PRINCIPLES OF BEHAVIORAL ECONOMICS

Prof. Sujata Kar

Department of Management Studies
IIT Roorkee

Week 30

Lecture 30

Hello everyone, this is Lecture 30 of the course on Principles of Behavioral Economics. We concluded the discussion on prospect theory as presented by Kahneman and Tversky in their two papers published in 1979 and 1992 in the previous module. Now, in this module, we will discuss some of the criticisms raised against prospect theory. PT has now been around for over three decades. During that time, not surprisingly—considering its unconventional nature and radical implications—it has attracted numerous criticisms.

It is also important to distinguish between PT and CPT, that is, prospect theory and cumulative prospect theory, since different criticisms may apply in each case. Some of the criticisms have been theoretical, and some have been empirical. So, in the first case, it has been claimed that the theory contains contradictions, is incomplete, or lacks clarification, whereas in the second case, that is, in the case of CPT—here, the first case refers to PT, and the second case refers to CPT—the claims are that empirical data violates the assumptions of the theory.

and that the theory makes incorrect predictions. But it should also be noted here that some of the criticisms raised against PT are addressed in CPT. Also, some of the criticisms raised against both PT and CPT are observed to be invalid by other authors. Nevertheless, we list out the criticisms that have been pointed out. So, in terms of theory, there are four main criticisms here.

The lack of normative status, internal contradictions, incompleteness, and the determination of reference points. Empirical criticisms relate to violations of the combination principle, violations of stochastic dominance, the failure to explain the Allais paradoxes, violations of gain-loss separability, the nature of the utility function, endowment effects (one such example), the discovered preference hypothesis and misconceptions, and the nature of framing effects. Now, as you can see, some of these

criticisms are highlighted. That is because we are going to discuss only these criticisms, as some of the other criticisms that are not highlighted are either no longer relevant

or perhaps the criticisms are contradicted by other authors, The third possibility is that, due to paucity of time and space, I just need to focus on some of the important criticisms. So, we begin with internal contradictions. The criticism comes particularly from Birnbaum, who drew attention to the editing rules, which he stated as imprecise, contradictory, and in conflict with the equations of PT. These contradictions mean that PT can explain most empirical results simply by applying the editing rules selectively.

So, here, 'exposed' implies that once the results are observed, we can always, by using selective editing rules, ensure that the empirical observations are explained by our theory. Unfortunately, the downside of this characteristic is that the theory is claimed to be difficult to use for prediction purposes on an ex-ante basis. Ex-ante basis implies forecasting when things have not yet happened. So, it is not able to predict behavior much, rather, the behaviors that are observed can, to some extent, be explained by using selective editing rules. This has unfortunate consequences also concerning the explanation of the Allais paradoxes.

Now, we will consider an experiment to explain these internal contradictions. So, in this experiment, each gamble is represented by an urn containing 100 marbles that are identical except for their colors. The urn for gamble A contains one red and one blue marble, each of which pays $100, and it has 98 white marbles that pay nothing. So, this is an important description of gamble A. Urn B, which refers to the urn under gamble B, contains one red marble paying $100, two green marbles that pay $45,

and 97 white marbles that pay nothing. Now, the prospects are given as, first of all, choice 1. In choice 1, you have gamble A and gamble B. You have to choose one of them. So, as far as the description goes, gamble A can be written as you win $100 with a probability of 0.01 because there is only one red ball.

You also win $100 with a probability of 0.01 because there is only one blue ball, and there are 98 white balls for which you do not win anything. So, it is basically 0. Similarly, for B, there is one red ball, which gives you $100 with a probability of 0.01 because there is one red ball. There are two green balls, each giving you $45. Since there are two green balls, the probability is 0.02, and then there are 97 white balls, which do not give you anything, so you get 0. Now, applying combination, choice 2 can be obtained as A prime.

Well, I am actually combining the first two prospect, right. So, this is the combination rule. If I combine the first two outcomes of prospect A, then I will have 0.02 probability of winning 100. And then 0.98 to win nothing. B remains as it is.

So this is how we can make some changes—basically, some editing is done to the first set of gambles. Now, if I apply cancellation, then choice 3 can be obtained by applying cancellation to the very first two prospects. Now, you can see that the first two components of A and B are the same. So, they can be ignored altogether because they are common to both of them. This is what the cancellation principle implies.

If that is done, then there is a 0.01 probability of winning 100 and and then the rest 99 you win nothing similarly we have 0.02 to win 45 and then there is 98 that you win nothing so you can see that first combination and then cancellation is applied however choices between 1 and 2 that is choice 1 and choice 2 and the one between 2 and 3 or the one between 1 and 3 need not be the same and may result in preference reversal. So, this is what the internal contradictions in terms of the way prospect theory suggest applications of editing rules,

which may itself lead to some internal contradictions or inconsistent preferences. The next one is the failure to explain Allais paradoxes. Different theories explain the Allais paradoxes in different ways. We have already seen that PT was originally constructed to explain these paradoxes in terms of editing principles. To test the implications of PT, Birnbaum devised some tests that dissected the Allais paradoxes by presenting them in gambles of different forms,

some using the combination principle, some using the cancellation principle. Based on the empirical data from 200 participants, Birnbaum concluded that neither OPT, that is the original prospect theory, nor CPT, that is cumulative prospect theory, with or without their editing principles of cancellation and combination, can account for the dissection of the Allais paradoxes. So, just the example we took in the context of internal contradictions, that itself is an example of the theory failing to explain some of the Allais paradoxes.

