Dopamine and Reward Prediction Error

Professor Andrew Caplin, NYU
The Center for Emotional Economics, University of Mainz

June 28-29, 2010
The Proposed Axiomatic Method

- Work joint with Mark Dean, Paul Glimcher, and Robb Rutledge
- The potential for neuroeconomic researchers lies in **complementarities**
  - Neuroscientific measurement and biological understanding
  - Economic modelling and organizational principles
- Crucial that there is a methodological framework which allows communication across disciplines
  - "Utility" should not mean different things to neuroscientist than to economist
The axiomatic approach developed by decision theorists can provide a unifying role in mature theories

- Not "abstract" axioms connecting non-observables to one another
- But "testable axioms connecting latent variables to (ideal) experimental data
- Novelty of neuroeconomics lies in hybrid nature of data and theories

Economists introduced methods out of frustration with other treatments of latent variables

- The same reasons apply with even more force in neuroscience

Ideally leads to a progressive experimental agenda:

- When data fail to satisfy, amend theory
- When data do satisfy, refine e.g. to expected utility theory
Intuitions concerning utility also important in neuroeconomics

Dopamine is a neurotransmitter: transmits information from one part of the brain to another

 Neurobiological studies have associated dopamine with:

- **Choice:** Dopamine manipulations affect choice behavior in animals
- **Preference:** Dopamine encodes information on ‘revealed preferences’
- **Beliefs:** Changes in expectations modify dopamine activity
- **Learning:** Dopamine manipulations affect the way people learn
- **Addiction:** Many drugs of addiction act directly on dopamine

Understanding dopamine may give valuable insight into economic behavior
Our Approach

• There remain barriers to incorporating understanding from dopamine into economics
  • Competing theories of what dopamine does
  • No common language between economics and neuroscience
  • Treatment of unobservables

• We take an axiomatic approach to testing a model of dopamine
  • Provide a complete list of testable predictions
  • Provide a common language between disciplines by defining unobservables
  • Failure of particular axioms will aid model development

• Our aim is to systemize current neurobiological understanding

• Initially this may prove to be as much an ‘export from,’ as an ‘import to’ economics
• Dopamine fires only on receipt of *unpredicted* rewards

• Otherwise will fire at *first predictor* of reward

• If an expected reward is not received, dopamine firing will pause
• Reward Prediction Error hypothesis: Dopamine responds to the difference between *experienced* and *anticipated* reward
  • Information on preferences?
  • Information on beliefs?
  • Involved in reinforcement learning?

• We provide an axiomatic basis for this model of dopamine activity

• Use these axioms to develop parsimonious, non-parametric tests of the hypothesis

• Use experimental data from humans to perform these tests
• We consider the simplest possible environment in which we can think about reward prediction error

• Consists of prizes and lotteries:
  • $Z$: A metric space of prizes with typical elements $z, w$
  • $\Lambda$: Set of all simple probability distributions (lotteries) on $Z$ with typical element $p, q$
  • $\Lambda(z)$: Set of all probability distributions whose support includes $z$

• Let $e_z$ be the lottery that gives prize $z$ with certainty
In our idealized data set, we assume we observe a function:

\[ \delta : \mathcal{M} \rightarrow \mathbb{R} \]

\[ \mathcal{M} = \{(z, p) | z \in Z, \ p \in \Delta(z)\} \]

where \( \delta(z, p) \) is the amount of dopamine released when a prize \( z \) is obtained from a lottery \( p \in \Delta(z) \)
A Graphical Representation

Dopamine released when prize 2 is obtained

Dopamine released when prize 1 is obtained

p = 0.2
A Formalization of the RPE hypothesis

- Need to find some way of defining predicted reward from lotteries and experienced reward for prizes such that:
  - Contains all the information which determines dopamine response
  - Dopamine response is increasing in experienced and decreasing in predicted reward
  - Dopamine always responds to ‘no surprise’ in the same way
- Demand predicted reward of $e_z$ equal to the experienced reward of $z$
- We say that a dopamine release function has an RPE representation if we can find functions $r : \Lambda \rightarrow \mathbb{R}$ and $E : r(Z) \times r(\Lambda)$ such that:
  - $\delta(z, p) = E[r(e_z), r(p)]$
  - $E$ is strictly increasing in its first argument and decreasing in its second argument
  - $E(x, x) = E(y, y)$ for all $x, y \in r(Z)$
Axiom A1: Coherent Prize Dominance

\[
\forall (z, p), (w, p), (z, q), (w, q) \in M \quad \delta(z, p) > \delta(w, p) \Rightarrow \delta(z, q) > \delta(w, q)
\]
Necessary Condition 2: Coherent Lottery Dominance

