

Stochastic Choice: An Optimizing Neuroeconomic Model

Michael Woodford*

January 28, 2014

Abstract

A model is proposed in which stochastic choice results from noise in cognitive processing rather than random variation in preferences. The mental process used to make a choice is nonetheless optimal, subject to a constraint on available information-processing capacity that is partially motivated by neurophysiological evidence. The optimal information-constrained model is found to offer a better fit to experimental data on choice frequencies and reaction times than either a purely mechanical process model of choice (the drift-diffusion model) or an optimizing model with fewer constraints on feasible choice processes (the rational inattention model).

Experimental studies often find that choice has an apparently random element (e.g., Hey, 1995; Ballinger and Wilcox, 1997; Cheremukhin, Popova and Tutino, 2011). It is plausible, especially in light of the similarity between the stochasticity of choice behavior and the stochastic responses observed in perceptual discrimination tasks (e.g., Green and Swets, 1966), to think of this as reflecting noise in the cognitive processing of the choice situation, rather than variation in the subject's genuine preferences. Such a theory, however, makes specific predictions, even of a probabilistic nature, only under a specific hypothesis about the processing errors.

Here we offer such a theory, in which the mental process used to make a choice is optimal, in the sense of maximizing expected utility, subject to a constraint on the information-processing capacity of the neural pathways that supply information about the current choice situation. In proposing such an information constraint, the model is in the spirit of Sims' (2010) theory of "rational inattention." However, unlike Sims' theory, the model proposed here imposes additional constraints on the class of choice

*Department of Economics, Columbia University, 420 W. 118th Street, New York, NY 10027 (email: mw2230@columbia.edu). I thank Ian Krajbich for sharing his data, Ian Krajbich, Stephen Morris, Antonio Rangel, and Michael Shadlen for helpful discussions, Stéphane Dupraz for research assistance, and the Institute for New Economic Thinking for supporting this research.

algorithms that are assumed to be feasible. These are motivated by both behavioral and neurophysiological data suggesting that decisions between discrete alternatives are made using a form of “accumulation-to-bound” algorithm.¹

1 Accumulation-to-Bound Algorithms

Suppose that a decisionmaker must choose among a discrete set of alternatives a , where the reward $U(a, x)$ for choosing a depends also on a state x . Suppose further that the decision must be made by a mechanism with two elements. First, there is a *sensor* that produces a signal s about the state (i.e., the nature of the current choice situation) each time it is employed. And second, there is a *decoder* that decides, each time another signal is received from the sensor, either to halt the process and announce a decision a , or to engage the sensor another time. The actions of the decoder cannot depend on the state x other than through their dependence on the history of signals received from the sensor.

The “drift-diffusion model” (DDM) of mathematical psychology is an example of an algorithm of this form. The decision problem is one with two possible choices, which we may call L (left) and R (right), with values $U_L(x)$ and $U_R(x)$ respectively, depending on the state. With each additional use of the sensor, one of two possible signals is received, which we may also call L and R , as each signal provides additional evidence in favor of one of the actions or the other.² The probability $\lambda(x)$ of an R signal is assumed to be given by

$$\ln \frac{\lambda(x)}{1 - \lambda(x)} = \alpha [U_R(x) - U_L(x)] \quad (1.1)$$

each time, independent of the signal history, for some $\alpha > 0$.³ The decoder continues to collect additional signals from the sensor as long as n , the number of R signals minus the number of L signals received, has an absolute value less than some threshold N ; it halts as soon as n reaches the value N (in which case it chooses R) or $-N$ (in which case it chooses L).

Algorithms of this kind have been shown to be able to account for the way in which both the probability of choosing one alternative over the other and the average time required to make a decision vary with changes in the relative value of the two

¹See Shadlen *et al.* (2007) for a review of evidence for such processes in the context of motion perception, and Fehr and Rangel (2011) for a review of applications of similar models to value-based choice.

²In the neural implementation of such a mechanism for a motion-discrimination task discussed by Shadlen *et al.* (2007), these signals are additional spikes generated either by “left-preferring” or “right-preferring” neurons.

³In perceptual applications, the log odds are assumed to be a linear function of some objective characteristic of the presented stimulus, such as the “motion strength” in Shadlen *et al.* (2007), rather than a value difference.

options.⁴ In the context of motion perception, Shadlen *et al.* (2007) argue that such an algorithm describes neural mechanisms involved in perceptual judgments; both the signals s produced by the sensor and the evolving state variable n of the decoder can be identified with the firing rates of particular populations of neurons in the brain.

