Choice and Process: Theory Ahead of Measurement*

Jess Benhabib
NYU

Alberto Bisin
NYU

March 2007

Abstract

Are models and data on *choice processes* useful as a complement to *revealed preferences* in decision theory? We answer this methodological question in the affirmative. We also argue, however, that progress in neuroeconomics is likely to require relying more clearly on *structural empirical methods*. We illustrate our arguments by means of examples from *inter-temporal decision theory*.

*Many of the ideas exposed in the paper evolved in the course of discussions with Aldo Rustichini, who appeared as co-author in previous versions of the paper. Thanks to Andrew Caplin, Chris Flinn, Ariel Rubinstein and Yuval Salant for useful discussions.

1 A new method

The traditional method of decision theory, founded on *revealed preferences*, restricts its focus on predicting and explaining *choice* and is agnostic about the *process* underlying choice itself. Recent research in economics (typically under the heading of *neuroeconomics* or of *behavioral economics*) aims instead at developing joint implications on *choice* as well as on *processes*.

In this note we shall argue that models of decisions designed to produce joint implications on both choice and process constitute a new exciting area of research for decision theory.

We also argue, however, that the literature would gain from the adoption of structural empirical methods to guide the analysis and test the models. The conceptual hiatus between axiomatic decision theory in economics and models of decision processes in neuroscience in fact necessitate a structural approach to better lay out and clarify the often implicit identifying assumptions adopted in either discipline.

We finally attempt to illustrate, by means of examples in the context of intertemporal decision theory, how structural a structural analysis of choice and process can add explanatory and predictive power to decision theory.

1.1 Why a new method?

The traditional method adopted in economics has its foundations on *revealed preferences*. It is the product of the ordinal revolution in I. Fisher [1918] and V. Pareto [1906].[1] This method has been most recently discussed by Gul-Pesendorfer [2006 - in this volume] in a lucid and provocative manner as alternative to behavioral economics and neuroeconomics.

In the pure example of this method a decision maker is presented with choices among acts. An act describes the consequence that the decision maker will obtain for every realization of a state of nature. A theory of choice is a complete list of what the decision maker will choose for any possible option that he is facing. Since listing the choices obscures their systematic nature, the list is summarized in a set of axioms. The typical decision theory is then a representation theorem, that is, the statement that the decision maker chooses according to the axioms if and only if he chooses *as if* he were maximizing a certain value function.

[1] Intellectual interest in the study of decision processes can however be traced to the classic period, from W.S. Jevons' *Brief Account*, [1866], or of chapters II and III of his *Theory*, [1871] to J. Bentham's *Principle of Morals and Legislation*, [1780] or A. Smith's *Theory of Moral Sentiments*, [1759]

In this view, the representation of the preferences has purely the nature of a conceptual construct and has no independent informational value in addition to what is already contained in the axioms. For the method of revealed preferences the only admissible evidence to test a decision theory over a set of options is the agent's real choice from subsets of that set.

It is certainly good methodological practice not to test a theory based on data the theory itself was not design to fit. As we do not reject representative agent macroeconomic models based on the observation of agents' heterogeneity, we should not reject the standard axiomatic decision theory if we fail to identify a $\max U$ process in the brain.

Nonetheless, we will argue that a clear-cut methodology making use of explicit models of process as well as psychological or neurophysiological data can provide the decision theorist with useful tools to explain choice.

More specifically we claim that, even if we agree that our objective as economist is to explain choice per se, not process, nonetheless the study of choice processes has in principle additional explanatory power for decision theory.

If however we find ourselves to date hard pressed to name existing models and data on decision processes which have fundamentally contributed to our understanding of choice, this is, in our opinion, because most empirical work in neuroeconomics is not tightly guided by models making joint predictions on choice and process. We therefore observe that in practice progress in producing explanatory power to understand choice is likely to require *structural analysis*, more *theory ahead of measurement.*[2]

Before proceeding with the arguments, we wish to clarify two terms we shall repeatedly use in this note, *explain* and *structural analysis*. By *explain* we mean, with Gul-Pesendorfer [2006 - this volume, p. 6], "identify choice parameters [...] and relate these parameters to future behavior and equilibrium variables." In other words, an explanation is is such when it has predictive power, outside of sample. By structural analysis we mean the specification of a formal model which maps a set of parameters, assumed stable, to a set of observable (measurable) variables, and the estimation of the parameters by statistical inference methods. The structure of the model explcitly represents the a priori assumptions underlying the analysis. In the context of the study of choice and process, a structural model has implications for both choice as well a process variables.

[2]The classic formulation of the methodology of structural empirical analysis is in Koopmans [1947] and Marschak [1952]. A recent discussion is contained in Keane [2006].

2 When is theory and measurement of process useful?

The aim of this section is to examine the conditions under which economic analysis may be improved by theorizing about and observing psychological states and decision process.

