Behavioral Economics

Colin F. Camerer

Div HSS 228-77
Caltech
Pasadena CA 91125 USA
camerer@hss.caltech.edu

Abstract

Behavioral economics uses evidence from psychology and other disciplines to create models of limits on rationality, willpower and self-interest, and explore their implications in economic aggregates. This paper reviews the basic themes of behavioral economics: Sensitivity of revealed preferences to descriptions of goods and procedures; generalizations of models of choice over risk, ambiguity, and time; fairness and reciprocity; non-Bayesian judgment; and stochastic equilibrium and learning. A central issue is what happens in equilibrium when agents are imperfect but heterogeneous; sometimes firms "repair" limits through sorting, but profit-maximizing firms can also exploit limits of consumers. Frontiers of research are careful formal theorizing about psychology and studies with field data. Neuroeconomics extends the psychological data use to inform theorizing to include details of neural circuitry. It is likely to support rational choice theory in some cases, to buttress behavioral economics in some cases, and to suggest different constructs as well.

~~February 3, 2011~~~~March 1, 2006.~~ This paper was prepared for the World Congress of the Econometric Society, 2005, London 18-24 August 2005. Thanks to audience members and Ariel Rubinstein for comments, and to Joseph Wang for editorial help.

I. The themes and philosophy of behavioral economics

Behavioral economics applies models of systematic imperfections in human rationality, to the study and engineering of organizations, markets and policy. These imperfections include limits on rationality, willpower and self-interest (Rabin, 1998; Mullainathan and Thaler, 2000), and any other behavior resulting from an evolved brain with limited attention. The study of individual differences in rationality, and learning, is also important for understanding whether social interaction and economic aggregation minimizes effects of rationality limits.

In one sense, behavioral economics is the inevitable result of relaxing the assumption of perfect rationality. Like perfect competition and perfect information, the assumption of perfect agent rationality is a useful limiting case in economic theory. Generalizing those assumptions to account for imperfect competition and costly information was challenging, slow, and proved to be powerful; weakening the assumption of perfect rationality will be too.

One property of models of human rationality, which largely distinguishes them from studies of economic competition, is that other social sciences have cumulated a lot of ideas and empirical facts about human rationality. The approach to behavioral economics that I will describe chooses to pay careful attention to those constructs and facts. In this empirically-driven approach to behavioral economics, assumptions are chosen to fit what is known from other sciences. This approach can be thought of as scientifically humble, or it can be thought of as efficient and respectful of comparative advantage across disciplines.

Other than trying to “get the psychology right” in choosing assumptions, the empirically-driven approach to behavioral economics shares the methodological emphases of other kinds of analysis: The goal is to have simple formal models and themes which apply across many domains, which make predictions about naturally-occurring data (as well as experimental data).

The behavioral economics approach I describe in this essay is a clear departure from the "as if" approach endorsed by Milton Friedman. His "F-twist" argument combines two criteria:

1. Theories should be judged by the accuracy of their predictions;

2. Theories should not be judged by the accuracy of their assumptions.

The empirically-driven approach to behavioral economics agrees with criterion (1) and rejects criterion (2). In fact, criterion 2 is rejected *because of* the primacy of criterion 1, based on the belief that replacing unrealistic assumptions with more psychologically realistic ones should lead to better predictions. This approach has already had some success: This paper reports many examples of how behavioral theories grounded in more reasonable assumptions can account for facts about market outcomes which are anomalies under rational theories. More empirical examples are emerging rapidly.

The empirically-driven approach to behavioral economics combines two practices: (i) Explicitly modeling limits on rationality, willpower and self-interest; and (ii) using established facts to suggest assumptions about those limits. A different, "mindless", approach (Gul and Pesendorfer, 2005) follows elements of practice (i) but not (ii), modeling limits but enthusiastically ignoring empirical details of psychology. The argument for the mindless approach is Friedmanesque: Since theories that infer utility from observed choices were not originally intended to be tested by any data other than choices[1], evidence about assumptions does not count.

But theories are not copyrighted. So neuroscientists, for example, are free to assume that utilities actually are numbers which correspond to the magnitude of some process in the brain (e.g., neural firing rates) and search for utilities using neuroscientific methods (knowing full well their results will be ignored by "mindless"-type economists). Such a search doesn't 'misunderstand economics', it just takes the liberty of defining economic variables as neural constructs. The hope is also that new neural constructs will be discovered

[1] The doctrine that choices are the only possible data is a modern one, however (see footnote 3 below).

that are most gracefully accommodated only if the standard language of preference, belief and constraint is stretched by some new vocabulary.

Before proceeding, let me clarify two points. First, the discussion above should make clear that behavioral economics is *not* a distinct subfield of economics. It is a style of modeling, or a school of thought, which is meant to apply to a wide range of economic questions in consumer theory, finance, labor economics and so on. Second, while the psychological data that fueled many developments in behavioral economics are largely experimental, behavioral economics is an *approach* and experimental economics is a *method*. It is true that early in modern behavioral economics, experiments proved to be useful as a way of establishing that anomalies were not produced by factors that are hard to rule out in field data-- transaction costs, risk-aversion, confusion, self-selection, etc.— but are easy to rule out with good experimental control. But the main point of these experiments was just to suggest regularities that could be included in models to make predictions about naturally-occurring field data.

Section II is a brief digression reminding us that behavioral economics is something of a return to old paths in economic thought which were not taken. Section III reviews the tools and ideas that are the current canon of what is best established (see also Conlisk, 1996, Camerer, Loewenstein, and Rabin, 2004). Section IV is a reminder that aggregate outcomes—behavior in firms, and markets—matter and considers how imperfections in rationality cumulate or disappear at those levels. Section V discussing “franchises” of behavioral economics in applied areas, and some examples of growth in theory and field empirics. Section VI discusses neuroeconomics and section VII concludes.

II. Behavioral paths not taken

Why did behavioral economics not emerge earlier in the history of economic thought? The answer is that *it did*: Jeremy Bentham, Adam Smith, Irving Fisher, William Jevons and many others drew heavily on psychological intuitions. But those intuitions were largely left behind in the development of mathematical tools of economic analysis, consumer theory and general equilibrium (e.g., Ashraf, Camerer and Loewenstein, 2005; Colander, 2005).

For example, Adam Smith believed there was a disproportionate aversion to losses which is a central feature of Kahneman and Tversky's prospect theory. Smith wrote (1759, III, *ii*, pp. 176-7):

Pain ... is, in almost all cases, a more pungent sensation than the opposite and correspondent pleasure. The one almost always depresses us much more below the ordinary, or what may be called the natural state of our happiness, than the other ever raises us above it.

Smith (1759, II, ii, *ii*, p. 121) also anticipates Thaler's (1980) seminal[2] analysis of the insensitivity to opportunity costs, compared to out-of-pocket costs:

…breach of property, therefore, theft and robbery, which take from us what we are possessed of, are greater crimes than breach of contract, which only disappoints us of what we expected.

Why did behavioral insights like these get left out of the neoclassical revolution? A possible answer, suggested by Bruni and Sugden (2005), is that Vilfredo Pareto won an argument among economists in the early 1900's about how deeply economic theories should be anchored in psychological reality. Pareto thought ignoring psychology was not only acceptable, but was also necessary. In an 1897 letter he wrote:

> It is an empirical fact that the natural sciences have progressed only when they have taken secondary principles as their point of departure, instead of trying to discover the essence of things. ... Pure

[2] Many people regard Thaler's 1980 paper as the starting point of behavioral economics per se, since it drew on psychology but was clearly focused on the economics of consumer choice (see Thaler, 1999 for an update on the same topic).

> political economy has therefore a great interest in relying as little as possible on the domain of psychology.

Pareto advocated divorcing economics from psychology by simply *assuming* that unobserved Benthamite utility (“the subjective fact”) is revealed by choice (“the objective fact”). He justifies this equation (in modern terms, that choices necessarily reveal true preferences) by restricting attention “only [to] repeated actions”, so that consistency results from learning.

The Paretian equation of choice and true preference is neither a powerful proof nor a robust empirical regularity. It is a philosophical stance, pure and simple. And because Pareto clearly limits the domain of revealed preference to “repeated actions” in which learning has taught people what they want, he leaves out important economic decisions that are rare or difficult to learn about from trial-and-error (e.g., Einhorn, 1982)—corporate mergers, fertility and mate choice, partly-irreversible education and workplace choices, planning for retirement, buying houses, and so forth.

Could economic theory have taken another path? Many economists such as Edgeworth, Ramsey, and Fisher speculated about how to measure utility directly, but lacked modern tools and gave up[3]. What seemed an impossible task a hundred years ago might be possible now, given developments in experimental psychology, neuroscience and genetics. So this is a good time in history to revisit the ideas of Adam Smith and others, and the paths not taken by neoclassical economists due to Pareto’s bold move.

[3] Colander (2005) notes that Edgeworth described a “hedonimeter” which would measure momentary fluctuations in pleasure, and eventually provide a basis for utilitarian adding-up across people. Irving Fisher also speculated about how to measure utility in his 1892 dissertation. Ramsey wrote about a “psychogalvanometer”. It is interesting to speculate about whether at least some economists might have taken a different path in the early 20th century if fMRI, genetic methods, single-unit recording, and other tools were available which allowed more optimism about measuring utility directly. Would any of them have become neuroeconomists? Even if most did not, it is hard to believe that *none* of them would have, given the curiosity evident in all their writing.

III. The basic ideas and tools of behavioral economics

Much of behavioral economics emerged as the study of deviations from rational-choice principles. (The fact that clear deviations are permitted is one way the rational-choice approach is powerful.) Deviations and anomalies are not merely counterexamples, which any simplified theory permits; they are clues about new or more general theories.[4] I prefer alternative theories which include rational-choice as a limiting special case. These generalizations provide a clear way to measure the parametric advantage of extending the theory. They also make it easy to search empirically for conditions under which rational-choice principles hold.

Table 1 lists some central rational-choice modeling principles in economic theory, emerging behavioral alternative models, and some representative citations (see McFadden, 1999, for a longer list). I will describe each briefly, and highlight domains in which competing alternatives are emerging.

