

PRINCIPLES OF BEHAVIORAL ECONOMICS

Prof. Sujata Kar

Department of Management Studies

IIT Roorkee

Week 09

Lecture 09

This is Lecture 9 of the course on behavioral economics, and in Lecture 9, we are going to discuss the standard model in neoclassical economics. Basically, with this, we are beginning the second topic. The second topic will talk about values, preferences, judgments, or alternatively, heuristics, biases, etc. Before we talk about why we need to understand heuristics, biases, and all, it is important to understand what the basic premises of the standard model in neoclassical economics are, or what the neoclassical tradition actually offers us, against which we are going to build several other theories. So, we need to talk about violations.

And before we talk about violations, we need to focus on the specifics of the neoclassical tradition. So, the standard economic model is basically something that posits that individuals are rational, self-interested, and utility-maximizing agents, with complete information, consistent preferences, and perfect foresight. So, this part we have already discussed. We call these kinds of individuals ECONs.

Now, we will get into the details of what the standard model of economics actually assumes in order to define an ECON. So, we begin with the key assumptions. First of all, rationality—we have already talked about it. Individuals make decisions based on a logical and consistent evaluation of costs and benefits. Then it assumes that individuals are driven by self-interest, so they primarily act to maximize their own personal gain or utility.

The next one is utility maximization. The neoclassical tradition always assumes that individuals strive to achieve the highest possible level of satisfaction or well-being from their choices. Furthermore, we assume that there is complete information, which implies that individuals have access to all the relevant information about prices, products, and potential outcomes. So, it is not always only about prices, products, and potential outcomes.

Generally, any economic model would assume that both parties or all parties involved have perfect information.

If there is no perfect information, which is most often the case around us, then we have this problem of information asymmetry. So, economists have also worked on the concept of information asymmetry. But for the most part, we assume that information is perfectly available to all parties involved in most transactions. Then we assume consistent preferences. Individuals have well-defined and consistent preferences, which implies that their choices are logical and predictable, pretty much similar to the assumption of rationality.

Basically, all of them together imply that individual decision-makers are rational. Individuals also have perfect foresight, which implies that they can accurately predict future outcomes and make decisions based on their predictions. The most accepted mathematical model is the augmented neoclassical model. Now, we will consider a stylized version of the standard model with three components of rationality. The expression of the neoclassical model, which we in short write as NM in mathematical terms, enables us to achieve three important objectives.

And first of all, they are a concise description of the relevant factors affecting decision-making. Number two, an illustration of the various components of the model that can be examined. And number three, a general consideration of the assumptions underlying the model in terms of how they relate to the various components. So, as you can see, neoclassical modeling techniques are extremely systematic. And that is probably why they have been a favorite of economists.

You can go very mathematically, very systematically, for that you may need to use certain assumptions. Nevertheless, that systematization is actually extremely attractive. Now, the model can be stated in the following terms. We are now going to define a model, and to begin with, we use or explain some notations. So, an individual i at time t equals 0, which most often refers to the present time, maximizes expected utility subject to a probability distribution of the states of the world.

Where small s belongs to capital S . So, this S consists of all possible states, and small s refers to a single state. P refers to the probability associated with the single state. Now, we will be giving examples of that, but then you can consider many different situations around us. Where we basically have alternatives, and there are certain uncertainties associated with

them—what is going to happen in the future—and accordingly, we can take a call. So, for example, for a student, one needs to allocate study time to different subjects.

You can always associate probabilities with your performance in the exam so different states would be your different performances in different subjects associate probabilities with them. Some subjects you do not like much so even if you study hard maybe you don't understand much you don't like it much or you may not perform well in those subjects the probability of scoring high is much low as compared to those subjects in which you have an expertise, you like them, you enjoy studying them and that is why the probability of scoring high is also high. So, different states have different probabilities.

Now, you need to understand how much time or you want to calculate how much time you would like to devote for individual subjects. So, the possible scores or the probability of getting a particular score in a subject multiplied by that score and then in a similar fashion, I can have the probability of different subjects and the probability and the scores in those subjects. all added them together is going to give me the total utility. I can try to solve them in order to find out how much time I am going to devote to individual subject.

