Economic Theory and Psychological Data: Bridging the Divide*

Andrew Caplin
New York University

1 Introduction

Classical decision theory is a picture of harmony. Models are designed to predict economically consequential choices, and empirical tests concern their ability to explain these choices. The models themselves are typically solved using optimization techniques with solid axiomatic foundations. These foundations allow for a complete understanding of the qualitative and quantitative relationship of the models to relevant choice data. The consensus within economics concerning this decision theoretic methodology has had powerful positive externalities, ensuring strong communication across professional subdisciplines. Economics stands out in this respect from other social sciences, in which the approach to model-building is more fragmented, and in which communication problems are correspondingly more profound.

Despite its harmony, many researchers view classical decision theory as limited in its ability to incorporate psychological factors. As detailed in section 15.2, this has led to ever growing interest in non standard economic models that sit outside the choice theoretic tradition. At the same time, diverse new forms of "psychological" data are being developed, concerning features of informational search, eye movements, neurological responses, and answers to non-standard survey questions. The possibilities for new insights to be gleaned by appropriately combining non-standard theories with novel psychological data excite many prominent researchers (e.g. Benhabib and Bisin [this volume], Camerer [this volume], Crawford [this volume], and Schotter [this volume]).

While the recent expansion in the scope of economic theory and measurement opens up intriguing new vistas, it also raises profound challenges. In particular, it has disturbed the harmonic situation that previously prevailed. In striking contrast with the classical theory, there are expanding gaps between the new economic models and the psychological data that are used to inspire and to test them. Should these gaps continue to expand, it will pose a significant threat to

*I thank Mark Dean for his many contributions to the ideas expressed herein, and his collaboration on the associated research projects. I have also learned much from conversations with Paul Glimcher, Miles Kimball, John Leahy, Andy Schotter, and Ruth Wyatt. I thank the NSF and the C.V. Starr Center at NYU for research support.

the hard-won unity of our discipline. The challenge before us concerns how best to open up new avenues of exploration while retaining the essential coherence of economic thought.

In direct response to this challenge, Faruk Gul and Wolgang Pesendorfer [this volume] (henceforth GP) have proposed a highly disciplined method by which to expand the domain of economics. As detailed in section 15.3, they propose the use of choice theoretic (generally axiomatic) methods to model the potentially intricate impact of psychological factors on behavior, while restricting the data used to test these models to relate only to choices per se.

While sympathetic to the GP proposal with respect to theoretical methodology, I have an entirely different view concerning the data. Section 15.4 explains my belief concerning the great value of incorporating enriched psychological data into economic models. In light of this, I propose in section 15.5 an alternative methodology that encourages essentially limitless expansion in the psychological data that economists model, while setting high ideals concerning model-building techniques. The central pillar in the proposed methodology is the use of axiomatic methods along the lines envisaged by Samuelson [1938]. Rather than identifying theories of choice with unobservable "folk" concepts such as utility and its maximization, Samuelson sought axioms by which to characterize the observable implications of these theories. The methodology I propose herein operates in precisely this same manner on far richer data sets. It calls for axiomatic characterizations of potentially intricate psychological data sets, typically including various forms of standard choice data.

The identifying feature of the proposed methodology is that it is "minimalist". As in standard axiomatic choice theory, one characterizes the precise implications of a given theory for a data set of interest, and identifies as equivalent all models that produce the same predictions. Such tight characterizations reveal in stark for the essential features of the underlying model. Psychological and neurological constructs modeled in this manner will end up being characterized precisely in terms of observable counterparts, as has been the case in utility theory. In this sense, minimalism represents no more than a broadening of Samuelson's original proposal to richer settings. It places the axiomatic method at the center of theoretical enquiry, and imposes no restriction on the formal content of theories.

My deepest hope is that the minimalist methodology will prove of value interdisciplinary research, which is of rapidly growing importance. Typically, folk concepts are imbued with meaning through repeated use within disciplinary confines. As Gul and Pesendorfer [this volume] make crystal clear in their discussion of "risk aversion," interdisciplinary differences of interpretation cause communication problems that may be impossible to understand, let alone to overcome. The defining virtue of the minimalist methodology in this regard is that it removes the ambiguities associated with such folk definitions. Any intuitive constructs that minimalist theories introduce must be characterized directly in terms of observables. In this respect the current proposal is designed not only to retain the unity of economics, but also to permit conceptual unification across current disciplinary boundaries, including the new neuroscientific boundaries between social and natural science.

The importance that I attach to a methodology that allows for social scientific unification may be best understood using language borrowed from the theory of random graphs. Individual pieces of research can be conceptualized as nodes in an evolving social scientific "research graph." The stronger are the links between the nodes in this graph, the longer are the research paths over which each individual researcher can traverse. If social scientists can adopt common linguistic conventions, the end result will be a great strengthening of the connections between the various disparate nodes in this graph. With agreement around a connective methodology such as minimalism, future social scientists may be able to link up and form the kinds of "massive components" needed to overcome challenges that will increasingly disrespect field boundaries. Absent some such agreement, social science as a whole, and economics in particular, will become increasingly fragmented.

In terms of practical implementation, it is clear that a certain maturity of understanding will be necessary before it is appropriate to develop axiomatic models. There may be need both for preliminary investigation of new data sources and for exploration of theories that import concepts such as anxiety and fear directly from folk psychology (e.g. Caplin [2003]). The minimalist methodology is appropriate only when it comes time to unify theory and data. As gaps are found in early models and as new sources of insight open up, theory and measurement will ideally co-evolve. As potentially insightful sources of new data emerge, complementary models will be developed that make predictions concerning these data in addition to standard choices. Conversely, new theoretical approaches will be matched with complementary innovations in measurement.

One possible area of tension concerns the implicit downgrading of constrained maximization. While theories involving optimization may be of particular interest due to their simplicity and their connection to the long stream of prior economic analysis, they are not essential to the methodology. Development of axiomatic models that do not rely on optimization is likely to introduce communication problems with those branches of the discipline in which the principle of constrained optimization has been placed on a higher plane than axiomatic reasoning. My hope is that facing such problems head on will result in the methodology being sharpened rather than discarded.

A crucial issue in applying the methodology is how best to take advantage of the implied freedom to construct the data concerning which we develop our theories. Section 15.6 illustrates two cases in which this question has been adequately answered. The first example involves a minimalist approach to understanding the impact of subjective informational constraints on decision making, and is detailed in Caplin and Dean [2007A]. The second example involves a minimalist approach to neuroscientific data that have been hypothesized to register the degree of surprise as the environment changes for the better or for the worse. Theoretical foundations for this research are detailed in Caplin and Dean [2007B], with the empirical implementation being detailed in Caplin, Dean, Glimcher, and Rutledge [2007].

2 Paradise Lost

Sections 2.1 and 2.2 provide doctrinal background on the revealed preference approach to choice theory proposed by Samuelson [1938]. Section 2.3 outlines how and why model builders learned to by-pass the constraints that this approach entails. Section 2.4 reviews some of the new psychological data that is increasingly available, and the largely extra-theoretical approaches that have taken to discriminating among the new models based on these data. The review highlights the strains that are currently being placed on professional unity.

2.1 Preferences and Utility

Water is essential to life yet is worth far less than such an obvious non-necessity as a diamond. Efforts to explain this "diamond-water" paradox led to various strained efforts to purge theories of "exchange value" of any and all references to "use value", or utility. For example, the labor theory of value explained prices as resulting from the direct and indirect inputs of labor required to produce the object of exchange. Water is cheap relative to diamonds because it is so much easier to produce.

It was the marginalist revolution that indicated how to reinsert utility considerations into price theory. According to the marginalist interpretation, the reason that water is so inexpensive is that beyond the units that are readily available, incremental units do little to improve well-being. Were water in short supply, it would have far higher marginal utility and would surely sell for a very high price. Following this logic to its seemingly natural end-point, many economists of the late nineteenth century took the low price of water as evidence for the principle of diminishing marginal utility, whereby marginal utility declines with quantity consumed. The apparent need for utility to diminish at the margin justified lengthy investigations into how one might best quantify utility, with various measurement devices being proposed.

Speculations on how best to measure utility were cut short by the advent of the ordinalist approach to choice. According to this approach, choice relevant likes and dislikes are summarized by a complete and coherent (transitive) preference relationship among prizes. The observation that one can summarize a choice procedure that respects this relationship as equivalent to assigning utility numbers to prizes and maximizing this function is the first step in realizing the futility of the search for "true" utility. The clincher is the equally transparent observation that one could use any other utility function that preserved the order as an alternative to describe precisely the same pattern of choice.

In terms of understanding prices, the ordinalist approach showed that the low price of water must be explained based not on diminishing marginal utility, but rather on a diminishing marginal rate of substitution among distinct commodities. This is a property that is reflected in the shape of the indifference curves, which in turn derives entirely from the underlying preference ordering. The ordinalist revolution effectively ended the search for "true" internal utility.

2.2 Choices and Utility

In ordinalist logic, the starting point for demand theory is a coherent personal ranking of the potential objects of choice. Samuelson asked a simple question. What exactly are the limits that this theory imposes on demand functions themselves? This question and the subtle answer that it produced gave rise to an implicit critique of the standard ordinalist approach to choice theory. If observations are ultimately limited to choices alone, what is the point in theorizing about inner comparisons? If two distinct theories are identical in terms of the choices that result, are they not to all (economic) intents and purposes equivalent? The revealed preference approach to choice theory follows the logic of this critique to its natural end-point, and fully identifies theories with their behavioral implications.[1]

In technical terms, the move from ordinalism to revealed preference is relatively minor. Ordinalism takes as its starting point a preference ordering over elements of the choice set, X, taken for simplicity to be finite. Revealed preference takes as its starting point a choice correspondence that details all elements that might end up being chosen from a finite set $A \subset X$. The minimal assumption in this approach is that the choice correspondence satisfies the weak axiom of revealed preference (WARP): given $A, B \in 2^X$ and $x, y \in A \cap B$,

$$x \in C(A), y \in C(B) \Longrightarrow x \in C(B)$$

WARP implies that we can use the choice correspondence to define an implied binary relation,

$$x \succsim y \Longleftrightarrow x \in C(\{x, y\}),$$

that is itself complete and transitive, and therefore permits of standard utility representations. The converse is also true: if we derive the choice correspondence from an underlying preference ordering on X, then WARP will be satisfied.