Complete preferences: Completeness and transitivity of preference (which implies that choices can be represented by real-valued utilities) is an extremely powerful simplification. But the power comes precisely from excluding the many variables that a good's utility could depend upon. Thinking of choice as a result of cognition suggests obvious ways in which completeness of preference will be violated (e.g., Kahneman, 2003). The way in which choices are described, or "*framed*", can influence choice by directing attention to different features. The psychophysics of adaptation suggests that changes from a point of reference (*reference-dependence*) are likely to be a carrier of utility. A long-standing empirical problem is what the natural point of reference is (and how reference points change). Koszegi and Rabin (in press) suggest a resolution that should charm game theorists: The point of reference is the expectation of actual choice (which determines choice

[4] Lucas (1986) notes because rational expectations often permits multiple equilibria, theories based on limited rationality might actually be more precise than theories based on full rationality. This is also true in game theory, where theories with rationality limits can be more precise than equilibrium theories (e.g., Camerer, Ho, and Chong 2004). So if the goal is precision, behavioral alternatives may prove even better than rational theories in some cases.

recursively, since preferences depend on utilities relative to the reference point).[5] This approach creates multiple equilibria, which permits a supply-side role for marketing, advertising, and sale prices to influence preferences by creating reference points (e.g., Koszegi and Heidhues, 2005). This approach also provides a language in which to understand how small changes in instructions or repeated trading experience could change behavior—namely, through the reference point.[6]

Slovic and Lichtenstein (1968) were the first to notice that *reversals of expressed preference* could result when people choose between two gambles, relative to pricing the gambles separately, a violation of procedure-invariance (see also Grether and Plott, 1979). This insight lays the groundwork for using pricing institutions (such as different auctions) to influence expressions of preference.

Human perception and cognition is heavily influenced by contrast. A circle looks larger when surrounded by smaller circles than when it is surrounded by larger circles (the Titchener illusion). Since choices undoubtedly involve basic perceptual and cognitive neural circuitry, it would be surprising if choice evaluation were not sensitive to contrast as well. Indeed, there is ample evidence that the appeal of choices depends on the set of choices they are part of (e.g., Simonson and Tversky, 1992; Shafir, Osherson and Smith, 1989). Similarly, psychological comparison of outcomes with unrealized outcomes (disappointment) or with outcomes from foregone choices (regret) imply that the utility of a gamble is not separable into a sum of its expected component utilities, but there are workable formal models of these phenomena (e.g., Gul, 1991; Loomes, Starmer and Sugden, 1989).

---

[5] Denote the reference point by r (which may be probabilistic). Koszegi and Rabin assume utility depends on a combination of absolute outcomes, m(x) and a function μ(m(x)-m(r)) which is reference-dependent, depending on the difference m(x)-m(r)) between consumption utility and the reference utility. When goods have deterministic utility and the reference point is the same as the bundle chosen, then x=r so the second term disappears, and the model reduces to standard consumer theory.

[6] List (2003) finds that experienced sports-card dealers do not exhibit an “endowment effect” (while novice traders do). A natural interpretation is that dealers do not expect to hold on to goods they receive. Since their reference point does not include the goods, they do not feel less of a loss when selling them. Kahneman, Knetsch and Thaler (1990:1328) clearly anticipated this effect of experience, noting 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."

Choice over risk: Many applications in economics require a specification of preferences over gambles which have probabilistic risk, when probabilities may be subjective and when costs and benefits are spread over time. Independence axioms assume that people implicitly cancel common outcomes of equal probability in comparing risky choices (contrary to gestalt principles of perception, which resist cancellation), which leads mathematically to expected utility (EU) and subjective expected utility.

In contrast to EU, prospect theory assumes reference-dependence and diminishing psychophysical sensitivity, which together imply a "*reflection*" of risk preferences around the reference point (i.e., ,since the hedonic sensation of loss magnitude is decreasing at the margin, the utility function for loss is convex). Many other non-EU theories have been proposed and studied (Starmer, 2000), but prospect theory is more clearly rooted in psychology than most other theories, which are generally based on ingenious ways of weakening the independence axiom. Prospect theory also survives well in careful empirical comparisons among many theories aggregating many different studies, and adjusting for degrees of freedom (Harless and Camerer, 1994; cf. Hey and Orme, 1994).

The other components of prospect theory are disproportionate disutility for losses (compared to equal-sized gains)- "*loss-aversion*"—and nonlinear sensitivity to probability, probably due to nonlinearity in attention to low probabilities (e.g. Prelec, 1998).[7] Coefficients of loss-aversion[8] **λ**—the ratio of marginal loss to gain utilities around zero—have been estimated from a wide variety of data to fall in a range around 2 (see Table 2). The striking feature of the table is that the studies cover such a wide range of types of data and levels of analysis.

---

[7] The one-parameter version of Prelec's axiomatically-derived weighting function is $\pi(p)=1/\exp((\ln(1/p)^{\gamma})$ (where $\exp(x)=e^{x}$). In this remarkable function, the ratio of overweighting $\pi(p)/p$ grows very large as p becomes very small, as if there is a quantum of attention put on *any* probability, no matter how low. For example, with γ=.7 (an empirical estimate from experiments), $\pi(1/10)$=.165, $\pi(1/100)$=.05, and $\pi(1/1{,}000{,}000)$=.002. This type of extreme relative overweighting of very low probabilities is useful for explaining overreaction to rare diseases (mad cow disease), and the huge popularity of high-prize Powerball lotteries.

[8] The coefficient of loss-aversion is defined as the ratio of the limits of marginal utilities at the reference point, where marginal utilities approach from below and above, respectively. This definition allows a "kink" at the reference point which exhibits "first-order risk-aversion" (i.e., the utility loss from a gamble is proportional to the standard deviation, so that agents dislike even small-stakes gambles; Segal and Spivak, 1990).

Choice over ambiguity: Subjective expected utility (SEU) assumes that subjective (or, in Savage's term, "personal") probabilities are revealed by the willingness to bet on events. However, as Ellsberg's famous 1961 paradox showed (following Keynes and Knight), bet choices could depend both on subjective likelihood and the "weight of evidence" or confidence one has in the likelihood judgment; when bets are "ambiguous" decision weight is lower. In SEU, subjective probabilities are a slave with two masters—likelihood and willingness to bet (or decision weight). As Schmeidler (1989) pointed out, a simple resolution is to assume that decision weights are *nonadditive*. Then the nonadditivity is a measure of "reserved belief", or the strength of the unwillingness to bet on *either* color in the face of missing relevant information. Mukerji and Tallon (2004) describes many theoretical applications of ambiguity-aversion models to contracting, game theory and other domains.

Choice over time: If choices are dynamically consistent, then the discount weight put on future utilities must be exponential ($u(x_t)=\delta^t$). While dynamic consistency is normatively appealing, it seems to be contradicted by everyday behavior like procrastination and succumbing to temptations created by previous choices.[9] To understand these phenomena, Laibson (1997) borrowed a two-piece discounting function from work on intergenerational preference. His specification puts a weight of 1 on immediate rewards, and weights $u(x_t)=\beta\delta^t$ on rewards at future times t. This "quasi-hyperbolic" form is a close approximation to the mountains of evidence that animal and human discount functions are hyperbolic, $d(t)=1/(1+kt)$, and is easy to work with analytically. (Rubinstein, 2003 suggests an alternative based on temporal similarity.) The β-δ model has been calibrated to explain regularities in aggregate savings and borrowing patterns (Angeletos, Laibson, Tobacman, 2001), and applied to the study of procrastination and deadlines by O'Donoghue and Rabin (2001).

Self-interest: The idea that people only care about their own monetary or goods payoffs is not a central tenet of rational choice theory, but it is a common simplifying assumption. Economists also tend to be skeptical that people will sacrifice to express a concern for the payoffs of others. As Stigler (1981) wrote, "when self-

[9] See also Gul and Pesendorfer, 2001, who model a distaste for flexibility when choice sets include tempting goods.

interest and ethical values with wide verbal allegiance are in conflict, much of the time, most if the time in fact, self-interest theory…will win."

Despite skepticism like Stigler's, there is a long history of models that attempt to formalize when people trade off their own payoffs for payoffs of others (e.g., Edgeworth, 1881; "equity theory" in social psychology; and Loewenstein, Bazerman and Thompson, 1989).

Sensible models of this type face a difficult challenge: Sometimes people sacrifice to increase payoffs of others, and sometimes they sacrifice to lower the payoffs of others. The challenge is to endogenize when the weights placed on payoffs of others switch from positive to negative. A breakthrough paper is Rabin's (1993), based on psychological game theory, which includes beliefs as a source of utility. In Rabin's approach, players form a judgment of kindness or meanness of another player, based on whether the other player's action gives the belief-forming player less or more than a reference point (which can depend on history, culture, etc.). Players prefer to reciprocate in opposite directions, acting kindly toward others who are kind, and acting meanly toward others who are mean. As a result, in a coordination game like "chicken", there is an equilibrium in which both players expect to treat each other well, and they actually do (since doing so gives higher utility, but less money). But there is another equilibrium in which players expect each other to act meanly, and they also do. Rabin's model shows the thin line between love and hate. Falk and Fischbacher (2005) and Dufwenberg and Kirchsteiger (2004) extend it to extensive-form games, which is conceptually challenging.

A different approach is to assume that players have an unobserved type (depending on their social preferences), and their utilities depend on their types and how types are perceived (e.g., Levine, 1998 and Rotemberg, 2004). These models are more technically challenging but can explain some stylized facts.

Simpler models put aside judgments of kindness based on intentions, and just assume that people care about both money and inequity, either measured by absolute payoff deviations (Fehr and Schmidt, 1999) or by the deviation between earnings shares and equal shares (Bolton and Ockenfels, 2000). Charness and Rabin

(2002) introduce a “Rawlsitarian” model in which people care about their own payoff, the minimum payoff (Rawlsian) and the total payoff (utilitarian). In all these models, self-interest emerges as a special case when the weight on one’s own payoff swamps the weights on other terms.

These models are *not* an attempt to invent a special utility function for each game. They are precisely the opposite. The challenge is to show that the same general utility function, up to parameter values, can explain a wide variety of data that vary across games and institutional changes (e.g., Fischbacher, Fong and Fehr, 2003).

Bayesian statistical judgment: The idea that people’s intuitive judgments of probability obey statistical principles, and Bayes’ rule, is used in many applied microeconomics models (e.g., in games of asymmetric information). Tversky and Kahneman (see Kahneman, 2003) used deviations between intuitive judgments and normative principles (“biases”) to suggest heuristic principles of probability judgment. Their approach is explicitly inspired by theories of perception, which use optical illusions to suggest principles of vision (Tversky and Kahneman, 1982), without implying that everyday visual perception is badly maladaptive. Similarly, heuristics for judging probability (like availability of examples, and representativeness of samples to underlying processes) are not necessarily maladaptive. The point of studying biases is just to illuminate the heuristics they reveal, not to indict human judgment. Thus, their original view is consistent with the critique that heuristics can be ecologically rational.

The Bayesian approach is so simple and useful that is has taken some time to craft equally simple formal alternatives which are consistent with the heuristics Kahneman and Tversky suggested. An appealing way to do is to use the Bayesian framework but assume that people misspecify or misapply it in some way. Rabin and Schrag (1999) give a useful model of “confirmation bias”. They define confirmation bias as the tendency to overperceive data as more consistent with a prior hypothesis than they truly are. The model is fully Bayesian except for the mistake in encoding of data. Rabin (2002) models representativeness as the (mistaken) expectation that samples are drawn without replacement, and shows some fresh implications of that model (e.g, perceiving more skill among managers than truly exists). Barberis, Shleifer and Vishny (1998) show how a similar misperception among stock investors, that corporate earnings which actually follow a random walk either exhibit momentum or mean-reversion, can generate short-term underreaction ("earnings drift") and long-term overreaction in stock returns.

