The Idea of Neuroeconomics

with

Wolfgang Pesendorfer

Introduction

Objectives

Identify (define) two approaches to behavioral economics:

– the standard (revealed preference) approach and

– neuroeconomics

Standard approach is the extension of standard choice theoretic methods to analyze preference related variables that are often ignored.

Neuroeconomics refers to research in line with recent proposals for radical change in the methods economics.

Neuroeconomic principles:

(I) Psychological and physiological evidence refutes standard economic models and economic methodology.

(II) A normative model (welfare economics) cannot be based on behavior because behavior and well-being (i.e., true utility) are distinct.

Purpose of this talk: (0) convince you that this is a useful classification; (1) argue that (I) is wrong; (2) show that neuroeconomics offers no alternative to the standard welfare notion.

Key References:

1. Camerer, Loewenstein, and Prelec, "Neuroeconomics: How Neuroscience Can Inform Economics" (CLP)
2. Camerer, Loewenstein, and Prelec, "Neuroeconomics: Why Economics Needs Brains" (CLP B)
3. Köszegi and Rabin, "A Model of Reference Dependent Preferences"
4. Gruber and Köszegi, "Is Addiction Rational?"
5. Kreps, "A Representation Theorem for Preference for Flexibility"

Which is (more) true?

"Much risk averse behavior is driven by immediate fear responses to risks, and fear, in turn, seems to be largely traceable to the amygdala." (CLP).

"A decision-maker is (globally) risk averse, (...) if and only if his von Neumann-Morgenstern utility at the relevant (all) wealth levels is concave." Ingersoll (1987).

Depends on the discipline. (i) The meaning of the abstraction "risk aversion" depends on the discipline. (ii) Relation between abstractions and observables depends on the discipline. (iii) Disciplines differ in the questions they ask.

Argument 1: Neuroscience evidence cannot refute economic models because to latter make no assumptions or draw no conclusions about the physiology of the brain.

Two errors: (1) the standard model is treated as if it is a model of the brain; (2) it is assumed that brain science delivers parameters of a utility function.

Argument 2: The economists notion of welfare is a definition, not a theory. It cannot be disproved; it can only be replaced by an alternative definition. Neuroeconomics offers no such alternative.

Examples

Rabin (1996) asks *"Is maximizing utility the right model?"*. We interpret all statements such as "do people maximize utility?" "Are people rational?" as versions of (I).

"But models with time-inconsistent agents extend the role of government policy by breaking down revealed preference concepts of consumer choice. The argument that people act in their best interests, so-barring well-known qualifications–the government should leave them alone, is immediately invalidated in our setting. Therefore, although our models are explicitly of the no-externality type, a benevolent social planner would want to intervene in this economy." Gruber and Köszegi (2001).

The Standard Approach to Behavioral Economics

"A Representation Theorem for Preference for Flexibility" Kreps (1979)

X is a finite set

$\mathcal{D}$ is the set of all nonempty subsets of X.

$\succeq$ is a complete transitive binary $\mathcal{D}$.

$$A \succeq B \text{ implies } A \sim A \cup B \qquad (D)$$

Proposition: *(Demand-theory decision-maker) The preference relation $\succeq$ satisfies (D) if and only if there exists a utility function $u : X \to \mathbb{R}$ such that the function $U(A) := \max_{x \in A} u(x)$ represents $\succeq$.*

Example:

DM may or may not win lottery; (w denotes win and l denotes not win.)

Actions: x: buy expensive car in period 2; y: buy nothing.

Assume that $u_w(x) > u_l(y)$ and $u_w(x) < u_l(y)$

$$\begin{aligned} U(\{x\}) &= pu_w(x) + (1-p)u_l(x) \\ U(\{y\}) &= pu_w(y) + (1-p)u_l(y) \\ U(\{x,y\}) &= pu_w(x) + (1-p)u_l(y) \end{aligned} \tag{1}$$

$\Rightarrow \quad U(\{x,y\}) > \max\{U(\{x\}, U(\{y\})\}$ violating (D).