To set a frame of reference for the ensuing arguments, consider the typical decision problem in the spirit of revealed preference. A decision maker chooses from subsets A_1, A_2 of an abstract choice set $\mathbf{A}$. His/her behavior is formally represented by a choice function $C(.)$ whose domain is the set of subsets of $\mathbf{A}$ and whose range is $\mathbf{A}$. Standard decision theory would study the implications, in terms of choice, of a series of axioms on choice itself taken to define rationality. The fundamental axiom of rational choice is *consistency* (a.k.a. independence of irrelevant alternatives):

$$\text{if } A_1 \subseteq A_2 \subseteq \mathbf{A} \text{ and } C(A_2) \in A_1, \text{ then } C(A_1) = C(A_2)$$

The choice function $C(.)$ satisfies consistency if and only if it can be represented by a preference ordering $\geq$ such that $C(A)$ contains the maximal element of $\geq$ in A.

Notice that the explanatory power of standard decision theory results from an axiom like *consistency* underlying (having predictive power regarding) choice in many different environments, e.g., from intertemporal choice to choice under uncertainty.

2.1 Psychological states and preferences

Let S denote a set of *psychological states*, a list of emotions, for instance.[3] Suppose that in fact the decision maker's choices are affected by emotions, and his/her behavior consists in fact of a list of choice functions $C_s(.)$, for any state s in S. Then a decision theorist, observing choices while varying $A \in \mathbf{A}$ but oblivious of the fact that states are also varying, might conclude the observed choice function does not satisfy the consistency requirement necessary for representation, even though consistency might instead be satisfied for each $C_s(.)$. In this context, psychological states might count as primitives of the decision problem and observation of the states might represent a necessary constitutive component of a coherent revealed preferences exercise: preferences

[3]See Kahneman-Krueger [2006] for a survey of the theoretical constructs and the measurement issues behind the notion of psychological states.

are *revealed* at the varying of the choice set A (typically through variations in prices and income) as well as at the varying of the psychological state s.

But suppose instead that, even though the decision maker choice depends on psychological states s which are not observed in the revealed preference exercise, a representation of $C(.)$ is obtained. In a certain sense this is the show-case of the *as if* argument: no matter that choice might be related to the decision maker's psychological states, he or she chooses as if he or she were maximizing a well defined preference ordering. We claim that even in this case a structural analysis of psychological states states $s \in S$ is useful to improve the explanatory power of decision theory.

To illustrate this argument it is useful to introduce a simple formal example. Consider the typical decision problem induced by the choice of a consumption allocation $x \in \Re^n_+$ in a budget set $px \leq I$, where $p \in \Re^n_+$ is the price vector and $I \in \Re_+$ is income. The decision theorist observes a sequence of choices associated to different prices p and income I. Assuming linearity (for simplicity; our argument does not depend on this assumption) and allowing for observation errors ϵ, the decision theorist observes

$$x = \beta \begin{bmatrix} p \\ I \end{bmatrix} + \epsilon \tag{2.1}$$

where β is a vector of unknown parameters and ϵ a random variable.

The aim of the decision theorist is to explain choice (for clarity we'll say explain/predict choice), that is x. This requires a regression analysis to estimate β.

Psychological states are summarized by a variable $s \in S$ which depends on the data of the decision problem, $\begin{bmatrix} p \\ I \end{bmatrix}$. For instance, the consumer can feel excited if his income is large (increases) or if a series of goods are sold at a bargain (a low price).[4] In summary, states satisfy

$$s = \gamma \begin{bmatrix} p \\ I \end{bmatrix} \tag{2.2}$$

(adding noise does not change the argument).

Suppose that the psycho-neural process that is summarized by the variable s interacts directly (and interestingly) with choice. Suppose for instance that

$$\beta = \alpha s, \text{ that is } x = \alpha\gamma \begin{bmatrix} p \\ I \end{bmatrix} + \epsilon \tag{2.3}$$

[4] Evidence through skin conductance measurement

Notice that in this formulation the decision maker's choice function has a preference representation even if the decision theorist is oblivious to states s. This is manifested in the fact that the decision theorist can certainly estimate $\beta = \alpha\gamma$ without independently estimating the specific values of α and γ *per se*.

But, could the decision theorist improve his or her explanation/prediction of x by making use of his observation of s? The answer is certainly affirmative, in a statistical sense, since an independent estimate of γ can be in general efficiently used to improve the estimate of $\beta = \alpha\gamma$ and therefore to deepen our understanding of the relationship between the determinants of choice, $\begin{bmatrix} p \\ I \end{bmatrix}$, and choice x itself. While measurement of psychological states is not an easy task, proxies like heart rate and Galvanic skin response have been successfully employed in neuroscience; see e.g., Damasio [1999].

