Intermediate Micro Lecture Notes

Dirk Bergemann
Yale University
Spring 2010

Contents

1 Gains from Trade

1.1 Bilateral Trading

Suppose that a seller values a single, homogeneous object at c (opportunity cost), and a potential buyer values the same object at v (willingness to pay). Trade could occur at a price p, in which case the payoff to the seller is $p-c$ and to the buyer is $v-p$. We assume for now that there is only one buyer and one seller, and only one object that can potentially be traded. If no trade occurs, both agents receive a payoff of 0.

Whenever $v > c$ there is the possibility for a mutually beneficial trade at some price $c \leq p \leq v$. Any such allocation results in both players receiving non-negative returns from trading and so both are willing to participate ($p-c$ and $v-p$ are non-negative).

There are many prices at which trade is possible. And each of these allocations, consisting of whether the buyer gets the object and the price paid, is efficient in the following sense:

Definition 1.1. An allocation is Pareto efficient if there is no other allocation that makes at least one agent strictly better off, without making any other agent worse off.

1.2 Experimental Evidence

This framework can be extended to consider many buyers and sellers, and to allow for production. One of the most striking examples comes from international trade. We are interested, not only in how specific markets function, but also in how markets should be organized or designed.

There are many examples of markets, such as the NYSE, NASDAQ, E-Bay and Google. The last two consist of markets that were recently created where they did not exist before. So we want to consider not just existing markets, but also the creation of new markets.

Before elaborating on the theory, we will consider three experiments that illustrate how these markets function. We can then interpret the results in relation to the theory. Two types of cards (red and black) with numbers between 2 and 10 are handed out to the students. If the student receives a red card they are a seller, and the number reflects their cost. If the student receives a black card they are a buyer, and this reflects their valuation. The number on the card is private information. Trade then takes place according to the following three protocols.

1. Bilateral Trading: One seller and one buyer are matched before receiving their cards. The buyer and seller can only trade with the individual they are matched with. They have 5 minutes to make offers and counter offers and then agree (or not) on the price.

2. Pit Market: Buyer and seller cards are handed out to all students at the beginning. Buyers and sellers then have 5 minutes to find someone to trade with and agree on the price to trade.

3. Double Auction: Buyer and seller cards are handed out to all students at the beginning. The initial price is set at 6 (the middle valuation). All buyers and sellers who are willing to trade at this price can trade. If there is a surplus of sellers the price is decreased, and if there is a surplus of buyers then the price is increased. This continues for 5 minutes until there are no more trades taking place.

The outcomes of these experiments are interpreted in the first problem set

1.3 Multiple Buyers and Multiple Sellers

Suppose now that there are N potential buyers and M potential sellers. We can order their valuations and costs by

$$\begin{aligned} v_1 > v_2 > \cdots > v_N > 0 \\ 0 < c_1 < c_2 < \cdots < c_M \end{aligned}$$

We still maintain the assumption that buyers and sellers only demand or supply at most one object each. It is possible that $N \neq M$, but we assume that the number of potential traders is finite.

It is possible to realize the efficient trades through a simple procedure. We match the remaining highest valuation buyer with the lowest cost remaining seller until trade is no longer efficient. There will be n^* trades where $v_{n^*} \geq c_{n^*}$ and $v_{n^*+1} < c_{n^*+1}$ (see Figure 1).

The value generated by the market when n people trade is aggregate value of those who buy minus the costs of producing. The gross surplus of the buyers is given by

$$V(n) = \sum_{k=1}^{n} v_k$$

and the gross cost of producing is

$$C(n) = \sum_{k=1}^{n} c_k$$

The efficient number of trades $n^* = \arg\max_n \{V(n) - C(n)\}$. The marginal decision then is to find n^* such that

$$\begin{aligned} V(n^*) - V(n^* - 1) &\geq C(n^*) - C(n^* - 1) \\ V(n^* + 1) - V(n^*) &< C(n^* + 1) - C(n^*) \end{aligned}$$

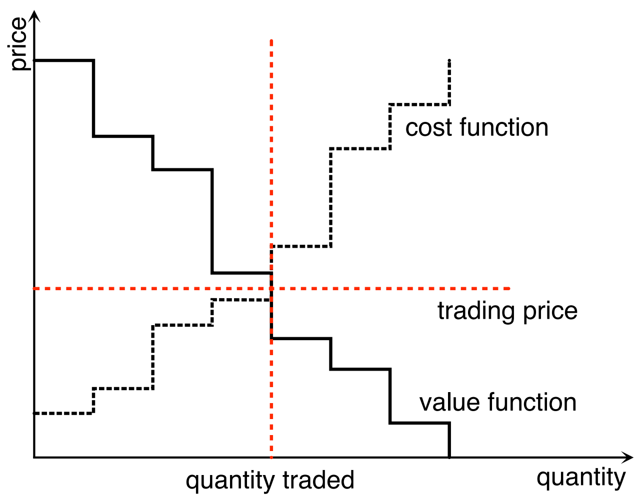

Figure 1: Trading with multiple buyers and sellers

which is, of course, equivalent to

$$v_{n^*} \geq c_{n^*}$$
$$v_{n^*+1} < c_{n^*+1}$$

What about price? If $p^* \in [c_{n^*}, v_{n^*}]$ then the market will exhaust all efficient trading opportunities. If there is a market maker who sets price p^* then we have that for all $n \leq n^*$

$$v_n - p^* \geq 0$$
$$p^* - c_n \geq 0$$

and for all $n > n^*$

$$v_n - p^* < 0$$
$$p^* - c_n < 0$$

So it is possible to achieve the efficient outcome with a single price. In this way we can use prices to decentralize trade.

Some of the matches that make sense in a bilateral trading market will not make sense in a pit market. Consider, for example, the match of v_1 and c_{n^*+1}. Since $v_1 - c_{n^*+1} > 0$, if that buyer and seller were alone together on an island, it would make sense for them to trade. However, once that island is integrated with the rest of the economy (perhaps a ferry service allows the traders on this island to reach other islands), that pair would break down in favor of a more efficient trade. Specifically, the buyer would be able to find someone to sell them the object at a price lower then c_{n^*+1}.

There are several important issues associated with the description of pit market.

- How are prices determined?
- The buyers and sellers can only buy/provide one unit of the good.
- There is only one type of good for sale, and all units are identical.
- There is complete information about the going price and also the valuations (the seller does not know anything that would affect the buyer's valuation).

Above, we considered the case where there were a finite number of buyers and sellers. We could instead consider the case where there are a continuum of buyers and sellers, and replace consideration of the number of trades with the fraction of buyers who purchase the object. The marginal decision then goes from considering $n-(n-1)$ to considering dn: a small change in the proportion buyers who buy. We can then write the aggregate valuations and costs as

$$V(n) = \int_0^n v_k dk$$

$$C(n) = \int_0^n c_k dk$$

and the optimality condition becomes

$$\frac{dV(n^*)}{dn} = \frac{dC(n^*)}{dn}$$

which is, of course, true if and only if $v_{n^*} = c_{n^*}$. This is then a "serious marginal condition" that the marginal buyer and marginal seller must have the same valuation. This, in turn, uniquely ties down the price as $p^* = v_{n^*} = c_{n^*}$. We can then consider "life at the margin" and see what the effects of small changes are on the economy.

2 Choice

In the decision problem in the previous section, the agents had a binary decision: whether to buy (sell) the object. However, there are usually more than two alternatives. The price at which trade could occur, for example, could take on a continuum of values. In this section we will look more closely at preferences, and determine when it is possible to represent preferences by "something handy," which is a utility function.

Suppose there is a set of alternatives $X = \{x_1, x_2, \ldots, x_n\}$ for some individual decision maker. We are going to assume, in a manner made precise below, that two features of preferences are true.

- There is a complete ranking of alternatives.
- "Framing" does not affect decisions.

We refer to X as a choice set consisting of n alternatives, and each alternative $x \in X$ is a consumption bundle of k different items. For example, the first element of the bundle could be food, the second element could be shelter and so on. We will denote preferences by $\succ$, where $x \succ y$ means that "x is strictly preferred to y." All this means is that when a decision maker is asked to choose between x and y they will choose x. Similarly, $x \succsim y$, means that "x is weakly preferred to y" and $x \sim y$ indicates that the decision maker is "indifferent between x and y." The preference relationship $\succsim$ defines an ordering on $X \times X$. We make the following three assumptions about preferences.

Axiom 2.1. *Completeness. For all $x, y \in X$ either $x \succ y$, $y \succ x$, or $y \sim x$.*

This first axiom simply says that, given two alternatives the decision maker can compare the alternatives, and will strictly prefer one of the alternatives, or will be indifferent.

Axiom 2.2. *Transitivity. For all triples $x, y, z \in X$ if $x \succ y$ and $y \succ z$ then $x \succ z$.*

Very simply, this axiom imposes some level of consistency on choices. For example, suppose there were three potential travel locations, Tokyo (T), Beijing (B), and Seoul (S). If a decision maker, when offered the choice between Tokyo and Beijing chose to go to Tokyo, and when given the choice between Beijing and Seoul choose to go to Beijing, then this axiom simply says that if they were offered a choice between a trip to Tokyo or a trip to Seoul, they would choose Tokyo. This is because they have already demonstrated that they prefer Tokyo to Beijing, and Beijing to Seoul, so preferring Seoul to Tokyo would mean that their preferences are inconsistent.

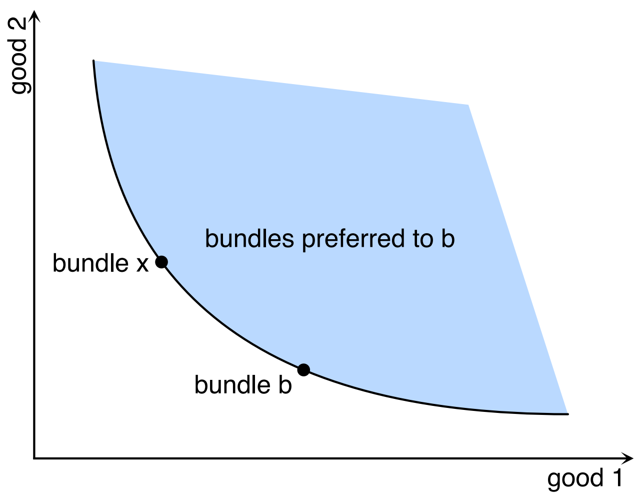

Figure 2: Indifference curve

Axiom 2.3. *Reflexivity. For all $x \in X$, $x \succsim x$ (equivalently, $x \sim x$).*

The final axiom is made for technical reasons, and simply says that a bundle cannot be strictly preferred to itself. Such preferences would not make sense.

These three axioms allow for bundles to be ordered in terms of preference. In fact, these three conditions are sufficient to allow preferences to be represented by a utility function.

Before elaborating on this, we consider an example. Suppose there are two goods, Wine and Cheese. Suppose there are four consumption bundles $z = (2, 2), y = (1, 1), a = (2, 1), b = (1, 2)$ where the two elements of the vector represent the amount of wine or cheese. Most likely, $z \succ y$ since it provides more of everything (i.e., wine and cheese are "goods"). It is not clear how to compare a and b. What we can do is consider which bundles are indifferent with b. This is an indifference curve (see Figure 2). We can define it as

$$I_b = \{x \in X | b \sim x)$$

We can then (if we assume that more is better) compare a and b by considering which side of the indifference curve a lies on: bundles above and to the right are more preferred, bundles below and to the left are less preferred. This reduces the dimensionality of the problem. We can speak of the "better than b" set as the set of points weakly preferred to b. These preferences are "ordinal:" we can ask whether x is in the better than set, but this does not tell us how much x is preferred to b. It is common to assume that preferences are monotone: more of a good is better.

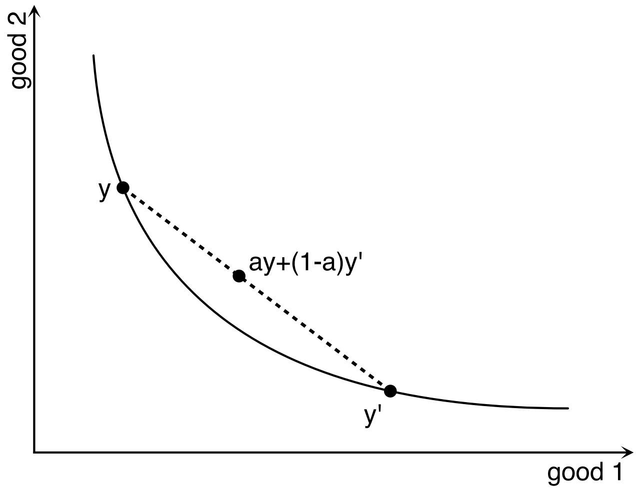

Figure 3: Convex preferences

Definition 2.4. The preferences $\succsim$ are said to be (strictly) monotone if $x \geq y \Rightarrow x \succsim y$ ($x \geq y, x \neq y \Rightarrow x \succsim y$ for strict monotonicity).[1]

Suppose I want to increase my consumption of good 1 without changing my level of well-being. The amount I must change x_2 to keep utility constant, $\frac{dx_2}{dx_1}$ is the marginal rate of substitution. Most of the time we believe that individuals like moderation. This desire for moderation is reflected in convex preferences. A mixture between two bundles, between which the agent is indifferent, is strictly preferred to either of the initial bundle (see Figure 3).

Definition 2.5. A preference relation is convex if for all y and y' with $y \sim y'$ and all $\alpha \in [0, 1]$ we have that $\alpha y + (1 - \alpha)y' \succsim y \sim y'$.

While convex preferences are usually assumed, there could be instances where preferences are not convex. For example, there could be returns to scale for some good.

Examples: perfect substitutes, perfect complements (see Figure 4). Both of these preferences are convex.

Notice that indifference curves cannot intersect. If they did we could take two points x and y, both to the right of the indifference curve the other lies on. We would then have $x \succ y \succ x$, but then by transitivity $x \succ x$ which contradicts reflexivity. So every bundle is associated with one, and only one, welfare level.

[1] If $x = (x_1, \ldots, x_N)$ and $y = (y_1, \ldots, y_N)$ are vectors of the same dimension, then $x \geq y$ if and only if, for all i, $x_i \geq y_i$. $x \neq y$ means that $x_i \neq y_i$ for at least one i.

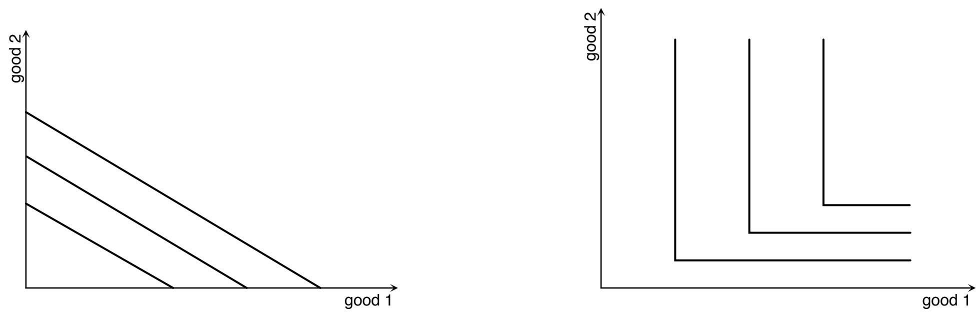

Figure 4: Perfect substitutes (left) and perfect complements (right)

2.1 Utility Functions

What we want to consider now is whether we can take preferences and map them to some sort of utility index. If we can somehow represent preferences by such a function we can apply mathematical techniques to make the consumer's problem more tractable. Working with preferences directly requires comparing each of a possibly infinite number of choices to determine which one is most preferred. Maximizing an associated utility function is often just a simple application of calculus. If we take a consumption bundle $x \in \mathbb{R}_+^N$ we can take a utility function as a mapping from $\mathbb{R}_+^N$ into $\mathbb{R}$.

Definition 2.6. A utility function (index) $u : X \to \mathbb{R}$ represents a preference profile $\succsim$ if and only if, for all $x, y \in X$: $x \succsim y \Leftrightarrow u(x) \geq u(y)$.

We can think about a utility function as an "as if"-concept: the agent acts "as if" she has a utility function in mind when making decisions.

Is it always possible to find such a function? The following result, due to Gerard Debreu, shows that such a function exists under the three assumptions about preferences we made above.

Proposition 2.7. *Every (continuous) preference ranking can be represented by a (continuous) utility function.*

This result can be extended to environments with uncertainty, as was shown by Leonard Savage. Consequently, we can say that individuals behave as if they are maximizing utility functions, which allows for marginal and calculus arguments. There is, however, one qualification. The utility function that represents the preferences is not unique.

Remark 2.8. If u represents preferences, then for any increasing function $f : \mathbb{R} \to \mathbb{R}$, $f(u(x))$ also represents the same preference ranking

In the previous section, we claimed that preferences usually reflect the idea that "more is better," or that preferences are monotone.

Definition 2.9. The utility function (preferences) are monotone increasing if $x \geq y$ implies that $u(x) \geq u(y)$ and $x > y$ implies that $u(x) > u(y)$.

One feature that monotone preferences rule out is (local) satiation, where one point is preferred to all other points nearby. For economics the relevant decision is maximizing utility subject to limited resources. This leads us to consider constrained optimization.

2.2 Budget Constraints

A budget constraint is a constraint on how much money (income, wealth) an agent can spend on goods. We denote the amount of available income by $M \geq 0$. $x_1, \ldots, x_N$ are the quantities of the goods purchased and $p_1, \ldots, p_N$ are the according prices. Then the budget constraint is

$$\sum_{i=1}^{N} p_i x_i \leq M.$$

As an example, we consider the case with two goods. With monotone preferences we get that $p_1 x_1 + p_2 x_2 = M$, i.e., the agent spends her entire income on the two goods. The points where the budget line intersects with the axes are $x_1 = M/p_1$ and $x_2 = M/p_2$ since these are the points where the agent spends her income on only one good. Solving for x_2, we can express the budget line as a function of x_1:

$$x_2(x_1) = \frac{M}{p_2} - \frac{p_1}{p_2} x_1,$$

where $\frac{dx_2}{dx_1} = -\frac{p_1}{p_2}$ is the slope of the budget line.

