac{ZL(1+g)}{K}\right)^{\alpha}\right)}\left(-\widehat{C}_{t+1}+[1-\alpha][\widehat{Z}_{t+1}+\widehat{L}_{t+1}-\widehat{K}_{t+1}]\right)\right] \tag{2.9.51}$$

$$-\widehat{C}_t = E_t\left[\left(-\widehat{C}_{t+1}+[1-\alpha][\widehat{Z}_{t+1}+\widehat{L}_{t+1}-\widehat{K}_{t+1}]\right)\right] \tag{2.9.52}$$

And finally

$$Z_t = \bar{Z}^{1-\rho}Z_t^{\rho}(1+g)^{(1-\rho)t}\exp(\epsilon_t) \tag{2.9.53}$$

$$\bar{Z}_t+\widetilde{Z}_t = \bar{Z}^{1-\rho}(\bar{Z}_{t-1}+\widetilde{Z}_{t-1})^{\rho}(1+g)^{(1-\rho)t}\exp(\epsilon_t) \tag{2.9.54}$$

$$Z(1+\widehat{Z}_t) = \bar{Z}^{1-\rho}Z^{\rho}(1+\widehat{Z}_{t-1})^{\rho}(1+g)^{(1-\rho)t}\exp(\epsilon_t) \tag{2.9.55}$$

$$(1+\widehat{Z}_t) = (1+\widehat{Z}_{t-1})^{\rho}(1+g)^{(1-\rho)t}(1+\epsilon_t) \tag{2.9.56}$$

$$\widehat{Z}_t = \rho\widehat{Z}_{t-1}+\epsilon_t \tag{2.9.57}$$

We can now calibrate the model with data-driven choices for α, β, η, g and the rest. (By the way the process is called DSGE, dynamic stochastic general equilibrium.) It turns out we can eliminate $\widehat{L}$ completely and go down to three equations.

The first solution method is then to simply write out the matrix equation of the form

$$E_t[\begin{bmatrix}\widehat{K}_{t+1}\\ \widehat{Z}_{t+1}\\ \widehat{C}_{t+1}\end{bmatrix}] = \mathbf{A}\begin{bmatrix}\widehat{K}_t\\ \widehat{Z}_t\\ \widehat{C}_t\end{bmatrix} \tag{2.9.58}$$

with $\mathbf{A}$ a 3×3 matrix. One can also include shocks with an extra vector of the form $[0, \epsilon_{t+1}, a\epsilon_{t+1}]^{\intercal}$ added to the right hand side and removing expectation values.

So we look at the eigenvalues and eigenvectors of $\mathbf{A}$. When we do such an analysis, we find that one eigenvalue is outside the unit circle and the other two are inside, so we don't have guaranteed stability (saddle path restrictions will give us the required stability). Essentially, because we have a free parameter (control variable), we can choose it to remain on a stable dynamic path. Saddle stability requires the number of stable roots is equal to the number of state variables, or Blanchard-Kahn condition.

The other solution is simply the method of undetermined coefficients, or good guessing. Therefore we know

$$\widehat{C}_t = \eta_{CK}\widehat{K}_t + \eta_{CZ}\widehat{Z}_t \tag{2.9.59}$$

with the η the elasticities of C with respect to the other variable. Then just substitute this into the other equations and solve for the ηs.

The strengths of RBC are that C, I, and Y are procyclical, the variances of them is in the correct ordering $\mathrm{Var}(I) > \mathrm{Var}(Y) > \mathrm{Var}(C)$, the interest rates are procyclical, and the model (propagation) is relatively easy to understand.

Some weaknesses are it predicts a constant wage share, it does not have persistence [not enough endogenous propagation, so things only really happen when the shock is on] (all persistence really comes from technology shock), and it predicts that government spending impoverishes people as it is completely wasted. In addition, unemployment is purely voluntary and there is no money in the model.

2.10 New-Keynesian Macroeconomics

This allows monetary neutrality to fail and is the most current widespread methodology. Essentially, it looks at why prices are "sticky". Remember the criticisms of the RBC model. Remember that RBC says the best response possible is from private agents to stochastic production possibilities.

So why isn't money neutral over shorter terms? The reason is nominal prices or wages are predetermined (state variables) when the money supply, liquidity or interest rates change. Prices are "sticky". Monetary neutrality is true in the steady state.

