EC 490

Lecture 12: Isolating decision variables in the brain

1. In this lecture, instead of reviewing evidence we consider methods for representing decision processes, and implications of these methods for bringing neuroscientific evidence to bear.

2. It is often claimed (e.g., see Gul & Pesendorfer at http://www.princeton.edu/~pesendor/mindless.pdf) that economics fundamentally *is* the science of rational decision. I think that claim implies a drastically over-narrow conception of economics. Nevertheless, as soon as we're considering processes by which people decide on consumption baskets, trade-offs and strategies, we're clearly in the heartland of – at least – most microeconomics since 1870.

3. Corado *et al* say that "decision variables … link the processes of option evaluation and action selection". Traditionally, such variables have been regarded as necessarily not *directly* observable. They are inferred, either from choices (*revealed preference theories*), or from economists' analysis of options conjoined with assumptions that agents are rational in some specific normative sense. More recently, some theorists have focused on procedural constraints imposed by time pressures and hypothesized mechanical models of cognition to develop computational models. These models invariably include decision variables. For every possible view about the extent to which these variables should be imagined to correspond to real neural processes, some cognitive scientist or philosopher has defended that view.

4. One way of trying to identify decision variables that almost no one regards as adequate (or even useful) is to ask people

to introspect. Mountains of evidence show that when people are asked to do this, what they do is confabulate stories that fit whatever cultural conventions they've learned.

5. A standard method of economists is to design incentive-compatible experimental mechanisms. This means: put subjects in games (typically, auctions over simple or complex lotteries) in which there are no Nash equilibria in which subjects would conceal information. Then, it is supposed, one can assume that subjects' bids will reveal their true expected-value functions. Corrado *et al* criticize this method on grounds that, like reliance on introspection, it depends on subjects knowing what their decision variables are. This criticism seems to me to be misplaced. The use of incentive-compatible mechanisms relies only on the much weaker assumption that subjects can learn to converge on strategies that are (often) subconsciously evaluated (presumably by parts of their brain) as optimal. They need have no self-aware access to the variables the economist would write into models of their optimization, or that a computational model would write into algorithms for finding optima. The evidence from the Glimcher lab reviewed previously, that single neurons in monkeys' visuomator system can learn to track Nash equilibria, seem relevant here. Glimcher and his associates didn't ask their monkeys to do anything resembling reporting the contents of introspection.

6. Experiments using incentive-compatible mechanisms aren't *alternatives* to designs aimed at eliciting revealed preferences, as Corrado *et al* suggest. They *are* elicitations of revealed preference. Bidding in an auction isn't so much a kind of *report* as a kind of *behavior* (provided the experimenter uses real and meaningful prizes).

7. I'll now show you an example of a very simple incentive-compatible mechanism that is commonly used to elicit preferences, a *multiple price list design.*

8. Corrado *et al* claim that methods based on revealed preference assume that preferences are *stationary*: "[b]ecause these methods generally rely on combining data across many trials to compute indifference points, they typically assume that the decision variables summarizing an option are identical every time the option is presented." *Some* experimental designs assume this, but by no means all do. One can allow for non-stationarity by (i) making valuations *conditional* on various theoretically or empirically motivated contingencies; (ii) controlling these contingencies in one's experiment (i.e., eliciting observations in all variants of the contingencies hypothesized to be relevant, then extracting the observed differences in difference through regression analyses). For example, it is becoming clear that people often trade off drastically more future utility for present rewards than they do future utility for rewards delivered very soon but not immediately. Various hypotheses have been developed to explain this. Whichever of them may be true, experimental economists have an increasingly well established method for taking this form of non-stationarity into account when estimating utility functions from data. (See Andersen, Harrison, Lau and Rutström, "Eliciting Risk and Time Preferences" *Econometrica*, 76(3), May 2008, 583-618.)

9. Corrado *et al* then consider what they call 'model-based approaches'. You should not infer from this that approaches relying on revealed preference have no models. But it is true that their models aren't what Corrado *et al* call 'mechanistic'. It is also true that revealed preference models aim to characterize average or aggregate choices rather than processes by which individuals make decisions. We now

consider approaches to the latter.

