The behavioral challenge to economics: Understanding normal people

Colin F. Camerer
Div HSS 228-77
Caltech, Pasadena CA 91125
camerer@hss.caltech.edu

June 4, 2003. This paper was prepared for the Federal Reserve of Boston meeting on "How Humans Behave", June 8-10, 2003. Comments welcome, especially before June 8 (meeting date). Note that references are to be added.

Economic behavior begins and ends in the brain. Like economics, astronomy is about complex systems composed of simpler components— double stars, solar systems, planets and their satellites. While astronomical theories abstract from detailed descriptions of stars and planets, these theories are also sharply constrained by lower-level concepts from chemistry and physics, like Newtonian mechanics and gravity. In a similar way, economics should be constrained by facts about how people think, feel and behave. Behavioral economics takes this constraint seriously-- *very* seriously—and grounds economic models in psychological regularity.

This essay is about what behavioral economics has established, what the new research frontiers are, and what can be said about welfare and consequently about policy.

The idea that people are boundedly rational has been around for a long time (e.g., Herb Simon) and is not in genuine dispute. Since Simon defined bounds on rationality as the antithesis of hyperrationality, and hyperrationality was never taken seriously as a cognitive model, the concept of bounded rationality should not be controversial. The debate is therefore not about whether people are hyper-rational or not. The debate is about precisely how ideas from psychology can inform economic models of savings, unemployment, consumer demand, market-clearing, organizational design, financial market fluctuations, economic growth, and so on. Furthermore, rational outcomes can be the limiting behavior of systems in which people (or organizations) learn, imitate, and seek advice, which are all psychological concepts originating in the brain.

Behavioral economists have been actively and carefully researching how to usefully complicate models based on rationality, by incorporating Simonian bounds, for more than a decade or two. (I'll speak presumptuously on behalf of "the field" starting here, and throughout, since the conference organizers asked me to.[i])

The growth rate of behavioral economics, as a serious alternative to models rooted in strong rationality, has been more like a rocket launch, building up a powerful thrust for years and suddenly taking off vertically, than a smooth ascent of an airplane. Now we are in takeoff phase and aren't sure how far the rocket can go or in which direction. But the flight is sure to change economics profoundly and for the better.

Behavioral economists believe in the usefulness of a few central principles. We know that people value changes from reference points (but know less about what points of reference are, and how multiple gains and losses from reference points are integrated), that expressed preferences are constructed much as people heuristically approximate solutions to problems (but have no parsimonious alternative to utility-maximization so far), that categories of wealth are not added up when people make spending decisions (but would like to know more about the detailed cognitive structure of their "mental accounting" of wealth categories). These distinctions provide a new set of questions that will be answered, eventually, in a phase of behavioral economics that should be (happily) Kuhnian "normal science".

A collateral contribution of the behavioral economics enterprise is to remind economists how little is truly known about the basic facts needed to shape policy. Is privatizing Social Security good? It depends on whether forcing saving is better than allowing unfettered choice among a wide range of choices (perhaps too many); we don't know. The fact that widely-used graduate textbooks "prove" that more choice is preferred to less— merely by assuming that preferences are complete-- should not be taken too seriously as a guide for policy without further corroboration and thought.

Does current drug policy correctly internalize public externalities from drug use, or perhaps implement a nearly-harmless paternalism, and balance social cost and benefit? We don't know, absent a sensible theory of choice about substances that people become addicted to.

These examples suggest that behavioral economics can also contribute much more when a dialogue with policymakers is created which poses specific questions that can answered by what we know from psychology, and points research in policy-relevant directions which we don't pretend to have answers to but which are undeniably important.

To stick close to the challenge of stated themes, the paper is organized around questions posed by the conference organizers. I also add a few more questions which are those that behavioral economists are currently asking.

I. How does research by psychologists and other behavioral scientists challenge the economist's model of individual decision making?

Table 1 lists some of the central building blocks in economic analysis. These principles are taught (often implicitly) in every required graduate course. Some combination of these assumptions, along with assumptions about institutional rules for determining prices and allocations, underlie most of what economists do.

Table 1: Building blocks of economic theory and behavioral alternatives

Rational-choice assumption	**Behavioral alternative**
Complete, transitive, smooth preferences	constructed preferences
over risks (EU, SEU)	prospect theory
	uncertainty-aversion
	case-based
over time	-
	multiple systems (e.g. hot-cold)
	internalities (habits, preference for increases)
Dynamic programming	heuristics (a la chess)
Preferences are "asocial"	herding, cascades
Bayesian updating	quasi-Bayesian
separation of belief, value	wishful thinking, self-serving bias
separation of prior, likelihood	encoding bias
Self-interest	social utilities
Profit-maximization	trial-and-error adjustment
Market-clearing	nonprice rationing (e.g., queues, nepotism) wage-price stickiness
Game-theoretic equilibrium	cognitive hierarchy, learning
Submerged axioms	
Description-invariance	framing
Procedure-invariance	compatibility effects
Context-invariance	comparative models
Portfolio view	decision isolation
Fungibility	mental accounting

Basic building blocks of economic theory

Virtually all economic analyses rely on combinations of these assumptions.[ii] Each assumption is a simplification and hence permits counterexamples (or "anomalies"). The

point of behavioral economics is not to produce anomalies, but to create new theory inspired by those anomalies.

At this point, it is useful to reminder the reader about the basic steps in the historical point-counterpoint debate between behavioral and rational-choice economics about how empirical weakness in rationality assumptions should be evaluated. Some of the most frequently-asked questions which have been largely asked and answered over the years are collected in an FAQ Appendix at the end of the paper.

The first line of defense: "as if"

The standard defense of rationality assumptions is that people may not explicitly calculate as the theory implicitly assumes, but they act "as if" they do. The "as if" mantra has been repeated so often in defense of rational modelling that it is easy to loose track of what it really means: Namely, that a theory which refuses to specify a detailed cognitive mechanism by which rational choice is achieved is inherently incomplete. Possible mechanisms which could lead to rational choice include natural or cultural selection against limited-rationality, less-rational agents buying advice (e.g., financial planning) which leads to more rational choice, and learning from experience or imitation of more-successful (and presumably more-rational) agents.

Rational-choice theory could be completed, rather than relying on the incomplete "as if" defense, by investigating the conditions under which these mechanisms lead to rational behavior or don't. Behavioral economics is about completing the theory by specifying more detailed mechanisms grounded in psychological regularity. Most analyses of that sort show that rational outcomes are not guaranteed by plausible details of these theoretically-unspecified mechanisms (e.g., Russell and Thaler, 1985).

