

Speed, Accuracy, and the Optimal Timing of Choices*

Drew Fudenberg[†] Philipp Strack[‡] Tomasz Strzalecki[§]

This version: March 20, 2017, First version: May 6, 2015

Abstract

We model the joint distribution of choice probabilities and decision times in binary decisions as the solution to a problem of optimal sequential sampling, where the agent is uncertain of the utility of each action and pays a constant cost per unit time for gathering information. We show that choices are more likely to be correct when the agent chooses to decide quickly provided that the agent’s prior beliefs are correct. This better matches the observed correlation between decision time and choice probability than does the classical drift-diffusion model (DDM), where the agent knows the utility difference between the choices.

*This paper was submitted to the AER on June 8, 2015 under the title “Stochastic Choice and Optimal Sequential Sampling.” We thank Stefano DellaVigna and Antonio Rangel for very stimulating conversations, a co-editor, three anonymous referees, as well as Jerome Busemeyer, Jan Drugowitsch, Arash Khodadadi, Josh Tenenbaum and Gonzalo Cisternas for helpful comments, Ian Krajbich, Carriel Armel and Antonio Rangel for sharing their experimental data with us, Kevin Kainan Li, In Young Cho and Jonathan Libgober for expert research assistance, and the Sloan Foundation and NSF grants SES-0954162, SES-1123729, SES-1258665 and CAREER grant SES-1255062 for support.

[†]Department of Economics, MIT

[‡]Department of Economics, University of California, Berkeley

[§]Department of Economics, Harvard University

1 Introduction

In laboratory experiments where individuals are repeatedly faced with the same choice set, the observed choices are stochastic—individuals don’t always choose the same item from a given choice set, even when the choices are made only a few minutes apart.¹ In addition, individuals don’t always take the same amount of time to make a given decision—response times are stochastic as well. Our goal here is to model the joint distribution of choice probabilities and decision times in choice tasks, which we call a *choice process*.

We restrict attention to the binary choice tasks that have been used in most neuroscience choice experiments, and suppose that the agent is choosing between two items that we call left (l) and right (r). In this setting, we can ask how the probability of the correct choice varies with the time taken to make the decision. If the agent is learning during the decision process, and is stopped by the experimenter at an exogenous time, we would expect the data to display a negative correlation between speed and accuracy, in the sense that the agent makes more accurate decisions when given more time to decide. However, in many choice experiments the agent chooses when to make their decision, and there is instead the opposite correlation: accuracy decreases with decision time in the sense that slower decisions are less likely to be correct (Swensson, 1972; Luce, 1986; Ratcliff and McKoon, 2008).²

To explain this, we develop a new variant of the drift diffusion model (DDM); other versions of the DDM have been extensively applied to choice processes in the neuroscience and psychology literatures.³ The specification of a DDM begins with a diffusion process Z_t that represents information the agent is receiving over time, and two disjoint stopping regions S_t^l and S_t^r . The agent stops at time t if $Z_t \in S_t^l$ (in which case she chooses l) or $Z_t \in S_t^r$ (in which case she chooses r); otherwise the agent continues. Because the evolution of the diffusion depends on which choice is better, the model predicts a joint probability

¹See Hey (1995, 2001), Camerer (1989), Ballinger and Wilcox (1997), Cheremukhin, Popova and Tutino (2011).

²The choice experiments we reference elicit ordinal rankings of all of the items from the subjects before having them make a series of binary choices, and identify “correct” with “more highly ranked.” When these ordinal rankings are not available, we would follow the revealed preference literature and say that the more-often-chosen item is the correct choice. This is consistent with the solution to the uncertain-difference DDM that we define in this paper and also the perturbed utility representation of stochastic choice in Fudenberg, Iijima and Strzalecki (2015).