The metaphor we adopted, decision theory as a regression, is certainly a stretch but it demonstrates effectively, in our opinion, that data on process is certainly useful as a complement to data on choice as soon as we require decision theory to provide explanation/prediction for behavior out of sample. Outside of the metaphor, structural models of the interaction between choices and psychological states could have in principle implications for diverse areas of decision theory, from intertemporal choice to choice under risk and uncertainty, and therefore contribute to a unifying explanation of several experimental puzzles.

The regression metaphor allows us also to stress that it is the structural model of the interaction of choice and process which really potentially adds to our abstract understanding of choice itself. Consider a formulation of the decision problem in which s is a random variable correlated with choice x, that is with ϵ. Then observing s would certainly reduce the noise in the prediction of x, like the introduction of any new explanatory variable in a regression will do. But it would not contribute to our analysis of choice more than the observation that "hot weather increases the demand for ice-cream". We are not after adding s to as many regressions as possible, but rather we are after the structural explanation of choices. It is structural models of the determination of psychological states and emotions, as well as of their interaction with choice, which can in principle deepen our understanding of choice.

2.2 Procedural rationality

While decision theory stands traditionally on choice axioms, an interesting literature has produced axiomatic analysis of choice processes, inspired by the work of H. Simon in the 50's (collected in Simon [1982]; see Rubinstein [1998] for a fascinating introduction to this literature).

As an illustration (we follow Rubinstein [1998] here), consider again the the standard decision problem. Rather than directly postulating properties of choice (through axioms) and then deriving an *as if* representation, a theory of procedural rationality describes a choice process (a *procedure*) and then derives its implied restrictions on choice. A typical procedure is e.g., *satisficing*:

let O denote an ordering on $\mathbf{A}$, let T be a subset of $\mathbf{A}$, and let $A \in \mathbf{A}$ denote a choice set; then $C(A)$ contains the first element of A, according to the order O, which belongs to T (and the last element of A if A and T are disjoint.

The primitives of the procedure include T and O. The properties of the choice function $C(.)$ obviously depend on a crucial manner on T and O. For instance, allowing T and O to depend on the choice set A implies that $C(.)$ might not satisfy the *consistency* axiom and a standard representation in terms of preference maximization might not exist. The same if T and O depend on psychological states $s \in S$. Behavior like e.g., framing (the dependence of choice in experiments from the way choice problems are posed) might be rationalized by a satisficing procedure, by letting the order O depend on unobservable psychological states induced by the way choice problems are posed.[5]

A decision theory formulated in terms of a procedure like *satisficing* has most explanatory power inasmuch as it includes a model of the determination of the set T and of the ordering O, for all $A \subseteq \mathbf{A}$ and psychological states $s \in S$. *Identifying choice parameters* in this context might require then data on process, like those collected through the eye-tracking procedures for saccade tasks commonly adopted in vision and attention studies in neuroscience (see for instance Deubel-Schneider, [1996]) or like the mouse tracking procedure used e.g., by Camerer-Johnson-Rymon-Sen [1993].

To better illustrate the kind of structural models we claim are useful in this context, consider the following abstract class of procedural environments:

let $A \in \mathbf{A}$ denote a choice set; two distinct procedures P_1 and P_2 map A into respective elements of A; a third procedure selects which of P_1 and P_2 controls choice, that is, if $C(A) = P_1(A)$ or $C(A) = P_2(A)$.

A procedure of this sort can abstractly capture a large class of cognitive decision processes which involve competing procedures and a selection mechanism. Often the competing procedures are modeled to represent the classic

[5]Relatedly, see Rubinstein-Salant [2006] for the axiomatic treatment of choices from lists.

automatic-controlled (or visceral/cognitive) dichotomy and the selection mechanism is represented by some form of attention control.

Procedures of this sort could represent well *choice mistakes*, "systematic phenomena which disappear once decision makers learn of their existence" (Rubinstein [1998], p. 22). A typical example is tourists looking left when crossing the street in the U.K., but many other examples flood the experimental psychology literature. *Choice mistakes* provide in fact a useful clarifying example of the *new method* we delineate in this paper. Gul-Pesendorfer [2006 - this volume, p. 27] suggest that according to the standard method of revealed preferences *mistakes* of this sort can be rationalized simply by means of "subjective constraints on" their "feasible strategies." This is because such *mistakes* would be "relevant only if they could be identified through economic data." We claim instead that a structural model of the choice process leading to such *mistakes* would constitute a better methodological practice. It would avoid adding a "subjective constraint" everytime needed and it could provide explanatory power to understand choice in several different interesting contexts in which attention control might be relevant (see also the discussion in the next section regarding intertemporal choice).

A model of the procedures P_1 and P_2 as well as of the selection mechanism, the primitives of the procedure, is necessary to provide a unifying explanations for choice mistakes. But what are the component such a model? What is *automatic*, what is *cognitive*? And, what is *attention*?