Axiom A2: Coherent Lottery Dominance

for all \((z, p), (w, p), (z, q), (w, q) \in M\)
\[\delta(z, p) > \delta(z, q) \Rightarrow \delta(w, p) > \delta(w, q)\]
Necessary Condition 3: Equivalence of Certainty

- **Axiom A3: No Surprise Equivalence**

\[ \delta(z, e_z) = \delta(w, e_w) \quad \forall \ z, w \in Z \]
In general, these conditions are necessary, but not sufficient for an RPE representation.

However, in the special case where we look only at lotteries with two prizes they are:

**Theorem 1:**
If $|Z| = 2$, a dopamine release function $\delta$ satisfies axioms A1-A3 if and only if it admits an RPE representation.

Thus, in order to test RPE in case of two prizes, we need only to test A1-A3.
• Generate observations of $\delta$ in order to test axioms
• Use a data set containing:
  • Two prizes: win $5, lose $5
  • Five lotteries: $p \in \{0, 0.25, 0.5, 0.75, 1\}$
• Do not observe dopamine directly
  • Use fMRI to observe activity in the Nucleus Accumbens
  • Brain area rich in dopaminergic neurons
Experimental Design

Fixation: 12 seconds

Options: 3.6 seconds
to view options

Choice selection: 1.2 seconds
to make a choice by button press
(fixation cross extinguished)

Choice: 8.4 seconds to
view the choice just made

Outcome: 3.6 seconds to
view outcome of choice
(outcome illuminated)

Task design

Lottery 1 EV: $0
Lottery 2 EV: -$1.25

Trial length:
~29 seconds
14 subjects (2 dropped for excess movement)

‘Practice Session’ (outside scanner) of 4 blocks of 16 trials

2 ‘Scanner Sessions’ of 8 blocks of 16 trials

For Scanner Sessions, subjects paid $35 show up fee, + $100 endowment + outcome of each trial

In each trial, subject offered one option from ‘Observation Set’ and one from a ‘Decoy Set’
• Need to determine which area of the brain is the Nucleus Accumbens

• Two ways of doing so:
  • Anatomical ROIs: Defined by location
  • Functional ROIs: Defined by response to a particular stimulus

• We concentrate on anatomical ROI, but use functional ROIs to test results
Constructing Delta
Anatomical Regions of Interest [Neto et al. 2008]
• We now need to estimate the function $\bar{\delta}$ using the data

• Use a between-subject design
  • Treat all data as coming from a single subject

• Create a single time series for an ROI
  • Average across voxels
  • Convert to percentage change from session baseline

• Regress time series on dummies for the revelation of each prize/lottery pair
  • $\bar{\delta}(x, p)$ is the estimated coefficient on the dummy which takes the value 1 when prize $x$ is obtained from lottery $p$
• Axioms hold
• Nucleus Accumbens activity in line with RPE model
• Experienced and predicted reward ‘sensible’
Time Paths

A

Options  Choice  Delay  Outcome

Percent signal change (%)

Anatomical

Time (s)
Two Different Signals?
fMRI activity in Nucleus Accumbens does satisfy the necessary conditions for an RPE encoder. However, this aggregate result may be the amalgamation of two separate signals:

- Vary in temporal lag
- Vary in magnitude
Axioms + experimental results tell us we can assign numbers to events such that NAcc activity encodes RPE according to those numbers.

Can we use these numbers to make inferences about beliefs and rewards?

- Are they ‘beliefs’ and ‘rewards’ in the sense that people usually use the words?
- Can we find any ‘external validity’ with respect to other observables?
  - Behavior?
  - Obviously rewarding events?

Can we then generalize to other situations?
Economic Applications

- New way of observing beliefs
- Makes ‘surprise’ directly observable
- Insights into mechanisms underlying learning
- Building blocks of ‘utility’
• We provide evidence that NAcc activity encodes RPE
• Can recover consistent dopaminergic ‘beliefs’ and ‘rewards’
• Potential for important new insights into human behavior and ‘state of mind’