The next criticism is the determination of reference points. Some economists view it as a weakness of PT that reference points are not determined endogenously. The determination of reference points is necessary to estimate the incidence and effects of loss aversion. A good example of this problem relates to the endowment effect. If subjects are given something for free in an experiment, they are likely to value this object differently than if they had earned it in some way.

So basically when we earn something, they are dearer to us as compared to when we get something as a windfall in later topics also we will come across the differences in terms of our treatment towards monetary gains or any kind of gains one which is coming to us as a gift which is as a windfall and something which we actually earn with lot of efforts so hard earned money is always valued more than something which comes to us freely, easily, as a gift or windfall. Kahneman, Tversky, and other supporters of PT generally use either the existing situation as a reference point or some anticipated or expected situation.

However, more precision regarding the determination of reference points would be an aid to constructing better models of behavior in the sense that if we actually pay attention to the reference point that from where it is coming, is the reference point relate to my hard earned money? Or if the reference point relates to some sudden gains, then it's quite possible that when the reference points vary, my attitude toward different uncertain or risky situations would change. For example, it would be useful to know if different types of savers or investors have the same reference points or consider the dynamic adjustment of reference points over time, like the writing off of sunk costs and things like that.

The next criticism refers to the concept of loss aversion. This is a key component of prospect theory and implies that most decision-makers find symmetric bets, such as lotteries of the form 0.5x, 0.5 minus x, unattractive. This is something we also discussed while discussing the prospect theory that symmetric bets are actually not much preferred when you win certain amount and lose the same amount with equal probabilities. However, a study by Adam and Kroll in 2012 indicated that a majority of risk-averse subjects prefer similar symmetric lotteries to a sure thing.

Having established that a sample of, say, 50 subjects were all risk-averse by comparing their certainty equivalence to a lottery of the form 0.5, 50-0.5, 0, the study then examined willingness to accept a lottery of 0.5 y plus x, 0.5 y minus x in preference to a sure thing of y. So, first of all, they established the fact that all the subjects were risk-averse by comparing their certainty equivalence with the expected value from a lottery like this. And then it examined the willingness to accept a lottery of 0.5 y plus x. y plus x 0.5, y minus x—this is a symmetric bet in preference to a sure thing of y. PT predicts that such subjects would prefer the sure thing because of loss aversion.

But 96% of the subjects played at least 1 out of 25 lotteries, with an average of 15.9 lotteries played. Adam and Kroll suggested that these results provide evidence in favor of a theory of attraction to chance caused by emotional factors. The next criticism is the nature of the

utility function. Another criticism relating to the utility function has been raised by Levy and Levy, in short LL, in 2002 arguing that the original KT data did not provide a reliable indicator of the shape of utility functions because it always asked subjects to compare prospects that were both either positive or negative.

So, basically, to begin with, they were strictly positive or strictly negative prospects. In reality, the LL study claimed most prospects are mixed, involving situations where there is a possibility of either gain or loss —for example, in investing in the stock market. So, when you invest in the stock market, there is always a possibility of making some losses and making some gains —maybe the probabilities vary. Sometimes the winning probabilities are higher compared to losing probabilities; sometimes it's the reverse.

Nevertheless, most often these are mixed prospects. The study included a number of experiments asking respondents to choose between such mixed prospects. The main objective was to test whether the data supported the PT model with the S-shaped utility function or the Markowitz model with the reversed S-shaped function throughout most of the range. So, the Markowitz utility function—also in an earlier module—I have just shown that this is not typically the S-shaped utility function as suggested by prospect theory.

Rather, it actually has some portions where it is convex and some portions where it is concave. So, in some way, these are inverse—rather, these are the S-shaped utility function suggested by Markowitz, the reversed S-shaped utility function suggested by Markowitz. The study used a total of 260 subjects, consisting of business students and faculty from a number of institutions, along with a number of professional practitioners. The subjects were asked to consider that they invested $10,000 in stock and were evaluating possible returns, choosing between the following two mixed prospects.

In the first, that is prospect F, you lose $3,000 with a probability of 0.5, and there is an equal probability of winning $4,500. The other prospect, G, gives you a probability of 0.25 to lose $6,000, while you win or gain $3,000 with a probability of 0.75. Now, if I compare these two prospects, we see that both prospects are, first of all, mixed. Their expected values are the same.