While the technical switch implied by the revealed preference approach seems minor, it is methodologically crucial. Models in this tradition characterize conditions under which observable choices can be calculated using a simple algorithm, such as maximization of a utility function. It is in this sense that economic theories of decision making are "as if" theories. Standard theory does not claim that an agent maximizes a utility function, or possesses a complete, transitive preference ordering. The claim is only that one can model decision makers as if they are utility maximizers if and only if their choices satisfy WARP.

As explored below, revealed preference models have now been developed for choice among many different types of object, with various different assumptions on the nature of the choices, and correspondingly different methods for computing choices. A proponent of the approach axiomatizes specific properties of choice that are to be captured. All remaining objects in the theory are then derived from the corresponding axioms. I take the following general structure to characterize the class of revealed preference theorems.

[1] The reader will find the term "revealed preference" defined in many different ways in this volume. My own definition is entirely technical: I use it to characterize a class of axiomatic models of observables, as detailed below.

> Choices among objects of type * have appealing properties ** if and only if determined by calculations/operations of the form ***.

In this general structure, calculating *** may represent an entirely unrealistic picture of the decision process. Yet the principle of revealed preference places the empirical bar for amending the theory higher. If ** is empirically violated, the task is to identify changes that accommodate this failing. If violations of ** are empirically unimportant, then choice data per se cannot force rejection of the theory. Note also that the underlying "theory" that is being captured in the axioms can be described either as ** or as ***. In the original formulation, the starting point was the theory that individuals maximize utility (***), and the point of the theorem was to characterize choices that are consistent with this class of theory. In many current formulations, it is the property of the choices (**) that is treated as the starting point, and the conclusion concerns the algorithm that this pattern of choice identifies. A final point to note is that the class of such theories is very much tied up with technology, since technological expansion opens up entirely new forms of calculation with which to underpin such theorems.

In addition to the basic characterization theorems, axiomatic models typically include statements concerning equivalent representations. These results provide precise limits on what the underlying theories are able to identify from data on the corresponding choices.

> There exist equivalence classes of calculations/operations of the form *** such that choices among objects of type * are identical within any such class, yet differ between classes.

2.3 Enter the Unobservable Abstractions

The revealed preference approach to choice pushed economic theory in the direction of observationally-based streamlining. All elements that are superfluous to choice are excluded from the theory by definition. This holds in particular for psychological intuitions concerning the manner in which choices are made. However, the asymmetric information and game theoretic revolutions introduced into economics many model elements that have no obvious basis in observation. With respect to presumed information asymmetries, what observation teaches one whether or not an agent is informed? And while out-of-equilibrium strategies are inherently unobservable, there have been long and sometimes fruitful debates concerning how they are chosen. As game theory has increasingly become the dominant approach to microeconomic interactions, so economic theorists have become accustomed to the gap between what can be theorized about and what can be observed.

It was in this theoretical environment that research in economics and psychology began in earnest. The early developments, while formally motivated by the empirical failures of standard theory (e.g. the Allais paradox), were informed by psychological intuitions concerning missing factors of first order importance.

Kahneman and Tversky [1979] were motivated by various violations of the predictions of the standard model, and their prospect theoretic utility function incorporates such standard objects as losses, gains, and objective probabilities. However their broader goal was to develop an approach to choice that reflected the "heuristics", such as anchoring and adjustment, that they hypothesized to be at work in real world decisions (Kahneman and Tversky [1974]).

A similar combination of formal and informal motives underlies the pioneering work of Strotz [1956] on time inconsistency, which forms a cornerstone of modern economics and psychology. The formal motivation for his work derives from the failure of standard economic models to capture use of commitment devices to reduce future flexibility. The informal motivation concerns the desire to introduce psychologically realistic self control problems into economics. Models with time inconsistency continue to have such powerful appeal not only because of their predictive power, but also because of their psychological content. In terms of predictions, David Laibson [1997] and Ted O'Donoghue and Matthew Rabin [1999] have done pioneering research on the many behaviors that are easier to explain with models of this form than with standard exponential discounting. In psychological terms, Loewenstein [1987] and Caplin and Leahy [2001] have shown how time inconsistency can be produced by psychological forces seemingly unrelated to self control problems, such as anticipatory feelings about future outcomes.

The phenomenon of time inconsistency powerfully illustrates the extent of the departure from revealed preference in economics and psychology. The field has by and large adopted the proposal of Peleg and Yaari [1973] that models with time inconsistency are best solved using game theoretic concepts, such as perfect equilibrium. Once an individual is regarded as involved in a strategic interaction with future selves, questions are opened up on many levels. What is the appropriate solution concept for the resulting game, and what does this imply about out-of-equilibrium beliefs? Are agents aware of their future preference changes? Finally, while the original Strotz model involves a unified individual who changes preferences over time, alternative conceptualizations are now available in which the individual makes cue-triggered mistakes ([Bernheim and Rangel [2004]) or suffers from intra-personal conflict within a given period (e.g. Benhabib and Bisin [2005], Fudenberg and Levine [2005], and Loewenstein and O'Donogue [2004]). Once one regards individual decisions as the outcome of a dynamic game with more than one player in each period, it is clear that the principle of revealed preference has been left far behind.

2.4 Psychological Data and Economic Theory: The Divide

As noted above, models in economics and psychology typically operate both on the behavioral level and on the level of psychological. With regard to the purely behavioral aspects of the underlying psychology, it is clear how to bring these models to the data. What is less clear is how to bring the more intuitive content, such as heuristics and self control problems, to the data. In this regard it is perhaps fortunate that the growth of the field has coincided with a revolution in data availability. This flood of new data is reflected in various empirical papers in economics and psychology.

In terms of measuring heuristics, Payne, Bettman and Johnson [1988] developed a program, Mouselab, precisely to provide new evidence on the process of decision making. The Mouselab system is a computer based way of tracking the order in which information is processed in a choice problem. It presents the subject with a choice amongst several alternatives, each of which is characterized by a number of attributes. All the attributes are given numerical values. The vector of attributes for each of the options is presented on a screen in a grid form. However, each of the numbers in the grid is obscured. In order to view a number, the subject has to move the mouse cursor over the relevant box. By recording these movements, the experimenter is able to view the order in which information is acquired and the speed with which it is acquired.

One experiment outlined by Payne, Bettman and Johnson [1988] used Mouselab while varying the choice environment along two dimensions. First, they varied the dispersion of importance of the different attributes from low to high. Second, they varied the degree to which the subject is subjected to time pressure as they make their decision. From the Mouselab system, they recorded various measures of the way that subjects collect information from the grid. They focused on identifying qualitative features of the resulting data, observing for example that the high dispersion environment saw more time spent on the most important attribute, and higher variance in time spent on each attribute and alternative. Payne, Bettman and Luce [1996] ran experiments using Mouselab in which they opportunity cost to time was allowed to vary.

A second new form of data that is increasingly available is response data of various sorts, including survey responses. The ever decreasing cost of internet surveys suggests that availability of such data is soon to mushroom. One potentially interesting set of psychological questions that have been posed concern planning activities. Ameriks, Caplin, and Leahy [2003] developed a new survey instrument to get insights into planning behaviors and forethought, an area of economics and psychology in which Lusardi [2001] pioneered. They provided results suggesting that those who with a high "propensity to plan" generally accumulated greater wealth. Another burgeoning area of research in economics and psychology concerns the analysis of data derived from surveys of happiness (Kahneman et al. [2004]).

A third important new source of psychological data derives from neuroscientific evidence (Glimcher and Rustichini [2004]). McClure, Laibson, Loewenstein, and Cohen [2004] have pioneered in the use of data on brain activation to get evidence on self control. Inspired in large part by the findings of Schultz [1998], the literature on dopaminergic responses has also been used to shed light both on preferences and on beliefs. Other experimental work suggests that the firing of dopamine neurons may record a reward prediction error (the difference between received and expected reward) rather than just the pure expectation of reward (Berns et al. [2001] and Bayer and Glimcher [2005]).

Unfortunately, these exciting new forms of data are often poorly integrated into the underlying models. Models of heuristic decision making are not rich enough to predict mouse clicks, rendering tests based on these behaviors more indicative than formally useful. The survey questions that are being explored concerning such variables as planning and happiness deal with objects that are undefined in the underlying models. Finally, of all areas, it is the neuroscientific data that is hardest to interpret. Self control based models of choice do not include a cortical or limbic BOLD signal function. Models of expectations are similarly silent on the dopaminergic response. The fact that neuroscientific data are being interpreted as relevant to models of choice is particularly disturbing to proponents of revealed preference. Hence it is no surprise that it was a proposal to expand the use of these data that triggered the current methodological debate (Camerer, Loewenstein, and Prelec [2005]). How can we prevent economics from being swamped by masses of new data of dubious interest to our progress in understanding choice?

3 Choice Theory and Psychology

There is nothing per se irrational about having brain scans, mouse clicks, and survey responses move one toward acceptance of models in which they do not formally appear. A Bayesian can update personal priors based on evidence that is relevant to a "large world" model, yet is not present in the "small world" theory that is being tested. Yet updating beliefs based on such a personal world view has its dangers. Those who are hold different priors over the large world may update in an entirely different manner, and no amount of evidence is guaranteed to provide convergence of posteriors. This process of updating based on extra-theoretical evidence may have as its end-point a highly fractured profession. Absent discipline, our proud and unified social science risks devolving into many little fiefdoms, each with its own private language concerning how to interpret evidence.