Another principle implicit in Bayesian reasoning is informational irreversibility—if you find out a piece of evidence is mistaken, the brain should reverse its impact on judgment. (For example, juries are instructed to ignore certain statements after they have been heard.) But the brain is an organ, as is human skin. When skin is grafted onto skin, the old and new merge and eventually it is impossible to undo the graft. Information in the brain is probably organically irreversible in a similar way. For example, when people find out that an event occurred, it is hard to resist a "hindsight bias", which biases recollection of ex ante probability in the direction of new information (Fischhoff and Beyth, 1975; Camerer, Loewenstein, Weber, 1989).

Equilibrium: Moving beyond the level of individual choice and judgment, behavioral economics has also contributed to a shift in the study of equilibrium at the market or game-theoretic level. Game theorists, in particular, have never been comfortable with simply assuming that beliefs and choices are in equilibrium—i.e., that players correctly anticipate what others will do—without clearly specifying mechanisms that generate equilibration. Evolutionary game theory (e.g., Weibull, 1995; Samuelson, 1997), and the sensible extension to the study of imitation (e.g., Schlag, 1998), are important approaches which show how equilibria might emerge from limited rationality and selection pressures.

Empirical models of learning in games have also been carefully calibrated on many different types of experimental data. One approach is reinforcement of chosen strategies (Arthur, 1991; Erev and Roth, 1998). A seemingly different approach is updating of beliefs based on experience, as in fictitious play (e.g., Fudenberg and Levine, 1998). However, Camerer and Ho (1999) noted that fictitious play is simply a generalized kind of reinforcement in which unchosen strategies are reinforced as strongly as chosen strategies are. That recognition inspired a hybrid "dual process theory" (EWA) in which reinforcement of actual and foregone outcomes can differ, nesting choice reinforcement and fictitious play as boundary cases. The hybrid model tends to fit about as well as each of the boundary cases, and sometimes fits substantially better when one of the models misses a central feature of the data. Ho, Camerer and Chong (2005) introduce a "self-tuning" version of their hybrid theory in which the key parameters adjust flexibly to experience, which economizes on parameters allows changes in the rate of learning after "surprise".[10]

Another approach to game-theoretic equilibrium maintains the assumption of equilibrium beliefs, but substitutes stochastic choice for best-response, creating "quantal response equilibrium" (QRE) models (McKelvey and Palfrey, 1998). Weakening best-response explains many of the experimental deviations from Nash equilibrium, but also approximates Nash play in games where the Nash equilibrium tends to be accurate (Goeree and Holt, 2001).

An alternative non-equilibrium approach, rooted in principles of limited cognition, assumes a "cognitive hierarchy" (CH) in which more thoughtful players best-respond to their perceptions that others do less thinking (Nagel, 1995; Stahl and Wilson, 1995; Costa-Gomes, Crawford, Broseta, 2001). These CH approaches are more precise than Nash equilibrium because they always predicts a single statistical distribution of play, and are generally more accurate than equilibrium in predicting behavior in one-shot games.

Before proceeding, note that the rational principles which are listed in Table 1 are *normative*. They describe behavior of an idealized agent with unlimited cognitive resources and willpower. As we are beginning to understand (e.g., Robson, 2001), it is unlikely that evolution would have sculpted us to satisfy these principles for all important economic decisions. As a result, it is a scientific error in judgment to always privilege normative principles in the search for the best descriptive principles across all decisions people make

[10] The self-tuning approach is similar to Erev, Bereby-Meyer and Roth (1999)'s use of "payoff variability"; and Marcet and Nicoli (2003)'s theory of regime-shifts in response to hyperinflation. Self-tuning also creates shift in parameter values, as if players are switching rules throughout the game, akin to direct learning across rules (cf. Stahl, 2000, on "rule learning").

(see also Starmer, 2004). Normative principles are, of course, useful in raising our children, teaching students, judging welfare, and as limiting cases of how some people behave or learn to behave. Or normative principles might be enforced by aggregation of decisions and market discipline, a crucial topic we consider next.

IV: Aggregation: From individuals to firms and markets

The previous section described behavioral economics alternatives to rational-choice microfoundations. But the central question is: What happens in a political economy where agents have limited rationality (e.g., Camerer and Fehr, 2006)?

Asking about market and political outcomes forces behavioral economics to confront two classes of questions that have not been the central focus of research so far: First, how heterogeneous are agents? And how detectable is heterogeneity? (This question is important because heterogeneity drives the division of labor in organizations, the development of expertise and human capital, and market interaction of rational and limitedly-rational agents.) And second, how do institutions sort heterogeneous agents, supply market substitutes for individual irrationality, and create organizational outcomes on the supply side?

Early theory: Some early papers tackled the issue of market aggregation theoretically. A pioneering paper is Thaler and Russell (1985)[11] who emphasized constraints that prevent rationality limits from being erased. Haltiwanger and Waldman (1989) noted that whether individual mistakes would be erased or magnified depends on whether behaviors are strategic *substitutes* or strategic *complements*. When behaviors are complements, a small proportion of irrational traders can force others to behave irrationally (as Keynes wrote about the stock market). The “limits to arbitrage” literature in finance is an extension of this general theme (e.g., Shleifer, 2000).

[11] See also the correction in Thaler and Russell (1987).

Sorting and constraint: Aggregation issues are central in labor economics. The fact that workers have different skills leads to sorting (self-selection and firms' allocation of workers to jobs), specialization, and division of labor.

Recent evidence shows substantial effects of basic intelligence on the tendency to make the kind of judgment mistakes documented in the heuristics literature, and on risk-aversion and immediacy preference (Benjamin and Shapiro, 2005; Frederick, 2005). This kind of evidence invites the possibility that "smarter" people will be sorted into jobs where their decisions minimize or repair mistakes by others. In a magazine interview Gary Becker opined that "division of labor…'strongly attenuates if not eliminates' any effects caused by bounded rationality" (Stewart, 2005).[12]

Becker's conjecture should be explored theoretically and empirically. The power of division of labor to necessarily produce organizational efficiency may be limited by various factors. First of all, large organizations demand some skills at a very high level (e.g., extreme honesty when there are huge opportunities for theft). A limited supply of agents with enough skill will then limit the size of the firm.

Second, the sorting process requires a human resources department or other mechanism to identify talent. If the ability to spot talent is itself a scarce talent, or self-selection is limited by optimism (for example), those forces will limit how much talent is spotted.

Third, what happens if managers are biased in one dimension but excellent at another? Hard-driving CEO's, for example, may be superb at motivating people and creating an inspiring vision, precisely *because* they are wildly optimistic and genuinely convinced they can't fail. So it is possible that the sorting process of managerial selection actually *selects* for optimism rather than selects for realism. The organizational challenge is to design job structure that harnesses a CEO's optimism as motivation, but keeps that optimism from making bad investments.

[12] Becker conjectures that "it doesn't matter if 90 percent of people can't do the complex analysis required to calculate probabilities. The 10 percent of people who can will end up in jobs where it's required". A good example is insurance actuaries or analysts who price derivative assets.

Finally, note that sorting is difficult to study in the field, but it is *easy* to study experimentally—because agents' characteristics can be measured, and self-selection can be measured too (e.g., Lazear, Malmendier, and Weber, 2005).

Organizational repairs: An interesting supply-side response to managerial rationality limits is what Heath, Larrick and Klayman (1998) call "organizational repairs". They suggest that some organizational practices can be seen as responses to managerial errors. Microsoft had a hard time getting its programmers to take customer complaints seriously (despite statistical evidence from customer help-lines), because the programmers thought the software was easy to use and couldn't believe that customers found it difficult (a "curse of knowledge"). So Microsoft created a screening room with a one-way mirror, so programmers could literally see for themselves how much trouble normal-looking consumers had using software. The trick was to use one judgment bias—the power of visually "available" evidence, even in small samples— to overcome another bias (the curse of knowledge).

Experiments on rationality aggregation: Experiments are ideally suited to studying how rationality aggregates. In an experiment, one can measure the degree of individual bias and market-level bias, and compute whether biases in market prices or quantities is smaller than the average (or dollar-weighted) individual bias. Anderson and Sunder (1995), and Camerer (1987) studied errors in abstract Bayesian judgments designed to test whether traders would overreact to likelihood evidence (and underweight priors) when a small sample of balls drawn from a bingo cage was "representative" of the cage's contents. They found small biases in market prices, which were reduced by hours of trading, but not eliminated. Ganguly, Kagel, and Moser (2000) found much larger pricing errors when the event was a hypothetical word problem rather than a bingo cage draw. Camerer, Loewenstein and Weber (1989) studied the "curse of knowledge" (mistakenly assuming other subjects have your private information) and Kluger and Wyatt (2004) studied the famous "Monty Hall" three-door problem. Both found that market trading reduced, but did not eliminate, mistakes. Maciejovsky and Budescu

(2005) found that markets for information in Wason 4-card logic problems do guide agents toward rational solutions.

The rationality tug-of-war between consumers and firms: Suppose you struggle with a gambling problem, and type “pathological gambling” into the Google search engine looking for help.[13] When I did this in April 2005, one of the entries on the first page is shown in Figure 1 (leading to http://www.casinolasvegas.com/currency-us-dollars/lang-en/skins/noscript.html).

This exercise illustrates the rationality tug-of-war between consumers and firms: If heterogeneity and sorting enables firms to weed out poorly-suited workers, is the result a larger supply of products and techniques for taking *advantage* of limited consumer rationality, or a larger supply of products that *help* consumers?

Figure 1: A first-page entry in an April 2005 Google search for “pathological gambling”

1. **GAMBLING PROBLEMS** - TOP RATED ONLINE CASINO SITES. FREE KENO MASSAGE SANDALS BONUS

   ... is licensed and **gambling problems** regulated ! Here you will find **gambling problems** more information about all ...

   www.casino-startup.com/**gambling-problems**.html - 17k - Cached - More from this site - Save - Block

Whether markets will correct rationality depends on factors like whether consumers know their own limits (and hence are receptive to advice), and whether there is more profit in protecting consumers or taking advantage of them. The result for any particular rationality limit is likely to depend sensitively on self-

[13] Thanks to George Loewenstein’s office door for this example.

awareness, industrial structure, regulation and law, the role of education in educating consumers, household dynamics between spouses, and many other factors.

One result might be an arms race in which consumer protection and exploitation *both* increase. For example, in the recent rise of obesity among Americans, industries selling cheap caloric food (such as pizzas with cheese *inside the crust*) have flourished. But healthier food, diet books, personal training, plastic surgery, and eating disorders have flourished too.

