

Prof. Bryan Caplan
bcaplan@gmu.edu
<http://www.bcaplan.com>
Econ 812

Week 10: Behavioral Economics and Irrationality, I

- I. The Behavioral Approach and Choice Theory
 - A. Most economists never even think about empirically testing fundamental micro choice theory. Why?
 - B. Elementary consumer theory almost seems true by definition.
 - C. If however we assume that preferences are *stable* - as almost all economists do in empirical work - there are a lot of testable implications.
 - D. Moreover, if we assume that preferences are selfish in the ordinary language sense of the word - another standard auxiliary assumption - there are a great many other testable implications.
 - E. Once we move from basic consumer theory to expected utility theory, there are lots of testable implications.
 - F. A rapidly expanding literature - often called "behavioral economics" - conducts precisely the empirical tests that most economists never think about running.
 - G. The product of this literature is a long list of "anomalies" - robust evidence that people sometimes violate basic axioms of choice theory.
 - H. These violations of choice theory are sometimes equated with "irrationality." Economists who earned their Ph.D.'s prior to the RE revolution are particularly likely to talk this way.
 - I. Before surveying some of the main documented anomalies, it is worth pre-answering a few objections.
 - J. Objection #1: "All theories are false. What matters is prediction."
 1. Reply: It is usually just as easy to provide evidence that the **predictions** of basic micro fail as it is to show that the assumptions are false.
 - K. Objection #2: "Deviations cancel out."
 1. Reply: They don't! They are systematic.
 - L. Objection #3: "Anomalies arise due to weak incentives."
 1. Reply: Stronger incentives often don't matter. And anomalies appear even in financial markets, where incentives would appear to be great.
 - M. To some extent I will be playing devil's advocate. Most economists familiar with behavioral economics either dismiss it or see it as highly significant. I personally often take an intermediate position.
- II. Preference Reversals
 - A. For practical purposes, economists almost always assume that people have *constant* preferences over *outcomes*.

- B. But behavioral economists have uncovered a number of what appear to be counter-examples. A single individual will prefer A to B or B to A for apparently irrelevant reasons. This is known as a "preference reversal."
 - C. Logically equivalent descriptions of the same choice problem elicit different choices: Doctors will select one form of surgery if you tell them it has a 90% chance of success, but make a different choice if you tell them it has a 10% chance of failure.
 - D. People sometimes select the choice they put less monetary value on: Given a choice of two bets (H: 8/9 chance of \$4; L:1/9 chance of \$40), most subjects choose H over L. But if they own the gamble and you want to buy it, most subjects demand more money to sell L than H.
 - E. Preference reversals have received an enormous amount of attention. But it is hard to me to see the real-world significance. True, if you aim to persuade others, you probably do better by saying "The glass is half full." But could you persuade a lot more people to play Russian roulette by saying "You have a 5/6 chance of surviving"?
 - F. Most preference reversal experiments focus on "close" choices. Could you induce a heavy metal fan to reverse his preference for Ozzy Osbourne over Bach? Even the choice of surgery may be fairly "close" - both routes look pretty bad, but the discrete structure of the problem masks this.
- III. The Endowment Effect and Status Quo Bias
- A. Another well-documented way that people deviate from basic choice theory is that their endowments somehow interact with their preferences.
 - B. Simplest anomaly: the endowment effect. People seem to put more value on objects merely because they are *their* objects. In a classic experiment, agents who were given a coffee cup had a markedly higher willingness to accept than willingness to pay. A few critics appealed to wealth effects, but that is a pretty lame objection.
 - C. Aside: This has some interesting implications for the Coase Theorem.
 - D. A more complex anomaly: status quo bias. People are somewhat reluctant to both buy AND sell. Alternatives become more popular purely by being designated as the status quo, even when transactions costs are negligible.
 - E. In the real world, it is easy to attribute apparent instances of the endowment effect and status quo bias to transactions costs or information costs. Usually the status quo option, for example, gets to be the status quo because most people prefer it that way. (Think toppings on a hamburger). And again, most experiments focus on

"close" choices. If you gave someone 100 coffee cups, I strongly suspect he would sell the vast majority without a premium.

IV. Selfishness and Cooperation

- A. While pure theorists rarely mention it, almost all real-world applications of economic theory assume that people are narrowly selfish.
 - 1. Slight variation: Inclusive fitness.
- B. Of course, perfectly selfish non-relatives can sustain cooperation through repeated play. So mere cooperation is hardly evidence against human selfishness.
- C. But: There is plenty of experimental evidence that people will cooperate even in one-shot games. Why? They care about strangers to some degree. (See General Zod in *Superman II*).
- D. Much of this evidence comes from public goods experiments. Even in one-shot games, agents contribute 40-60% of the socially optimal level.
- E. Cooperation declines with repeated play, but if you "re-start" a tournament with experienced players, cooperation initially jumps up again.
- F. There is plenty of extra-experimental evidence, too.
 - 1. Tipping
 - 2. Charity
 - 3. Voting
- G. Real-world significance? Clearly it is there to some degree. We offer and accept small favors from strangers all the time. We pass up and expect others to pass up small chances to take advantage. Still, people on average keep 98% of their income for themselves.