³The DDM was first proposed as a model of choice processes in perception tasks, where the subjects are asked to correctly identify visual or auditory stimuli. (For recent reviews see Ratcliff and McKoon (2008) and Shadlen et al. (2006).) DDM-style models have also been applied to choice experiments, where subjects are choosing from a set of consumption goods presented to them: Roe, Busemeyer and Townsend (2001), Clithero and Rangel (2013); Krajbich, Armel and Rangel (2010); Krajbich and Rangel (2011); Krajbich et al. (2012); Milosavljevic et al. (2010); Reutskaja et al. (2011).

distribution on choices and response times conditional on the true state of the world, which is unknown to the agent.

Our first set of results relate the slope of an arbitrary boundary to the correlation between speed and accuracy. We then turn to our main focus, which is to provide learning-theoretic foundations for a form of DDM where the agent’s behavior is the solution to a sequential sampling problem with a constant cost per unit of time as in the two-state model used to provide a foundation for the simple DDM, but with a different prior. In the *uncertain-difference DDM*, the agent believes that the utilities $\theta = (\theta^l, \theta^r)$ of the two choices are independent and normally distributed; this allows her to learn not only which alternative is better, but also by how much. In this model an agent with a large sample and Z_t close to zero will decide the utility difference is likely to be small, and so be more eager to stop than an agent with the same Z_t but a small sample, who is less certain of her point estimates and so has a higher option value of continuing to sample.

Our main insight is that the specification of the agent’s prior is an important determinant of the optimal stopping rule and thus of whether accuracy increases or decreases with stopping time. In particular, we show that in the uncertain-difference DDM it is optimal to have the range of Z_t for which the agent continues to sample collapse to 0 as time goes to infinity, and moreover that it does so asymptotically at rate $1/t$.⁴ Using the optimality of the stopping rule, we show that the optimal boundary is pointwise non-increasing in the cost so that increasing cost decreases both accuracy and decision time. We also show that an analyst who is uncertain of the agent’s true utilities and so aggregates data across many utility realizations will see that accuracy declines with the time taken to make a decision. In addition, we show that when the agent can split its attention between the two alternatives, with the rate of learning proportional to the attention paid, it is optimal to pay equal attention to each alternative at each point in time, so that the solution is the same as in our original model, where the attention levels are implicitly required to be equal.

We then provide a functional form that approximates the boundary; the functional form fits very well numerically for large t , and lets us show that for large costs the boundary is

⁴The intuition that the boundary should converge to 0 has been put forward both as a heuristic in various related models and as a way to better fit the data (see, e.g., [Shadlen and Kiani, 2013](#)). In addition, [Chernoff \(1961\)](#) proves that boundaries collapse in a model we show is equivalent, and [Drugowitsch et al. \(2012\)](#) announces this result for a related decision problem with a known utility difference between the choices but unknown signal strength. Time-dependent stopping thresholds also arise if the cost or utility functions are time-dependent or if there is a fixed terminal date, see e.g. [Rapoport and Burkheimer \(1971\)](#) and [Drugowitsch et al. \(2012\)](#); [Frazier and Yu \(2007\)](#) extend this to the case of a deadline with a time-dependend hazard rate.

essentially constant. Finally, we investigate the consequences of allowing the flow cost to vary arbitrarily with time. Intuitively, if the cost decreases quickly enough, this might outweigh the diminishing effectiveness of learning and lead to an increasing boundary. We show that this intuition is correct, and more strongly that any stopping region at all can be rationalized by a suitable choice of a cost function. Thus optimal stopping on its own imposes essentially no restrictions on the observed choice process, and so it is compatible with the boundaries assumed in Ratcliff and McKoon (2008) and Shadlen and Kiani (2013). However, the force of the model derives from its joint assumptions about the evolution of beliefs and the cost function, and the cost functions needed to rationalize various specifications of the stopping boundary may or may not seem plausible in the relevant applications.