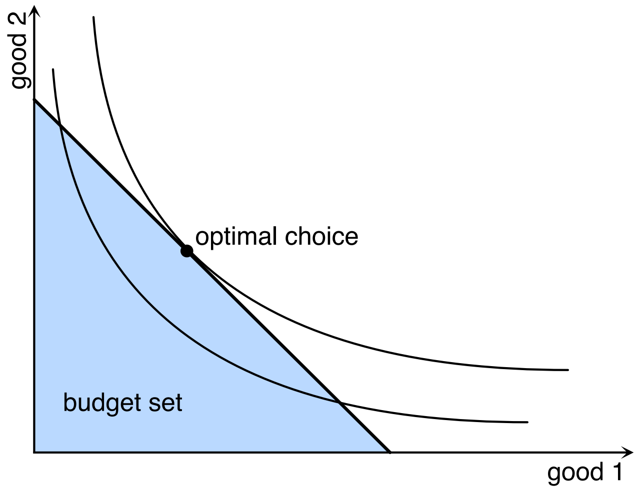

Figure 5: Indifference curve and budget set

3 Utility Maximization

Consumers are typically endowed with money M, which determines which consumption bundles are affordable. The budget set consists of all consumption bundles such that $\sum_{i=1}^{N} p_i x_i \leq M$. The consumer's problem is then to find the point on the highest indifference curve that is in the budget set. At this point the indifference curve must be tangent to the budget line, $\frac{dx_2}{dx_1} = -\frac{p_1}{p_2}$ which defines how much x_2 must decrease if the amount of consumption of good 1 is increased by or dx_1 for the bundle to still be affordable. It reflects the opportunity cost, as money spent on good 1 cannot be used to purchase good 2 (see Figure 5).

The marginal rate of substitution reflects the relative benefit from consuming different goods. We can define it as

$$MRS = \frac{MU_1}{MU_2} = \frac{\frac{\partial u(x)}{\partial x_1}}{\frac{\partial u(x)}{\partial x_2}},$$

where MU_i refers to the marginal utility of consuming good i. The slope of the indifference curve is $-MRS$ so the relevant optimality condition, where the slope of the indifference curve equals the slope of the budget line, is

$$\frac{p_1}{p_2} = \frac{\frac{\partial u(x)}{\partial x_1}}{\frac{\partial u(x)}{\partial x_2}}.$$

We could equivalently talk about equating marginal utility per dollar. If

$$\frac{\frac{\partial u(x)}{\partial x_2}}{p_2} > \frac{\frac{\partial u(x)}{\partial x_1}}{p_1}$$

then one dollar spent on good 2 generates more utility then one dollar spent on good 1. So shifting consumption from good 1 to good 2 would result in higher utility. So, to be at an optimum we must have the marginal utility per dollar equated across goods.

Does this mean then that we must have $\frac{\partial u(x)}{\partial x_i} = p_i$ at the optimum? No. Such a condition wouldn't make sense since we could rescale the utility function. We could instead rescale the equation by a factor $\lambda \geq 0$ that converts "money" into "utility." We could then write $\frac{\partial u(x)}{\partial x_i} = \lambda p_i$. Here, λ reflects the marginal utility of money.

3.1 Optimization by Substitution

The consumer's problem is to maximize utility subject to a budget constraint. There are two ways to approach this problem. The first approach involves writing the last good as a function of the previous goods, and then proceeding with an unconstrained maximization. Consider the two good case. The budget set consists of the constraint that $p_1 x_1 + p_2 x_2 \leq M$. So the problem is

$$\max_{x_1, x_2} u(x_1, x_2) \text{ subject to } p_1 x_1 + p_2 x_2 \leq M$$

But notice that whenever u is (locally) non-satiated then the budget constraint holds with equality since there in no reason to hold money that could have been used for additional valued consumption. So, $p_1 x_1 + p_2 x_2 = M$, and so we can solve for $x_2 = \frac{M - p_1 x_1}{p_2}$. Now we can maximize $u\left(x_1, \frac{M - p_1 x_1}{p_2}\right)$ as the standard single variable maximization problem.

Example 3.1. We consider a consumer with Cobb-Douglas preferences. Cobb-Douglas preferences are easy to use and therefore commonly used. The utility function is defined as (with two goods)

$$u(x_1, x_2) = x_1^{\alpha} x_2^{1-\alpha}, \ \alpha > 0$$

The goods' prices are $p_{1,} p_2$ and the consumer if endowed with income M. Hence, the constraint optimization problem is

$$\max_{x_1, x_2} x_1^{\alpha} x_2^{1-\alpha} \text{ subject to } p_1 x_1 + p_2 x_2 = M.$$

We solve this maximization by substituting the budget constraint into the utility function so that the problem becomes an unconstrained optimization with one choice variable:

$$u(x_1) = x_1^{\alpha} \left(\frac{M - p_1 x_1}{p_2} \right)^{1-\alpha}. \tag{1}$$

In general, we take the total derivative of the utility function

$$\frac{du(x_1, x_2(x_1))}{dx_1} = \frac{\partial u}{\partial x_1} + \frac{\partial u}{\partial x_2}\frac{dx_2}{dx_1} = 0$$

which gives us the condition for optimal demand

$$\frac{dx_2}{dx_1} = -\frac{\frac{\partial u}{\partial x_1}}{\frac{\partial u}{\partial x_2}}.$$

The right-hand side is the marginal rate of substitution (MRS).

In order to calculate the demand for both goods, we go back to our example. Taking the derivative of the utility function (1)

$$\begin{aligned} u'(x_1) &= \alpha x_1^{\alpha-1}\left(\frac{M - p_1x_1}{p_2}\right)^{1-\alpha} + (1-\alpha)x_1^{\alpha}\left(\frac{M - p_1x_1}{p_2}\right)^{-\alpha}\left(-\frac{p_1}{p_2}\right) \\ &= x_1^{\alpha-1}\left(\frac{M - p_1x_1}{p_2}\right)^{-\alpha}\left[\alpha\frac{M - p_1x_1}{p_2} - (1-\alpha)x_1\frac{p_1}{p_2}\right] \end{aligned}$$

so the FOC is satisfied when

$$\alpha(M - p_1x_1) - (1-\alpha)x_1p_1 = 0$$

which holds when

$$x_1^* = \frac{\alpha M}{p_1}. \tag{2}$$

Hence, we see that the budget spent on good 1, p_1x_1, equals the budget share αM, where α is the preference parameter associated with good 1.

Plugging (2) into the budget constraint yields

$$x_2^* = \frac{M - p_1x_1}{p_2} = \frac{(1-\alpha)M}{p_2}.$$

Several important features of this example are worth noting. First of all, x_1 does not depend on p_2 and vice versa. Also, the share of income spent on each good $\frac{p_ix_i}{M}$ does not depend on price or wealth. What is going on here? When the price of one good, p_2, increases there are two effects. First, the price increase makes good 1 relatively cheaper ($\frac{p_1}{p_2}$ decreases). This will cause consumers to "substitute" toward the relatively cheaper good. There is also another effect. When the price increases the individual becomes poorer in real terms, as the set of affordable consumption bundles becomes strictly smaller. The Cobb-Douglas utility function is a special case where this "income effect" exactly cancels out the substitution effect, so the consumption of one good is independent of the price of the other goods.

3.2 Optimization Using the Lagrange Approach

While the approach using substitution is simple enough, there are situations where it will be difficult to apply. The procedure requires that we know, before the calculation, that the budget constraint actually binds. In many situations there may be other constraints (such as a non-negativity constraint on the consumption of each good) and we may not know whether they bind before demands are calculated. Consequently, we will consider a more general approach of Lagrange multipliers. Again, we consider the (two good) problem of

$$\max_{x_1,x_2} u(x_1,x_2) \text{ s.t. } p_1x_1 + p_2x_2 \leq M$$

Let's think about this problem as a game. The first player, let's call him the kid, wants to maximize his utility, $u(x_1,x_2)$, whereas the other player (the parent) is concerned that the kid violates the budget constraint, $p_1x_1 + p_2x_2 \leq M$, by spending too much on goods 1 and 2. In order to induce the kid to stay within the budget constraint, the parent can punish him by an amount λ for every dollar the kid exceeds his income. Hence, the total punishment is

$$\lambda(M - p_1x_1 - p_2x_2).$$

Adding the kid's utility from consumption and the punishment, we get

$$\mathcal{L}(x_1,x_2,\lambda) = u(x_1,x_2) + \lambda(M - p_1x_1 - p_2x_2). \tag{3}$$

Since, for any function, we have max $f = -\min f$, this game is a zero-sum game: the payoff for the kid is $\mathcal{L}$ and the parent's payoff is $-\mathcal{L}$ so that the total payoff will always be 0. Now, the kid maximizes expression (3) by choosing optimal levels of x_1 and x_2, whereas the parent minimizes (3) by choosing an optimal level of λ:

$$\min_{\lambda} \max_{x_1,x_2} \mathcal{L}(x_1,x_2,\lambda) = u(x_1,x_2) + \lambda(M - p_1x_1 - p_2x_2).$$

In equilibrium, the optimally chosen level of consumption, x^*, has to be the best response to the optimal level of λ^* and vice versa. In other words, when we fix a level of x^*, the parent chooses an optimal $\lambda^*$and when we fix a level of λ^*, the kid chooses an optimal x^*. In equilibrium, no one wants to deviate from their optimal choice. Could it be an equilibrium for the parent to choose a very large λ? No, because then the kid would not spend any money on consumption, but rather have the maximized expression (3) to equal λM.

Since the first-order conditions for minima and maxima are the same, we have the following first-order conditions for problem (3):

$$\frac{\partial \mathcal{L}}{\partial x_1} = \frac{\partial u}{\partial x_1} - \lambda p_1 = 0 \tag{4}$$

$$\frac{\partial \mathcal{L}}{\partial x_2} = \frac{\partial u}{\partial x_2} - \lambda p_2 = 0 \tag{5}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = M - p_1 x_1 - p_2 x_2 = 0.$$

Here, we have three equations in three unknowns that we can solve for the optimal choice x^*, λ^*.

Before solving this problem for an example, we can think about it in more formal terms. The basic idea is as follows: Just as a necessary condition for a maximum in a one variable maximization problem is that the derivative equals 0 ($f'(x) = 0$), a necessary condition for a maximum in multiple variables is that all partial derivatives are equal to 0 ($\frac{\partial f(x)}{\partial x_i} = 0$). To see why, recall that the partial derivative reflects the change as x_i increases and the other variables are all held constant. If any partial derivative was positive, then holding all other variables constant while increasing x_i will increase the objective function (similarly, if the partial derivative is negative we could decrease x_i). We also need to ensure that the solution is in the budget set, which typically won't happen if we just try to maximize u. Basically, we impose a "cost" on consumption (the punishment in the game above), proceed with unconstrained maximization for the induced problem, and set this cost so that the maximum lies in the budget set.

Notice that the first-order conditions (4) and (5) imply that

$$\frac{\frac{\partial u}{\partial x_1}}{p_1} = \lambda = \frac{\frac{\partial u}{\partial x_2}}{p_2}$$

or

$$\frac{\frac{\partial u}{\partial x_1}}{\frac{\partial u}{\partial x_2}} = \frac{p_1}{p_2}$$

which is precisely the "MRS = price ratio" condition for optimality that we saw before.

Finally, it should be noted that the FOCs are necessary for optimality, but they are not, in general, sufficient for the solution to be a maximum. However, whenever $u(x)$ is a concave function the FOCs are also sufficient to ensure that the solution is a maximum. In most situations, the utility function will be concave.

Example 3.2. We can consider the problem of deriving demands for a Cobb-Douglas utility function using the Lagrange approach. The associated Lagrangian is

$$\mathcal{L}(x_1, x_2, \lambda) = x_1^{\alpha} x_2^{1-\alpha} + \lambda(M - p_1 x_1 - p_2 x_2),$$

which yields the associated FOCs

$$\frac{\partial \mathcal{L}}{\partial x_1} = \alpha x_1^{\alpha-1} x_2^{1-\alpha} - \lambda p_1 = \alpha \left(\frac{x_2}{x_1}\right)^{1-\alpha} - \lambda p_1 = 0 \tag{6}$$

$$\frac{\partial \mathcal{L}}{\partial x_2} = (1-\alpha) x_1^{\alpha} x_2^{-\alpha} - \lambda p_2 = (1-\alpha) \left(\frac{x_1}{x_2}\right)^{\alpha} - \lambda p_2 = 0 \tag{7}$$

$$\lambda(M - p_1 x_1 - p_2 x_2) = 0. \tag{8}$$

We have three equations with three unknowns (x_1, x_2, λ) so that this system should be solvable. Notice that since it is not possible that $\frac{x_2}{x_1}$ and $\frac{x_1}{x_2}$ are both 0 we cannot have a solution to equations (6) and (7) with $\lambda = 0$. Consequently we must have that $p_1 x_1 + p_2 x_2 = M$ in order to satisfy equation (8). Solving for λ in the above equations tells us that

$$\lambda = \frac{\alpha}{p_1} \left(\frac{x_2}{x_1}\right)^{1-\alpha} = \frac{(1-\alpha)}{p_2} \left(\frac{x_1}{x_2}\right)^{\alpha}$$

and so

$$p_2 x_2 = \frac{1-\alpha}{\alpha} p_1 x_1.$$

Combining with the budget constraint this gives

$$p_1 x_1 + \frac{1-\alpha}{\alpha} p_1 x_1 = \frac{1}{\alpha} p_1 x_1 = M.$$

So the Marshallian[2] demand functions are

$$x_1^* = \frac{\alpha M}{p_1}$$

and

$$x_2^* = \frac{(1-\alpha)M}{p_2}.$$

So we see that the result of the Lagrangian approach is the same as from approach that uses substitution. Using equation (6) or (7) again along with the optimal demand x_1^* or x_2^* gives us the following expression for λ:

$$\lambda^* = \frac{1}{M}.$$

Hence, λ^* equals the derivative of the Lagrangian $\mathcal{L}$ with respect to income M. We call this derivative, $\frac{\partial \mathcal{L}}{\partial M}$, the marginal utility of money.

[2] After the British economist Alfred Marshall.

When we take a monotone transformation of a utility function the underlying preferences represented are not changed. Consequently the consumer's demands are unaffected by such a transformation. The following example shows this using a monotone transformation of the Cobb-Douglas utility function.

Example 3.3. Notice that

$$\ln\left(x_1^\alpha x_2^{1-\alpha}\right) = \alpha \ln x_1 + (1-\alpha)\ln x_2$$

so we could write the utility function for Cobb-Douglas preferences by $u(x_1, x_2) = \alpha \ln x_1 + (1-\alpha)\ln x_2$. The associated Lagrangian is

$$\mathcal{L}(x_1, x_2, \lambda) = \alpha \ln x_1 + (1-\alpha)\ln x_2 + \lambda(M - p_1x_1 - p_2x_2)$$

with associated FOCs

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial x_1} &= \frac{\alpha}{x_1} - \lambda p_1 = 0 \\
\frac{\partial \mathcal{L}}{\partial x_2} &= \frac{1-\alpha}{x_2} - \lambda p_2 = 0 \\
\lambda(M - p_1x_1 - p_2x_2) &= 0.
\end{aligned}$$

Since $\frac{\alpha}{x_1} > 0$ we must have $\lambda > 0$ so the budget constraint must hold with equality. Solving for λ gives

$$\lambda = \frac{\alpha}{p_1x_1} = \frac{1-\alpha}{p_2x_2}$$

so

$$p_2x_2 = \frac{1-\alpha}{\alpha}p_1x_1.$$

and

$$\frac{1-\alpha}{\alpha}p_1x_1 + p_1x_1 = \frac{1}{\alpha}p_1x_1 = M.$$

So, we can conclude, as in the previous example, that

$$x_1 = \frac{\alpha M}{p_1}$$

and

$$x_2 = \frac{(1-\alpha)M}{p_2}.$$

After we calculate the optimal choice (or Marshallian demand functions), we can plug them into the original utility function to get the *indirect utility function*, which is a function of prices and income only:

$$V(p_1, p_2, M) = u\left(x_1^*(p_1, p_2, M), x_2^*(p_1, p_2, M)\right).$$

It turns out that $\frac{\partial V}{\partial M} = \lambda^* = \frac{\partial \mathcal{L}}{\partial M}$, which is therefore also called the *marginal value of income*.

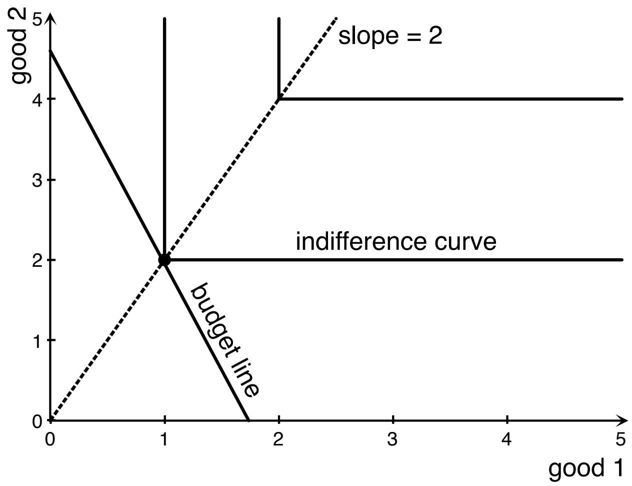

Figure 6: Leontief preferences

4 Demand

So far we have considered only the decisions of single agents trying to make the optimal decisions given some exogenous parameters and constraints. In the next sections we will integrate these individual decision makers with other decision makers and consider when the resulting system is in "equilibrium." First, we will consider the behaviour of demands.

4.1 Comparative Statics

After having derived the optimal demand functions $x_1^*(p_1, p_2, M)$ and $x_2^*(p_1, p_2, M)$ we can ask how a change in the exogenous variables (p_1, p_2, and M) affects the endogenous variables (x and y). This exercise is known as *comparative statics*. For example, a firm might be interested in how its revenue $p_i x_i$ changes when it increases the price of its good, p_i.

Example 4.1. Consider the following utility function, also known as Leontief utility:[3]

$$u(x_1, x_2) = \min\{2x_1, x_2\}.$$

A consumer with this type of utility function only achieves a higher utility level when the quantity of both goods increases. An example is left and right shoes: many left shoes do not increase utility unless they are matched by the same number of right shoes. The indifference curves for the utility function are depicted in Figure 6. The utility maximization

[3]After Wassily Leontief, a Russian economist, sought after in vain by the CIA, who used this functional form in his analysis of production technology and growth.

problem is

$$\max_{x_1,x_2} \min\{2x_1, x_2\} \text{ s.t. } p_1x_1 + p_2x_2 = M. \tag{9}$$

Looking at the indifference curves, we see that optimal consumption bundles satisfy the condition $2x_1 = x_2$, i.e., optimal bundles are located at the kinks of the indifference curves. Starting at one of the kinks, consuming more of good 1 or more of good 2 does not increase utility. Hence we can find the solution to problem (9) as the intersection of the line $2x_1 = x_2$ with the budget line. We solve by substituting the optimality condition into the budget constraint:

$$p_1x_1 + p_2(2x_1) = M.$$

Hence, the optimal demand functions are

$$x_1^*(p_1, p_2, M) = \frac{M}{p_1 + 2p_2}$$

and, by plugging x_1^* into the optimality condition,

$$x_2^*(p_1, p_2, M) = \frac{2M}{p_1 + 2p_2}.$$

Note that we cannot equate the MRS with the slope of the budget line here, because the MRS is not defined at the point where $2x_1 = x_2$. Next, we do comparative statics by taking the derivative of the endogenous variables with respect to the exogenous variables (using the quotient rule for derivatives):

$$\frac{\partial x_1^*}{\partial p_1} = -\frac{M}{(p_1 + 2p_2)^2} < 0,$$

which is called the *own-price effect on demand*, and

$$\frac{\partial x_1^*}{\partial p_2} = -\frac{2M}{(p_1 + 2p_2)^2} < 0,$$

which is the *cross-price effect on demand*. Note that the demand for good 1 decreases when p_1 increases and when p_2 increases.