The subconscious line of defense: Normative and descriptive

There is a subconcious line of defense for rational modelling which is mentioned surprisingly rarely. Most rational models are normative by definition—weighting future rewards other than exponentially is misforecasting one's own future behavior[iii]; updating probabilities according to Bayes' rule is sensible; and smoothing consumption over one's life cycle is the only way to satisfy certain appealing axioms. But why should normative

models be privileged as the most likely descriptions of behavior? One explanation is that academic economists gravitate toward these assumptions because they seem descriptive of their own behavior—either that's why they chose to study economics rather than physics or sociology in the first place or, once drawn into the discipline, the assumptions begin to sound plausible because of acculturation. In any case, academic economists who succeed and shape prevailing thought are likely to be or think more foresightfully (buying costly education to gain future wages) and analytically than average folks, and to be more persuaded by t-statistics than anecdote, rhetoric, and popular opinion. These academics projecting their own tastes and skills on to the general population is a kind of "projection bias" or "curse of knowledge" (cites).

Aren't approximations useful?

The precision of rational modelling inherently implies that such models are simplified approximations. There is no doubt among behavioral economists that rational modelling is often a useful approximation (an issue cogently raised by Roth, 19xx, but anticipated from the start by behavioral economists). The usefulness of rationality as an approximation neve*r was*, is *not now*, and never *will be* a matter of debate; so it is a waste of time to frame the debate in those terms. The critiques of rationality and the interest in anomalies are not meant to heckle the rational model by showing its imperfections—since we know well that all models have are simplifications and hence permit counterexamples (including behavioral anomalies). Instead, the critiques are designed to point toward new alternative theories and constructs.

So the only question is whether behavioral economics can complicate the rational model in a way that is parsimonious, useful and fruitful. As Mullainathan and Thaler, 200?) point out, relaxing perfect rationality assumptions is a natural step in progress in economic modelling after relaxing perfect competition and perfect information. Both of those two "perfect"'s are useful, and are therefore still used in introductory teaching to illustrate basic principles (because models based on them are easy to grasp, and informative). But they are also looked back on by modern researchers as quaint, like a Model T car or early fax machines[iv], compared to the souped-up models of game-

theoretic competition and signaling that are routinely taught and used today. Hopefully we will someday—soon-- look back on perfect rationality in the same way.

There is also no doubt that having models which operate at fundamentally different levels and are only weakly linked can be useful. For example, microfoundations of macroeconomics (e.g., growth theory). ; it is hard to build up aggregate demand curves from individual ones (Sonnenschein) or social utility functions from individual ones (Condorcet, Arrow). But the advantage of weak linkage should not be an excuse for ignoring any possible inspiration and empirical discipline from a "lower level" model. If it makes sense to complicate a higher-level model, like models of life-cycle savings or asset price changes, by considering alternative assumptions about lower-level assumptions about individual behavior, then the first place to look should be evidence from fields that study these phenomena carefully at the lower level (e.g., patterns of discounting of cash flows by humans and other animals, or evidence of mistakes in forecasting a time series of earnings).

The promise of the "as if" approach is that theories based on approximate mechanisms—i.e., assumptions that are technically false-- might still make good predictions. But theories based on strong rationality assumptions have *not* always made especially good predictions. Much of this is the familiar half-full/half-empty debate: Rational theories are not unreasonable, but in every major field of economics there are large puzzles which have not been resolved by decades of debate. Behavioral alternatives have the advantage of being better-grounded in regularity and being able to potentially explain these puzzles. Good science is progressive, which is precisely the idea that expanding theories to explain regularities that puzzling from the point of view of previous theorizing is progress.

Here are some examples: The assumption that wealth categories are fungible (or compiled into a net worth figure) is violated by clear evidence that marginal propensities to consume vary with wealth categories (Shefrin-Thaler). Market-clearing theories of labor supply cannot explain unemployment. In macroeconomics, there is no consensus about why there are business cycles and why prices and wages are so sticky (especially as the technology for rapidly changing prices and wages has advanced rapidly). In finance, the debate about whether markets are informationally efficient is more unsettled than ever

(even though there are literally about 100 million observations of daily stock prices widely available on the CRSP tapes); and the chronic poor performance and proliferation of mutual funds has not been clearly explained. In corporate finance and business strategy, it is well-established that firms often destroy stockholder value with optimistic mergers, especially when poorly governed; but why is governance so poor? In studies of life-cycle savings, it is not known why people in modern America save so little, and why consumption is so smooth and drops sharply upon retirement (which it should not if people plan for retirement).

My point is not that behavioral economics can clearly resolve all these puzzles rapidly— we can't, yet. The point is just that defending "as if" rationality assumptions, on the basis that their predictions can be surprisingly accurate, is severely undermined by the plain fact that many of the predictions are clearly wrong.

II. How should we amend our views of individual consumption and investment decisions?

The right side of Table 1 lists alternatives to the conventional rational choice principles on the left side of the table. The profession has not achieved a consensus on which of these principles are most realistic or analytically fruitful. The point of the table is that a lot of progress has been made rapidly in suggesting some formal alternatives which extend rationality principles to explain anomalies.

Most of the alternatives are extensions of conventional models which add a parameter or two or suggest new types of functions. These extensional models are useful because they include rationality as a boundary or limiting case; and they lead naturally to econometric tests in which the "amount" of limited rationality can be precisely estimated in the form of parameter values. For examples, models of hyperbolic discounting extend exponential (i.e., dynamically consistent) discounting by adding a preference for immediacy, codified by a parameter . When =1 the theory reduces to exponential discounting. Empirical estimates of then give an idea of "how irrational" (or present-biased) people are.

Another example is social utility functions which assume that people care about their own payoffs, plus weights on "envy" (earning less than others) and on "guilt"

(earning more than others), which have proved useful in explaining regularities in bargaining and public goods contribution (Fehr and Schmidt, 1999; Fehr and Gächter, 2000). When = =0 players are self-interested. Measuring values of and therefore provides a sharp metric for how much of a simplification self-interest is and, consequently, how much is gained by extending the self-interest assumption.

Unlike hyperbolic discounting and envy-guilt theories, many of the alternative theories listed on the right side of Table 1 have not yet been sanded and polished into clearly specified theories that can be taken to data. But there is every reason to think that sort of formalization will happen rapidly and prove useful.

III. What more do we need to know?

One way to think about policy is to ask what we need to know before we can make prescriptions. There are at least three hotly-debated topics which lie at the core of policy debates and are now active areas of research in behavioral economics (or should be): Field tests, self-awareness, endogeneous institutions, missing psychology, and neuroscience.