Average time between price changes is 6 to 12 months in the US and 13 months in the Euro. It varies a lot by sector. Goods change around 3 months and services about 8 months. These are due to contracts, inability to store, etc.

The original Keynesian idea is that some agents have set money prices or wages that don't change when aggregate demand or money supply changes. So price level is sticky because individual prices are set in advance and don't change with new information. (One could also use that psychologically, people may punish a price raise). The question of why this happens is still open.[11]

A central assumption of New-Keynesian approaches is imperfect competition and the ability of firms (households) to set prices (wages). For specifics, one can look at a Mankiw-Romer diagram which shows how a monopolist sets pricing to get the most profit.

We'll start with the Rotemberg (1983) model, following a clean presentation from Roberts (1995). It models a linear-quadratic cost minimization problem in log of prices (p), given the optimal price (p^*), the prices charged by other (p') and a shock to demand (ϵ). By imposing symmetry p and p' are enough to convey an impression of the model in general.

Firms want to set their own price close to the optimal

$$p_t^* = p_t' + \beta y_t + \epsilon_t \tag{2.10.1}$$

but they have competitors that might not cooperate (here y_t is like an output gap). Here y_t is the exogenous demand for the firm's output, and β is an elasticity of the optimal price with respect to the price with $\beta > 0$.

Cost changes are costly and use a penalty term $c(p_t - p_{t-1})^2$ with $c > 0$ for a convex function. This is customers resenting price increases. So small steps are better than big ones. The firm will choose a price policy/sequence. For now the firm chooses a policy based on maximizing the expected discounted value of profits (i.e., maximize profits knowing that big changes today mean not needing to change prices a lot in the future and that things are worth less as time goes on).

We can find the optimal plan by minimizing the present discounted value of total expected costs

$$E_0 \left[\sum_{t=0}^{\infty} R^t \left[(p_t^* - p_t)^2 + c(p_t - p_{t-1})^2 \right] \right] \tag{2.10.2}$$

We have p_t^* so this actually says

$$E_0 \left[\sum_{t=0}^{\infty} R^t \left[(p_t' + \beta y_t - p_t)^2 + c(p_t - p_{t-1})^2 \right] \right] C \tag{2.10.3}$$

There are no constraints, so we can simply take derivatives to see if there are extrema. Thus we take $\frac{\partial}{\partial p_t}$ for every t and find

$$E_0 \left[R^t \left[2(p_t' + \beta y_t - p_t)(-1) + 2c(p_t - p_{t-1})(1) \right] + 2R^{t+1} c(p_{t+1} - p_t)(-1) \right] = 0 \tag{2.10.4}$$

$$E_0 \left[-(p_t' + \beta y_t - p_t) + c(p_t - p_{t-1}) - Rc(p_{t+1} - p_t) \right] = 0 \tag{2.10.5}$$

(2.10.6)

Now we impose symmetry so that every one is acting like their competitors and $p_t = p_t'$. Then inflation is $\pi_t = p_t' - p_{t-1}'$. One can then solve for inflation and find

$$\pi_t = R E_t \pi_{t+1} + \frac{\beta y_t + \epsilon_t}{c} \tag{2.10.7}$$

[11] Some models can have humorous descriptions. Calvo's "fairy" is the "fairy" that randomly gives firms the chance to change prices in their model, for example.

the Rotemberg New Keynesian Phillips curve.

Now let's look at Calvo, where firms cannot change prices whenever they want. The fairy must give them permission to do so (in a time discrete model, this occurs with probably $1-\varphi$ for $0<\varphi<1$ where φ is the probability of getting stuck at its current price). Here φ is a measure of price rigidity. Because of this, firms will frontload their price changes. The probability of a firm being able to change price on the $(n+1)$th period is exactly $(1-\varphi)\varphi^n$.

We can find that this is normalized and also the expected value, given by $(1-\varphi)^{-1}$. (Use that we can take a derivative term by term).