Kreps (1979): can equation (1) above be used to represent the preferences of a DM with an **intrinsic preference for flexibility?**

What are the behavioral implications of maximizing a function U, where

$$U(A) = \sum_{s \in S} \max_{x \in A} u_s(x) \qquad (PF)$$

$$U(A) = \sum_{s \in \{w,l\}} \max_{x \in A} u_s(x) \qquad (1)$$

(PF) and (1) are not the same: States in S are subjective, w, l are objective

What assumptions characterize the formula PF?

$$A \subset B \text{ implies } B \succeq A \qquad (M)$$

$$A \subset B,\ A \succeq A \cup C \text{ implies } B \succeq B \cup C \qquad (ODMU)$$

Proposition: *There exists a utility function of the form (PF) that represents* $\succeq$ *if and only if* $\succeq$ *satisfies* (M) *and* $(OMDU)$.

The Methodology

1. Kreps starts with a novel psychological phenomenon: Preference for Flexibility (PF).
2. The psychological phenomenon PF is defined in terms of its implied behavior (over sets). The preferences over sets are the primitive.
3. The theorem invites the reader consider the psychological process underlying its axiom and the formula. This is the *suggestiveness* of the model.
4. Axioms summarize the empirical content of the theory. The *formal content* of the theory is determined by the empirical validity of its assumptions and the usefulness and measurability of the parameters.

Neuroeconomics

"'*This new approach, which I consider a revolution, should provide a theory of how people decide in economic and strategic situations,'* said Dr. Aldo Rustichini,.... *'So far, the decision process has been for economists a black box.'* "

"....Dr. Cohen and his colleagues,..., took images of people's brains as they played the ultimatum game, ... Most people in the shoes of Player 2 refuse to take amounts under $2 or $3, Dr. Cohen said. They would rather punish the first player than feel cheated. *'But this makes no economic sense,'* he said. *'You're better off with something than nothing.'*" New York Times, June 17, 2003.

Neuroeconomic research: Köszegi and Rabin (2005) (KR)

X is a finite set; Y is the set of all nonempty subsets of X.

$U : X \times X \to \mathbb{R}$ is a reference-dependent utility function; $U(x, z)$ is the utility of x given the reference z.

Reference point: the x that ultimately gets chosen.

Choice: $x \in A$ is a possible choice if for all $x' \in A$

$$U(x, x) \geq U(x', x) \tag{2}$$

Additional assumption

$$U(x, z) \geq U(z, z), x \neq z \text{ implies } U(x, x) > U(z, x) \tag{3}$$

A function U satisfying (2)−(3) is a reference-dependent utility function. KR assume that U has the following form:

$$U(x,z) = \sum_{j=1}^{k} u_j(x) + \sum_{j=1}^{k} \mu(u_j(x) - u_j(z)) \tag{4}$$

where μ is an increasing function with $\mu(0) = 0$. They insist that the k relevant dimensions of consumption "should be specified based on psychological principles".

Observable Implications of the KR Theory

Choice function: $c(A) \subset A$ and $c(A) \neq \emptyset$

Choice function induced by $\succeq$*:* $C_{\succeq}(A) = \{x \in A \,|\, x \succeq z \forall z \in A\}$

KR choice function (induced by U):

$$C_{kr}(A, U) = \{x \in A \,|\, U(x, x) \geq U(x', x) \forall x' \in A\}$$

Special KR choice function: Choice function where U has the form

$$U(x, z) = \sum_{j=1}^{k} u_j(x) + \sum_{j=1}^{k} \mu(u_j(x) - u_j(z)) \tag{4}$$

$C_{\succeq}(A) = \emptyset$ unless certain assumptions are made on $\succeq$.

$C_{\succeq}$ is a choice function whenever $\succeq$ is complete and transitive

transitivity is not necessary for $C_{\succeq}$ to be a choice function.

The following proposition is not in KR:

Proposition: *The following three conditions are equivalent:*

(i) c is a KR choice function

(ii) c is induced by some complete binary relation

(iii) c is a special KR choice function

Summary

KR would say that the proposition is besides the point: Their point has nothing to do with relaxing transitivity. They focus on the suggestiveness of their theory.