One motivation for modeling the joint distribution of decision times and choices is that the additional information provided by decision times can let us rule out theories that predict similar choice patterns but fail to match the data on this richer domain. In addition, it can lead to models that are closer to the underlying neural mechanisms,⁵ which may be why it leads to better out-of-sample predictions of choice probabilities (Clithero and Rangel, 2013). In other settings than the simple choice tasks we study here, decision times have been used to classify decisions as “automatic/instinctive/heuristic” or “cognitive/considered/reflective,” as in Rubinstein (2007), Rand, Greene and Nowak (2012), and Caplin and Martin (2014), though Krajbich et al. (2015) argue that declining accuracy makes these classifications suspect.⁶

The oldest and most commonly used version of the DDM (which we will refer to as *simple* DDM) specifies that the stopping regions are constant in time, i.e., $S_t^l = S^l$ and $S_t^r = S^r$, and that Z_t is a Brownian motion with drift equal to the difference in utilities of the items. This specification corresponds to the optimal decision rule for a Bayesian agent who believes that there are only two states of the world corresponding to whether action l or action r is optimal, pays a constant flow cost per unit of time, and at each point in time decides whether to continue gathering the information or to stop and take an action.⁷

The constant stopping regions of the simple DDM imply that the expected amount of time that an agent will gather information depends only on the current value of Z_t , and

⁵See Shadlen and Kiani (2013) and Bogacz et al. (2006) for discussions of how DDM-type models help explain the correlation between decision times and neurophysiological data such as neuronal firing rates.

⁶In addition, there is a literature that uses reaction times and other observables to understand behavior in games: Costa-Gomes, Crawford and Broseta (2001), Johnson et al. (2002), and Brocas et al. (2014).

⁷Wald (1947) stated and solved this as a hypothesis testing problem; Arrow, Blackwell and Girshick (1949) solved the corresponding Bayesian version. These models were brought to the psychology literature by Stone (1960) and Edwards (1965).

not on how much time the agent has already spent observing the signal process, and that the probability of the correct choice is independent of the distribution of stopping times.⁸ In contrast, in many psychological tasks (Churchland, Kiani and Shadlen, 2008; Ditterich, 2006) accuracy decreases with response time in the sense that reaction times tend to be higher when the agent makes the incorrect choice. For this reason, when the simple DDM is applied to choice data, it predicts response times that are too long for choices in which the stimulus is weak, or the utility difference between them is small. Various authors have extended the simple DDM to better match the data, by allowing more general processes Z_t or stopping regions S_t^i , see e.g., Laming (1968); Link and Heath (1975); Ratcliff (1978), and by allowing the signal process to be mean-reverting (the decision field theory of (Busemeyer and Townsend, 1992, 1993; Busemeyer and Johnson, 2004)). However, with the exceptions cited above, past work has left open the question of whether these generalizations correspond to any particular learning problem, and if so, what form those problems take.

Gabaix and Laibson (2005), Branco, Sun and Villas-Boas (2012) and Ke, Shen and Villas-Boas (2013) look at decisions derived from optimal stopping rules where the gains from sampling are exogenously specified as opposed to being derived from Bayesian updating, as they are here; neither paper examines the correlation between decision time and accuracy. Vul et al. (2014) studies the optimal predetermined sample size for an agent whose cost of time arises from the opportunity to make future decisions; they find that for a range of parameters the optimal sample size is one.

Natenzon (2013) and Lu (2016) study models with an exogenous stopping rule. They treat time as a parameter of the choice function, and not as an observable in its own right. The same is true of Caplin and Dean’s (2011) model of sequential search. Accumulator models such as (Vickers, 1970) specify an exogenous stopping rule; Webb (2013) shows that the distribution of choices induced by these models is consistent with random utility.

2 Choice Processes and DDMs

2.1 Observables

Let $A = \{l, r\}$ be the set of alternatives, which we will call left (l) and right (r). Let $T = [0, +\infty)$ be the set of decision times—the times at which the agent is observed to state a choice. The analyst observes a joint probability distribution $P \in \Delta(A \times T)$; we call this a

⁸Stone (1960) proved this independence directly for the simple DDM in discrete time. Our Theorem 1 shows that the independence is a consequence of the stopping boundaries being constant.