A simple example of how to analyze the impact of consumer rationality on markets is Gabaix and Laibson (2006)'s model of products with "add-ons". Add-ons are typically marginal goods or services whose prices can be easily hidden or "shrouded" (like bank transaction fees or the cost of printer ink cartridges). If enough consumers don't think about the shrouded add-on price, then in a competitive market firms will compete by offering very low prices on base goods (below marginal cost) and will charge high markups on add-ons. Sophisticated consumers who know the add-on price, but can cheaply substitute away from the add-ons (avoiding bank ATM fees, for example) will *prefer* products with expensive add-ons, because they benefit from the low base-good price produced by competition. (The myopic consumers who don't think about the add-on cost are subsidizing the sophisticated consumers.) As a result, competition does not theoretically lead to revealing the add-on price, because a firm that reveals its add-ons will not attract either myopic consumers (who will mistakenly think the price-revealing firm is too expensive) *or* sophisticates (who benefit from the below-cost base-good price). This paper is a good example of why careful analysis is needed to be able to make sharp conclusions about whether markets will erase or exploit limits on consumer rationality. Two other examples are Della Vigna and Malmendier (2005)'s analysis of gym memberships, and Grubb (2005) 's analysis of overconfident planning of cell phone usage of minutes, and pricing of packages.

V. Some Frontiers of Behavioral Economics

This section is about some new frontiers in behavioral economics: Franchising (applying behavioral economics to traditional subfields, like finance and labor); formal foundations; field studies; and importing different kinds of psychology.

A. The franchising of behavioral economics

Much of the power of economic analysis comes from models used in different applications areas which rely on shared general principles—consistent preferences and equilibrium-- but are customized to the special questions in different application areas. A thriving part of behavioral economics is similar—the application of basic ideas to various subfields, or "franchising". Besides the areas discussed in more detail below, other franchises have been established in law (Jolls, Sunstein and Thaler, 1998; Jolls, in press) and development (Mullainathan, in press).

Finance: The central hypothesis in financial economics for the last thirty years is that stock markets are informationally efficient. Faith in this claim comes from a simple argument: Any semi-strong-form inefficiency (detectable using cheaply-acquired data) would be noticed by wealthy investors and erased. Market efficiency was therefore thought to provide a stiff challenge to models which assume investors have limited rationality. But "behavioral finance" based on rationality limits has emerged rapidly and might be the clearest empirical franchise success for behavioral economics (e.g., Barberis and Thaler, in press). One advantage is that theories of asset pricing often provide sharp predictions. Another big advantage is that there are many cheaply-available data which can be used to test theories.

Behavioral finance got its biggest early boost from DeBondt and Thaler's (1985) discovery that portfolios of "loser' stocks (stocks whose market value had dropped the most in the previous year) outperformed portfolios of winners in subsequent years. Their paper was published in the proceedings of the

Journal of Finance and immediately drew attention and counterargument. Note that DeBondt and Thaler *predicted* this anomaly, based on the idea that investors would be surprised by reversion to the mean in unusually high- and low-performing firms (an application of the "representativeness heuristic").

An important theoretical attack on market efficiency was showing that if investors have limited horizons (due to quarterly evaluation of institutional portfolio managers, for example) then even if prices wander away from fundamental values, investors might not have enough aggregative incentive to trade prices back to the fundamentals, which allows mispricing to persist. (As Keynes, noted, markets might stay irrational longer than you can stay liquid

Comment [JW1]:

A central point here is that an attack on the proposition that prices would fully reveal information caused the finance profession to carefully examine the microstructural and institutional reasons why such revelation might, or might not occur. So the behavioral critique, whether right or wrong, did lead to a sharper focus on institutional details, which eventually led to better financial economics.

A recent trend is extending some of these ideas to corporate finance— how companies raise and spend financing from capital markets (see Baker, Ruback, Wurgler 2004). Behavioral influences might be even stronger here than in asset pricing because large decisions are made by individuals or small groups, and discipline is only exerted by boards of directors, career concerns, sorting for talented decision makers, and so forth. So it is possible that very large corporate mistakes are made by a combination of limitedly-rational managers and weak governance.

An interesting feature of the evolution of academic finance is how some early behavioral ideas which were largely dismissed are now taken seriously. For example, Miller (1977) suggested that divergence of opinion, combined with restrictions on short-selling could lead to inflated stock valuations. Miller's paper was rarely cited at first, but the same idea was used, twenty-five years later, to explain the American dot-com bubble (Ofek and Richardson, 2003). Similarly, Modigliani and Cohn (1979) advanced the radical idea that stock market investors did not distinguish between nominal and inflation-adjusted ("real") rates of return. Decades later, their radical theory is consistent with tests by Cohen, Polk and Vuolteenaho (2005) and Campbell and Vuoltenaho (2004).

Game theory: Game theory is a taxonomy of canonical strategic interactions and a collection mathematical theories of how players with varying degrees of rationality are likely to play in games as they are perceived. Since many of the games are complicated, and equilibrium theories often assume a high degree of mutual rationality and complicated Bayesian inference, game theory is ripe for introduction of behavioral alternatives that weaken equilibrium assumptions in a disciplined way. Many theoretical papers have explored the implications of weakened assumptions of rationality. Many predictions of game theory depend delicately on what players commonly know and on assumptions about the utility derived from outcomes. As a result, experiment which carefully control strategies, information, and payoffs have been unusually helpful in clarifying conditions under which equilibrium predictions are likely to hold or not (Crawford, 1997; Camerer, 2003).

Two central contributions of behavioral game theory are worth mentioning. One is the study of limits on strategic thinking. One type of theory studies how finite automata that implement strategies with limited calculation and memory will behave (e.g., Rubinstein, 1998). Empirically-driven theories posit some distribution of steps of thinking (the cognitive hierarchy theories discussed in section III). The other important contribution is precise theories of how monetary payoffs to one player and others map onto the focal player's utility (also discussed in section III).

Behavioral game theory has largely been shaped by experimental observation of educated people playing games in experiments for money. Here, equilibrium predictions do not always fare well compared to learning theories, and to QRE and cognitive hierarchy approaches. But equilibrium theory might apply at other levels of analysis, especially low and high levels, such as animal behavior sculpted by evolution (e.g., optimal foraging), and decisions of firms and nation-states which are widely-deliberated and analyzed carefully.

Labor and organizational economics: Labor economics is certainly ripe for behavioral analysis (see Camerer and Malmendier, in press). Most workers do not have much chance to learn from experience before making important decisions with irreversibility— choosing education, and a first job that often determines a career track. The goods that workers sell—their time—is also likely to involve more social comparison, optimism, emotion and identity than when firms sell cars or iPods. In many cases, workers appear to care about a range of nonpecuniary incentives besides money, such as fair treatment and being appreciated.

Inside the firm, evaluation of worker performance is imperfect in all but the simplest organizations in which piece rates can be tied to individual productivity (like fruit-picking and car repair); imperfect evaluation leads to scope for biases in judgment. For example, hindsight bias—the tendency to think, ex post, that outcomes were more ex-ante predictable than they actually were—creates second-guessing and complicates implementation of the idealized contracts in agency theory.

Many experiments have studied reciprocity (or gift-exchange) in simple versions of labor markets. In the simplest case, firms prepay a wage and workers then choose effort which is costly for them but valuable for firms. If there is an excess supply of workers and no scope for reputation-building[14], self-interested workers should be happy to get jobs but should also shirk; firms should anticipate this and offer a minimum wage. Empirically, however, when effort is very valuable to firms and not too costly to workers, firms pay wages far above the minimum, and workers reciprocate by exerting more effort when they were paid a higher wage. When workers are identified to firms, and firms can repeatedly hire good workers, Brown, Falk and Fehr (2004) show how a "two-tier" insider-outsider economy can emerge experimentally.

[14] Healy (2004) shows that the amount of reciprocity by workers is sensitive to the shared gains from effort. Charness, Frechette and Kagel (2004) show that framing of the instructions can lower reciprocity. Healy also shows in a simple model how a perception of correlation of reciprocal worker types can induce gift exchange even when the wage-effort game is repeated only finitely. His important insight is that type correlation induces "group reputation".

Data like these are a reminder that intrinsic motivations like reciprocity matter and can be quite strong. Furthermore, adding extrinsic incentives can be harmful if they "crowd out" intrinsic incentives (a phenomenon long-studied in psychology), so that standard models get the sign wrong in predicting effects of extrinsic incentive changes. Benabou and Tirole (2003) approach crowding out in a different way. They show that higher incentives can induce lower effort because high wages signal that a job is very hard, or a worker is unskilled.

Public finance: Behavioral public finance asks how limits on consumer and voter rationality influence taxation and public spending. Two pioneering examples are Krishna and Slemrod (2003) and McCaffrey's (1994) paper on cognitive psychology and taxation. The central principle is that some taxes are more visible than others. Politicians exploit these differences in searching for ways to increase tax receipts. A full theory of taxation and spending therefore depends on a good account of which types of taxes are easy and hard to impose (well-organized interest group competition will matter too, of course), and how astute revenue-seeking politicians are at understanding investor tax psychology.

Behavioral public economics is also likely to be the franchise that most squarely confronts issues of welfare analysis in behavioral economics. In the standard theory, what consumers choose is taken as a tautological definition of welfare (i.e., if consumers are rational, then what they choose is also what is best for them). Thinking about psychology permits the possibility that private choices do not maximize welfare. For example, Berridge and Robinson (2003) suggest that separate brain areas control "wanting"—choice—and "liking"—hedonic evaluation. If liking is true welfare, then neural separability of these processes implies that it is possible for choice and welfare to be different. The obvious places to look are decisions by adolescents and addicts (Bernheim and Rangel,2004), and potential mistakes in rare decisions, or when it is difficult to learn from experience.

B. Formal foundations

The goal of behavioral economics is not just to create a list of anomalies. The anomalies are used to inspire and constrain formal alternatives to rational-choice theories. Many such theories have emerged in recent years; a few of them were mentioned in section III.

Tremendous progress has been made in going from deviations and anomalies to general theories which are mathematically and can be applied to make fresh predictions. The general theories that economists are justifiably proud of only emerged over many decades of careful attention and refinement. Behavioral economics theories will become refined, and more general and useful, now that it has attracted the attention of an army of smart theorists and graduate students.

Excluded from Table 1, and from the discussion in of basic ideas in section III, are a rapidly-emerging variety of formal "dual system" models, drawing on old dichotomies in psychology. These models generally retain optimization by one of the systems and make behavior of another system automatic (or myopic) and nonstrategic, so that extensions of standard tools can be used. (Intuitively, think of part of the brain as optimizing against a new type of constraint—an internal constraint from another brain system, rather than a budget constraint or an external constraint from competition.) In Kahneman (2003) the systems are intuitive and deliberative systems ("systems 1 and 2"). In Loewenstein and O'Donoghue, 2004) the systems are deliberative and affective; in Benhabib and Bisin (2005) the systems are controlled and automatic; in Fudenberg and Levine (2004) the systems are "long-run" (and controlling) and "short-run"; in Bernheim and Rangel (2005) the systems are "hot" (automatic) and "cold". In Brocas and Castillo (2005) a myopic "agent" system has private information about utility, so a farsighted "principal" (who cares about the utility of all agents) creates mechanisms for the myopic agents to reveal their information.