So, this is how we write it. We are going to maximize basically a stream of utility with each output subject to a particular state, there is certain probability associated with it. We tend to discount the future events and that is why a discount rate is associated here or given here. See here in this case, everything is going to happen in future.

But it is possible that something is going to happen today, something is going to happen tomorrow, something is going to happen day after tomorrow. So, there is a stream of utility forthcoming. You need to decide how to maximize your utility by choosing, say, the level of consumption in different time periods. So, accordingly, there could be discount rates associated with future consumption. I previously mentioned that discount rates are like we do not equally value present and future.

So, something which is coming in future is valued less as compared to something that is here at present. Very simple examples are, I have 100 rupees today. If I want to lend it to someone else I would be expecting an interest rate. Why is it so? Because when I lend it out then it implies that I am foregoing my present consumption, so I need a reward for them or reward for this compromise or sacrifice. As a result of which what is hundred rupees to me today -- today hundred rupees is actually not equal to tomorrow's hundred rupees tomorrow

Since I compromised it today, I need something larger tomorrow. So, tomorrow I could be happy with not 100 rupees but say rather 110 rupees. Alternatively, if I ask someone that I want to give you some money either today or say 30 days after today. So, you can take 100 rupees today or you can take 110 rupees tomorrow. So on the 30th day which one you would prefer? If somebody says that I want 100 rupees today instead of 110 rupees after 30 days which means he or she discounts the future.

110 rupees after 30 days is not equal to 100 rupees. Of course, understandably, 100 rupees today will also not be equal to 100 rupees on the 30th day. So, since these valuations are different, we say that the future is discounted. And we are basically summing them up—the future stream of consumptions. Future streams multiplied by the probability of their occurrences, and then we maximize that sum.

$$\max_{x_i^t \in X} \sum_{t=0}^{\infty} \delta^t \sum_{s_t \in S_t} p(s_t) U(x_i^t | s_t)$$

This is a standard exposition of the augmented neoclassical model. The utility function $U(x)$ given (s) is defined over the payoff x_i^t , which is individual i 's outcome or consumption at time t , given as x_i^t . So, it is mentioned here: payoff x_i^t of individual i , and future utility is discounted with a time-consistent discount factor, which is δ . Now, the time-consistent discount factor is very specific to the neoclassical model, but we will talk about it much later when we discuss intertemporal choice models. So, for the time being, we just talk about a discount factor, which is denoted by δ .

Now, let us break the model into four components.

$$\underbrace{1. \max}_{x_i^t \in X} \quad \underbrace{2. \sum_{t=0}^{\infty} \delta^t} \quad \underbrace{3. \sum_{s_t \in S_t} p(s_t)} \quad \underbrace{4. U(x_i^t | s_t)}$$

The first one is we are maximizing where we are choosing the possible output x_i^t , which belongs to x , consisting of all possible x_i^t . Then, we have the second component, which is the summation of the discount factor t running from 0 to infinity. So, for t equals 0, we do not have any discount. At the time—at present—our δ raised to the power 0 is equal to 1.

From next time onwards, period 1 will have delta equal to discount, which is equal to delta 1. Period 1 discount will be delta. So, at t equals 1. At t equals 2, we will have delta squared and so on. Then, we have these probabilities summed up, and that is basically for each state; we have an individual probability that the state belongs to the set of all states at time period t .

And finally, the fourth component is the summation of all the utilities. The main assumption underlying the neoclassical model can now be stated in terms of how it relates to these components. First of all, economic agents are rational, which is reflected in all the components: 1, 2, 3, 4. Economic agents are maximizing; they are in an optimization problem, maximizing some of the possible utilities. Then, they are using a consistent discount factor, which is also one characteristic of rationality.

They can predict the future reasonably well. So, they have probabilities associated with each and every possible outcome. And then finally, a very clear idea about the stream of utility. Economic agents are motivated by expected utility maximization. Again, they are going for maximization.