In general, we say that goods i and j are *substitutes* if the cross-price effect is positive, i.e., demand for good i increases when the price of good j increases:

$$\frac{\partial x_i^*}{\partial p_j} \geq 0$$

and they are called *complements* if the cross-price effect is negative:

$$\frac{\partial x_i^*}{\partial p_j} < 0.$$

Hence, in the Leontief example, the two goods are complements. In fact, they are called *perfect complements* in this case. Besides the price effects, we can also calculate the income effect, which is defined as

$$\frac{\partial x_i^*}{\partial M} \gtrless 0.$$

4.2 Elasticities

When calculating price effects, the result depends on the units used. For example, when considering the own-price effect for gasoline, we might express quantity demanded in gallons or liters and the price in dollars or euros. The own-price effects would differ even if consumers in the U.S. and Europe had the same underlying preferences. In order to make price effects comparable across different units, we can normalize them by dividing by $\frac{x_i}{p_i}$. This is the *own-price elasticity of demand* and denoted by ϵ:

$$\epsilon = -\frac{\frac{\partial x_i}{\partial p_i}}{\frac{x_i}{p_i}} = -\frac{\partial x_i}{\partial p_i}\frac{p_i}{x_i}.$$

It is common to multiply the price effect by -1 so that ϵ is a positive number since the price effect is usually negative. Of course, the *cross-price elasticity of demand* is defined similarly

$$-\frac{\frac{\partial x_i}{\partial p_j}}{\frac{x_i}{p_j}} = -\frac{\partial x_i}{\partial p_j}\frac{p_j}{x_i}.$$

Price elasticities measure by how many percent demand changes in response to a one percent change in price.

Price elasticities play an important role in economics. Consider a firm trying to maximize the revenue generated by demand for its good. Revenue is the product of quantity demanded and price:

$$R = x_i(p)p_i.$$

The firm chooses the price of its good, p_i, in order to maximize R. The effect of price on revenue is, using the chain rule:

$$\frac{dR}{dp_i} = x_i(p) + \frac{dx_i}{dp_i}p_i.$$

Factoring out $x_i(p)$ we can rewrite this as

$$\frac{dR}{dp_i} = x_i(p)\left[1 + \frac{dx_i}{dp_i}\frac{p_i}{x_1(p)}\right] = x_i(p)\,(1-\epsilon)\,.$$

Hence, the effect of price on revenue depends on the own-price elasticity of demand. We have that

$$\frac{dR}{dp_i}\begin{cases} >0 & \text{if } \epsilon < 1 \\ =0 & \text{if } \epsilon = 1 \\ <0 & \text{if } \epsilon > 1 \end{cases}.$$

In the first case, demand is called *inelastic*, in the second case, it has unit elasticity, and in the third case, it is called *elastic*. Intuitively, price has two effects on revenue: a higher price increases revenue since the latter is the product of quantity and price, but a higher price also decreases the quantity demanded as long as the own-price effect is negative. Hence, the elasticity measures which of the two opposite effects dominates.

4.3 Income and Substitution Effects, Expenditure Minimization

For given prices and wealth an individual has demands $(x_1^*(p_1,p_2,M),\ x_2^*(p_1,p_2,M))$. As we've seen above, the demand for good 1 can be affected by changes in its own price, $\frac{\partial x_1}{\partial p_1}$, or by a cross-price effect $\frac{\partial x_1}{\partial p_2}$.

When the price of good 2 changes there are two effects. First, it makes the prices ratio $\frac{p_1}{p_2}$ change. This is known as the *substitution effect*. Second, the change in p_2 causes a change in the consumption bundles that are feasible, so an increase in p_2 makes the consumer poorer, while a decrease makes individuals richer. This effect is known as the *income effect*. Together these two effects constitute a price effect.

Example 4.2. One example where the substitution effect does not matter is the Leontieff utility function. Goods are perfect complements and so there is no substitution. Here, $u(x) = \min\{ax_1, bx_2\}$ so it will be optimal to set $ax_1 = bx_2$ whatever the prices of the two goods may be. So $x_2 = \frac{ax_1}{b}$ and from the budget set $p_1x_1 + p_2x_2 = \left(p_1 + \frac{ap_2}{b}\right)x_1 = M$. So

$$x_1 = \frac{bM}{bp_1 + ap_2}$$
$$x_2 = \frac{aM}{bp_1 + ap_2}$$

If the price of good 1 were to fall the consumption of both goods will increase, since the consumer is made wealthier (more bundles are now affordable).

Now we express income and substitution effects analytically. To do so, we introduce expenditure minimization first. Recall the general utility maximization problem from section 3. With many goods and prices we can write it as (writing x and p are vectors with I components each)

$$\max_x u(x) \text{ s.t. } p \cdot x \leq M.$$

The resulting demand function is denoted by $x^*(p, M)$ and also called *uncompensated demand* or Marshallian demand. Plugging in the optimal demand into the utility function yields *indirect utility* as a function of prices and income:

$$V(p, M) = u(x^*(p, M)).$$

Instead of maximizing utility subject to a given income we can also minimize expenditure subject to achieving a given level of utility $\bar{u}$. In this case, the consumer wants to spend as little money as possible to enjoy a certain utility. Formally, we write

$$\min_x p \cdot x \text{ s.t. } u(x) \geq \bar{u}. \tag{10}$$

The result of this optimization problem is a demand function again, but in general it is different from $x^*(p, M)$. We call the demand function derived from problem (10) *compensated demand* or Hicksian demand[4] and denote it by $x^c(p, \bar{u})$. Note that compensated demand is a function of prices and the utility level whereas uncompensated demand is a function of prices and income. Plugging compensated demand into the objective function $(p \cdot x)$ yields the *expenditure function* as function of prices and $\bar{u}$

$$E(p, \bar{u}) = p \cdot x^c(p, \bar{u}).$$

Hence, the expenditure measures the minimal amount of money required to buy a bundle that yields a utility of $\bar{u}$.

Uncompensated and compensated demand functions usually differ from each other, which is immediately clear from the fact that they different arguments. There is a special case where they are identical. First, note that indirect utility and expenditure function are related by the following relationships

$$V(p, E(p, \bar{u})) = \bar{u}$$
$$E(p, V(p, M)) = M.$$

That is, if income is exactly equal to the expenditure necessary to achieve utility level $\bar{u}$, then the resulting indirect utility is equal to $\bar{u}$. Similarly, if

[4]After the British economist Sir John Hicks, co-recipient of the 1972 Nobel Prize in Economic Sciences.

the required utility level is set equal to the indirect function when income is M, then minimized expenditure will be equal to M. Using these relationships, we have that uncompensated and compensated demand are equal in the following two cases:

$$\begin{aligned} x^*(p, M) &= x^c(p, V(p, M)) \\ x^*(p, E(p, \bar{u})) &= x^c(p, \bar{u}). \end{aligned} \tag{11}$$

Now we can express income and substitution effects analytically. Start with one component of equation (11) (recall that x and p are vectors):

$$x_i^c(p, \bar{u}) = x_i^*(p, E(p, \bar{u}))$$

and take the derivative with respect to p_j using the chain rule

$$\frac{\partial x_i^c}{\partial p_j} = \frac{\partial x_i^*}{\partial p_j} + \frac{\partial x_i^*}{\partial M}\frac{\partial E}{\partial p_j}. \tag{12}$$

Now we have to find an expression for $\frac{\partial E}{\partial p_j}$. Start with the Lagrangian associated with problem (10) evaluated at the optimal solution ($x^c(p, \bar{u})$, $\lambda^*(p, \bar{u})$):

$$\mathcal{L}(x^c(p, \bar{u}), \lambda^*(p, \bar{u})) = p \cdot x^c(p, \bar{u}) + \lambda^*(p, \bar{u})[\bar{u} - u(x(p, \bar{u}))].$$

Taking the derivative with respect to any price p_j and noting that $\bar{u} = u(x(p, \bar{u}))$ at the optimum we get

$$\begin{aligned} \frac{\partial \mathcal{L}(x^c(p, \bar{u}), \lambda^*(p, \bar{u}))}{\partial p_j} &= x_j^c + \sum_{i=1}^{I} p_i \frac{\partial x_i^c}{\partial p_j} - \lambda^* \sum_{i=1}^{I} \frac{\partial u}{\partial x_i}\frac{\partial x_i}{\partial p_j} \\ &= x_j^c + \sum_{i=1}^{I} \left(p_i - \lambda^* \frac{\partial u}{\partial x_i}\right) \frac{\partial x_i}{\partial p_j}. \end{aligned}$$

But the first -order conditions for this Lagrangian are

$$p_i - \lambda \frac{\partial u}{\partial x_i} = 0 \text{ for all } i.$$

Hence

$$\frac{\partial E}{\partial p_j} = \frac{\partial \mathcal{L}}{\partial p_j} = x_j^c(p, \bar{u}).$$

This result also follows form the Envelope Theorem. Moreover, from equation (11) it follows that $x_j^c = x_j^*$. Hence, using these two facts and bringing the second term on the RHS to the LHS we can rewrite equation (12) as

$$\frac{\partial x_i^*}{\partial p_j} = \underbrace{\frac{\partial x_i^c}{\partial p_j}}_{SE} - x_j^* \underbrace{\frac{\partial x_i^*}{\partial M}}_{IE}.$$

This equation is known as the *Slutsky Equation*[5] and shows formally that the price effect can be separated into a substitution (SE) and an income effect (IE).

4.4 Corner Solutions

Consider the utility function $u(x_1, x_2) = \ln x_1 + x_2$, where p_1, p_2 and M are all strictly greater than 0. The first thing to notice is that while the marginal utility of good 2 is constant $\frac{\partial u}{\partial x_2} = 1$, the marginal utility of good 1, $\frac{\partial u}{\partial x_1} = \frac{1}{x_1}$ is a decreasing function of x_1. The second derivative is negative, $\frac{\partial^2 u}{\partial x_1^2} = -\frac{1}{x_1^2} < 0$ because the natural logarithm is a concave function. That is, there is diminishing marginal utility from consuming good 1. Suppose that we consider the problem of maximizing $u(x) = \ln x_1 + x_2$ subject to the budget constraint, $p_1 x_1 + p_2 x_2 \leq M$. We would expect that the size of the budget would greatly affect the shape of the demands. The consumer will spend everything on good 1 as long as

$$\frac{MU_1}{p_1} \geq \frac{MU_2}{p_2}.$$

Once x_1 is large enough that

$$\frac{MU_1}{p_1} = \frac{MU_2}{p_2}$$

we reach a "critical point" and it becomes more profitable to spend the remaining money on good 2. This point occurs when

$$\frac{\frac{1}{x_1}}{p_1} = \frac{1}{p_2}$$

So $x_1^* = \frac{p_2}{p_1}$. Hence, the first $p_1 x_1^*$ dollars will be spent on good 1, and any remaining money will be spent on good 2. The maximum amount to be spent on good 1 is simply $m^* = p_1 x_1^* = p_2$. So we can now formulate the demand.

$$\left\{ \begin{matrix} x_1 = \frac{M}{p_1} \\ x_2 = 0 \end{matrix} \right\} \text{ if } M < m^* = p_2$$

$$\left\{ \begin{matrix} x_1 = \frac{p_2}{p_1} \\ x_2 = \frac{M - p_2}{p_1} \end{matrix} \right\} \text{ if } M \geq m^* = p_2$$

[5] After the Russian statistician and economist Eugen Slutsky.

In the above example, with a small budget it would be preferable to consume negative amounts of good 2. This negative consumption is called "short selling." In most situations this is not reasonable. What would it mean, for example, to consume -2 apples? Consequently, we often need to add additional constraints to the problem: that the consumption of each good must be non-negative. When such constraints bind we have what is known as a "corner solution." We can deal with these non-negativity constraints exactly the same way as the budget constraint: by adding Lagrange multipliers associated with each constraint. Unlike with the budget constraint which will hold with equality whenever utility functions are monotone, it may not be clear before beginning the calculation which non-negativity constraint will bind. Instead of just having the budget constraint

$$p_1x_1 + p_2x_2 \leq M$$

we also have the constraints

$$\begin{aligned} x_1 &\geq 0 \\ x_2 &\geq 0. \end{aligned}$$

Typically the Lagrange multipliers associated with these constraints are denoted by μ_1 and μ_2.

So the associated Lagrangian is

$$\mathcal{L}(x_1, x_2, \lambda, \mu_1, \mu_2) = u(x_1, x_2) + \lambda(M - p_1x_1 - p_2x_2) + \mu_1x_1 + \mu_2x_2.$$

Note that the Lagrangian has 5 arguments now (2 goods and 3 multipliers). This leads to the following FOCs:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial x_1} &= \frac{\partial u}{\partial x_1} - \lambda p_1 + \mu_1 = 0 \\ \frac{\partial \mathcal{L}}{\partial x_2} &= \frac{\partial u}{\partial x_2} - \lambda p_2 + \mu_2 = 0 \end{aligned}$$

and the complimentary slackness conditions (see section 5.5 for more on this)

$$\lambda(M - p_1x_1 - p_2x_2) = 0 \tag{13}$$

$$\mu_1x_1 = 0 \tag{14}$$

$$\mu_2x_2 = 0. \tag{15}$$

This means we are left with 5 equations in 5 unknowns. In the previous example with $u(x_1, x_2) = \ln x_1 + x_2$, the first two conditions become

$$\frac{1}{x_1} - \lambda p_1 + \mu_1 = 0 \tag{16}$$

$$1 - \lambda p_2 + \mu_2 = 0. \tag{17}$$

For small M, $x_1 > 0$ we have that $x_2 = 0$ and so $\mu_1 = 0$ for condition (14) to hold. But the budget constraint (13) guarantees that

$$x_1 = \frac{M}{p_1}$$

which combined with condition (16) says that $\lambda = \frac{1}{M}$. So, from condition (17)

$$\mu_2 = \frac{p_2}{M} - 1$$

This is the shadow price of short-selling. That is, how valuable relaxing the non-negativity constraint on good 2 would be.

5 A Brief Review of Optimization

Let us start with a function on the real line

$$f : \mathbb{R} \to \mathbb{R}$$

which is continuous and differentiable. We want to identify local and global maxima in a world without constraints, i.e., we are looking at unconstrained optimization:

$$\max_{x \in \mathbb{R}} f(x)$$

x^* is a global maximum if

$$\text{for all } x, \;\; f(x^*) \geq f(x).$$

The necessary conditions for a local or a global maximum at $x = x^*$ are given by the first order condition (FOC):

$$f'(x^*) = 0,$$

and the second order condition

$$f''(x^*) \leq 0.$$

A sufficient condition for a global maximum is given by

$$f'(x^*) = 0,$$

and here comes the new part for the sufficiency condition

$$\text{for all } x, \;\; f''(x) \leq 0.$$

Definition 5.1. A univariate function $f : \mathbb{R} \to \mathbb{R}$ is a concave function if $u''(x) < 0$.

A concave function has only a single (global) maximum and no minimum.

5.1 Constraints

We consider again a continuous and differentiable function

$$f : \mathbb{R} \to \mathbb{R}$$

but now impose constraints on the function, or on the search

$$x \in [-\infty, \bar{x}]$$

for some $\underline{x} < \overline{x}$. In other words we are looking for

$$\max \; f(x)$$

subject to

$$\overline{x} - x \geq 0$$

The inequalities represent the constraint which simply say that we can look for the x which maximizes $f(x)$ but only in the range of the interval $[-\infty, \bar{x}]$.

If the optimal $x^* = \bar{x}$, then we say that x^* is a *corner solution*. It is easy to see that a corner solution may not satisfy the first order condition:

$$f'(x^*) = 0,$$

because the decision maker would like to move his choice in a direction with a positive gradient but the constraint does not allow him to go beyond $\overline{x}$. For example if $x^* = \overline{x}$, then we might have at $x = x^*$:

$$f'(x^*) \geq 0.$$

We now would like to introduce a general method so that the first order conditions, namely the idea that there are no further net gains to be made, is reestablished. Second we would like to integrate the constraints directly in the objective function so that the problem becomes to look more like an unconstrained problem, and hence more familiar with us. Finally, we would like to reinterpret the role of constraint in terms of a price.

5.2 Lagrange

Joseph-Louis Lagrange, comte de l'Empire (January 25, 1736 – April 10, 1813) was an Italian-French mathematician and astronomer who made important contributions to all fields of analysis and number theory and to classical and celestial mechanics as arguably the greatest mathematician of the 18th century. It is said that he was able to write out his papers complete without a single correction required. Before the age of 20 he was professor of geometry at the royal artillery school at Turin. By his mid-twenties he was recognized as one of the greatest living mathematicians because of his papers on wave propagation and the maxima and minima of curves.

He was born, of French and Italian descent, in Turin. His father, who had charge of the Kingdom of Sardinia's military chest, was of good social position and wealthy, but before his son grew up he had lost most of his property in speculations, and young Lagrange had to rely on his own abilities for his position. He was educated at the college of Turin, but it was not until he was seventeen that he showed any taste for mathematics – his interest in the subject being first excited by a paper by Edmund Halley which he came across by accident. Alone and unaided he threw himself into mathematical studies; at the end of a year's incessant toil he was already an accomplished mathematician, and was made a lecturer in the artillery school.

In mathematical optimization problems, Lagrange multipliers, named after Joseph Louis Lagrange, is a method for finding the local extrema of a function of several variables subject to one or more constraints. This method reduces a problem in n variables with k constraints to a solvable problem in $n + k$ variables with no constraints. The method introduces a new unknown scalar variable, the Lagrange multiplier, for each constraint and forms a linear combination involving the multipliers as coefficients.

5.3 Constrained Optimization

We are looking for a solution of the following problem

$$\max_{x \in \mathbb{R}} \; f(x)$$

subject to

$$\overline{x} - x \geq 0.$$

We approach this problem by associating a Lagrange multiplier to the constraint, and it is typically a Greek letter, say $\lambda \in \mathbb{R}_+$, and define a function on x and λ, called the Lagrangian function, $\mathcal{L}$:

$$\mathcal{L}(x, \lambda) = f(x) + \lambda(\overline{x} - x).$$

We now have a new function, in two (x, λ) rather than one (x) variable. We impose the following constraint on the shape of the Lagrangian

$$\lambda(\overline{x} - x) = 0 \tag{18}$$

for all x and all λ. The new constraint, (18), is often called the *complementary slackness constraint*. We now maximize the Lagrange an with respect to x (and as it turns out also minimize it with respect to λ):

$$\max_{x} \; \mathcal{L}(x, \lambda).$$

We now look at the first order conditions of the unconstrained problem and find the first order condition

$$f'(x^*) - \lambda x^* = 0$$

and the auxiliary complementary slackness constraint:

$$\lambda\left(\overline{x} - x^*\right) = 0.$$

If x^* is a corner solution with $x^* = \overline{x}$, then $\lambda \geq 0$ and we can interpret λ as the price (of violating the constraint). In the process we have replaced the strict constraint with a price for the constraint (equal to the Lagrange multiplier). At the optimum the price is equal to the marginal value of the objective function. For this reason we refer to it as the *shadow price* of the constraint. It is the smallest price we can associate to the constraint so that the decision maker, facing this price would respect the constraint.