"By all intuition and evidence, the feeling of loss when giving up a mug is a real hedonic experience, and making choices reflecting that real hedonic experience is partly rational. But as interpreted by Kahneman (2001) and Loewenstein, O'Donoghue, and Rabin (2003), people seem to over-attend to this experience because they ignore that the sensation of loss will pass very quickly-behaving as if they would spend much time longing for the mug they once had."

The suggestiveness' is real; not an "as if" statement. Utility indices (u_j's) and attachment disutilities μ are not unique up to some transformation, they are unique and measurable. KR view the identification of the dimension of hedonic utility as a matter of craft.

"In ... the previous version of this paper, we argue at length.... that the consumption dimensions used in our framework should be specified based on psychological principles, and not necessarily correspond directly to quantities of different products."

"Several aspects of our theory, however, render it short of fully general and formulaically applicable. Many or our specific assumptions are based on intuition rather than direct evidence."

"Is Addition 'Rational"' Gruber and Köszegi (2001) (GK)

Model 1: DM maximizes U_t in period t

$$U_t(\bar{c}) = \sum_{i=0}^{T-t} u_i(\bar{c})\delta^i$$

where $\bar{c} = (c_1, \ldots, c_T)$, $c_t = (a_t, n_t)$,

$$u_t(\bar{c}) = v(a_t, S_t(\bar{c})) + w(n_t)$$

with $S_t(\bar{c})$ defined inductively: $S(0) = a_0$; $S_t(\bar{c}) = (1-d)(S_{t-1}(\bar{c}) + a_t)$.

Dynamical consistency: For $\bar{c}$ and $\bar{c}'$ such that $c_i = c'_i$ for all $i \leq t$, $U_t(\bar{c}) \geq U_t(\bar{c}')$ implies $U_{t+1}(\bar{c}) \geq U_{t+1}(\bar{c}')$

Model 2: Replace the expression for U_t with

$$U_t(\bar{c}) = u_t(\bar{c}) + \beta \sum_{i=1}^{T-i} u_i(\bar{c})\delta^i \qquad (5\beta)$$

where $0 < \beta < 1$.

Model 2 is not dynamically consistent

Given additional assumptions both versions that for $t' > t$ and increase in the period t price of $a_{t'}$ decreases the consumption of a_t.

Data can't distinguish between Model 1 and 2.

But the two versions have very different welfare implications.

For both versions, GK say U_1 measures welfare.

First Problem with the Welfare Criterion

Let $d = .5$, $a_0 = 0$, $v(a, S) = a - \lambda S$, $w(n) = n$.

Let $\bar{c} = (c_1, \ldots, c_T)$, $c_t = (a_t, n_t)$, $a_1 = 1$, $a_t = 0$ for all $t > 1$ and

$n_t = 0$ for all t.

By choosing λ appropriately, we can make $U_1(\bar{c}) = 0$.

Let $\bar{c}' = (c'_1, \ldots, c'_t)$, where $c'_t = (0, 0)$ for all t.

Then $U_1(\bar{c}) = U_1(\bar{c}')$ and $U_t(\bar{c}) < U_t(\bar{c}')$.

$\bar{c}'$ Pareto-dominates $\bar{c}$ but welfare criterion ranks the two as equivalent.

Second Problem with the Welfare Criterion

Alternative specification of utility:

$$\hat{U}_t(\bar{c}^*) = \sum_{i=0}^{T-t} (\lambda_t a_t^* + n_t^*)\delta^i \tag{9}$$

We can choose λ_t's so that $\hat{U}_1 = U_1$ and $\hat{U}_t(\bar{c}^*) \geq \hat{U}_t(\bar{c}^{**})$ if and only if $U_t(\bar{c}^*) \geq U_t(\bar{c}^{**})$.

However, for all $t \geq 1$

$$\hat{U}_t(\bar{c}') = U_t(\bar{c})$$

Behaviorally equivalent utility functions yield different welfare conclusions.