While we do not know of an axiomatic analysis of this class of procedures, models of this kind of behavior have been developed and their structural empirical implications studied in neuroscience. The *language* of modeling in neuroscience is different than in economics: it involves simulating the dynamic activation properties of a neuronal network rather than deriving the logical implications of axiomatic relations. Nonetheless these models provide natural constructs for the structural analysis of choices and of choice processes.

As an illustration consider Cohen-Dunbar-McClelland [1990]'s cognitive control model of automatic and controlled processes in the Stroop task, after the experiments by Stroop in the 30's.[6] The Stroop task consists in naming the ink color of either a conflicting word or a non-conflicting word (e.g., respectively, saying 'red' to the word 'green' written in red ink; and saying 'red' to the word 'red' written in red ink). Cohen et al. [1990]'s model is based on *parallel distributed processing* (PDP).[7] PDP models consist of a collection of

[6] The distinction between automatic and controlled processing is common in neuroscience, and is articulated, e.g., in Schneider-Shiffrin [1977] and in Norman-Shallice [1980].

[7] See Rumelhard-McClelland [1989] for an extensive presentation and discussion of PDP models in neuroscience.

different processing units distributed over a network whose architecture represents processing modules and pathways. Each processing unit is characterized by a pattern of activation which dynamically propagates over the network, from the exogenous inputs to the output. For instance, letting the index i run over all the processing units which input into element j, the activation of j at time t, $a_j(t)$ is written as follows:

$$a_j(t) = \tau \sum_i a_i(t) w_{ij} + (1 - \tau) \sum_i a_i(t-1) w_{ij}$$

where w_{ij} represent the weight (or strength) of the connection between unit i and j.

Cohen et al. [1990] model word-reading and color-naming as competitive processing pathways in the network, which are simultaneously activated by the word image. Furthermore, a different pathway is activated by the explicit goal of the cognitive task, say color-naming, which cognitively controls the output of the task by differentially activating the appropriate processing pathway and inhibiting the other one.

This model is crucially supplemented by a specific learning model which determines endogenously the weights of the connections w_{ij} in the network. An *automatic* pathway is defined as one which has been repeatedly active in the past, so that the learning process has generated high connecting weights, and the output is quickly generated from the input. In Cohen et al. [1990] word-reading, which is a much common task in the subjects' practice outside the lab, is modeled as an automatic process which produces a rapid response. The controlled processing aspect of the task can however override the stronger word-reading process by inhibiting the automatic reading association.[8]

The model implies a pattern of reaction times to conflicting and non-conflicting words which is consistent with the pattern observed in experiments with Stroop: *i)* reaction times for reading tasks are unaffected by the ink color, *ii)* reaction times for conflicting words are higher than for non-conflicting words, *iii)* reaction times are higher, for either conflicting and non-conflicting words, than the reaction times of simple reading tasks.[9]

While these models has been developed to understand behavior in cognitive rather than decision theoretic tasks, we suggest that this class of models could very usefully be adapted to study choice processes. We believe that formal models of automatic and controlled processes can provide a unifying explanation of choice mistake as well as of several other puzzling choice phenomena

[8]See Miller-Cohen [2001] for a general introduction to attention control models and their structural empirical analysis.

[9]See also Braver-Cohen-Servan Schreiber [1995].

documented in the experimental lab, in several decision theoretic environments ranging from intertemporal choice (see the next section) to choice under uncertainty (see Loewenstein-O'Donoghue [2005]).

3 Intertemporal choice: The method in practice

In the previous section we have argued that models and data on psychological states and choice processes are *in principle* useful to decision theory. In this section we survey as an illustration the literature concerning intertemporal choice. We shall identify in the the lack of structural analysis a bottleneck of neuroeconomics in this context.

The standard economic approach to the study of intertemporal decisions involves agents maximizing their present *exponentially* discounted utility. Exponential discounting postulates that the present discounted value of a reward u received with a t-period delay is $\delta^t u$ for some $\delta < 1$.[10] Recently, however behavioral economists have criticized this approach on the basis of a vast amount of behavioral regularities (called *anomalies*) documented in experimental psychology which indicate that agents may have a preference for present consumption, a *present bias*, that cannot be rationalized with exponential discounting.[11] The most important of such regularities is called *reversal of preferences*. It occurs when

> a subject prefers $\$x$ now rather than $\$x + \Delta$ in a day, but he/she prefers $\$x + \Delta$ in a year plus a day rather than $\$x$ now.