These models are more alike than they are different. In the years to come, careful thought will probably sharpen our understanding of the similarities and differences among models. More thought will probably point to more general formulations that include models like those above as special cases, narrowing the focus of attention. And of course, empirical work is needed to see which predictions of different models hold up best, presumably inspiring some refinements that might eventually lead to a single model which could occupy a central place in microeconomics.

Herbert Simon was a towering figure in the development of behavioral economics. Simon coined the terms "bounded rationality" and "procedural rationality" and sowed the seeds for the analyses of rationality bounds that are the substance of this paper. Despite the influence of Simon's language, he had in mind a style of theorizing that has not caught on in economics. Influenced by cognitive science and the information processing model of human decision making, Simon thought good theories might take the form of algorithms which describe the procedures that people and firms use.

The economist in modern times who carries Simon's methodological torch is Ariel Rubinstein (e.g., see his 1998 book). Rubinstein's models are often stylized to a particular economic application and describe the mathematical result of particular algorithms which embody rationality limits. While these models are widely-known, in many cases they have not led to a sustained program of research, as his seminal work on bargaining has. Rubinstein's frustration with inattention to models driven by similarity judgment, a central concept in psychology, is evident in his 2003 discussion of models of time preference.

C. Field studies

Many new studies look for the influences of rationality limits in naturally-occurring field data. A good example that highlights interest in time preference is Della Vigna and Malmendier's (2005) study of health club memberships. The health clubs they study allow people to spend a fixed sum for an annual membership, or pay for each visit separately. People who discount hyperbolically, but are "naïve" about their future hyperbolic preferences, will sign up for large-fee annual plans with per-visit fees that are below marginal cost (typically free). They find that even though per-visit fees average $10, the typical consumer who bought the annual-fee package ended up going rarely enough that the per-visit cost was $19. They also show theoretically that this contract is optimal for firms: Naïve hyperbolics like it because they misforecast how often they will go (they don't realize they are choosing a suboptimal contract), and "sophisticated" hyperbolic consumers like it because the low per-visit fee provides external self-control (which they know they will need).

An early example of a field study inspired by behavioral economics is Camerer, Babcock, Loewenstein and Thaler's (1997) study of cab driver labor supply. New York City cab drivers typically rent their cabs by the day, for a fixed fee, keep all the revenues they earn, and can drive up to 12 hours. The standard theory of upward-sloping labor supply, and intertemporal substitution, predicts that drivers will drive longer on high-wage days. But suppose drivers take a short horizon, e.g., one day at a time, and have an aspiration level or reference point they dislike falling short of (i.e., they are averse to a perceived revenue "loss" relative to their reference point or daily target). Myopic target-driven drivers will drive more hours on low-wage days, the opposite of the standard prediction. (This is a case where behavioral economics made a clear prediction of a new phenomenon, rather than just explaining an established anomaly.) Camerer et al found that inexperienced drivers appear to have a negative labor supply elasticity—they drove more hours on low-wage days—and the elasticity of experienced drivers was around zero. Farber (2004) replicated this study with a smaller fresh sample using a hazard rate model of hourly quitting decisions. He found no evidence of daily targeting in general and weak evidence for three of five drivers for whom there are a lot of data. A subsequent study (Farber, 2005) finds effects of targeting which are significant but small in magnitude.

Conlin, O'Donoghue and Vogelsang (2005) estimate how often items ordered from mail-order catalogues are returned. Their study is motivated by evidence of "projection bias"—the idea that one's current emotional state exerts too much influence on a projection of one's future state (e.g., people buy more groceries when they are hungry). They show theoretically that returns of cold-weather items (e.g., jackets or gloves) on a particular day depend on whether the return-day weather is warm, and also depend on weather the ordering-day weather was cold. (The intuition is that people who order on a cold day mistakenly forecast it will be equally cold in the future, so they are systematically surprised.) Their result is striking because people are well aware of seasonality in weather (most people can tell you whether a day is unseasonably warm or cold). It is not as they are misforecasting their tastes for exotic novelties like sea urchin or funnypunk music.

A booming and important area of field study is experimentation in field settings. Field experiments can range (Harrison and List, 2004) from abstract simple experiments done outside university labs, to measurement of treatment effects in field sites where those effects are of special interest (see Cardenas and Carpenter, 2005). These studies combine the value of measuring an effect directly in a population of interest with the gain from experimental control. The gain comes from randomized assignment of treatments, which avoids self-selection effects that are challenging to control econometrically in field data.[15]

D. Importing 'new' psychology

The workhorse models in Table 1 draw on a narrow range of cognitive psychology, mostly from decision research. Other psychological concepts, which are hardly new in psychology but new to economists, are starting to be applied as well (such as memory, see Wilson 2004).

Attention is perhaps the ultimate scarce cognitive resource. A few studies have started to explore its implications for economics. Odean and Barber (2005) show that attention-getting events—abnormal trading volumes or returns, or news events—correlate with purchases by individual investors. Della Vigna and Pollett (2005) find that markets react less to earnings announcements made on Fridays than on other days; firms seem to know this and are more likely to release bad news on a Friday. Falkinger (2005) develops a rich model in which firms must choose signal strength for their products to get the attention of consumers.

[15] Tanaka, Camerer and Nguyen (2005) is one study that measures multiple dimensions of time, risk and trust preferences corresponding to models in Table 1.

Attribution theory describes how people intuitively infer causes from effects. Many studies indicate systematic misattributions, such as the tendency to overattribute cause to personal actions rather than exogeneous structural features (Weber et al, 2001). For example, Bertrand and Mullainathan (2001) find that oil company executives are rewarded when oil prices go up, but are not penalized equally penalized when prices go down. Einav and Yariv (2005) note that authors of economics papers whose names come earlier in a list of authors benefit disproportionately by various measures, even though the order is almost always alphabetical.

Categorization refers to the way in which the brain forms categories. Mullainathan (2002) shows how categorization can generate non-Bayesian effects. An important property of categories is that likelihood evidence which is weak can tip interpretations from one category to another, producing large effects from small causes. Fryer and Jackson (2004) develop a model of optimal categorization and discuss its application to labor market discrimination.

E. Neuroeconomics

Neuroeconomics is the grounding of microeconomics in details of neural functioning. It is natural to be skeptical about whether economists need to know precisely where in the brain computations occur to make predictions about economic behavior such as responses to prices. But keep in mind that the revealed preferences approach which deliberately avoided "trying to discover the essence of things" (in Pareto's phrase) was adopted about a hundred years ago. At that time it really *was* impossible to make all the measurements and causal interventions that can be made today, with PET, TMS, MEG, pharmacological and hormone changes, genetic testing in all species and gene knockouts in mice (actually engineering the genes), and fMRI. The fact that there are so many tools means that limits of one method can be compensated for by strengths of other methods (they are complements). Technological substitution from 100 years ago to now suggest economists might learn something from these new measurements about choices.

Some basic facts about the brain can guide economic modeling (and already have, in "dual-process" models). The brain is divided into four lobes—frontal, parietal, occipital and temporal. Regions of these lobes are interconnected and create specialized "circuits" for performing various tasks.

The human brain is a primate brain with more neocortex. To deny this important fact is akin to creationism. The fact that many human and animal brain structures are shared means that human behavior generally involves interaction between "old" brain regions and more newly-evolved ones. The descent of humans from other species also means we might learn something about human behavior from other species. For example, rats become addicted to all drugs that humans become biologically addicted to, which implies that old reward circuitry shared by rat and human brains is part of human addiction.

While we often think of complex behavior as deliberate, resources for "executive function" or "cognitive control" are rather scarce (concentrated in the cingulate). As a result, the brain and body are very good at delegating components of complex behavior into automatic processes. For example, a student driver is overwhelmed by visual cues, verbal commands, memory required for navigation, and mastery of motor skills. Many accidents result during this learning process. But within a few years, driving becomes so effortless that drivers can eat and talk (perhaps on a cell phone) while driving safely.[16]

Methodologically, neuroeconomics is not intended to test economic theory in a traditional way (particularly under the view that utilities and beliefs are only revealed by choices). Instead, the goal is to establish the neural circuitry underlying economic decisions, for the eventual purpose of making better predictions.

Seen this way, neuroeconomics is likely to produce three types of findings: Evidence for rational-choice processes; evidence supporting behavioral economics processes and parameters (as in Table 1); and evidence of different types of constructs which do not fit easily into standard modeling categories.

[16] However, as activities become automatic, they often become harder to remember and difficult to teach to others, an important fact for the division of labor in large firms where learning-by-doing creates automaticity.

Results consistent with rational choice: In choice domains where evolution has had a long time to sculpt pan-species mechanisms that are crucial for survival (food, sex, and safety), neural circuits which approximate Bayesian rational choice have probably emerged. For example, Platt and Glimcher (1999) find neurons in monkey lateral intraparietal cortex (LIP) which fire at a rate that is almost perfectly correlated with the expected value of an upcoming juice reward, triggered by a monkey eye movement (saccade). Monkeys can also learn to approximate mixed-strategies in games, probably using generalized EWA-type reinforcement algorithms (Lee, McGreevy and Barraclough, 2005). Neuroscientists are also finding neurons that appear to express values of choices (Padoa-Schioppa and Assad, 2005) and potential locations of “neural currency” that create tradeoffs (Shizgal, 1999).

Results consistent with behavioral economics: Other neural evidence is already vaguely consistent with behavioral economics ideas like those in Table 1. McClure et al (2004) find evidence of two systems involved in time discounting, consistent with a quasi-hyperbolic β-δ theory. Sanfey et al (2003) find that low offers in ultimatum games (compared to near-equal offers) differentially activate emotional areas (insula), planning and evaluation areas (dorsolateral prefrontal cortex, DLPFC) and conflict resolution areas (anterior cingulate). Relative activity in the insula and DLPFC predicts whether offers will be rejected or not. This result is consistent with social preferences models in which money and distaste for unfairness or inequality are traded off (by the cingulate). Hsu et al (2005) compared decisions under ambiguity and risk (using Ellsberg-paradox examples). Ambiguity differentially activates the orbitofrontal cortex (OFC, just above the eye sockets) and the amygdala, a “vigilance” area which responds rapidly to fearful stimuli and is important in emotional processing and learning. The fact that OFC activity is stronger and longer-lasting for ambiguous choices implies that people with damage to the OFC might not exhibit typical patterns of ambiguity-aversion. Indeed, Hsu et al find that they do not.

New constructs and ideas: The biggest impact of neuroeconomics will probably not come from adjudicating debates between rational-choice and behavioral economics; it will come from establishing a detailed empirical basis for constructs which are new in economics (although some of them could be defined in familiar terms).