choice process. For simplicity we assume that P has full support so that in particular there is a positive probability of stopping in any interval $[t, t + dt]$, and conditional probabilities are well defined; our working paper (Fudenberg, Strack and Strzalecki, 2015) does not make this assumption. We will decompose P as $p^i(t)$ and $F(t)$, where $p^i(t)$ is probability of choosing $i \in A$ conditional on stopping at time t and $F(t) = P(A \times [0, t])$ is the cdf of the marginal distribution of decision times. It will also be useful to decompose P the other way, as P^i and $F^i(t)$, where $P^i = P(\{i\} \times T)$ is the overall probability of choosing $i \in A$ at any time, and $F^i(t) = \frac{P(\{i\} \times [0, t])}{P^i}$ is the cdf of time conditional on choosing $i \in A$.

It is easy to define accuracy in perceptual decision tasks, since in such settings the analyst knows which option is ‘correct.’ In choice tasks the agents’ preferences are subjective and may be unknown to the researcher; however, many experiments that measure decision time in choice tasks independently elicit the subjects’ rating of the items, and we will define “correct” in these tasks to mean picking the item that was given a higher score.⁹ We let $p(t)$ denote the probability of making the correct choice conditional on stopping at time t .¹⁰

Definition 1. There is *increasing accuracy* when p is monotone increasing, *decreasing accuracy* when p is monotone decreasing, and *constant accuracy* when p is constant.¹¹

2.2 DDM representations

DDM representations have been widely used in the psychology and neuroscience literatures (Ratcliff and McKoon, 2008; Shadlen et al., 2006; Fehr and Rangel, 2011). The two main ingredients of a DDM are the stimulus process Z_t and a boundary $b(t)$.

In the DDM representation, the stimulus process Z_t is a Brownian motion with drift δ and volatility $2\alpha^2$:

$$Z_t = \delta t + \alpha\sqrt{2} B_t, \tag{1}$$

⁹In those experiments the rating task is done before the choice tasks. In any given choice task agents are more likely to choose the item with the higher rating score, with the choice probability increasing in the score difference of the items. See, e.g., Krajbich, Armel and Rangel (2010), Milosavljevic et al. (2010), Krajbich et al. (2012).

¹⁰In cases where the correct choices are not observable, one can use the modal choice as the definition of a correct choice. Here we mean the average choice frequency, aggregated over all decision times. Our working paper (Fudenberg, Strack and Strzalecki, 2015) proceeds along these lines.

¹¹Our working paper (Fudenberg, Strack and Strzalecki, 2015) considers two additional, closely related, measures of how accuracy varies with time: (i) How the probability of making the modal choice conditional on stopping in the interval $[0, t]$ varies with t , and (ii) whether the distribution of stopping times for one choice is sooner than the other in the sense of first-order stochastic dominance.

where B_t is a standard Brownian motion, so in particular $Z_0 = 0$.¹² In early applications of DDM such as Ratcliff (1978), Z_t was not observed by the experimenter. In some recent applications of DDM to neuroscience, the analyst may observe signals that are correlated with Z_t ; for example the neural firing rates of both single neurons (Hanes and Schall, 1996) and populations of them (e.g., Ratcliff, Cherian and Segraves, 2003). In the later sections we interpret the process Z_t as a signal about the utility difference between the two alternatives.

Following the literature, we focus on symmetric boundaries so that a *boundary* is a function $b : \mathbb{R}_+ \rightarrow \mathbb{R}$, and assume that $\delta > 0$ if left is the correct choice and $\delta < 0$ if right is the correct choice. Define the hitting time τ

$$\tau = \inf\{t \geq 0 : |Z_t| \geq b(t)\}, \quad (2)$$

i.e., the first time the absolute value of the process Z_t hits the boundary.¹³ Let $P(\delta, \alpha, b) \in \Delta(A \times T)$ be the choice process induced by τ and a decision rule that chooses l if $Z_\tau = b(\tau)$ and r if $Z_\tau = -b(\tau)$.¹⁴

Definition 2. A choice process P has a *DDM representation* (δ, α, b) if $P = P(\delta, \alpha, b)$.