Surveying the scene with their eyes fixed on observables, Gul and Pesendorfer [2008] posed anew the question of Samuelson. If two distinct psychological theories produce the same implications for choice in all situations, are they really different in a way that matters to economists? Believing this question to have a negative answer, GP propose that the standard revealed preference approach to choice theory be restored to its traditional centrality. They propose replacing all theories that invoke direct assumptions on unobservable psychological phenomena with theories that are phrased directly in terms of choice alone. This proposal, and its seemingly devastating implications for economics and psychology, have profoundly caught the attention of the profession.[2]

[2] It is hard to understand why issues of observability have attracted so little attention in interactive contexts. How could one ever know whether equal and immediate division in a bargaining problem reflected subtle mutually held beliefs concerning out-of-equilibrium behavior, or instead reflected simple application of conventions? It might be argued that the former provide the simplest generalizable account of such outcomes, since conventions are fragile. Hausman [2000] highlights this issue in discussing weak points in the conception of weak preference. For those who are deeply attached to the latter, it may be seen as highlighting weaknesses in game theory. Yet this debate has not been joined, since the subject

I illustrate in this section the potential psychological power of the methodology proposed by GP. Section 3.1 indicates the general method by which expansions in the domain of choice are used to increase the psychological range of choice theory. It is precisely the fact that this approach focuses direct attention on the domain of choice that accounts for its advantages in terms of parsimony and elegance over psychologically-driven theories of behavior. Sections 3.2 and 3.3 present an example that GP themselves stress as a perfect exemplar of their approach. The underlying subject concerns the impact of mental states such as anxiety and suspense on choice of information. Section 3.2 develops a directly psychological model of these phenomena due to Caplin and Leahy [2001]. Section 3.3 show how to apply the axiomatic model of Kreps and Porteus [1977] to capture the key behavioral implications of these psychological forces without ever having to define them or to measure them. The broad point is that one does not need to specify psychological phenomena directly if one's only goal is to study behavior. This point is illustrated again in section 3.4 and 3.5 in the context of models of bounded rationality and of the decision process. These examples also make clear why maximization is less essential to choice theory than is the axiomatic approach. Current axiomatic models of boundedly rational decision making are explicitly procedural, and stand or fall on the behavioral implications of these procedures regardless of whether or not they can be phrased as in some sense optimal.

3.1 Domain of Choice and Psychological Range

It was the theory of choice under uncertainty due to von Neumann and Morgenstern [1947] that first indicated the potential psychological power of the choice theoretic approach. Their expected utility theory model is explicit in conceptualizing objects of economic choice as psychological constructs rather than as physical objects. The implicit hypothesis is that choice under uncertainty is equivalent to choice among objective lotteries. This represents a huge psychological advance over viewing prizes as physical objects. Of course, there is no direct way of knowing if uncertain winnings are in fact perceived as objective lotteries, which is what makes the connection to implied patterns of behavior so crucial.

Savage [1954] took the process of abstraction from the physical prize to a psychological construct one step further, applying axiomatic methods to choice among acts that associate consequences with all states of nature. This removed the need to assume acceptance of commonly understood objective probabilities. The results that Savage demonstrated based on his conceptualization of the domain of choice stand as among the most remarkable achievements in social science. With his axioms, one and the same process of inference from concepts that are purely choice-based give rise not only to the implied structure of preferences, both also to the implied structure of subjective beliefs.

The above models illustrate the connection between the domain on which

of observability has not struck a chord in the strategic context.

a choice theory operates and the psychology that the theory can encompass. Using as the domain a purely physical prize space disallows psychological phenomena connected with uncertainty. The domain of objective lotteries is clearly richer in this regard, and allows one to capture some of the financial implications of the psychology of risk (e.g. interest in insurance). However the range of psychological questions that can be shoe-horned into a model with such a simple domain remains limited. For example, the model of Kreps and Porteus [1978] outlined below shows that a richer domain is needed account for non-instrumental information preferences. The fundamental point is that no theory of choice can be more psychologically sophisticated than is allowed for by the domain on which it is defined. If objects are not differentiated in the domain, there can be no distinction between them at any later stage of the theory.

A domain that is particularly rich in terms of the psychology it allows to be modeled is the set of subsets of a given physical prize space, rather than the prizes themselves. Kreps [1979] used this domain to model preference for flexibility. Yet it is the GP model of temptation and self control that most profoundly illustrates the rich psychology that this domain liberates (Gul and Pesendorfer [2001]). As noted above, Strotz [1956] argued that time inconsistent preferences that over-emphasize the present give rise to demand for commitment devices. GP turned this around, and asked what interest in commitment devices reveals about preferences over choice sets. The resulting model implies that larger choice sets may not be for the better, yet does not allow for time changing preferences.

3.2 Belief-Based Utility and PEU

Traditional expected utility theory has the property that information is purely instrumental. Hence more information can never be either a bad thing, or desirable for its own sake. Yet many psychological arguments suggest that factors such as anxiety, fear of disappointment, love of suspense, and possible regret may impact preferences over information for reasons that go beyond pure instrumentality. To formalize the modeling, consider a medical setting with two possible health states, θ_1 and θ_2, in which outcome θ_2 is having Huntingdon's disease (an incurable degenerative disease with onset at some future date) and outcome θ_1 is not so having. Let $p \in [0, 1]$ denote the probability of not so having, as assessed in the period prior to the actual final resolution of the uncertainty, yet following receipt of a signal $s \in S$ concerning this outcome. Each signal is associated with information that impacts the probability that state θ_1 will eventuate: it gives rise to a probability distribution over beliefs.

In the case of Huntingdon's disease, there is a simple genetic test that can be taken to resolve the $p = 0.5$ chance of having the disease that exists among the offspring of prior victims. Yet many who have been offered the right to uncover their genetic fate have chosen not to take the test. The primary motivating factor in these cases appears to be psychological rather than monetary (commonly the blood has already been sampled, so that the only question is whether or not the patient is passed the information that the responsible medical authorities already know).

One way to model the forces that give rise to information rejection is to directly model the impact of belief-based prizes on utility. This is the approach adopted in the psychological expected utility model of Caplin and Leahy [2001] (henceforth PEU), which explicitly connects such preferences with a particular psychological response to information. In the context of Huntingdon's disease, the model relies on the psychological hypothesis that subjective contentment in the pre-outcome period depends on the operative belief state, p. To capture this formally, the PEU model posits existence of a function that calibrates the cost in expected utility terms of the aversive state associated with pre-outcome beliefs p. While the name is unimportant, "anxiety" is a reasonable psychological label for the aversive mental state in the pre-outcome period.

A simple way to analyze the choice theoretic implications of the PEU model is to work directly in a prize space comprising all lotteries over pre-outcome beliefs and actual outcomes. Pure prizes are belief-outcome pairs of the form (p, θ_j) with $p \in [0,1]$ and $j \in \{1,2\}$, and the PEU function is defined on the set of all such pure prizes,

$$Z = \{(p, \theta_j) : 0 \leq p \leq 1; j \in \{1,2\}\}.$$

According to the model, one lottery over these prizes is preferred to another if and only if it has higher expected utility according to some utility function $u^{ANX} : Z \to R$. In the context of Huntingdon's disease, the natural psychological hypothesis is that the best prize is good news early, while the worst is bad news early. Normalizing as usual, we know that for all belief-prize pairs (p, θ_j) with $p \in [0,1]$ and $j \in \{1,2\}$,

$$1 = u^{ANX}(1, \theta_1) \geq u^{ANX}(p, \theta_j) \geq u^{ANX}(0, \theta_2) = 0.$$

There are two reasons that $(1, \theta_1)$ is so good. First, the ultimate outcome is good. Second, receiving the good news early alleviates the anxiety associated with living with fear of impending doom. Similarly, $(0, \theta_2)$ is so bad not only because the ultimate outcome is bad, but also because it forces the agent to live in fear.

To understand the impact of the model on choice of information, note that the probability that belief p will be paired with outcome θ_1 is p. The other possibility, that it will be paired with outcome θ_2, has probability $1-p$. Hence one can define the function $K^{ANX}(p)$ as the overall expected utility corresponding to this belief concerning final outcomes,

$$K^{ANX}(p) \equiv p u^{ANX}(p, \theta_1) + (1-p) u^{ANX}(p, \theta_2).$$

It is this function that determines choice among signals. Given signals $s, s' \in S$, one signal is preferred to another if and only if it produces a lottery over beliefs p that yields higher expected utility according to the function $K^{ANX}(p)$.

$$s \succsim s' \text{ if and only if } E_s(K^{ANX}) \geq E_{s'}(K^{ANX})$$

Note that information is rejected if this function is concave, and that such a rejection permits of a psychological interpretation. Information is rejected if increased pessimism in the face of a bad signal is regarded by the agent as more aversive at the margin than increased optimism in the face of a good signal is beneficial.

3.3 Avoiding Anxiety or Preserving Surprise?

I show now that the PEU model contains redundant elements from a purely choice theoretic perspective. Consider a PEU model designed to capture positive surprise. Outcome θ_1 is that a birthday party is thrown for a close friend. Outcome θ_2 is that no such party is thrown. The friend's post signal pre-outcome belief that the party will be thrown is p. To model this as a preference over prizes and beliefs, one makes the psychologically natural assumption that the best prize is good news late, while the worst is bad news that is disappointing. Normalizing again, this gives rise to the following restrictions on the underlying expected utility function on belief-prize pairs (p, θ_k) with $p \in [0,1]$ and $k \in \{1,2\}$,

$$1 = u^{SUR}(0, \theta_1) \geq u^{SUR}(p, \theta_k) \geq u^{SUR}(1, \theta_2) = 0.$$

As before, informational choice depends on the appropriate weighted average as captured by $K^{SUR}(p)$,

$$K^{SUR}(p) \equiv p u^{SUR}(p, \theta_1) + (1-p) u^{SUR}(p, \theta_2).$$

Note that the attitude to information derives from an almost polar opposite psychology in the case of the surprise party than in the case of Huntingdon's disease. In the former case pessimism is beneficial, while in the latter it is per se unpleasant. Yet it is easy to construct cases in which revealed preferences among signals are identical. In particular, the following linear case meets all of the intuitive criteria that can be imposed on the psychology of anxiety and of surprise,

$$\begin{aligned}
u^{SUR}(p, \theta_1) &= 1 - \alpha^{SURP} p, \text{ with } \alpha^{SURP} \in (0,1); \\
u^{SUR}(p, \theta_2) &= (1-p)\beta^{SURP}, \text{ with } \beta^{SURP} \in (0,1); \\
u^{ANX}(p, \theta_1) &= \alpha^{ANX} p + (1 - \alpha^{ANX}), \text{ with } \alpha^{ANX} \in (0,1); \\
u^{ANX}(p, \theta_2) &= \beta^{ANX} p, \text{ with } \beta^{ANX} \in (0,1).
\end{aligned}$$

With $\beta^{SURP} = 0.15$, $\alpha^{SURP} = 0.35$, $\alpha^{ANX} = 0.1$, and $\beta^{ANX} = 0.5$, the two functions are increasing affine transforms, and hence behaviorally equivalent,

$$K^{SURP}(p) \equiv 0.15 - 0.2p^2 + 0.7p = 0.15 + 0.5 K^{ANX}(p).$$

The fact that two entirely different psychologies produce identical choices implies immediately that they are indistinguishable from a standard choice theoretic perspective. From the perspective of choice theory, psychological motivations are just as hidden as are decision processes. In both cases the decision theorist can look for the projections of the underlying factors onto the domain of choice, and provide models that mimic in this domain any behaviors that may be uncovered through other forms of research. In this interpretation, psychology may be inspiring, but it is not to be modelled.