For example, in game theory players are in equilibrium when their beliefs about what other players will do are accurate, and they choose best responses given those accurate beliefs. A neural analogue of this mathematical is that brain activity in equilibrium will be highly overlapping when players are making their own choices, compared to when they are forming beliefs about choices of others, because creating accurate beliefs requires them to simulate choices by others. Indeed, Bhatt and Camerer (2005) found very little difference in brain activity between choosing and guessing in periods in which players' choices and beliefs were in equilibrium. Thus, game-theoretic equilibrium is a "state of mind" as well as a restriction on belief accuracy and best response.

***Causing* preferences:** *Some* areas in the brain are active during economic decision making. So what is learned from knowing precisely *where* those regions are? The answer is that regions develop at different rates across the life cycle, are different across species, use different neurotransmitters, have different types of neurons, and participate in decisions that might seem superficially different. (For example, the insula which is activated by low ultimatum offers, is also activated by bodily discomforts like pain and disgust; so when a person says an offer is "disgustingly low" they may be speaking rather literally.)

Knowing which regions are part of the neural circuit for a particular decision enables us to use other knowledge about specialization to make new types of predictions. Valuation of a good—a utility—, which is often thought of as basic preference, might actually be the *middle* phase of a biological process. Valuations are an *input* to a more complex downstream process which incorporates prices, budget constraint, and possibly social concerns (e.g., peer pressure or rational conformity). But valuations are also the *output* of an earlier upstream process, which should perhaps be considered the "primitive" in modeling preferences.

A behavioral way to demonstrate an understanding of the process that creates expressed preferences is to show how changing variables can cause or influence preferences. In standard economic terms, preferences are "state-dependent", where the states are internal biological states (that can also be changed exogeneously). Then the important questions are: What are those states? And does an executive cortical process understand how the state-dependence works, and influences it or compensates for exogeneous shocks?

For example, the oxytocin hormone is involved in social bonding and is implicated in studies of trust games (Zak et al, 2005). It follows that if oxytocin can be increased exogeneously, and the brain does not undo the effect of the exogeneous change, then adding oxytocin might *create* trust. Kosfeld et al (2005) showed exactly this effect. They administered synthetic oxytocin to subjects, which increased the amount those subjects invested in a trust game. The capacity to change behavior (traditionally interpreted as revelation of preferences) is routine for neuroscientists. Direct stimulation of single neurons is conjectured to create preferences for one choice or another, by intervening upstream.

This approach suggests a general recipe for causing changes in behavior. As noted earlier in section B, most dual-process models posit two processes: (1) A controlled, long-run, deliberative, or "cold" process which accepts inputs and tries to constrain or override another (2) process which is automatic, short-run, affective, or "hot". The recipe for changing behavior is to either stimulate the second process directly, and see whether the first type of deliberative process undoes the exogeneous change, or to place cognitive overload on the first process (tying up its scarce resources) and see whether its ability to constrain the second process suffers. Lerner, Small and Loewenstein (2004) stimulate the second process. They induced emotional states which affected how people priced goods they were endowed with (reversing the typical "endowment effect" in which owned goods are valued more highly). Shiv and Fedorikhin (1999) constrained the first (controlled) process. They asked subjects remember either simple (2-digit) or difficult (7-digit) strings of numerical digits as they walked by foods that were tempting (potato chips) or virtuous (fruit). Overloading the controller system with the more taxing 7-digit memory task led to more consumption of the tempting foods. The simplest language of preference theory would say that the difficult 7-digit memory task “changed preferences”. A more detailed view, and a more useful one, is that resistance to temptation requires scarce cognitive resources; multitasking which consume these resources lowers resistance and leads people to eat more chips.

VI. Conclusions

Empirically-driven behavioral economics uses evidence from psychology and other disciplines to inform models of limits on rationality, willpower and self-interest, to explain anomalies and make new predictions. This approach deliberately rejects the “F-twist” premise that theories should not be judged by their assumptions, on the grounds that models based on more realistic assumptions will make better predictions.

Many concepts have already been proposed, which generally add one or more parameters to models of choice, including risk, ambiguity and time (Table 1).

This essay highlights a few areas of active research. A central question is the market implications of limits on rationality, willpower, and self-interest. While experience and sorting might weaken the impact of limited individual rationality on firm behavior, these firms also supply goods to a demand side of the market where institutional and social forces are not as strong as erasing the effects of limits. Whether market forces therefore limit the impact of mistakes, or exaggerate them (by creating hyper-rational firms that are optimized to exploit consumer limits) is therefore an open question.

Important trends in behavioral economics including “franchising” of ideas to application areas (such as finance and labor economics), development of theoretical models, field studies, and including new types of psychology (such as attention, attribution, categorization, and limited memory). Another small emerging field is "neuroeconomics", a subfield of behavioral economics which uses details of neural activity to inform microfoundations. Some of these studies are likely to show neural evidence consistent with rational choice, others have already shown circuitry consistent with behavioral economics constructs, and still others will point to constructs that indicate state-dependence of preference (where the states are internal brain states).

References

Anderson, M. J. and S. Sunder (1995). "Professional Traders as Intuitive Bayesians." Organizational Behavior and Human Decision Processes 64(2): 185-202.

Angeletos, G.-M., D. Laibson, A. Repetto, J. Tobacman and S. Weinberg (2001). "The Hyperbolic Consumption Model: Calibration, Simulation, and Empirical Evaluation." Journal of Economic Perspectives 15(3): 47-68.

Arrow, K. J. (1963). Social Choice and Individual Values. New Haven, Yale University Press.

Arthur, W. B. (1991). "Designing Economic Agents that Act like Human Agents: A Behavioral Approach to Bounded Rationality." The American Economic Review 81(2): 353-359.

Ashraf, N., C. F. Camerer and G. Loewenstein (2005). "Adam Smith, Behavioral Economist." Journal of Economic Perspectives 19(3): 131-145.

Baker, M. P., R. S. Ruback and J. Wurgler (2004). Behavioral Corporate Finance: A Survey, NBER.

Barberis, N., A. Shleifer and R. Vishny (1998). "A model of investor sentiment." Journal of Financial Economics 49(3): 307-343.

Barberis, N. and R. Thaler (in press). A Survey of Behavioral Finance. Handbook of the Economics of Finance. G. Constantinides, M. Harris and R. Stulz.

Bateman, I., D. Kahneman, A. Munro, C. Starmer and S. Robert (2005). "Is There Loss Aversion in Buying? An Adversarial Collaboration." Journal of Public Economics 89(8): 1561-1580.

Bateman, I., A. Munro, B. Rhodes, C. Starmer and R. Sugden (1997). "A Test of the Theory of Reference-Dependent Preferences." Quarterly Journal of Economics 112(2): 479-505.

Benabou, R. and J. Tirole (2003). "Intrinsic and Extrinsic Motivation." Review of Economic Studies 70: 489-520.

Benartzi, S. and R. H. Thaler (1995). "Myopic Loss Aversion and the Equity Premium Puzzle." The Quarterly Journal of Economics 110(1): 73-92.

Benhabib, J. and A. Bisin (2005). "Modeling internal commitment mechanisms and self-control: A neuroeconomics approach to consumption-saving decisions." Games and Economic Behavior 52(2): 460-492.

Benjamin, D. J. and J. M. Shapiro (2005). Who is “Behavioral”? Cognitive Ability and Anomalous Preferences, University of Chicago.

Bernheim, B. D. and A. Rangel (2004). "Addiction and cue-triggered decision processes." American Economic Review 94(5): 1558-1590.

Bernheim, B. D. and A. Rangel (2005). Behavioral Public Economics: Welfare and Policy Analysis With Fallible Decision-Makers. Economic Institutions and Behavioral Economics. P. Diamond and H. Vartiainen. Princeton, Princeton University Press.

Berridge, K. C. and T. E. Robinson (2003). "Parsing Reward." Trends in Neurosciences 26(9): 507-513.

Bertrand, M. and S. Mullainathan (2001). "Are CEOs Rewarded for Luck? The Ones Without Principals Are." Quarterly Journal of Economics 116(3): 901-932.

Bertrand, M. and S. Mullainathan (2004). "Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination." American Economic Review 94(4): 991-1013.

Bhatt, M. and C. F. Camerer (2005). "Self-referential thinking and equilibrium as states of mind in games: fMRI evidence." Games and Economic Behavior 52(2): 424-459.

Bolton, G. E. and A. Ockenfels (2000). "ERC: A Theory of Equity, Reciprocity, and Competition." American Economic Review 90(1): 166-193.

Brocas, I. and J. Carrillo (2005). The Brain as a Hierarchical Organization, University of Southern California.

Brown, M., A. Falk and E. Fehr (2004). "Relational contracts and the nature of market interactions." Econometrica 72(3): 747-780.

Bruni, L. and R. Sugden (2005). "The Road Not Taken. Two Debates on Economics and Psychology?" Economic Journal forthcoming.

Camerer, C., L. Babcock, G. Loewenstein and R. Thaler (1997). "Labor Supply of New York City Cabdrivers: One Day at a Time." The Quarterly Journal of Economics 112(2, In Memory of Amos Tversky (1937-1996)): 407-441.

Camerer, C. and T. H. Ho (1999). "Experience-weighted attraction learning in normal form games." Econometrica 67(4): 827-874.

Camerer, C., G. Loewenstein and M. Weber (1989). "The Curse of Knowledge in Economic Settings - an Experimental-Analysis." Journal of Political Economy 97(5): 1232-1254.

Camerer, C. F. (1987). "Do Biases in Probability Judgment Matter in Markets - Experimental-Evidence." American Economic Review 77(5): 981-997.

Camerer, C. F. (2003). Behavioral game theory: Experiments on strategic interaction. Princeton, Princeton University Press.

Camerer, C. F. and E. Fehr (2006). "When does economic man dominate social interaction?" Science forthcoming.

Camerer, C. F., T. H. Ho and J. K. Chong (2004). "A cognitive hierarchy model of games." Quarterly Journal of Economics 119(3): 861-898.

Camerer, C. F., G. F. Loewenstein and M. Rabin (2004). Advances in Behavioral Economics. Princeton, Princeton University Press.

Camerer, C. F. and U. Malmendier (in press). Behavioral Organizational Economics. Yrjo Jahnsson Foundation 50th Anniversary Conference on Economics Institutions and Behavioral Economics. P. Diamond and H. Vartiainen. Princeton, Princeton University Press.

Campbell, J. Y. and T. Vuolteenaho (2004). "Inflation Illusion and Stock Prices." American Economic Review 94(2): 19-23.

Cardenas, J. C. and J. P. Carpenter (2005). Experiments and Economic Development: Lessons from field labs in the developing world, Middlebury College.

Charness, G., G. R. Frechette and J. H. Kagel (2004). "How robust is laboratory gift exchange?" Experimental Economics 7(2): 189-205.

Charness, G. and M. Rabin (2002). "Understanding social preferences with simple tests." Quarterly Journal of Economics 117(3): 817-869.

Chen, K., V. Lakshminarayanan and L. Santos (2005). How Basic are Behavioral Biases? Evidence from Capuchin-Monkey Trading Behavior, Yale University.

Chua, Z. and C. F. Camerer (2004). Experiments on Intertemporal Consumption with Habit Formation and Social Learning., Caltech.

