Behavioral Economics: Past, Present, Future

Colin F. Camerer

Division of Humanities and Social Sciences 228-77
Caltech
Pasadena, CA 91125
camerer@hss.caltech.edu

George Loewenstein

Department of Social and Decision Sciences
Carnegie-Mellon University
Pittsburgh PA 15213
gl20+@andrew.cmu.edu

draft: 10/25/02

Behavioral economics increases the explanatory power of economics by providing it with more realistic psychological foundations. This book consists of representative recent articles in behavioral economics.[1] This chapter is intended to provide an introduction to the approach and methods of behavioral economics, and to some of its major findings, applications, and promising new directions. It also seeks to fill some unavoidable gaps in the chapters' coverage of topics.

What Behavioral Economics Tries To Do

At the core of behavioral economics is the conviction that increasing the realism of the psychological underpinnings of economic analysis will improve economics *on its own terms* -- generating theoretical insights, making better predictions of field phenomena, and suggesting better policy. This conviction does not imply a wholesale rejection of the neoclassical approach to economics based on utility maximization, equilibrium, and efficiency. The neoclassical approach is useful because it provides economists with a theoretical framework that can be applied to almost any form of economic (and even non-economic) behavior, and it makes

[1] Since it is a book of advances, many of the seminal articles which influenced those collected here are not included, but are noted below and are widely reprinted elsewhere.

refutable predictions. Many of these predictions are tested in the chapters of this book, and rejections of those predictions suggest new theories.

Most of the papers modify one or two assumptions in standard theory in the direction of greater psychological realism. Often these departures are not radical at all because they relax simplifying assumptions that are not central to the economic approach. For example, there is nothing in core neoclassical theory that specifies that people should not care about fairness, that they should weight risky outcomes in a linear fashion, or that they must discount the future exponentially at a constant rate.[2] Other assumptions simply acknowledge human limits on computational power, willpower, and self-interest. These assumptions can be considered 'procedurally rational' (Herbert Simon’s term) because they posit functional heuristics for solving problems that are often so complex that they cannot be solved exactly by even modern computer algorithms.

Evaluating Behavioral Economics

Stigler (1965) says economic theories should be judged by three criteria: congruence with reality, generality, and tractability. Theories in behavioral economics should be judged this way too. We share the positivist view that the ultimate test of a theory is the accuracy of its predictions.[3] But we also believe that, ceteris paribus, better predictions are likely to result from theories with more *realistic* assumptions.

Theories in behavioral economics also strive for *generality* – e.g., by adding only one or two parameters to standard models. Particular parameter values then often reduce the behavioral model to the standard one, and the behavioral model can be pitted against the standard model by estimating parameter values. And once parameter values are pinned down, the behavioral model can be applied just as widely as the standard one.

---

[2]While the papers in this book largely adhere to the basic neoclassical framework, there is nothing inherent in behavioral economics that *requires* one to embrace the neoclassical economic model. Indeed, we consider it likely that alternative paradigms will eventually be proposed which have greater explanatory power. Recent developments in psychology, such as connectionist models that capture some of the essential features of neural functioning, bear little resemblance to models based on utility maximization, yet are reaching the point where they are able to predict many judgmental and behavioral phenomena.

[3]Contrary to the positivistic view, however, we believe that predictions of feelings (e.g., of subjective well-being) should also be an important goal.

Adding behavioral assumptions often *does* make the models less tractable. However, many of the papers represented in this volume show that it can be done. Moreover, despite the fact that they often add parameters to standard models, behavioral models, in some cases, can even be more *precise* than traditional ones which assume more rationality, when there is dynamics and strategic interaction. Thus, Lucas (1986) noted that rational expectations allows multiple inflationary and asset price paths in dynamic models, while adaptive expectations pins down one path. The same is true in game theory: Models based on cognitive algorithms (e.g., Camerer, Ho & Chong, 2001) often generate precise predictions in those games where the mutual consistency requirement of Nash permits multiple equilibria.

The realism, generality and tractability of behavioral economics can be illustrated with the example of loss-aversion. Loss-aversion is the disparity between the strong aversion to losses relative to a reference point and the weaker desire for gains of equivalent magnitude. Loss aversion is more *realistic* than the standard continuous, concave, utility function over wealth, as demonstrated by hundreds of experiments. Loss aversion has proved useful in identifying where predictions of standard theories will go wrong: Loss-aversion can help account for the equity premium puzzle in finance and asymmetry in price elasticities. (We provide more examples below.) Loss aversion can also be parameterized in a general way, as the ratio of the marginal disutility of a loss relative to the marginal utility of a gain at the reference point (i.e., the ratio of the derivatives at zero); the standard model is the special case in which this "loss-aversion coefficient" is one. As the foregoing suggests, loss-aversion has proved *tractable*—although not always simple-- in several recent applications (e.g., Barberis, Huang & Santos, 2001).

The Historical Context Of Behavioral Economics

Most of the ideas in behavioral economics are not new; indeed, they return to the roots of neoclassical economics after a century-long detour. When economics first became identified as a distinct field of study, psychology did not exist as a discipline. Many economists moonlighted as the psychologists of their times. Adam Smith, who is best known for the concept of the "invisible hand" and The Wealth of Nations, wrote a less well-known book The Theory of Moral Sentiments, which laid out psychological principles of individual behavior that are arguably as profound as his economic observations. The book is bursting with insights about human psychology, many of which presage current developments in behavioral economics. For example, Adam Smith commented (1759/1892, 311) that "we suffer more... when we fall from a better to a worse situation, than we ever enjoy when we rise from a worse to a better." Loss aversion! Jeremy Bentham, whose utility concept formed the foundation of neoclassical economics, wrote extensively about the psychological underpinnings of utility, and some of his insights into the determinants of utility are only now starting to be appreciated (Loewenstein 1999). Francis Edgeworth's Theory of Mathematical Psychics, which introduced his famous "box" diagram showing two-person bargaining outcomes, also included a simple model of social utility, in which one person's utility was affected by another person's payoff, which is a springboard for modern theories (see chapters 9 and 10 for two examples).

The rejection of academic psychology by economists, perhaps somewhat paradoxically, began with the neoclassical revolution, which constructed an account of economic behavior built up from assumptions about the nature—that is, the *psychology*—of homo-economicus. At the turn of the 20th century, economists hoped their discipline could be like a natural science. Psychology was just emerging at that time, and was not very scientific. The economists thought it provided too unsteady a foundation for economics. Their distaste for the psychology of their period, as well as dissatisfaction with the hedonistic assumptions of Benthamite utility, led to a movement to expunge the psychology from economics.[4]

Expunging psychology from economics happened slowly. In the early part of the 20th century, the writings of economists such as Irving Fisher and Vilfredo Pareto still included rich speculations about how people feel and think about economic choices. Later John Maynard Keynes very much appealed to psychological insights, but by the middle of the century discussions of psychology had largely disappeared.

Throughout the second half of the century, many criticisms of the positivistic perspective took place in both economics and psychology. In economics, researchers like George Katona, Harvey Leibenstein, Tibor Scitovsky, and Herbert Simon wrote books and articles suggesting the

[4]The economists of the time had less disagreement with psychology than they realized. Prominent psychologists of the time were united with the economists in rejecting hedonism as the basis of behavior. William James, for example, wrote that "psychologic hedonists obey a curiously narrow teleological superstition, for they assume without foundation that behavior always aims at the *goal* of maximum pleasure and minimum pain; but behavior is often impulsive, not goal-oriented," while William McDougall stated in 1908 that "it would be a libel, not altogether devoid of truth, to say that classical political economy was a tissue of false conclusions drawn from false psychological assumptions." (Both quotes from Lewin (1996).)

importance of psychological measures and bounds on rationality. These commentators attracted attention, but did not alter the fundamental direction of economics.

Many coincident developments led to the emergence of behavioral economics as represented in this book. One development was the rapid acceptance by economists of the expected utility and discounted utility models as normative and descriptive models of decision making under uncertainty and intertemporal choice, respectively. Whereas the assumptions and implications of generic utility analysis are rather flexible, and hence tricky to refute, the expected utility and discounted utility models have numerous precise and testable implications. As a result, they provided some of the first "hard targets" for critics of the standard theory. Seminal papers by Allais (1953), Ellsberg (1961) and Markowitz (1952) pointed out anomalous implications of expected and subjective expected utility. Strotz (1955) questioned exponential discounting. Later scientists demonstrated similar anomalies using compelling experiments that were easy to replicate (Kahneman & Tversky, 1979, on expected utility, and Thaler, 1981, and Loewenstein & Prelec, 1992, on discounted utility).

As economists began to accept anomalies as counterexamples that could not be permanently ignored, developments in psychology identified promising directions for new theory. Beginning around 1960, cognitive psychology became dominated by the metaphor of the brain as an information-processing device replacing the behaviorist conception of the brain as a stimulus-response machine. The information-processing metaphor permitted a fresh study of neglected topics like memory, problem solving and decision making. These new topics were more obviously relevant to the neoclassical conception of utility maximization than behaviorism had appeared to be. Psychologists such as Ward Edwards, Duncan Luce, Amos Tversky and Daniel Kahneman, began to use economic models as a benchmark against which to contrast their psychological models. Perhaps the two most influential contributions were published by Tversky and Kahneman. Their 1974 Science article argued that heuristic short-cuts created probability judgments which deviated from statistical principles. Their 1979 paper "Prospect theory: decision making under risk" documented violations of expected utility and proposed an axiomatic theory, grounded in psychophysical principles, to explain the violations. The latter was published in the technical journal Econometrica and is one of the most widely cited papers ever published in that journal.

A later milestone was the 1986 conference at the University of Chicago, at which an extraordinary range of social scientists presented papers (see Hogarth & Reder, 1987). Ten years later, in 1997, a special issue of the Quarterly Journal of Economics was devoted to behavioral economics (three of those papers are reprinted in this volume).

Early papers established a recipe that many lines of research in behavioral economics have followed. First, identify normative assumptions or models that are ubiquitously used by economists, such as Bayesian updating, expected utility and discounted utility. Second, identify anomalies—i.e., demonstrate clear violations of the assumption or model, and painstakingly rule out alternative explanations (such as subjects' confusion or transactions costs). And third, use the anomalies as inspiration to create alternative theories that generalize existing models. A fourth step is to construct economic models of behavior using the behavioral assumptions from the third step, derive fresh implications, and test them. This final step has only been taken more recently but is well represented in this volume of advances.

The Methods Of Behavioral Economics

The methods used in behavioral economics are the same as those in other areas of economics. At its inception, behavioral economics relied heavily on evidence generated by experiments. More recently, however, behavioral economists have moved beyond experimentation and embraced the full range of methods employed by economists. Most prominently, a number of recent contributions to behavioral economics, including several included in this book (Chapters 21, 25 and 26, and studies discussed in chapters 7 and 11) rely on field data. Other recent papers utilize methods such as field experiments (Gneezy and Rustichini (this volume) computer simulation (Angeletos et al., 2001), and even brain scans (McCabe et al, 2001).

Experiments played a large role in the initial phase of behavioral economics because experimental control is exceptionally helpful for distinguishing behavioral explanations from standard ones. For example, players in highly anonymous one-shot take-it-or-leave-it "ultimatum" bargaining experiments frequently reject substantial monetary offers, ending the game with nothing (see Camerer & Thaler, 1995). Offers of 20% or less of a sum are rejected about half the time, even when the amount being divided is several weeks' wages or $400 in the

US (e.g., Camerer, 2002). Suppose we observed this phenomenon in the field, in the form of failures of legal cases to settle before trial, costly divorce proceedings, and labor strikes. It would be difficult to tell whether rejection of offers was the result of reputation-building in repeated games, agency problems (between clients and lawyers), confusion, or an expression of distaste for being treated unfairly. In ultimatum game experiments, the first three of these explanations are ruled out because the experiments are played once anonymously, have no agents, and are simple enough to rule out confusion. Thus, the experimental data clearly establish that subjects are expressing concern for fairness. Other experiments have been useful for testing whether judgment errors which individuals commonly make in psychology experiments also affect prices and quantities in markets. The lab is especially useful for these studies because individual and market-level data can be observed simultaneously (e.g., Camerer, 1987; Ganguly, Kagel & Moser, 2000).

Although behavioral economists initially relied extensively on experimental data, we see behavioral economics as a very different enterprise from experimental economics (see Loewenstein, 1999). As noted, behavioral economists are methodological eclectics. They define themselves, not on the basis of the research methods that they employ, but rather their application of psychological insights to economics. Experimental economists, on the other hand, define themselves on the basis of their endorsement and use of experimentation as a research tool. Consistent with this orientation, experimental economists have made a major investment in developing novel experimental methods that are suitable for addressing economic issues, and have achieving a virtual consensus among themselves on a number of important methodological issues.

This consensus includes features that we find appealing and worthy of emulation (see Hertwig & Ortmann, in press). For example, experimental economists often make instructions and software available for precise replication, and raw data are typically archived or generously shared for reanalysis. Experimental economists also insist on paying performance-based incentives, which reduces response noise (but does not typically improve rationality; see Camerer & Hogarth, 1999), and also have a virtual prohibition against deceiving subjects.

However, experimental economists have also developed rules that many behavioral economists are likely to find excessively restrictive. For example, experimental economists rarely collect data like demographics, self-reports, response times, and other cognitive measures which behavioral economists have found useful. Descriptions of the experimental environment are usually abstract rather than evocative of a particular context in the outside world because economic theory rarely makes a prediction about how contextual labels would matter, and experimenters are concerned about losing control over incentives if choosing strategies with certain labels is appealing because of the labels themselves. Psychological research shows that the effect of context on decision making can be powerful (see, e.g., Goldstein & Weber, 1995; Loewenstein, 2001) and some recent experimental economics studies have explored context effects too (e.g., Cooper, Kagel, Lo & Gu, 1999; Hoffman et al, 1994). Given that context is likely to matter, the question is whether to treat it as a nuisance variable or an interesting treatment variable. It is worth debating further whether helping subjects see a connection between the experiment and the naturally-occurring situations the experiments is designed to model, by using contextual cues, is helpful or not.

Economics experiments also typically use "stationary replication"—in which the same task is repeated over and over, with fresh endowments in each period. Data from the last few periods of the experiment are typically used to draw conclusions about equilibrium behavior outside the lab. While we believe that examining behavior after it has converged is of great interest, it is also obvious that many important aspects of economic life are like the first few periods of an experiment rather than the last. If we think of marriage, educational decisions, and saving for retirement, or the purchase of large durables like houses, sailboats, and cars, which happen just a few times in a person's life, a focus exclusively on "post-convergence" behavior is clearly not warranted.[5]

All said, the focus on psychological realism and economic applicability of research promoted by the behavioral-economics perspective suggests the immense usefulness of both empirical research outside the lab and of a broader range of approaches to laboratory research.

[5]We call the standard approach "Groundhog Day" replication, after the Bill Murray movie in which the hero finds himself reliving exactly the same day over and over. Murray's character is depressed until he realizes that he has the ideal opportunity to learn by trial-and-error, in a stationary environment, and uses the opportunity to learn how to woo his love interest.

Basic Concepts and Research Findings

The field of Behavioral Decision Research, on which behavioral economics has drawn more than any other subfield of psychology, typically classifies research into two categories: judgment and choice. Judgment research deals with the processes people use to estimate probabilities. Choice deals with the processes people use to select among actions, taking account of any relevant judgments they may have made. In this section, we provide a background on these two general topics to put the contributions of specific chapters into a broader context.

Probability judgment

Judging the likelihood of events is central to economic life. Will you lose your job in a downturn? Will you be able to find another house you like as much as the one you must bid for right away? Will the Fed raise interest rates? Will an AOL-TimeWarner merger increase profits? Will it rain during your vacation to London? These questions are answered by some process of judging likelihood.