Field tests: Early evidence in behavioral economics came largely from laboratory experiments. Experiments are particularly useful for ruling out alternative explanations. For example, if we observed people in the field keeping the cars they bought or spouses they married, even if they grew dissatisfied, we can't be sure whether they exhibit an endowment effect or the transactions costs of switching are large. So Knetsch (cite) ran experiments in which subjects are given a good randomly (e.g., a chocolate bar or pen) and allowed to switch. The transactions costs were made as low as possible—an experimenter would pick the chocolate bar off their desk, if the subject wanted, and plunk down a pen, or vice versa. Most subjects still stick with what they were endowed with, which rules out the economic transaction cost explanation. A true believer would stretch the definition of transaction costs to include a psychic transaction cost-- which is right where we want them, on behavioral territory, since this relabelling begs the question of which giving something up is psychically costly.

Now that so many persistent anomalies have been demonstrated experimentally, and shown to be robust to incentives, repetition, educational background, and so forth, it

is useful to switch to looking for these effects in field. In fact, field evidence has been slowly accumulating for some time (particularly since the late 1980s). Evidence of asymmetric consumer responses to price changes, reluctance to sell large losing assets (stocks and houses), sensitivity of New York cab drivers to daily income targets, strong preferences for low-probability lotteries, and overreaction of markets to stocks with bad returns (among others; see Camerer, 2000, for a summary) are all consistent with phenomena established earlier in the lab-- reference-dependence, decision isolation, overweighting of low probabilities, and failure to expect regression toward the mean.

The latest studies are even better because they often compare rational-choice predictions with behavioral alternatives more sharply. They have shown the power of defaults in determining pension allocations (Madrian et al), - discounting in savings (Laibson et al), and mistakes in forecasting future health-club use (della Vigna and Malmendier). In each case, it is possible to tell a complicated (perhaps cockamamie) rational-choice "story" of why the established regularity exists. But doing so in an ad hoc way for each regularity, undermines the claim for parsimony that rational choice theorists are so rightly proud of. In any case, more studies of this sort are needed and many are on the way. Some will disprove earlier behavioral explanations or establish limits that the earlier experimental data would not have anticipated. That's fine with us because we care about finding out how the economy works rather than foolishly defending a faith.

Field studies are particularly useful for establishing the boundaries of phenomena and their practical importance. For example, if people have an exaggerated distaste for giving up goods they own or consumption patterns they are used to—the "endowment effect" shown by Knetsch's experiments (Thaler, 1980; Kahneman, Knetsch and Thaler, 1990; cf. Plott and Zeiler, 2003)--, the question arises of whether the experience of buying and selling over a lifetime will diminish any such endowment effect. For example, if the endowment effect is due to misforecasting how badly one will feel after losing a good, as I suspect it does, people may learn that life goes on and quit misforecasting. Indeed, studies comparing amateur and experienced (active) traders of sports paraphernalia showed that more experienced traders show no endowment effect for the goods they regularly trade (List QJE in press). Furthermore, experienced traders of paraphernalia show smaller endowment effects for everyday goods used in lab studies

than novice traders, as if the experience of buying and selling spills over to minimize aversion to losses of everyday goods (List, 2003). Our study of cab drivers showed that more experienced drivers did not show the distaste for falling short of a target that inexperienced drivers did (Camerer et al, 1997). These facts are raw material for refinements: An appropriately rich theory should explain when phenomena like endowment effects exist, and when they don't. Learning over the life cycle that losses aren't so bad is one candidate for how effects shrink over time, but there is plenty of other room for other variables like advertising, education, and so forth.

Finally, field studies are useful because they permit sharp empirical showdowns. They force behavioral economists to be very precise about what their models predict. And they *also* force rational-choice theorists to be precise about what sorts of evidence would unequivocally refute predictions based on rationality.[v]

Self-awareness: Economic theorists who are (over?)trained to think about rational choice instinctively think that even if people make mistakes, they must be aware of those mistakes. This presumption leads immediately to meta-rational model in which people are self-aware about their limits and strive to overcome them (sometimes called "sophistication" in the intertemporal-choice research).

It is certainly useful to explore the implications of models of this sort based on self-awareness.[vi] But it is not the only path.

To a psychologist, the idea of human self-awareness is closely related to the "homunculus fallacy". The homunculus is an idealized "creature" in the brain which behaves optimally—a brain-endowed creature within the brain which has "executive control". In cognitive science, it is considered an embarassing failure to include a homunculus as a cheap short-cut to explain behavior because positing a homunculus creates a recursive problem: Does the homunculus have a brain? And if so, does *that* brain have an even tinier homunculus? And what about *that* homunculus's brain? Ad infinitum. Assuming self-awareness creates a homunculus problem and as a result, cognitive scientists are much less inclined to quickly work toward presume self-awareness than some modern economic theorists are.

Furthermore, many examples of behavior suggest limited self-awareness. (An economist might model this as the brain having separate functions which interact; but no

single homunculistic function understands all those interactions.) A famous example of limited self-awareness is the behavior of split-brain patients—patients whose "corpus callosum" (the thin band of tissue connecting the left and right hemispheres) have been severed, usually to prevent epileptic fits from spreading from one hemisphere to another. It is well-known that language comprehension and recognition are normally located in the left hemisphere (Broca's and Wernicke's areas). It is also well-known that if you see an image on the left part of a screen, that image is seen by the left eye but processed in the right side of the brain. Suppose you show a split-brain patient a screen which says "wave" on the left side, and ask the patient to do what the screen says to do. Patients will wave their hand. The right side of the brain sees the instruction and instructs the right hand to wave.

Now you ask the patient *why* she waved. The left hemisphere (the language area) has to think up a verbal answer and say it. Often subjects say "Well, I saw somebody I know and waved to them". We know this explanation is wrong but the brain is so overtrained to make sense of behavior that the language area picks words which make sense to "it" (the left hemisphere). Other patients say "I don't know"—what they mean is "I, the left side, don't know". The left side literally doesn't know what the right side was doing.

This example shows that we might *act* as if we are self-aware, but "we" (the patients' left sides) are just making it up—rationalizing. The human brain is like a monkey brain with a press secretary. The fact that an apparently-self-aware explanation is so typically given by the press secretary suggests that much of our faith in self-awareness is illusory and should be taken skeptically. As they say in neuroscience, "Don't ask the person; ask the brain".

These split-brain patients are exceptional but there are many other examples of this sort of failure of self-awareness. For example, show normal subjects a smiling or frowning face very rapidly (30 milliseconds) followed by a neutral-face "mask". Most subjects say they have no idea whether saw a smiling or frowning face initially. They aren't lying or holding back—"they" (their cortex which grasps for an answer) truly doesn't know. But if you measure movements of facial muscles (cheekbones for smiling and between the eyebrows for frowning) the muscle movements are correlated with the

type of face seen in the initial 30 msec exposure; the person who saw a happy face smiled, and the person who saw a sad face frowned, but this rapid emotional processing does not reach the press secretary in the cortex. So the face knows but the "person"-- i.e, the cortex which tries to interpret signals from elsewhere in the brain-- doesn't.