This is the expected thing and this is utility. I have reproduced the equation at the top. So, further an agent's utility is governed by purely selfish concerns in the narrow sense that it does not take into consideration the utility of others. So, as you can see that when we talk about individual's utility, individual I 's utility, it actually concerns itself about that individual only. The utility generated by the consumption of other individual's are not coming into picture here, which might not be the case always.

I may get certain pleasure by donating some money. In that case, consumption by others are also adding to my utility. Yesterday or in the last lecture, we talked about the vicarious consumption of Veblen. Here he talked about what kind of pleasure the family head gets when he sees that his family members are enjoying luxury. They are, in case of rich people, even the rich people is actually earning money or the person, the family head is earning money, the other family members are also enjoying.

So that is vicarious consumption. Similarly, even for middle class individuals, when they see that he is able to or an individual is able to arrange money for his family, he gets some pleasure. So, it is not that we get pleasure out of only our own consumption, we get pleasure from the consumption of others as well, but that is not considered in the neoclassical framework. Agents are Bayesian probability operators.

So, we have these probabilities here. Agents have consistent time preferences according to the discounted utility model. I repeat, the discounted utility model—we are going to talk about it much later while discussing intertemporal choice models. All income and assets are completely fungible. Fungibility is again one concept we will talk about at length or focus on while discussing the concept of mental accounting.

For the time being, fungibility implies that money does not have any labels associated with it. So, it is not that the salary I receive at the beginning of the month or at the end of the month has any specification about how much should be spent on food, how much should be spent on children's education, how much should be saved, and things like that. So, the idea is that the entire amount can be spent on any possible goods and services as I wish, which apparently sounds perfectly fine.

This is how we, of course, look at money, but when it comes to our behavior, we will discuss later that this is not true. We often violate the assumption of fungibility. Now, let us take an example of the neoclassical model. Suppose a student is going to attend a lecture or a course on behavioral economics. Now, he does not have much idea about how the course is going to be.

Now, the state of the world is the course can be interesting, the course can be boring. He develops certain probabilities or Bayesian priors by consulting seniors. And gets to know that the probability that the course will be, or the class will be interesting, is 0.8, and there is a 20% chance that the class will be boring. Now, the class is in the morning, and before going to the class, he has two options: either he can drink coffee or he can drink beer.

So now, we have the situation; this is the decision one needs to make. So, we can call it x_i , where i takes two values: whether to drink coffee or beer. The state of the world, s , has again two alternative states: interesting and boring. Probabilities are as we were denoting them by ψ . So, the probability of the class being interesting is 80 percent, and the class being boring is 20 percent, and this is the payoff.

So, this is the expected utility associated with each and every decision. So, along with the state. So, x_i given the state s_i or we can write it s_j . What is happening here is that if the class is interesting, then he expects a payoff of 10, provided he drank coffee before coming to the class. Because coffee keeps you awake. Class is interesting, you have taken coffee, so now you are able to concentrate better.

As a result, you have high utility at 10. The problem is that if you have taken coffee and the class turns out to be boring, then you are not able to even sleep during the class. So that's a lot of trouble. And that's why you have the lowest possible utility in such a situation.

Say you perceive these are again individual perception. This may vary from individual to individual. So one individual perceives that the payoff would be as low as 2. Now, in case you have taken beer in the morning and the class turns out to be interesting, then what happens is that you are not able to focus much.

So, as a result, your utility is actually somewhat low. You get a payoff of 6. The class is interesting, which is why you have a slightly higher payoff. But since you are not able to focus, that reduces your payoff. And if the class is boring and you had taken beer, then of course it is a pretty good situation.

You do not like it much because the class is boring, which disappoints you and reduces your utility. But since you are able to sleep during the class, that adds to your utility. So, you can expect a payoff of 4. These are all hypothetical numbers, and I repeat that these are subjective evaluations that may vary from individual to individual. Now, having said that, what we understand is what my expected utility is from the two situations.

So, in this situation, if I drink coffee, then I will have 0.8 multiplied by 10. Plus 0.2 multiplied by 2. So, this implies I will have a utility of 8.4. In a similar fashion, if I drink beer, then from this, like 0.8 multiplied by 6 plus 0.2 multiplied by 4, it will give me something like 5.6. So, since 5.6 is less than 8.4, given the information, you decide to drink coffee before going to the class.