It should be said that the probability process is a Poisson process. An advantage of the Calvo method is then that we can get a variance of the price level as it is a probabilistic process. An assumption of this model is that

$$\begin{aligned} p_t &= (1-\varphi)\sum_{\tau=0}^{\infty}\varphi^\tau E_t p^*_{t+\tau} = (1-\varphi)p^*_t + (1-\varphi)\sum_{\tau=1}^{\infty}\varphi^\tau E_t p^*_{t+\tau} \\ &= (1-\varphi)p^*_t + \varphi(1-\varphi)\sum_{\tau=0}^{\infty}\varphi^\tau E_t p^*_{t+\tau+1} \\ &= (1-\varphi)p^*_t + \varphi E_t p_{t+1} \end{aligned} \tag{2.10.8}$$

where again we use $p^*_t = p'_t + \beta y_t + \epsilon_t$. Here p'_t is the aggregate price index

$$\begin{aligned} p'_t &= (1-\varphi)\sum_{\tau=0}^{\infty}\varphi^\tau p_{t-\tau} = (1-\varphi)p^*_t + (1-\varphi)\sum_{\tau=1}^{\infty}\varphi^\tau p_{t-\tau} \\ &= (1-\varphi)p_t + \varphi p'_{t-1} \end{aligned} \tag{2.10.9}$$

Now we can get the inflation rate

$$\pi_t \equiv p'_t - p'_{t-1} = E_t\pi_{t+1} + \frac{(1-\varphi)^2}{\varphi}(\beta y_t + \epsilon_t) \tag{2.10.10}$$

We can see this via

$$p'_t - p'_{t-1} = (1-\varphi)p_t + \varphi p'_{t-1} - p'_{t-1} = (1-\varphi)(p_t - p'_{t-1}) \tag{2.10.11}$$

$$= (1-\varphi)(p_t - (1-\varphi)p_{t-1} - \varphi p'_{t-2}) \tag{2.10.12}$$

$$= (1-\varphi)(p_t - (1-\varphi)p_{t-1} - \varphi(1-\varphi)p_{t-2} - \varphi^2 p_{t-3'}) \tag{2.10.13}$$

$$= (1-\varphi)p_t - (1-\varphi)^2\sum_{j=1}^{N}\varphi^{j-1}p_{t-j} - (1-\varphi)\varphi^{N+1}p_{t-N-1} \tag{2.10.14}$$

as $N\to\infty$ we can ignore the last term as negligible and so we find

$$= (1-\varphi)p_t - \frac{(1-\varphi)^2}{\varphi}\sum_{j=1}^{\infty}\varphi^j p_{t-j} \tag{2.10.15}$$

$$= (1-\varphi)^2 p^*_t + \varphi(1-\varphi)E_t p_{t+1} - (1-\varphi)^2\left[\frac{p'_t}{1-\varphi} - p_t\right] \tag{2.10.16}$$

and etc. using $p'_{t+1} = (1-\varphi)p_{t+1} + \varphi p'_t$ and expectation values for future events.

We then set up a maximization problem over a possibly continuously infinite set of goods as

$$E_0\left[\sum_{t=0}^{\infty}\beta^t\left(U(C_t) - V(L_t)\right)\right] \tag{2.10.17}$$

with $U', V' > 0$ and $U'', V'' < 0$ and C_t defined as

$$C_t = \left\{\int_0^1 \mathrm{d}i\ [C_t(i)]^{\frac{\epsilon-1}{\epsilon}}\right\}^{\frac{\epsilon}{\epsilon-1}} \tag{2.10.18}$$

Then C_t is called a consumption index, or subutility or Armington aggregator. For $\epsilon > 1$ then C_t has the properties that it is homothetic, so doubling $C_t(i)$ for all i then C_t doubles, and as $\epsilon \to 1$ each good becomes essential and commands a constant budget share. When $\epsilon \to \infty$ the goods become perfect substitutes. A firm's market power is determined by ϵ. A place like Amazon has an ϵ near 1 as it has strong market power.

So a consumer faces a two stage decision. How much to buy today and how much to save for tomorrow. Introduce $P_t(j)$ for $0 \le j \le 1$ which is the price of good with index j.

Suppose the budget constraint says one has Z_t for consumption in time t. Then we want to maximize C_t subject to the constraint

$$\int_0^1 \mathrm{d}i\ P_t(i)C_t(i) = Z_t \tag{2.10.19}$$

Thus it is a Lagrangian form problem of the form

$$L = \left\{\int_0^1 \mathrm{d}i\ [C_t(i)]^{\frac{\epsilon-1}{\epsilon}}\right\}^{\frac{\epsilon}{\epsilon-1}} - \lambda\left[\int_0^1 \mathrm{d}i\ P_t(i)C_t(i) - Z_t\right] \tag{2.10.20}$$