One might argue that the problem here is simply rapidity of exposure. It's true that in most economic decisions, people have longer to deliberate and process information than 30 milliseconds. But they are also dealing with much more complex stimuli. It is not clear that having longer to think about how much to save, whether to quit a job, and so forth will necessarily produce self-awareness about all the brain activity which determined the decision. It might be, as in the split-brain patients, that some rapid motor or emotional activity quickly "decides" and the cortex spends some time rationalizing or trying to figure why the emotions decided what they did. For example, it is well-established that political affiliations are highly-correlated across generations (cite). But if you ask a young Republican, "Why do you vote Republican?" she typically won't say, "Because my Dad did". More likely he will (a) create a plausible tale about ideology, liking or disliking particular candidates, fiscal irresponsibility of government, a fetus's right to life, and so on. It may even be that the process of talking with his Dad about these issues shaped his viewpoint; but even then, a truly self-aware young Republican should say, (b) "I spent a lot of time talking with my Dad about these issues and he convinced me". Explanation (b) permits an observer to infer that if the young Republican's Dad had not talked politics she might be a Democrat, an inference which explanation (a) doesn't invite.

The point is not that people are stupid, lie to survey questioners, or are always and entirely clueless about what really caused their behavior. The point is just that the potential exists for poor self-awareness because the self is not a self— it is (they are?) an interaction of complicated mechanisms: The cortex is eager to make up a rationalization; and the rationalization sounds reasonable even when outside observers know it's wrong (e.g., the experimenters know why the split-brain patients really waved).

One way to think of the opposite of sophistication—naivete—is that it requires persistent misforecasting of future behavior. Economic theorists seems to instinctively regard this as impossible. But social psychologists have found precisely such persistent misforecasting about future emotions and understand a lot about it (Dan Gilbert, et al) . Another common claim is that when people misforecast repeatedly, they will learn that

they do so, and wise up. Whether they do is an empirical question that shouldn't be prejudged by pure faith in learning from experience. Learning is guided by hippocampal processing of information into long-term memories; when this process works poorly people may fail to learn. More importantly, "top-down" encoding of information can create biases in what is perceived which lead people to ignore unexpected surprises, which means the feedback about mistakes necessary for learning is filtered out.

Furthermore, many economic decisions are made once or are largely irreversible, and have long-term effects (fertility choices, education choices, early savings choice). So even if forecasting mistakes are learned over a person's life, it may be too late for older folks to reverse the long-run effects of their youthful mistakes (cf. "youth is wasted on the young").

Endogeneous institutions: A theme which may emerge rapidly from thinking about cognitive limits is the idea that collective actions or institutions emerge to both limit and exploit these mistakes. An example is advertising. The standard theories of advertising are embarassingly simple and make advertising executives burst out laughing. One theory is that ads convey information about product price and quality. The most expensive ads—on TV— often do not.[vii] Another theory is that advertising "burns money" and signals to consumers that the advertising firm's products are so good that future sales will pay for the ad costs many times over. But advertisements studiously avoid mentioning how much the ad cost, which is the one thing the signaling theory predicts the advertisement should mention, and the one thing it does *not.* A useful way to think about advertising is that it reveals an implicitly behavioral theory advertisers have about consumer decision making. The nature of the advertising is a clue about how rational advertisers think people behave. Rationalizing institutional behavior of this sort will be useful.

Another example illustrates how clever institutions understand and "debias" the "disposition effect". The disposition effect is a name for the fact that house owners and people who bought stock shares tend to be reluctant to sell at an accounting loss (i.e., at a selling price less than their purchase price; Shefrin & Statman, Odean, Weber and Camerer,). Except for small tax effects (which usually encourage selling losing stocks, not keeping them), rational house and stock owners should ignore their purchase price

and concentrate on their own future costs and benefits relative to a prevailing market price.

A large trading firm understands the powerful tug loss-aversion has on inhibiting its traders from closing out trading positions which have lost money, clouding their judgment.[viii] So the firm takes two traders and swaps their positions: Sharon's portfolio at day-end on Tuesday is given to Colin to manage in the morning on Wednesday, and vice versa. The new traders feel less emotional attachment to the new position they inherited from the other trader (and ideally, don't even know the prices at which the position was initiated) and exhibit less or no disposition effect. The beauty of this scheme is that, from the firm's viewpoint, the swap is completely neutral—it does not change their aggregate portfolio at all. But by swapping the positions the disposition effect is "debiased". The fact that the firm does this implies that they believe there is a disposition effect among individual traders and create an institutional rule change to limit the firm's harm from the effect. Furthermore, the firm's rule creats a novel type of scale economy: A firm with more traders can do more firm-neutral swaps and, if the individual disposition effect leads to mistakes, larger firms will earn more than smaller ones (ceteris paribus).

A third and final example comes from Microsoft. Software designers and other experts often suffer from a "curse of knowledge" (Camerer et al): They find it hard to imagine how little others know about their area of expertise. So Microsoft found that its software experts were surprised, and skeptical, about reports from help-line employees about how buggy callers found Microsoft software to be. They installed a room with a one-way mirror in which designers could watch typical users struggle with their software. Microsoft cleverly pitted one bias against another: The designers' curse of knowledge was overcome by the power of vivid, salient information as they watched normal-looking users struggle in front of their very eyes. The result was that designers realized how difficult their software was for average people to comprehend (Heath, Larrick and Klayman, cite).

These examples show how firms' investments in advertising, position-swapping, and making consumer complaints evident, led to institutional changes which reveal an implicit behavioral economics theory that explains the changes. So behavioral economics

guides a positive analysis of what these firms are doing which might otherwise be difficult to comprehend.

Missing psychology: An important area of active research is drawing new ideas from psychology into behavioral economics. The research from 1980-2000 or so was heavily influenced by work in psychology which was, in turn, driven by using anomalies relative to rational choice models of judgment and preference. As a result, central ideas in psychology which are not gracefully cast in terms of rational choice principles have been neglected.

One example is categorization. Knowledge is clearly organized in categories (cite). Furthermore, modern psychologists think categories are often organized around exemplar or prototypical examples rather than defined by a list of features. (For example, a perceived "scientist" is a person who shares traits with Albert Einstein, not a list of traits like a job description.) Concepts of categorization are useful for thinking about labor markets (Fryer and Jackson, 2003?) and creating alternatives to Bayesian updating (Mullainathan, 2003?). For example, Fryer and Jackson develop a theory of optimal categorization under constraint which implies that minorities will be lumped in coarser categories than majority members. Their theory is consistent with evidence of racial discrimination. Mullainathan's theory implies that new information can create changes in beliefs that are much sharper than they would be in smoother Bayesian updating; this can account for sharp "regime shifts" in perceptions of business cycles, financial market volatility and so forth.