10. We begin by considering the modeling of *option evaluation mechanisms*, algorithms that take features of alternatives as inputs and send decision variables as outputs to be delivered to *action selection mechanisms*. The (very typical) problem used as an example by Corrado *et al* is that of comparatively evaluating two lotteries *A* and *B* in terms of their probabilities of payoff *p* and their dollar payoffs *x*:

$$(D_A, D_B = g(p_A, x_A, p_B, x_B)$$

If we assume that *A* and *B* are evaluated independently by separate mechanisms then this reduces to

$$D_i = f(p_i, x_i) \quad i \in \{A, B\}.$$

11. Our key choices of course concern *f*. We can try for a *parameter-free* function not derived from fitting to data; or use a *parametric* model in which a small number of free parameters are estimated from data; or resort to a *reconstructed* model in which we make few assumptions about the functional form of *f* and allow many free parameters we'll estimate from data. In general, the more free parameters we allow in a model the more independent observations we need in our data.

12. As examples, in modeling the hypothetical decision over lotteries above, we'd have a parameter-free model if we represented our subject as simply multiplying each probability by its associated dollar amount to derive expected values; we'd have a parametric model if we represented our subject as mapping expected values into expected utilities described by a utility function with standard properties (with, say, marginal utility decreasing with wealth, and/or some risk

aversion); we'd have a reconstructed model if we thought that we needed to represent our subject as maximizing an idiosyncratic utility function we needed to derive from that subject's specific behavior.

13. Economists most often use parametric models, because reconstructed ones usually generalize poorly to aggregate data. We might expect neuroeconomists to turn more often to reconstructed models, to the extent that their target is individual behavior. On the other hand, neuroeconomists often have small data sets, because fMRI data are expensive. Thus they too most often use parametric models.

14. Once we've specified an option evaluation mechanism, we must complete the model with an action evaluation mechanism that takes us from decision variables to behavioral choice. Action evaluation mechanisms can be either *deterministic* or *probabilistic*. The most common deterministic mechanism in economics incorporates the assumption that agents choose whichever options maximize their utility. This isn't likely to yield reliable predictions in modeling an individual, however. Then one might more typically assume that the probability of an option's being chosen increases with its ranking in the agent's utility function. A common example is the *softmax* function:

$$pchose(i) = \frac{\exp\{\kappa D_i\}}{\sum_j \exp\{\kappa D_j\}}$$

where *i* and *j* are the options and κ is the steepness of the relationship between utility and probability.

To complete this model, we'd then plug our option evaluation mechanism in for *D*. Suppose we'd chosen the parametric model

$D_A : EU_A = p_A U(x_A)$

$U(x) \approx x^{\alpha}$

(so, $D_A : p_A x_A^{\alpha}$)

Then

$$pchose(i) = \frac{\exp\{\kappa p_i x_i^{\alpha}\}}{\sum_j \exp\{\kappa p_j x_j^{\alpha}}$$

Then two parameters must be estimated from data: α, the concavity of the subject's utility function, and κ, the steepness of the decision function.

15. How does one go about estimating parameters? One finds values for them that optimize an *objective function* measuring goodness of fit between some decisions made by the subject and decisions predicted by the model. *Maximum likelihood* (ML) is the most common such method in neuroeconomics. You can find this summarized at http://mathworld.wolfram.com/MaximumLikelihood.html, and if you want to learn it, free software is available for download.

16. These same methods are used to assess a model's accuracy in predicting choices. It is, however, result cooking to use the same data for both parameter estimation and model assessment. Thus one needs enough data to be able to use a different subset of it for each of these tasks. (The subsets should be chosen using a procedure that isn't pre-rigged. Be especially watchful for this in the neuroeconomics literature!) In neuroeconomics, one standardly gets a large enough dataset not by scanning many subjects, but by gathering many observations from each of a small group of subjects.

17. What is usually optimized is the log-likelihood function

$$LogL = \sum_t \log pchose_t(x_t)$$

where $pchose_t(x)$ is the model's predicted probability that the subject will choose option x on trial t, and x_t is the option chosen by the subject on t.

18. How accurate must a model's predictions be that we regard it as confirmed? What matters is the difference between our ability to generate the same prediction accuracy with and without the model. Some leading experiments in social neuroscience have recently been heavily criticized for using techniques of deriving predictions that grossly inflate this difference; see http://www.pashler.com/Articles/Vul_etal_2008inpress.pdf Some data we've reviewed in this course are now being re-checked in light of this critique; it will be interesting to see which results survive this exercise.

19. Here's a complication. Often a subject's decision at a time t_2 is partly a function of the subject's decision at t_1. This is rational if the relevant state of the environment at t_2 is a function of its state at t_1. Thus this relationship is characteristic of learning mechanisms and processes. However, this relationship in principle diminishes the capacity of a model of such a process to generate informative predictions. In this circumstance, a better way to test the model is to have it *generate* simulated data, and then compare these with real data. Since the brain is a learning system through and through, this is an approach we should expect to see more and more often from methodologically improved neuroeconomics.