3.4 A Choice Theoretic Approach

In fitting with GP proposal, what is needed is a choice theoretic model that captures the behavioral implications of psychological forces such as anxiety and love of surprise. The model of Kreps and Porteus [1978] provides just such foundations, and is correctly regarded by GP as an exemplar of their approach to economics and psychology. Kreps and Porteus (henceforth KP) take on exactly the same challenge as does the PEU model: that of how to allow for beliefs that may directly impact choices. However they do not make any direct references to psychology. Instead they take a choice theoretic approach, looking to summarize properties of choice among signals. In order to do this, KP apply axiomatic methods of choice theory to the domain of temporal lotteries ("lotteries over lotteries over ...") rather than simple lotteries as in expected utility theory.

The KP model involves axioms such that signals are ranked in a simple and transparent fashion. Specifically, these axioms are equivalent to existence of a function $K : [0, 1] \to R$ such that given $s, s' \in S$,

$$s \succsim s' \text{ if and only if } E_s(K) \geq E_{s'}(K).$$

One signal is preferred to another if and only if it produces a lottery over beliefs p that yields higher expected utility according to this function. This is exactly as in the PEU model, with the only difference being that the function $K(p)$ preserves consistency in the order as between beliefs and outcomes: if prize θ_1 is strictly preferred to prize θ_2 so that $K(1) > K(0)$, then the K function is strictly increasing; it is strictly decreasing if the inequality is reversed; while if $K(1) = K(0)$ all beliefs are equally ranked and all signals are indifferent. Again, as in the PEU model the function $K(p)$ fully identifies choice among signals. The concave case produces what is known as preference for late resolution of uncertainty, in which a signal that provides no information is preferred to a signal that resolves any of the uncertainty up front. Similarly the convex case produces preference for early resolution of uncertainty. Linearity is equivalent to neutrality with respect to the temporal resolution of uncertainty.

While at first appearance it may be a source of comfort that the KP model and PEU theory have such strong similarities, the differences between them are profound. The critical distinction is that the domain on which preferences are defined is strictly smaller in the KP model. The KP utility function has in its domain any and all lotteries that may be available to an agent choosing among information gathering strategies. By construction, this domain exhausts all possibilities for private choice. In contrast, the PEU function is constructed on the domain of lotteries over beliefs and final prizes. This space contains objects to which Bayes' rule does not apply, such as belief-prize pair $(0.5, \theta_1)$

for sure: an ex ante belief that θ_1 and θ_2 are equally likely, yet the realization of outcome θ_1 for sure. Unfortunately, a guarantee that a currently uncertain future will turn out well is not available as an object of choice. It is this extension of the domain of preferences into the realm of fantasy that makes the PEU model fall afoul of the principle of revealed preference.

As noted above, one issue that the enriched domain gives rise to is interpretational ambiguity. An individual who will pay to avoid information is certainly displaying a preference for late resolution of the form modeled by Kreps and Porteus. Yet one cannot be similarly confident that the psychological motivation is avoidance of an aversive prior mental state induced by pessimistic beliefs, since precisely the same pattern of behavior can be observed in cases in which pessimistic beliefs are beneficial as opposed to damaging.

3.5 The Decision Process and Bounds on Rationality

One motivation for models of bounded rationality is that they are more intuitively appealing than the standard optimizing model. In fact the first formal model of bounded rationality, the "satisficing" model of Herbert Simon [1955], was introduced as a direct response to the perceived implausibility of the prevailing model of economic man. Simon proposed a theory of choice by which the decision maker has some concept of the characteristics that would make an option satisfactory. The choice set is searched sequentially until an option is found which fulfils those criteria, which is then chosen. A second attempt at designing a model of the decision making process based on plausibility is Sauermann and Selten's aspiration adaptation model (cited in Selten [1999]). As its name suggests, the model describes a set of rules by which a decision maker adjusts the aspiration level in the process of search until only one choice is left.

While the early bounded rationality theories had an aura of plausibility, it was not clear whether or not they were different in terms of implied choices from the standard optimizing approach. Later theories were more explicit in differentiating the implied behaviors from those that would be produced in standard choice theory. The "elimination by aspects" model of Tversky [1972] specifies an environment in which a choice object has a number of well specified attributes, or "aspects". The procedure that Tversky posits is one in which the decision maker orders these aspects, and sequentially rules out alternatives that are not satisfactory on the basis of each aspect in turn, until left with a unique choice. It is easy to produce examples whereby application of this decision making process gives rise to violations of WARP: the theory is behaviorally distinct from standard rationality based choice theory. Similarly such heuristic rules as the "search for reasons" (Rubinstein [1998]) can cause behaviors that violate standard principles of rational decision making.

Heuristics are rational in the sense that they appeal to intuition and avoid deliberation cost, but boundedly rational in the sense that they often lead to biased choices. While there are examples in which various heuristics describe behavior better than does rational decision theory, the question arises as to whether a heuristic would be applied in an environment in which massive losses would result. Many doubt that this would be the case, and as a result have proposed models that allow adaptive selection among heuristics. Payne Bettman and Johnson [1993] have used Mouselab to provide evidence suggesting that the choice environment indeed impacts the heuristic employed, thereby edging toward a theory of rational choice among these new primitives.

An entirely different approach to bounded rationality involves formalizing specific constraints on the decision maker's abilities, and solving for optimal strategies subject to these constraints. Wilson's [2002] analysis of an agent with bounded memory and Radner and Rothschild's [1975] examination of the allocation of effort solve optimization problems of this variety. Variations on this optimality based theme include models that are designed to provide simple and plausible solutions for models that are too complex for an agent realistically to solve. Gabaix and Laibson [2000] propose just such a decision rule based on an assumption of myopia, whereby agents choose sequentially among cognitive operations to maximize expected benefit as if they are going to stop and choose in the next period.

3.6 Axioms Please

There are by now a bewildering number of different models of the decision making process, and there is little consensus concerning how best to discriminate among them. In fact many models of process have the defect that it is unclear what restrictions they place on choice data. The typical way this issue is currently addressed is by designing an experimental environment in which predictions are easy to derive. However, when these models perform poorly, one does not know whether to adjust the assumptions that were made to operationalize the models or the overall modeling approach. In contrast, test of standard rational choice theory based on the Weak Axiom of Revealed Preference operate at a high level of generality. The tests explore whether the data is consistent with a class of decision making models, rather than just with a specific process in a particular narrow setting.

Given this background, it is not surprising that there is currently an attempt to provide choice theoretic models of bounds on rationality and the decision making process.

Tyson [2005] provides an axiomatic treatment of a version of the satisficing procedure. Manzini and Mariotti [2004] consider what they term the "rational shortlist" method of choice. Their axioms characterize a class of decision making algorithms associated with two orderings rather than one. For any given choice problem, the decision maker uses the first ordering to construct a shortlist comprising all maximal that are maximal. The agent then selects from this "shortlist" all elements that are maximal according to the second ordering. A third paper in a similar vein is Ergin's [2003] model of costly contemplation. Ergin considers decisions in which people have to choose choice sets from which they will have to pick an element at a second stage. Thinking harder may provide a finer partition, and thus allow a better first stage decision, but also may require more effort, represented by a cost function c.

While the choice theoretic modeling of bounded rationality is at an early stage, it is quickly reshaping the theoretical landscape. The parsimony of these early choice theoretic approaches stands in pleasant contrast to the prior models that were explicitly procedural, yet whose empirical content has proven so hard to divine. Moreover, what is of interest in these models is precisely how they depart from the standard theory of utility maximization. This makes the models particularly important in methodological terms, in showing the value of providing axiomatic models that capture how psychological forces may give rise to departures from standard treatments of constrained optimization.

4 Modeling Psychological Variables

The GP proposal involves an implicit prohibition on modeling psychological variables. Their proposal in this regard is outlined in section 4.1. The remaining sections show that there are available many models that allow better use to be made of the new psychological data. Section 4.2 makes this argument in the context of models that can be fit using survey responses. Section 4.3 repeats the argument for models that make direct predictions of various non-standard data items related to the decision making process. Section 4.4 turns to models that are predictive of data on internal states, such as PEU. In all such cases, enriching the theoretical and evidentiary base allows the associated models to perform to a higher standard than those that are restricted to exclude the non-choice data. Direct modeling of psychological data allows one to fit data on pure choice better than would be possible were these data to be excluded on the basis of a methodological prohibition. Many of the other papers in this Handbook provide rich examples along similar lines, in particular those of Benhabib and Bisin [2008], Camerer [2008], Crawford [2008], and Schotter [2008].

4.1 The GP Proposal

GP argue that modeling choice is the central task of economic theory, and those engaged in research in economics and psychology have yet to provide a strong counterargument. Yet even if these are shared priorities, they do not provide an answer to the question of what to do about the new psychological data that is increasingly available. Certainly, one can criticize as ad hoc the common procedure of modeling only choice data, yet commenting on the model fit based on data that is not included in the model. But the GP proposal in this regard is also subject to criticism.