Calibration

"We will ignore all these, and assume that the only disutility from smoking is in the increased chance of early death. Viscusi [1993] reviews the literature on life valuation and suggests a consensus range of 3-7 million 1990 dollars for the value of a worker's life; choosing the midpoint value and expressing it in current dollars gives a figure of $6.4 million..... Making reasonable....we find that an extra year at the end of a smoker's life is worth $99,110.... At these figures, the cost in terms of life years lost per pack of cigarettes is $30.4."

KR extract a parameter that cannot be identified through choice behavior from Viscusi's revealed preference estimates.

Neuroeconomics Critique of the Standard Approach

"The first is that the methodological strictures against a hedonic notion of utility are a relic of an earlier period in which a behavioristic philosophy of science held sway. Subjective states are now a legitimate topic of study, and hedonic experiences such as pleasure, pain, satisfaction or discomfort are considered open to useful forms of measurement. The second observation is that it may be rash to assume as a general rule that people will later enjoy what they want now. The relation between preferences and hedonic consequences is better studied than postulated." Kahneman (1994).

First corresponds to (I) and second corresponds to (II).

"The foundations of economic theory were constructed assuming that details about the functioning of the brain's black box would not be known... But now neuroscience has proved Jevons' pessimistic prediction wrong; the study of the brain and nervous system is beginning to allow direct measurement of thoughts and feelings." (CLP)

Mistakes, Multiple-Selves, and Multiple-Processes

Argument: People make systematic mistakes. These mistakes cannot be treated as noise. Economics needs a theory of mistakes.
Mistakes point to a general flaw in the way economists model preferences.

"American visitors to the U.K. suffer numerous injuries and fatalities because they often look only to the left before stepping into streets, even though they know traffic approaches from the right. One cannot reasonably attribute this to the pleasure of looking left or to masochistic preferences. The pedestrian's objectives - to cross the street safely - are clear, and the decision is plainly a mistake." (Bernheim and Rangel (2005))

"Human behavior thus requires a fluid interaction between controlled and automatic processes, and between cognitive and affective systems. However, many behaviors that emerge from this interplay are routinely and falsely interpreted as being the product of cognitive deliberation alone..... Since we see only the top of the automatic iceberg, we naturally tend to exaggerate the importance of control." CLP

Wanting versus Liking, Welfare versus Happiness:

"... ability of economists to make normative statements is premised on the idea that giving people what they want makes them better off. But, there is considerable evidence from neuroscience and other areas of psychology that the motivation to take an action is not always closely tied to hedonic consequences. Berridge ... finds that certain lesions and pharmacological interventions can selectively enhance a rat's willingness to work for a food, without changing the pleasure of eating the food, as measured, admittedly somewhat questionably, by the animal's facial expression...." CLP.

The Standard Economics Responses:

Response 1: *Neuroeconomics offers no data that can be used to test economic theories. Instead, neuroeconomics provides evidence that economic models of utility maximization are inadequate for the analysis of the brain. Since the economic models were designed to deal with economic phenomena such evidence is irrelevant.*

• Economics and psychology ask different questions.

• Therefore, they have different abstractions.

• Preference, choice function, GDP, utility, etc. have proven to be useful abstraction in economics.

• Brain mechanisms can't offer evidence against transitivity.

• Or against purposeful behavior.

• The abstractions of economics were not intended for analyzing the brain.

"For example, when economists think about gambling they assume that people combine the chance of winning (probability) with an expectation of how they will value winning and losing ("utilities"). If this theory is correct, neuroeconomics will find two processes in the brain-one for guessing how likely one is to win and lose, and another for evaluating the hedonic pleasure and pain of winning and losing-and another brain region which combines probability and hedonic sensations."

Camerer (2005)

Consider the reverse procedure:

We find that drug addicts generally satisfy SARP.

Conclude that the addicts' behavior maximizes some utility function.

Argue that the concept of the "limbic system" is misguided:

Consumer maximizes some utility implies no separate brain functions.