The standard principles used in economics to model probability judgment in economics are concepts of statistical sampling, and Bayes' rule for updating probabilities in the face of new evidence. Bayes' rule is unlikely to be correct descriptively because it has several features that are cognitively unrealistic. First, Bayesian updating requires a prior.[6] Second, Bayesian updating requires a separation between previously-judged probabilities and evaluations of new evidence. But many cognitive mechanisms use previous information to filter or interpret what is observed, violating this separability. For example, in perception experiments, subjects who expect to see an object in a familiar place—such as a fire hydrant on a sidewalk—perceive that object more accurately than subjects who see the same object in an unexpected place—such as on a coffeeshop counter. Third, subjective expected utility assumes separability between probability judgments of states and utilities which result from those states. Wishful thinking and other self-serving motivations violate this separation (see Babcock & Loewenstein, 1997, and this volume). Fourth, the Bayesian updating predicts no effects of the order of arrival of information. But order effects are common in memory due to the strength of recent information in working memory

[6]Because it does not specify where the prior comes from, however, it leaves room for psychological theory on the front end of the judgment process.

(recency effects), and increased "rehearsal" of older memories (primacy effects). These order effects mean that how information is sequenced distorts probability judgment (see Hogarth & Einhorn, 1992).

Cognitive psychologists have proposed heuristic mechanisms that will lead to judgments which sometimes violate either sampling principles or Bayes' rule (see Kahneman & Frederick, 2002). For example, people may judge the probabilities of future events based on how easy those events are to imagine or to retrieve from memory. This "availability heuristic" contributes to many specific further biases. One is "hindsight bias": Because events which actually occurred are easier to imagine than counterfactual events that did not, people often overestimate the probability they previously attached to events which later happened. This bias leads to "second-guessing" or Monday-morning quarterbacking and may be partly responsible for lawsuits against stockbrokers who lost money for their clients. (The clients think the brokers "should have known") A more general bias is called the "curse of knowledge"—people who know a lot find it hard to imagine how little others know. The development psychologist Jean Piaget suggested that the difficulty of teaching is caused by this curse. (Why is it so hard to explain something "obvious" like consumer indifference curves or Nash equilibrium to your undergraduate students?[7]) Anybody who has tried to learn from a computer manual has seen the curse of knowledge in action.

Another heuristic for making probability judgments is called "representativeness": People judge conditional probabilities like P(hypothesis|data) or P(example|class) by how well the data represents the hypothesis or the example represents the class. Like most heuristics, representativeness is an economical shortcut that delivers reasonable judgments with minimal cognitive effort in many cases, but sometimes goofs badly and is undisciplined by normative principles. Prototypical exemplars of a class may be judged to be more likely than they truly are (unless the prototype's extremity is part of the prototype). For example, in judging whether a certain student described in a profile is, say, a psychology major or a computer science major, people instinctively dwell on how well the profile matches the psychology or computer science

---

[7]Here is an example from the business world: When its software engineers refused to believe that everyday folks were having trouble learning to use their opaque, buggy software, Microsoft installed a test room with a one-way mirror so that the engineers could see people struggling before their very eyes (Heath, Larrick, & Klayman, 1998).

major stereotype. Many studies show how this sort of feature-matching can lead people to underweigh the "base rate" – in this example, the overall frequency of the two majors.[8]

Another byproduct of representativeness is the "law of small numbers": Small samples are though to represent the properties of the statistical process that generated them (as if the law of large numbers, which guarantees that a large sample of independent draws *does* represent the process, is in a hurry to work). If a baseball player gets hits 30% of his times at bat, but is 0 for 4 so far in a particular game, then he is "due" for a hit in his next at bat in this game, so that this game's hitting profile will more closely represent his overall ability. The so-called "gambler's fallacy", whereby people expect a tail after a coin landed heads three times in a row, is one manifestation of the law of small numbers. The flip side of the same misjudgment (so to speak) is surprise at the long streaks which result if the time series is random, which can lead people to conclude that the coin must be unfair when it isn't. Field and experimental studies with basketball shooting and betting on games show that people, including bettors, believe that there is positive autocorrelation—that players experience the "hot hand"— when there is no empirical evidence that such an effect exists (see Camerer, 1989a; Gilovich, Vallone & Tversky, 1985).

Many studies explore these heuristics and replicate their "biases" in applied domains (such as judgments of accounting auditors, consumers buying products, and students in classroom negotiations). It is important to note that a "heuristic" is both a good thing and a bad thing. A good heuristic provides fast, close to optimal, answers when time or cognitive capabilities are limited, but it also violates logical principles and leads to errors in some situations. A lively debate has emerged over whether heuristics should be called irrational if they were well-adapted to domains of everyday judgment ("ecologically rational"). In their early work, Kahneman, Tversky, and others viewed cognitive biases as the judgmental kin of speech errors ("I cossed the toin"), forgetting, and optical illusions: These are systematic errors which, even if rare, are useful for illuminating how cognitive mechanisms work. But these errors do not imply the mechanisms fail frequently or are not well-adapted for everyday use. But as Kahneman and Tversky (1982, p. 494) wrote, "Although errors of judgment are but a method by

[8]However, this "base-rate fallacy" is being thoughtfully re-examined (e.g., Koehler, 1996). The fact that base rates are more clearly included when subjects are asked what fraction of 100 hypothetical cases fit the profile is an important clue about how the heuristic operates and its limits (Gigerenzer, Hell & Blank, 1988; Tversky and Kahneman, 1983).

which some cognitive processes are studied, the method has become a significant part of the message." The shift in emphasis from the heuristics to the biases they sometimes create happened gradually as research moved to applied areas; the revisionist view that heuristics may be near-optimal is largely a critique (a reasonable one) of the later applied research.

Progress in modeling and applying behavioral models of judgment has lagged behind other areas, such as loss aversion and hyperbolic time discounting. A promising recent modeling approach is “quasi-Bayesian”—viz., assume that people misspecify a set of hypotheses, or encode new evidence incorrectly, but otherwise use Bayes’ rule. For example, Rabin and Schrag (1999) model "confirmation bias" by assuming that people who believe hypothesis A is more likely than B will never encode pro-A evidence mistakenly, but will sometimes encode pro-B evidence as being supportive of A.[9] Rabin (2002) models the "law of small numbers" in a quasi-Bayesian fashion by assuming that people mistakenly think a process generates draws from a hypothetical "urn" *without replacement*, although draws are actually independent (i.e., made *with* replacement). He shows some surprising implications of this misjudgment. For example, investors will think there is wide variation in skill of, say, mutual-fund managers, even if there is no variation at all. (A manager who does well several years in a row is a surprise if performance is mistakenly thought to be mean-reverting due to "nonreplacement", so quasi-Bayesians conclude that the manager must be *really* good.)

Barberis, Shleifer and Vishny (1998) adopt such a quasi-Bayesian approach to explain why the stock market under-reacts to information in the short-term and overreacts in the long-term. In their model, earnings follow a random walk but investors believe, mistakenly, that earnings have positive momentum in some regimes and regress toward the mean in others. After one or two periods of good earnings, the market can’t be confident that momentum exists and hence expects mean-reversion; but since earnings are really a random walk, the market is too pessimistic and is underreacting to good earnings news. After a long string of good earnings, however, the market believes momentum is building. Since it isn’t, the market is too optimistic and overreacts.

---

[9]This encoding asymmetry is related to "feature-positive" effects and perceptual encoding biases well documented in research on perception. After buying a Volvo you will suddenly "see" more Volvos on the road, due purely to heightened familiarity.

While other approaches that find ways of formalizing some of the findings of cognitive psychology are possible, our guess is that the quasi-Bayesian view will quickly become the standard way for translating the cognitive psychology of judgment into a tractable alternative to Bayes' rule. The models mentioned in the two paragraphs above are parameterized in such a way that the Bayesian model is embedded as a special case, which allows theoretical insight and empirical tests about how well the Bayesian restriction fits.

Preferences: Revealed, constructed, discovered, or learned?

Standard preference theory incorporates a number of strong and testable assumptions. For example, it assumes that preferences are "reference independent" – i.e., are not affected by the individual's transient asset position. It also assumes that preferences are invariant with respect to superficial variations in the way that options are described, and that elicited preferences do not depend on the precise way that preferences are measured as long as the method used is "incentive compatible" – i.e., provides incentives for people to reveal their "true" preferences. All of these assumptions have been violated in significant ways (see Slovic, 1995).

For example, numerous "framing effects" show that the way that choices are presented to an individual often determine the preferences that are "revealed." The classic example of a framing effect is the "Asian disease" problem in which people are informed about a disease that threatens 600 citizens and asked to choose between two undesirable options (Tversky & Kahneman, 1981). In the "positive frame" people are given a choice between (A) saving 200 lives for sure, or (B) a 1/3 chance of saving all 600 with a 2/3 chance of saving no one. In the "negative frame" people are offered a choice between (C) 400 people dying for sure, or (D) a 2/3 chance of 600 dying and a 1/3 chance of no one dying. Despite the fact that A and C, and B and D, are equivalent in terms of lives lost or at risk, most people choose A over B but D over C.

Another phenomenon that violates standard theory is called an "anchoring effect." The classic demonstration of an anchoring effect (Tversky & Kahneman, 1974, and in this volume) was identified in the context of judgment rather than choice. Subjects were shown the spin of a wheel of fortune that could range between 0 and 100 and were asked to guess whether the number of African nations in the United Nations was greater than or less than this number. They were then asked to guess the true value. Although the wheel of fortune was obviously random, subjects' guesses were strongly influenced by the spin of the wheel. As Kahneman and Tversky interpreted it, subjects seemed to "anchor" on the number spun on the wheel and then adjusted for whatever else they thought or knew, but adjusted insufficiently. Of interest in this context is that anchoring effects have also been demonstrated for choices as opposed to judgments. In one study, subjects were asked whether their certainty equivalent for a gamble was greater than or less than a number chosen at random and then were asked to specify their actual certainty equivalent for the gamble (Johnson & Schkade, 1989). Again, the stated values were correlated significantly with the random value.

In a recent study of anchoring, Ariely, Loewenstein and Prelec (in press) sold valuable consumer products (a $100 wireless keyboard, a fancy computer mouse, bottles of wine, and a luxurious box of chocolate) to postgraduate (MBA) business students. The students were presented with a product and asked whether they would buy it for a price equal to the last two digits of their own social security number (a roughly random identification number required to obtain work in the United States) converted into a dollar figure– e.g., if the last digits were 79 the hypothetical price was $79. After giving a yes/no response to the question "Would you pay $79?, subjects were asked to state the most they would pay (using a procedure that gives people an incentive to say what they really would pay). Although subjects were reminded that the Social Security number is essentially random, those with high numbers were willing to pay more for the products. For example, subjects with numbers in the bottom half of the distribution priced a bottle of wine-- a 1998 Cotes du Rhone Jaboulet Parallel '45' – at $11.62, while those with numbers in the top half priced the same bottle at $19.95.

Many studies have also shown that the *method* used to elicit preferences can have dramatic consequences, sometimes producing "preference reversals"-- situations in which A is preferred to B under one method of elicitation, but A is judged as inferior to B under a different elicitation method (e.g., Grether & Plott, 1979). The best known example contrasts how people choose between two bets versus what they separately state as their selling prices for the bets. If bet A offers a high probability of a small payoff and bet B offers a small probability of a high payoff, the standard finding is that people choose the more conservative A bet over bet B when asked to choose, but are willing to pay more for the riskier bet B when asked to price them separately. Another form of preference reversal occurs between joint and separate evaluations of pairs of goods (Hsee et al, 1999; see Hsee & LeClerc, 1998, for an application to marketing). People will often price or otherwise evaluate an item A higher than another item B when the two are evaluated independently, but evaluate B more highly than A when the two items are compared and priced at the same time.

"Context effects" refer to ways in which preferences between options depend on what other options are in the set (contrary to "independence of irrelevant alternatives" assumptions). For example, people are generally attracted to options that dominate other options (Huber, Payne & Puto, 1982). They are also drawn disproportionately to "compromise" alternatives whose attribute values lie between those of other alternatives (Simonson & Tversky, 1992).

All of the above findings suggest that preferences are not the pre-defined sets of indifference curves represented in microeconomics textbooks. They are often ill-defined, highly malleable and dependent on the context in which they are elicited. Nevertheless, when required to make an economic decisions—to choose a brand of toothpaste, a car, a job, or how to invest—people do make some kind of decision. Behavioral economists refer to the process by which people make choices with ill-defined preferences as "constructing preferences" (Payne, Bettman & Johnson, 1992; Slovic, 1995).

A theme emerging in recent research is that, although people often reveal inconsistent or arbitrary preferences, they typically obey normative principles of economic theory when it is transparent how to do so. Ariely, Loewenstein and Prelec (in press) refer to this pattern as "coherent arbitrariness" and illustrate the phenomenon with a series of studies in which the amount subjects demanded to listen to an annoying sound is sensitive to an arbitrary anchor, but they also demand much more to listen to the tone for a longer period of time. Thus, while expressed valuations for one unit of a good are sensitive to an anchor which is clearly arbitrary, subjects also obey the normative principle of adjusting those valuations to the quantity – in this case the duration -- of the annoying sound.

Most evidence that preferences are constructed comes from demonstrations that some feature that should not matter actually does. The way gambles are "framed" as gains and losses from a reference outcome, the composition of a choice set, and whether people choose among objects or value them separately, have all been shown to make a difference in expressed preference. But admittedly, a list of a theory's failings is not an alternative theory. So far, a parsimonious alternative theory has not emerged to deal with all of these challenges to utility maximization.[10]

Overview of the Book

In what follows, we review different topic areas of behavioral economics to place chapters of the book into context. The book is organized so that early chapters discuss basic topics such as decision making under risk and intertemporal choice, while later chapters provide applications of these ideas.

Reference-dependence and loss aversion

In classical consumer theory, preferences among different commodity bundles are assumed to be invariant with respect to an individual's current endowment or consumption. Contrary to this simplifying assumption, diverse forms of evidence point to a dependence of preferences on one's reference point (typically the current endowment). Specifically, people seem to dislike losing commodities from their consumption bundle much more than they like gaining other commodities. This can be expressed graphically as a kink in indifference curves at the current endowment point (Knetsch, 1992; Tversky & Kahneman, 1991).

In the simplest study showing reference-dependence, Knetsch (1992) endowed some subjects randomly with a mug, while others received a pen.[11] Both groups were allowed to switch their good for the other at a minimal transaction cost, by merely handing it to the experimenter. If preferences are independent of random endowments, the fractions of subjects swapping their mug for a pen and the fraction swapping their pen for a mug should add to roughly one. In fact, 22% of subjects traded. The fact that so few chose to trade implies an exaggerated preference for the good in their endowment, or a distaste for losing what they have.

A seminal demonstration of an "endowment effect" in buying and selling prices was conducted by Kahneman et al (1990). They endowed half of the subjects in a group with coffee mugs. Those who had mugs were asked the lowest price at which they would sell. Those who did

---

[10]Some specialized models have been proposed to explain particular phenomena, such as Hsee, Loewenstein, Blount & Bazerman, 1999; Prelec, Wernerfelt & Zettelmeyer, 1997; Tversky, Slovic & Kahneman, 1990.

[11]Note that any possible information value from being given one good rather than the other is minimized because the endowments are random, and subjects knew that half the others received the good they didn't have.

not get mugs were asked how much they would pay. There should be essentially no difference between selling and buying prices. In fact, the median selling price was $5.79 and the median buying price was $2.25, a ratio of more than two: one which has been repeatedly replicated. Although calibrationally entirely implausible, some economists were concerned that the results could be driven by “wealth effects”—those given mugs are wealthier than those not given mugs, and this might make them value mugs more and money less. But in a different study reported in the same paper, the selling prices of one group were compared to the "choosing" prices of another: For a series of money amounts, subjects chose whether they would prefer to have a mug or money. The median choosing price was half the median selling price ($3.50 versus $7.00). Choosers are in *precisely* the same wealth position as sellers—they choose between a mug or money. The only difference is that sellers are "giving up" a mug they "own," whereas choosers are merely giving up the right to have a mug. Any difference between the two groups cannot be attributed to wealth effects.