GP allow that one can use any form of data one wants to guide intuitions concerning choice. Thinking of psychological forces such as anxiety and self control problems may provide inspiration concerning the behaviors to model. Yet psychological forces have no place in the model itself, so that psychological data have no role in formally fitting the models. Similarly it may be that survey answers are suggestive of forces that drive behavior, and that thinking about these forces improves understanding of behavior. One may even use neurological data and data on the process of choice to derive conjectures concerning the likely pattern of choice in various contingencies. However, their proposal is that one should limit the use of this data to precisely such a suggestive role. One can seek choice theoretic inspiration from psychological data, but psychological variables themselves must ultimately be excluded from the theory.

The GP proposal privileges data on choice above all other forms of data. In fact, the approach is built up as if no additional data were available. If one were indeed to have data only on standard choices, it is hard to see why one would want to construct two distinct explanatory frameworks. When measurement and theory are not tightly matched, differences among competing theories will reside in the unobservables. Those looking for common theoretical ground will remove as many references as possible to these variables. In this sense, the GP proposal represents the logical limit of the theorist's natural desire to remove superfluous elements in a world in which only choice data is available. Yet this leaves their approach ill-suited to a world in which rich new sources of data do in fact exist.

4.2 Strategic Survey Responses

Surveys responses provide a particularly rich form of data that may demonstrably help in fitting models of choice. After all, models of behavior are constructed in large part to help answer questions about future contingent choices. The fact that exactly the same situation has not been seen previously is what makes it important to have an all-purpose picture of decision making, in particular for those interested in making policy interventions. The standard method that revealed preference allows for filling in this data is extremely limited: moment-matching based on data from past choices. It is easy to find cases in which these data turn out to provide little of the information one is seeking. This is particularly so when one is looking to explore the impact of policy changes that are likely to result in something of a break from past patterns of behavior. In many such contexts the complications involved in designing a realistic experiment make it far too expensive a proposition, at least until and unless some preliminary evidence has been gathered using less costly means, such as conducting surveys.

A simple question that has proven very hard to answer concerns the impact of an estate tax reduction on wealth accumulation. Kotlikoff and Summers [1981] argued that the bequest motive is the primary driver of wealth accumulation, which implies that the estate tax is a crucial determinant of spending in the retirement period. Hurd [1987] has argued that precautionary motives provide an equally attractive explanation for the observed pattern spending among the elderly. Since that time efforts to distinguish in the data between bequest and precautionary motives have proven so unsuccessful that such prominent researchers as Dynan, Skinner, and Zeldes [2002] have argued that they may be empirically indistinguishable. Hence standard economic models based on behavior alone provide at best weak evidence on which to base any changes in estate tax policy.

In current research, a “strategic survey” methodology is being developed to improve inference in dynamic optimization models of this sort (Ameriks et al. [2007]). This methodology exploits the fact that the decision maker’s strategy is potentially far more revealing than is behavior alone.[3] The goal of a strategic survey question is to explore informative aspects of the agent’s strategy that may not have been revealed in the prior history. Seeking data on strategies is consistent with the spirit of revealed preference, according to which choice behavior is more revealing than any statement of priorities could ever be. In philosophical terms, the strategic survey methodology is a minimum departure from standard observation of behavior, and as such represents the natural next step when purely behavioral data are unrevealing.

In Ameriks et al. we confirm technically the limitations of behavioral data in separating bequest and precautionary motives, and design specific strategic survey questions to shed new light on their relative importance. In technical terms, the critical issue is how to fit a model of behavior and survey responses to data of both different forms. This calls for models of the survey error process: the gap between survey responses and what would in fact occur in the given contingency. As do Kimball, Sahm, and Shapiro [2005] and Ljunqvist [2003], we develop an explicit model of the survey error process in estimating our model. While our ultimate interest lies only in behavior, modeling the survey data is a necessary means to achieving this standard behavioral end. Fitting the model based on past behavior alone would needlessly limit our understanding of future behavior.

4.3 The Decision Process

The resources that go into making a decision are scarce. To fully assess all the possible alternatives in a decision problem is going to take both time and effort. In some cases, if the alternatives are relatively simple and the stakes for making a good decision are large, this may be a good idea. However, in other cases, where the alternatives are more complicated and the cost of the wrong decisions is relatively small, this may be not such a good idea. In these cases, the decision maker may chose to cut some corners, either by choosing to ignore some of the options available, or by not enquiring into some of the aspects of some of the options. If this is the case, a model that hopes to well explain choice must explain how the decision maker searches for information, how they decide to stop searching and how they make a decision based on the information that they have. For this reason, models of the decision making process may provide insight into many aspects of the decision in addition to what is finally chosen.

A simple variable that has often been used to provide additional insight into the decision making process is time to decide. For example, Gabaix et al. [2006] have used decision time and other additional observables to examine the success of their directed cognition model. Within psychology, the work of Dean

[3] Barsky, Juster, Kimball, and Shapiro [1997], Kimball and Shapiro [2003] and Kimball, Sahm, and Shapiro [2005] have done much to blaze this methodological trail.

et al. [2007] has similarly indicated the impact of decision time on choice. They consider a class of models of aiming at a fixed physical target in which the degree of accuracy depends on the time available to point at this target. Their model produces a time dependent error around a clear objective, with convergence to an irreducible individual level of error. Their model predicts accuracy far better than one would be able to if one excluded reference to the time constraint and looked merely at the projection down of the decision strategy onto the choice space as standardly conceived, which is the geometric region in which the target is constrained to lie. Even more basic in this regard is evidence concerning the time it takes for sensory information such as sights and sounds to effectively register, and the differential speed with which various parts of the brain are now known to operate. A classical illustration of this effect at work lies in the finding that it takes longer to read the word red in green letters than in red letters (the "Stroop" effect discussed by Camerer [2008]).

When explicit models of the time taken in making a decision are developed, richer data can in principle be used to better fit the entire set of available data, potentially offering a more robust model of the decision itself. This is the spirit in which Gabaix et al. [2006] use Mouselab to fit non-choice variables that are predicted by their directed cognition model of information acquisition. The non-choice variables they model and fit are the number of boxes opened in each column, the number of boxes opened in each row, and the time taken to make a decision. In an analogous manner, Johnson, Camerer, Sen and Rymon [2002] use Mouselab data in fitting models that characterize the impact of strategic sophistication on the outcomes of various games. Crawford [2008] provides convincing evidence of how richly search data of this form can inform prediction. Camerer [2008] provides a very rich description of psychological data sets that either currently exist, or are clearly emergent. He presents rich evidence indicating how much these data may enrich our understanding of decisions. It is quite out of the question for these rich new sources of data and of insight to be ignored.

4.4 PEU and Psychological States

While the KP model has advantages over PEU in the purely decision theoretic context, PEU has complementary strengths. In fact the direct inclusion of psychological factors in the PEU model opens the door to rich analysis of what determines preferences over signals. To capture this, the model includes a production function relating the evolution of psychological states to the external environment. Making psychologically natural assumptions on this function, Caplin [2003] models the impact of attention grabbing messages on medical choices. Health authorities often use vivid "fear appeals" to stimulate learning and prevention among at-risk agents who might otherwise neglect possible preventive acts, such as cancer screening. PEU theory can capture this by making the level of anxiety depend not only on objective risk, but also on attentional resources that the messages are designed to commandeer.

The empirical research agenda associated with such an application is clear.

Since the psychological production function is responsible for driving choice relevant psychological states, empirical work must focus not only on fitting choices, but also on fitting the production function for choice relevant psychological states. Just as physical production functions represent the stable underlying structure that survive parameter shifts, so does its psychological counterpart. Only by recovering the underlying production function will we be able to make robust predictions in the face of the many policy changes the model is designed to explore.

Once one takes account of the psychological production function, it is clear that PEU theory may make predictions for data other than standard choices. In this sense, it has a richness that is lacking in the KP model, which seems all to the good. For example if psychological states can be measured and used in fitting the theory, then it becomes easy in principle to distinguish anxiety-avoidance motives for informational choice from those that are based on the desire to be surprised. While such measurement is very hard with present technologies, the idea that this research should be ruled out on the grounds that it can only be tested based on non-standard data seems strained.

There are normative as well as positive reasons for developing models in which the domain extends beyond that required for standard private choice. Even though it was shown in section 3 above that PEU models of anxiety and of surprise may be characterized by identical private decisions, the incentive of an empath to pass on good news is very different (see Caplin and Leahy [2004]). In the former case, the choice will be made to pass on good news early, since optimistic beliefs are always to the good, while in the latter it will be withheld since low expectations are beneficial. In this respect it seems that it is the domain of the KP model that is too limited, rather than that of PEU being excessive.

5 Revealed Preference and Minimalism

The research outlined above into the joint properties of standard choice data and psychological data will move forward whether or not research methodology adjusts. Yet moving forward with this research absent methodological discipline presents severe dangers to professional unity, as outlined in section 5.1. The good news is that just such discipline is at hand if we adapt the techniques pioneered in revealed preference theory to cover psychological data. To realize this, all one has to do is to open up the question of what constitutes a choice for the purposes of revealed preference theory. Section 5.2 argues that modeling verbal evidence as chosen is consistent with a broad interpretation of the principle of revealed preference. Section 5.3 goes further and argues that any and all measurements of human activity can in principle be modeled as chosen, and therefore be incorporated into the mainstream of economic methodology. Section 5.4 details the resulting "minimalist" methodology, which retains the insistence on streamlined modeling matched tightly to well-defined data that characterizes revealed preference. However the approach is entirely open in terms of the variables that can be modeled as chosen. Rejection of vast amounts of previously unmeasurable individual data is not only inconceivable in practice, it is also methodologically incoherent.

5.1 Dangers of Model Proliferation

There is an unstoppable momentum behind efforts to build rich new models to predict survey responses, mouse clicks, and neurological responses in addition to standard choices. What is unclear is how much economic content these models will have, and whether or not they will achieve the high standards of coherence and parsimony that previously characterized economic model-building. One problem is that economists have historically tended to focus on behaviors close to our natural concerns, with a clear concept of the motivations driving behavior. The new psychological data that we are trying to model are entirely different. We have little or no experience modeling these data, we have very limited understanding of underlying measurement problems, and we do not have a clear conceptual framework relating the new measurements to standard choices. Given that the new variables are so unfamiliar to us, there is profound risk of ill-discipline in the modeling process and of professional fragmentation. What discipline can be added to keep this expansion from giving rise to a proliferation of different models that are mutually incompatible not only in terms of the details, but also in terms of the over-arching approach?