5.4 Example

Consider the following function

$$f\left(x\right) = a - \left(x - b\right)^2$$

with $a, b > 0$.

1. Graphically display the function, label a and b on the respective axis. Solve analytically for the unconstrained maximum of this function.

2. Suppose now that the choice of the optimal x is constrained by $x \leq \overline{x}$, where $\overline{x}$ is an arbitrary number, satisfying

$$0 < \overline{x} < b.$$

   (a) Graphically display the nature of the new constraint relative to the optimization problem and the function f.

   (b) Solve via the Lagrangian method for the optimal solution.

   (c) **Lagrange multiplier.** Suppose that we increase the bound $\overline{x}$ by dx. What is the marginal effect this has on the value of the function to be optimized in terms of a and b. How does it compare to the value of the Lagrangian multiplier which you just computed.

   (d) **Interpretation of Lagrangian multiplier.** Imagine now that we cannot insist that the decision maker respect the constraint but that we can ask a penalty, say λ, for every unit of x over and above $\overline{x}$. What is the price that we would have to charge so that the decision maker would just be happy to exactly choose $x^* = \overline{x}$ and thus in fact respect the constraint even so did not face the constraint directly.

In the process we have replaced the strict constraint with a price for the constraint (equal to the Lagrange multiplier). At the optimum the price is equal to the marginal value of the objective function. For this reason we refer to it as the *shadow price* of the constraint. It is the smallest price we can associate to the constraint so that the decision maker, facing this price would respect the constraint.

5.5 Understanding Lagrangians

Recall that we can solve the general utility maximization problem by using the Lagrange approach as follows:

$$\max_{x_1,x_2}\min_{\lambda}\ \mathcal{L}(x_1,x_2,\lambda),$$

where the Lagrangian $\mathcal{L}$ is defined as

$$\mathcal{L}(x_1,x_2,\lambda)=u(x_1,x_2)+\lambda(M-p_1x_1-p_2x_2),$$

so that the resulting problem is an unconstrained optimization problem. Graphically we can interpret this approach as follows: Imagine we can choose a value of x on the north-south axis to maximize $\mathcal{L}$, i.e., we are "climbing a mountain" along this axis. Along the east-west axis we pick a λ to minimize $\mathcal{L}$. The solution to this problem is a "saddle point:" a point that is the lowest (along the east-west axis) among the highest point along the north-south axis.

Before we proceed we consider the simplest possible constrained problem, the one-dimensional maximization problem, to develop a feel for how this technique works. Let $f(x)$ be a function from the real line into itself (for example the solid black one depicted in Figure 7). The problem is to choose x to maximize $f(x)$. In the unconstrained problem, the FOC $f'(x^*)=0$ is a sufficient condition for a local extremum. To ensure that said extremum is in fact a local maximum we must consider the second-order condition $f''(x^*)\leq 0$. What if we are interested in a global maximum? Then we could replace the above SOC with $f''(x)\leq 0$ for all x. This will guarantee that the function is concave and so any local maximum must be a global maximum.

Let $\hat{x}=\arg\max f(x)$, and consider some $\bar{x}<\hat{x}$ and suppose instead that we are considering the problem of maximizing $f(x)$ subject to $x\leq\bar{x}$. Assume that f is concave at $\hat{x}$. We know that $f'(\hat{x})=0$ but $\hat{x}$ is not available in the new problem. Is there an analogous expression for the solution to the constrained problem?

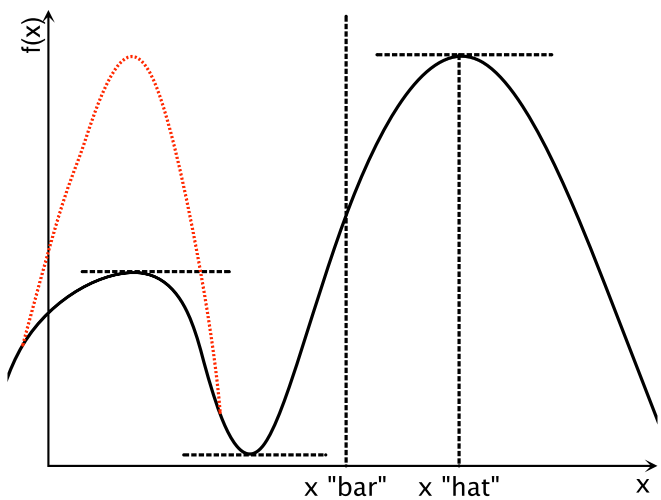

Figure 7: A one-dimensional maximization problem

If we consider the constrained problem we know that f is increasing at $\bar{x}$, so we expect the the solution to be $x^* = \bar{x}$, a "corner solution." That is, although $\bar{x}$ is the solution to the constrained maximization problem, the FOC is not satisfied at this point: $f(\bar{x}) \neq 0$. We can write this problem as the associated Lagrangian:

$$\mathcal{L}(x, \lambda) = f(x) - \lambda(x - \bar{x}).$$

Again, we can interpret λ as a penalty. If x exceeds $\bar{x}$ by one unit the objective function $\mathcal{L}$ decreases by λ. As before, we maximize the Lagrangian with respect to x and minimize with respect to λ:

$$\max_x \min_\lambda \ \mathcal{L}(x, \lambda) = f(x) - \lambda(x - \bar{x}).$$

The first-order conditions are

$$\frac{\partial \mathcal{L}}{\partial x} = f'(x) - \lambda = 0 \tag{19}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = x - \bar{x} = 0 \text{ or } \lambda = 0. \tag{20}$$

The two expressions in equation (20) can also be combined and written as

$$\lambda\,(x - \bar{x}) = 0,$$

which is known as the *complementary slackness condition.* In other words, either $x = \bar{x}$, in which case $\lambda > 0$ is necessary as a penalty, or $\lambda = 0$, which means that the constraint is not binding so that there is no need to penalize for exceeding the constraint (for example if we replace part of the solid black function in Figure 7 by the dashed red fragment). Hence, we have to distinguish two cases after determining x^* and λ^*:

$$x^* < \bar{x} \Rightarrow \lambda^* = 0$$
$$x^* = \bar{x} \Rightarrow \lambda^* = f'(x^*).$$

In both cases, equations (19) and (20) are satisfied.

How do we know that we won't have $x^* > \bar{x}$? Under the original formulation individuals were simply not allowed to choose a bundle that violates the constraint, but under the Lagrangian approach the maximization is unconstrained. The λ term introduces a "price," that ensures that individuals do not choose to consume more then is feasible. If more was consumed than available then this price would have been set too low. So rather than setting unbreakable constraints, the economics approach is to allow for "prices" that make individuals not want to choose impossible amounts given the resource constraints in the economy. This allows for "decentralizing via prices."

The Lagrange multiplier λ is also called a shadow price or marginal value of the constraint. The latter term becomes clear when we consider the marginal value of increasing $\bar{x}$, i.e., of relaxing the constraint. Suppose we have solved the constraint maximization problem

$$\max_{x \leq \bar{x}} f(x)$$

and derived the optimal values $x^*(\bar{x})$ and $\lambda^*(\bar{x})$. Then the value of the Lagrangian is

$$\mathcal{L}\left(x^*(\bar{x}), \lambda^*(\bar{x})\right) = f\left(x^*(\bar{x})\right) - \lambda^*(\bar{x})\left[x^*(\bar{x}) - \bar{x}\right].$$

Now we can calculate the marginal value of relaxing (i.e., increasing) $\bar{x}$ by taking the derivative

$$\frac{\partial \mathcal{L}\left(x^*(\bar{x}), \lambda^*(\bar{x})\right)}{\partial \bar{x}} = f'\left(x^*(\bar{x})\right) x^{*\prime}(\bar{x}) - \lambda^*(\bar{x})\, x^{*\prime}(\bar{x}) + \lambda^*(\bar{x}),$$

which can be rewritten as

$$\left[f'\left(x^*(\bar{x})\right) - \lambda^*(\bar{x})\right] x^{*\prime}(\bar{x}) + \lambda^*(\bar{x}).$$

But from equation (19) we know that the term in square brackets is equal to zero. Hence

$$\frac{\partial \mathcal{L}\left(x^*(\bar{x}), \lambda^*(\bar{x})\right)}{\partial \bar{x}} = \lambda^*(\bar{x}).$$

In other words, the marginal value of relaxing the constraint is equal to the shadow price λ.

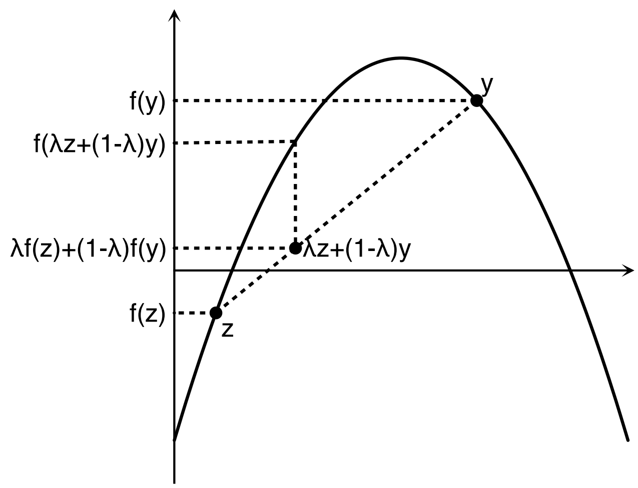

Figure 8: A concave function

5.6 Concavity and Quasi-Concavity

A function $f : \mathbb{R}^N \to \mathbb{R}$ is called concave if for all $x, y \in \mathbb{R}^N$ and all $\lambda \in [0, 1]$, we have

$$f(\lambda x + (1-\lambda)y) \geq \lambda f(x) + (1-\lambda) f(y).$$

See Figure 8 for an example. Concave functions have a useful characteristic for optimization problems. All of their extrema (i.e., critical points x such that $f'(x) = 0$) are global maxima. Hence, first order conditions are both necessary and sufficient for finding the global maxima of concave functions in unconstrained optimization problems.

A function $f : \mathbb{R}^N \to \mathbb{R}$ is called quasi-concave if for all $x, y \in \mathbb{R}^N$ and all $\lambda \in [0, 1]$:

$$f(\lambda x + (1-\lambda)y) \geq \min\{f(x), f(y)\}.$$

See Figure 9 for an example. A function $f : \mathbb{R}^N \to \mathbb{R}$ is called strictly quasi-concave if for all $x \neq y \in \mathbb{R}^N$ and all $\lambda \in (0, 1)$:

$$f(\lambda x + (1-\lambda)y) > \min\{f(x), f(y)\}.$$

The function in Figure 9 is also strictly quasi-concave. Whereas quasi-concave functions can have many global maxima, strictly quasi-concave functions have at most one global maximum. An alternative definition of a strictly quasi-concave function is that the upper contour set of f is convex, .i.e. for all $z \in \mathbb{R}$

$$UC(z) = \left\{x \in \mathbb{R}^N \,|\, f(x) > z\right\}. \tag{21}$$

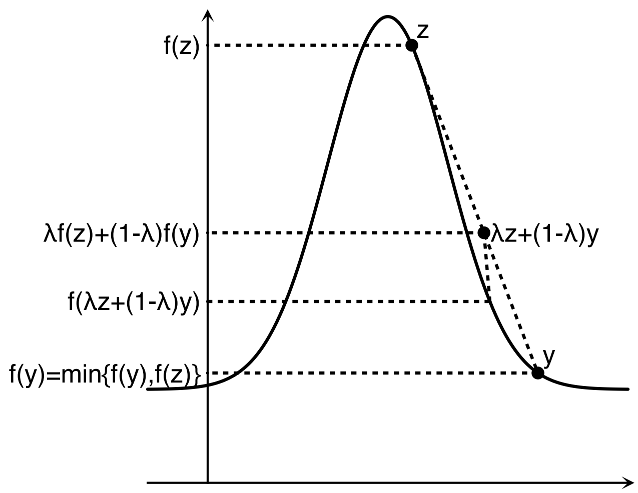

Figure 9: A (strictly) quasi-concave function

Quasi-concave utility functions are widely used because definition (21) implies that a consumer with a quasi-concave utility function has convex preferences. In other words, the set of consumption bundles preferred to a point y is convex (see Figure 3 for an illustration). Another implication of quasi-concave utility functions is the fact that the own-price effect of the compensated demand i is always negative, i.e.

$$\frac{\partial x_i^c\left(p, \overline{U}\right)}{\partial p_i} < 0.$$

This can be seen from the negative slope of indifference curves when preferences are convex. This is not true for the own-price effect of uncompensated demand however.

6 Competitive Equilibrium

So far we have been concerned with a single agent decision problem. Using the tools developed, we can now consider the more interesting situation where many agents are interacting. One way of modelling the interaction between many agents is Competitive Equilibrium, or General Equilibrium. In a competitive equilibrium framework the agents are "endowed" with some vector of commodities, and then proceed to trade at some prices. These prices must be such that the amount offered for trade (the sum of the agents' endowments) equals the amount demanded. This trading results in a new allocation, where all mutually beneficial trades have been exhausted.

We now consider the example of an economy with two agents and two goods. The framework is far more general and can be extended without complications to any arbitrary number of agents and goods, but we consider only this case to simplify notation. We assume that there are two people in the economy, Robinson and Friday. Since they live on an island they can only trade with each other. The budget set is simply $p_1x_1 + p_2x_2 \leq M$ as usual. However, the amount of money an individual has is simply the value of their endowment. If the initial endowment of good 1 and 2 respectively is $e = (e_1, e_2)$ then the budget set is simply

$$p_1x_1 + p_2x_2 \leq p_1e_1 + p_2e_2. \tag{22}$$

Suppose there are two goods, bananas and coconuts, then we can write the endowments for Robinson and Friday as $e^R = (e^R_B, e^R_C)$ and $e^F = (e^F_B, e^F_C)$ as the endowments of bananas and coconuts respectively.

Each agent maximizes his utility subject to the budget constraint (22). The problem for Robinson is to solve

$$\max_{x^R_B, x^R_C} u^R\left(x^R_B, x^R_C\right) \text{ s.t. } p_Bx^R_B + p_Cx^R_C \leq p_Be^R_B + p_Ce^R_C \tag{23}$$

and similarly Friday's problem is to solve

$$\max_{x^F_B, x^F_C} u^F\left(x^F_B, x^F_C\right) \text{ s.t. } p_Bx^F_B + p_Cx^F_C \leq p_Be^F_B + p_Ce^F_C \tag{24}$$

We can then define a competitive equilibrium. This definition can, of course, be extended to more goods, and more agents.

Definition 6.1. A Competitive Equilibrium is a set of prices $p = (p_B, p_C)$ and a set of consumption choices $x^R = \left(x^R_B, x^R_C\right)$ and $x^F = \left(x^F_B, x^F_C\right)$ such that at these prices x^R solves the utility maximization problem in (23) and x^F solves (24), and markets clear so that

$$\begin{aligned} x^R_B + x^F_B &= e^R_B + e^F_B \\ x^R_C + x^F_C &= e^R_C + e^F_C \end{aligned}$$

So competitive equilibrium consists of an individual problem and a social problem: agent optimization and market clearing. Basically, this definition means that in a competitive equilibrium all agents take the prices as given, and choose the optimal bundle at these prices. The prices must be set so the aggregate demand for each good, (the sum of how much each agent demands) is equal to the aggregate supply (the sum of all agents' endowments). We can denote the total endowment of each good by $e_B = e_B^R + e_B^F$ and $e_C = e_C^R + e_C^F$.

Here we state two important theorems that we will consider in more detail later on.

Theorem 6.2. *(First Welfare Theorem) Every competitive equilibrium is Pareto efficient.*

Theorem 6.3. *(Second Welfare Theorem) Every Pareto efficient allocation can be decentralized as a competitive equilibrium. That is, every Pareto efficient allocation is the equilibrium for some endowments.*

We now consider a simple example, where Friday is endowed with the only (perfectly divisible) banana and Robinson is endowed with the only coconut. That is $e^F = (1, 0)$ and $e^R = (0, 1)$. To keep things simple suppose that both agents have the same utility function

$$u(x_B, x_C) = \alpha\sqrt{x_B} + \sqrt{x_C}$$

and we consider the case where $\alpha > 1$, so there is a preference for bananas over coconuts that both agents share. We can determine the indifference curves for both Robinson and Friday that correspond to the same utility level that the initial endowments provide. The indifference curves are given by

$$u^F\left(e^F_{B,}x^F_C\right) = \alpha\sqrt{e^F_B} + \sqrt{e^F_C} = \alpha = u^F(1, 0)$$
$$u^R\left(e^R_{B,}e^R_C\right) = \alpha\sqrt{e^R_B} + \sqrt{e^R_C} = 1 = u^R(0, 1)$$

All the allocations between these two indifference curves are Pareto superior to the initial endowment.

This is depicted in the Edgeworth box in Figure 10. Note that we have a total of four axes in the Edgeworth box. The origin for Friday is in the south-west corner and the amount of bananas he consumes is measured along the lower horizontal axis whereas his amount of coconuts is measured along the left vertical axis. For Robinson, the origin is in the north-east corner, the upper horizontal axis depicts Robinson's

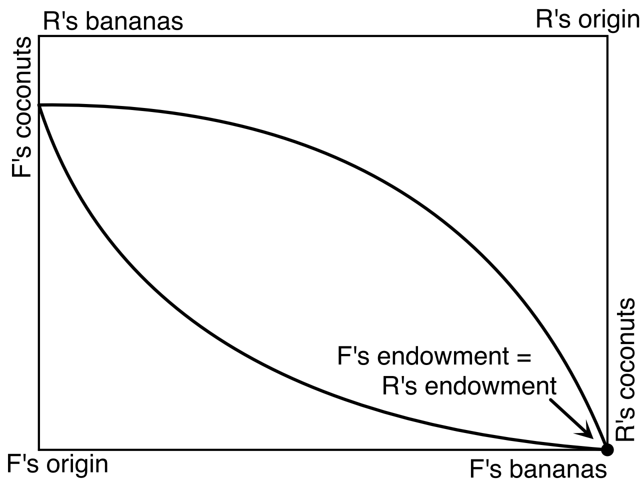

Figure 10: endowment in the Edgeworth box

banana consumption, and the right vertical axis measures his coconut consumption. The height and width of the Edgeworth box are one each since there are one banana and one coconut in this economy. Hence, the endowment bundle is the south-east corner where the amount of Friday's bananas and Robinson's coconuts are both equal to one. This also implies that Friday's utility increases as he moves up and right in the Edgeworth box, whereas Robinson is better off the further down and left he gets.