Cochrane, J. (1989). "The Sensitivity of Tests of the Intertemporal Allocation of Consumption to Near-Rational Alternatives." American Economic Review 79(3): 319-337.

Cohen, R. B., C. Polk and T. Vuolteenaho (2005). "Money Illusion in the Stock Market: The Modigliani-Cohn Hypothesis." Quarterly Journal of Economics 120(2): 639-668.

Colander, D. (2005). Neuroeconomics, the Hedonimeter, and Utility: Some Historical Links, Middlebury College.

Conlin, M., T. O'Donoghue and T. Vogelsang (2005). Projection bias in catalog orders, Cornell University.

Conlisk, J. (1996). "Why Bounded Rationality?" Journal of Economic Literature 34(2): 669-700.

Costa-Gomes, M., V. P. Crawford and B. Broseta (2001). "Cognition and behavior in normal-form games: An experimental study." Econometrica 69(5): 1193-1235.

Coy, P. (2005). Why Logic Often Takes A Backseat BusinessWeek. March 28, 2005.

Crawford, V. P. (1997). Theory and experiment in the analysis of strategic interaction. Advances in economics and econometrics: Theory and applications. D. Kreps and K. Wallis. Cambridge, Cambridge University Press: 1-52.

Debondt, W. F. M. and R. Thaler (1985). "Does the Stock-Market Overreact." Journal of Finance 40(3): 793-805.

Degeorge, F., J. Patel and R. Zeckhauser (1999). "Earnings Management to Exceed Thresholds." Journal of Business 72(1): 1-33.

Della Vigna, S. and U. Malmendier (2005). "Paying Not To Go To The Gym." American Economic Review forthcoming.

Della Vigna, S. and J. Pollet (2005). Investor Inattention, Firm Reaction, and Friday Earnings Announcements, University of California, Berkeley.

Dufwenberg, M. and G. Kirchsteiger (2004). "A Theory of Sequential Reciprocity." Games and Economic Behavior 47: 268-298.

Edgeworth, F. Y. (1881). Mathematical Psychics: An Essay on the Application of Mathematics to the Moral Sciences.

Einav, L. and L. Yariv (2005). "What's in a Surname? The Effects of Surname Initials on Academic Success." Journal of Economic Perspectives forthcoming.

Einhorn, H. J. (1982). Learning from Experience and Suboptimal Rules in Decision Making. Judgement under Uncertainty: Heuristics and Biases. D. Kahneman, P. Slovic and A. Tversky, Cambridge University Press.

Erev, I., Y. Bereby-Meyer and A. E. Roth (1999). "The effect of adding a constant to all payoffs: experimental investigation, and implications for reinforcement learning models." Journal of Economic Behavior & Organization 39(1): 111-128.

Erev, I. and A. E. Roth (1998). "Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria." American Economic Review 88(4): 848-881.

Falk, A. and U. Fischbacher (2005). "A Theory of Reciprocity." Games and Economic Behavior in press.

Falkinger, J. (2005). Limited Attention as the Scarce Resource in an Information-Rich Economy, IZA Discussion Papers.

Farber, H. (2004). Reference-Dependent Preferences and Labor Supply: The Case of New York City Taxi Drivers, Princeton University.

Farber, H. (2005). "Is Tomorrow Another Day? The Labor Supply of New York City Cabdrivers." Journal of Political Economy 113(1): 46-82.

Fehr, E. and K. M. Schmidt (1999). "A Theory of Fairness, Competition, and Cooperation." Quarterly Journal of Economics 114(3): 817-868.

Fischbacher, U., C. M. Fong and E. Fehr (2003). Fairness, Errors and the Power of Competition, IEW Working Paper No. 133, University of Zurich 2003.

Fischhoff, B. and R. Beyth (1975). "I Knew it Would Happen: Remembered Probabilities of Once-Future Things." Organizational Behavior and Human Performance 13: 1-16.

Frederick, S. (2005). "Cognitive Reflection and Decision Making." J. Economic Perspectives 19(4): 25-42.

Fryer, R. and M. Jackson (2004). A Categorical Model of Cognition and Biased Decision-Making, California Institute of Technology.

Fudenberg, D. and D. Levine (1998). Theory of Learning in Games. Cambridge, MA, MIT Press.

Fudenberg, D. and D. Levine (2004). A Dual Self Model of Impulse Control, Harvard.

Gabaix, X. and D. Laibson (2006). "Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets." Quarterly Journal of Economics 121(2): forthcoming.

Ganguly, A. R., J. H. Kagel and D. V. Moser (2000). "Do asset market prices reflect traders' judgment biases?" Journal of Risk and Uncertainty 20(3): 219-245.

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

Goeree, J. K. and C. A. Holt (2001). "Ten Little Treasures of Game Theory and Ten Intuitive Contradictions." American Economic Review 91(5): 1402-1422.

Grether, D. M. and C. R. Plott (1979). "Economic Theory of Choice and the Preference Reversal Phenomenon." The American Economic Review 69(4): 623-638.

Grubb, M. (2005). Screening Overconfident Consumers, Stanford University.

Gul, F. (1991). "A Theory of Disappointment Aversion." Econometrica 59(3): 667-686.

Gul, F. and W. Pesendorfer (2001). "Temptation and Self-Control." Econometrica 69(6): 1403-1435.

Gul, F. and W. Pesendorfer (2005). The Case for Mindless Economics, Princeton University.

Haltiwanger, J. and M. Waldman (1989). "Limited Rationality and Strategic Complements: The Implications for Macroeconomics." Quarterly Journal of Economics 104(3): 463-483.

Hardie, B. G. S., E. J. Johnson and P. S. Fader (1993). "Modeling Loss Aversion and Reference Dependence Effects on Brand Choice." Marketing Science 12(4): 378-394.

Harless, D. W. and C. F. Camerer (1994). "The Predictive Utility of Generalized Expected Utility Theories." Econometrica 62(6): 1251-1289.

Harrison, G. W. and J. A. List (2004). "Field Experiments." Journal of Economic Literature 42(4): 1009-1055.

Healy, P. J. (2004). Group Reputations and Stereotypes as a Contract Enforcement Device, Caltech.

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

Hey, J. D. and C. Orme (1994). "Investigating Generalizations of Expected Utility Theory Using Experimental Data." Econometrica 62(6): 1291-1326.

Hines, J., James R. and R. Thaler (1995). "Anomalies: The Flypaper Effect." Journal of Economic Perspectives 9(4): 217-226.

Ho, T., C. F. Camerer and J.-K. Chong (2005). The Economics of Learning Models: A Self-tuning Theory of Learning in Games, Caltech.

Ho, T. H., N. Lim and C. F. Camerer (2005). Modeling the Psychology of Consumer and Firm Behavior with Behavioral Economics, Caltech.

Ho, T.-H. and J.-J. Zhang (2004). Does the Format of Pricing Contracts Matter?, UC- Berkeley.

Hsu, M., M. Bhatt, R. Adolphs, D. Tranel and C. F. Camerer (2005). "Neural Systems Responding to Degrees of Uncertainty in Human Decision-Making." Science 310: 1680-1683.

Jolls, C., C. R. Sunstein and R. Thaler (1998). "A behavioral approach to law and economics." Stanford Law Review 50(5): 1471-1550.

Jolls, C. M. (in press). Behavioral Law and Economics. Yrjo Jahnsson Foundation 50th Anniversary Conference on Economics Institutions and Behavioral Economics. P. Diamond and H. Vartiainen. Princeton, Princeton University Press.

Kahneman, D. (2003). "Maps of bounded rationality: Psychology for behavioral economics." American Economic Review 93(5): 1449-1475.

Kahneman, D., J. L. Knetsch and R. H. Thaler (1990). "Experimental Tests of the Endowment Effect and the Coase Theorem." Journal of Political Economy 98(6): 1325-1348.

Kahneman, D. and A. Tversky (1979). "Prospect Theory - Analysis of Decision under Risk." Econometrica 47(2): 263-291.

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

Kluger, B. D. and S. B. Wyatt (2004). "Are Judgment Errors Reflected in Market Prices and Allocations? Experimental Evidence Based on the Monty Hall Problem." Journal of Finance 59(3): 969-998.

Kosfeld, M., M. Heinrichs, P. J. Zak, U. Fischbacher and E. Fehr (2005). "Oxytocin increases trust in humans." Nature 435(7042): 673-676.

Koszegi, B. and P. Heidhues (2005). The Impact of Consumer Loss Aversion on Pricing, University of California, Berkeley.

Koszegi, B. and M. Rabin (forthcoming). "A Model of Reference-Dependent Preferences." Quarterly Journal of Economics forthcoming.

Krishna, A. and J. Slemrod (2003). "Behavioral Public Finance: Tax Design as Price Presentation." International Tax and Public Finance 10(2): 189-203.

Laibson, D. (1997). "Golden Eggs and Hyperbolic Discounting." The Quarterly Journal of Economics 112(2, In Memory of Amos Tversky (1937-1996)): 443-477.

Lazear, E. P., U. Malmendier and R. Weber (2005). Sorting in Experiments with Application to Social Preferences, Stanford University.

Lee, D., B. P. McGreevy and D. J. Barraclough (2005). "Learning and decision making in monkeys during a rock-paper-scissors game." Cognitive Brain Research 25(2): 416-430.

Lerner, J., D. A. Small and G. Loewenstein (2004). "Heart Strings and Purse Strings: Carryover Effects of Emotions on Economic Decisions " Psychological Science 15(5): 337-341.

Levine, D. K. (1998). "Modeling Altruism and Spitefulness in Experiments." Review of Economic Dynamics 1(3): 593-622.

List, J. A. (2003). "Does market experience eliminate market anomalies?" Quarterly Journal of Economics 118(1): 41-71.

Loewenstein, G. and T. O'Donoghue (2004). Animal Spirits: Affective and Deliberative Influences on Economic Behavior, Carnegie Mellon University.

Loewenstein, G. F., M. H. Bazerman and L. Thompson (1989). "Social Utility and Decision-Making in Interpersonal Contexts." Journal of Personality and Social Psychology 57(3): 426-441.

Loomes, G., C. Starmer and R. Sugden (1989). "Preference Reversal: Information-Processing Effect or Rational Non-Transitive Choice?" Economic Journal 99(395): 140-151.

Lucas, R. E., Jr. (1986). "Adaptive Behavior and Economic Theory." Journal of Business 59(4, Part 2: The Behavioral Foundations of Economic Theory): S401-S426.

Maciejovsky, B. and D. V. Budescu (2005). Is Cooperation Necessary? Learning and Knowledge Transfer in Cooperative Groups and Competitive Auctions, Univ Illinois, Urbana-champaign.

Marcet, A. and J. P. Nicolini (2003). "Recurrent Hyperinflations and Learning." American Economic Review 93(5): 1476-1498.

McCaffery, E. J. (1994). "Cognitive Theory and Tax." Ucla Law Review 41(7): 1861-1947.