And so we find

$$\frac{\epsilon}{\epsilon-1}\left\{\int_0^1 \mathrm{d}i\ [C_t(i)]^{\frac{\epsilon-1}{\epsilon}}\right\}^{\frac{\epsilon}{\epsilon-1}-1}\int_0^1 \mathrm{d}i\ \left[\frac{\epsilon-1}{\epsilon}C_t(i)^{\frac{\epsilon-1}{\epsilon}-1}\right] - \int_0^1 \mathrm{d}i\ \lambda P_t(i) = 0 \tag{2.10.21}$$

$$\left\{\int_0^1 \mathrm{d}i\ [C_t(i)]^{\frac{\epsilon-1}{\epsilon}}\right\}^{\frac{1}{\epsilon-1}}\int_0^1 \mathrm{d}i\ \left[C_t(i)^{\frac{-1}{\epsilon}}\right] - \int_0^1 \mathrm{d}i\ \lambda P_t(i) = 0 \tag{2.10.22}$$

And if we restrict to a single good then we'd find

$$\left\{\int_0^1 \mathrm{d}i\ [C_t(i)]^{\frac{\epsilon-1}{\epsilon}}\right\}^{\frac{1}{\epsilon-1}}\left[C_t(i)^{\frac{-1}{\epsilon}}\right] - \lambda P_t(i) = 0 \tag{2.10.23}$$

And using the definition of bare C_t we see that this says

$$\left[\frac{C_t(i)}{C_t}\right]^{-1/\epsilon} = \lambda P_t(i) \tag{2.10.24}$$

To get rid of λ, our multiplier we can use a neat trick

$$C_t(i) = \lambda^{-\epsilon} P_t(i)^{-\epsilon} C_t \tag{2.10.25}$$

$$P_t(i)C_t(i) = \lambda^{-\epsilon} P_t(i)^{1-\epsilon} C_t \tag{2.10.26}$$

$$\int_0^1 \mathrm{d}i\ P_t(i)C_t(i) = \int_0^1 \mathrm{d}i\ \lambda^{-\epsilon} P_t(i)^{1-\epsilon} C_t = \lambda^{-\epsilon} \int_0^1 \mathrm{d}i\ P_t(i)^{1-\epsilon} C_t \tag{2.10.27}$$

$$Z_t == \lambda^{-\epsilon} \int_0^1 \mathrm{d}i\ P_t(i)^{1-\epsilon} C_t \tag{2.10.28}$$

$$\lambda = \left[\frac{C_t \int_0^1 \mathrm{d}i\ P_t(i)^{1-\epsilon}}{Z_t} \right]^{1/\epsilon} \tag{2.10.29}$$

So we find as FONCs

$$C_t(j) = \frac{Z_t P_t(j)^{-\epsilon} \cancel{C_t}}{\cancel{C_t} \int_0^1 \mathrm{d}i\ P_t(i)^{1-\epsilon}} \tag{2.10.30}$$

One can then define the CES (constant elasticity of substitution) price index as $P_t \equiv [\int_0^1 \mathrm{d}i\ P_t(i)^{1-\epsilon}]^{(1-\epsilon)^{-1}}$ so that this now says

$$C_t(j) = Z_t P_t(j)^{-\epsilon} P_t^{\epsilon-1} = \left[\frac{P_t(j)}{P_t} \right]^{-\epsilon} \frac{Z_t}{P_t} \tag{2.10.31}$$

The demand good j has unit real income elasticity, own price elasticity ϵ, and is homogeneous of degree 0 in all nominal prices (doubling all prices means nothing happens in real terms). This is the demand curve faced by the price setter.

Another method to find this is to use ratios of FONCs so that

$$\frac{C_t(i)}{C_t(j)} = \frac{\lambda^{-\epsilon} P_t(i)^{-\epsilon} C_t}{\lambda^{-\epsilon} P_t(j)^{-\epsilon} C_t} = \left[\frac{P_t(i)}{P_t(j)} \right]^{-\epsilon} \tag{2.10.32}$$

Now plug this into the budget constraint

$$Z_t = \int_0^1 \mathrm{d}i\ P_t(i)C_t(i) = \int_0^1 \mathrm{d}i\ P_t(i) \left[\frac{P_t(i)}{P_t(j)} \right]^{-\epsilon} C_t(j) \tag{2.10.33}$$

$$Z_t = \int_0^1 \mathrm{d}i\ P_t(i)C_t(i) = C_t(j)P_t(j)^{\epsilon} \int_0^1 \mathrm{d}i\ P_t(i)^{1-\epsilon} \tag{2.10.34}$$

which with the same price index definition yields

$$Z_t = C_t(j)P_t(j)^{\epsilon} P_t^{1-\epsilon} \tag{2.10.35}$$

the same as we found before.