2 The Optimal Sensor Problem

Here we restrict attention to algorithms in which the set of possible signals and the operation of the decoder are as in the DDM, but let the probability of receiving a signal R when the sensor is employed be an arbitrary function $\lambda(h, x)$ of the state x and the signal history h . Our hypothesis is that the sensor is chosen so as to maximize the expected reward $E[U]$, where the expectation is taken both over possible choice situations and possible signal histories in a particular situation, subject to an upper bound on the information-processing capacity required by the sensor. As in Sims (2010), we measure required processing capacity by the *mutual information* between the input state (x, h) and the output signal s , an information-theoretic measure of the degree to which observation of s is informative about the current input state, or alternatively, of the degree to which knowledge of the input state reduces uncertainty about the signal that will be sent (Cover and Thomas, 2006).

Calculation of both the expected reward and the mutual information implied by a given algorithm depends on a prior over possible choice situations. We assume the algorithm is optimized for an environment in which a choice problem specified by values (U_L, U_R) is initially drawn from some distribution π ; then, each time the sensor is employed, if a decision is not made by the algorithm, an immediate action is forced by the environment with probability $1 - \rho$, while with probability ρ , the algorithm is allowed to continue, collecting another signal from the sensor and having another opportunity to choose. In the event of a forced action, it is assumed that each action will be taken with equal probability; this implies a loss (on average) from delaying decision. We can without loss of generality restrict attention to choice problems in which $(U_L, U_R) = (-V, V)$ for some real number V , and a reward of zero is obtained if action is forced.

We further assume that the prior π is symmetric, in the sense that $V = -v$ has the same probability as $V = v$, for any real number v . In this case, Woodford (2014) shows that the optimal operation of the sensor in a choice situation with a particular value of V can be determined as a function of V , the persistence parameter ρ , and the Lagrange multiplier $\theta > 0$ on the information constraint, independently of the probability assigned by the prior to any other values of V .⁵ This allows us to obtain

⁴See, e.g., Figure 1 of Shadlen *et al.* (2007) for the fit of this model to experimental data on motion discrimination, and Figure 1 below for its fit to experimental data on choice between pairs of food items. See also Clithero and Rangel (2013) and Krajbich, Oud and Fehr (2014) for further comparisons with discrete choice data.

⁵Of course, the Lagrange multiplier θ depends on the entire distribution π ; but the way in which

testable predictions about choices and response times for the particular values of V that are experimentally observed, without requiring any assumption (beyond the symmetry assumption) about the possible values of V (under the prior for which the algorithm is optimally adapted) that are not observed.

Conditional on a particular value of V , the optimal sensor problem reduces to the choice of a function $\lambda(h)$ defined for each possible finite signal history h such that $-N < n < N$, so as to maximize

$$\mathbb{E} \left[U - \theta \sum_{t=1}^T D(\lambda(h_t)) \right], \quad (2.2)$$

where the expectation is now only over possible signal histories, T is the random number of uses of the sensor until either a decision is made or action is forced, h_t is the signal history of length $t - 1$ prior to the t th use of the sensor, and

$$D(\lambda) \equiv \lambda \ln \lambda + (1 - \lambda) \ln(1 - \lambda) + \ln 2$$

is a relative-entropy measure of the difference between the conditional signal probabilities $(\lambda, 1 - \lambda)$ and the unconditional probabilities $(1/2, 1/2)$.

The optimal probability of receiving a signal R from the sensor can then be shown to be a function $\lambda(n)$ of the excess of R signals over L signals already received, given by

$$\ln \frac{\lambda(n)}{1 - \lambda(n)} = \frac{\rho}{\theta} [W(n + 1) - W(n - 1)], \quad (2.3)$$

where $W(n)$ is the value function for the problem (2.2). The value function can in turn be characterized by a difference equation of the form

$$W(n) = \Phi(W(n - 1), W(n + 1))$$

for each $-N < n < N$, together with boundary conditions

$$W(-N) = -V, \quad W(N) = V. \quad (2.4)$$

This implies random signals (and a corresponding random path of the counter n) similar to those postulated by the DDM. Indeed, if one replaces $W(n + 1) - W(n - 1)$ in (2.3) by the *average* gradient of the function W implied by (2.4), one obtains a logistic relation between the signal frequency and the relative value of the options of exactly the kind specified in (1.1). The model is not equivalent to the DDM, however, in that the value of $\lambda(n)$ implied by (2.3) varies with n .

π matters can be summarized by the single number θ , which also reflects the degree of tightness of the information constraint.