So, neoclassical economists say that individuals may make decisions in this way. Therefore, the optimal decision, which maximizes expected utility, depends on the probability estimates of the state of the world. The student should verify that the expected payoff or utility of drinking coffee is 8.4, while the expected utility of drinking beer is 5.6. Thus, the best decision in this situation is to drink coffee. However, if the probabilities were reversed, meaning the probability that the class would be a good one is just 20% or 0.2,

Then the optimal decision would be to drink beer, or we can recalculate given the information. We can see from this example that the estimation of Bayesian prior probabilities has an important effect on decision-making. Bayesian prior probabilities are the probabilities assigned to individual states. The rational person will update these in light

of new information, so if the class turns out to be boring, this will reduce the estimated probability of the next class being interesting and may affect the student's drink decision next time around.

So, this is what is called the Discovered Preference Hypothesis. This was suggested by Plott in 1966. The standard economic model is a good predictor if people have had ample opportunity to learn from experience. So Plott said that one can always discover preferences over repeated trials. So on the first day of class, if my expectation doesn't align with reality, then in the next class I go and update it further

and so after several classes, I may finally understand whether the class is actually interesting or boring, or what the probabilities of the class being interesting or boring are and accordingly, whether I should drink coffee or beer. But of course, there are certain situations where we cannot have repeated trials for example, people generally get married only once. Then there are several other decisions, like retiring only once. So certain decisions actually come into your life only once and cannot be repeated In that case, the discovered preference hypothesis may not be applicable.

Now, talking about the decision-making process, let us understand decision-making by examining the process involved in most decisions. Of course, some decisions do not really appear to be decisions at all because we act on instinct without conscious thought. This is often because there is no time for using the 'cold' analytical cognitive system. Instead, we rely on the 'hot' affective system, which is often dependent on instantaneous emotional or visceral responses to a situation. For example, ducking an object flying at our head or lashing out in anger—physically or verbally—at someone who has offended us.

There are certain reactions that are very much automatic. We call them reactions in a hot state. So, basically, when we are not able to process things properly. However, even in cases where this hot system is used and we are not aware of making a conscious decision, there is still a situation where some kind of stimulus results in an action. Whenever this situation arises, we deem that a decision-making process is involved.

What is needed, therefore, is a general model of the decision-making process that applies to automatically reaching for the Oreo cookies at the supermarket, deciding where to buy a lottery ticket, investing in a particular stock, or choosing a particular medical treatment. So, decision-making processes also entail certain automatic reactions, right? So, maybe we need to understand both situations. When we are making decisions by thinking about them or thinking them over several times, then the decisions are supposed to be highly rational.

We cannot question them. But the decisions which are taken out of our natural reflexes or automatically. The automatic part of our brain, when it gets activated, leads to certain reactions and decisions. They are probably of more interest to us in terms of how people behave in such situations and whether we then conform to the assumptions of neoclassical understanding. Obviously, these decisions are very different in nature and involve different factors and even decision systems, but there are certain elements that are common.

There are three fundamental characteristics of decision-making. These are described below along with their corresponding components in the neoclassical model as expressed previously. So, first of all, I will rewrite those components of the neoclassical model. The first component of the decision-making process is preference. These are the rankings people have over a set of options or gambles that are based on attitudes and values related to the outcomes of these options.

That refers to component 4. Number 2 is beliefs. These relate to the probabilities with which people think various outcomes will occur, conditional on available information. So, as you understand, that refers to component 3. Then we have rationality.

This involves all four components of the standard model, referring to the ways in which people first determine preferences based on attitudes and values. So, component four. Basically, our utilities are dependent on our attitudes and values. The value we associate with different outcomes or different situations, states, etc. Appropriately modify their beliefs in light of new information.

So, this is component 3. Discount values of future outcomes. We are talking about component 2. And succeed in choosing optimal actions given their preferences and beliefs. So, given all this information, how we choose the optimal action by maximizing or following the maximization principle.