Thus Calvo and Rotemberg lead to New Keynesian Phillips Curves (NKPCs) (really supply curves). These are

$$\pi_t \overset{\text{Calvo}}{=} E_t \pi_{t+1} + \frac{(1-\varphi)^2}{\varphi} (\beta y_t + \epsilon) \tag{2.10.36}$$

$$\pi_t \overset{\text{Rotemmberg}}{=} E_t \pi_{t+1} + \frac{\beta y_t + \epsilon}{c} \tag{2.10.37}$$

$$\pi_t \overset{\text{general}}{=} \widetilde{\pi}_{t+1} + b_t Y_t + s_t \tag{2.10.38}$$

This shows that some combination of nominal rigidity (small c or large φ) and real rigidity (small β) are necessary for real effects of monetary policy.

Now for the critiques of these models. The MacCullam critique says that it should not be possible for a country to enrich itself by permanently lowering or raising inflation. Also, it doesn't give enough inflation persistence (models say past inflation doesn't matter, but they do). Relatedly, firms that don't change prices keep their prices constant when maybe the firms that don't "change" should assume a rule of thumb such as last period's inflation.[12] The capital stock is completely ignored in the model. Indeed, New Keynesian models without capital stock have trouble reproducing the accelerator-multiplier effect (like China). Finally, the Lucas critique, which is that the parameters of the model should not be time-independent since people in the model should react to what is happening in the model.

One partial response to the persistence problem is using a Calvo-Fischer contract.[13] Now we have sticky information rather than sticky prices. Thus, a firm can change prices at any time, but to set a new policy it has to do the research which will take time. Basically businesses (in general) aren't keeping track of inflation, but are trying to maximize profits and so miss out on inflation trends unless they're very salient. So firms set a price policy with a contract but then can't change it until the contract ends (a contract ends with a probability of doing more research).

Remember that central bank doesn't directly control the money supply, they control interest rates. New Keynesian models say not money, but nominal interest rates are decisive holding conditional expectations of inflation constant. Taylor "rule" (hypothesis) was central bank sets interest rates to fight inflation and to close the output gap.[14] These models also assume cashless economy is good enough approximation. There are models where markets clear (non-Walrasian) but also have Keynesian properties of cycles.

When we add rational expectations, even if the underlying model dynamics would be unstable (there are fewer eigenvalues inside the unit circle is less than the number of state variables) the rational expectations can allow stability (though the path is not determined except by exogenous information). Models with increasing rates of production often have multiple equilibria and so extra information determines which equilibrium you are going to. These are called sunspot models because sunspots were somewhat correlated with markets. The sunspots will then in reality be technology shocks or something similar. An increasing return production function would be $Y_t = Z_t K_t^{\alpha}(1 - L_t)^{\beta}$ with $\alpha + \beta > 1$.

2.11 IS-TR-PC Model

This comes from the IS-TR model, also called the IS-LM model,[15] and we will consider more modern variants. We have an IS (investment saving) curve from goods market, intertemporal forces. The TR (Taylor rule) curve comes from monetary policy, money market, Taylor rule, nominal interest rates.

Then IS + TR yields the AD curve. Then the NKPC (New Keynesian Phillips Curve) yields us the AS curve. We also add rational expectations in the form $E_t[E_{t+1}[x_{t+j}]] = E_t[x_{t+j}]$.

[12] Some get around this by having some agents not be so clever and just use the past inflation rate.

[13] Basically, a different type of "fairy".

[14] The rule is central banks raise interest rates when inflation rises above target and when output is above trend.

[15] The LM is liquidity-preference-money supply.

This is summarized as

$$y_t = E_t y_{t+1} - b_1(i_1 - E_t \Delta p_{\,t+1}) + v_t \tag{2.11.1}$$

$$i_t = E_t \Delta p_{\,t+1} + d_1 \left(E_t \Delta p_{\,t+1} - \pi\right) + d_2 y_t + w_t \Delta p_{\,t} \quad = (1 - c_1)\Delta p_{\,t-1} + c_1 E_t \Delta p_{\,t+1} + c_2 y_t + u_t \tag{2.11.2}$$

Can be written as a matrix equation and then solved. Thus, the AS-AD model is not completely unmoored from more microeconomic foundations.