We can define the net trade for Friday (and similarly for Robinson) by

$$z_B^F = x_B^F - e_B^F$$
$$z_C^F = x_C^F - e_C^F$$

Notice that since initially Friday had all the bananas and none of the coconuts

$$z_B^F \leq 0$$
$$z_C^F \geq 0$$

The First Welfare Theorem tells us that in a competitive equilibrium trade will take place until we have found a Pareto efficient allocation, where all mutually beneficial trades have been exhausted. At this point the indifference curves of the two agents must be tangent (see Figure 11 again).

There could be many Pareto efficient allocations (e.g., Friday gets everything, Robinson gets everything, etc.), but we can calculate which

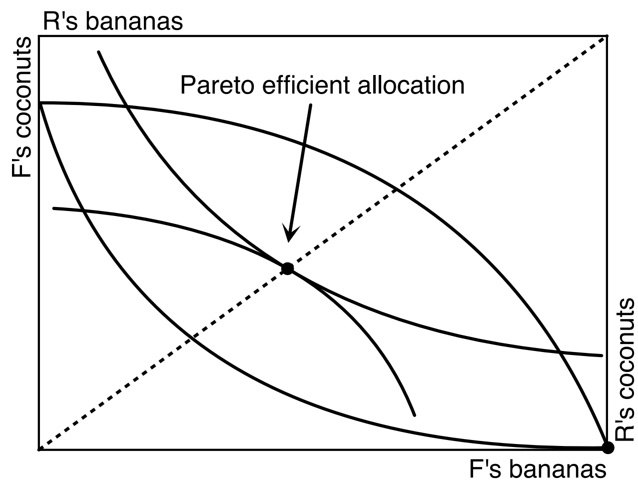

Figure 11: tangent indifference curves in the Edgeworth box

allocations are Pareto optimal. If the indifference curves at an allocation are tangent then the marginal rates of substitution must be equated. In this case, the resulting condition is

$$\frac{\frac{\partial u^F}{\partial x_B^F}}{\frac{\partial u^F}{\partial x_C^F}} = \frac{\frac{\alpha}{2\sqrt{x_B^F}}}{\frac{1}{2\sqrt{x_C^F}}} = \frac{\frac{\alpha}{2\sqrt{x_B^R}}}{\frac{1}{2\sqrt{x_C^R}}} = \frac{\frac{\partial u^R}{\partial x_B^R}}{\frac{\partial u^R}{\partial x_C^R}}$$

which simplifies to

$$\frac{\sqrt{x_C^F}}{\sqrt{x_B^F}} = \frac{\sqrt{x_C^R}}{\sqrt{x_B^R}}$$

and, of course, since there is a total of one unit of each commodity, for market clearing we must have

$$\begin{aligned} x_C^R &= 1 - x_C^F \\ x_B^R &= 1 - x_B^F \end{aligned}$$

so

$$\frac{\sqrt{x_C^F}}{\sqrt{x_B^F}} = \frac{\sqrt{1 - x_C^F}}{\sqrt{1 - x_B^F}}$$

and squaring both sides

$$\frac{x_C^F}{x_B^F} = \frac{1 - x_C^F}{1 - x_B^F}$$

which implies that

$$x_C^F - x_C^F x_B^F = x_B^F - x_C^F x^B$$

and so

$$\begin{aligned}x_C^F &= x_B^F \\ x_C^R &= x_B^R.\end{aligned}$$

So in this example, the Pareto efficient allocations are precisely the 45 degree line in the Edgeworth box (see Figure 11).

What are the conditions necessary for an equilibrium? First we need the conditions for Friday to be optimizing. We can write Robinson's and Friday's optimization problems as the corresponding Lagrangian, where we generalize the endowments to any $e^R = (e_B^R, e_C^R)$ and $e^F = (e_B^F, e_C^F)$:

$$\mathcal{L}\left(x_B^F, x_C^F, \lambda^F\right) = \alpha\sqrt{x_B^F} + \sqrt{x_C^F} + \lambda\left(p_B e_B^F + e_C^F - p_B x_B^F - x_C^F\right), \quad (25)$$

where we normalize $p_C = 1$ without loss of generality. A similar Lagrangian can be set up for Robinson's optimization problem. The first-order conditions for (25) are

$$\frac{\partial \mathcal{L}}{\partial x_B^F} = \frac{\alpha}{2\sqrt{x_B^F}} - \lambda^F p_B = 0 \quad (26)$$

$$\frac{\partial \mathcal{L}}{\partial x_C^F} = \frac{1}{2\sqrt{x_C^F}} - \lambda^F = 0 \quad (27)$$

$$\frac{\partial \mathcal{L}}{\partial \lambda^F} = p_B e_B^F + e_C^F - p_B x_B^F - x_C^F = 0. \quad (28)$$

Solving as usual by taking the ratio of equations (26) and (27) we get the following expression for the relative (to coconuts) price of bananas

$$p_B = \alpha\frac{\sqrt{x_C^F}}{\sqrt{x_B^F}}$$

so that we can solve for x_C^F as a function of x_B^F

$$x_C^F = \left(\frac{p_B}{\alpha}\right)^2 x_B^F.$$

Plugging this into the budget constraint from equation (28) we get

$$p_B x_B^F + \left(\frac{p_B}{\alpha}\right)^2 x_B^F = p_B e_B^F + e_C^F.$$

Then we can solve for Friday's demand for bananas

$$x_B^F = \frac{p_B e_B^F + e_C^F}{p_B + \left(\frac{p_B}{\alpha}\right)^2}$$

and for coconuts

$$x_C^F = \left(\frac{p_B}{\alpha}\right)^2 \frac{p_B e_B^F + e_C^F}{p_B + \left(\frac{p_B}{\alpha}\right)^2}.$$

The same applies to Robinson's demand functions, of course.

Now we have to solve for the equilibrium price p_B. To do that we use the market clearing condition for bananas, which says that demand has to equal supply (endowment):

$$x_B^F + x_B^R = e_B^F + e_B^R.$$

Inserting the demand functions yields

$$\frac{p_B e_B^F + e_C^F}{p_B + \left(\frac{p_B}{\alpha}\right)^2} + \frac{p_B e_B^R + e_C^R}{p_B + \left(\frac{p_B}{\alpha}\right)^2} = e_B^F + e_B^R = e_B,$$

where e_B is the social endowment of bananas and we define $e_C = e_C^F + e_C^R$. We solve this equation to get the equilibrium price of bananas in the economy:

$$p_B^* = \alpha \sqrt{\frac{e_C}{e_B}}.$$

So we have solved for the equilibrium price in terms of the primitives of the economy. This price makes sense intuitively. It reflects relative scarcity in the economy (when there are relatively more bananas than coconuts, bananas are cheaper) and preferences (when consumers value bananas more, i.e., when α is larger, they cost more). We can then plug this price back into the previously found equations both for agents' consumption and have an expression for consumption in terms of the primitives.

We will conclude this section on competitive equilibrium by reiterating and proving the most famous result in economics, the First Welfare Theorem.

Theorem 6.4. *(First Welfare Theorem) Every Competitive Equilibrium allocation x^* is Pareto Efficient.*

Proof. Suppose not. Then there exists another allocation y, which is feasible, such that

$$\text{for all } n\colon u^n(y) \geq u^n(x^*)$$
$$\text{for some } n\colon u^n(y) > u^n(x^*).$$

If $u^n(y) \geq u^n(x^*)$, then the budget constraint (and monotone preferences) implies that

$$\sum_{i=1}^{I} p_i y_i^n \geq \sum_{i=1}^{I} p_i x_i^{*n} \tag{29}$$

and for some n

$$\sum_{i=1}^{I} p_i y_i^n > \sum_{i=1}^{I} p_i x_i^{*n}. \tag{30}$$

Equations (29) and (30) imply that

$$\sum_{n=1}^{N} \sum_{i=1}^{I} p_i y_i^n > \sum_{n=1}^{N} \sum_{i=1}^{I} p_i x_i^{*n} = \sum_{i=1}^{I} p_i e_i,$$

where the left-most term is the aggregate expenditure and the right-most term the social endowment. This is a contradiction because feasibility of y means that

$$\sum_{n=1}^{N} y_i^n \leq \sum_{n=1}^{N} e_i^n = e_i$$

for any n and hence

$$\sum_{n=1}^{N} \sum_{i=1}^{I} p_i y_i^n \leq \sum_{i=1}^{I} p_i e_i.$$

$\square$

7 Producer Theory

We can use tools similar to those we used in the consumer theory section of the class to study firm behaviour. In that section we assumed that individuals maximize utility subject to some budget constraint. In this section we assume that firms will attempt to maximize their profits given a demand schedule and production technology.

Firms use inputs or commodities $x_1, \ldots, x_I$ to produce an output y. The amount of output produced is related to the inputs by the production function $y = f(x_1, \ldots, x_I)$, which is formally defined as follows:

Definition 7.1. A *production function* is a mapping $f : \mathbb{R}_+^I \longrightarrow \mathbb{R}_+$.

The prices of the inputs/commodities are $p_1, \ldots, p_I$ and the output price is p_y. The firm takes prices as given and independent of its decisions.

Firms maximize their profits by choosing the optimal amount and combination of inputs.

$$\max_{x_1,\ldots,x_I} p_y f(x_1, \ldots, x_I) - \sum_{i=1}^{I} p_i x_i. \tag{31}$$

Another way to describe firms' decision making is by minimizing the cost necessary to produce an output quantity $\bar{y}$.

$$\min_{x_1,\ldots,x_I} \sum_{i=1}^{I} p_i x_i \text{ s.t. } f(x_1, \ldots, x_I) \geq \bar{y}.$$

The minimized cost of production, $C(\bar{y})$, is called the cost function.

We make the following assumption for the production function: positive marginal product

$$\frac{\partial f}{\partial x_i} \geq 0$$

and declining marginal product

$$\frac{\partial^2 f}{\partial x_i^2} \leq 0.$$

The optimality conditions for the profit maximization problem (31) and the FOCs for all i

$$p_y \frac{\partial f}{\partial x_i} - p_i = 0.$$

In other words, optimal production requires equality between marginal benefits and marginal cost of production. The solution to the profit maximization problem then is

$$
\begin{gathered}
x_i^*(p_1, \ldots, p_I, p_y), i = 1, \ldots, I \\
y^*(p_1, \ldots, p_I, p_y),
\end{gathered}
$$

i.e., optimal demand for inputs and optimal output/supply.

The solution of the cost minimization problem (), on the other hand is

$$x_i^*\left(p_1, \ldots, p_I, \bar{y}\right), i = 1, \ldots, I,$$

where $\bar{y}$ is the firm's production target.

Example 7.2. One commonly used production function is the Cobb-Douglas production function where

$$f(K, L) = K^\alpha L^{1-\alpha}$$

The interpretation is the same as before with α reflecting the relative importance of capital in production. The marginal product of capital is $\frac{\partial f}{\partial K}$ and the marginal product of labor is $\frac{\partial f}{\partial L}$.

In general, we can change the scale of a firm by multiplying both inputs by a common factor: $f(tK, tL)$ and compare the new output to $tf(K, L)$. The firm is said to have *constant returns to scale* if

$$tf(K, L) = f(tK, tL),$$

it has *decreasing returns to scale* if

$$tf(K, L) > f(tK, tL),$$

and *increasing returns to scale* if

$$tf(K, L) < f(tK, tL).$$

Example 7.3. The Cobb-Douglas function in our example has constant returns to scale since

$$f(tK, tL) = (tK)^\alpha (tL)^{1-\alpha} = tK^\alpha L^{1-\alpha} = tf(K, L).$$

Returns to scale have an impact on market structure. With decreasing returns to scale we expect to find many small firms. With increasing returns to scale, on the other hand, there will be few (or only a single) large firms. No clear prediction can be made in the case of constant returns to scale. Since increasing returns to scale limit the number of firms in the market, the assumption that firms are price takers only makes sense with decreasing or constant returns to scale.

8 Decisions under Uncertainty

So far, we have assumed that decision makers have all the needed information. This is not the case in real life. In many situations, individuals or firms make decisions before knowing what the consequences will be. For example, in financial markets investors buy stocks without knowing future returns. Insurance contracts exist because there is uncertainty. If individuals were not uncertain about the possibility of having an accident in the future, there would be no need for car insurance.

8.1 Lotteries

To conceptualize uncertainty, we define the notion of a *lottery* as follows:

$$L = \{c_1, \ldots, c_N; p_1, \ldots, p_N\},$$

where $c_n \in \mathbb{R}$ is a money award (positive or negative) that will be paid out if the state of the world is n and $p_n \in [0, 1]$ is the probability that state n occurs. For example, a state of the world could be the occurrence of an accident. Then p_n measures the probability that an accident occurs and c_n is the loss associated with the accident.

Lotteries are nothing else than random variables, which means that we can evaluate them using tools from statistics. The *expected value* of a lottery (its mean) is defined as

$$E[L] = \sum_{n=1}^{N} p_n c_n.$$

For a valid random variables we need the probabilities to be non-negative and to sum up to one: $p_n \geq 0$ and $\sum_{n=1}^{N} p_n = 1$. Moreover, the $c_1, \ldots, c_N$ have to be mutually exclusive. A safe lottery or deterministic outcome can be represented by a degenerate probability distribution, for example $(p_1, \ldots, p_N) = (1, 0, \ldots, 0)$. In this case, state $n = 1$ occurs with certainty. Another way to represent a safe outcome is by having $c_1 = \cdots = c_N$; i.e., all awards are the same. The risk of a lottery can be represented by its *variance*, which is defined as

$$\operatorname{var}(L) = \sum_{n=1}^{N} (c_n - E[L])^2 \, p_n.$$

Example 8.1. An famous lottery is the following one, also known as the *St. Petersburg paradox*. A fair coin is tossed until head comes up for the first time. Then the reward paid out is equal to 2^n, where n is the number of coin tosses that were necessary for head to come up once. This lottery is described formally as

$$L_{SP} = \left\{2, 4, 8, \ldots, 2^n, \ldots; \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \ldots, \frac{1}{2^n}, \ldots\right\}.$$

Its expected value is

$$E\left[L_{SP}\right] = \sum_{n=1}^{\infty} p_n c_n = \sum_{n=1}^{\infty} \frac{1}{2^n} 2^n = \sum_{n=1}^{\infty} 1 = \infty.$$

Hence the expected payoff from this lottery is infinitely large and an individual offered this lottery should be willing to pay an infinitely large amount for the right to play this lottery. This is not what people do, however, hence the paradox.

8.2 Expected Utility

Agents usually do not just consider the mean of lotteries, but also their risk. This leads to concave utility functions.

Definition 8.2. A utility function $u : \mathbb{R}_+ \longrightarrow \mathbb{R}_+$ is *concave* if for x, y and all $\lambda \in [0, 1]$

$$\lambda u(x) + (1-\lambda) u(y) \leq u\left(\lambda x + (1-\lambda) y\right),$$

where the LHS is the expected values of utility and the RHS is utility from the expected payoff.

In other words, an agent with a concave utility function prefers the expected value of the lottery to the lottery itself. Figure 12 illustrates this. We call agents with concave utility function *risk avers.*

Definition 8.3. The *expected utility* of a decision maker with utility function $u : \mathbb{R}_+ \longrightarrow \mathbb{R}_+$ for a lottery $L = \{c_1, \ldots, c_N; p_1, \ldots, p_N\}$ is defined by

$$U(L) = \sum_{n=1}^{N} p_n u\left(c_n\right).$$

Note the difference between the upper-case U in $U(L)$ and the lower-case u in $u\left(c_n\right)$; they are two different functions.

Definition 8.4. A decision maker is called *risk averse* if the utility function $u : \mathbb{R}_+ \longrightarrow \mathbb{R}_+$is concave and she is called *risk loving* or a risk seeker if u is convex.

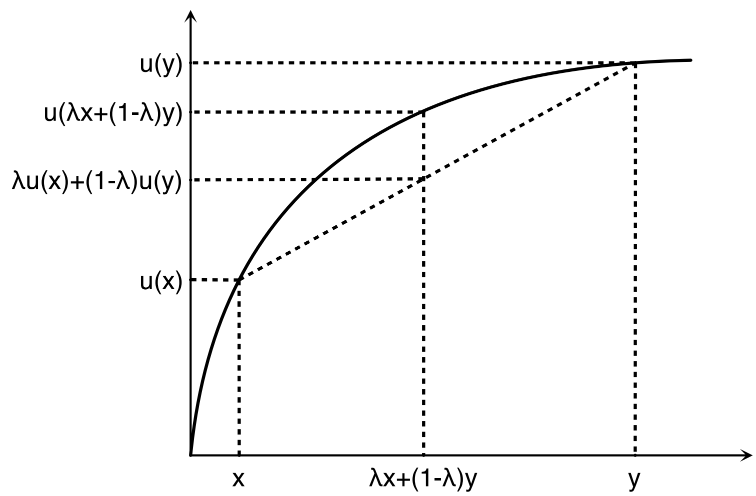

Figure 12: A concave utility function

8.3 Insurance

An important application for uncertainty is insurance contracts. Consider a world with two states, one where an accident occurs and one without an accident. In both states, the agent's wealth is w. An accident occurs with probability $\pi \in (0, 1)$ and the damage in case of an accident is $D > 0$. The agent can buy insurance of \$1 at the price of q. The amount of insurance bought is α. Then we can summarize the payoffs as follows:

state of the world	payoff
no accident	$w - \alpha q$
accident	$w - D - \alpha q + q$

The goal is to determine the optimal level of insurance, α^*. To do that, we maximize the agent's expected utility over α:

$$\max_{\alpha} \left\{ \pi u\left(w - D - \alpha q + q\right) + (1 - \pi) u\left(w - \alpha q\right) \right\}.$$

The FOC is

$$\pi u'\left(w - D - \alpha q + q\right)(1 - q) - (1 - \pi) u'\left(w - \alpha q\right) q = 0. \tag{32}$$

In words, the marginal utility in the bad state (benefit of insurance) has to equal the marginal utility in the good state (cost of insurance).

Now, we assume that the insurance company sells "fair insurance," which means that it sets an actuarially fair price so that it has an expected payoff of zero.[6] Hence the expected payment if an accident occurs,

[6] This is obviously not a very realistic assumption.

π, has to equal the premium q. The FOC (32) at $q = \pi$ becomes

$$\begin{gathered}\pi u'\left(w - D - \alpha q + q\right)(1-\pi) - (1-\pi)u'\left(w - \alpha q\right)\pi = 0 \Longleftrightarrow \\ u'\left(w - D - \alpha q + q\right) = u'\left(w - \alpha q\right),\end{gathered}$$

which implies the solution $\alpha^* = D$, assuming a strictly concave utility function with $u''(w) < 0$. Hence, in order to equalize marginal utility in both states, the wealth in both states is the same. This is also known as full insurance.