Simple DDMs are ones with constant boundaries. *Hitting time models* generalize DDM by not requiring that the signal process Z_t is Brownian. The assumption that the process Z_t is Brownian is an important one, as without it the model is vacuous.¹⁵

When the stopping time is given by an exogenous distribution that is independent of the signal process, the agent will have more information when it stops later, and so is more likely to make the modal decision. Endogenous stopping when the signal hits the boundary

¹²Recall that standard Brownian motion is a process that starts at time 0, has continuous sample paths, and has independent normally distributed increments, meaning that $B_{t+s} - B_t$ is distributed $N(0, s)$. We normalize the volatility by 2 to fit with our later interpretation of Z_t as the difference in two signals.

¹³There are boundaries for which there is a positive probability that $\tau = \infty$. We consider only boundaries that lead the agent to stop in finite time with probability 1. This property will be satisfied in any model where the stopping time comes from a statistical decision problem in which never stopping incurs an infinite cost and the value of full information is finite.

¹⁴Note that the parameter α can be removed here by setting $\alpha' = 1$, $\delta' = \delta/\alpha$, and $b' = b/\alpha$. By a similar argument, δ can be assumed to be $-1, 0$, or 1 . We nonetheless retain α and δ here as we will use them in the next section to distinguish between utility and signal strength.

¹⁵Fudenberg, Strack and Strzalecki (2015) show that any choice process P has a hitting time representation where the stochastic process Z_t is a pair of fully revealing Poisson signals with the appropriate time-varying arrival rates. Moreover, Jones and Dzhafarov (2014) show that any choice data can be matched realization-by-realization with a signal process that has linear paths and is deterministic in that trial but is allowed to vary in an arbitrary way from trial to trial. The most general diffusion model of Ratcliff (1978) also allows the starting point and drift to be variable. Its predictions are not vacuous since it retains the Gaussian structure, but the added degrees of freedom lack a natural interpretation.

generates a selection effect that pushes the other way, as when the agent stops early the signals were relatively strong and thus relatively informative. Just how strong this selection is depends on the slope of the stopping boundary; the next theorem shows that a constant boundary is the case where these two effects exactly balance out.

Theorem 1. *Suppose that P has a DDM representation (α, δ, b) . Then P displays increasing, decreasing, or constant accuracy if and only if $b(t)$ is increasing, decreasing, or constant respectively.*

The intuition behind the proof is as follows: Suppose that $\delta > 0$ (so the correct action is l) and that the process stopped at time t . The odds that a correct decision is made in this situation are

$$\frac{p^l(t)}{p^r(t)} = \frac{\mathbb{P}[Z_t = b(t) | \{\tau = t\} \cap \{|Z_t| = b(t)\}]}{\mathbb{P}[Z_t = -b(t) | \{\tau = t\} \cap \{|Z_t| = b(t)\}]}, \quad (3)$$

where $\{\tau = t\}$ is the event that the process Z has not crossed the barrier before time t . From Bayes rule and the formula for the density of the normal distribution

$$\frac{\mathbb{P}[Z_t = b(t) | \{|Z_t| = b(t)\}]}{\mathbb{P}[Z_t = -b(t) | \{|Z_t| = b(t)\}]} = \exp\left(\frac{\delta b(t)}{\alpha^2}\right), \quad (4)$$

which is a decreasing function of t whenever b is. Moreover, a symmetry argument using the Brownian bridge shows that the conditioning event $\{\tau = t\}$ does not matter as it enters the numerator and denominator in (3) in the same way.

One way to generate decreasing or increasing accuracy with a constant boundary is to allow random initial positions for the signal process, a stochastic delay before the first possible response, and/or random drift, as in (e.g. Ratcliff and Smith, 2004), but that model does not have a foundation in optimal learning theory. Section 3 derives the stopping boundary from optimal learning, and shows how this leads to decreasing accuracy.

2.3 Aggregate DDM representations

In many settings with human subjects, analysts will have few observations of a given individual facing exactly the same choice problem, and will need to aggregate data across problems to estimate choice probabilities. If the decisions in each trial are made by a DDM, and the drift δ varies from trial to trial, we say that the aggregate data is an average DDM.