Various alternative decision theories have been developed which rationalize such data, and reversal of preferences in particular. For instance Laibson [1996] and O'Donoghue-Rabin [1999] favor a quasi-hyperbolic specification of discounting, which posits that the present discounted value of a reward u received with a t-period delay is $\beta\delta^t u$ for some $\beta, \delta < 1$. Others, e.g., Ainsle (1992) favor an hyperbolic specification, implying a discounted value of the form $\frac{1}{1+\delta t}u$. Finally, Gul-Pesendorfer [2001] develop an axiomatic theory of temptation and self-control which rationalizes present bias by extending the domain of choice to sets of actions.

[10]See Koopmans [1960] and Fishburn-Ribinstein [1982] for the classic axiomatic treatment of intertemporal choice.

[11]See, e.g., Ainsle [1992] and Frederick-Loewenstein-O'Donoghue [2002] for comprehensive surveys.

All these are standard decision theoretic models in that they induce restrictions/implications only on choice, and not on processes, and they are formulated as a preference representation.[12] They can therefore be tested in the experimental lab with choice data.[13] As we argued in the previous sections, however, this is not enough to conclude that data on process are not useful: data on process can in principle help explain explain choice. We next survey selectively some attempts at doing just in practice in the context of intertemporal choice.

For instance, Wilson-Daly [2004] document higher discounting for men after having observed photographs of women which they reported as attractive. This finding can be appealingly interpreted as a manifestation of dependence of discounting on the psychological states, the photographs are "inducing a 'mating opportunity'" in the words of the authors (p. S177). No model of such dependence is however developed that can be tested with choice or with process data. Consequently, the paper is silent on the possible relationship between psychological states and the discounting anomalies which we seek an explanation for.[14]

Much more important and central to our understanding of the interaction between choice and process are two recent studies at the forefront of neuroeconomics, McClure-Laibson-Loewenstein-Cohen [2004] and Glimcher-Kable-Louie [2006]. who develop and study brain imaging data to explicitly distinguish between some of the different preference representations .

Both McClure et al. [2004] and Glimcher et al. [2006] produce and study brain imaging data during intertemporal choice tasks. They then interpret this data as evidence in favor of different preference representations of discounting which rationalize present bias anomalies. In particular, McClure et al. [2004] claim evidence for the quasi-hyperbolic representation of discounting. They postulate that such representation results from the influence of two distinct neural processes, one which is differentially activated in the presence of immediate rewards and one which is commonly activated when the decision maker engages in intertemporal choices. They then measure brain activation of several subjects in an intertemporal choice task by functional magnetic imaging (fMRI) techniques and identify econometrically areas of differential activation when the choice task involves an immediate reward. McClure et al. [2004] then categorize such areas of the brain as *β-areas*, interpreting them as representing

[12]See Masatlioglu-Ok [2004] for general axiomatic representations of these discounting preferences.

[13]This is done e.g., by Benhabib-Bisin-Schotter [2005].

[14]See Smith-Dickhaut [2004] for empirical evidence on the effect of emotions on bidding in auctions.

present bias, and more generally interpret the existence of such areas as evidence in favor of the existence of two-neural processes involved in intertemporal choice.

McClure et al. [2004] provide no structural analysis of choice and process underlying the different neural processes. No choice data is reported. No formal model of the neural processes which are postulated to underlie choice is proposed and hence no formal implications are derived regarding the pattern of activation of different areas of the brain. Furthermore, no clear a priori theoretical presumption links quasi-hyperbolic discounting with two-distinct neural processes.[15] As a consequence, the empirical results of McClure et al. [2004] are prone to different interpretations and can hardly identify the properties of the underlying choice process that they observe in their fMRI study, including the existence of two-neural processes involved in intertemporal choice. In this sense, Glimcher et al. [2006] argue that the activation patterns found in McClure et al. [2004] are in fact consistent with the hypothesis that brain activation simply correlates with an hyperbolic representation of the decision maker's discounting preferences; the implicit choice process in this case simply being represented by discounted utility maximization (without recourse to two neural processes).

Glimcher et al. [2006] proceed then to estimate discounting preferences with choice data, finding that an hyperbolic representation is not statistically rejected. They measure brain activation by fMRI and find a clear correlation pattern of activation measurements in areas of the brain typically associated with option valuation with the discounting representation estimated with choice data. This is interpreted as evidence for an explicit preference maximization procedure underlying intertemporal choice. Lacking a structural model of choice and process, however, the data can once again hardly identify the (single or multiple) neural processes involved in intertemporal choice.

In summary, the neuroeconomics literature has made great progress in studying brain imaging data of decision makers engaged in intertemporal choice tasks. Structural models to guide the empirical analysis are still lacking, however, which can provide the decision theorist with the explanatory power to distinguish between different choice representations of discounting and hence, ultimately, to explain intertemporal choice anomalies.