9 Pricing Power

So far, we have considered market environments where single agent cannot control prices. Instead, each agent was infinitesimally small and firms acted as price takers. This was the case in competitive equilibrium. There are many markets with few (oligopoly) or a single firms, (monopoly) however. In that case firms can control prices to some extent. Moreover, when there are a few firms in a market, firms make interactive decisions. In other words, they take their competitors' actions into account. In Section 10, we will use game theory to analyse this type of market structure. First, we cover monopolies, i.e., markets with a single producer.

9.1 Monopoly

If a firm produces a non-negligible amount of the overall market then the price at which the good sells will depend on the quantity sold. Examples for firms that control the overall market include the East India Trading Company, Microsoft (software in general because of network externalities and increasing returns to scale), telecommunications and utilities (natural monopolies), Standard Oil, and De Beers.

For any given price there will be some quantity demanded by consumers, and this is known as the demand curve $x : \mathbb{R}_+ \longrightarrow \mathbb{R}_+$ or simply $x(p)$. We assume that consumers demand less as the price increases: the demand function is downward sloping or $x'(p) < 0$. We can invert this relationship to get the inverse demand function $p(x)$ which reveals the price that will prevail in the market if the output is x.

If the firm is a monopolist that takes the demand data $p(x)$ as given then its goal is to maximize

$$\pi(x) = p(x)x - c(x) \tag{33}$$

by choosing the optimal production level. For the cost function we assume $c'(x) > 0$ and $c''(x) \geq 0$, i.e., we have positive and weakly increasing marginal costs. For example, $c(x) = cx$ satisfies these assumptions (a Cobb-Douglas production function provides this for example). The monopolist maximizes its profit function (33) over x, which leads to the following FOC:

$$p(x) + xp'(x) - c'(x) = 0. \tag{34}$$

Here, in addition to the familiar $p(x)$, which is the marginal return from the marginal consumer, the monopolist also has to take the $xp'(x)$ into account, because a change in quantity also affects the inframarginal consumers. For example, when it increases the quantity supplies, the monopolist gets positive revenue from the marginal consumer, but the inframarginal consumers pay less due to the downward sloping demand function. At the optimum, the monopolist equates marginal revenue and marginal cost.

Example 9.1. A simple example used frequently is $p(q) = a - bq$, and we will also assume that $a > c$ since otherwise the cost of producing is higher than any consumer's valuation so it will never be profitable for the firm to produce and the market will cease to exist. Then the firm want to maximize the objective

$$\pi(x) = (a - bx - c)x.$$

The efficient quantity is produced when $p(x) = a - bx = c$ because then a consumer buys an object if and only if they value it more than the cost of producing, resulting in the highest possible total surplus. So the efficient quantity is

$$x^* = \frac{a-c}{b}.$$

The monopolist's maximization problem, however, has FOC

$$a - 2bx - c = 0$$

where $a - 2bx$ is the marginal revenue and c is the marginal cost. So the quantity set by the monopolist is

$$x^M = \frac{a-c}{2b} < x^*.$$

The price with a monopoly can easily be found since

$$\begin{aligned} p^M &= a - bx^M \\ &= a - \frac{a-c}{2} \\ &= \frac{a+c}{2} \\ &> c. \end{aligned}$$

Figure 13 illustrates this.

A monopoly has different welfare implications than perfect competition. In Figure 13, consumers in a monopoly lose the areas A and B compared to perfect competition. The monopolist loses area C and wins area A. Hence, there are distributional implications (consumers lose and the producer gains) as well as efficiency implications (overall welfare decreases).

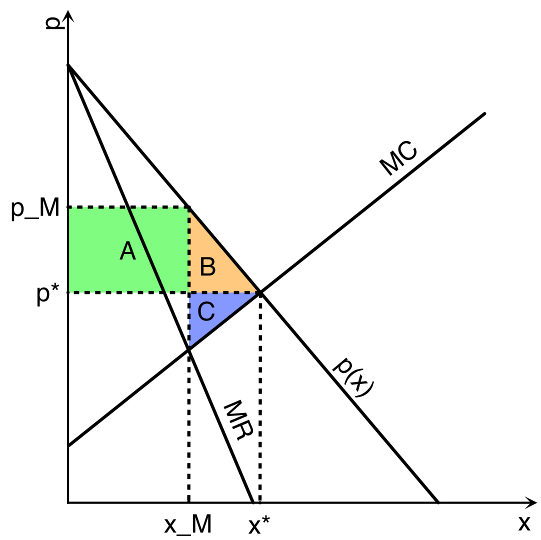

Figure 13: Monopoly

We can write the monopolist's FOC (34) in terms of the demand elasticity introduced in Section 4.2 as follows:

$$\begin{aligned} p(x^*) + x^* p'(x^*) &= c'(x^*) \iff \\ p(x^*)\left[1 + \frac{x^* p'(x^*)}{p(x^*)}\right] &= c'(x^*) \iff \\ p(x^*) &= \frac{c'(x^*)}{1 + \epsilon_p^{-1}}. \end{aligned}$$

Since $\epsilon_p < 0$, we have that $p(x^*) > c'(x^*)$, in other words, the monopolist charges more than the marginal cost. This also means that if demand is very elastic, $\epsilon_p \to \infty$, then $p(x^*) \approx c'(x^*)$. On the other hand, if demand is very inelastic, $\epsilon_p \approx -1$, then $p(x^*) \gg c'(x^*)$.

9.2 Price Discrimination

In the previous section we saw that the monopolist sets an inefficient quantity and total welfare is decreased. Is there a mechanism, which allows the monopolist to offer the efficient quantity and reap the entire possible welfare in the market? The answer is yes if the monopolist can set a two-part tariff, for example. In general, the monopolist can extract consumer rents by using price discrimination.

First degree price discrimination (perfect price discrimination) means discrimination by the identity of the person or the quantity ordered (non-linear pricing). It will result in an efficient allocation. Suppose there is a single buyer and a monopoly seller where the inverse demand is given by

$p = a - bx$. If the monopolist were to set a single price it would set the monopoly price. As we saw in the previous section, however, this does not maximize the joint surplus, so the monopolist can do better. Suppose instead that the monopolist charges a fixed fee F that the consumer has to pay to be allowed to buy any positive amount at all, and then sells the good at a price p, and suppose the monopolist sets the price $p = c$. The fixed fee will not affect the quantity that a participating consumer will choose, so if the consumer participates then they will choose quantity equal to x^*. The firm can then set the entry fee to extract all the consumer surplus and the consumer will still be willing to participate. This maximizes the joint surplus, and gives the entire surplus to the firm, so the firm is doing as well as it could under any other mechanism. Specifically, using the functional form from Example 9.1 the firm sets

$$F = \frac{(a-c)x^*}{2} = \frac{(a-c)^2}{2b}$$

In integral notation this is

$$F = \int_0^{x^*} (p(x) - c)dx.$$

This pricing mechanism is called a two-part tariff, and was famously used at Disneyland (entry fee followed by a fee per ride), greatly increasing revenues.

Now, let's assume that there are two different classes of consumers, type A with utility function $u(x)$ and type B with $\beta u(x), \beta > 1$, so that the second class of consumers has a higher valuation of the good. If the monopolist structures a two-part tariff $(F, p = c)$ to extract all surplus from type B consumers, type A consumers would not pay the fixed fee F since they could not recover the utility lost from using the service. On the other hand, if the firm offers two two-part tariffs $(F_A, p = c)$ and $(F_B, p = c)$ with $F_A < F_B$, all consumers would pick the cheaper contract $(F_A, p = c)$. A solution to this problem would be to offer the contracts $(F_A, p_A > c)$ and $(F_B, p = c)$. Type A consumers pick the first contract and consume less of the good and type B consumers pick the second contract, which allows them to consume the efficient quantity. This is an example for *second degree price discrimination*, which means that the firm varies the price by quantity or quality only. It offers a menu of choices and lets the consumers self-select into their preferred contract.

In addition, there is *third degree price discrimination*, in which the firm varies the price by market or identity of the consumers. For example, Disneyland can charge different prices in different parks. Let's assume there are two markets, $i = 1, 2$. The firm is a monopolist in both markets and its profit maximization problem is

$$\max_{x_1, x_2} x_1 p_1(x_1) + x_2 p_2(x_2) - c(x_1 + x_2).$$

The FOC for each market is

$$p_i(x_i) + x_i p_i'(x_i) = c'(x_1 + x_2),$$

which leads to optimal solution

$$p_i(x_i^*) = \frac{c'(x_1^* + x_2^*)}{1 + \frac{1}{\epsilon_p^i}} \text{ for } i = 1, 2.$$

Hence, the solution depends on the demand elasticity in market i. The price will be different as long as the structure of demand differs.

10 Oligopoly

Oligopoly refers to environments where there are few large firms. These firms are large enough that their quantity influences the price and so impacts their rivals. Consequently each firm must condition its behaviour on the behaviour of the other firms. This strategic interaction is modelled with game theory.

10.1 Duopoly

Probably the most important model of oligopoly is the Cournot model of quantity competition. We will first consider the duopoly case, where there are only two firms. Suppose the inverse demand function is given by $p(q) = a - bq$, and the cost of producing is constant and the same for both firms $c_i(q) = cq$. The quantity produced in the market is the sum of what both firms produce $q = q_1 + q_2$. The profits for each firm is then a function of the market price and their own quantity,

$$\pi_i(q_i, q_j) = q_i \left(p\left(q_i + q_j\right) - c\right).$$

The strategic variable that the firm is choosing is the quantity to produce q_i.

Suppose that the firms' objective was to maximize their joint profit

$$\pi_1(q_1, q_2) + \pi_2(q_1, q_2) = (q_1 + q_2)\left(p\left(q_1 + q_2\right) - c\right)$$

then we know from before that this is maximized when $q_1 + q_2 = q^M$. We could refer to this as the collusive outcome. One way the two firms could split production would be $q_1 = q_2 = \frac{q^M}{2}$.

If the firms could write binding contracts then they could agree on this outcome. However, that is typically not possible (such an agreement would be price fixing), so we would not expect this outcome to occur unless it is stable/self-enforcing. If either firm could increase its profits by setting another quantity, then they would have an incentive to deviate from this outcome. We will see below that both firms would in fact have an incentive to deviate and increase their output.

Suppose now that firm i is trying to choose q_i to maximize its own profits, taking the other firm's output as given. Then firm i's optimization problem is

$$\max_{q_i} \pi_i(q_i, q_j) = q_i \left(a - b\left(q_i + q_j\right) - c\right),$$

which has the associated FOC

$$\frac{\partial \pi_i(q_i, q_j)}{\partial q_i} = a - b\left(2q_i + q_j\right) - c = 0.$$

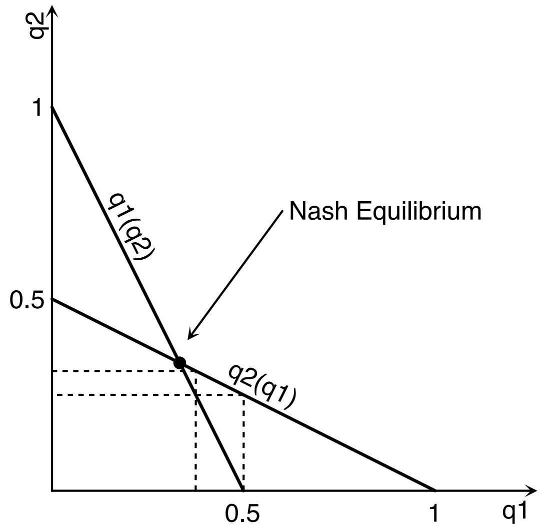

Figure 14: Cournot equilibrium

Then the optimal level q_i^* given any level of q_j is

$$q_i^*\left(q_j\right) = \frac{a - bq_j - c}{2b}.$$

This is firm i's *best response* to whatever firm j plays. In the special case when $q_j = 0$ firm i is a monopolist, and the observed quantity q_i corresponds to the monopoly case. In general, when the rival has produced q_j we can treat the firm as a monopolist facing a "residual demand curve" with intercept of $a - bq_j$. We can write firm i's best response function as

$$q_i^*(q_j) = \frac{a-c}{2b} - \frac{1}{2}q_j.$$

Hence,

$$\frac{dq_i}{dq_j} = -\frac{1}{2}.$$

This has two important implications. First, the quantity player i chooses is decreasing in its rival's quantity. This means that quantities are *strategic substitutes*. Second, if player j increases their quantity player i decreases their quantity by less than player j increased their quantity (player i decreases his quantity by exactly $\frac{1}{2}$ for every unit player j's quantity is increased). So we would expect that the output in a duopoly would be higher than in a monopoly.

We can depict the best response function graphically. Setting $a = b = 1$ and 0, Figure 14 shows the best response functions. Here, the best response functions are $q_i^*(q_j) = \frac{1-q_j}{2}$.

We are at a "stable" outcome if both firms are producing a best response to their rivals' production. We refer to such an outcome as an equilibrium. That is, when

$$\begin{aligned} q_i &= \frac{a - bq_j - c}{2b} \\ q_j &= \frac{a - bq_i - c}{2b}. \end{aligned}$$

Since the best responses are symmetric we will have $q_i = q_j$ and so we can calculate the equilibrium quantities from the equation

$$q_i = \frac{a - bq_i - c}{2b}$$

and so

$$q_i = q_j = \frac{a - c}{3b}$$

and hence

$$q = q_i + q_j = \frac{2(a - c)}{3b} > \frac{a - c}{2b} = q^M.$$

There is a higher output (and hence lower price) in a duopoly then a monopoly.

More generally, both firms are playing a best response to their rival's action because for all i

$$\pi_i(q_i^*, q_j^*) \geq \pi_i(q_i, q_j^*) \text{ for all } q_i$$

That is, the profits from the quantity are (weakly) higher then the profits from any other output. This motivates the following definition for an equilibrium in a strategic setting.

Definition 10.1. A *Nash Equilibrium in the duopoly game* is a pair (q_i^*, q_j^*) such that for all i

$$\pi_i(q_i^*, q_j^*) \geq \pi_i(q_i, q_j^*) \text{ for all } q_i.$$

This definition implicitly assumes that agents hold (correct) expectations or beliefs about the other agents' strategies.

A Nash Equilibrium is ultimately a stability property. There is no profitable deviation for any of the players. In order to be at equilibrium we must have that

$$\begin{aligned} q_i &= q_i^*(q_j) \\ q_j &= q_j^*(q_i) \end{aligned}$$

and so we must have that

$$q_i = q_i^*(q_j^*(q_i))$$

so equilibrium corresponds to a fixed-point of the mapping $q_1^*(q_2^*(\cdot))$. This idea can also be illustrated graphically. In Figure 14, firm 1 initially sets $q_1 = \frac{1}{2}$, which is not the equilibrium quantity. Firm 2 then optimally picks $q_2 = q_2^*\left(\frac{1}{2}\right) = \frac{1}{4}$ according to its best response function. Firm 1, in turn, chooses a new quantity according to its best response function: $q_1 = q_1^*\left(\frac{1}{4}\right) = \frac{3}{8}$. This process goes on and ultimately converges to $q_1 = q_2 = \frac{1}{3}$.

10.2 General Cournot

Now, we consider the case with I competitors. The inverse demand function (setting $a = b = 1$) is

$$p(q) = 1 - \sum_{i=1}^{I} q_i$$

and firm i's profit function is

$$\pi(q_i, q_{-i}) = \left(1 - \sum_{i=1}^{I} q_i - c\right) q_i, \tag{35}$$

where the vector q_{-i} is defined as $q_{-i} = (q_1, \dots, q_{i-1}, q_{i+1}, \dots, q_I)$, i.e., all quantities excluding q_i.

Again, we can define an equilibrium in this market as follows:

Definition 10.2. A *Nash Equilibrium in the oligopoly game* is a vector $q^* = (q_1^*, \dots, q_I^*)$ such that for all i

$$\pi_i(q_i^*, q_{-i}^*) \geq \pi_i(q_i, q_{-i}^*) \text{ for all } q_i.$$

We simply replaced the quantity q_j by the vector q_{-i}.

Definition 10.3. A Nash equilibrium is called *symmetric* if $q_i^* = q_j^*$ for all i and j.

The FOC for maximizing the profit function (35) is

$$1 - \sum_{j \neq i} q_j - 2q_i - c = 0$$

and the best response function for all i is

$$q_i = \frac{1 - \sum_{j \neq i} q_j - c}{2}. \tag{36}$$

Here, only the aggregate supply of firm i's competitors matters, but not the specific amount single firms supply. It would be difficult to solve for I separate values of q_i, but due to symmetry of the profit function we get that $q_i^* = q_j^*$ for all i and j so that equation (36) simplifies to

$$q_i^* = \frac{1 - (I-1)q_i^* - c}{2},$$

which leads to the solution

$$q_i^* = \frac{1-c}{I+1}.$$

As I increases (more firms), the market becomes more competitive. Market supply is equal to

$$\sum_{i=1}^{I} q_i^* = Iq_i^* = \frac{I}{I+1}(1-c).$$

As the number of firms becomes larger, $I \longrightarrow \infty$, $q_i^* \longrightarrow 0$ and

$$\sum_{i=1}^{I} q_i^* \longrightarrow 1-c,$$

which is the supply in a competitive market. As each player plays a less important strategical role in the market, the oligopoly outcome converges to the competitive market outcome.

Note that we used symmetry in deriving the market outcome from the firms' best response function. We cannot invoke symmetry when deriving the FOC. One might think that instead of writing the profit function as (35) one could simplify it to

$$\pi(q_i, q_{-i}) = (1 - Iq_i - c)\, q_i.$$

This is wrong, however, because it implies that firm i controls the entire market supply (acts as a monopolist). Instead, in an oligopoly market, firm i takes the firms' output as given.

10.3 Stackelberg Game

In the Counot game in Section 10.1 two firms choose their quantities simultaneously. In many markets, however, one firm (the leader or first mover) might make a decision before the other (the follower or second mover). The most natural interpretation is that the leader is the incumbent in the market, and the second firm is a new entrant. Then, the first mover can anticipate the second mover's action and pick its quantity taking the other firm's best response into account. This situation differs from the Cournot game because the leader can commit to a quantity before the follower makes a decision. Since firm 1 must take into account the effect of their quantity on firm 2, we use "backward induction" to solve this game. We first determine what firm 2 will produce in response to any quantity from quantity 1 in the first period. The first firm takes firm 2's response to its quantity into account when it sets its quantity in the first period.

As before the profit function for firm $i = 1, 2$ is

$$\pi_i(q_i, q_j) = [a - b(q_i + q_j) - c]\, q_i.$$

Using backward induction, we first determine firm 2's (the follower) choice given any quantity supplied by firm 1. In other words, we have to derive firm 2's best response as before. The follower's profit function is

$$\pi_2(q_1, q_2) = [a - b(q_1 + q_2) - c]\, q_2$$

and the FOC is

$$\frac{d\pi_2}{dq_2} = a - bq_1 - 2bq_2 - c = 0,$$

which leads to firm 2's best response

$$q_2^*(q_1) = \frac{a-c}{2b} - \frac{q_1}{2}.$$

Again we see that

$$q_2'(q_1) = -\frac{1}{2} < 0$$

which implies that q_1 and q_2 are strategic substitutes.