McClure, S. M., D. I. Laibson, G. Loewenstein and J. D. Cohen (2004). "Separate Neural Systems Value Immediate and Delayed Monetary Rewards." Science 306(5695): 503-507.

McFadden, D. L. (1999). "Rationality for Economists?" Journal of Risk and Uncertainty 19(1-3): 73-105.

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

Miller, E. M. (1977). "Risk, Uncertainty, and Divergence of Opinion." Journal of Finance 32(4): 1151-1168.

Modigliani, F. and R. Cohn (1979). "Inflation, rational valuation, and the market." Financial Analysts Journal 35(3): 24-44.

Mukerji, S. and J. M. Tallon (2004). An Overview of Economic Applications of David Schmeidler's Models Of Decision Making Under Uncertainty. Uncertainty in Economic Theory: A Collection of Essays in Honor of David Schmeidler's 65th Birthday. I. Gilboa, Routledge.

Mullainathan, S. (2002). Thinking Through Categories, MIT.

Mullainathan, S. (in press). Psychology and Development Economics. Yrjo Jahnsson Foundation 50th Anniversary Conference on Economics Institutions and Behavioral Economics. P. Diamond and H. Vartiainen. Princeton, Princeton University Press.

Mullainathan, S. and R. Thaler (2000). Behavioral Economics Entry in International Encyclopedia of the Social and Behavioral Sciences. International Encyclopedia of the Social and Behavioral Sciences, Massachusetts Institute of Technology.

Nagel, R. (1995). "Unraveling in Guessing Games: An Experimental Study." The American Economic Review 85(5): 1313-1326.

Odean, T. (1998). "Are Investors Reluctant to Realize Their Losses?" Journal of Finance 53(5): 1775-1798.

Odean, T. and B. M. Barber (2005). All that Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors, University of California, Berkeley.

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

Ofek, E. and M. Richardson (2003). "DotCom Mania: The Rise and Fall of Internet Stock Prices." Journal of Finance 58(3): 1113-1138.

Padoa-Schioppa, C. and J. Assad (2005). Neuronal Processing of Economic Value in Orbitofrontal Cortex, Harvard Medical School.

Platt, M. L. and P. W. Glimcher (1999). "Neural correlates of decision variables in parietal cortex." Nature 400(6741): 233-238.

Prelec, D. (1998). "The Probability Weighting Function." Econometrica 66(3): 497-527.

Rabin, M. (1993). "Incorporating Fairness into Game-Theory and Economics." American Economic Review 83(5): 1281-1302.

Rabin, M. (1998). "Psychology and Economics." Journal of Economic Literature 36(1): 11-46.

Rabin, M. (2002). "Inference by believers in the law of small numbers." Quarterly Journal of Economics 117(3): 775-816.

Rabin, M. and J. L. Schrag (1999). "First Impressions Matter: A Model of Confirmatory Bias." Quarterly Journal of Economics 114(1): 37-82.

Robson, A. J. (2001). "The Biological Basis of Economic Behavior." J. Economic Literature 39(1): 11-33.

Rotemberg, J. J. (2004). Minimally Acceptable Altruism and the Ultimatum Game, Harvard Business School.

Rubinstein, A. (1998). Modelling Bounded Rationality. Cambridge, MA, MIT Press.

Rubinstein, A. (2003). ""Economics and Psychology"? The Case of Hyperbolic Discounting." International Economic Review 44(4): 1207-1216.

Russell, T. and R. Thaler (1985). "The Relevance of Quasi Rationality in Competitive Markets." American Economic Review 75(5): 1071-1082.

Russell, T. and R. Thaler (1987). "The Relevance of Quasi Rationality in Competitive Markets: Reply." American Economic Review 77(3): 499-501.

Samuelson, L. (1997). Evolutionary Games and Equilibrium Selection. Cambridge, MIT Press.

Sanfey, A. G., J. K. Rilling, J. A. Aronson, L. E. Nystrom and J. D. Cohen (2003). "The Neural Basis of Economic Decision-Making in the Ultimatum Game." Science 300(5626): 1755-1758.

Schlag, K. H. (1998). "Why Imitate, and If So, How?, : A Boundedly Rational Approach to Multi-armed Bandits." Journal of Economic Theory 78(1): 130-156.

Schmeidler, D. (1989). "Subjective Probability and Expected Utility without Additivity." Econometrica 57(3): 571-587.

Segal, U. and A. Spivak (1990). "First Order Versus Second Order Risk Aversion." Journal of Economic Theory 51(1): 111-125.

Shafir, E., D. N. Osherson and E. E. Smith (1989). "An Advantage Model of Choice." Journal of Behavioral Decision Making 2(1): 1-23.

Shiv, B. and A. Fedorikhin (1999). "Heart and mind in conflict: The interplay of affect and cognition in consumer decision making." Journal of Consumer Research 26(3): 278-292.

Shizgal, P. (1999). On the Neural Computation of Utility: Implications from Studies of Brain Reward. Well-Being: The Foundations of Hedonic Psychology. D. Kahneman, E. Diener and N. Schwarz: 502-526.

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

Simonson, I. and A. Tversky (1992). "Choice in Context - Tradeoff Contrast and Extremeness Aversion." Journal of Marketing Research 29(3): 281-295.

Slovic, P. and S. Lichtenstein (1968). "Importance of Variance Preferences in Gambling decisions." Journal of Experimental Psychology 78: 646-654.

Smith, A. (1759/1892). The Theory of Moral Sentiments. Indianapolis, Liberty Fund.

Stahl, D., O. and P. Wilson, W. (1995). "On Players' Models of Other Players: Theory and Experimental Evidence." Games and Economic Behavior 10(1): 218-254.

Stahl, D. O. (2000). "Rule Learning in Symmetric Normal-Form Games: Theory and Evidence." Games and Economic Behavior 32(1): 105-138.

Starmer, C. (2000). "Developments in Non-Expected Utility Theory: The Hunt for a Descriptive Theory of Choice under Risk." Journal of Economic Literature 38(2): 332-382.

Starmer, C. (2004). Friedman's Risky Methodology, University of Nottingham.

Stewart, S. A. (2005). Can Behavioral Economics Save Us from Ourselves? U. Chicago Magazine. 97.

Stigler, G. (1981). Economics or ethics? Tanner Lectures on Human Values. S. McMurrin. Cambridge, Cambridge University Press.

Tanaka, T., C. F. Camerer and Q. Nguyen (2005). Politics, Poverty and Preferences: Field Experiments and Survey Data from Vietnam, California Institute of Technology.

Thaler, R. (1980). "Toward a Positive Theory of Consumer Choice." J. Econ. Behavior & Org. 1(1): 39-60.

Thaler, R. H. (1999). "Mental accounting matters." Journal of Behavioral Decision Making 12(3): 183-206.

Thompson, L. and G. Loewenstein (1992). "Egocentric Interpretations of Fairness and Interpersonal Conflict." Organizational Behavior and Human Decision Processes 51(2): 176-197.

Tovar, P. (2004). The Effects of Loss-aversion on Trade Policy and the Anti-trade Bias Puzzle, U. Maryland.

Tversky, A. and D. Kahneman (1992). "Advances in Prospect-Theory - Cumulative Representation of Uncertainty." Journal of Risk and Uncertainty 5(4): 297-323.

Weber, M. and C. F. Camerer (1998). "The disposition effect in securities trading: an experimental analysis." Journal of Economic Behavior & Organization 33(2): 167-184.

Weber, R., C. Camerer, Y. Rottenstreich and M. Knez (2001). "The illusion of leadership: Misattribution of cause in coordination games." Organization Science 12(5): 582-598.

Weibull, J. (1995). Evolutionary Game Theory. Cambridge, MA, MIT Press.

Wilson, A. (2004). Bounded Memory and Biases in Information, University of Chicago.

Zak, P. J., K. Borja, W. T. Matzner and R. Kurzban (2005). "The Neuroeconomics of Distrust: Sex Differences in Behavior and Physiology." American Economic Review 95(2): 360-363.

Table 1: Some rational-choice principles and behavioral economics alternatives

Rational (or simplifying) assumption	**Behavioral alternative model**	**Representative citation**
Complete preferences		
description-invariance	Framing, reference-dependence	Kahneman-Tversky 1979 Koszegi-Rabin (in press)
procedure-invariance	Contingent weighting	Slovic-Lichtenstein, 1968, Grether-Plott 1979
context-independence	Comparative utility	Tversky-Simonson 1992
separable u(x)	Regret, disappointment	Loomes- Starmer-Sugden 1989, Gul 1991
Choice over		
Risk	Prospect theory	Kahneman-Tversky 1979
Ambiguity	Nonadditive decision weight	Schmeidler 1989
Time	Hyperbolic β-δ discounting	Laibson 1997
Self-interest	Inequality-aversion, fairness	Rabin, 1993, Fehr-Schmidt 1999
Bayesian judgment	Overconfidence	Odean 1998
	Encoding bias	Rabin-Schrag 1999
Equilibrium	Learning	Erev-Roth 1998 Camerer-Ho 1999
	Quantal response, cognitive hierarchy	McKelvey-Palfrey, 1998, Camerer-Ho-Chong 2004

Table 2: Evidence of loss-aversion from different studies using field and experimental data

Economic domain	**citation(s)**	**Type of data**	**Estimated λ**
Instant endowment effects for goods	Kahneman-Knetsch-Thaler (1990)	Field data (survey), goods experiments	2.29
Choices over money gambles	Kahneman and Tversky (1992)	Choice experiments	2.25
Asymmetric price elasticities for consumer product increases & decreases	Putler (1992), Hardie-Johnson-Fader (1993)	Consumer purchases (supermarket scanner data)	2.40, 1.63
Loss-aversion for goods relative to money	Bateman et al (2005)	Choice experiments	1.30
Loss-aversion relative to initial seller "offer"	Chen, Lakshminarayanan, Santos (2005)	Capuchin monkeys trading tokens for stochastic food rewards	2.70
Reference-dependence in two-part distribution channel pricing	Ho and Zhang (2004)	Bargaining experiments	2.71
Aversion to losses from international trade	Tovar (2004)	Non-tariff trade barriers, US 1983	1.95-2.39
Surprisingly few announcements of negative EPS and negative year-to-year EPS changes	DeGeorge-Patel-Zeckhauser (1999)	Earnings per share (EPS) changes from year to year for US firms	
Disposition effects in housing	Genesove and Mayer (2001)	Boston condo prices 1990-97	
Disposition effects in stocks	Odean (1998)	Individual investor stock trades	
Disposition effects in stocks	Weber and Camerer (1998)	Stock trading experiments	

Daily income targeting by NYC cab drivers	Camerer-Babcock-Loewenstein-Thaler (1997)	Dailyhours-wages observations (three data sets)	
Equity premium puzzle	Benartzi-Thaler (1995)	US stock returns	
Consumption: Aversion to period utility loss	Chua and Camerer (2004)	Savings-consumption experiments	