Kahneman et al's work was motivated in part by survey evidence from "contingent valuation" studies that attempt to establish the dollar value of goods which are not routinely traded. Contingent valuation is often used to do government cost-benefit analysis or establish legal penalties from environmental damage. These surveys typically show very large differences between buying prices (e.g., paying to clean up oily beaches) and selling prices (e.g., having to be paid to allow beaches to be ruined). Sayman and Öncüler (1997) summarize 73 data sets which show selling-to-buying ratios ranging from .67 (for raspberry juice) to 20 or higher (for density of trees in a park and health risks).

Loss aversion has already proved to be a useful phenomenon for making sense of field data (see Camerer, 2000, and this volume). Asymmetries in demand elasticities after price increases and decreases (Hardie, Johnson, & Fader, 1993), the tendency for New York City cab drivers to quit early after reaching a daily income target (producing surprising upward-sloping labor supply curves; see Camerer et al, 1997 and in this volume), and the large gap between stock and bond returns (the "equity premium"; see Benartzi & Thaler, 1995, in this volume) can all be explained by models in which agents have reference-dependent preferences and take a short planning horizon, so that losses are not integrated against past or future gains.

A particularly conclusive field study by Genoseve and Mayer (2001, and this volume) focuses on the real estate market. (Housing is a huge market—worth $10 trillion at the time of their study, a quarter of the wealth in the US—and full of interesting opportunities to do behavioral economics.) They find that list prices for condominiums in Boston are strongly affected by the price at which the condominium was purchased. Motivated sellers should, of course, regard the price they paid as a sunk cost and choose a list price that anticipates what the market will pay. But people hate selling their houses at a nominal loss from the purchase price. Sellers' listing prices and subsequent selling behavior reflects this aversion to nominal losses. Odean (1998) finds the same effect of previous purchase price in stock sales.[12]

At least three features of endowment effects remain open to empirical discussion. First, do people *anticipate* the endowment effect? The answer seems to be "No": Loewenstein and Adler (1995) found that subjects did not anticipate how much their selling prices would increase after they were endowed with mugs.[13] Van Boven, Dunning and Loewenstein (2000) and Van Boven, Loewensstein and Dunning (2000) found that agents for buyers also underestimated how much sellers would demand.

Second, Kahneman, Knetsch and Thaler (1990:1328) note that "there are some cases in which no endowment effect would be expected, such as when goods are purchased for resale rather than for utilization." However, the boundary of commercial non-attachment has not been carefully mapped. Do art or antique dealers "fall in love" with pieces they buy to resell? What about surrogate mothers who agree to bear a child for a price paid in advance? Evidence on the degree of commercial attachment is mixed. In their housing study, Genesove and Mayer (2001 and this volume) note that investors who don't live in their condos exhibit less loss-aversion than owners. A field experiment by List (in press) found that amateur sports paraphernalia collectors who do not trade very often showed an endowment effect, but professional dealers and amateurs who trade a lot did not.[14] An example where attachment seemed important even among

---

[12]Though it is harder to unambiguously interpret as loss aversion in the sense we are discussing here, reference points can also serve as social focal points for judging performance. Degeorge, Patel & Zeckhauser (1999) document an interesting example from corporate finance. Managers whose firms face possible losses (or declines from a previous year's earnings) are very reluctant to report small losses. As a result, the distribution of actual losses and gains shows a very large spike at zero, and hardly any small reported losses (compared to the number of small gains). Wall Street hates to see a small loss. A manager who does not have the skill to shift accounting profits to erase a potential loss (i.e., "has some earnings in his pocket") is considered a poor manager. In this example, the market's aversion to reported losses can serve as a signaling device which tells the markets about managerial ability.

[13]Failure to anticipate the strength of later loss-aversion is one kind of "projection bias" (Loewenstein, O'Donoghue & Rabin, 1999), in which agents are make choices as if their current preferences or emotions will last longer than they actually do.

[14]By revisiting the same traders a year later, List showed that it was trader experience which reduced endowment effects, rather than self-selection (i.e., people who are immune to such effects become dealers.)

experienced traders with high incentives was described by an investment banker who said his firm combats loss-aversion by forcing a trader to periodically switch his "position" (the portfolio of assets the trader bought and is blamed or credited for) with the position of another trader. Switching ensures that traders do not make bad trades because of loss-aversion and emotional attachment to their past actions (while keeping the firm's net position unchanged, since the firm's total position is unchanged).

Third, it is not clear the degree to which endowment effects are based solely on the current endowment, rather than past endowments or other reference points. Other reference points, such as social comparison (i.e., the possessions and attainments of other people) and past ownership, may be used to evaluate outcomes. How multiple reference points are integrated is an open question. Strahilevitz and Loewenstein (1998) found that the valuation of objects depended not only on whether an individual was currently endowed with an object, but on the entire past history of ownership – how long the object had been owned or, if it had been lost in the past, how long ago it was lost and how long it was owned before it was lost. These "history-of-ownership effects" were sufficiently strong that choice prices of people who had owned for a long period but just lost an object were higher than the selling prices of people who had just acquired the same object.

If people are sensitive to gains and losses from reference points, the way in which they combine different outcomes can make a big difference. For example, a gain of $150 and a loss of $100 will seem unattractive if they are evaluated separately, if the utility of gains is sufficiently less than the disutility of equal-sized losses, but the gain of $50 that results when the two figures are added up is obviously attractive. Thaler (1980, 1999 and this volume) suggests that a useful metaphor for describing the rules which govern gain/loss integration is "mental accounting"—people set up mental accounts for outcomes which are psychologically separate, much as financial accountants lump expenses and revenues into separated accounts to guide managerial attention. Mental accounting stands in opposition to the standard view in economics that "money is fungible"; it predicts, accurately, that people will spend money coming from different sources in different ways (O'Curry, 1999), and has wide-ranging implications for such policy issues as how to promote saving (see, e.g., Thaler, 1994).

A generalization of the notion of mental accounting is the concept of "choice bracketing," which refers to the fashion in which people make decisions narrowly, in either a piece-meal fashion, or broadly – i.e., taking account of interdependencies between decisions (Read, Loewenstein & Rabin, 1999). How people bracket choices has far-reaching consequences in diverse areas, including finance (see Bernartzi & Thaler, 1995 and chapter 22), labor supply (Camerer, Babcock, Loewenstein & Thaler, 1997, and chapter 19), and intertemporal choice (Frederick, Loewenstein & O'Donoghue, in press and chapter 6, section 5.3.4). For example, when making many separate choices between goods, people tend to choose more diversity when the choices are bracketed broadly than when they are bracketed narrowly. This was first demonstrated by Simonson (1990), who gave students their choice of one of six snacks during each of three successive weekly class meetings. Some students chose all three snacks in the first week, although they didn't receive their chosen snack until the appointed time, and others chose each snack on the day that they were to receive it (narrow bracketing; sequential choice). Under broad bracketing, fully 64% chose a different snack for each week, as opposed to only 9% under narrow bracketing. Follow-up studies demonstrated similar phenomena in the field (e.g., in purchases of yogurt; Simonson & Winer, 1992).

Bracketing also has implications for risk-taking. When people face repeated risk decisions, evaluating those decisions in combination can make them appear less risky than if they are evaluated one at a time. Consequently, a decision maker who refuses a single gamble may nonetheless accept two or more identical ones. By assuming that people care only about their overall level of wealth, expected utility theory implicitly assumes broad bracketing of risky decisions. However, Rabin (2000) points out the absurd implication which follows from this assumption (combined with the assumption that risk aversion stems from the curvature of the utility function): A reasonable amount of aversion toward risk in small gambles implies a dramatic aversion to reduction in overall wealth. For example, a person who will turn down a coin flip to win $11 and lose $10 at all wealth levels must also turn down a coin flip in which she can lose $100, *no matter how large the possible gain is*.[15] Rabin's proof is a mathematical

[15]The intuition behind Rabin's striking result is this: In expected-utility theory, rejecting a (+$11,-$10) coin flip at wealth level W implies that the utility increase from the $11 gain is smaller than the total utility decrease from the $10 loss, meaning that the marginal utility of each dollar gained is at most 10/11 of the marginal utility of each dollar lost. By concavity, this means that the marginal utility of the W+11th dollar is at most 10/11 the marginal utility of the W-10th dollar—a sharp 10% drop in marginal utility for small change in overall wealth of $21. When the curvature of the utility function does not change unrealistically over ranges of wealth levels, this means the marginal utility plummets quickly as wealth increases—the marginal utility of the W+$32 dollar (= W+11 + 21) can be at most (10/11)(10/11), which is around 5/6 of the marginal utility of the W-10th dollar. Every $21 decrease in wealth yields another 10% decline in marginal utility. This implies, mathematically, that implying a person's value for a dollar if he were $500 or $1,000 wealthier would be tiny compared to how much he values dollars he might lose in a

demonstration that people who are averse to small risks are probably not integrating all their wealth into one source when they think about small gambles.

Preferences over risky and uncertain outcomes

The expected-utility EU hypothesis posits that the utility of a risky distribution of outcomes (say, monetary payoffs) is a probability-weighted average of the outcome utilities. This hypothesis is normatively appealing because it follows logically from apparently reasonable axioms, most notably the independence (or "cancellation") axiom. The independence axiom says that if you are comparing two gambles, you should cancel events which lead to the same consequence with the same probability; your choice should be independent of those equally-likely common consequences. Expected utility also simplifies matters because a person's taste for risky money distributions can be fully captured by the shape of the utility function for money.

Many studies document predictive failures of expected utility in simple situations in which subjects can earn substantial sums of money from their choices.[16] Starmer's (2000) contribution to this volume reviews most of these studies, as well as the many theories that have been proposed to account for the evidence (see also Camerer, 1989b, 1992; Hey, 1997; Quiggin, 1993). Some of these new theories alter the way in which probabilities are weighted, but preserve a "betweenness" property which says that if A is preferred to B, then any probabilistic gamble between them must be preferred to B but dispreferred to A (i.e., the gambles lie "between" A and B in preference). Other new theories suggest that probability weights are "rank-dependent"—outcomes are first ranked, then their probabilities are weighted in a way which is sensitive to how they rank within the gamble that is being considered. One mathematical way to do this is transform the *cumulative* probabilities of outcomes (i.e., the chance that you will win X or less) nonlinearly and weight outcome utilities by the differences of

---

bet. So if a person's attitude towards gambles really came from the utility-of-wealth function, even incredibly large gains in wealth would not tempt her to risk \$50 or \$100 losses, if she really dislikes losing \$10 more than she likes gaining \$11 at every level of wealth.

[16]Some of the earlier studies were done with hypothetical payoffs, leading to speculation that the rejection of EU would not persist with real stakes. Dozens of recent studies show that, in fact, paying real money instead of making outcomes hypothetical appears either fails to eliminate EU rejections, or *strengthens* the rejections of EU (because sharper results which come from greater incentive imply that rejections are more statistically significant; Harless & Camerer, 1994).

those weighted cumulative probabilities.[17] The best known theory of this sort is cumulative prospect theory (Tversky & Kahneman, 1992).

There are three clear conclusions from the experimental research (Harless & Camerer, 1994). One is that of the two new classes of theories that allow more general functional forms than expected utility, the new rank-dependent theories fit the data better than the new betweenness class theories. A second conclusion is that the statistical evidence against EU is so overwhelming that it is pointless to run more studies testing EU against alternative theories (as opposed to comparing theories with one-another). The third conclusion is that EU fits worst when the two gambles being compared have different sets of possible outcomes (or "support"). Technically, this property occurs when one gamble has a unique outcome. The fact that EU does most poorly for these comparisons implies that nonlinear weighting of low probabilities is probably a major source of EU violations. Put differently, EU is like Newtonian mechanics, which is useful for objects traveling at low velocities but mispredicts at high speeds. Linear probability weighting in EU works reasonably well except when outcome probabilities are very low or high. But low-probability events are important in the economy, in the form of “gambles” with positive skewness (lottery tickets, and also risky business ventures in biotech and pharmaceuticals), and catastrophic events which require large insurance industries.

Prospect theory (Kahneman & Tversky, 1979) explains experimental choices more accurately than EU because it gets the psychophysics of judgment and choice right. It consists of two main components: a probability weighting function, and a 'value function' which replaces the utility function of EU. The weighting function $\pi(p)$ combines two elements: (1) The level of probability weight is a way of expressing risk tastes (if you hate to gamble, you place low weight on any chance of winning anything); and (2) the curvature in $\pi(p)$ captures how sensitive people are to differences in probabilities. If people are more sensitive in the neighborhoods of possibility and certainty—i.e., changes in probability near zero and 1— than to intermediate gradations, then their $\pi(p)$ curve will overweight low probabilities and underweight high ones.

---

[17]A technical motivation for “rank dependent” theories-- ranking outcomes, then weighting their probabilities-- is that when separate probabilities are weighted, it is easy to construct examples in which people will violate dominance by choosing a “dominated” gamble A which has a lower chance of winning at each possible outcome amount, compared to the higher chance of winning the same outcome amount for a dominant gamble B. If people rarely choose such dominated gambles, they are acting as if they are weighting the *differences* in cumulated probabilities, which is the essence of the rank-dependent approaches.

The value function reflects the insight, first articulated by Markowitz (1952), that the utility of an outcome depends not on the absolute level of wealth that results but on whether the outcome is a gain or a loss. Prospect theory also assumes reflection of risk-preferences at the reference point: People are typically averse to risky spreading of possible money gains, but will take gambles where they could lose big or break even rather than accept a sure loss. Prospect theory also assumes "loss-aversion": The disutility of a loss of x is worse than the utility of an equal-sized gain of x.

Expected utility is restricted to gambles with known outcome probabilities. The more typical situation in the world is "uncertainty", or unknown (subjective, or personal) probability. Savage (1954) proposed a subjective expected utility (SEU) theory in which choices over gambles would reveal subjective probabilities of states, as well as utilities for outcomes. Ellsberg (1961) quickly pointed out that in Savage's framework, subjective probabilities are slaves to two masters-- they are used as decision weights applied to utilities, and they are expressions of likelihood. As a result, there is no way to express the possibility that, because a situation may have lots of "ambiguity," one is reluctant to put much decision weight on *any* outcome. Ellsberg demonstrated this problem in his famous paradox: Many people prefer to bet on black drawn from an urn with 50 black and 50 red balls, rather than bet on black drawn from an urn with 100 balls of unknown black and red composition, and similarly for red (they just don't want to bet on the unknown urn). There is no way for the two sets of red and black subjective probabilities from each urn to both add to one (as subjective probabilities require), and still express the distaste for betting neither color in the face of ambiguity.

Many theories have been proposed to generalize SEU to allow for ambiguity-aversion (see Camerer & Weber, 1992, for a review). One approach, first proposed by Ellsberg, is to let probabilities be <u>sets</u> rather than specific numbers, and assume that choices over gambles reveal whether people pessimistically believe the worst probabilities are the right ones, or the opposite. Another approach is to assume that decision weights are nonadditive. For example, the weights on red and black in the Ellsberg unknown urn could both be .4; the missing weight of .2 is a kind of "reserved belief" which expresses how much the person dislikes betting when she knows that important information is missing.