In the quest to explain intertemporal choice anomalies, a few models of choice and process have been developed and studied. Unfortunately, until now these models have been studied empirically only with choice data from

[15]Intuitively, however, the power of the test relies on their finding that β-*areas* are mostly located in the limbic system, that is, an area of the brain typically associated with impulsive choice rather than cognitive processing.

the experimental lab and not yet with data on process itself. We illustrate this point by explicitly discussing two such models, Rubinstein [2003] and Benhabib-Bisin [2004].

Rubinstein [2003] considers a simple procedural model of the binary choice over rewards x at different delays t. The procedure studied is a *similarity* procedure:[16]

> facing the binary choice over (x, t) and (x', t'), the decision maker first looks for dominance, e.g., choosing (x, t) if $x \geq x'$ and $t \geq t'$ (with at least one strict inequality); lacking dominance the decision maker looks for *similarity*, e.g., choosing (x, t) if x *is similar to* x' and $t > t'$; if the two previous procedure are inconclusive choice is made using yet another procedure.

Several choice experiments are designed which can distinguish hyperbolic discounting from the *similarity* procedure (once complemented with a specific notion of the relation *is similar to)* and choice data is provided that supports the *similarity* procedure.

Note also that, in accordance with to the methodological claim we are exposing, i.e., that models of process have the potential of providing unifying explanation of various phenomena in experimental choice behavior, a related *similarity procedure* has been studied by Rubinstein [1988] in the context of choice over risky lotteries.

While rich implications of the *similarity* procedure can certainly be derived on process data, e.g., on reaction times, we know of no research along these lines.[17]

Benhabib-Bisin [2004] provide instead an intertemporal choice model in which choice is the result of the interaction of automatic and controlled processing.[18] When specialized to the binary choice over rewards at different delays, the typical choice experiment which gives rise to the *anomalies* in experiments, Benhabib-Bisin [2004]'s model induces a present discounting representation of a reward u at delay t of the form $\delta^t u - b$, where b represents the psychological cost induced by the need to exercise self-control, interpreted as a psychological restraint from the impulse of choosing the immediate reward. The cost of delay b is a fixed cost, that is, is independent both of the size of the reward and of

[16]See Tversky [1977] for the early introduction of similarity relations in decision making.

[17]But see Rubinstein [2006], which relates cognition and reaction times in several strategic environments.

[18]While Benhabib-Bisin [2004] model the choice process directly, without providing its axiomatic foundation, see Nehring [2006].

the amount of the delay.[19] Benhabib-Bisin-Schotter [2005] estimate discount preferences with experimental choice data and find statistical evidence which in fact favors this representation over both quasi-hyperbolic and hyperbolic discounting.

A different test of Benhabib-Bisin [2004]'s model can be performed, once again, with experimental choice data, by considering a simple environment in which an agent has to decide how much to consume today out of available income z. When their analysis is specialized to this simple environment they obtain the following representation:

$$\max \left\{ \max_{x \leq z} u(x) - b, u\left(\arg\max_{x \leq z} v(x)\right)\right\} \tag{3.4}$$

where $u(x)$ and $v(x)$ are smooth, concave real functions representing, respectively, the cognitive and automatic components of preferences. The cognitive component of preferences controls choice if and only if its valuation minus the self-control cost b is larger than the automatic valuation $u\left(\arg\max_{x \leq z} v(x)\right)$. Under some regularity assumption,[20] this representation has the identifying behavioral implication:

> the choice $x(z)$ is not increasing in z, but rather has a decreasing jump: small temptations are not controlled, while large ones are.

In work in progress, Benhabib-Bisin-Kariv are exploring this implication in the experimental lab. About a half of the 20 subjects for which data has been collected display the behavior predicted by the model.

While Benhabib-Bisin [2004]'s model of intertemporal choice and process has been tested with choice data, as we have reported above, the model also has potentially several clear-cut implications about process which can be derived by formulating a parallel distributing processing model along the lines of the Stroop model in Cohen-Dunbar-McClelland [1990] and which can be tested e.g., by recording reaction times data during an intertemporal choice task. This has not yet been done.

---

[19] Note that the quasi-hyperbolic representation can be written $\delta^t u - (1-\beta)\delta^t u$ and hence implies instead a variable cost associated to non-immediate rewards, that is, a cost proportional to the value of the reward u.

[20] In particular, assuming

$$x^* = \arg\max_{x \in X} u(x) < \arg\max_{x \in X} v(x)$$

allows the interpretation, essentially without loss of generality, of "temptations" as "preferences for a larger x;" so that z measures, parametrically, as the "size" of the temptation.

The theory of intertemporal choice is a fascinating laboratory: it has *i)* choice *anomalies* to explain, *ii)* sophisticated models, from axiomatic to algorithmic, to put to data, and *iii)* a wealth of data, from the experimental choice data to brain imaging data to test its models. All the ingredients for the application of the *new method* we have discussed are ready to be mixed.