I believe that it is relevant to ask of the new richer models the precise question that was originally asked in the context of models of choice, and that gave rise to the original revealed preference proposal. Suppose one fits a rich model that predicts standard choices in addition to new psychological data, and one wishes to improve this model. Are the problems fundamental, or do they relate to the underlying conceptualization? If we are to answer questions of this form, we will need to discipline the modeling process to capture the implications for the data of general structural assumptions concerning the data generating process. It is precisely this drive for generality that gave rise to revealed preference theory in the narrower domain of classical choice data. Why not apply precisely the same discipline in the context of the richer psychological data? In order to explore such an approach, we first need to reconsider the limits of the original principle of revealed preference.

5.2 Words are Chosen

Revealed preference modeling may be applied whenever one conceptualizes a choice as having made from among a set of feasible objects. But when is and when is not a choice being made? During the discussion of behaviorism in psychology the point was made forcefully that verbal statements of all kinds may be conceptualized as chosen behaviors. Ericsson and Simon [1984] proposed recording subjects' "talk aloud" procedures when performing complex decision tasks. They felt that the systematic collection of these types of observations could be used to test models of human reasoning and of decisions. To counter the obvious behavioral counter-argument, they argued that there is no reason for a behaviorist to rule out such "verbal protocols" as evidence for their theories.

> "We see verbal behavior as one type of recordable behavior, which should be observed and analyzed like any other behavior.." (Ericsson and Simon, [1984], p. 7).

Willingness to consider verbal statements as evidence does not convey a blind assumption that they are to be taken at face value,

> "The report 'X' need not be used to infer that X is true" (Ericsson and Simon [1984], p. 9).

The argument that answers to the "strategic survey questions" outlined in section 4 above are chosen is particularly in keeping with the spirit of revealed preference theory, rooted as it is in hypothetical rather than actual choices. The essentially hypothetical nature of revealed preference is stressed in particular by Aumann [1998],

> "The essence of the axiomatic approach is that it ... relates the given, "real" situation to a whole lot of other situations, all of them hypothetical. This happens ... even in Savage's development of probabilities themselves." (Aumann [1998], p. 935.)

In a letter to Aumann, Savage himself expressed comfort not only with the need to consider hypotheticals, but further with the need to consider situations that are regarded as impossible.

> "It is quite usual in this theory to contemplate acts that are not actually available..... I can contemplate the possibility that the lady dies medically and yet is restored in good health to her husband." (Quoted in Aumann [1998], p. 935.)

In light of this history, it is particularly hard to understand how the line in revealed preference theory between permissible and impermissible evidence can rule out statements concerning intended future behavior.

5.3 Of Human Activity

Any attempt to restrict our attention to "choice data" rests on a shared "folk" understanding of what is and what is not a choice: "I know when I am choosing". Consciousness is a key ingredient in this folk conception of choice. Unfortunately consciousness itself is a very poorly understood phenomenon. Experiments in social psychology show that many apparently conscious choices are made automatically, and that ex post explanations for apparently conscious choices may be demonstrably false (Bargh and Chartrand [1999]). The difficulty in separating human behavior as between automatic and chosen was highlighted by Selten [1999] in the course of a discussion of bounded rationality.

> "Much of human behavior is automatized in the sense that it is not connected to any conscious deliberation. In the process of walking one does not decide after each step which leg to move next and by how much. Such automatized routines can be interrupted and modified by decisions but while they are executed they do not require any decision making. ... . One might want to distinguish between bounded rationality and automatized routine. However, it is difficult to do this. Conscious attention is not a good criterion. Even thinking is based on automatized routine. We may decide what to think about but not what to think."(Selten [1999], p.4.)

No clear line separates human activities into mechanistic and chosen. The goal of decision theory is after all to model "choice" itself mechanically. Hence there is no reason to limit a priori the outputs of human activity that can be included in axiomatic models. All such activities may be modeled as having been chosen from a feasible set, and the issue of what theory accounts for the observed realization is open. Any separation of human activities as between pre-programmed and chosen is best characterized on a case by case basis rather than in the over-arching methodology.

A pragmatic issue relating to modeling many non-standard psychological data is our relative lack of knowledge of, and possibly even interest in, the set of feasible alternatives to given observations. For this reason they may be modeled as tightly complementary with some external feature. One may even wish to keep the language of production theory, while recognizing that this is really just a way in which to remove too many active margins at this stage in our understanding of behavior.

5.4 Minimalism

I view the principle of revealed preference as an approach to modeling rather than a set of rigid constraints on measurement and theory. In this interpretation, it imposes no restrictions on the data that can be employed in fitting economic models, but rather guides the modeling of the given data. Specifically it suggests a very streamlined, or "minimalist" approach to modeling whatever data is ultimately selected.

In loose terms, this approach involves characterization of the exact empirical content of a model for a particular data set. As in standard choice theory, development of a single axiomatization of all measured data is the ideal, since it reduces to an absolute minimum the number of model elements that are needed. Again as with standard revealed preference, models producing the same data are to be regarded as equivalent: this is the generalized version of the "as if" principle. In fact the general theorem structure in a minimalist model is unchanged in essentials from that described above for standard revealed preference, with the only difference being the greater range of data that can be measured.

> "Measured outputs of type * have appealing joint properties ** if and only if these data are determined by calculations of the form ***. "

The need to analyze equivalence classes of representations is just as pressing in these models as in standard choice theoretic models, and the theorem structure is essentially as it was in the case of standard choice theory. In practice, results placing precise limits on what the richer theories are able to identify from the new psychological data may be even more important than in the standard case, since no little is as yet known about how to characterize these data in a concise manner.

> There exist equivalence classes of calculations of the form *** such that measured outputs of type * are identical within any such class, yet differ between classes.

What is needed to effectively differentiate the minimalist methodology from standard revealed preference are direct examples in which the data go beyond standard choices. Some examples are sketched in the next section. In each case, conventional concepts as studied in choice theory serve as the starting point for a theory that encompasses specific additional items of psychological data.

6 Two Applications

Section 6.1 outlines a minimalist approach to understanding the impact of subjective informational constraints on decision making, as detailed in Caplin and Dean [2007A]. Sections 6.2 outlines a minimalist approach to neuroscientific data hypothesized to register the degree of surprise as the environment changes for the better or for the worse. Theoretical foundations for this research are detailed in Caplin and Dean [2007B], with the empirical implementation being detailed in Caplin, Dean, Glimcher, and Rutledge [2007].

6.1 Information and the Choice Process

How best to model bounds on rationality and apparently poorly adapted decisions? One possible interpretation of such decisions is that they stem from subjective constraints on information of the form associated with the standard microeconomic model of constrained optimization. An entirely different interpretation is that they stem from application of decision heuristics of the sort identified by Kahneman and Tversky [1987]. Greater consensus is needed on how best to model the internal search process before the theory of bounded rationality can attain the standard of coherence and centrality that characterizes standard search theory, as pioneered by Stigler [1961] and McCall [1970].

Caplin and Dean [2007a] argue that a critical barrier holding back agreement on how best to model the internal search process is evidentiary. To overcome this constraint, they measure and model a new source of information on the search process. In the basic choice process data set, each possible non-empty subset of the commodity set, $A \subset X$, is imagined as being presented one and only one time to the DM. The model specifies not only the final choice that a DM makes, but also how the most preferred object changes in the course of arriving at the final decision. The idealized "choice process" data identifies for any choice set and any length of time what the DM would choose from that choice set after that length of time. For each discrete time $t \in \mathbb{N}$ (the natural numbers, or strictly positive integers), the DM is observed selecting a subset of the elements of A, with the interpretation that if the choice situation was to be terminated at this time, an identified member of this list would be received.

Definition 1 *A* ***choice process environment*** (X, C) *comprises a finite set* X *and a correspondence* $C : 2^X/\phi \times \mathbb{N} \to 2^X/\phi$, $C(A,t) \subseteq A$, *referred to as the* ***choice process****.*

A theory of the choice process identifies all successive approximations to the object chosen in a naturalistic setting with no external time constraints. Caplin and Dean characterize the extent to which this enriched data makes observable aspects of the internal search that would otherwise be kept private. This characterization matches in many respects the theory of sequential search with perfect recall. The DM is endowed with a complete, transitive preference ordering over the objects in the commodity space and explores options sequentially, retaining in hand the best object identified to date. The vision underlying the model is that the set of available alternatives is hard to identify, not their characteristics once identified as available. The process of search is then the process of uncovering precisely what is feasible in a given decision problem.

Definition 2 *A choice process environment* (X, C) *permits of an alternative-based search (****ABS****) representation* (u, S) *if there exists a utility function* $u : X \to \mathbb{R}$ *and a correspondence* $S : 2^X/\phi \times \mathbb{N} \to 2^X/\phi$*, the* ***search process****, that is expanding (i.e.* $S(A,s) \subseteq S(A,t) \subseteq A$ *for all* $A \in 2^X/\phi$ *and* $t \geq s$*) and such that,*

$$C(A,t) = \arg \max_{x \in S(A,t)} u(x).$$

This model has the property that only a DM who takes the time to search through the entire choice set can be guaranteed to satisfy WARP. Hence the theory is boundedly rational, in the sense that an "inferior" object may be selected over a "superior" one if search is incomplete. Yet it shares much in common with standard optimal search theory. The model can be specialized to cases in which search terminates when, and only when, a "reservation" utility is achieved. The resulting mode of choice has obvious links not only to the theory of optimal sequential search, but also to Simon's models of satisficing behavior. Moreoever, by allowing for a possible impact of search order on choice, the model suggests a largely standard search theoretic interpretation for such phenomena as reliance on default options and framing effects.