List of Terms

acyclical Acyclical variables are variables that do not fluctuate in correlation with business cycle fluctuations in GDP plural. 61

adverse selection This is when a statement/offer conveys negative information. 23

average cost This is the mean or average cost of all the products produced after producing the nth one. That is, if you sell 50 units, then the average cost is the sum of the costs of all 50 units plus fixed costs divided by 50.. 19

Cambridge equation This is a relationship between quantity of goods Y in a period of time, price P in that period, the money supply M and money velocity V. The Cambridge equation states $MV = PY$. plural. 49

capital stock This refers to all goods that aid in producing output (in the form of products, hence GDP). Examples include machines and factories. Conventionally, it is given by $K(t)$ with t time plural. 27

club good This is a good that is non-rivalrous and excludable. Cinemas, zoos, golf courses, etc., are examples.. 22

Coase's theorem A theorem that says that externalities can be managed by market forces by giving property rights to a party when the following conditions hold.

- Property rights must be clearly defined.
- There must be little to no transaction costs.
- There must be few affected parties (transaction costs are high if one must navigate many parties)
- There are no wealth effects. The efficient solution will be the same regardless of which party gets the property rights.

There are behavioral economic critiques of whether these apply in normal circumstances. 17

comparative statics This is the analysis of a model or system where we hold all variables but one constant. That is, we look at a system via its independent variables one at a time to get an understanding for how the system changes as we change parameters/variables. 35

complement good A complement good is a good (in comparison to another good A) whose demand increases when good A's price decreases. 11

constant costs industry An industry where the supply curve slopes horizontally to the right on a price vs quantity graph. 19

constant returns to scale This is usually an assumption/hypothesis for models that states that an increase in inputs (capital or labor) leads to an equivalent (proportional) increase in output. Thus increasing inputs by 10% leads to an outputs increasing by 10%. 29, 33

consumer surplus For an individual, this is the difference between the maximum price a consumer is willing to pay and what they actually pay. For all consumers, this is the sum of each individual's consumer surplus, and in the continuous case, the upper part of the left most region of a supply-demand curve plot in price vs quantity. 10

control variable See costate variable plural. 66

costate variable A costate variable is a variable that one does have control over. It can be thought of as something one changes to have effects on state variables, and so the state of the economy. These are often also called control variables plural. 44, 74

countercyclical Countercyclical variables are variables that fluctuate negatively with business cycle fluctuations in GDP (that is, they change in the opposite direction as the GDP fluctuations) plural. 61

deadweight loss These are missed opportunities or excessive trading because of a price/quantity discrepancy in a market. 13

decreasing costs industry An industry where the supply curve slopes downward and to the right on a price vs quantity graph. 19

demand Demand in economics means how much of a good (the quantity) is wanted by consumers at a particular price. (One can interpret this the other way around, as the price willing to be paid given a certain quantity of goods.) An increase in demand means changing the curve, while a change in quantity demanded is along the curve. It is usually a monotonically decreasing function of quantity. 7

economic surplus This is the sum of the consumer and producer surpluses. It is the left most region of a supply-demand curve plot in price vs quantity. Sometimes it is called total welfare or Marshallian surplus. 10

elasticity This is usually defined as ϵ as the percent change in quantity over the percent change in price, and characterizes a demand or supply curve. $|\epsilon| < 1$ is said to be inelastic and $|\epsilon| > 1$ is said to be elastic with $|\epsilon| = 1$ said to be unit elastic. The definition is often given as $\epsilon = \frac{\Delta Q}{\Delta P}\frac{P}{Q}$, though given the ambiguity of this, the "midpoint rule" definition is usually used given by $\epsilon = \frac{\Delta Q}{\Delta P}\frac{\langle P\rangle}{\langle Q\rangle}$ where $\langle q\rangle = \frac{q_1+q_2}{2}$ and $\delta q = q_2 - q_1$. In the continuum case, we get $\epsilon = \frac{\mathrm{d}\ln Q}{\mathrm{d}\ln P} = \frac{P}{Q}\frac{\mathrm{d}Q}{\mathrm{d}P}$. 10, 27, 46