Another interesting psychological concept, neglected in behavioral economics until recently, is limited attention. Attention is the ultimate scarce cognitive resource. It is possible to learn to attend to many stimuli at the same time—as busy financial-market floor traders and cell-phone-using drivers do—but long-term memory suffers. Scarce attention might be useful for explaining economic phenomena like organizational structure (division of labor expands organizational attention but is constrained by the need to coordinate) and advertising (which "grabs" attention).

Neuroscience: "I are we": A small group of behavioral economists are increasingly convinced that economics could be usefully informed by neuroscience, a field which is

exploding with insight because of advances in imaging brain activity and genetics (see Camerer, Loewenstein, Prelec, 2003).

Two simple insights are key: First, human are just primates with a thin layer of extra cortex (used for long-term planning, understanding complex social structures and symbolic systems like language and mathematics, and creating refined social emotions like guilt and embarassment). This means we can learn a lot from similarities between humans and other animals. For example, rats become addicted to every drug of abuse humans become addicted to (cite). But human cortex means we can inhibit drug use even when rats, for example, do not. (Rats voluntarily choose cocaine doses over food until they die of starvation; most humans do not.) The human-animal similarity also means the explanation for distinctly human behavior must lie in the cortex; so we know where to look in imaging studies.

Second, neuroscience reminds us that the brain is complex. Look under the hood of a car. There is a battery, carburetor, fan belt, oil pan, and so on. The car produces a certain amount of basic behavior—e.g., moving forward rapidly—which does not require intimate knowledge of what's under the hood. But if you want to know why a car broke down, or how to build a better one, you have to look under the hood.

The brain is the same way— it is an evolved (and developed, and socialized) collection of modular subsystems, which interact to produce behavior. As a result, it is unlikely that this brain would maximize any single function, like a utility function over health-work-leisure-money tradeoffs. "I" is not an "I"— I are we.

The breakthroughs in using neuroscience to do economics will come from simple approaches which model behavior as the interaction of a small number of systems, perhaps only two (which is, to a neuroscientist, grossly oversimplifed, but which nonetheless might be the right level of abstraction for economics).

The challenge is to pick two systems (a neurally plausible dichotomy), or more, and generate precise predictions about the behavior which would result. One such model could focus on the difference between "wanting" (preference revealed by choices) and "liking" (how the brain response to actual consumption). I'll say more about this below in discussing welfare.

Another model focusses on "hot" and "cold" systems—loosely speaking, emotions and cognition (Zajonc, etc). Loewenstein has argued that in "hot" states—anger, hunger, fatigue, PMS in women, drug craving, depression— people become more self- and immediately-focussed, and their "heat" inhibits thoughtful reasoning. (And as noted above, in hot states the brain is probably not fully self-aware that the hot state is altering behavior.)

Bernheim and Rangel (cites) use this distinction to create neurally-grounded models of addiction and savings. They take the view that a person's choices in the cold state are the correct measure to be used for evaluating welfare. With this working assumption, they are able to make precise welfare statements about policies (e.g., comparing taxation, regulated dispensation, and bans of drugs, and laissez-faire). Laibson's (2001) model of how environmental "cues" trigger demand for habit-forming consumption is another elegant example showing that modelling of this sort *can* be done and is insightful.

Another useful two-system distinction is automatic ("habit") vs. controlled systems (see Benhabib and Bisin, 200?). Much of our behavior is clearly automatic (involuntary reflexes like knee-jerks, breathing, and startle responses to fear), or becomes automatized with practice (driving a car). Controlled systems require cortical deliberation to override the automatic responses.[ix] Automaticity is probably important in our reactions to advertising, rapid ("quick and dirty") reactions to racial and gender cues, and other examples of economic interest. For example, Duflo and Saenz report that corporate employees who accompany their friends to a forum to make health-program decisions are more likely to change their own plans at the same time (cite). They don't push this interpretation (and their data are insufficient to validate my explanation) but is plausible that going somewhere with a friend is an automatic response ("Hey, wanna come to lunch? I have to stop at this health care thing and fill out a form real quick, it will only take a few minutes"). Once faced with a form, some controlled deliberation results in a choice. These employees made changes to their health care programs they *could* have made in a controlled way ("I should really look at these new programs") but which they only made because they automatically went with a friend to the health-program signup desk.

There is no reason to be anything but optimistic that *something* interesting will come from neuroeconomic models of the interaction of hot-cold and automatic-controlled systems. Why? First, because such models are scientifically deep. Technology changes will generate an enormous amount of new data about the economizing brain and, after perhaps after years of confusion and spurious results, some eventual insight. Second, economic theorists love a technical challenge and rise to the occasion. The brain facts provide a challenge because there a lot of modelling choices and will require some ingenuity to choose models that are neurally reasonable, but solvable, and which lead to new implications and eventually to policy. Once the right way of posing questions if found, models of this sort will surely prove no more difficult than the sort of deep general equilibrium and delicate game-theoretic questions that are routinely answered in modern economic theory. (Keep in mind, too, that the pace of research has already been dramatic— problems in economic theory which were unsolved 20 years ago, and not even imagined 40 years ago, are now assigned in problem sets to first-year graduate students.)

IV. What about welfare?

A central issue for policy is what behavioral economics can say about human welfare, and which policies are welfare-enhancing. I'll offer some precise ideas about welfare, but start with two caveats.

The first caveat is that virtually everyone doing behavioral economics agrees we should go slowly in advocating policy change unless forced to make quick recommendations. While behavioral economics in its modern form is about 20 years old, it is a small field and much of our energy, until recently, has been consumed fending off rational explanations of behavioral "anomalies". Our thinking has also not designed to precisely answer questions about welfare and policy but this is a good time in the intellectual history of the field to say <u>something</u>.

The second caveat is that rational theory is also an incomplete guide to policy. Many policy prescriptions boil down to the law of demand— changing prices affects behavior—which is agreed upon by neoclassical and behavioral economists because even random behavior in the face of budget constraint will generate downward-sloping

demand (Becker, 196*). The problem is not that the theory is wrong, but that it is incomplete. It doesn't tell us about how rapidly responses occur, which is important when political pressures make it hard to impose short-run losses for long-run gains (think of privitization efforts in former Communist countries, the appeal of Latin American populists trying to push for painful reforms, etc.) The undeniable response of behavior to prices also doesn't tell us whether behavior is changed by policy variables other than prices (e.g., anti-smoking campaigns, shifts in social norms, advertising, the influence of catastrophes like 9/11 on risk perceptions, etc.)

There are at least three ways to think about welfare from a behavioral economics viewpoint.