In both situations, we would be getting an expected value of $750, and the same applies to both the gains and the loss components. That is, 3000 multiplied by 0.5 gives us 1500, so here minus 3000 becomes minus 1500. Similarly, minus 6000 multiplied by 0.25 gives minus 1500. On the other hand, 4500 multiplied by 0.5 gives 2250, and similarly, I will be

getting 2250 by multiplying 3000 with 0.75, right? So, the gain and loss expected values are the same, and the overall expected values are the same for the mixed prospects.

Now, according to PT, people are risk-averse in the domain of gains. Therefore, they should prefer the gain of 3000 with a probability of 0.75 in prospect G to a gain of 4500 with a probability of 0.5 in prospect F. Here, the gains are higher, but the probabilities are much higher here. So that's why risk-averse individuals should choose 3000 over 4500. Similarly, the PT model proposes that people are risk-seeking in the domain of losses.

Thus, they should prefer a loss of 6000 with a smaller probability like 0.25 in prospect G to a loss of 3000 with a higher probability of 0.5 in prospect F. So, combining these two, we see that prospect G dominates prospect F. However, LL found that 71% of their subjects preferred F, while only 27% preferred G, which supports the Markowitz model. So, as predicted, the observations did not support Prospect Theory's predictions. They interpreted this finding as strong evidence against the PT model, combined with the results of other tasks in their experiments.

They concluded that the Markowitz model was better supported. Next, we have violations of gain-loss separability. For mixed prospects, CPT describes the valuation of gains and losses separately, and then they are added together. This assumes the feature of gain-loss separability—that we can always separate gains from losses. In simple terms, this means that if you prefer the good part of B to the good part of A and if you prefer the bad part of B to the bad part of A, then you should prefer B to A.

However, various empirical studies have contradicted this feature. For example, Wu and Markle in 2008 reported a reversal between preferences for mixed gambles and the associated gain and loss gambles. Such that the mixed gamble A is preferred to mixed gamble B, but the gain and loss portions of B are preferred to the gain and loss portions of A. So, what it simply tries to say is that B and A both are mixed prospects. So, B has some negative component and some positive component, and similarly, A also has some negative component and some positive component.

So, it has been observed that the negative component of B is preferred to the negative component of A. In a similar fashion, the positive component of B is preferred to the positive component of A, which implies that B should ideally be preferred to A. But then, people actually preferred overall A compared to B, so that's why gain-loss separability might not be valid all the time. So, if you separate gains and losses, they may individually

look attractive, but when you combine them together, then a combination or combined prospect may have completely different preferences or choices.

The next one is the nature of framing effects. The inconsistent results have been reported regarding the ability of PT to explain framing effects. Different types of framing effects have been demonstrated, such as standard risky choice, attribute framing, and goal framing. It is claimed that PT probably best explains the first type of effect, i.e. standard risky choice, but not the other two.

It is also doubted that PT could interpret the empirical evidence of risky choices in different contexts. or that a framing effect depends on task, content, and context variables inherent in the choice problem. In the context of the Asian disease problem, which we have already discussed, people are informed about a disease that threatens 600 citizens and asked to choose between two undesirable options. In the positive frame, people are given a choice between A, which saves 200 lives with certainty, or B, a one-third chance of saving all 600 with a two-thirds chance of saving nobody.

So, you can see that, as we previously mentioned, this is the positive or the gain frame. Most people preferred A to B here. In the negative frame, people are asked to choose between C, which is 400 people dying with certainty, or D, a two-thirds chance of 600 dying and a one-third chance of nobody dying. In this case, most people preferred D to C, in spite of the fact that A and C are identical results or prospects, and B and D are again identical results or prospects. Several studies have argued that this approach actually confounded two different effects: a framing effect and a reflection effect.

So, as previously mentioned while discussing the Asian disease problem, it contains several effects. Two of them were framing and reflection. So, these impacts are considered to be confounded in the sense that they together impact the decision. So, things cannot be particularly ascribed to any single effect. A framing effect depends on whether the problem is framed in a positive or negative way, which depends on the negation 'not'.

The reflection effect depends on the domain of the problem, meaning whether it relates to a gain or loss. To illustrate this difference, statement A, '200 people will be saved,' represents both a positive frame and a positive domain, whereas statement C, '400 people will die,' involves both a negative frame and a negative domain. It is therefore argued that because frame and domain correlate perfectly in Tversky and Kahneman's treatment of the Asian disease problem, it is impossible to disentangle the framing and reflection effects.

On the other hand, it can be claimed that the statement '400 people will not be saved,' although identical in meaning to statement A, involves a negative frame but a positive domain. Similarly, the statement '200 people will not die' is identical in meaning to statement C but involves a positive frame with a negative domain. Thus, by restating A and C, it is possible to test PT against other theories as far as explaining framing effects is concerned.

So, with this, I conclude this module on the criticism of prospect theory, and this also concludes the third topic on prospect theory. Thank you.