Neuroeconomic Evidence

"When the core body temperature falls below the 98.6F set-point, almost anything that raises body temperature (such as placing one's hand in warm water) feels good, and the opposite is true when body temperature is too high. Similarly, monitors for blood sugar levels, intestinal distention and many other variables trigger hunger. Homeostasis means preferences are 'state-dependent' in a special way..."

Economics models are silent on body temperature and intestinal distention; facts regarding these cannot refute them. Standard economics is not committed to any particular theory of what makes people feel good. Nor does it assume that feeling good is all that people care about.

"Inferring preferences from a choice does not tell us everything we need to know, and may tell us very little.... Al and Naucia, who both refuse to buy peanuts at a reasonable price (cf. Romer, 2000).... Al turned down the peanuts because he is allergic: consuming peanuts ..could even be fatal. Naucia turned down the peanuts because she ate a huge bag of peanuts at a circus....While Al and Naucia both revealed an identical disutility, a neurally-detailed account tells us more. Al has an inelastic demand for peanuts-you can't pay him enough to eat them!-while Naucia would try a fistful for the right price.....

Instead of providing a neurally-detailed account, a standard economist would vary the price.

"A fourth problem with preference is that people are assumed to value money for what it can purchase-that is, the utility of income is indirect.... But ... it appears that similar brain circuitry-dopaminergic neurons in the midbrain–is active for a wide variety of rewarding experiences– drugs, food, attractive faces...–and money rewards. This means money may be directly rewarding, and it's loss painful...."

Economic models are not (mathematical) description of physiological and psychological mechanisms that underlie choice.

"Addiction is an important topic for economics because it seems to resist rational explanation. .... Laibson (2001) created a pioneering formal model of cue-dependent use, showing that there are multiple equilibria in which cues either trigger use or are ignored. The more elaborate model of Bernheim and Rangel (in press), is a paradigmatic example of how economic theory can be deeply rooted in neuroscientific details. They assume that when a person is in a hot state they use drugs; in a cold state, whether they use is a rational choice. A variable S, from 0 to N, summarizes the person's history of drug use. When he uses, S goes up; when he abstains S goes down.... By assuming the cold state reflects the person's true welfare, they can also do welfare analysis..."

This is research preference not a critique.

What the authors find valuable is a description that is suggestive of "neuroscientific details."

They like the suggestiveness of the models; the formal content of these models is in neuroeconomics.

Direct Evidence from Neuroscience

"asking the brain, not the person," and hence

"add precision to functions and parameters in standard .. models."

There is no example of this ever being done, no suggestions as to how it could be done.

No neuroscience criterion for distinguishing $\delta = .97$ vs $\delta = .98$.

The only conclusions CLP draw regarding the discount factor and the risk aversion parameter is that these may not be useful constructs.

Response 2: *Neuroeconomics offers no coherent alternative to the standard definition of welfare in economics. It offers no criteria by which alternative welfare concepts could be evaluated. Therefore, neuroeconomics can offer no serious challenge to standard welfare economics.*

"A third problem with preferences is that there are different types of utilities which do not always coincide.....For example, Berridge and Robinson (1998) have found distinct brain regions for 'wanting' and 'liking', which correspond roughly to choice utility and experienced utility. The fact that these areas are dissociated allows a wedge between those two kinds of utility... If the different types of utility are produced by different regions, they will not always match up."

Neuroeconomics never provides a procedure as to how true utilities are to be assessed.

Models posit separate but related utilities for behavior and welfare but never describe how we could identify the latter explain the relationship between the two:

$$U_t(\bar{c}) = u_t(\bar{c}) + \beta \sum_{i=1}^{T-i} u_i(\bar{c})\delta^i$$

$$U_t(\bar{c}) = \sum_{i=0}^{T-t} u_i(\bar{c})\delta^i$$

Subjective readings of facial expressions can't be a legitamite alternative.

Neuroeconomists only consider situations where the good choice is clear. We never learn true utilities.

A useful welfare criterion has to be distinguishable from preference statements of the researcher.

Welfare in economics is a *definition* and *not a theory*.

Neuroeconomists treat this definition *as if* it were a theory of happiness and proceed to find evidence against this theory.