Now in the first period, the leader knows how their rival will respond to whatever quantity they produce and will take this into account when setting quantity. So firm 1 chooses quantity to maximize

$$\begin{aligned}
\hat{\pi}_1(q_1) &= \pi_1\left(q_1, q_2^*(q_1)\right) \\
&= \left[a - b\left(q_1 + q_2^*(q_1)\right) - c\right] q_1 \\
&= \left[a - b\left(q_1 + \frac{a-c}{2b} - \frac{q_1}{2}\right) - c\right] q_1 \\
&= \left(\frac{a-c}{2} - \frac{b}{2} q_1\right) q_1,
\end{aligned}$$

which has the associated FOC

$$\frac{d\hat{\pi}_1}{dq_1} = \frac{a-c}{2} - bq_1 = 0$$

and then

$$q_1^* = \frac{a-c}{2b} = q_M > q_1^C.$$

The leader chooses the same quantity as a monopolist in this market. In particular, it chooses a larger quantity than each firm in the Cournot game. The follower acts upon firm 1's chosen quantity by playing its best response:

$$q_2^* = \frac{a-c}{2b} - \frac{q_1^*}{2} = \frac{a-c}{2b} - \frac{a-c}{4b} = \frac{a-c}{4b} < q_2^C.$$

Hence we notice that firm 1 exploits its first mover advantage to produce more than its rival. Notice also that $q_1 + q_2 = \frac{3a}{4b}$ which is higher then the total output of the Cournot game. This is because firm 1, anticipating its affect on its rival, produces more than the Cournot output, and firm 2 only cuts back production by $\frac{1}{2}$ for each addition unit firm 1 produces.

The key here was that firm 1 was able to commit to the high output. If it could go back and adjust its output at the end it would undermine this commitment and remove the first mover advantage. So the restriction that the firm cannot go back and adjust its output is actually beneficial. In the single agent decision problem constraints are never beneficial, but in a strategic environment they may be. Firm 2 which is at a disadvantage could try to threaten to produce more to try to prevent firm 1 from producing so much, but such a claim would not be credible (since if firm 2 does not play a best response it decreases its own profits). Of course, we have assumed that the follower is motivated solely by the desire to maximize its profits in this period (i.e., no future periods where their behaviour this period can influence the future payoff).

10.4 Vertical Contracting

In many markets we can observe more than two kinds of participants. It is common, for example, to have a producer or wholesaler who sells goods or services to a retailer or distributor who in turn sells them to the final consumer. The wholesaler is also known as the upstream firm and the retailer is called the downstream firm. This situation can lead to double marginalization, which means that both the upstream and the downstream firms act as monopolists thereby adding a margin to their costs and decreasing the quantity sold. The following simple model illustrates this.

Suppose the final consumer's inverse demand function is

$$p = a - bq.$$

Then the retailer's profit function is

$$\pi_R(q_R) = (a - bq_R - p_W)q_R,$$

where p_W is the price set by the wholesaler (the unit cost for the retailer) and q_R is the quantity set by the retailer. The FOC is

$$a - 2bq_R - p_W = 0$$

which can be solved as

$$q_R^*(p_W) = \frac{a - p_W}{2b}. \tag{37}$$

Hence the price charged by the retailer is

$$p_R^*(p_W) = \frac{a + p_W}{2} > p_W$$

since we assume that $a > p_W$ so that at least the first marginal consumer demands the good. The "pass through rate" is $\frac{1}{2}$: for a \$1 price increase by the wholesaler, the retailer increase the final price by \$0.50. Since price increases are passed through imperfectly, the upstream firm has a larger incentive to increase p_W.

Now we calculate the optimal price p_W given the retailer's optimal supply function (37). The wholesaler's profit is

$$\pi_W(p_W) = p_W q_R^*(p_W) = p_W \frac{a - p_W}{2b}$$

assuming that the wholesaler's cost of producing the good is 0. The FOC is

$$a - 2p_W = 0$$

so that the upstream firm sets the optimal price

$$p_W^* = \frac{a}{2}.$$

Plugging this into the retailer's best response functions we get

$$p_R^*(p_W^*) = \frac{3a}{4}$$

and

$$q_R^*(p_W^*) = \frac{a}{4b} < q_M^* = \frac{a}{2b}.$$

Hence, the quantity under this market structure is even lower than in a monopoly. In general, deeper distribution makes the outcome less efficient.

One solution to alleviate this inefficiency would be vertical integration: By merging the upstream and downstream firm we could get rid of the retailer. Note that this is different from horizontal integration, which refers to mergers between firms on the same level of the distributional chain. Since vertical integration leads to social surplus due to a less inefficient quantity it is important to distinguish between these two types of integration when evaluating a potential merger.

Another solution to avoid double marginalization is a two-part tariff (F_W, r_W), where the retailer pays F_W to obtain the right to distribute the good (participation fee) and then pays r_W for every unit purchased from the wholesaler. An example for this strategy is the market for rental videos. Optimally, the the wholesaler sets $r_W^* = 0$ to induce the retailer to sell as many units as possible (q_R^* in equation (37) depends negatively on p_W) and then sets F_W^* so that the retailer just participates, but the wholesaler can capture the entire payoff of the downstream firm. In our model,

$$F_W^* = \pi_R^* = \left(\frac{a}{2b}\right)^2.$$

The distributional implications of the two-part tariff are as follows: the consumers' payoff increases since they get a higher quantity at a lower price and the wholesaler's payoff increases because it extracts all the surplus from the retailer. In our case, the retailer loses all its surplus, but this is not true in general.

11 Game Theory

In the previous section we introduced game theory in the context of firm competition. In this section, we will generalize the methods used above and introduce some specific language. The specification of (static) game consists of three elements:

1. The players, indexed by $i = 1, \dots, I$. In the duopoly games, for example, the players were the two firms.

2. The strategies available: each player chooses strategy s_i from the available strategy set S_i. We can write $s_{-i} = (s_1, \dots, s_{i-1}, s_{i+1}, \dots, s_I)$ to represent the strategies of the other $I-1$ players. Then, a strategy profile of all players is defined by $s = (s_1, \dots, s_I) = (s_i, s_{-i})$. In the Cournot game, the player's strategies were the quantities chose, hence $S_i = \mathbb{R}_+$.

3. The payoffs for each player as a function of the strategies of the players. We use game theory to analyze situations where there is strategic interaction so the payoff function will typically depend on the strategies of other players as well. We write the payoff function for player i as $u_i(s_i, s_{-i})$. The payoff function is the mapping

$$u_i : S_1 \times \cdots \times S_I \longrightarrow \mathbb{R}.$$

The definition we have given above for a Nash Equilibrium in the oligopoly game extends to games in general.

Definition 11.1. A strategy profile $s^* = (s_1^*, \dots, s_I^*)$ is a *Nash Equilibrium* if, for all i

$$u_i\left(s_i^*, s_{-i}^*\right) \geq u_i\left(s_i, s_{-i}^*\right) \text{ for all } s_i \in S_i.$$

We refer to strategy s_i^* as a best response to s_{-i}^*.

11.1 Pure Strategies

We can represent games (at least those with a finite choice set) in normal form. A normal form game consists of the matrix of payoffs for each player from each possible strategy. If there are two players, 1 and 2, then the normal form game consists of a matrix where the (i, j)th entry consists of the tuple (player 1's payoff, player2's payoff) when player 1 plays their ith strategy and player 2 plays their jth strategy. We will now consider the most famous examples of games.

Example 11.2. (Prisoner's Dilemma) Suppose two suspects, Bob and Rob are arrested for a crime and questioned separately. The police can prove the committed a minor crime, and suspect they have committed a more serious crime but can't prove it. The police offer each suspect that they will let them off for the minor crime if they confess and testify against their partner for the more serious crime. Of course, if the other criminal also confesses the police won't need his testimony but will give him a slightly reduced sentence for cooperating. Each player then has two possible strategies: Stay Quiet (Q) or Confess (C) and they decide simultaneously. We can represent the game with the following payoff matrix:

		Rob	
		Q	C
Bob	Q	3, 3	−1, 4
	C	4, −1	0, 0

Each entry represents (Bob, Rob)'s payoff from each of the two strategies. For example, if Rob stays quiet while Bob confesses Bob's payoff is 4 and Rob's is -1. Notice that both players have what is known as a dominant strategy; they should confess regardless of what the other player has done. If we consider Bob, if Rob is Quiet then confessing gives payoff $4 > 3$, the payoff from staying quiet. If Rob confesses, then Bob should confess since $0 > -1$. The analysis is the same for Rob. So the only stable outcome is for both players to confess. So the only Nash Equilibrium is (Confess, Confess). Notice that, from the perspective of the prisoners this is a bad outcome. In fact it is Pareto dominated by both players staying quiet, which is not a Nash equilibrium.

The above example has a dominant strategy equilibrium, where both players have a unique dominant strategy.

Definition 11.3. A strategy is s_i is *dominant* if

$$u_i(s_i, s_{-i}) > u_i(s_i', s_{-i}) \text{ for all } s_i' \in S_i, s_{-i} \in S_{-i}.$$

If each player has a dominant strategy, then the only rational thing for them to do is to play that strategy no matter what the other players do. Hence, if a dominant strategy equilibrium exists it is a relatively uncontroversial prediction of what will happen in the game. However, it is rare that a dominant strategy will exist in most strategic situations. Consequently, the most commonly used solution concept is Nash Equilibrium, which does not require dominant strategies.

Note the difference between Definitions 11.1 and 11.3: A Nash Equilibrium is only defined for the best response of the other players, s^*_{-i}, whereas dominant strategies have to hold for strategies $s_{-i} \in S_{-i}$. A strategy profile is a Nash Equilibrium if each player is playing a best response to the other players' strategies. So a Nash Equilibrium is a stable outcome where no player could profitably deviate. Clearly when dominant strategies exist it is a Nash Equilibrium for all players to play a dominant strategy. However, as we see from the Prisoner's Dilemma example the outcome is not necessarily efficient. The next example shows that the Nash Equilibrium may not be unique.

Example 11.4. (Coordination Game) We could represent a coordination game where Bob and Ann are two researcher both of whose input is necessary for a project. They decide simultaneously whether to do research (R) or not (N).

		Bob	
		R	N
Ann	R	3, 3	−1, 0
	N	0, −1	1, 1

Here (R,R) and (N,N) are both equilibria. Notice that the equilibria in this game are Pareto ranked with both players preferring to coordinate on doing research. Both players not doing research is also an equilibrium, since if both players think the other will play N they will play N as well.

A famous example of a coordination game is from traffic control. It doesn't really matter if everyone drives on the left or right, as long as everyone drives on the same side.

Example 11.5. Another example of a game is a "beauty contest." Everyone in the class picks a number on the interval $[1, 100]$. The goal is to guess as close as possible to $\frac{2}{3}$ the class average. An equilibrium of this game is for everyone to guess 1. This is in fact the only equilibrium. Since no one can guess more than 100, $\frac{2}{3}$ of the mean cannot be higher than $66\frac{2}{3}$, so all guesses above this are dominated. But since no one will guess more than $66\frac{2}{3}$ the mean cannot be higher than $\frac{2}{3}(66\frac{2}{3}) = 44\frac{4}{9}$, so no one should guess higher than $44\frac{4}{9}$. Repeating this n times no one should guess higher than $\left(\frac{2}{3}\right)^n 100$ and taking $n \longrightarrow \infty$ all players should guess 1. Of course, this isn't necessarily what will happen in practice if people solve the game incorrectly or expect others too. Running this experiment in class the average guess was approximately 12.

11.2 Mixed Strategies

So far we have considered only pure strategies: strategies where the players do not randomize over which action they take. In other words, a pure strategy is a deterministic choice. The following simple example demonstrates that a pure strategy Nash Equilibrium may not always exist.

Example 11.6. (Matching Pennies) Consider the following payoff matrix:

		Bob H	Bob T
Ann	H	1, −1	−1, 1
	T	−1, 1	1, −1

Here Ann wins if both players play the same strategy, and Bob wins if they play different ones. Clearly there cannot be pure strategy equilibrium, since Bob would have an incentive to deviate whenever they play the same strategy and Ann would have an incentive to deviate if they play the differently. Intuitively, the only equilibrium is to randomize between H and T with probability $\frac{1}{2}$ each.

While the idea of a matching pennies game may seem contrived, it is merely the simplest example of a general class of zero-sum games, where the total payoff of the players is constant regardless of the outcome. Consequently gains for one player can only come from losses of the other. For this reason, zero-sum games will rarely have a pure strategy Nash equilibrium. Examples would be chess, or more relevantly, competition between two candidates or political parties. Cold War power politics between the US and USSR was famously (although probably not accurately) modelled as a zero-sum game. Most economic situations are not zero-sum since resources can be used inefficiently.

Example 11.7. A slight variation is the game of Rock-Paper-Scissors.

		Bob R	Bob P	Bob S
	R	0, 0	−1, 1	1, −1
Ann	P	1, −1	0, 0	−1, 1
	S	−1, 1	1, −1	0, 0

Definition 11.8. A *mixed strategy* by player i is a probability distribu-

tion $\sigma_i = \left(\sigma_i\left(S_i^1\right), \ldots, \sigma_i\left(S_i^K\right)\right)$ such that

$$\sigma_i\left(s_i^k\right) \geq 0$$
$$\sum_{k=1}^{K} \sigma_i\left(s_i^k\right) = 1.$$

Here we refer to s_i as an action and to σ_i as a strategy, which in this case is a probability distribution over actions. The action space is $S_i = \left\{s_i^1, \ldots, s_i^K\right\}$.

Expected utility from playing action s_i when the other player plays strategy σ_j is

$$u_i\left(s_i, \sigma_j\right) = \sum_{k=1}^{K} \sigma_j\left(s_j^k\right) u_i\left(s_i, s_j^k\right).$$

Example 11.9. Consider a coordination game (also knows as "battle of the sexes" similar to the one in Example 11.4 but with different payoffs

			Bob	
			σ_B	$1-\sigma_B$
			O	C
Ann	σ_A	O	1, 2	0, 0
	$1-\sigma_A$	C	0, 0	2, 1

Hence Bob prefers to go to the opera (O) and prefers to go to a cricket match (C), but both players would rather go to an event together than alone. There are two pure strategy Nash Equilibria: (O, O) and (C, C). We cannot make a prediction, which equilibrium the players will pick. Moreover, it could be the case that there is a third Nash Equilibrium, in which the players randomize.

Suppose that Ann plays O with probability σ_A and C with probability $1 - \sigma_A$. Then Bob's expected payoff from playing O is

$$2\sigma_A + 0(1-\sigma_A) \tag{38}$$

and his expected payoff from playing C is

$$0\sigma_A + 1(1-\sigma_A). \tag{39}$$

Bob is only willing to randomize between his two pure strategies if he gets the same expected payoff from both. Otherwise he would play the pure strategy that yields the highest expected payoff for sure. Equating (38) and (39) we get that

$$\sigma_A^* = \frac{1}{3}.$$

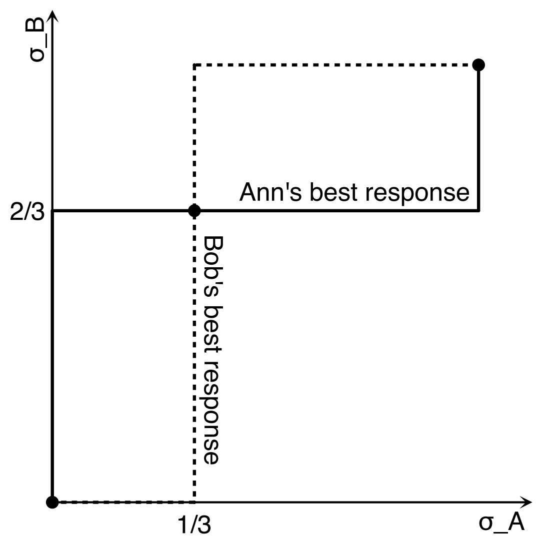

Figure 15: Three Nash Equilibria in battle of the sexes game

In other words, Ann has to play O with probability $\frac{1}{3}$ to induce Bob to play a mixed strategy as well. We can calculate Bob's mixed strategy similarly to get

$$\sigma_B^* = \frac{2}{3}.$$

Graphically, we can depict Ann's and Bob's best response function in Figure 15. The three Nash Equilibria of this game are the three intersections of the best response functions.

12 Asymmetric Information

Asymmetric information simply refers to situations where some of the players have relevant information that other players do not. We consider two types of asymmetric information: adverse selection, also known as hidden information, and moral hazard or hidden action.

A leading example for adverse selection occurs in life or health insurance. If an insurance company offers actuarially fair insurance it attracts insurees with above average risk whereas those with below average risk decline the insurance. (This assumes that individuals have private information about their risk.) In other words, individuals select themselves into insurance based on their private information. Since only the higher risks are in the risk pool the insurance company will make a loss. In consequence of this adverse selection the insurance market breaks down. Solutions to this problems include denying or mandating insurance and offering a menu of contracts to let insurees self-select thereby revealing their risk type.

Moral hazard is also present in insurance markets when insurees' actions depend on having insurance. For example, they might exercise less care when being covered by fire or automobile insurance. This undermines the goal of such insurance, which is to provide risk sharing in the case of property loss. With moral hazard, property loss becomes more likely because insurees do not install smoke detectors, for example. Possible solutions to this problem are copayments and punishment for negligence.

12.1 Adverse Selection

The following model goes back to George Akerlof's 1970 paper on "The market for lemons." The used car market is a good example for adverse selection because there is variation in product quality and this variation is observed by sellers, but not by buyers.

Suppose there is a potential buyer and a potential seller for a car. Suppose that the quality of the car is denoted by $\theta \in [0, 1]$. Buyers and sellers have different valuations/willingness to pay v_b and v_s, so that the value of the car is $v_b\theta$ to the buyer and $v_s\theta$ to the seller. Assume that $v_b > v_s$ so that the buyer always values the car more highly then the seller. So we know that trade is always efficient. Suppose that both the buyer and seller know θ, then we have seen in the bilateral trading section that trade can occur at any price $p \in [v_s\theta, v_b\theta]$ and at that price the efficient allocation (buyer gets the car) is realized (the buyer has a net payoff of $v_b\theta - p$ and the seller gets $p - v_s\theta$, and the total surplus is $v_b\theta - v_s\theta$).

The assumption that the buyer knows the quality of the car may be reasonable in some situations (new car), but in many situations the seller will be much better informed about the car's quality. The buyer of a used car can observe the age, mileage, etc. of a car and so have a rough idea as to quality, but the seller has presumably been driving the car and will know more about it. In such a situation we could consider the quality θ as a random variable, where the buyer knows only the distribution but the seller knows the realization. We could consider a situation where the buyer knows the car is of a high quality with some probability, and low quality otherwise, whereas the seller knows whether the car is high quality. Obviously the car could have a more complicated range of potential qualities. If the seller values a high quality car more, then their decision to participate in the market potentially reveals negative information about the quality, hence the term adverse selection. This is because if the car had higher quality the seller would be less willing to sell it at any given price. How does this type of asymmetric information change the outcome?