endogenous An endogenous variable is a variable that is specified "inside" of the model. That is, it is a variable where one cannot choose its form. The model does not take the variable as an input (except maybe an initial condition), but determines the form of the endogenous variable via evolution equations (or the equivalent) plural. 30

excludable good An excludable good is a good where it is easy to preven people who do not pay for the good to not get any benefits from the good. For example, if I eat an apple, no one else can then easily benefit from it. A non-excludable good is a good where if someone uses the good, it is difficult to prevent it being used (in some sense) by people who do not pay. For example, timber in a forest (where no one owns the forest) is non-excludable since it is hard if not illegal or impossible to prevent others from harvesting the wood. 22

exogenous An exogenous variable is a variable that is specified "outside" of the model. That is, it is a variable where one can choose its form. It can be thought of as saying the model does not force any specific form for the variable, but tells you what occurs once you do choose a form plural. 30, 51

externality An externality is a cost that is imposed neither on the consumer or the producer, but on bystanders. This warps the incentives for optimal use of goods in a market. 16

factor price frontier This uses neoclassical theory to understand the income distribution. Essentially, it is the possibilities for the division of factor production between wages and captital. 33, 38

finished good A finished good is a good that will not be sold again as a part of another good. A used good being resold is also not considered a finished good. Sometimes this is called a final good.. 25

First and Second Fundamental Welfare theorems These are theorems that basically show that an optimum planner and competitive markets will yield the same results in economies. Obviously, competitive markets has some assumptions about consumers and the way the markets operate plural. 48

first order necessary condition These are the conditions that are required to give an optimum for some economic problem. They are often abbreviated FONCs plural. 47, 75

fixed cost This is the cost for a business that does not change when increasing or decreasing production. 18

FONC See first order necessary condition plural. 47, 50, 51, 70

GDP per capita This is the gross domestic product divided by the total population within the geographic area/country. 26

Giffen good This is a good with an positive slope demand curve. This means an increase in price actually causes an increase in demand. The way this works is that if you rely on bread for meals, but at its current low price you can afford some meat (much more expensive than bread, let's say), then you will spend money on both. If the bread increases some, then it could be that bread is now too expensive for you to buy the meat. But you still need the calories, so you are forced to buy more bread to fill up the calorie deficit. 7

Goldsmith equation This is an equation relating the time derivative of the capital stock to the production function Y and current capital stock K with constant investment γ and constant depreciation δ given by $\mathrm{d}K/\mathrm{d}t = \gamma Y(t) - \delta K(t)$ plural. 30

gross domestic income See gross domestic product. 26

gross domestic product Gross domestic product (GDP) is the market value of all finished/final goods produced in a geographic area (usually a country) in one year. 25, 75, 77, 78

gross national product Gross national product (GNP) is the market value of all goods and services produced in one year by labor and property supplied by the citizens of a country. It is now often called GNI. 25

increasing costs industry An industry where the supply curve slopes upward and to the right on a price vs quantity graph. 19

inferior good An inferior good is a good whose demand increases when consumer income decreases. 11

inflation Inflation is an increase in prices of goods not associated with any improvements (increases) in goods or services provided. 26

institution An economic institution is a set of shared norms/ideas that are used by a society. Thus, appreciation for creating new knowledge, the system of granting patents, and prizes for innovation are examples of institutions. Other types of instiutions include good governance, such as a fair (or at least consistent) system of laws plural. 28

intermediate good A intermediate good is a good that will be sold again as a part of another good. 25

Keynes-Ramsey rule This is a rule that tells us consumption changes c_t when the production function $y(k)$ changes and we have population changes n, deprecation δ, and temporal discounting θ. The elasticity of substition $\sigma = -cu''/u'$ for u utility is also used and the rule states

$$\frac{\dot{c}_t}{c_t} = \frac{y'(k_t) - (n + \theta + \delta)}{\sigma}$$

plural. 46

Lucas critique The Lucas critique is a criticism of any model that uses time-fixed coefficients when those coefficients represent human behavior. The Lucas critique says people react and learn, so that the time-independent parameters should actually be time dependent plural. 57

marginal cost The marginal cost is the extra cost crated by selling one additional unit of a product. 19

marginal product This is the change in output resulting from employing one more unit of input. Thus given an output/production function $Y = Y(K, L)$ then the marginal product of capital K is $\frac{\partial Y}{\partial K}$ and the marginal product of labor is $\frac{\partial Y}{\partial L}$. 33