3 Explaining Logistic Choice

The model can be used to predict the probability of choosing each of the two options, as functions of the relative value of the two options to the DM. In the special case $\rho = 1$, Woodford (2014) shows that

$$\text{Prob}(R) = \frac{e^{(U_R - U_L)/\theta}}{1 + e^{(U_R - U_L)/\theta}}. \quad (3.5)$$

in each possible state. This kind of logistic relationship is very commonly fit to data on binary choices (both from laboratory experiments and the field); if the difference in the values of the two alternatives is assumed to be a linear function of some vector of measured characteristics, one obtains the familiar logistic regression model.

The standard interpretation given to this statistical specification is in terms of a random-utility model (RUM), in which R is chosen over L if and only if $U_R + \epsilon_R > U_L + \epsilon_L$, where ϵ_L, ϵ_R are two independent draws from an extreme value distribution of type I (McFadden, 1974). But while the RUM provides a possible justification for the econometric practice, there is no obvious reason to expect that the additive random terms in people’s valuations (even supposing that they vary randomly from each occasion of choice to the next) should be drawn from this particular type of distribution, so that there is little reason to expect logistic regressions to be correctly specified, under this interpretation.

The DDM provides an alternative interpretation of logistic choice. If the signal frequencies are given by (1.1), one can also derive a relation of the form (3.5), in which $\theta = 1/(\alpha N)$. But the prediction of a logistic relationship still depends on the particular functional form (1.1), which has no motivation deeper than analytical convenience. The optimal information-constrained model (OICM) instead provides an explanation for the empirical fit of a logistic relationship (or something close to it, in any case with ρ close enough to 1) that does not depend on an *a priori* assumption of either a special probability distribution or any special functional forms (apart from the information-theoretic measure of the information required by different sensors).

Logistic choice is also predicted by the theory of rational inattention (RI), without any need for restriction to the class of accumulation-to-bound algorithms considered here. As formulated by Sims (2010), RI postulates that decisions are made on the basis of a random signal (or sequence of signals) so as to maximize expected reward, subject to an upper bound on the mutual information between the external state and the sequence of signals; no further restrictions on the space of possible signals or the ways in which the decision may depend on the signal are imposed. When applied to a binary decision problem, this hypothesis also implies choice frequencies given by (3.5), as shown by Woodford (2008), Cheremukhin, Popova, and Tutino (2011), and Matějka and McKay (2013), and so might seem a more parsimonious explanation for logistic choice. However, the unconstrained RI hypothesis fails to explain the observed relation between the relative value of the available options and the time

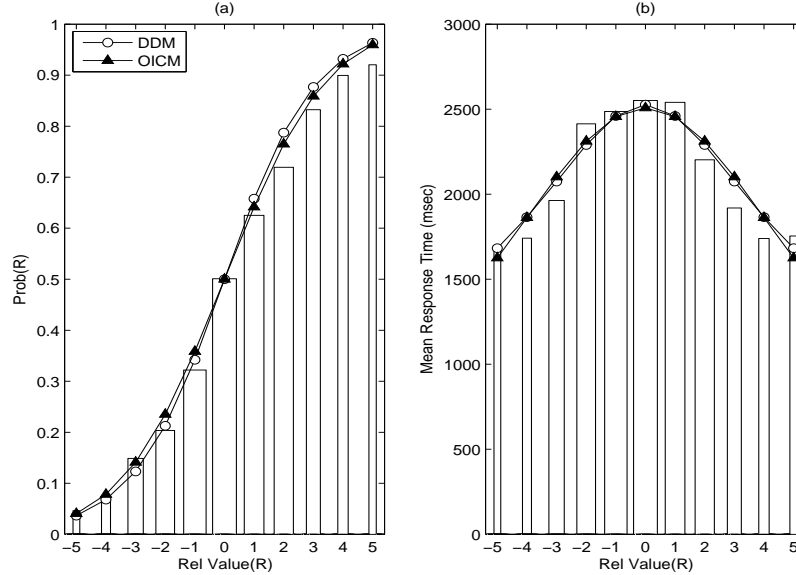


Figure 1: Data and theoretical predictions for (a) probability of choosing R and (b) mean time required to make a decision, each plotted as a function of the subject's relative valuation of option R . Data source: Krajbich, Armel and Rangel (2010).

required for decision, as discussed next.⁶

4 Explaining Variation in Response Times

The OICM can also be used to predict the average time required for a choice to be made, under the assumption that each additional use of the sensor to produce another signal requires a fixed time interval $\Delta > 0$. Here we consider the consistency of these predictions with data from the experiment of Krajbich, Armel and Rangel (2010, hereafter KAR), a leading example of application of a version of the DDM to the explanation of value-based choice. KAR ask subjects to choose between pairs of food items that they have previously ranked on an integer scale in terms of how much they would like to eat the item. After making 100 such decisions, the subject is allowed to actually eat the item chosen in one of the binary choices, selected at random.