Suppose instead that the buyer only knows that $\theta \sim U[0,1]$. That is that the quality is uniformly distributed between 0 and 1. The the seller is willing to trade if

$$p - v_s\theta \geq 0 \tag{40}$$

and the buyer, who does not know θ, but forms its expected value, is willing to trade if

$$E[\theta]v_b - p \geq 0. \tag{41}$$

However, the buyer can infer the car's quality from the price the seller is asking. Using condition (40), the buyer knows that

$$\theta \leq \frac{p}{v_s}$$

so that condition (41) becomes

$$E\left[\theta \middle| \theta \leq \frac{p}{v_s}\right] v_b - p = \frac{p}{2v_s} v_b - p \geq 0, \tag{42}$$

where we use the conditional expectation of a uniform distribution:

$$E[\theta|\theta \leq a] = \frac{a}{2}.$$

Hence, simplifying condition (42), the buyer is only willing to trade if

$$v_b \geq 2v_s.$$

In other words, the buyer's valuation has to exceed twice the seller's valuation for a trade to take place. If

$$2v_s > v_b > v_s$$

trade is efficient, but does not take place if there is asymmetric information.

In order to reduce the amount of private information the seller can offer a warranty of have a third party certify the car's quality.

If we instead assumed that neither the buyer or the seller know the realization of θ then the high quality cars would not be taken out of the market (sellers cannot condition their actions on information they do not have) and so we could have trade. This indicates that it is not the incompleteness of information that causes the problems, but the asymmetry.

12.2 Moral Hazard

Moral hazard is similar to asymmetric information except that instead of considering hidden information, it deals with hidden action. The distinction between the two concepts can be seen in an insurance example. Those who have pre-existing conditions that make them more risky (that are unknown to the insurer) are more likely, all else being equal, to buy insurance. This is adverse selection. An individual who has purchased insurance may become less cautious since the costs of any damage are covered by insurance company. This is moral hazard. There is a large literature in economics on how to structure incentives to mitigate moral hazard. In the insurance example these incentives often take the form of deductibles and partial insurance, or the threat of higher premiums in response to accidents. Similarly an employer may structure a contract to include a bonus/commission rather then a fixed wage to induce an employee to work hard. Below we consider an example of moral hazard, and show that a high price may signal an ability to commit to providing a high quality product.

Suppose a cook can choose between producing a high quality meal ($q = 1$) and a low quality meal ($q = 0$). Assume that the cost of producing a high quality meal is strictly higher than a low quality meal ($c_1 > c_0 > 0$). For a meal of quality q, and price p the benefit to the customer is $q - p$ and to the cook is $p - c_i$. So the total social welfare is

$$q - p + p - c_i = q - c_i$$

and assume that $1 - c_1 > 0 > -c_0$ so that the high quality meal is socially efficient. We assume that the price is set beforehand, and the cook's choice variable is the quality of the meal. Assume that fraction α of the consumers are repeat clients who are informed about the meal's quality, whereas $1-\alpha$ of the consumers are uniformed (visitors to the city perhaps) and don't know the meal's quality. The informed customers will only go to the restaurant if the meal is good (assume $p \in (0,1)$). These informed customers allow us to consider a notion of reputation even though the model is static.

Now consider the decision of the cook as to what quality of meal to produce. If they produce a high quality meal then they sell to the entire market so their profits (per customer) are

$$p - c_1$$

Conversely, by producing the low quality meal, and selling to only $1-\alpha$ of the market they earn profit

$$(1-\alpha)(p-c_0)$$

and so the cook will provide the high quality meal if

$$p - c_1 \geq (1-\alpha)(p-c_0)$$

or

$$\alpha p \geq c_1 - (1-\alpha)c_0$$

where the LHS is the additional revenue from producing a high quality instead of a low quality meal and the RHS is the associated cost. This corresponds to the case

$$\alpha \geq \frac{c_1 - c_0}{p - c_0}.$$

So the cook will provide the high quality meal if the fraction of the informed consumers is high enough. So informed consumers provide a positive externality on the uninformed, since the informed consumers will monitor the quality of the meal, inducing the chef to make a good meal.

Finally notice that price signals quality here: the higher the price the smaller the fraction of informed consumers necessary ensure the high quality meal. If the price is low ($p \approx c_1$) then the cook knows he will lose $p-c_1$ from each informed consumer by producing a low quality meal instead, but gains $c_1 - c_0$ from each uninformed consumer (since the cost is lower). So only if almost every consumer is informed will the cook have an incentive to produce the good meal. As p increases so does $p - c_1$, so the more is lost for each meal not sold to an informed consumer, and hence the lower the fraction of informed consumers necessary to ensure that the good meal will be provided. An uninformed consumer, who also may not know α, could then consider a high price a signal of high quality since it is more likely that the fraction of informed consumers is high enough to support the good meal the higher the price.

12.3 Second Degree Price Discrimination

In Section 9.2 we considered first and third degree price discrimination where the seller can identify the type of potential buyers. In contrast, second degree price discrimination occurs when the firm cannot observe to consumer's willingness to pay directly. Consequently they elicit these preferences by offering different quantities or qualities at different prices. The consumer's type is revealed through which option they choose. This is known as screening.

Suppose there are two types of consumers. One with high valuation of the good θ_h, and one with low valuation θ_l. θ is also called the buyers' marginal willingness to pay. It tells us how much a buyer what be willing to pay for an additional unit of the good. Each buyer's type is his private information. That means the seller does not know ex ante what type a buyer he is facing is. Let α denote the fraction of consumers who have the high valuation. Suppose that the firm can produce a product of quality q at cost $c(q)$ and assume that $c'(q) > 0$ and $c''(q) > 0$.

First, we consider the efficient or first best solution, i.e., the case where the firm can observe the buyers' types. If the firm knew the type of each consumer they could offer a different quality to each consumer. The condition for a consumer of type $i = h, l$ buying an object of quality q for price p voluntarily is

$$\theta_i q - p(q) \geq 0$$

and for the firm to participate in the trade we need

$$p(q) - c(q) \geq 0.$$

Hence maximizing joint payoff is equivalent to

$$\max_q \theta_i q - p(q) + p(q) - c(q)$$

or

$$\max_q \theta_i q - c(q).$$

The FOC for each quality level is

$$\theta_i - c'(q) = 0,$$

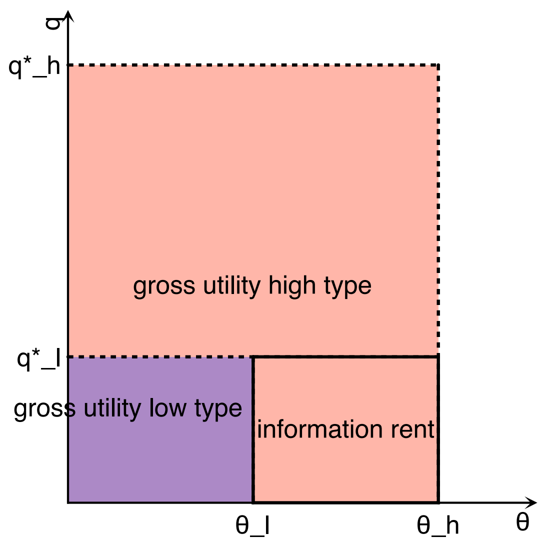

Figure 16: Price discrimination when types are known to the firm

from which we can calculate the optimal level of quality for each type, $q^*(\theta_i)$. Since marginal cost is increasing by assumption we get that

$$q^*(\theta_l) < q^*(\theta_h),$$

i.e., the firm offers a higher quality to buyers who have a higher willingness to pay in the first best case. In the case of complete information we are back to first degree price discrimination and the firm sets the following prices to extract the entire gross utility from both types of buyers:

$$p_h^* = \theta_h q^*(\theta_h) \quad \text{and} \quad p_h^* = \theta_l q^*(\theta_l)$$

so that buyers' net utility is zero. In Figure 16, the buyers' gross utility, which is equal to the price charged, is indicated by the rectangles $\theta_i q_i^*$.

In many situations, the firm will not be able to observe the valuation/willingness to pay of the consumers. That is, the buyers' type is their private information. In such a situation the firm offers a schedule of price-quality pairs and lets the consumers self-select into contracts. Thereby, the consumers reveal their type. Since there are two types of consumers the firm will offer two different quality levels, one for the high valuation consumers and one for the low valuation consumers. Hence there will be a choice of two contracts (p_h, q_h) and (p_l, q_l) (also called a menu of choices). The firm wants high valuation consumers to buy the first contract and low valuation consumers to buy the second contract. Does buyers' private information matter, i.e., do buyers just buy the first best contract intended for them? High type buyers get zero net utility from buying the high quality contract, but positive net utility of $\theta_h q^*(\theta_l) - p_l > 0$. Hence, high type consumers have an incentive to pose as low quality consumers and buy the contract intended for the low type. This is indicated in Figure 16 as "information rent," i.e., an increase in high type buyers' net utility due to asymmetric information.

The firm, not knowing the consumers' type, however, can make the low quality bundle less attractive to high type buyers by decreasing q_l or make the high quality contract more attractive by increasing q_h or decreasing p_h. The firm's profit maximization problem now becomes

$$\max_{p_h, p_l, q_h, q_l} \alpha \left(p_h - c(q_h)\right) + (1-\alpha)\left(p_l - c(q_l)\right). \tag{43}$$

There are two type of constraints. The consumers have the option of walking away, so the firm cannot demand payment higher than the value of the object. That is, we must have

$$\theta_h q_h - p_h \geq 0 \tag{44}$$

$$\theta_l q_l - p_l \geq 0. \tag{45}$$

These are known as the individual rational (IR) or participation constraints that guarantee that the consumers are willing to participate in the trade. The other type of constraints are the self-selection or incentive compatibility (IC) constraints

$$\theta_h q_h - p_h \geq \theta_h q_l - p_l \tag{46}$$

$$\theta_l q_l - p_l \geq \theta_l q_h - p_h, \tag{47}$$

which state that each consumer type prefers the menu choice intended for him to the other contract. Not all of these four constraints can be binding, because that would determine the optimal solution of prices and quality levels. The IC for low type (47) will not binding because low types have no incentive to pretend to be high types: they would pay a high price for quality they do not value highly. On the other hand high type consumers' IR (44) will not be binding either because we argued above that the firm has to incentivize them to pick the high quality contract. This leaves constraints (45) and (46) as binding and we can solve for the optimal prices

$$p_l = \theta_l q_l$$

using constraint (45) and

$$p_h = \theta_h(q_h - q_l) + \theta_l q_l$$

using constraints (45) and (46). Substituting the prices into the profit function (43) yields

$$\max_{q_h, q_l} \alpha \left[\theta_h (q_h - q_l) + \theta_l q_l - c(q_h)\right] + (1-\alpha)\left(\theta_l q_l - c(q_l)\right).$$

The FOC for q_h is simply

$$\alpha \left(\theta_h - c'(q_h)\right) = 0,$$

which is identical to the FOC in the first best case. Hence, the firm offers the high type buyers their first best quality level $q_R^*(\theta_h) = q^*(\theta_h)$. The FOC for q_l is

$$\alpha(\theta_l - \theta_h) + (1-\alpha)\left(\theta_l - c'(q_l)\right) = 0,$$

which can be rewritten as

$$\theta_l - c'(q_l) - \frac{\alpha}{1-\alpha}(\theta_l - \theta_h) = 0.$$

The third term on the LHS, which is positive, is an additional cost that arises because the firm has to make the low quality contract less attractive for high type buyers. Because of this additional cost we get that $q_R^*(\theta_l) < q^*(\theta_l)$: the the quality level for low types is lower than in the first best situation. This is depicted in Figure . The low type consumers' gross utility and the high type buyers' information rent are decreased, but The optimal level of quality offered to low type buyers is decreasing in the fraction of high type consumer α:

$$\frac{dq_R^*(\theta_l)}{d\alpha} < 0$$

since the more high types there are the more the firm has to make the low quality contract unattractive to them.

This analysis indicates some important results about second degree price discrimination:

1. The low type receives no surplus.

2. The high type receives a positive surplus of $q_l(\theta_h - \theta_l)$. This is known as an information rent, that the consumer can extract because the seller does not know his type.

3. The firm should set the efficient quality for the high valuation type.

4. The firm will degrade the quality for the low type in order to lower the rents the high type consumers can extract.

13 Auctions

Auctions are an important application of games of incomplete information. There are many markets where goods are allocated by auctions. Besides obvious examples such as auctions of antique furniture there are many recent application. A leading example is Google's sponsored search auctions. Google matches advertiser to readers of websites and auctions advertising space according to complicated rules.

Consider a standard auction with I bidders, and each bidder i from 1 to I has a valuation v_i for a single object which is sold by the seller or auctioneer. If the bidder wins the object at price p_i then he receives utility $v_i - p_i$. Losing bidders receive a payoff or zero. The valuation is often the bidder's private information so that we have to analyze the uncertainty inherent in such auctions. This uncertainty is captured by modelling the bidders' valuations as draws from a random distribution:

$$v_i \sim F(v_i).$$

We assume that bidders are symmetric, i.e., their valuations come from the same distribution, and we let b_i denote the bid of player i.

There are many possible rules for auctions. They can be either sealed bid or open bid. Examples of sealed bid auctions are the first price auction (where the winner is the bidder with the highest bid and they pay their bid), and the second price auction (where the bidder with the highest bid wins the object and pays the second highest bid as a price). Open bid auctions include English auctions (the auctioneer sets a low price and keeps increasing the price until all but one player has dropped out) and the Dutch auction (a high price is set and the price is gradually lowered until someone accepts the offered price). Another type of auction is the Japanese button auction, which resembles an open bid ascending auction, but every time the price is raised all bidders have to signal their willingness to increase their bid. Sometimes, bidders hold down a button as long as they want to increase their bid and release when they want to exit the auction.

Let's think about the optimal bidding strategy in a Japanese button auction, denotes by $b_i(v_i) = t_i$, where $t_i = p_i$ is the price the winning bidder pays for the good. At any time, the distribution of valuations, F, the number of remaining bidders are known to all players. As long as the price has not reached a bidder's valuation it is optimal for him to keep the button pressed because he gets a positive payoff if all other players exit before the price reaches his valuation. In particular, the bidder with the highest valuation will wait longest and therefore receive the good. He will only have to pay the second highest bidder's valuation, however, because he should release the button as soon as he is the only one left. At that time the price will have exactly reached the second highest valuation. Hence, it is optimal for all bidders to bid their true valuation. If the price exceeds v_i they release the button and get 0 and the highest valuation bidder gets a positive payoff. In other words, the optimal strategy is

$$b_i^*(v_i) = v_i.$$

What if the button auction is played as a descending auction instead? Then it is no longer optimal to bid one's own valuation. Instead, $b_i^*(v_i) < v_i$ because only waiting until the price reaches one's own valuation would mean that there might be a missed chance to get a strictly positive payoff.

In many situations (specifically when the other players' valuations does not affect your valuation) the optimal behavior in a second price auction is equivalent to an English auction, and the optimal behaviour in a first price auction is equivalent to a Dutch auction. This provides a motivation for considering the second price auction which is strategically very simple, since the English auction is commonly used. It's the mechanism used in the auction houses, and is a good first approximation how auctions are run on eBay.

How should people bid in a second price auction? Typically a given bidder will not know the bids/valuations of the other bidders. A nice feature of the second price auction is that the optimal strategy is very simple and does not depend on this information: each bidder should bid their true valuation.

Proposition 13.1. *In a second price auction it is a Nash Equilibrium for all players to bid their valuations. That is $b_i^* = v_i$ for all i is a Nash Equilibrium.*

Proof. Without loss of generality, we can assume that player 1 has the highest valuation. That is, we can assume $v_1 = \max_i\{v_i\}$. Similarly, we can assume without loss of generality that the second highest valuation is $v_2 = \max_{i>1}\{v_i\}$. Define

$$\mu_i(v_i, b_i, b_{-i}) = \left\{ \begin{array}{c} v_i - p_i, \text{ if } b_1 = \max_j\{b_j\} \\ 0, \text{ otherwise} \end{array} \right\}$$

to be the surplus generated from the auction for each player i. Then under the given strategies $(b = v)$

$$\mu_i(v_i, v_i, v_{-i}) = \left\{ \begin{array}{c} v_1 - v_2, \ i = 1 \\ 0, \text{ otherwise} \end{array} \right\}$$

So we want to show that no bidder has an incentive to deviate.

First we consider player 1. The payoff from bidding b_1 is

$$\mu_1(v_1, b_1, v_{-1}) = \begin{Bmatrix} v_1 - v_2, \text{ if } b_1 > v_2 \\ 0, \text{ otherwise} \end{Bmatrix} \leq v_1 - v_2 = \mu_1(v_1, v_1, v_{-1})$$

so player 1 cannot benefit from deviating.

Now consider any other player $i > 1$. They win the object only if they bid more than v_1 and would pay v_1. So the payoff from bidding b_1 is

$$\mu_i(v_i, b_i, v_{-i}) = \begin{Bmatrix} v_i - v_1, \text{ if } b_i > v_1 \\ 0, \text{ otherwise} \end{Bmatrix} \leq 0 = \mu_i(v_i, v_i, v_{-i})$$

since $v_i - v_1 \leq 0$. So player i has no incentive to deviate either.

We have thus verified that all players are choosing a best response, and so the strategies are a Nash Equilibrium. □

Note that this allocation is efficient. The bidder with the highest valuation gets the good.

Finally, we consider a first price sealed bid auction. There, we will see that it is optimal for bidders to bid below their valuation, $b_i^*(v_i) < v_i$, a strategy called bid shedding. Bidder i's expected payoff is

$$\max_{b_i} (v_i - b_i) \Pr(b_i > b_j \text{ for all } j \neq i) + 0 \Pr(b_i < \max\{b_j\} \text{ for all } j \neq i). \tag{48}$$

Consider the bidding strategy

$$b_i(v_i) = cv_i$$

i.e., bidders bid a fraction of their true valuation. Then, if all players play this strategy,

$$\Pr(b_i > b_j) = \Pr(b_i > cv_j) = \Pr\left(v_j < \frac{b_i}{c}\right). \tag{49}$$

With valuations having a uniform distribution on $[0, 1]$, (49) becomes

$$\Pr\left(v_j < \frac{b_i}{c}\right) = \frac{b_i}{c}$$

and (48) becomes

$$\max_{b_i} (v_i - b_i) \frac{b_i}{c} + 0$$

with FOC

$$\frac{v_i - 2b_i}{c} = 0$$

or

$$b_i^* = \frac{v_i}{2}.$$

Hence, we have verified that the optimal strategy is to bid a fraction of one's valuation, in particular, $c = \frac{1}{2}$.