Compared to non-EU theories, relatively little empirical work and applications have been done with these uncertainty-aversion theories so far. Uncertainty-aversion might explain phenomena like voting "rolloff" (when a voter, once in the voting booth, refuses to vote on obscure elections in which their vote is most likely to prove pivotal; Ghirardato & Katz, 2000), incomplete contracts (Mukherji, 1998) and "home country bias" in investing: People in every country overinvest in the country they are most familiar with-- their own. (Finnish people invest in firms closer to their own town, see Grinblatt & Keloharju, 2001.)

In asset pricing, ambiguity-aversion can imply that asset prices satisfy a pair of Euler inequalities, rather than an Euler equation, which permits asset prices to be more volatile than in standard theory (Epstein & Wang, 1994). Hansen, Sargent and Tallarini (1999) have applied related concepts of "robust control" to macroeconomic fluctuations. Finally, uncertainty-averse agents will value information even if it does not change the decisions they are likely to make after becoming better-informed (simply because information can make nonadditive decision weights closer to additive, and make agents "feel better" about their decision). This effect may explain demand for information in settings like medicine or personal finance, where new information usually does not change choices, but relieves anxiety people have from knowing there is something they could know but do not (Asch, Patton and Hershey, 1990).

Intertemporal choice

The discounted-utility (DU) model assumes that people have instantaneous utilities from their experiences each moment, and that they choose options which maximize the present discounted sum of these instantaneous utilities. Typically it is assumed that the instantaneous utility each period depends solely on consumption in that period, and that the utilities from streams of consumption are discounted exponentially, applying the same discount rate in each period. Samuelson (1937) proposed this particular functional form because it was simple and similar to present value calculations applicable to financial flows. But in the article in which he proposed the DU model, he repeatedly drew attention to its psychological implausibility.[18] Decades of empirical research substantiated his doubts (see Loewenstein & Prelec, 1992, and Frederick, Loewenstein & O'Donoghue, in press and in this volume).

---

[18] The notion of discounting utility at a fixed rate was first mentioned, in passing, in an article on intergenerational saving by Ramsey (1928).

It is useful to separate studies dealing with intertemporal choice into those that focus on phenomena that can be explained on the basis of the discount function and those that can be explained on the basis of the utility function. The following two subsections cover these points.

Time Discounting

A central issue in economics is how agents trade off costs and benefits that occur at different points in time. The standard assumption is that people weight future utilities by an exponentially-declining discount factor $d(t) = \delta^t$, where $1>\delta > 0$. Note that the discount factor δ is often expressed as $1/(1+r)$, where r is a discount **rate**.

However, a simple hyperbolic time discounting function of $d(t) = 1/(1+kt)$ tends to fit experimental data better than exponential discounting. The early evidence on discounting came from studies showing that animals exhibit much large discounting when comparing immediate rewards and rewards delayed t periods, compared to the tradeoff between rewards k and k+t periods in the future. Thaler (1981) was the first to empirically test the constancy of discounting with human subjects. He told subjects to imagine that they had won some money in a lottery held by their bank. They could take the money now or earn interest and wait until later. They were asked how much they would require to make waiting just as attractive as getting the money immediately. Thaler then estimated implicit (per-period) discount rates for different money amounts and time delays under the assumption that subjects had linear utility functions. Discount rates declined linearly with the duration of the time delay. Later studies replicated the basic finding that discount rates fall with duration (e.g., Benzion, Rapoport & Yagil, 1989; Holcomb & Nelson, 1992). The most striking effect is an "immediacy effect" (Prelec & Loewenstein, 1991); discounting is dramatic when one delays consumption that would otherwise be immediate.

Declining discount rates have also been observed in experimental studies involving real money outcomes. Horowitz tested the constancy of discounting by auctioning "bonds" in a Vickrey (highest-rejected-bid) auction. The amount bid for a bond represented how much a subject was willing to give up at the time of the auction for certain future payoffs. Discount rates again decreased as the horizon grew longer. Pender (1996) conducted a study in which Indian farmers made several choices between amounts of rice that would be delivered either sooner or later. Fixing the earlier rice ration and varying the amount of rice delivered later gives an estimate of the discount rate. To avoid immediacy effects, none of the choices were delivered immediately. Per-period discount rates decline with the increasing horizon: the mean estimated discount rate was .46 for 7 months and .33 for 5 years.

Hyperbolic time discounting implies that people will make relatively far-sighted decisions when planning in advance – when all costs and benefits will occur in the future – but will make relatively short-sighted decisions when some costs or benefits are immediate. The systematic changes in decisions produced by hyperbolic time discounting create a time-inconsistency in intertemporal choice not present in the exponential model. An agent who discounts utilities exponentially would, if faced with the same choice and the same information, make the same decision prospectively as he would when the time for a decision actually arrives. In contrast, somebody with time-inconsistent hyperbolic discounting will wish prospectively that in the future he will take far-sighted actions; but when the future arrives he will behave against his earlier wishes, pursuing immediate gratification rather than long-run well-being.

Strotz (1955) first recognized the planning problem for economic agents who would like to behave in an intertemporally consistent fashion, and discussed the important ramifications of hyperbolic time discounting for intertemporal choice. Most big decisions – e.g., savings, educational investments, labor supply, health and diet, crime and drug use-- have costs and benefits which occur at different points in time. Many authors such as Thaler (1981), Thaler and Shefrin (1981), and Schelling (1978) discussed the issues of self control and stressed their importance for economics. Laibson (1997) accelerated the incorporation of these issues into economics by adopting a "quasi-hyperbolic" time discounting function (first proposed by Phelps and Pollak (1968) to model intergenerational utility). The quasi-hyperbolic form approximates the hyperbolic function with two-parameters, β and δ, in which the weight on current utility is 1 and the weight on period-t instantaneous utility is $\beta\delta^t$ for $t>0$. The parameter β measures the immediacy effect — if $\beta = 1$ the model reduces to standard exponential discounting. When delayed rewards are being compared, the immediacy premium β divides out so the ratio of discounted utilities is solely determined by δ^t (consistent with the observations of Benzion, Rapoport & Yagil, 1989)

Thus, quasi-hyperbolic time discounting is basically standard exponential time discounting plus an immediacy effect; a person discounts delays in gratification equally at all moments except the current one—caring differently about well-being now versus later. This functional form provides one simple and powerful model of the taste for immediate gratification.

In his paper reprinted in this volume, Laibson (1997 and chapter 15) applies the quasi-hyperbolic model to a model of lifetime consumption-savings decisions. He emphasizes the role that the partial illiquidity of an asset plays in helping consumers constrain their own future consumption. If people can withdraw money immediately from their assets, as they can with simple savings or checking accounts, they have no way to keep their temptation to over-consume under control. Assets that are less liquid, despite their costly lack of flexibility or even lower yield, may be used as a commitment device for those consumers who at least partially understand their tendency to over-consume. In this paper (and in more recent papers by Laibson, Repetto and Tobacman (1998)) and others it has been demonstrated how quasi-hyperbolic discounting potentially provides a better account than does conventional exponential discounting of various savings and consumption phenomena, such as different marginal propensities to consume out of different forms of savings, and the ways that financial innovation (typically in the form of increased liquidity) may lead to damaging decreases in savings.

An important question in modelling self-control is whether agents are aware of their self-control problem ("sophisticated") or not ("naï ve"). The work in macroeconomics described above assumes agents are sophisticated, but have some commitment technologies to limit how much the current self can keep the future self from overspending.[19] However, there are certainly many times in which people are partially unaware of their own future misbehavior, and hence overly optimistic that they will behave in the future the way that their "current self" would like them to. O'Donoghue and Rabin (1999 and this volume; cf. Akerlof, 1991) show how awareness of self-control problems can powerfully moderate the behavioral consequences of quasi-hyperbolic discounting.

Naivete typically makes damage from poor self-control worse. For example, severe procrastination is a creation of over-optimism: One can put off doing a task repeatedly if the perceived costs of delay are small—"I'll do it tomorrow, so there is little loss from not doing it today"—and hence accumulate huge delay costs from postponing the task many times. A sophisticated agent aware of his procrastination will realize that if they put if off they will put if off in the future, and hence will do the task immediately. However, in some cases, being sophisticated about one's self-control problem can *exacerbate* yielding to temptation. If you are

[19]Ariely and Wertenbroch (in press) report similar self-commitment—deadline-setting—in an experiment.

aware of your tendency to yield to a temptation in the future, you may conclude that you might as well yield now; if you naively think you will resist temptation for longer in the future, that may motivate you to think it is worthwhile resisting temptation now. More recently, O'Donoghue and Rabin (2001) have developed a model of "partial naivete" that permits a whole continuum of degree of awareness, and many other papers on quasi-hyperbolic discounting have begun to clarify which results come from the quasi-hyperbolic preferences *per se* and which come from assumptions about self-awareness of those preferences.

Many of the most striking ways in which the classical DU model appears to fail stem not from time discounting, but from characteristics of the utility function. Numerous survey studies (Benzion et al. 1989; Loewenstein, 1988; Thaler 1981) have shown that gains and losses of different absolute magnitudes are discounted differently. Thaler's (1981) subjects were indifferent between $15 immediately and $60 in a year (a ratio of .25) and between $250 immediately and $350 in a year (a ratio of .71). Loewenstein and Prelec (1992) replicated these "magnitude effects," and also show that estimated discount rates for losses tend to be lower than those for gains. Again, these effects are inconsistent with DU. A third anomaly is that people dislike "temporal losses" – delays in consumption -- much more than they like speeding up consumption (Loewenstein, 1988).

None of these effects can be explained by DU, but they are consistent with a model proposed by Loewenstein and Prelec (1992). This model departs from DU in two major ways. First, as discussed in the previous subsection, it incorporates a hyperbolic discount function. Second, it incorporates a utility function with special curvature properties that is defined over gains and losses rather than final levels of consumption. Most analyses of intertemporal choice assume that people integrate new consumption with planned consumption. While such integration is normatively appealing, it is computationally infeasible and, perhaps for this reason, descriptively inaccurate. When people make decisions about new sequences of payments or consumption, they tend to evaluate them in isolation – e.g., treating negative outcomes as losses, rather than as reductions to their existing money flows or consumption plans. No model that assumes integration can explain the anomalies just discussed.

The anomalies just discussed are sometimes mislabeled as discounting effects. It is said that people "discount" small outcomes more than large, gains more than losses, and that they exhibit greater time discounting for delay than for speed-up. Such statements are misleading. In fact, all of these effects are consistent with stable, uniform, time discounting once one measures discount rates with a more realistic utility function. The inconsistencies arise from misspecification of the utility function, not from differential time discounting of different types of outcomes.

A second anomaly is apparent *negative* time discounting. If people like savoring pleasant future activities they may postpone them to prolong the pleasure (and they may get painful activities over with quickly to avoid dread). For example, Loewenstein (1987) elicited money valuations of several outcomes which included a "kiss from the movie star of your choice," and "a nonlethal 110 volt electric shock" occurring at different points in time. The average subject paid the most to delay the kiss three days, and was eager to get the shock over with as quickly as possible (see also Carson and Horowitz, 1990; MacKeigan et al, 1993). In a standard DU model, these patterns can only be explained by discount factors which are greater than one (or discount **rates** which are negative). However, Loewenstein (1987) showed that these effects can be explained by a model with positive time discounting, in which people derive utility (both positive and negative) from anticipation of future consumption.

A closely related set of anomalies involves sequences of outcomes. Until recently, most experimental research on intertemporal choice involved single outcomes received at a single point in time. The focus was on measuring the correct form of the discount function and it was assumed that once this was determined the value of a sequence of outcomes could be determined by simply adding up the present values of its component parts. The sign and magnitude effects and the delay/speed-up asymmetry focused attention on the form of the utility function that applies to intertemporal choice, but retained the assumption of additivity across periods. Because they only involved single outcomes, these phenomena shed no light on the validity of the various assumptions that involve multiple time periods, and specifically about the different independence assumptions.

Research conducted during the past decade, however, has begun to examine preferences toward sequences of outcomes and has found quite consistently that preferences for sequences do not follow in a simple fashion from preferences for their component parts (Loewenstein & Prelec, 1993). People care about the “gestalt” or overall pattern of a sequence, in a way that violates independence.

A number of recent studies have shown that people generally favor sequences that improve over time. Loewenstein and Sicherman (1991) and Frank and Hutchens (1993 and this volume), for example, found that a majority of subjects prefer an increasing wage profile to a declining or flat one, for an otherwise identical job. Preference for improvement appears to be driven in part by savoring and dread (Loewenstein, 1987), and in part by adaptation and loss aversion. Savoring and dread contribute to preference for improvement because, for gains, improving sequences allow decision makers to savor the best outcome until the end of the sequence. With losses, getting undesirable outcomes over with quickly eliminates dread. Adaptation leads to a preference for improving sequences because people tend to adapt to ongoing stimuli over time, and to evaluate new stimuli relative to their adaptation level (Helson, 1964), which means that people are sensitive to *change*. Adaptation favors increasing sequences, which provide a series of positive changes – i.e., *gains*, over decreasing sequences, which provide a series of negative changes – i.e., *losses*. Loss aversion (Kahneman & Tversky, 1979) intensifies the preference for improvement over deterioration.

The idea that adaptation and loss aversion contribute to the preference for sequences, over and above the effects of savoring and dread, was suggested by a study conducted by Loewenstein and Prelec (1993). They asked subjects to first state a preference between a fancy French restaurant dinner for two either on Saturday in one month or Saturday in two months. Eighty percent preferred the more immediate dinner. Later the same respondents were asked whether they would prefer the sequence fancy French this month, mediocre Greek next month or mediocre Greek this month and fancy French next month. When the choice was expressed as one between sequences, a majority of respondents shifted in favor of preferring the improving sequence – which delayed the French dinner for two months. The same pattern was observed when the mediocre Greek restaurant was replaced by "eat at home," making it even more transparent that the sequence frame was truly changing people’s preferences. The conclusion of this research is that, as in visual perception, people have a "gestalt" notion of an ideal distribution of outcomes in time, which includes interactions across time periods that violate simple separability axioms.

Fairness and social preferences

The assumption that people maximize their own wealth and other personal material goals (hereafter, just "self-interest") is a widely correct simplification that is often useful in economics. However, people may sometimes choose to "spend" their wealth to punish others who have harmed them, reward those who have helped, or to make outcomes fairer. Just as understanding demand for goods requires specific utility functions, the key to understanding this sort of social preferences is a parsimonious specification of "social utility" which can explain many types of data with a single function.

An experimental game which has proved to be a useful workhorse for identifying departures from self interest is the "ultimatum" game, first studied by Güth et al. (1982). In an ultimatum game, a Proposer has an amount of money, typically about $10, from which he must propose a division between himself and a Responder. (The players are anonymous and will never see each other again.) If the Responder accepts the offered split, they both get paid and the game ends. If she rejects the offer they get nothing and the game ends. In studies in more than 20 countries, the vast majority of Proposers offer between a third and a half of the total, and Responders reject offers of less than a fifth of the total about half the time. A responder who rejects an offer is spending money to punish somebody who has behaved unfairly.

A "trust" game can be used to explore the opposite pattern, "positive reciprocity." Positive reciprocity means that players are disposed to reward those who have helped them, even at a cost to themselves. In a typical trust game, one player has a pot of money, again typically around $10, from which he can choose to keep some amount for himself, and to invest the remaining amount X, between $0 and $10, and their investment is tripled. A trustee then takes the amount 3X, keeps as much as she wants, and returns Y. In standard theory terms, the investor-trustee contract is incomplete and the investor should fear trustee moral hazard. Self-interested trustees will keep everything (Y = 0) and self-interested investors who anticipate this will invest nothing (X = 0). In fact, in most experiments investors invest about half and trustees pay back a little less than the investment. Y varies positively with X, as if trustees feel an obligation to repay trust.