In each panel of Figure 1, trials are grouped according to the extent to which the right option was ranked higher than the left option by that particular subject. Panel (a) shows the fraction of trials on which the option R was chosen, as a function of the relative rank of option R ; panel (b) shows the average time taken to decide (in

⁶On the value of models that jointly predict choice and response time, even for economists interested only in predicting choice, see for example Clithero and Rangel (2013) and Krajbich, Oud and Fehr (2014).

milliseconds), again as a function of the relative value. In each panel, the height of the bar indicates the value of the statistic for a given bin, while the width of the bar is proportional to the number of trials assigned to that bin.

The open circles plot the predictions of the DDM for these two statistics, if the relative reward $U_R - U_L$ is assumed to be proportional to the subject's reported relative ranking of the two options, and the parameters of the model are chosen to minimize a weighted sum of squared prediction errors.⁷ In fitting the DDM to the data, measured response time is assumed to equal $A + T\Delta$, where T is the number of signals received before the decision, and A is a positive constant. Hence there are three free parameters, α , Δ , and A . The black triangles plot the corresponding predictions of the OICM, if the parameters are chosen to minimize the same criterion. Arbitrarily fixing $N = 100$,⁸ the best-fitting parameter values are found to be $\rho = 0.9996$, $\theta = 2.08$, and $\Delta = 0.25$ msec.⁹

As discussed above, the OICM, like the DDM, is able to account for the logistic relationship between the value difference and choice frequency; panel (b) shows that it is also able to do nearly as well in accounting quantitatively for the degree to which decisions take longer, on average, when the value difference is smaller. In this respect, the OICM is superior to the simple (unconstrained) RI hypothesis. If the response time T and the action a are both outcomes of the decision process, then the information about the external state used by the process must be at least the mutual information between the pair (a, T) and the state.

Because only a is payoff-relevant, RI would require that T contain no additional information about the state beyond that revealed by the decision a ; hence the probability distribution of possible response times, conditional upon a given choice (say, R), would have to be independent of the values of the options on that trial. This would require average response time (plotted in Figure 1(b)) to be an affine function of the probability of choosing R (plotted in Figure 1(a)), which is clearly not the case. Under the further symmetry postulate (consistent with Figure 1(b)) that average response time should be the same for a given pair of goods, regardless of which is offered on the left and which on the right, RI would require average response time to be independent of the values of the options, again clearly not the case.¹⁰ Hence the additional constraint proposed here is required in order to simultaneously explain both panels of Figure 1.

While the OICM and DDM fit the statistics shown in Figure 1 to a roughly similar degree, the models are not equivalent. The OICM predicts that the probability of receiving a signal R depends on the signal history (specifically, the value of n) as

⁷See Woodford (2014) for details.

⁸As discussed in Woodford (2014), the model's predictions are nearly independent of the value of N , for any large enough value of N , once the numerical values of ρ and Δ are suitably adjusted.

⁹No allowance is made in this case for a fixed time requirement A in addition to the time $T\Delta$ used by the algorithm, so that again three free parameters are used to fit the data.

¹⁰See Woodford (2014) for further discussion.

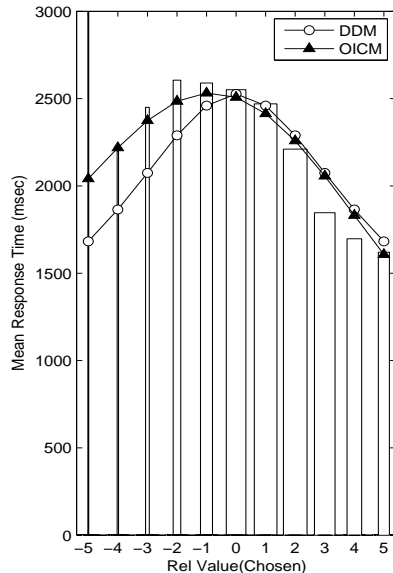


Figure 2: Data and theoretical predictions for mean response time, now plotted as a function of the relative valuation of the option chosen. Data source: Krajbich, Armel and Rangel (2010).

well as upon the relative value of the two options. One of the more noteworthy implications is shown in Figure 2, where the mean response time is plotted not as a function of the relative value assigned to the R option, but rather as a function of the relative value of whichever option was actually chosen by the subject in that particular trial.