"Economics proceeds on the assumption that satisfying people's wants is a good thing. This assumption depends on knowing that people will like what they want. If likes and wants diverge, this would pose a fundamental challenge to standard welfare economics. Presumably welfare should be based on 'liking.' But if we cannot infer what people like from what they want and choose, then an alternative method for measuring liking is needed, while avoiding an oppressive paternalism." CLP

The only challenge to a normative definition can come from another, (better), definition of welfare. CLP never describe sensations and emotions are to be quantified as true utility.

Paternalism

Once we base welfare judgements on not what individuals want but on what we think is good for them, paternalism can't be avoided.

Camerer, Issacharov, Loewenstein, O'Donoghue, and Rabin (CILOR) "Regulation for Conservatives: Behavioral Economics and the Case for Asymmetric Paternalism," introduce the following criterion:
"A regulation is asymmetrically paternalistic if it creates large benefits for those who make errors, while imposing little or no harm on those who are fully rational."

Thaler and Sunstein (TS) suggest *libertarian paternalism*:

"First, the libertarian paternalist might select the approach that majority would choose if explicit choices were required and revealed."

"Second, the libertarian paternalist might select the approach that would force people to make their choices explicit."

"Third, the libertarian paternalist might select the approach that minimizes the number of opt-outs."

Motivated by evidence on retirement plan choice,

The specification of the default option affects ultimate choices.

CILOR use this example to motivate their principle,

TS principles work only for such situations.

TS offer no arguments for their principles.

What to do when the principles conflict or when they don't apply.

CILOR don't say which preferences are bounded rationality.

When benefits are large and when there is little or no harm.

Conclusion

"More ambitiously, students are often bewildered that the models of human nature offered in different social sciences are so different, and often contradictory. ... An identical question on a final exam in each of the fields about trust, for example, would have different "correct" answers in each of the fields. It is possible that a biological basis for behavior in neuroscience, perhaps combined with all-purpose tools like learning models or game theory, could provide some unification across the social sciences." CLP B.

Neuroeconomists assume that three is a single set of questions.

Their questions resemble the current questions of psyhology more than the current questions of economics.

Psychology/neuroscience abstractions are viewed as "real".
It is more useful to think of the 'choice' as resulting from the interaction of multiple systems–an automatic biological one which homeostatically shuts down the body when it is tired, and a controlled cognitive one which fights off sleep when closing your eyes can be fatal. CLP B.

It also highlights the need for economic models of the primitive reward circuitry, which would apply equally to man and rat. CLP B.

Economics constructs are “all purpose”; i.e., language.

Approved uses of game theory and other economic tools amounts to a restatement of recent research in neuroscience with different words.

“In economic language, the experiments create a situation where the utility of food and disutility of work remain the same, but the amount of work-for-reward goes up. This implies that it is possible to be motivated to take actions that bring no pleasure.” CLP

What Camerer, Loewnstein, and Pralec fail to show is that neuroeconomics can "improve economics on its own terms,"

Earlier work on Psychology and Economics offers experimental evidence against common assumptions.

Neuroeconomics offers no central result or defining experiment.

Only the premise that a complete understanding of behavior must entail a complete understanding of what goes on in the brain.

The problem with this (reductionist) view that it leaves no criteria for economist to evaluate work;

Until that complete understanding comes, the only way to decide which economics paper is better is to see which one is more firmly based in neuroscientific detail.

A Thought Experiment

Would incentive theory be irrelevant in a world with perfect lie-detectors?

Lie-detectors may cost too much,

There may be institutional restrictions on their use etc.

A theory relating hormonal activity to stock market behavior need not supercede economic theory; economists may not have access to hormonal data or they view hormonal factors as intervening variables rather than the causes of behavior.

A Thought Experiment

Does incentive theory need lie-detectors? A truth-economist might argue as follows:

A thorough understanding of incentives requires a complete understanding of when people lie. This cannot be done only by analysis of incentives but requires direct analysis of the human behavior. Modern technology allows us to directly measure when people lie. This will (i) facilitate a deeper understanding of economic contracting; (ii) lead to more effective contracts that in turn rely on lie-detectors.