First, we could draw some inspiration from other fields which comfortably accept the idea that what people choose for themselves is not always in their best interest (whether the latter is prescribed socially or personally). For example, psychiatrists have concepts of mental health which they use to classify disorders which are treated with therapies or drugs. They are implicitly using standards of ideal consumer welfare and are trying to push behavior in that direction, often backed by the force of the law (which should be understood in this context as a welfare standard endorsed by society). An example in law is the "prudent man" rule in judging whether fudiciaries have acted responsibly. This concept implicitly advocates a socially-agreed-upon ideal and punishes deviations.

Another example is self-improvement. Suppose an overweight person loves In-and-Out Burgers and pizza delivered in large, fatty portions. A revealed preference view would insist that since the burger and pizza aficionado chose them, he must have higher utility for those foods than for salad. But suppose our hero presents at a clinic and says he wants to cut back on burgers and pizza. Showing up reveals a meta-preference for changing his revealed preferences. We should use his revealed meta-preference to override day-to-day inferences from what he eats about what he really wants.

Second, a useful language for complicating discussions about welfare was revived by neo-Benthamite types of utility (Kahneman, ?). A neuroscientist would distinguish at at least four separate kinds of utility numbers that could, in principle, be measured in the brain with techniques that are here or on the way: Remembered utility, anticipated utility,

choice utility, and experienced utility. Remembered utility is what people recall liking; anticipated utility is what they expect to like; choice utility is what they reveal by choosing (classical revealed preference); and experienced utility is what they actually like when they consume. Using revealed preferences to judge welfare is the same as assuming that choice utility and experienced utility match up.

In general, it might be exceptional that all four types of utilities match up. At Caltech there are only four nearby places to eat lunch. Over a ten-year period my remembered utilities, anticipated utilities, choice utilities, and experienced utilities for different lunch meals have become rather well-matched (subject to a taste for variety-seeking) If I think about getting a bowl of Vietnamese pho at the Avery Center, I remember accurately how it tasted the last time and, using that memory to imagine how it will taste today, correctly anticipate what it will taste like today. So I choose it. The taste experience corresponds to how I remembered it did and anticipated it would.

The four types of utility will not match up when decisions are rare or the world changes. In fact, neural evidence shows that "wanting" (anticipated utility) and "liking" (experienced utility) are located in separate areas of the brain. This dissociation means that in the short run, people might want what they won't actually like (or won't want what they would actually like). Examples are easy to find. Some infants reveal a preference for eating dirt: do they rationally anticipate liking it, or are they just exploring a possible preference which wiser parents know their children don't have? Addicts often report drug craving (wanting) which lead to consumption (choosing) they do not find pleasurable (don't like). Compulsive shoppers buy goods (choose them) which they never use (don't get experienced utility from).

Are average consumers more like me, rationally buying pho for lunch when they are in the mood, or like dirt-eating infants, addicts and shopaholics? Presumably most people in many choices are somewhere in between. When the decisions are momentous and rare, like getting pregnant, deciding whether to go to college, signing up for pension contributions, or buying a house, there is no reason to think the four types of utility will necessarily match up.

Furthermore, there is a clear welfare criterion among the four kinds of utility: What people *should* want is to maximize a combination of experienced utility and remembered utility (because "consumption of experiences" matters too).

Distinguishing types of utility also suggests a way to explain why society permits institutions or a person's relatives to overrule consumer sovereignty—denying that choices match actual welfare—in interesting exceptions, like children and the mentally ill. The idea that minors cannot make choices for themselves is rooted in an implicit developmental psychology. The implicit idea is that brains which are not 18 years old yet (for example) have not learned what is best for the owners of those young brains. Psychiatric diagnosis of incompetence appeals to an idea that mental health experts can overrule consumer sovereignty. The 18-year old and "psychiatrists said so" rules are obviously coarse bright lines. A more refined analysis would presumably depend on an expanded model in which people figure out what is good for them at some point, and we (policymakers) can figure out in a reasonably unbiased way whether people have reached that point of consumer sovereignty.

For example, one could imagine a diagnostic tool which presented people with images of products and asked them to choose which they like most (measuring "wanting"). Then we separately measure how much they like actually consuming those goods ("liking"). If wanting and liking brain activation are different, a consumer could be judged as incompetent (or not-yet-competent). This would probably work well for children, and possibly for addicts and other classes of patients society currently denies free choice to. These test cases aren't proof that such a method would work more widely, but they are proof that a restricted version of the method could scientifically support what society does currently and might lead to reasonable changes in the current bright-line procedures we apply.

A third way to think about welfare is provoked by the neural models described above which recognize that the human brain is an interaction of multiples systems. In these models welfare is usually defined as the choices made by one of the many systems. For example, Bernheim and Rangel (cite) assume that the "cold" (anticipatory) system's choices are those which reflect a person's true welfare. In systems with controlled and automatic components, the controlled or deliberative choices reflect true welfare. It is

undoubtedly hard for an economist steeped in the tradition of revealed preference to adjudicate between which system's choices should be used to judge welfare. But this is only because the classical model gives you the luxury of avoiding such a choice, by denying the multiple-system distinction.

V. Conclusions

This essay lays out the case for shifting from conventional assumptions of rationality to behavioral alternatives in doing economic analysis. Many of these alternatives already extend the rationality model precisely by adding one or two parameters so the instinctive fear that behavioral models are unfalsifiable or inevitably unparsimonious is immediately proved wrong. A more difficult challenge provoked by this conference is what the behavioral models might say about welfare and policy.

Here is a summary of the key points:

- Many rationality principles—the nature of preferences (including over gambles and intertemporal streams), Bayesian probability, strategic thinking—can be replaced by viable behavioral alternatives. These alternatives have the properties economists like—formal structure, parsimony, and the ability to make predictions about field phenomena and shape policy.
- The "as if" defense of rationality principles is inherently incomplete, because it does not prescribe precise mechanisms of learning, selection, advice-taking, imitation, etc., which lead the principles to be true of human behavior. So it is impossible to "disprove" the as-if view (since it is not a complete theory); the goal, instead, is to flesh out conditions under which the as-if view holds and, when it does not, what else will happen.
- The idea that incorrect assumptions can lead to accurate predictions does not vindicate "as if" rationality models because many predictions based on those assumptions are clearly wrong (or at the very least, widely debated). Every important area of economics has large anomalies that invite behavioral explanations.
- The next wave of debates will include: Predictions from behavioral alternatives about field data; the degree of self-awareness in human choice; the value of

behavioral economics for explaining institutional choices (like advertising and organizational design); incorporating ideas in psychology missing from the first wave of behavioral economics (like categorization and limited attention); and neuroscientific "multiple-system" approaches.