The first attempt to model these sorts of patterns was Rabin (1993, and this volume). Fixing Player A's likely choice, Player B's choice determines A's payoff. From A's point of view, B's choice can be either kind (gives A a lot) or mean (gives A very little). This enables A to form a numerical judgment about B's kindness, which is either negative or positive (zero represents kindness-neutrality). Similarly, A's action is either kind or mean toward B. In Rabin's approach, people earn a utility from the payoff in the game and a utility from the product of their kindness and the kindness of the other player. Multiplying the two kindness terms generates both negative and positive reciprocity, or a desire for emotional coordination: If B is positively kind, A *prefers* to be kind too; but if B is mean (negative kindness), then A prefers to be mean. Rabin then uses concepts from game-theory to derived consequences for equilibrium, assuming people have fairness-adjusted utilities.[20]

Besides explaining some classic findings, Rabin's kindness-product approach makes fresh predictions: For example, in a prisoner's dilemma (PD), mutual cooperation can be a "fairness equilibrium." (Cooperating is nice; therefore, reciprocating anticipated cooperation is mutually nice and hence utility-maximizing.) But if player A is *forced* to cooperate, then player A is not being kind and player B feels no need to behave kindly. So players B should defect in the "involuntary" PD.

Other approaches posit a social utility function which combines one's own payoff with their relative share of earnings, or the difference between their payoffs and the payoffs of others. One example is Fehr and Schmidt (1999, and this book), who use the function $u_i(x_1, x_{2,\ldots}x_n)= x_i-\alpha\ \Sigma_k[x_k-x_i]_0/(n-1)-\beta\Sigma_k[x_i-x_k]_0/(n-1)$, where $[x]_0$ is x if $x > 0$ and 0 otherwise. The coefficient α is the weight on envy or disadvantageous inequality (when $x_k > x_i$) and β is the weight on guilt or advantageous inequality ($x_i>x_k$). This inequality-aversion approach matches ultimatum rejections because an offer of $2 from a $10 pie, say, has utility $2-(8-2)\alpha$ while rejecting yields 0. Players who are sufficiently envious ($\alpha>1/3$) will reject such offers. Inequality-aversion also mimics the effect of positive reciprocity because players with positive values of βwill feel sheepish about earning more money than others do; so they will repay trust and feel bad about defecting in PDs and free-riding in public goods contribution games. Bolton and Oeckenfels (2000) propose a similar model.

Charness and Rabin (forthcoming) propose a "Rawlsitarian" model which integrates three factors—one's own payoff, and a weighted average of the lowest payoff anyone gets (a la Rawls) and the sum of everyone's payoff (utilitarian). This utility function explains new results from

[20]He used the theory of psychological games, in which a player's utilities for outcomes can depend on their **beliefs** (Geanakopolos, Pearce & Stacchetti, 1989). (For example, a person may take pleasure in being surprised by receiving a gift, aside from the gift's direct utility.)

three-person games which are not explained by the inequality-aversion forms, and from a large sample of two-person games where the inequality-aversion approaches often predict poorly.

The key point is that careful experimental study of simple games in which social preferences play a role (like ultimatum and trust) has yielded tremendous regularity. The regularity has, in turn, inspired different theories that map payoffs to all players into each player's utility, in a parsimonious way. Several recent papers compare the predictions of different models (see Camerer, 2002, chapter 2). The results show that some form of the intentionality incorporated in Rabin (1993, and this volume) (players care about whether another player *meant* to harm them or help them), combined with inequality aversion or Rawlsitarian mixing will explain a lot of data. Models like these also make new predictions and should be useful in microeconomics applications as well.

Kahneman, Knetsch and Thaler (1986 and this volume) studied consumer perceptions of fairness using phone surveys. They asked people about how fair they considered different types of firm behavior to be. In a typical question, they asked people whether a hardware store that raised the price of a snow shovel after a snowstorm was behaving fairly or not. (People thought the store was unfair.) Their results can be neatly summarized by a "dual-entitlement" hypothesis: Previous transactions establish a reference level of consumer surplus and producer profit. Both sides are "entitled" to these levels of profit, so price changes which threaten the entitlement are considered unfair.

Raising snow-shovel prices after a snowstorm, for example, reduces consumer surplus and is considered unfair. But when the cost of a firm's inputs rises, subjects said it was fair to raise prices-- because <u>not</u> raising prices would reduce the firm's profit (compared to the reference profit). The Kahneman et al framework has found surprisingly little application, despite the everyday observation that firms do not change prices and wages as frequently as standard theory suggests. For example, when the fourth Harry Potter book was released in summer 2000, most stores were allocated a small number of books that were pre-sold in advance. Why not raise prices, or auction the books off? Everyday folks, like the subjects in KKT surveys, find actions which exploit excess demand to be outrageous. Concerned about customer goodwill, firms limit such price increases.

An open question is whether consumers are really willing to express outrage at unfairness by boycotts and other real sacrifices. Even if most aren't, a little threat of boycott may go a long way toward disciplining firms. (In the ultimatum game, for example, many subjects *do* accept low offers; but the fraction that reject such offers is high enough that it pays for Proposers to offer almost half.) Furthermore, even if consumer boycotts rarely work, offended consumers are often able to affect firm behavior by galvanizing media attention or provoking legislation. For example,"scalping" tickets for popular sports and entertainment events (reselling them at a large premium over the printed ticket price) is constrained by law in most states. Some states have "anti-gouging" laws penalizing sellers who take advantage of shortages of water, fuel, and other necessities by raising prices after natural disasters. A few years ago, responding to public anger at rising CEO salaries when the economy was being restructured through downsizing and many workers lost their jobs, Congress passed a law prohibiting firms from deducting CEO salary, for tax purposes, beyond $1 million a year (Rose & Wolfram, 2000). Explaining where these laws and regulations come from is one example of how behavioral economics might be used to expand the scope of law and economics (see Sunstein, 2000).

Behavioral game theory

Game theory has rapidly become an important foundation for many areas of economic theory, such as bargaining in decentralized markets, contracting and organizational structure, as well as political economy (e.g., candidates choosing platforms and congressional behavior). The descriptive accuracy of game theory in these applications can be questioned because equilibrium predictions often assume sophisticated strategic reasoning, and direct field tests are difficult. As a result, there have been many experiments testing game-theoretic predictions. "Behavioral game theory" uses this experimental evidence and psychological intuition to generalize the standard assumptions of game theory in a parsimonious way. Some of the experimental evidence, and its relation to standard ideas in game theory, is reviewed by Crawford (1997, and this volume). Newer data and theories which explain them are reviewed briefly by Goeree and Holt (1999) and at length by Camerer (this volume).

One component of behavioral game theory is a theory of social preferences for allocations of money to oneself and others (discussed above). Another component is a theory of how people choose in one-shot games or in the first period of a repeated game. A simple example is the `p-beauty contest game": Players choose numbers in [0,100] and the player whose number is closest in absolute value to p times the average wins a fixed prize. (The game is named after a well-known passage in which Keynes compared the stock market to a `beauty contest' in which investors only care about what stocks others think are `beautiful'.) There are many experimental studies for p=2/3. In this game the unique Nash equilibrium is zero. Since players want to choose 2/3 of the average number, if players think others will choose 50, for example, they will choose 33. But if they think others use the same reasoning and hence choose 33, they will want to choose 22. Nash equilibrium requires this process to continue until players beliefs' and choices match. The process only stops, mathematically, when x=(2/3)x, yielding an equilibrium of zero.

In fact, subjects in p-beauty contest experiments seem to use only one or two steps of iterated reasoning: Most subjects best-respond to the belief that others choose randomly (step 1), choosing 33, or best-respond to step-1 choices (step-2), choosing 22. (This result has been replicated with many subject pools, including Caltech undergraduates with median math SAT scores of 800 and corporate CEOs.)

Experiments like these show that the mutual consistency assumed in Nash equilibrium— players correctly anticipate what others will do— is implausible the first time players face a game, so there is room for a theory which is descriptively more accurate. A plausible theory of this behavior is that players use a distribution of decision rules, like the steps which lead to 33 and 22, or other decision rules (Stahl and Wilson, 1995; Costa-Gomes, Crawford & Broseta, 2001). Camerer, Ho and Chong (2001) propose a one-parameter cognitive hierarchy (CH) model in which the frequency of players using higher and higher steps of thinking is given by a one-parameter Poisson distribution). If the mean number of thinking steps is specified in advance (1.5 is a reasonable estimated), this theory has zero free parameters, is just as precise as Nash equilibrium (sometimes more precise), and always fits experimental data better (or equally well).

. A less behavioral alternative which maintains the Nash assumption of mutual consistency of beliefs and choices is a stochastic or "quantal-response" equilibrium (QRE; see Goeree and Holt (1999); McKelvey and Palfrey (1995, 1998); cf. Weiszacker, 2000). In a QRE players form beliefs about what others will do, and calculate the expected payoffs of different strategies, but they do not always choose the best response with the highest expected payoff (as in Nash equilibrium). Instead, strategies are chosen according to a statistical rule in which better responses are chosen more often. QRE is appealing because it is a minimal (one-parameter) generalization of Nash equilibrium, which avoids many of the technical difficulties of Nash[21] and fits data better.

A third component of behavioral game theory is a model of learning. Game theory is one area of economics in which serious attention has been paid to the process by which an equilibrium comes about. A popular approach is to study the evolution of a population (abstracting from details of how different agents in the population learn). Other studies posit learning by individual agents, based on their own experience or on imitation (e.g., Schlag, 1998). Many learning theories have been proposed and carefully tested with experimental data. Theories about population evolution never predict as well as theories of individual learning (though they are useful for other purposes). In reinforcement theories, only chosen strategies get reinforced by their outcomes (e.g., Roth et. al., 2000). In belief learning theories, players change their guesses about what other players will do, based on what they have seen, and choose strategies which have high expected payoffs given those updated guesses (e.g., Fudenberg & Levine, 1998). In the hybrid EWA theory of Camerer and Ho (1999), players respond weakly to ``foregone payoffs" from unchosen strategies and more strongly to payoffs they actually receive (as if underweighting ``opportunity costs"; see Thaler, 1999 and this volume). Reinforcement and ``fictitious play" theories of belief learning are boundary cases of the EWA theory. In many games (e.g., those with mixed-strategy equilibria) these theories are about equally accurate, and better than equilibrium theories. However, EWA is more robust in the sense that it predicts accurately in games where belief and reinforcement theories don't predict well (see Camerer, Ho and Chong, 2002).

Some next steps are to explore theoretical implications of the theories that fit data well, understand learning in very complex environments. The most important direction is application to field settings. Two interesting examples are the industrial structure in the Marseilles fish market (Weisbuch, Kirman & Herreiner, 2000), and a massive sample (130,000) of consumer supermarket purchases (Ho & Chong, 2000).

---

[21] A classic problem is how players in a dynamic game update their beliefs off the equilibrium path, when a move which (in equilibrium) has zero probability occurs. (Bayes' rule cannot be used because P(event)=0, so any conditional probability P(state|event) divides by zero.) QRE sidesteps this problem because stochastic responses ensure that all events have positive probability. This solution is much like the "trembles" proposed by Selten, and subsequent refinements, except that the tremble probabilities are endogeneous.

APPLICATIONS

Macroeconomics and Saving

Many concepts in macroeconomics probably have a behavioral underpinning that could be elucidated by research in psychology. For example, it is common to assume that prices and wages are rigid (in nominal terms), which has important implications for macroeconomic behavior. Rigidities are attributed to a vague exogeneous force like “menu costs," shorthand for some unspecified process that creates rigidity. Behavioral economics suggests some ideas for where rigidity comes from. Loss-aversion among consumers and workers, perhaps inflamed by workers’ concern for fairness, can cause nominal rigidity but are rarely discussed in the modern literature (though see Bewley, 1998; Blinder et al, 1998).

An important model in macroeconomics is the life-cycle model of savings (or permanent income hypothesis). This theory assumes that people make a guess about their lifetime earnings profile, and plan their savings and consumption to smooth consumption across their lives. The theory is normatively appealing if consumption in each period has diminishing marginal utility, and preferences for consumptions streams are time-separable (i.e., overall utility is the sum of the discounted utility of consumption in each separate period). The theory also assumes people lump together different types of income when they guess how much money they’ll have (i.e., different sources of wealth are fungible).

Shefrin and Thaler (1992 and this volume) present a "behavioral life cycle" theory of savings in which different sources of income are kept track of in different mental accounts. Mental accounts can reflect natural perceptual or cognitive divisions. For example, it is possible to add up your paycheck and the dollar value of your frequent flyer miles, but it is simply unnatural (and a little arbitrary) to do so, like measuring the capacity of your refrigerator by how many calories it holds. Mental accounts can also be bright-line devices to avoid temptation: Allow yourself to head to Vegas after cashing an IRS refund check, but not after raiding the childrens’ college fund or taking out a housing equity loan. Shefrin and Thaler (1992, and this volume) show that plausible assumptions about mental accounting for wealth predict important deviations from life-cycle savings theory. For example, the measured marginal propensities to consume (MPC) an extra dollar of income from different income categories are very different. The MPC from housing equity is extremely low (people don’t see their house as a pile of cash).

On the other hand, the MPC from windfall gains is substantial and often close to 1 (the MPC from one-time tax cuts is around 1/3-2/3).

It is important to note that many key implications of the life-cycle hypothesis have *never* been well-supported empirically (e.g., consumption is far more closely related to current income than it should be according to theory). Admittedly, since empirical tests of the life-cycle model involve many auxiliary assumptions, there are many possible culprits if the theory’s predictions are not corroborated. Predictions can be improved by introducing utility functions with “habit formation,” in which utility in a current period depends on the reference point of previous consumption, and by more carefully accounting for uncertainty about future income (see e.g. Carroll, 2000). Mental accounting is only one of several behavioral approaches that may prove useful.

An important concept in Keynesian economics is “money illusion”— the tendency to make decisions based on nominal quantities rather than converting those figures into “real” terms by adjusting for inflation. Money illusion seems to be pervasive in some domains. In one study (Baker, Gibbs, & Holmstrom, 1994) of wage changes in a large financial firm, only 200 of more than 60,000 wage changes were nominal decreases, but 15% of employees suffered real wage cuts over a 10-year period, and in many years more than half of wage increases were real declines. It appears that employees don’t seem to mind if their real wage falls as long as their nominal wage does not fall. Shafir, Diamond, and Tversky (1997 and this volume) demonstrate the pervasiveness of money illusion experimentally and sketch ways to model it.

Labor economics

A central puzzle in macroeconomics is involuntary unemployment-- why can some people not find work (beyond frictions of switching jobs, or a natural rate of unemployment)? A popular account of unemployment posits that wages are deliberately paid above the market-clearing level, which creates an excess supply of workers and hence, unemployment. But why are wages too high? One interpretation, "efficiency wage theory," is that paying workers more than they deserve is necessary to ensure that they have something to lose if they are fired, which motivates them to work hard and economizes on monitoring. Akerlof and Yellen (1990 and this volume) have a different interpretation: Human instincts to reciprocate transform the employer-worker relation into a "gift-exchange". Employers pay more than they have to as a gift; and workers repay the gift by working harder than necessary. They show how gift-exchange can be an equilibrium (given reciprocal preferences), and show some of its macroeconomic implications.

In labor economics, gift-exchange is clearly evident in the elegant series of experimental labor markets described by Fehr and Gachter (2000, and this volume). In their experiments there is an excess supply of workers. Firms offer wages; workers who take the jobs then choose a level of effort, which is costly to the workers and valuable to the firms. To make the experiments interesting, firms and workers can enforce wages, but not effort levels. Since workers and firms are matched anonymously for just one period, and do not learn each other's identities, there is no way for either side to build reputations or for firms to punish workers who chose low effort. Self-interested workers should shirk, and firms should anticipate that and pay a low wage. In fact, firms deliberately pay high wages as gifts, and workers choose higher effort levels when they take higher-wage jobs. The strong correlation between wages and effort is stable over time.