It should be noted that there is significant interdependence and overlap between these three categories. Beliefs affect preferences, and both of these affect choices as reflected in decision-making. The beliefs, preferences, and rationality classification is useful in that the factors within each category are involved at different stages of the decision-making process. In a comprehensive survey of empirical findings in the field related to behavioral economics, DellaVigna examined Deviations from the neoclassical model according to the above classification.

He refers to these deviations as non-standard preferences, non-standard beliefs, and non-standard decision-making. So the deviations have been there, but as I told you in the beginning, this topic broadly covers several modules on the deviations. But before we discuss the deviations, we need to know about some important approaches and tools that the neoclassicals use. The standard model can be easily translated into a neoclassical model of consumer behavior. So now, let us become more specific.

The previous approach of the model is the broadest possible one representing neoclassical thoughts, but we can have different modeling structures for different segments or different problems. So, talking about consumer behavior, first of all, the objects of consumer choice are referred to as consumption bundles, consisting of a complete list of the goods and services that are being considered. In a generalized situation, the context also matters.

For example, people care about the future availability of goods and services, One would care more for a bottle of water in the middle of a desert rather than in the Antarctic. So, basically, our preferences or choices are heavily influenced by various factors, including time, context, etc. Most often, we may not talk about it, and we are rather going for a generalized version of the behavior—consumer behavior that we would get to see around us. So now, I will talk about some terminologies and notations.

Any bundle of goods X can be described in the simplest terms as X equals X_1, X_2 , where X_1 denotes the amount of one good and X_2 the amount of another good. Now, this bundle consists of only two commodities. But this bundle can be extended to as many commodities as possible. For example, we can write X as equal to X_1, X_2 up to X_n . But generally, we use two commodities because it facilitates the use of graphical representation and analysis.

In the neoclassical model, it is assumed that consumers can rank bundles according to their desirability. If a consumer definitely wants an X bundle rather than a Y bundle, then it is said that he or she strictly prefers the X bundle to the Y bundle. And this is denoted by (X_1, X_2) preferred to (Y_1, Y_2) . Alternatively, if a consumer is indifferent between the two bundles, this means that they have no preference for either bundle over the other. This relationship is usually described as $(X_1, X_2) \sim (Y_1, Y_2)$, where ' \sim ' indicates indifference.

Finally, if a consumer prefers or is indifferent between the two bundles, we say that he or she weakly prefers (X_1, X_2) to (Y_1, Y_2) . And this is the notation that is generally used, which is similar to the 'greater than or equal to' sign but often has a curvature. So, this is how we might write 'preferred to.' Preference relations are meant to be operational notions.

This is also assumed in the neoclassical model that choice is determined by preference. This is important since it is choice that is directly observable, not preference. And accordingly, we have this theory of revealed preference, which says that if a consumer chooses a particular bundle, it is assumed that this bundle is preferred to another bundle. If that other bundle was both available and affordable. So, the bundle was available for purchase, and the consumer had the money to buy that bundle as well.

So, this is what is meant by the concept of revealed preference. Revealed preference results basically in the choice; the choices are observed, we purchase them, and otherwise, if choices are not made, then preferences will also not be revealed. The theory of consumer choice is often illustrated using a graphical approach involving indifference curves. An indifference curve represents different combinations of two goods between which the consumer is indifferent, that is, which yield the same total utility. So, this is how the indifference curves look.

Those who have a background in economics would be very familiar with indifference curves. For those who are non-initiated, I will just briefly say that these curves are downward-sloping, as you understand that along an individual indifference curve, say I_1 , the level of utility is constant. So what is happening is that this is one combination point where the individual is, say, consuming Y_0 and X_0 . Now, if I increase the amount of X naught or consumption of X from X naught to X_1 , then, of course, utility increases.

That depends on some of the assumptions. We will talk about them later. But what we broadly assume is that if I get more of one commodity, I must be better off. The moment I am better off, my utility increases. The two points where only the amount of x increases, while the amount of y remains constant, cannot be on the same indifference curve.