marginal revenue The marginal revenue is the extra revenue crated by selling one additional unit of a product. 19

market power This is the ability to raise the price of a good above the marginal cost (without any competitors to come in and drive down the price).. 20

monetary neutrality Monetary neutrality means that the time derivative of the velocity of money is zero. That is, people spend money at roughly the same rate over time. It has empirically been found to be true over large spans of time plural. 50

moral hazard Moral hazard is when one has entered into an agreement while misleading the other party in some way. 23

nominal GDP This is the standard gross domestic product defined above. It is simply a tabulation using the current prices. 26

normal good A normal good is a good whose demand increases when consumer income increases. 11

open access common good This is a good that is rivalrous and non-excludable. These are things that are held in common but can be used up, so fish stocks in the ocean, timber in an unowned forest, etc.. 22

opportunity cost This is the cost incurred by doing one thing rather than another, even if there is no "actual cost". For example, suppose you could sell lemonade or sit on your couch. While sitting on your couch doesn't cost you anything, there is an opportunity cost to it since you could have been making money from selling lemonade. 18

Phillips curve The Phillips curve is an empirical relationship that states that unemployment is inversely related to inflation. It does not always hold, and events like stagflation (stagnation via high unemployment and high inflation) showed that it is not a general phenomenon plural. 55

price discrimination Price discrimination is setting different prices for the same (or very similar) product for different groups based on their willingness to pay. It is a way for a monopoly or firm to get profit by taking away some of the consumer surplus. 21

private good This is a good that is rivalrous and excludable. Markets work extremely well, and are what we usually think of when we think of things bought and sold, such as apples, gasoline, cars, etc. 22

procyclical Procyclical variables are variables that fluctuate positively with business cycle fluctuations in GDP (that is, they change in the same direction as the GDP fluctuations) plural. 61, 66

producer surplus For an individual, this is the difference between the price a seller actually sells for and the minimum price a seller would sell a good for. For all sellers/producers, this is the sum of each individual's producer surplus, and in the continuous case, the lower part of the left most region of a supply-demand curve plot in price vs quantity. 10

production function This is a function that gives the output of an economy. It can be given in real or nominal GDP, for example. It is conventionally given as $Y(t)$ with t time plural. 27

public good This is a good that is non-rivalrous and non-excludable. National defense and air are common examples. 22

purchasing power parity This is the gross domestic product when controlling for inflation. This is then using the same prices over time. 26

RBC model See real business cycle model plural. 60, 61

real business cycle model This is model of economies that uses stochastic difference equations. The business cycles are the cyclic ups and downs in the economy that appear random, but have a somewhat constant-looking frequency and shape plural. 60, 78

real GDP This is the gross domestic product when controlling for inflation. This is then using the same prices over time. 26

rivalrous good A rivalrous good is a good that is "used up" after its use, so that no one else can then use/consume it. For example, an apple is a rivalrous good when sold and eaten. After it is eaten, no one can use it again. Non-rivalrous goods are anything that are not consumed after use [they can benefit someone else] (for example, cinemas and beautiful paintings are non-rivalrous). 22

state variable A state variable is a variable that one does not have control over. It is simply a variable that relates the state of the economy (or whatever is being modeled) at a certain time plural. 44

substitute good A substitute good is a good (in comparison to another good A) whose demand increases when good A's price increases. 11

supply Supply in economics means the amount of goods (quantity) willing to be given a certain price (alternatively, given a quantity of goods, the price they can be sold for). It is usually a monotonically increasing function of quantity, but by no means always monotonically increasing. 7

transversality condition A condition imposed on optimization solutions so that the optimal solution satisfies certain constraints that occur in the real world. This usually involves setting conditions on initial conditions or boundary conditions of variables at infinity plural. 45

tying This is a form of price discrimination where one good is "tied" to another good such that the original purchase requires the tied good in order to function. Printers and printer ink are an example (printers require the printer ink cartridges), as are razors and razor blades. 21

variable cost This is the cost for a business that does change when increasing or decreasing production. Usually the variable costs increase as the amount of product produced increases. 18

Veblen good This is a good with an positive slope demand curve. This means an increase in price actually causes an increase in demand. The way this works is that the good is a luxury item like a Rolex. Then an increase in price may make it more exclusive (hence a better status signal) and so can increase the demand for such an exclusive good. Essentially, we are not measuring the benefit of status with price here. 7