V. Fairness and Vindictiveness

- A. The opposite of selfishness is altruism - caring directly about the well-being of others. But empirically, interpersonal motivation seems richer than either.
- B. For one thing, treating other people better than selfishness recommends often seems to be motivated by concern for "fairness" rather than directly caring about others. We seem more concerned about how we treat people that we directly interact with, and pay more attention to whether we behaved "fairly" than the actual welfare of others.
- C. Thus, in ultimatum and dictator experiments, first-movers often offer splits with the second-mover, but rarely share their winnings with the next stranger they meet. Fairness suggests the first, but not the second.
- D. A second motive that operates in the ultimatum game, but not the dictator game, is "vindictiveness." Especially when we have been treated unfairly, we often put a *negative* weight on the welfare of another person.

- E. Real-world significance? Again, it is easy to observe on some level. Even when there is no repeated interaction, we give up small personal benefits to do what fairness requires, and expect others to do the same. Prices and wages might be more volatile in the absence of fairness and vindictiveness motives. A few glaring shortages (concert tickets, for instance) would disappear. But how big is the overall effect?
- VI. Preference Heterogeneity
- A. Many economists not only assume that preferences are constant over time; they also assume that they are identical across individuals. Stigler and Becker made this a standard methodological position, with the slogan "you can explain anything with preferences."
 - B. There is ample empirical evidence, however, that this is simply not so. In my *JEBO* paper on personality and economics, I review a wealth of evidence from personality psychology indicating a high degree of preference heterogeneity.
 - C. Methodological point: Preferences only "explain everything" if preferences are not independently measured.
 - D. Applications.
- VII. Expected Utility Anomalies
- A. Recall that expected utility theory puts definite restrictions on choice under uncertainty. Many of these have been experimentally falsified.
 - B. People often seem risk-averse over utility, not just wealth.
 - C. People often seem risk-averse relative to a "reference point." E.g., if a wealthy person plays a low-stakes game of poker, he is likely to play as if he had a large risk premium, even though he remains rich if he loses.
 - D. Choice is not linear in probabilities, as EU theory predicts.
 - E. The Allais paradox. Consider the following choices of gambles:
 1. \$27,500 w/p=.33, \$24,000 w/p=.66, \$0 w/p=.01; \$24,000 w/p=1.
 2. \$27,500 w/p=.33, \$0 w/p=.67; \$24,000 w/p=.34, \$0 w/p=.66.
 - F. Most people take the second choice in the first case and the first choice in second case. EU theory says this is impossible. (Can you prove why?)
- VIII. Loss Aversion and Prospect Theory
- A. One general lesson that behavioral economists attempt to draw from various findings is that people are "loss averse." In basic consumer theory, this makes no sense, because preferences and endowments are separate. The intuition behind "loss aversion," though, is that having something you currently possess taken from you is worse than never having had it at all.
 - B. This insight gives rise to one of the major positive theoretical innovations of behavioral economics, known as "prospect theory."

Basic idea: replace EU w/risk aversion with an S-shaped curve that kinks at the current "reference point."

- C. Such a curve indicates that people are risk-averse in gains but risk-seeking in losses. Given a choice of an extra \$1000 or a gamble with the same expected value, they prefer the \$1000. Given a choice of a loss of \$1000 or a gamble with the same expected value, they prefer the gamble.
 - D. The simple version of prospect version raises as many problems as it solves. In particular, it seems to predict no one would want insurance against losses. In practice, almost all insurance works that way.
 - E. A more sophisticated version of prospect theory says that people are:
 1. Risk-averse in gains with high probabilities.
 2. Risk-seeking in gains with low probabilities.
 3. Risk-averse in losses with low probabilities.
 4. Risk-seeking in losses with high probabilities.
 - F. Many experts in behavioral economics find this highly illuminating, but obviously it has a lot of wiggle room: you can fiddle with the reference point as well as the cut-point between "high" and "low" probabilities.
- IX. Intertemporal Anomalies
- A. With perfect capital markets, basic micro says that everyone will discount future payments by the rate of interest. Even with imperfect capital markets, moreover, there are often definite predictions of intertemporal choice.
 - B. Behavioral economists have uncovered a variety of intertemporal anomalies. In many cases, consumers appear to have *negative* discount rates - choosing, for example, to be paid over 12 months rather than 9, or persistently receiving tax refunds.
 - C. In other cases, consumers appear to have discount rates far in excess of the interest rate. In buying major appliances, for example, they appear to put little weight on future energy cost savings.
 - D. Furthermore, consumers often appear to have different discount rates for gains versus losses.
 - E. Even more striking, discount rates often seem to vary with the total waiting time. They discount a benefit a year-and-a-day from now only slightly more than a benefit a year from now. But they discount a benefit tomorrow a great deal compared to a benefit today.
 - F. This implies "dynamic inconsistency." As time goes by, consumers actually regret previous decisions and want to change them.
 - G. It is natural to see this as a sort of "self-control" problem. A person may want to save for retirement, but face a constant urge to spend. Therefore, they might want to have money withheld from their

paycheck to overcome "temptation." Similarly, a person may want to lose weight, but at every particular moment they have an urge to eat. Therefore they might, for example, avoid having food in the house to avoid temptation.

- H. It is widely assumed that the long-term plan is somehow better or more reflective of the agent's true preferences than the short-term plan, but is there any warrant for this assumption?
- I. Many of these example are striking, but again, what do they mean in the real world? It hardly seems like there are abundant opportunities to loan money to people at 100% interest. And it is not at all clear that it would be worth my time at current interest rates if I investigated the energy efficiency of my appliances. It is somewhat interesting that people turn down free interest, but the dollar value does not seem that high.