Other chapters in this section explore different types of departures from the standard assumptions that are made about labor supply. For example, standard life-cycle theory assumes that, if people can borrow, they should prefer wage profiles which maximize the present value of lifetime wages. Holding total wage payments constant, and assuming a positive real rate of interest, present value maximization implies that workers should prefer declining wage profiles over increasing ones. In fact, most wage profiles are clearly rising over time, a phenomenon which Frank and Hutchens (1993, and this volume) show cannot be explained by changes in marginal productivity. Rather, workers derive utility from positive changes in consumption, but have self-control problems that would prevent them from saving for later consumption if wages were front-loaded in the life-cycle. In addition, workers seem to derive positive utility from increasing wage profiles, per se, perhaps because rising wages are a source of self-esteem; the desire for increasing payments is much weaker for non-wage income (see Loewenstein & Sicherman, 1991).

The standard life-cycle account of labor supply also implies that workers should intertemporally substitute labor and leisure based on the wage rate they face and the value they place on leisure at different points in time. If wage fluctuations are temporary, workers should work long hours when wages are high and short hours when wages are low. However, because changes in wages are often persisting, and because work hours are generally fixed in the short-run, it is in practice typically difficult to tell whether workers are substituting intertemporally (though see Mulligan, 1998). Camerer et al. (1997, and this volume) studied labor supply of cab drivers in New York City (NYC). Cab drives represent a useful source of data for examining intertemporal substitution because drivers rent their cabs for a half-day and their work hours are flexible (they can quit early, and often do), and wages fluctuate daily because of changes in weather, day-of-the-week effects, and so forth. Their study was inspired by an alternative to the substitution hypothesis: Many drivers say they set a daily income target, and quit when they reach that target (in behavioral economics language, they isolate their daily decision and are averse to losing relative to an income target). Drivers who target daily will drive longer hours on low-wage days, and quit early on high-wage days. This behavior is exactly the opposite of intertemporal substitution. Camerer et al (1997, and this volume) found that data from three samples of inexperienced drivers support the daily targeting prediction. But experienced drivers do not have negative elasiticies, either because target-minded drivers earn less and self-select out of the sample of experienced drives, or drivers learn over time to substitute rather than target.

Perhaps the simplest prediction of labor economics is that the supply of labor should be upward sloping in response to a transitory increase in wage. Gneezy and Rustichini (this volume) document one situation in which this is not the case. They hired students to perform a boring task and either paid them a low piece-rate, a moderately high piece-rate, or no piece-rate at all. The surprising finding was that individuals in the low piece-rate condition produces the lowest "output" levels. Paying subjects, they argued, caused subjects to think of themselves as working in exchange for money and, when the amount of money was small, they decided that it simply wasn't worth it. In another study reported in their chapter, they showed a similar effect in a natural experiment that focused on a domain other than labor supply. To discourage parents from picking their children up late, a day-care center instituted a fine for each minute that parents arrived late at the center. The fine had the perverse effect of *increasing* parental lateness. The authors postulated that the fine eliminated the moral disapprobation associated with arriving late (robbing it of its gift-giving quality) and replaced it with a simple monetary cost which some parents decided was worth incurring. Their results show that the effect of price changes can be quite different than in economic theory when behavior has moral components which wages and prices alter.

Finance

In finance, standard equilibrium models of asset pricing assume that investors only care about asset risks if they affect marginal utility of consumption, and they incorporate publicly available information to forecast stock returns as accurately as possible (the "efficient markets hypothesis"). While these hypotheses do make some accurate predictions—e.g., the autocorrelation of price changes is close to zero—there are numerous anomalies. The anomalies have inspired the development of "behavioral finance" theories exploring the hypothesis that some investors in assets have limited rationality. Important articles are collected in Thaler (1993) and reviewed in Shleifer (2000) and Barberis and Thaler (2001).

An important anomaly in finance is the "equity premium puzzle": Average returns to stocks are much higher than returns to bonds (presumably to compensate stockholders for higher perceived risks).[22] To account for this pattern, Benartzi and Thaler (1995 and this volume) assume a combination of decision isolation—investors evaluate returns using a 1-year horizon—and aversion to losses. These two ingredients create much more perceived risk to holding stocks than would be predicted by expected utility. Barberis, Huang and Santos (2001) use a similar intuition in a standard asset pricing equation. Several recent papers (e.g., Barberis, Shleifer & Vishny, 1998) show how empirical patterns of short-term underreaction to earnings surprises, and long-term overreaction, can arise from a quasi-Bayesian model.

Another anomaly is the magnitude of volume in the market. The so-called "Groucho Marx" theorem states that people should not want to trade with people who would want to trade with them, but the volume of stock market transactions is staggering. For example, Odean (1999 and this volume) notes that the annual turnover rate of shares on the New York Stock Exchange is greater than 75 percent, and the daily trading volume of foreign-exchange transactions in all currencies (including forwards, swaps, and spot transactions) is equal to about one-quarter of the total annual world trade and investment flow. Odean (1999, and this volume) then presents data on individual trading behavior which suggests that the extremely high volume may be driven, in part, by overconfidence on the part of investors.

---

[22]The idea of loss aversion has appeared in other guises without being directly linked to its presence in individual choice. For example, Fama (1991:1596) wrote that "consumers live in morbid fear of recessions." His conjecture can only be reasonably construed as a disproportionate aversion to a drop in standard of living, or overweighting the low probability of economic catastrophe. Both are features of prospect theory.

The rise of behavioral finance is particularly striking because, until recently, financial theory bet all its chips on the belief that investors are too rational to ignore observed historical patterns-- the "efficient markets hypothesis." Early heretics like Shiller (1981), who argued empirically that stock price swings are too volatile to reflect only news, and DeBondt and Thaler (1985), who discovered an important overreaction effect based on the psychology of representativeness, had their statistical work "audited" with special scrutiny (or worse, were simply ignored). In 1978 Jensen called the efficient markets hypothesis "the most well-established regularity in social science." Shortly after Jensen's grand pronouncement, however, the list of anomalies began to grow. (To be fair, anomaly-hunting is aided by the fact that market efficiency is such a precise, easily-testable claim). A younger generation are now eagerly sponging up as much psychology as they can to help explain anomalies in a unified way.

NEW FOUNDATIONS

In a final, brief section of the book, we include two papers that take behavioral economics in new directions. The first is case-based decision theory (Gilboa & Schmeidler, 1995 and this volume). Because of the powerful influence of decision theory (a la Ramsey, de Finetti, & Savage) economists are used to thinking of risky choices as inevitably reflecting a probability-weighted average of the utility of their possible consequences. The case-based approach starts from different primitives. It treats a choice situation as a "case" which has degrees of similarity to previous cases. Actions in the current case are evaluated by a sum or average of the outcomes of the same action in previous cases, weighted by the similarity of those previous cases to the current one. Cased-based theory substitutes the psychology of probability of future outcomes for a psychology of similarity with past cases.

The primitive process of case comparison is widely used in cognitive science and is probably a better representation of how choices are made in many domains than is probability-weighted utility evaluation. In hiring new faculty members or choosing graduate students, you probably don't talk in terms of utilities and probabilities. Instead, it is irresistible to compare a candidate to others who are similar and who did well or poorly. Case-based reasoning may be just as appealing in momentous decisions, like choosing a presidential ticket (Lloyd Bentsen's "I knew John Kennedy, and you're no John Kennedy") or managing international conflict ("Will fighting the drug war in Colombia lead to another Vietnam?"). Explicitly case-based approaches are also widely used in the economy. Agents base a list price for a house on the selling prices of nearby houses that are similar ("comparables"). "Nearest-neighbor" techniques based on similarity are also used in credit-scoring and other kinds of evaluations.

Another promising new direction is the study of emotion, which has boomed in recent years (see Loewenstein & Lerner, 2001, for a review of this literature with a special focus on its implications for decision making). Damasio (1994) found that people with relatively minor emotional impairments have trouble making decisions and, when they do, they often make disastrous ones. Other research shows that what appears to be deliberative decision making may actually be driven by gut-level emotions or drives, then rationalized as a thoughtful decision (Wegner & Wheatley, 1999). Loewenstein (1996, in this volume, and 2000) discusses the possibilities and challenges from incorporating emotions into economic models.

There are many other new directions that behavioral economics is taking that, we hope, will provide more than adequate content for a sequel to this volume in the not too distant future. One such thrust is the study of "hedonics" (e.g., Kahneman, Diener, & Schwartz, 1999; Kahneman, Sarin & Wakker, 1997). Hedonics begins by expanding the notion of utility. In the neoclassical view, utility is simply a number that codifies an expressed preference ("decision utility"). But people may also have memories of which goods or activities they enjoyed most ("remembered utility"), immediate momentary sensations of pleasure and pain ("instant utility"), and guesses about what future utilities will be like ("forecasted utility"). It would be remarkable coincidence if the human brain were built to guarantee that all four types of utility were exactly the same. For example, current utilities and decision processes both depend on emotional or visceral states (like hunger, fatigue, anger, sympathy, or arousal), and people overestimate the extent they will be in the same hedonic state in the future (Loewenstein, 1996, and this volume). As a result, forecasted utility is biased in the direction of instant utility (see Loewenstein, O'Donoghue & Rabin, 1998). The differences among these utilities is important because a deviation between decision utility and one of the other types of utility means there is a mismatch which could perhaps be corrected by policies, education, or social guidance. For example, addicts may relapse because their remembered utility from using drugs highlights pleasure and excludes the instant disutility of withdrawal. The new hedonics links survey ratings of happiness with economic measures. For example, Easterlin (1974) stressed that average expressed ratings of happiness rise over decades much less than income rose. He suggested that people derive much of their happiness from relative income (which, by definition, cannot rise over time). Studies of worker quit rates, suicide, and other behavioral measures show similar effects of relative income and tie the happiness research to important economic phenomena (Clark & Oswald, 1994, 1996; Frey & Stutzer, 2002; Oswald, 1997).

A second direction uses neuroscientific evidence to guide assumptions about economic behavior. Neuroscience is exploding with discoveries because of advances in imaging techniques which permit more precise temporal and spatial location of brain activity.[23] It is undoubtedly a large leap from precise neural activity to big decisions like planning for retirement or buying a car. Nonetheless, neuroscientific data may show that cognitive activities that are thought to be equivalent in economic theory actually differ, or activities thought to be different may be the same. These data could resolve years or decades of debate which are difficult to resolve with other sorts of experiments (see Camerer, Loewenstein & Prelec, in press)

A third direction acknowledges Herb Simon's emphasis on "procedural rationality" and model the procedures or algorithms people use (e.g., Rubinstein, 1998). This effort is likely to yield models which are not simply generalizations of standard ones. For example, Rubinstein (1988) models risky choice as a process of comparing the similarity of the probabilities and outcomes in two gambles, and choosing on dimensions which are dissimilar. This procedure has some intuitive appeal but it violates all the standard axioms and is not easily expressed by generalizations of those axioms.

Concluding Comments

As we mentioned above, behavioral economics simply rekindles an interest in psychology that was put aside when economics was formalized in the latter part of the neoclassical revolution. In fact, we believe that many familiar economic distinctions do have a lot of behavioral content—they are *implicitly* behavioral, and could surely benefit from more explicit ties to psychological ideas and data.

---

[23]A substantial debate is ongoing in cognitive psychology about whether knowing the precise details of how the brain carries out computations is necessary to understand functions and mechanisms at higher levels. (Knowing the mechanical details of how a car works may not be necessary to turn the key and drive it.) Most psychology experiments use indirect measures like response times, error rates, self-reports, and "natural experiments" due to brain lesions, and have been fairly successful in codifying what we know about thinking; pessimists think brain scan studies won't add much. The optimists think the new tools will inevitably lead to some discoveries and the upside potential is so great that they cannot be ignored. We share the latter view.

An example is the distinction between short-run and long-run price elasticity. Every textbook mentions this distinction, with a casual suggestion that the long run is the time it takes for markets to adjust, or for consumers to learn new prices, after a demand or supply shock. Adjustment costs undoubtedly have technical and social components, but probably also have some behavioral underpinning in the form of gradual adaptation to loss, and learning.

Another macroeconomic model which can be interpreted as implicitly behavioral is the Lucas "islands" model (1975). Lucas shows that business cycles can emerge if agents observe local price changes (on "their own island") but not general price inflation. Are the "islands" simply a metaphor for the limits of their own minds? If so, theory of cognition could add helpful detail.

Theories of organizational contracting are shot through with implicitly behavioral economics. Williamson (1985) and others motivate the incompleteness of contracts as a consequence of bounded rationality in foreseeing the future, but do not tie the research directly to work on imagery, memory, and imagination. Agency theory begins with the presumption that there is some activity the agent does not like to do -- usually called "effort" -- which cannot be easily monitored or enforced, and which the principal wants the agent to do. The term "effort" connotes lifting sides of beef or biting your tongue when restaurant customers are sassy. What exactly is the "effort" agents dislike exerting, which principals want them to? It's not likely to be time on the job-- if anything, workaholic CEOs may be working too hard! A more plausible explanation, rooted in loss-aversion, fairness, self-serving bias, and emotion, is that managers dislike making hard, painful decisions (such as large layoffs, or sacking senior managers who are close friends). Jensen (1993) hints at the idea that overcoming these behavioral obstacles is what takes "effort"; Holmstrom and Kaplan (2000) talk about why markets are better at making dramatic changes than managers but ascribe much of the resistance to ``influence costs''. Influence costs are the costs managers incur lobbying for projects they like or personally benefit from (like promotions or raises). A lot of influence costs are undoubtedly inflated by optimistic biases-- each division manager really does think their division desperately needs funds--, self-serving biases, and social comparison of pay and benefits (otherwise, why are salaries kept so secret?).

In all these cases, conventional economic language has emerged which begs the deeper psychological questions of where adjustment costs, rigidities, mental "islands", contractual incompleteness, effort-aversion, and influence costs come from. Cognitively detailed models of these phenomena could surely produce surprising testable predictions.

Is psychology regularity an assumption or a conclusion?

Behavioral economics as described in this chapter, and compiled in this book, generally **begins** with assumptions rooted in psychological regularity and asks what follows from those assumptions. An alternative approach is to work backward, regarding a psychological regularity as a **conclusion** that must be proved, an *explanandum* that must be derived from deeper assumptions before we fully understand and accept it.

The alternative approach is exemplified by a fashionable new direction in economic theory (and psychology too), which is to explain human behavior as the product of evolution. Theories of this sort typically describe an evolutionary environment, a range of behaviors, and precise rules for evolution of behavior (e.g., replicator dynamics), and then show that a particular behavior is evolutionarily stable. For example, overconfidence about skill is evolutionarily adaptive under some conditions (Postlewaite & Comte, 2001; Waldman, 1994). Loss-aversion can be adaptive (because exaggerating one's preference for an object improves one's outcome under the Nash bargaining solution and perhaps other protocols; e.g., Carmichael & MacLeod, 1999). Rejections of low offers in take-it-or-leave-it ultimatum games are often interpreted as evidence of a specialized adaptation for punishing partners in repeated interactions, which cannot be "turned off" in unnatural one-shot games with strangers (e.g., Samuelson, 2001).

We believe in evolution, of course, but we do not believe that behavior of intelligent, modern people immersed in socialization and cultural influence can only be understood by guessing what their ancestral lives were like and how their brains might have adapted genetically. Furthermore, a major challenge for evolutionary modeling is that ex post "just so stories" are easy to concoct because there many degrees of freedom permitted by our inability to travel back in time to the ancestral past. As a result, it is easy to figure out mathematically whether an evolutionary story is a **sufficient** explanation of behavior, and almost impossible to know whether a particular story is **necessarily** the right one.