So, to increase x , I must reduce the amount of y . When I reduce y , it is possible to have another point where utility remains the same. Now, how much I reduce y certainly depends on the slope of the indifference curve. If these curves are much flatter, it implies that I do not need to reduce y much to increase x . So, I increase x proportionally more while reducing y proportionally less. Alternatively, if my indifference curves are very steep, it implies that, given my initial consumption, if I increase x ,

I must reduce y more than proportionally. So, how much I reduce y to maintain the same level of utility depends on the slope of the indifference curves and they are dictated by individual consumer choices. This indifference map shows a number of indifference curves with curves farther from the origin representing bundles with greater quantities of goods.

These bundles are therefore preferred to those on lower indifference curves. Simply speaking, higher indifference curves give a higher level of utility, which means that the utility associated with I1 is less than the utility associated with I2, and the utility associated with I3, the indifference curve, is actually the highest. Thus, any combination of goods on curve I2 is preferred to any combination on curve I1. Indifference curves are normally drawn as downward-sloping and convex to the origin.

So, why are they normally downward-sloping? That I just discussed in the previous slide, and they are convex to the origin, which has certain implications also, but it is not mandatory for them to be convex to the origin. Indifference curve maps can be used to illustrate the notion of consumer equilibrium. It is assumed that consumers have a budget constraint relating to the amount of money they have available to spend on the relevant consumption bundles. This budget constraint can be represented by the following inequality, where P_x is the price of x multiplied by the amount of x consumed.

P_y is the price of y multiplied by y , basically the amount of y consumed. And that should be less than or equal to the total money available, M , to be spent on these two commodities. So this is how a budget constraint looks. The maximum amount that I can spend on commodity X is M divided by P_x . The maximum amount I can spend on commodity Y is M divided by P_y .

In between, I can spend any amount on both commodities. So here I am using Y naught, or I am purchasing Y naught and X naught. Here I am purchasing Y_1 and X_1 . So these are various combinations that can be purchased using the same amount of money. Any point within the budget constraint implies that the budget is not fully utilized.

I am purchasing some amount of Y and some amount of X . When they are added together, they would be strictly less than m . This is the indifference curves equilibrium graphically. So, what we are doing is that we initially gave a set of indifference curves. Now, this is a budget constraint. I superimpose the set of indifference curves on the budget constraint, and we arrive at equilibrium here.

Why? Because let us consider one point A . I have already mentioned that I1 is less preferred to I2, which is less preferred to I3. Higher indifference curves give a higher level of utility. Utility increases, right, as we move up. So, from point A , by consuming less of Y and more of X , the consumer can reach point C .

Similarly, from point B, by consuming less of X and more of Y, the consumer can again reach point C. Now, see that any point above C would lie on another indifference curve, giving a higher level of utility, but that point is not attainable given the consumer's budget constraint of M. So, C is the tangency point which gives the maximum possible utility or maximum attainable utility from two different commodities. We will be consuming X_1 and Y_1 . By consuming X_1 and Y_1 commodities at the given prices, we are maximizing utility.

So, formally, although the budget is sufficient to buy combinations A and B, these are not optimal combinations since they are not on the highest indifference curve that can be reached. Combination C is the optimal combination where the consumer purchases bundles X_1 and Y_1 . This situation can be generalized. Any optimal point of consumption will occur at a point of tangency between an indifference curve and the budget line.

Setting the slopes of these two curves equal gives the condition that the consumer should spend so that the marginal utility of the last dollar spent on each good is the same. This condition can be further generalized to apply to any number of goods in a consumption bundle. The mathematics associated with it may be done later. The exposition of consumer equilibrium in the neoclassical model often uses the term marginal utility. But in fact, this is not essential.

The slope of an indifference curve can also be expressed in terms of the marginal rate of substitution of one good for another. The MRS XY represents the amount of Y that a consumer is prepared to give up in order to get one more unit of X. Thus, the equilibrium condition amounts to consuming a combination such that the MRS is equal to the ratio of the prices, that is, MRS XY equals P_x upon P_y . However, the interesting part is that MRS XY is also equal to the marginal utility from X divided by the marginal utility of Y. So, they essentially boil down to the same thing.

However, the advantage of this form of exposition is that it avoids the thorny concept of utility. With this, I conclude this topic or the module. Thank you so much.