- The behavioral view inescapably complicates welfare analysis, but only because complete, transitive preferences inherently assume away the possibility that what people choose and what is truly best for them may differ.
- In the behavioral view, welfare might be judged much as psychiatrists do (or by using the meta-preferences people reveal by seeking treatment), by distinguishing types of utility and taking experienced and remembered utilities as true welfare, or by assuming that one of multiple brain systems—e.g., the "cold" or "controlled (deliberative)" system—represents true welfare.

References (To Be Added)

Appendix: FAQ's

Aren't mistakes transitory?

Not necessarily. Important mistakes can occur in rare decisions which have irreversible lifetime consequences—choosing to skip college, have a child at a young age, study hard as a child (in a highly path-dependent system like Japan's), career choices, early savings/consumption (bankruptcy early in life), crime (joining a gang), early childhood eating habits, etc.

Aren't models of rationality inherently less precise than behavioral ones?

No. There is a curious myth that weakening optimization means anything can happen. This is simply wrong because precision can come from other kinds of constraints. An obvious constraint is empirical: In choosing among alternative assumptions about how rationality is limited, look to neighboring disciplines which have explored alternatives carefully for decades. (To do otherwise is to succumb to the mathematical conceit that brilliant people can deduce principles of actual behavior from pure logic and deduction without actually watching what people do. Pure deduction has never worked in any science—chemistry, astronomy, physics—and it is especially unlikely to work in a social science, where the mathematicians doing the thinking are likely to dismiss constrained reasoning of average folks who aren't as smart as they are.) A good constructive example is Gabaix and Laibson's (2002?) model of limited rationality in problems of searching for optimal paths through decision trees. Their model is well-rooted in intuitions about procedurally rational search but is also precise.

More interesting examples come from domains in which multiplicity of equilibria arise from intertemporal considerations or strategic ones. E.g., Lucas (1986) notes that there are often many rational expectations paths in money-price models; and in game theory many interesting games have multiple equilibria (e.g., when players can coordinate on a risky "good" equilibrium or a less risky "bad" one). Models with adaptive expectations and limited iterated reasoning can actually be *more* precise in these cases (e.g. Lucas, 1986; Camerer, Ho and Chong, 2002).

Aren't behavioral addenda just epicycles?

The "epicycle" critique of behavioral economics is an unfortunate mischaracterization of both astronomical history *and* behavioral economics. Epicycles were created to rescue a theory of planetary motion which failed to account for subtle gravitational forces and hence made wrong predictions. The analogy to ideas in behavioral economics is misguided because the new parameters and ideas in behavioral economics were not concocted on the spot to explain small perturbations—instead, they are deeply rooted in decades of very careful work by psychologist on how people behave, and the departures from rational behavior they aim to explain are large, not small. To call the behavioral economic theories which put the psychologists' carefully-established data to work "epicycles" is ignorantly to dismiss the entire disciplines that provide raw material for behavioral modifications and to mischaracterize how behavioral economists search for independent empirical verification of modifications to theory rather than simply positing curve-fitting epicycles.

Ironically, the besst producers of epicycles are rational-choice theorists who are eager to explain away every lapse in rationality by an imputed "information cost" or an evolutionary adaptation to a hypothetical EEA problem. Because costs and evolution are not directly measurable, imputing missing costs and evolutionary adaptations in this way is inherently tautological. (A good example is Posner and Fremling (Cite). I am a big admirer of Judge Posner's intellectual creativity in general, which is on great display in this paper.)

An alternative scientific model to epicycles is the periodic table of the elements in chemistry. Nobody complains that new elements are epicyclic extensions of old ones because they are clearly distinguished chemically. Psychological regularity is similar—a list of effects are like elements which are believed by psychologists to be "chemically" distinct, while made up of common elements (atoms, electrons, protons, etc.). Behavioral economic theories work in a similar way. They use common "psychological elements" to derive conceptually distinct explanations, and, importantly, to commit to new predictions derived from the idea that chemical elements are distinct but combine in molecularly regular ways.

Aren't behavioral models inherently unparsimonious because they add parameters?

No. The key test is not parsimony but precision and falsifiability. Many behavioral models which weaken rationality do add parameters to measure a particular kind of limit on rationality, but usually in a minimal way (e.g., adding a strength of immediacy preference to discounting models or a number of thinking steps τ in game theory models). The prediction is that data will show some regularity in these new parameters.

At the same time, behavioral economists are searching for *more* parsimony by seeking regularities which provide a unified view of social science and natural science. For example, neuroscience uses evidence from humans with brain damage, PET scanning, fMRI imaging, EEG wave measurement, self-report, and the "animal model" (primates, rats) to triangulate on a single hypothesis that can explain regularities in all these types of data, which is extremely parsimonious.

The rational-choice explanations of different behavioral anomalies, in contrast, often explain different anomalies with contradictory assumptions. For example, utility over money gains is assumed to be convex to explain gambling, and concave to explain the large equity premium in stock returns compared to bonds. Adding specialized ad hoc assumptions to explain anomalies in different domains is precisely the opposite of what scientists generally mean by parsimony.

Won't markets correct individual irrationality?

Not necessarily. The crucial insight is that behavior in markets is a kind of intellectual and financial arms race in which forces punishing mistakes, or offering advice to those who make mistakes, are matched (or overmatched) by forces that exploit mistakes. Education, rehabilitation, advice markets, and so forth probably do improve the ability of people to save and spend wisely, and resist temptation; but the incentive of individuals to create corporations which pool knowledge about mistakes and exploit them may be even greater. Which force wins out depends on the relative profitability of selling

mistake-correction and exploiting mistakes, as well as scale and scope economies in creating organizations that provide corrective and exploitative services.

An example is food consumption. The diet-fad and nutrition-advice industry have grown dramatically in the US in recent years, so people surely know more about nutritional content and what they should eat. At the same time, fast food consumption and measured obesity (and associated illnesses, such as heart disease) have grown too (and will surely become more important as less-developed economies move beyond subsistence diets toward a point where their citizens can afford to get fat). The average 7-13 year old eats hamburgers about 6 times per week (more often than adults); McDonald's buys more pork than any company in America, even though pork is not served on their dinner and lunch menus (surprise: it's breakfast sausage; Schlossberg, 200?). From a welfare point of view, it is convenient to conclude that if people (pre-teens included?) eat so much fast food it must have high utility. But if the crucial goal is lifetime utility or, equivalently, about making the tradeoff between eating a juicy hamburger now and dying sooner several decades later, the conclusion that children eating six hamburgers a week are behaving optimally is wrong.

It is easy to construct examples in which the profits to exploitation of mistakes exceed those from advice about avoiding mistakes. If people are naive about their own future tastes—e.g., they act as if they will switch to a healthier diet in a few years--exploiters will earn more profits. In fact, exploiters of different sorts might get paid twice. Suppose a diet-fad ad promised you could eat absolutely anything you want and still lose weight (as some do). Gullible people would pay for the advice (paying once) and would pay again in future health costs when the diet doesn't work.