Another potential problem with evolutionary reasoning is that most studies posit a special brain mechanism to solve a particular adaptive problem, but ignore the effect of how that mechanism constrains solution of other adaptive problems. (This is nothing more than the general equilibrium critique of partial equilibrium modelling, applied to the brain.) For example, a fashionable interpretation of why responders reject ultimatum offers is that agents cannot instinctively distinguish between one-shot and repeated games. But agents who could not do this would presumably be handicapped in many other sorts of decisions which require distinguishing unique and repeated situations, or accurately forecasting horizons (such as life-cycle planning), unless they have a special problem making distinctions among types of games.

There are other, non-evolutionary, models that treat psychological regularity as a conclusion to be proved rather than an assumption to be used.[24] Such models usually begin with an observed regularity, and reverse-engineer circumstances under which it can be optimal. Models of this sort appeal to the sweet tooth economists have for deriving behavior from "first principles" and rationalizing apparent irrationality. Theories of this sort are useful behavioral economics, provided they are held to the same high standards all good models are (and earlier behavioral models have been held to): Namely, can they parsimoniously explain a <u>range</u> of data with one simple mechanism? And what <u>fresh predictions</u> do they make?

Final thoughts

Critics have pointed out that behavioral economics is not a unified theory, but is instead a collection of tools or ideas. This is true. It is also true of neoclassical economics. A worker might rely on a "single" tool-- say, a power drill-- but also use a wide range of drill bits to do various jobs. Is this one tool or many? As Arrow (1986) pointed out, economic models do not derive much predictive power from the single tool of utility-maximization. Precision comes from the drill bits—such as time-additive separable utility in asset pricing including a child's utility into a parent’s utility function to explain bequests, rationality of expectations for some applications and adaptive expectations for others, homothetic preferences for commodity bundles, price-taking in some markets and game-theoretic reasoning in others, and so forth.

---

[24]For example, one recent model (Benabou & Tirole, 1999) derives overconfidence from hyperbolic time discounting. Agents, at time 0, face a choice at time 1 between a task that requires an immediate exertion of effort and a payoff delayed till time 2 which depends on their level of some skill. Agents know that, due to hyperbolic time discounting, some tasks that are momentarily attractive at time 0 will become unattractive at time 1. Overconfidence arises because they persuade themselves that their skill level – i.e., the return from the task – will be greater than it actually will be so as to motivate themselves to do the task at time 1. There may, however, be far more plausible explanations for the same phenomenon, such as that people derive utility directly from self-esteem. Indeed the same authors later proposed precisely such a model (Benabou & Tirole, 2000).

Sometimes these specifications are even contradictory— for example, pure self-interest is abandoned in models of bequests, but restored in models of life-cycle savings; and risk-aversion is typically assumed in equity markets and risk-preference in betting markets. Such contradictions are like the "contradiction" between a Phillips-head and a regular screwdriver: They are different tools for different jobs. The goal of behavioral economics is to develop better tools that, in some cases, can do both jobs at once.

Economists like to point out the natural division of labor between scientific disciplines: Psychologists should stick to individual minds, and economists to behavior in games, markets, and economies. But the division of labor is only efficient if there is effective coordination, and all too often economists fail to conduct intellectual trade with those who have a comparative advantage in understanding individual human behavior. All economics rests on <u>some</u> sort of implicit psychology. The only question is whether the implicit psychology in economics is good psychology or bad psychology. We think it is simply unwise, and inefficient, to do economics without paying <u>some</u> attention to good psychology.

We should finally stress that behavioral economics is not meant to be a separate approach in the long run. It is more like a school of thought or a style of modeling, which should lose special semantic status when it is widely taught and used. Our hope is that behavioral models will gradually replace simplified models based on stricter rationality, as the behavioral models prove to be tractable and useful in explaining anomalies and making surprising predictions. Then strict rationality assumptions now considered indispensable in economics will be seen as useful special cases (much as Cobb-Douglas production functions or expected value maximization are now)—namely, they help illustrate a point which is truly established only by more general, behaviorally-grounded theory.

References

Akerlof, G. A. (1991). Procrastination and obedience. American Economic Review, 81(2), 1-19.

Akerlof, G. A. & Yellen, J. L. (1990). The fair wage-effort hypothesis and unemployment Quarterly Journal of Economics, 105(2), 255-83.

Allais, M. (1953). Le comportement de l'homme rationnel devant le risque, critique des postulats et axiomes de l'ecole americaine. Econometrica, 21, 53-526.

Angeletos, G-M., Repetto, A., Tobacman, J. & Weinberg, S. (2001) The hyperbolic buffer stock model: Calibration, simulation, and empirical evaluation. Journal of Economic Perspectives, 15(3), 47-68.

Ariely, D., Loewenstein, G. & Prelec, D. (in press). Coherent arbitrariness: Stable demand curves without stable preferences. Quarterly Journal of Economics.

Ariely, D. & Wertenbrock, K. (in press). Procrastination, deadlines, and performance: Using public commitments to regulate one's behavior. Psychological Science.

Arrow, Kenneth J. (1986). Rationality of self and others in an economic system. In R. M. Hogarth & M. W. Reder (eds.), Rational choice: The contrast between economics and psychology (pp. 201-215). Chicago, IL: University of Chicago Press.

Asch, D.A., Patton, J.P. and Hershey, J.C. (1990). Knowing for the sake of knowing. Medical Decision Making, 10, 47-57.

Babcock, L., & Loewenstein, G. (1997). Explaining bargaining impasse: the role of self-serving biases. Journal of Economic Perspectives. 11, 109-126.

Babcock, L., & Loewenstein, G. (this volume).

Baker, G., Gibbs, M., & Bengt H. (1994). The wage policy of a firm. Quarterly Journal of Economics, 109, 921-955.

Barberis, N., Huang, M., & Santos, T. (2001). Prospect theory and asset prices. Quarterly Journal of Economics, 116(1), 1-53.

Barberis, N., Shleifer, A. & Vishny, R. (1998). A Model of Investor Sentiment. Journal of Financial Economics, 49(3), 307-343

Barberis, N. & Thaler, R. (in press). A survey of behavioral finance. In G. Constantinides, M. Harris and R. Stulz, Handbook of the Economics of Finance. http://gsbwww.uchicago.edu/fac/nicholas.barberis/research/

Benabou, R. & Tirole, J. (1999). Self confidence: Intrapersonal strategies. Princeton working paper. .

Benabou, R. & Tirole, J. (2000). Self confidence and social interactions. Princeton working paper.

Benartzi, S. & Thaler, R. H. (1995). Myopic loss aversion and the equity premium puzzle. Quarterly Journal of Economics, 110(1), 73-92.

Benzion, U., Rapoport, A. & Yagil, J. (1989). Discount rates inferred from decisions: An experimental study. Management Science, 35, 270-284.

Bewley, T. F. (1998). Why not cut pay? European Economic Review, 42(3-5), 459-490.

Blinder, A. S., Canetti, E., Lebow, D. E. & Rudd, J. B. (1998). Asking about prices: A new approach to understanding price stickiness. New York: Russell Sage Foundation.

Bolton, G. E., & Ockenfels, A (2000). ERC: A theory of equity, reciprocity, and competition. American Economic Review, 90(1), 166-193.

Camerer, C. F. (1987). Do biases in probability judgment matter in markets? Experimental evidence. American Economic Review, 77, 981-997.

Camerer, C. F. (1989a). Does the basketball market believe in the 'hot hand'? American Economic Review, 79, 1257-1261.

Camerer, C. F. (1989b). An experimental test of several generalized utility theories. Journal of Risk and Uncertainty 2, 61-104.

Camerer, C. F. (1992). Recent tests of generalized utility theories. In W. Edwards (ed.), Utility theories: Measurement and applications (pp. 207-251). Cambridge: Cambridge University Press.

Camerer, C. F. (2000). Prospect theory in the wild: Evidence from the field. In D. Kahneman & A. Tversky (eds.), Choices, values, and frames (pp. 288-300). Cambridge: Cambridge University Press.

Camerer, C. F. (2002). Behavioral game theory: Experiments on strategic interaction. Princeton: Princeton University Press.

Camerer, C. F., Babcock, L., Loewenstein, G. & Thaler, R. (1997). Labor supply of New York City cab drivers: One day at a time. Quarterly Journal of Economics, 111, 408-441.

Camerer, C. F. & Ho, T. H. (1999). Experience weighted attraction learning in normal-form games. Econometrica, 67, 827-874.

Camerer, C., Ho, T. H. & Chong, K. (2001). Behavioral game theory: Thinking, learning and teaching. Caltech working paper. http://www.hss.caltech.edu/~camerer/camerer.html

Camerer, C. F., Ho. T. & Chong, K. (2002). Functional EWA: A one-parameter theory of learning in games. Caltech working paper. http://www.hss.caltech.edu/~camerer/camerer.html

Camerer, C. F. & Hogarth, R. M. (1999). The effects of financial incentives in economics experiments: A review and capital-labor-production framework. Journal of Risk and Uncertainty, 19, 7-42.

Camerer, C., Loewenstein, G., & Prelec, D. (in press). Gray matters: How neuroscience can inform economics. Journal of Economic Perspectives.

Camerer, C. F. & Thaler, R. (1995). Anomalies: Dictators, ultimatums, and manners. Journal of Economic Perspectives, 9, 209-219.

Camerer, C. F. & Weber, M. W. (1992). Recent developments in modelling preferences: Uncertainty and ambiguity. Journal of Risk and Uncertainty, 5, 325-370.

Carmichael, H. L. & MacLeod, B. (1999). Caring about sunk costs: A behavioral solution to hold-up problems with small stakes. USC Olin working paper no. 99-19.

Carroll, C. (2000). Requiem for the representative consumer? Aggregate implications of microeconomic consumption behavior. American Economic Review, Papers and Proceedings, 90(2), 110-115.

Carson, R. T., Horowitz, John. (1990). Discounting statistical lives. Journal of Risk and Uncertainty.

Charness, G. & Rabin, M. (forthcoming). "Understanding Social Preferences with Simple Tests" Quarterly Journal of Economics.

Clark, A. E. & Oswald, A. J. (1994). Unhappiness and unemployment. Economic Journal 104(424), 648-659.

Clark, A. E. & Oswald, A. J. (1996). Satisfaction and comparison income. Journal of Public Economics, 61(3), 359-381.

Cooper, D. J., Kagel, J. H., Lo, W. & Gu, Q. L. (1999). Gaming against managers in incentive systems: Experimental results with Chinese students and Chinese managers. American Economic Review, 89(4), 781-804.

Costa Gomes, M., Crawford, V. & Broseta, B. (in press). Experimental studies of strategic sophistication and cognition in normal-form games. Econometrica, 69(), 1193-1235.

Crawford, V. (1997). Theory and experiment in the analysis of strategic interaction. In D. Kreps & K. Wallis (eds.), Advances in economics and econometrics: Theory and applications (XX-XX). Cambridge: Cambridge University Press.

Damasio, A. R. (1994). Descartes' error: Emotion, reason, and the human brain. New York: Putnam.

De Bondt, W. F. M. & Thaler, R. (1985). Does the stock market overreact? Journal of Finance, 40(3), 793-805.

Degeorge, F., Patel, J., & Zeckhauser, R. (1999). Earnings management to exceed thresholds. Journal of Business, 72, 1-33.

Easterlin, R. A. (1974). Does economic growth improve the human lot? Some empirical evidence. In: P. A. David & M. W. Reder (eds), Nations and households in economic growth: Essays in honor of Moses Abramowitz (89-125). New York: Academic Press.

Ellsberg, D. (1961). Risk, ambiguity, and the savage axioms. Quarterly Journal of Economics, 75, 643-669.

Epstein, L. G. & Wang, T. (1994). Intertemporal asset pricing under Knightian uncertainty. Econometrica, 62(2), 282-322.

Fama, E. F. (1991). Efficient capital markets: II. Journal of Finance, 46(5), 1575-1617.

Fehr, E. & Gachter, S. (2000, and this volume). Fairness and retaliation: The economics of reciprocity. Journal of Economic Perspectives, 14(3), 159-181.

Fehr, E. & Schmidt, K. M. (1999). A theory of fairness, competition and cooperation. Quarterly Journal of Economics, 117(3), 817-868.

Frank, R. H. & Hutchins, R. M. (1993 & this volume). Wages, seniority, and the demand for rising consumption profiles. Journal of Economic Behavior and Organizations, 21, 251-276.

Frederick, S., Loewenstein, G. & O'Donoghue, T. (in press & this volume). Intertemporal Choice: A critical review. Journal of Economic Literature.

Frey, B. S. & Stutzer, A. (2002). What can economicsts learn from happiness research. Journal of Economic Literature, XL, 402-435.

Fudenberg, D. & Levine, D. (1998). The theory of learning in games. Cambridge: MIT Press.

Ganguly, A., Kagel, J. H. & Moser, D. (2000). Do asset market prices reflect traders' judgment biases? Journal of Risk and Uncertainty, 20, 219-246.

Geanakoplos, J., Pearce, D. & Stacchetti, E. (1989). Psychological games and sequential rationality. Games and Economic Behavior, 1, 1989, 60-80.

Genesove, D. & Mayer, C. (2001). Loss aversion and seller behavior: Evidence from the housing market. Quarterly Journal of Economics, 116(4), 1233-1260.

Ghirardato, P. & and Katz, J. (2000). Indecision theory: Explaining selective abstention in multiple elections. Caltech University working paper. http://masada.hss.caltech.edu/~paolo/ghiro.html

Gigerenzer, G., Hell, W. & Blank, H. (1988). Presentation and content: The use of base rates as a continuous variable. Journal of Experimental Psychology 14(3), 513-525.

Gilboa, I. & Schmeidler, D. (1995). Case-based decision theory. Quarterly Journal of Economics, 110, 605-39.

Gilovich, T., Vallone, R. & Tversky, A. (1985). The hot hand in basketball: On the misperception of random sequences. Cognitive Psychology 17, 295-314.

Gneezy & Rustichini (this volume).

Goeree, J. K., & Holt, C. A. (1999). Stochastic game theory: For playing games, not just for doing theory. Proceedings of the National Academy of Sciences, 96, 10564-10567.

Goldstein, W. M., & Weber, E. U. (1995). Content and its discontents: The use of knowledge in decision making. In J. R. Busemeyer, R. Hastie & D. L. Medin (eds.), Decision making from a cognitive perspective. The psychology of learning and motivation, Vol. 32 (pp. 83-136). New York, NY: Academic Press.

Grether, D. M. & Plott, C. (1979). Economic theory of choice and the preference reversal phenomenon. American Economic Review, 69(4), 623-638.

Grinblatt, M. & Keloharju, M. (2001). Distance, language and culture bias: The role of investor sophistication. Journal of Finance. 56(3), 1053-1073.

Güth, W., Schmittberger, R. & Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. Journal of Economic Behavior and Organization, 3, 367-388.

Hansen, L. P., Sargent, T., & Tallarini Jr., T. D. (2001). Robust permanent income and pricing. Review of Economic Studies, 66, 873-907.

Hardie, B. G. S., Johnson, E. J. & Fader, P. S. (1993). Reference dependence, loss-aversion, and brand choice. Marketing Science, 12, 1993, 378-394.

Harless, D. & Camerer. C. F. (1994). The predictive utility of generalized expected utility theories. Econometrica, 62, 1251-1290.

Heath, C., Larrick, R. P., & Klayman, J. (1998). Cognitive repairs: How organizational practices can compensate for individual shortcomings. Review of Organizational Behavior, 20, 1-38.

Heilbroner, R. L. (1982). The Socialization of the Individual in Adam Smith. History of Political Economy, 14,(3), 427-439.

Helson, H. (1964). Adaptation level theory: An experimental and systematic approach to behavior. New York: Harper & Row.