While intertemporal decision theory appears representative of other areas of neuroeconomics in terms of the methods used, structural models of choice and process are being developed at the frontier, for instance in the study of reward prediction errors in learning (Caplin-Dean [2007]) and in the study of random utility (Maccheroni-Marinacci-Rustichini [2006]).

4 Conclusions

In this note we have argued that, in our opinion, no logical reasons exist to exclude models and data on choice processes from decision theory. On the contrary, the structural analysis of models and data on process represents in our opinion the fascinating frontier of decision theory.

While standard decision theory has been very successful in rationalizing a rich set of behavioral data from lab experiments by a combination of weakening of the axioms and enlarging of the choice set,[21] it seems to us that this success has come at the expense of explanatory power, that is, of a unified theory of decision making. A new method exploiting the study of choice processes as well as choices, in our opinion, contains the promise of unifying the explanations of many behavioral puzzles observed in the experiments.

With respect specifically to the theory of intertemporal choice, we have noted that the structural analysis of choice and process we claim could advance our understanding of choice seems to be yet missing: the most advanced brain imaging techniques are adopted without the guide of theoretical analysis, resembling what economists call *fishing for factors*, while the few models which in fact focus on the interaction of choice and process are tested only with choice data, wasting much of their explanatory/predictive power.

We conclude this note by claiming an added rhetorical advantage for a methodology for decision theory which goes beyond *as if* representations by directly formulating models of choice processes which can be tested with data on process. The representation of preferences in the standard theory of revealed preferences is often, by its very nature, an informal description of a process.

[21] Notable examples include Kreps-Porteus [1978] on early resolution of uncertainty, Gilboa-Schmeidler [1989] on uncertainty aversion, Gul-Pesendorfer [2001] on present bias, and many others.

The elegance of the representation, its accordance with introspective beliefs about decision processes which *inspire* it,[22] in summary its *intuitive appeal* is typically crucial for a decision theory to be accepted. The necessarily informal (but important) role that concepts like *intuitive appeal* or *inspiration* end up performing in the method of *revealed preferences* is in our view a substantial limitation of the method itself. Wouldn't it be much better to lay all the cards down for inspection?

[22]In this sense, Gul-Pesendorfer [2006 - this volume] accept a reference to process in decision theory as *inspiration*; e.g., p. 41 and 44.

References

Ainsle, G. (1992), *Picoeconomics*, Cambridge University Press.

Benhabib, J. and A. Bisin (2004), 'Modelling Internal Commitment Mechanisms and Self-Control: A Neuroeconomics Approach to Consumption-Saving Decision,' *Games and Economic Behavior*, 52(2), 460-492, special issue on *Neuroeconomics*.

Benhabib, J., A. Bisin, and A. Schotter (2005), 'Hyperbolic Discounting: An Experimental Analysis,' mimeo, NYU.

Bentham, J. (1780), 'An Introduction to the Principles of Morals and Legislation', London.

Braver, T., J.D. Cohen, and D. Servan Schreiber (1995), 'A Computational Model of Prefrontal Cortex Function,' in *Advances in Neural Information Processing Systems*, ed., D.S. Touretzky, G. Tesauro, and T.K. Leen, Cambridge, MA: MIT Press.

Camerer, C., E. Johnson, T. Rymon, and S. Sen (1993), 'Cognition and Framing in Sequential Bargaining for Gain and Losses,' in *Frontiers of Game Theory*, ed., K. Binmore, A. Kirman, and P. Tani, 27-48, Boston: MIT Press.

Caplin, A. and M. Dean (2007), 'Dopamine and Reward Prediction Error: An Axiomatic Approach to Neuroeconomics,' *American Economic Review, Paper and Proceedings*, forthcoming.

Damasio, A.R. (1999), *The Feeling of What Happens,* Orlando, FL: Harcourt Brace.

Cohen, J.D., K. Dunbar, and J.L. McClelland (1990), 'On the Control of Automatic Processes: A Parallel Distributed Processing Model of the Stroop Effect,' *Philosophical Transactions of the Royal Society (London) B*, 351, 1515-27.

Deubel, H. and Schneider, W.X. (1996) 'Saccade Target Selection and Object Recognition: Evidence for a Common Attentional Mechanism,' *Vision Research*, 36, 1827-1837.

Fishburn, P. and A. Rubinstein (1982), 'Time Preference,' *International Economic Review*, 23, 677-94.

Fisher, I., (1892), *Mathematical Investigations in the Theory of Value and Prices*, Transactions of the Connecticut Academy of Sciences and Arts, Vol. IX.

Frederick, S.G. Loewenstein and T. O'Donoghue (2002), 'Time Discounting and Time Preference: A Critical Review,' *Journal of Economic Literature*, XL, 351-401.

Gilboa, I. and D. Schmeidler (1989), 'Maxmin Expected Utility with a Non-unique Prior,' *Journal of Mathematical Economics*, 18, 141-53.