Another example is depression and suicide, typically an extreme result of depression. Suppose suicide results from depression which includes a forecasting bias (those who have "suicidal ideation" mistakenly think life will always be terrible, and hence kill themselves in a cost-benefit calculation which is rational given their forecast). Who earns more: Sleeping pill salesmen, or a therapist who counsels against killing oneself? Therapists *should* earn more but won't always. Better products only sell if people are good shoppers. People contemplating suicide probably aren't.

Of course, the issue for policymakers is which regulatory interventions help. Beyond the usual justifications for government— internalizing or correcting externalities, supply of public goods, productive enforcement of property rights—we can add a category of "protecting the weak". Governments *already do this* by protecting children, the mentally ill, the uninformed, and by banning products. Behavioral economists suggest, very tentatively, that normal people are also weak in regular ways. Then the moral basis for protections might be extended further, when research clearly establishes the boundaries of weakness (succumbing to temptation, processing complex information badly) *and* points to simple robust fixups (e.g., Camerer et al, 03; Sunstein and Thaler, 03).

If mistakes are so common, why are developed economies so (relatively) productive, people so long-lived, etc.?

These successes may be partly a result of advantages of scale and diversity, rather than average intellectual capacities of participants; so success of entire economies does not imply much about how smart average participants are. We can also afford mistakes. If people in developed economies eat too much (relative to a lifetime-maximizing diet) they just die a few years earlier than they would have, but live longer than if they lived in poverty. This is not so in say poverty-line agricultural societies—and indeed, rates of business failure, death etc are higher in those societies. This doesn't mean we should not be trying to avoid unfortunate outcomes from rationality limits. Limits on rationality also reallocate wealth & welfare; the consequences of these reallocations may be deemed inherently unfair by societies.

Aren't we evolved to behave rationally?

No. We are evolved to reproduce our genes; rationality principles are based on logic, external arguments (e.g., Dutch book rationales for coherence of probabilities), s. Furthermore most biological evolution ocurred in an environment long ago (often called the "EEA", the environment of evolution of our ancestors). Ideally, evolutionary arguments about how our brains adapted should be an argument in favor of behavioral ideas about rationality limits on rationality, temptation, and self interest, rather than a

defense of hyper-rationality. In the EEA, the human brain didn't need to plan so far ahead (long life is a recent phenomenon) or to deal with monogamy, widespread availability of food, etc. This doesn't mean it is rational, in modern terms, to eat so much fat—it's not in terms of traditional criterion of achieving goals through best possible means, but it is in terms of using mechanisms adapted to an ancestral environment to cope in a modern one ("ecological rationality").

The central issue is whether it is necessary to have an evolutionary foundation for anything we discover in behavioral economics, and whether evolutionary modeling actually produces new insight and falsifiable predictions. There are many reasons to be skeptical. Effective evolutionary modelling requires creating premises (evolutionary mechanisms of transmission and a good description of the EEA) which can never be well grounded empirically because they rely so heavily on archeological speculation. Furthermore, much science progresses without knowing precisely where original mechanisms really came from (e.g., the origin of life, the origin of the universe, how evolution could have created such dramatically "punctuated" changes in species). Scientists in those fields realize how difficult it is to pin down true origins.

Endnotes

[i] I hope my degree of presumption to speak on behalf of a field is just about right (he said presumptively). Behavioral economists do not meet in a cabal to decide on our collective positions and vote. Hardly! It is true that many of us meet in intellectual forums regularly, at first as common-enemy dissenters from a rational-choice view, and later as scientists with eclectic overlapping interests. But we divide along many intellectual and methodological lines, and ebb and flow between them. A practical definition of what makes an economist behavioral is how much attention she or he pays to facts established in neighboring sciences as constraints on modelling. Behavioral economists pay a lot of attention because we believe in a division of labor (which is, ironically, a simple economic concept): We believe that somebody who knows a ton about, say, human memory can save you a lot of trouble in getting to the truth if you talk to them (mostly listening) before you lift pen to paper to write down a model of memory with economic implications. (Since good science is hard, ignoring cheaply-available advice cannot be economical in terms of the intellectual production function which creates good economics.) As a result, in a behavioral economics seminar you get extra points for actually knowing facts and using them to justify your approach; in many other economics seminars you don't get points: Instead, people regard such an empirical justification quizzically, or actually debate facts with you (usually brimming with a surprising combination of confidence and ignorance), or give extra points for a model which is so bizarrely detached from reality that it poses special technical challenges which show prowess, as if a triple flip was the better way to get into a pool than diving straight in. Triple flips are hard but the goal is to get into the pool.

[ii] An interesting exception is macro models which take economy-wide factors like capital and labor as inputs without tying them explicitly to microfoundations. Detached models of this sort are subject to their own criticisms (but the detachment in them is also undoubtedly useful for some purposes).

[iii] Sophisticated vs. naive forecasting notwithstanding.

[iv] In the movie "Almost Famous", set in the 1970s, a touring rock band crowds around a fax machine and gapes as a page prints out. One of them explains that the machine transmits a printed page over a phone line...``and it only takes 3 minutes a page!". The line is funny precisely because we forget how bad old technology is compared to new (a kind of assimilation or "overwriting" bias in memory similar to hindsight bias and curse of knowledge). If the same is true of behavioral and rational modelling, we will soon look back at quaint rational modelling with a mixture of embarassment and nostalgia.

[v] In the movie "Almost Famous", set in the 1970s, a touring rock band crowds around a fax machine and gapes as a page prints out. One of them explains that the machine transmits a printed page over a phone line...``and it only takes 3 minutes a page!". The line is funny precisely because we forget how bad old technology is compared to new (a kind of assimilation or "overwriting" bias in memory similar to hindsight bias and curse of knowledge). If the same is true of behavioral and rational modelling, we will soon look back at quaint rational modelling with a mixture of embarassment and nostalgia.

[vi] Laibson, epstein, andrea wilson limited memory.

[vii] A current ad for a fast-food chain's chicken sandwich describes a "chicken" husband cowering as his wife urges him to investigate the suspicious sounds downstairs, which sound like a burglar to her. The tail end of the ad mentions the chain's chicken sandwich briefly. Nothing is said about the sandwich's price or quality. Presumably the advertiser's goal was to associate the emotional valence of the term "chicken", in an amusing and memorable way, with the chain's new sandwich.

[viii] Chris Mayer told me this story.

[ix] In “Stroop tasks” an automatic response must be overidden and the failure to do so leads to errors. A classic example is listing a color word and asking the subject to name the color of the ink the word is written in. If the word is “red” people often answer “red” when the correct answer is black.