Hertwig, R. & Ortmann, A. (in press). Experimental practices in economics: A methodological challenge for psychologists. Behavioral and Brain Sciences.

Hey, J. (1997). Experiments and the Economics of Individual Decision Making under Risk and Uncertainty. In Kreps D.M. and Wallis K.F. (eds), *Advances in Economics and Econometrics: Theory and Applications*, vol 1, pp 173-205, Cambridge University Press, 1997.

Ho, T. & Chong, J. K. (2000). A parsimonious model of SKU choice. Wharton School, PA: Unpublished working paper.

Hoffman, E., McCabe, K., Shachat, K. & Smith, V. L. (1994). Preferences, property rights and anonymity in bargaining games. Games and Economic Behavior, 7, 346-380.

Hogarth, R. M. & Einhorn, H. J. (1992). Order effects in belief updating: The belief-adjustment model. Cognitive Psychology, 24(1), 1-55

Hogarth, R. M. & Reder, M. W. (1987). Rational choice : The contrast between economics and psychology. Chicago : University of Chicago Press.

Holcomb, J. H. & Nelson, P. S. (1992). Another experimental look at individual time preference. Rationality and Society, 4(2), 199-220.

Holmstrom, B. & Kaplan, S. N. (2000). .Corporate governance and merger activity in the U.S.: Making sense of the 80s and 90s. MIT: Unpublished working paper.

Hsee, C. K. & Leclerc, F. (1998). Will products look more attractive when evaluated jointly or when evaluated separately? Journal of Consumer Research, 25, 175-186.

Hsee, C. K., Loewenstein, G, Blount, S., & Bazerman, M. (1999). Preference reversals between joint and separate evaluations of options: a theoretical analysis. Psychological Bulletin, 125(5), 576-590.

Huber, J., Payne, J. W., & Puto, C. (1982). Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. Journal of Consumer Research, 9(1), 90-98.

Jensen, M. (1993). The modern industrial revolution, exit, and the failure of internal control sysstem. Journal of Finance, 48(3), 831-881.

Johnson, E. J. & Schkade, D. A. (1989). Bias in utility assessments: Further evidence and explanations. Management Science 35. 406-424.

Kahneman, D., Diener, E. & Schwarz, N. (eds.) (1999). Well-being: The foundations of hedonic psychology. New York: Russell Sage Foundation Press.

Kahneman, D. J. & Frederick, S. (2002). Representativeness revisited: Attribution substitution in intuitive judgment. In T. Gilovich, D. Griffin & D. Kahneman, Heuristics of intuitive judgment: Extensions and applications (pp. XX-XX). New York: Cambridge University Press.

Kahneman, D., Knetsch, J. L. & Thaler, R. (1990). Experimental tests of the endowment effect and the coase theorem. Journal of Political Economy, 98, 1325-1348.

Kahneman, D., Knetsch, J. L. & Thaler, R. (1986). Fairness as a constraint on profit seeking: Entitlements in the market. American Economic Review, 76, 728-741.

Kahneman, D. & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. Econometrica, 47, 263-291.

Kahneman, D. & Tversky, A. (1982). On the study of statistical intuitions. In Kahneman, D., Slovic, P. and Tversky, A. Judgment under Uncertainty: Heuristics and Biases. Cambridge: Cambridge University Press.

Kahneman, D., Wakker, P. P. & Sarin, R. (1997). Back to Bentham? Explorations of experienced utility. Quarterly Journal of Economics, 112, 375-405.

Knetsch, J. L. (1992). Preferences and nonreversibility of indifference curves. Journal of Economic Behavior and Organization, 17(1), 131-139.

Koehler, J. J. (1996). Issues for the next generation of base rate research. Behavioral and Brain Sciences, 19, 41-53.

Laibson, D. (1997). Golden eggs and hyperbolic discounting. Quarterly Journal of Economics, 112, 443-477.

Laibson, D., Repetto, A. & Tobacman, J. (1998). Self-control and saving for retirement. Brookings Papers on Economic Activity, 91-196.

Lewin , S. B. (1996). Economics and psychology: Lessons for our own day from the early twentieth century. Journal of Economic Literature, 34(3), 1293-1323.

List, J. (in press). The effect of market experience on the WTA/WTP disparity: Evidence from the field. Quarterly Journal of Economics.

Loewenstein, G. (1987). Anticipation and the valuation of delayed consumption. Economic Journal, 97, 666-684

Loewenstein, G. (1988). Frames of mind in intertemporal choice. Management Science, 34, 200-214.

Loewenstein, G. (1992). The fall and rise of psychological explanation in the economics of intertemporal choice. In G. Loewenstein & J. Elster (eds.), Choice over time (pp. 3-34). New York: Russell Sage.

Loewenstein, G. (1996 & this volume). Out of control: visceral influences on behavior. Organizational Behavior and Human Decision Processes, 65, 272-92.

Loewenstein, G. (1999). Experimental economics from the vantage-point of behavioral economic. Economic Journal Controversy Corner: "What's the Use of Experimental Economics," 109, 25-34.

Loewenstein, G. (2000). Emotions in economic theory and economic behavior. American Economic Review: Papers and Proceedings, 90, 426-32.

Loewenstein, G. (2001) The creative destruction of decision research. Journal of Consumer Research, 28(3), 499-505.

Loewenstein, G. & Adler, D. (1995). A bias in the prediction of tastes. Economic Journal, 105, 929-937.

Loewenstein, G. & Lerner, J. (in press). The role of emotion in decision making. In. R.J. Davidson, H.H, Goldsmith & K.R. Scherer, The handbook of affective science. Oxford, England: Oxford University Press.

Loewenstein, G., O'Donoghue T. & Rabin, M. (1999). Projection bias in the prediction of future utility. Working Paper.

Loewenstein, G. & Prelec D. (1992). Anomalies in intertemporal choice: Evidence and an interpretation. Quarterly Journal of Economics, May, 573-597.

Loewenstein, G. & Prelec, D. (1993). Preferences over outcome sequences. Psychological Review, 100(1), 91-108.

Loewenstein, G. & Sicherman, N. (1991). Do workers prefer increasing wage profiles? Journal of Labor Economics, 9, 67-84.

Lucas Jr., R. E. Jr. (1975). An equilibrium model of the business cycle. Journal of Political Economy, 83, 1133-1144.

Lucas, Robert E. Jr. (1986). Adaptive behavior in economic theory. Journal of Business, 59(4), S401-S426.

MacKeigan, L. D., Larson, L. N., Draugalis, J. R., Bootman, J. L., and Bruns, L. R. (1993). Time preference for health gains versus health losses. PharmacoEconomics, 3, 374-386.

Markowitz, H. (1952). The utility of wealth. Journal of Political Economy 60, 151-158.

McCabe, Kevin, Houser, Daniel, Ryan, Lee, Smith, Vernon and Trouard, Theodore (2001). A Functional Imaging Study of Cooperation in Two-Person Reciprocal Exchange," Proceedings of the National Academy of Sciences, (98)20, 11832-11835.

McKelvey, R. D. & Palfrey, T. R. (1995). Quantal response equilibria for normal form games. Games and Economic Behavior, 7, 6-38.

McKelvey, R. D. & Palfrey, T. R. (1998). Quantal response equilibria for extensive form games. Experimental Economics, 1, 9-41.

Mukherji, S. (1998). Ambiguity aversion and incompleteness of contractual form. American Economic Review, 88 (5), 1207-1231.

Mulligan, C. B. (1998). Substitution over time: Another look at life cycle labor supply. NBER Working Paper No.w6585.

O'Curry, S. (1999). Consumer budgeting and mental accounting. In P. E. Earl & S. Kemp (eds.) The Elgar companion to consumer research and economic psychology (pp. XX-XX). Northhampton, MA: Cheltenham.

Odean, T. (1998). Are investors reluctant to realize their losses? Journal of Finance, LIII(5), 1775-1798.

Odean, T. (1999, and this volume). Do investors trade too much? American Economic Review, 89, 1279-1298.

O'Donoghue, T. & Rabin, M. (1999, and this volume). Doing it now or later. American Economic Review, 89, 103-124.

O'Donoghue, T. & Rabin, M. (2001). Choice and procrastination. Quarterly Journal of Economics, 116(1), 121-160.

Oswald, A. (1997). Happiness and economic performance. Economic Journal, 107, 1815-1831.

Payne, J. W., Bettman, J. R. & Johnson, E. J. (1992). Behavioral decision research: A constructive processing perspective. Annual Review of Psychology, 43, 87-131.

Pender, John L. 1996. "Discount rates and credit markets: Theory and evidence from rural India." Journal of Development Economics, 50:2, pp. 257-96.

Phelps, E. S. & Pollak, R. A. (1968). On second-best national saving and game-equilibrium growth. Review of Economic Studies, 35, 185-199.

Postlewaite, A. & Compte, O. (2001). Confidence enhanced performance. University of Pennsylvania working paper. http://www.ssc.upenn.edu/%7Eapostlew/paper/working.html

Prelec, D. & Loewenstein, G. (1991). Decision making over time and under uncertainty: A common approach. Management Science, 37, 770-786.

Prelec, D. & Loewenstein, G. (1998). The red and the black: mental accounting of savings and debt. Marketing Science, 17, 4-28.

Prelec, D., Wernerfelt, B. & Zettelmeyer, F. (1997). The role of inference in context effects: Inferring what you want from what is available. Journal of Consumer Research, 24(1), 118-125.

Quiggin, J. (1993). Generalized expected utility theory: The rank-dependent expected utility model. Amsterdam: Kluwer-Nijhoff.

Rabin, M. (1993, and this volume). Incorporating fairness into game theory and economics. American Economic Review, 83(5), 1281-1302.

Rabin, M. (2000). Risk-aversion and expected utility theory: A calibration theorem. Econometrica, 68(5), 1281-1292.

Rabin, M. (2002). The law of small numbers. Quarterly Journal of Economics.

Rabin, M. & Schrag, J. (1999). First impressions matter: A model of confirmatory bias. Quarterly Journal of Economics, 114(1), 37-82/

Ramsey, F. P (1928). A mathematical theory of saving. Economic Journal(?), 38(152), 543-559

Read, D., Loewenstein, G. & Rabin, M. (1999). Choice bracketing. Journal of Risk and Uncertainty, 19, 171-97.

Rose, N. & Wolfram. C. (2000). Has the "million-dollar cap" affected the CEO pay? American Economic Review, 90(2), 197-202.

Roth, A., Erev, I., Slonim, R., & G. Barron. G. (2000). Equilibrium and learning in economic environments: The predictive value of approximations. Discussion paper, Harvard University, Department of Economics.

Rubinstein, A. (1988). Similarity and decision-making under risk (Is there a utility theory resolution to the Allais paradox?). Journal of Economic Theory, 46, 145-153.

Samuelson, Larry. 2001. Adaptations, analogies and anomalies. Journal of Economic Theory, 97, 320-66.

Samuelson, Paul (1937). A note on measurement of utility. Review of Economic Studies, 4, 155-161.

Savage, L. J. (1954). The foundations of statistics. New York: Wiley.

Sayman, S.A. and Öncüler, Ayse (1997). "A Meta Analysis of the Willingness- to-Accept and Willingness-to-pay disparity. Working paper, Wharton School, University of Pennsylvania. Available at: http://home.ku.edu.tr/~ssayman/research/meta.pdf

Schelling, T. C. (1978). Egonomics, or the art of self-management. American Economic Review, 68(2), 290-294.

Schlag, K. (1998). Why imitate, and if so, How? A bounded rational approach to multi-armed bandits. Journal of Economic Theory, 78(1), 130-156.

Shafir, E., Diamond, P. & Tversky, A. (1997). Money illusion. Quarterly Journal of Economics, 112(2), 341-74.

Shefrin, H. M. & Thaler, R. H. (1992). Mental accounting, saving, and self-control. In G. Loewenstein & J. Elster (eds.), Choice over time (pp. 287-329). New York: Russell Sage Foundation.

Shiller, R. (1981). Do stock prices move too much to be justified by subsequent changes in dividends? American Economic Review, 71, 421-436.

Shleifer, A. (2000). Inefficient markets. Oxford: Oxford University Press.

Simonson, I. (1990). The effect of purchase quantity and timing on variety-seeking behavior. Journal of Marketing Research, 17, 150-162.

Simonson, I. & Tversky, A. (1992). Choice in context: Tradeoff contrast and extremeness aversion. Journal of Marketing Research, 29, 281-295.

Simonson, I., & Winer, R. S. (1992). The influence of purchase quantity and display format on consumer preference for variety. Journal of Consumer Research, 19, 133-138.

Slovic, Paul. (1995). The construction of preferences. American Psychologist, 50, 364-371.

Smith, A. (1759/1892). The Theory of Moral Sentiments. New York: Prometheus Books.

Stahl, D. & Wilson, P. (1995). On players' models of other players: Theory and experimental evidence. Games and Economic Behavior, 10(1), 218-254

Starmer, C. (2000). Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk. Journal of Economic Literature. 38, 332-382.

Stigler G. (1965). “The development of utility theory,” Chap. 5 in Essays in the History of Economics. Chicago: University of Chicago Press.

Strahilevitz, M., & Loewenstein, G. (1998). The effects of ownership history on the valuation of objects. Journal of Consumer Research, 25, 276-289.

Strotz, R.H., (1955). Myopia and inconsistency in dynamic utility maximization. Review of Economic Studies, 23, 165-80.

Sunstein, C. (Ed.) (2000). Behavioral Law and Economics. Cambridge: Cambridge University Press.

Thaler, R. (1980). Toward A Positive Theory of Consumer Choice, Journal of Economic Behavior and Organization, 1, 39-60.

Thaler, R. (1981). Some empirical evidence on dynamic inconsistency. Economics Letters, 8, 201-207.

Thaler, R. (1993). (Ed.) Advances in behavioral finance. New York: Russell Sage Foundation.

Thaler, R. (1994). Psychology and savings policies. American Economic Review, 84(2), 186-192.

Thaler (1999, this volume). Mental accounting matters. Journal of Behavioral Decision Making, 12, 183-206.

Thaler, R. H. & Shefrin, H. M. (1981). An economic theory of self-control. Journal of Political Economy, 89(2). 392-406.

Tversky, A. & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. Science, 185(4157), 1124-1131

Tversky, A. & Kahneman, D. (1981). The framing of decisions and the psychology of choice. Science, 211, 453-458.

Tversky, A. and Kahneman, D. (1983). Extensional vs. intuitive reasoning: The conjunction fallacy in probability judgment. Psychological Review, 90, 293-315.

Tversky, A. & Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. Quarterly Journal of Economics 106, 1039-1061.

Tversky, A. & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. Journal of Risk and Uncertainty, 5, 297-323.

Tversky, A., Slovic, P., & Kahneman, D. (1990). The causes of preference reversal. American Economic Review, 80, 204-217.

VanBoven, L., Dunning, D. & Loewenstein, G. (2000). Egocentric empathy gaps between owners and buyers: Misperceptions of the endowment effect. Journal of Personality and Social Psychology, 79, 66-76.

VanBoven, L., Loewenstein, G. & Dunning D. (2000). Biased predictions of others' tastes: Underestimation of owners' selling prices by "buyer's agents." Working paper.

Waldman, M. (1994). Systematic errors and the theory of natural selection The American Economic Review, 84(3), 482-497.

Weisbuch, G., Kirman, A. & Herreiner, D. (2000). Market organisation and trading relationships. The Economic Journal, 110, 411-436.

Weizsacker, G. (2000). Ignoring the rationality of others: Evidence from experimental normal-form games. Harvard Business School working paper.

Wegner, D. M. & Wheatley, T. (1999). Apparent mental causation: Sources of the experience of will. American Psychologist, 54(7), 480-492.

Williamson, O. (1985). The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting. New York: The Free Press.