The theory is not intended to have universal applicability. The intricacies of realistic search processes are strongly hinted at in the work of Payne, Bettman and Johnson [1993]. In their terminology, our search model is "alternative based", with given options being compared with one another across many dimensions before new options are considered. An obvious alternative is "attribute based" search, in which the DM compares alternatives on an attribute-by-attribute basis. The goal of the axiomatization is only to characterize environments in which the modeled form of alternative-based search constitutes an adequate description of behavior. Ideally, the nature of observed violations in environments to which the model is poorly adapted will serve as inspiration for future process models.

Theories of the choice process data set introduced above represent only a first step in the process of understanding bounded rationality. One could in principle gather richer data on the evolution of preferences via an "indifference process" methodology in which one tracks over time sets of objects that are mutually indifferent. If theoretical considerations make clear that this is a potentially insightful data set, it will motivate the corresponding experimental design task in humans and potentially in other animals. While research into this question is in its very infancy, it illustrates the extent to which social scientific data and theories are endogenous, and can be simultaneously designed to bring out complementarities.

6.2 Dopamine and Choice

Dopamine, which is a neurotransmitter, plays a profound role in choice behavior. Experiments from the mid-1950's to the present show that rats and other mammals repeatedly return to locations that have been set up only to stimulate dopaminergic activity (Bernheim and Rangel [2004]). Rats make this choice over such alternatives as food, water, and opportunities to mate, even in the presence of such disincentives as painful electric shocks. Recent experiments suggest that dopamine's role in choice may be causal: rats given drugs that block dopamine receptors eventually stop feeding. Indeed an emerging theory of Parkinson's disease in humans is that, by compromising dopaminergic function, it radically alters the incentives to make even previously rewarding choices.

The best-developed theory of how dopamine impacts choice is the "dopaminergic reward prediction error" (DRPE) hypothesis. Schultz et al. [1993] demonstrated an increase in dopaminergic activity when a monkey was unexpectedly given a squirt of fruit juice, which aligned with the early theory that dopamine impacts choice directly through hedonic impact. Yet they went one step further, motivated in part by the Pavlovian learning model of Rescorla and Wagner [1972]. They found that if the monkey learned to associate a particular sound with the later receipt of the fruit juice, the dopaminergic response occurred when that sound was heard, not when the juice was received. Moreover when cued rewards failed to arrive, dopamine neurons exhibited a momentary pause in their background firing. These findings suggest that dopamine tracks the path over time of the difference between some form of anticipated and realized reward value. Intriguingly, Schultz, Dayan, and Montague [1997] noted that this "prediction error" signal is precisely what is needed in reinforcement learning algorithms designed by computer scientists to approximate standard dynamic programming value functions (Barto and Sutton [1982]). This has led many researchers to conclude that this similarity is no coincidence, and that dopamine does indeed measure a reward prediction error that is used to update an evolving value function.

Given the recent findings, it is increasingly clear that dopaminergic activity carries economically important information concerning how beliefs and preferences are formed, how they evolve, and how they play out in the act of choice. Yet communication across field boundaries remains largely out of reach. Many neuroscientific tests of the DRPE hypothesis take the perspective of "classical" or "Pavlovian" conditioning, in which choice plays no role, rendering economic interpretation difficult at best. Another fundamental problem is that the implicit assumption within the neuroscientific community is that the fMRI signal is related to the unobserved "reward" according to an affine transformation. No justification has ever been provided for limiting attention to this simple class of transformations. Neuroeconomic research needs to follow the lead of utility theory, and internalize the perils of treating an ordinal reward function as if it was cardinal.

To pursue a minimalist research agenda into dopaminergic function, Caplin and Dean [2007B] use axiomatic methods to characterize the empirical implications of this hypothesis for a data tape with combined information on choice and dopaminergic activity. This enables us to define neurological abstractions directly in terms of their empirical implications, removing the essential language barrier between neuroscientific and economic theory. We outline three economic applications of the resulting dopaminergic framework. First, we outline the potential use of dopamine measurements to provide insight into belief formation, a topic of great interest in experimental economics. The second application relates directly to learning theory, in which Erev and Roth (1998) pioneered application of the reinforcement model of animal learning. Finally we outline an application to addiction, strongly related to the work of Bernheim and Rangel [2004], and based on the observation that many addictive substances stand out as dopamine "agonists", stimulating high levels of dopaminergic activity upon consumption. While neuroscientists are currently taking the lead in exploring the interaction between dopamine and addiction, we believe that the interaction with economic reasoning is essential if the ultimate goal is to impact choice. An integrative theory such as ours is a necessary prelude to the required form of interdisciplinary research.

We develop the DRPE hypothesis for a case in which probabilities are objective and dopaminergic responses derive from realizations of specific lotteries over final prizes. The ideal data set comprises both any initial act of choice among lotteries, and the dopaminergic response when each possible prize is realized. Definition 1 lays out the fundamental building blocks of the theory.

Definition 3 *The set of prizes is a metric space Z with generic element $z \in Z$. The set of all simple lotteries over Z is denoted Λ, with generic element $p \in \Lambda$. We define $e_z \in \Lambda$ as the degenerate lottery that assigns probability 1 to prize*

$z \in Z$ and the set $\Lambda(z)$ as all lotteries with z in their support,

$$\Lambda(z) \equiv \{p \in \Lambda | p_z > 0\}.$$

*The function $\delta(z,p)$ defined on $M = \{(z,p)|z \in Z, p \in \Lambda(z)\}$ identifies the **dopamine response function (DRF)**, $\delta : M \to R$.*

The DRPE hypothesis hinges on dopaminergic responses being somehow determined by the relationship between "expected" and "experienced" rewards associated with any pair $(z,p) \in M$. The simplest technical translation of the hypothesis involves a function $r : \Lambda \to \mathbb{R}$ which defines the expected reward associated with each lottery and that simultaneously induces the reward function on prizes $z \in Z$ as $r(e_z)$. We define the DRPE hypothesis in three parts: the reward function should contain all information relevant to dopamine release; the dopaminergic response should be strictly higher for a more than for a less rewarding rewarding prize from a given lottery, and from a lottery with less rewarding that from a more rewarding lottery to a given prize; and, if expectations are met, the dopaminergic response should not depend on what was expected.

Definition 4 *Given a function $r : \Lambda \to \mathbb{R}$, define $r(\Lambda)$ as the range of the function and $r(Z)$ as the set of values taken by the function r across degenerate lotteries,*

$$r(Z) = \{r(p) \in \mathbb{R} | p = e_z, z \in Z).$$

A DRF $\delta : M \to R$ admits a dopaminergic reward prediction error (DRPE) representation if there exist functions $r : \Lambda \to \mathbb{R}$ and $E : r(Z) \times r(\Lambda) \to \mathbb{R}$ that:

1. *Represent the DRF $\delta : M \to R$, in that, given $(z,p) \in M$,*

$$\delta(z,p) = E(r(e_z), r(p)).$$

2. *Respect dopaminergic dominance, in that E is strictly increasing in its first argument and strictly decreasing in its second argument.*

3. *Satisfies **no surprise constancy** in that, given $x, y \in r(Z)$,*

$$E(x,x) = E(y,y).$$

Precise characterization theorems are presented in Caplin and Dean [2007B]. Note that the rewards in the above characterizations are defined only in relation to dopaminergic activity, just as in classical revealed preference theory utility is defined through choice. If the basic DRF permits a DRPE characterizations, there exists a definition of reward such that dopamine responds consistently to that definition, while it does not, no such reward definition can be found. Note also that the above theory will be of little interest to economists unless the reward function is somehow related to choice behavior. The strongest such relation would be if choices among lotteries could be modeled as deriving from maximization of the DRPE reward function. While this case is of obvious interest to economists, it represents an extreme form of the DRPE hypothesis. A more standard scenario involves dopamine as simply one component of a richer overall process of learning and of choice.

The axiomatic methodology may be even more important when modeling novel neuroscientific data than in its standard decision theoretic setting. It enables one to define theoretical abstractions such as rewards, predictions, and the DRPE hypothesis itself, in a direct and precise empirical language. If the data do not obey the proposed axioms, then the DRPE model is fundamentally wrong, not merely misspecified. Moreover, from a purely neuroscientific viewpoint, the minimalist agenda has value in suggesting experimental protocols directed to the central tenets of the theory, rather than to particular parametrizations. We explore this in Caplin, Dean, Glimcher, and Rutledge [2007].

7 Conclusion

Given the limitations of standard economic models and massive increases in data availability, research methods will surely be revolutionized over the coming decades. The minimalist methodology outlined above is designed to make room for new data while improving communication across sub-disciplines of social science. Successful research in line with this methodology has the potential to expand the domain of social science in ways that we are currently unable even to imagine.

8 Bibliography

References

[1] Ameriks, J., A. Caplin, and J. Leahy, 2003, "Wealth Accumulation and the Propensity to Plan", *Quarterly Journal of Economics,* vol. 118, 1007-1048.

[2] Ameriks, J., A. Caplin, S. Laufer, and S. Van Niewerburgh, 2007, "The Joy of Giving or Assisted Living: Using Strategic Surveys to separate Bequest and Precautionary Motives", mimeo., New York University.

[3] Aumann, Robert T., 1998, "Common Priors: A Reply to Gul", *Econometrica,* vol. 66, 929-938.

[4] Bargh, John, and Tania Chartrand, 1999, "The Unbearable Automaticity of Being", *American Psychologist,* LIV, 462-479.

[5] Barsky, Robert, Miles Kimball, Thomas Juster, and Matthew Shapiro, 1997, "Preference Parameters and Behavioral Heterogeneity: An Experimental Approach Using the Health and Retirement Survey", *Quarterly Journal of Economics* CXII, 537-579.

[6] Barto, Andrew, and Sutton, Richard, 1982, "Simulation of anticipatory responses in classical conditioning by a neuron-like adaptive element," *Behavioral Brain Research*, 4, 221–235.

[7] Bayer, H. and P. Glimcher, 2005, "Midbrain Dopamine Neurons Encode a Quantitative Reward Prediction Error Signal", *Neuron*, Vol. 47, 129-141.

[8] Benhabib, Jess, and Alberto Bisin, 2005, "Modeling internal commitment mechanisms and self-control: A neuroeconomics approach to consumption–saving decisions", *Games and Economic Behavior,* 52, 460-492.