Glimcher, P.W., J. Kable, and K. Louie (2006), 'Neuroeconomic Studies of Impulsivity: Now or Just As Soon As Possible?,' mimeo, NYU.

Gul, F. and W. Pesendorfer (2001), 'Temptation and Self-Control,' *Econometrica*, 69, 6, 1403-36.

Gul, F. and W. Pesendorfer (2006 - this volume), 'The Case for Mindless Economics.'

Jevons, W. S. (1866), 'Brief Account of a General Mathematical Theory of Political Economy', *Journal of the Royal Statistical Society*, 29, 282-87, (Read in Section F of the British Association, 1862).

Jevons, W. S., (1871), *The Theory of Political Economy*, London: Macmillan.

Kahneman, D. and A.B. Krueger (2006), 'Developments in the Measurement of Subjective Well-Being,' *Journal of Economic Perspectives*, 20, 1, 3-24.

Keane, M.P. (2006), 'Structural vs. Atheoretic Approaches to Econometrics,' Yale University, mimeo.

Kreps, D. and E.L. Porteus (1978), 'Temporal Resolution of Uncertainty and Dynamic Choice Theory,' *Econometrica*, 46, 1, 185-200.

Koopmans, T. C., (1947), 'Measurement without Theory,' *The Review of Economics and Statistics*, 29, 161-72.

Koopmans, T.C. (1960), "Stationary Ordinal Utility and Impatience," *Econometrica* 28, 287-309.

Laibson, D. (1996), 'Golden Eggs and Hyperbolic Discounting,' *Quarterly Journal of Economics*, CXII, 443-77.

Loewenstein, G. and E.D. O'Donoghue (2005), 'Animal Spirits: Affective and Deliberative Processes in Economic Behavior,' mimeo, Cornell University.

Maccheroni, F., M. Marinacci, and A. Rustichini (2006), 'Preference-based Decision Processes,' mimeo, University of Minnesota.

Marschak, J. (1952), 'Economic Measurement for Policy and Predictions,' in *Studies in Econometric Methods*, ed. W.C. Hood and T.C. Koopmans, 1-26, New York: Wiley.

McClure, S.M., D. Laibson, G. Loewenstein, and J.D. Cohen (2004), 'Separate Nueral Systems Value Immediate and Delayed Monetary Rewards,' *Science*, 306, 503-7.

Miller E.K. and J. Cohen (2001), 'An Integrative Theory of Prefrontal Cortex Function,' *Annual Review of Neuroscience*, 24, 167-202.

Nehring, K. (2006), 'Self-Control Through Second Order Preferences,' mimeo, UC Davis.

Norman, D.A. and Shallice, T. (1980), 'Attention to Action: Willed and Automatic Control of Behaviour,' Centre for Human Information Processing (Technical Report #99), University of California, San Diego.

O'Donoghue, E.D. and M. Rabin (1999), 'Doing It Now or Doing It Later,' *American Economic Review*, 89, 103-24.

Ok, E. and Y. Masatlioglu (2004), 'A General Theory of Time Preferences,' mimeo, NYU.

Pareto, V., (1906), *Manuale di Economia Politica, con una Introduzione alla Scienza Sociale*, Milano.

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

Rubinstein, A. (2003), '"Economics and Psychology"? The Case of Hyperbolic Discounting,' *International Economic Review*, 44, 4, 1207-16.

Rubinstein, A. (2006), 'Instinctive and Cognitive Reasoning: A Study of Response Times,' *Economic Journal*, forthcoming.

Rubinstein, A. and Y. Salant (2006), 'A Model of Choice from Lists,' *Theoretical Economics* 1, 1, 3-17.

Rumelhart, D.E. and J.L. McClelland (1986), ed., *Parallel Distributed Processing: Exploration in the Microstructure of Cognition*, Cambridge, MA: MIT Press.

Schneider, W. and R.M. Shiffrin (1977), 'Controlled and Automatic Human Information Processing: 1. Detection, Search, and Attention,' *Psychological Review*, 84, 1-66.

Simon, H., (1982), *Models of Bounded Rationality*, MIT Press, Cambridge, MA.

Smith, A. (1759), *The Theory of Moral Sentiments*, A. Millar, London, A. Kincaid and J. Bell, Edinburgh.

Smith, K. and J. Dickhaut (2004), 'Economics and Emotions: Institutions Matter,' *Games and Economic Behavior*, 52(2), 316-35, special issue on *Neuroeconomics*.

Tversky, A. (1977), 'Features of Similarity,' *Psychological Review*, 84, 327-52.

Wilson, M. and M. Daly, (2003), 'Do Pretty Women Inspire Men to Discount the Future?' *Biology Letters (Proc. The Royal Society. Lond. B*, S177-9.