Definition 3. Choice process P has an *average DDM representation* if $P = \int P(\delta, \alpha, b) d\mu(\delta)$,

where μ is normal with mean 0, and variance $2\sigma_0^2$.¹⁶

Averaging over trials with different drifts creates an additional selection effect beyond the one that occurs when the value of the drift is fixed: When the distribution over δ is also normal, and the boundary b does not increase too quickly, the selection effect implies that in trials where the agent chose quickly, the drift was probably larger than usual, so the agent received sharper signals and is more likely to be correct.

Proposition 1. *Suppose P has an average DDM representation and that $b(t) \cdot (\sigma_0^{-2} + t\alpha^{-2})^{-1/2}$ is non-increasing in t . Then for all $d > 0$ and $0 < t < t'$*

$$\mathbb{P}[|\delta| < d \mid \tau = t] < \mathbb{P}[|\delta| < d \mid \tau = t'] .$$

The next result shows that the combination of this selection effect and the one that comes from endogenous stopping is sufficient to imply that accuracy decreases with decision time even when the boundary is not decreasing, at least when the distribution of drifts is normally distributed.

Theorem 2. *Suppose that P has an average DDM representation $P(\mu_0, \alpha, b)$, where δ is normal with mean 0 and variance σ_0^2 . Then P displays increasing, decreasing, or constant accuracy if and only if $b(t) \cdot (\sigma_0^{-2} + t\alpha^{-2})^{-1/2}$ is increasing, decreasing, or constant in t respectively.*

3 Optimal Stopping

3.1 Statement of the model

Both the simple DDM used to explain data from perception tasks and our uncertainty-difference DDM are based on the idea of sequential learning and optimal stopping. As we will see, the models differ only in their prior, but this difference leads to substantially different predictions. In the learning model, the agent doesn't know the true utilities, $\theta = (\theta^l, \theta^r) \in \mathbb{R}^2$, but has a prior belief about them $\mu_0 \in \Delta(\mathbb{R}^2)$. The agent observes a two-dimensional signal $(Z_t^i)_{t \in \mathbb{R}_+}$ for $i \in \{l, r\}$ which as in the DDM has the form of a drift plus a Brownian motion; in the learning model we assume that the drift of each Z^i is equal to the corresponding state,

¹⁶We set the variance to be $2\sigma_0^2$ to match the case we consider later where the agent sees two independent signals.

so that

$$dZ_t^i = \theta^i dt + \alpha dB_t^i \quad (5)$$

where α is the noisiness of the signal and the processes $\{B_t^i\}$ are independent.¹⁷ The signals and prior lie in a probability space $(\Omega, \mathbb{P}, \{\mathcal{F}_t\}_{t \in \mathbb{R}_+})$, where the information \mathcal{F}_t that the agent observed up to time t is simply the paths $\{Z_s^i\}_{0 \leq s < t}$. We denote the agent's posterior belief about θ given this information by μ_t . Let $X_t^i = \mathbb{E}_{\mu_t} \theta^i = \mathbb{E}[\theta^i | \mathcal{F}_t]$ be the posterior mean for each $i = l, r$. As long as the agent delays the decision she has to pay flow cost, which for now we assume to be constant $c > 0$. (Section 3.6 explores the implications of time varying cost.) The agent's problem is to decide which option to take and at which time. Waiting longer will lead to more informed and thus better decisions, but also entails higher costs. What matters for this decision is the difference between the two utilities, so a sufficient statistic for the agent is

$$Z_t := Z_t^l - Z_t^r = (\theta^l - \theta^r)t + \alpha\sqrt{2}B_t,$$

where $B_t = \frac{1}{\sqrt{2}}(B_t^1 - B_t^2)$ is a Brownian Motion. Note that the signal is more informative (its drift is larger compared to the volatility) when $|\theta^l - \theta^r|$ is large, while close decisions generate a less informative signal.