20. Neuroeconomists seek neural correlates for decision variables, ultimately hoping to turn the unobservable observable. Corrado *et al* review three studies.

21. In the first study, Corrado *et al* began with a model of rhesus monkey foraging based on Herrnstein's *matching law*. According to this, animals allocate responses to two choice alternatives proportionally to the relative frequencies of reinforcement received from them. The empirical relation $B_1/B_2 = R_1/R_2$, where the Bs and Rs refer to responses allocated to and reinforcements received from the two alternatives, respectively, became known as the matching law because relative behavioral allocation matched relative reinforcement received. Herrnstein later showed that a choice rule that directly compares the local rates of reinforcement on the two alternatives and selects the alternative with the higher local rate describes the process that produces matching, which he termed *melioration*. Subsequent work building on these basic relations showed that we get accurate predictions of human and animal choice behavior if we model them as maximizing *local* rather than *overall expected* utility. The family of models that technically represent melioration specify that a person or animal will distribute their resources to a range of activities during a time interval in exact proportion to the value of the rewards they experience for each activity *in that time interval*. The interval over which local rates of reinforcement are computed and compared may vary. If variation in the interval over which rates are computed changes the estimate of those rates, this changes the valuation of the choice options and leads to a change in behavioral allocation. Thus the standard microeconomic assumption of choice that reveals consistent preferences may be violated.

22. The authors rather vaguely report that their model predicted firing rates in a part of area LIP. They say that "the map of option value in LIP was in spatial coordinates". I *assume* they mean by this that the part of area LIP in which firing rates were predicted by the model – though they don't

quite say that the model *did* predict firing rates – was a part that had previously been associated with either spatial perception or projection of action into space. At least, I don't know what else they could intend the claim to mean. They then say that they "used the large datasets collected in the experiment" to build a reconstructed model that "no longer bore any formal resemblance to the original matching law." They do not quite say that the large datasets in question were data on area LIP (or indeed were neural data at all). It is, in general, hard to entirely understand the report of this experiment in the text.

23. The next reported experiment is clearer. Knutson *et al* used the *temporal difference learning algorithm*, which is widely used as a model of learning in the dopamine reward circuit, to try to predict differences in activity levels in different parts of that circuit while human subjects chose among slot machines with varying payout rates and prize magnitudes.

24. Here is the TD algorithm:

$$V^{*}(s_t) = E[r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \gamma^3 r_{t+3} + \ldots]$$

which we close by writing

$$V^{*}(s_t) = E[r_t + \gamma\ V^{*}(s_{t+1})].$$

This describes the procedure by which the reward-system learning algorithm continuously inputs new information to keep refining its estimate of V^{*} to get a particular stream of actual temporal valuations V. From this we can define a measure δ of the extent to which the value estimates of two successive states and a reward experienced by the system are consistent with one another:

$$\delta(t) = r_t + \gamma\, V(s_{t+1}) - V(s_t),$$

where δ is an error signal that pushes $V(s)$ towards better estimates as it gets more data. If $V(s_{t+1})$ turns out to be better than expected, then $\delta(t)$ will be positive, thus indicating that $V(s_t)$ needs to be adjusted upward. If $V(s_{t+1})$ turns out to be worse than expected, $\delta(t)$ will be negative and $V(s_t)$ will be adjusted downwards. If $\delta(t) = 0$ then of course no learning occurs.

25. They suggest on this basis that vmPFC processing computes relative value, while striatal processing codes for errors in reward prediction. They then used the model's predictions to divide subjects' choices into two classes: exploitative (choosing the machine assigned the higher value by the model) and exploratory (choosing a lesser valued machine). They then found two areas – ACC and PFC – that were more active during exploratory decisions. No quantitative differences in magnitudes are mentioned.

26. In the third experiment, Hampton *et al* compared two learning models, a Bayesian Markov one and a simple RL one, on a task in which the former outperformed the latter. They found that reward circuit activity better matched the more sophisticated model, which in turn better predicted behavior. No quantitative differences in magnitudes are mentioned.

27. The article concludes by describing a new technique for neural investigation known as *optogenetics*. In this approach, genes (channel proteins) are selectively inserted into circuits that boost or lower the base rate of neural excitation in the circuit in question. This promises to allow for true controlled

experiments, as opposed to just observations across varying conditions, to which present technology restricts us.