The predictions of the DDM are the same in Figure 2 as in Figure 1(b), for this model predicts that the mean response time should be the same regardless of whether the lower-valued or higher-valued option is chosen. The data, instead, show that “correct” choices (cases in which the higher-valued option is chosen) are on average made more quickly than “incorrect” choices, for a given degree of difficulty of the choice (based on the absolute difference in rankings). This is in fact a common feature of experimental data on response times in binary perceptual classification tasks as well, and is an often-noted empirical failing of the DDM (Luce, 1986). The OICM instead correctly predicts that correct choices should be made more quickly, and by roughly the amount by which the mean response times are different in the data.

The quantitative predictions of the OICM regarding response times are thus more accurate than those of either simple (unconstrained) RI theory or the classic (constant-drift) version of the DDM. This suggests that it may indeed be useful to view the stochastic processing errors responsible for stochastic choice as a property of an optimal mechanism, subject to a constraint on information-processing capacity. However, it appears that an empirically adequate model will have to incorporate

additional constraints, reflecting the necessity of sequential information processing, that are not part of the classic Sims (2010) formulation of RI theory. The additional measurements made by neuroeconomists should make an important contribution to better understanding the correct specification of these constraints.

References

- [1] **Ballinger, T. Parker, and Nathaniel T. Wilcox.** 1997. “Decisions, Error and Heterogeneity.” *Economic Journal* 107: 1090-1105.
- [2] **Cheremukhin, Anton, Anna Popova, and Antonella Tutino.** 2011. “Experimental Evidence on Rational Inattention.” Federal Reserve Bank of Dallas Research Department Working Paper 1112.
- [3] **Clithero, John A., and Antonio Rangel.** 2013. “Combining Response Times and Choice Data Using a Neuroeconomic Model of the Decision Process Improves Out-of-Sample Predictions.” Unpublished.
- [4] **Cover, Thomas M., and Joy A. Thomas.** 2006. *Elements of Information Theory*. New York: Wiley-Interscience, 2d ed.
- [5] **Fehr, Ernst, and Antonio Rangel.** 2011. “Neuroeconomic Foundations of Economic Choice — Recent Advances.” *Journal of Economic Perspectives* 25(4): 3-30.
- [6] **Green, David M., and John A. Swets.** 1966. *Signal Detection Theory and Psychophysics*. New York: Wiley.
- [7] **Hey, John D.** 1995. “Experimental Investigations of Errors in Decision Making under Risk.” *European Economic Review* 39: 633-640.
- [8] **Krajbich, Ian, Carrie Armel, and Antonio Rangel.** 2010. “Visual Fixations and the Computation and Comparison of Value in Goal-Directed Choice.” *Nature Neuroscience* 13: 1292-1298.
- [9] **Krajbich, Ian, Bastiaan Oud, and Ernst Fehr.** 2014. “Benefits of Neuroeconomics Modeling: New Policy Interventions and Predictors of Preference.” *American Economic Review* xx: xxxx.
- [10] **Luce, R. Duncan.** 1986. *Response Times*. New York: Oxford University Press.
- [11] **Matějka, Filip, and Alisdair McKay.** 2013. “Rational Inattention to Discrete Choices: A New Foundation for the Multinomial Logit Model.” Unpublished.

- [12] **McFadden, Daniel.** 1974. “Conditional Logit Analysis of Qualitative Choice Behavior.” In *Frontiers in Econometrics*, ed. Paul Zarembka, 105-142. New York: Academic Press.
- [13] **Shadlen, Michael N., Timothy D. Hanks, Anne K. Churchland, Roozbeh Kiani, and Tianming Yang.** 2007. “The Speed and Accuracy of a Simple Perceptual Decision: A Mathematical Primer.” In *Bayesian Brain: Probabilistic Approaches to Neural Coding*, ed. Kenji Doya, Shin Ishii, Alexandre Pouget, and Rajesh P.N. Rao, 209-237. Cambridge, MA: MIT Press.
- [14] **Sims, Christopher A.** 2010. “Rational Inattention and Monetary Economics.” in B.M. Friedman and M. Woodford, eds., In *Handbook of Monetary Economics*, ed. Benjamin M. Friedman and Michael Woodford, vol. 3A, 155-181. Amsterdam: Elsevier.
- [15] **Woodford, Michael.** 2008. “Inattentiveness as a Source of Randomized Discrete Adjustment.” Unpublished.
- [16] **Woodford, Michael.** 2014. “An Optimizing Neuroeconomic Model of Discrete Choice.” NBER Working Paper.