[9] Benhabib, Jess, and Alberto Bisin, 2008, this volume,

[10] Bernheim, B. Douglas and Antonio Rangel, 2004, "Addiction and Cue-Triggered Decision Processes," *American Economic Review*, Vol. 94, pp. 1558-1590.

[11] Berns, G.S., McClure, S.M., Pagnoni, G., and Montague, P.R., 2001, "Predictability modulates human brain response to reward", *Journal of Neuroscience*. 21, 2793-2798.

[12] Berridge, Kent C. and Terry E. Robinson, 1998, "What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience?", *Brain Research Reviews* 28 .309-369.

[13] Camerer, Colin F., 2008, this volume,

[14] Camerer, Colin F., George Loewenstein and Drazen Prelec, 2005, "Neuroeconomics: How neuroscience can inform economics", *Journal of Economic Literature*, Vol. 34, No. 1.

[15] Caplin, A., "Fear as a Policy Instrument", 2003, in *Time and Decision*, edited by R. Baumeister, G. Loewenstein and D. Read, Russell Sage, New York.

[16] Caplin, A., and M. Dean, 2007A, "Choice-Tracking: A New Window on the Decision Process", mimeo, New York University.

[17] Caplin, A., and M. Dean, 2007B, "Dopamine and Reward Prediction Error: An Axiomatic Approach to Neuroeconomics", mimeo, New York University.

[18] Caplin, A., M. Dean, P. Glimcher, and R. Rutledge 2007, "Measuring Anticipated and Realized Rewards: A Neuroeconomic Approach", mimeo., New York University.

[19] Caplin, A., and J. Leahy, 2001. "Psychological Expected Utility Theory and Anticipatory Feelings", *Quarterly Journal of Economics*, 55-80.

[20] Caplin, A., and J. Leahy, 2004, "The Supply of Information by a Concerned Expert", *Economic Journal*, 487-505.

[21] Crawford, V., 2008, this volume, "Look-ups as the Windows of the Strategic Soul: Studying Cognition via Information Search in Game Experiments".

[22] Dean, M., 2007,

[23] Dynan, Karen, Jonathan Skinner, and Stephen Zeldes, 2002, "The Importance of Bequests and Life-Cycle Saving in Capital Accumulation: A New Answer", *American Economic Review Papers and Proceedings*, 274-278.

[24] Ericsson, K. & H. Simon, 1984, *Protocol Analysis*. Cambridge, MA, MIT Press.

[25] Ergin, H., 2003, "Costly Contemplation", Unpublished Manuscript.

[26] Friedman, M., 1953, *Essays in Positive Economics*, Chicago, Ill., University of Chicago Press.

[27] Fudenberg, D., and D. Levine, 2005, "A Dual Self Model of Impulse Control", mimeo.

[28] Gabaix, Xavier and David Laibson, 2000, "A boundedly rational decision algorithm", *American Economic Review Papers and Proceedings*, 90, p. 433-438.

[29] Gabaix, X., D. Laibson, G. Moloche, and S. Weinberg, 2006, "Information Acquisition: Experimental Analysis of a Boundedly Rational Model", forthcoming, *American Economic Review*.

[30] Glimcher, Paul and Aldo Rustichini, 2004, "Neuroeconomics: The Consilience of Brain and Decision", *Science* Vol. 306, pp. 447-452.

[31] Gul, F. and Pesendorfer, W., 2008, this volume, "The Case for Mindless Economics".

[32] Gul, F. and W. Pesendorfer, 2001, "Temptation and Self-Control", *Econometrica*, 69(6), 1403-1436.

[33] Hausman, D., 2000, "Revealed Preference, Belief, and Game Theory." *Economics and Philosophy* 16, 99-115.

[34] Hurd, Michael D., 1987, "Savings of the Elderly and Desired Bequests", *American Economic Review*, 1987, 77 (3), 298-312.

[35] Johnson, Eric J., Colin Camerer, Sankar Sen, and Talia Rymon, 2002, "Detecting Failures of Backward Induction: Monitoring Information Search in Sequential Bargaining", *Journal of Economic Theory* 104, 16-47.

[36] Kahneman, D., A. Krueger, D. Schkade, N. Schwarz, N. and A. Stone, 2004, "A survey method for characterizing daily life experience: The Day Reconstruction Method (DRM)", *Science*, 306, 1776-1780.

[37] Kahneman, Daniel and Amos Tversky, 1974, "Judgment Under Uncertainty: Heuristics and Biases", *Science*, 185, 1124-31.

[38] Kahneman, Daniel and Amos Tversky, 1979, "Prospect Theory: An Analysis of Decision under Risk", *Econometrica*, Vol. 47, No. 2., pp. 263-292.

[39] Kimball, Miles S. and Matthew D. Shapiro, 2003, "Labor Supply: Are income and Substitution Effects Both Large or Both Small?", Working Paper, University of Michigan.

[40] Kimball, Miles S., C. Sahm and Matthew D. Shapiro, 2005, "Using Survey-Based Risk Tolerance", Working Paper, University of Michigan.

[41] Kotlikoff, Laurence J. and Lawrence H. Summers, 1971, "The Role of Intergenerational Transfers in Aggregate Capital Accumulation", *Journal of Political Economy*, 89(4), pp. 706-32.

[42] Kreps, D., 1979, "A Preference for Flexibility", *Econometrica*, 47: 565—576.

[43] Kreps, D. and E. Porteus, 1978 "Temporal Resolution of Uncertainty and Dynamic Choice Theory", *Econometrica*, 46, 185-200.

[44] Laibson, David, 1997, "Golden Eggs and Hyperbolic Discounting", *Quarterly Journal of Economics* ,CXII, 443-477.

[45] Ljungqvist, Lars, 1993, "A Unified Approach to Measures of Privacy in Randomized Response Models: A Utilitarian Perspective", *Journal of the American Statistical Association*, Vol. 88, No. 421, pp. 97–103.

[46] Loewenstein, G., 1987, "Anticipation and the Valuation of Delayed Consumption", *Economic Journal*, vol.97, 666-684.

[47] Lowenstein, G., and T. O'Donoghue, 2004, "Affective and Deliberative Processes in Economic Behavior", mimeo.

[48] Lusardi, Annamaria, 2000, "Explaining why so many households do not save", mimeo, University of Chicago Harris School of Public Policy, 2000.

[49] Manzini, P. and M. Mariotti., 2004 , "Rationalizing Boundedly Rational Choice", IZA Discussion Paper No 1.

[50] McClure, Samuel, David I. Laibson, George Loewenstein, and Jonathan D. Cohen, 2004, "Separate Neural Systems Value Immediate and Delayed Monetary Reward", Neural Networks 15, Science 306, 603-616.

[51] von Neumann, J. and O. Morgenstern, 1947, *Theory of Games and Economic Behavior*, Princeton University Press, Princeton, NJ.

[52] O'Donoghue, Ted, and Matthew Rabin, 1999, "Doing it Now or Later", *American Economic Review*, vol. XCIX, 103-124.

[53] O'Doherty, J.P., P. Dayan, K. Friston, H. Critchley, and R.J. Dolan, 2003, "Temporal difference models and reward-related learning in the human brain", *Neuron* 38, 329-337.

[54] Payne, John W., James R. Bettman, and Eric J. Johnson, 1988, "Adaptive Strategy Selection in Decision Making", J*ournal of Experimental Psychology: Learning, Memory, and Cognition*, Vol. 14, No. 3, 534-552.

[55] Payne, John W., James R. Bettman, and Eric J. Johnson, 1993, *The Adaptive Decision Maker*, Cambridge: Cambridge University Press.

[56] Payne, J. W., J. R. Bettman, and M. F. Luce, 1996, "When Time is Money: Decision Behavior under Opportunity-Cost Time Pressure", *Organizational Behavior and Human Decision Process*, 66:2, pp 131-152.

[57] Peleg M.and M. E.Yaari, 1973, "On the Existence of a Consistent Course of Action when Tastes are Changing", *Review of Economic Studies* 40: 391-401.

[58] Radner, R. and M. Rothschild, 1975, "On the Allocation of Effort", *Journal of Economic Theory*, 10 (3), 358-76.

[59] Rescorla, R. A., and Wagner, A. R., 1972, "A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement," *Classical Conditioning II*, A. H. Black and W. F. Prokasy, Eds., pp. 64-99. Appleton-Century-Crofts.

[60] Samuelson, P., 1938, "A note on the pure theory of consumer's behavior", *Economica*, 5, 61 -71.

[61] Rubinstein, A., 1998, *Modeling Bounded Rationality*, Zeuthen Lecture Book Series, MIT Press, Cambridge, Mass.

[62] Samuelson, P., 1938, "A note on the pure theory of consumer's behavior", *Economica*, 5, 61 -71.

[63] Savage, L.J., 1954, *The Foundations of Statistics*, New York, Wiley and Sons.

[64] Schotter, A., 2008, this volume.

[65] Schultz, Wolfram, Apicella, Paul, and Tomas Ljungberg, 1993, "Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task," *Journal of Neuroscience*, vol. 13, 900–913.

[66] Schultz, Wolfram, Dayan, Peter, and Montague, P. Read, 1997, "A Neural Substrate of Prediction and Reward", *Science*, 275 :1593-1599.

[67] Selten, R., 1999, "What is Bounded Rationality?", SFB Discussion Paper B-454.

[68] Simon, Herbert, 1955, "A behavioral model of rational choice", *Quarterly Journal of Economics*, 69 (1), pp 99-118.

[69] Strotz, Robert, 1956, "Myopia and Inconsistency in Dynamic Utility Maximization", *Review of Economic Studies*, 23, 165-180.

[70] Tversky, A., 1972, "Elimination By Aspects: A Theory of Choice", *Psychological Review* 79:4, pp 281-299.

[71] Tyson, C., 2005, "Contraction Consistency, Cognitive Constraints and the Satisficing Criterion", Working Paper, Oxford University.

[72] Wilson, Andrea, 2002, "Bounded Memory and Biases in Information Processing", Working Paper, Princeton University.