When the agent stops, it is optimal to choose the option with the highest posterior expected value; thus, the value of stopping at time t is $\max_{i=l,r} X_t^i$. The agent decides optimally when to stop: she chooses a *stopping time* τ , i.e., a function $\tau : \Omega \rightarrow [0, +\infty]$ such that $\{\tau \leq t\} \in \mathcal{F}_t$ for all t ; let \mathcal{T} be the set of all stopping times. Hence, the problem of the agent at $t = 0$ can be stated as

$$\max_{\tau \in \mathcal{T}} \mathbb{E} \left[\max_{i=l,r} X_\tau^i - c\tau \right] \quad (6)$$

Before analyzing this maximization problem, we note that its solution is identical to a regret-minimization problem posed by Chernoff (1961), in which the agent's objective is to minimize the sum of his sampling cost $c\tau$ and his ex-post regret, which is the difference between the utility of the object chosen and the utility of the best choice. When the agent stops, he picks the object with the higher expected utility, so his expected regret for any stopping time τ is $\mathbb{E} \left[-\mathbf{1}_{\{x_\tau^l \geq x_\tau^r\}}(\theta^r - \theta^l)^+ - \mathbf{1}_{\{x_\tau^r > x_\tau^l\}}(\theta^l - \theta^r)^+ \right]$, and his objective function

¹⁷This process was also studied by Natenzon (2013) to study stochastic choice with exogenously forced stopping times; he allows utilities to be correlated, which can explain context effects.

¹⁸Following the literature, in cases where the optimum is not unique, we assume that the agent stops the first time she at least weakly prefers to do so. That is we select the *minimal* optimal stopping time.

is

$$\text{Ch}(\tau) := \mathbb{E} \left[-\mathbf{1}_{\{x_\tau^l \geq x_\tau^r\}}(\theta^r - \theta^l)^+ - \mathbf{1}_{\{x_\tau^r > x_\tau^l\}}(\theta^l - \theta^r)^+ - c\tau \right].$$

Let $\kappa = \max(\theta^l, \theta^r)$; if the agent knew the value of each choice from the start he would obtain this payoff.

Proposition 2. (i) For any stopping time τ

$$\text{Ch}(\tau) = \mathbb{E} \left[\max\{X_\tau^l, X_\tau^r\} - c\tau \right] - \kappa.$$

(ii) Therefore, these two objective functions induce the same choice process.

To gain some intuition for this result, recall that the agent's expected payoff in our model when stopping at a fixed time t is $\mathbb{E} \left[\max\{X_t^l, X_t^r\} - ct \right]$. Now suppose we treat κ as a known constant and subtract it from the agent's payoff, yielding $\mathbb{E} \left[\max\{X_t^l, X_t^r\} - ct \right] - \kappa$, which is equal to $-ct$ if the agent makes the ex-post optimal choice and equal to $-ct - (\max(\theta^l, \theta^r) - \min(\theta^l, \theta^r))$ when he makes a mistake. The proof consists of treating κ as a random variable and using iterated expectations and the fact that τ is a stopping time to show that the expected value of this mistake is $\mathbb{E} \left[-\mathbf{1}_{\{x_\tau^l \geq x_\tau^r\}}(\theta^r - \theta^l)^+ - \mathbf{1}_{\{x_\tau^r > x_\tau^l\}}(\theta^l - \theta^r)^+ \right]$.

Chernoff and following authors in the mathematical statistics literature have focused on the behavior of the optimal boundary for very small and very large values of t , and have not characterized the full solution. We have not found any results on increasing versus decreasing accuracy in this literature, nor any comparative statics, but we make use of [Bather's \(1962\)](#) asymptotic analysis of the Chernoff model in what follows. Our characterization of the solution to equation (5) contributes to the study of the Chernoff problem by establishing non-asymptotic properties.

3.2 Certain Difference

In the simple DDM the agent's prior is concentrated on two points: (θ'', θ') and (θ', θ'') , where $\theta'' > \theta'$. The agent receives payoff θ'' for choosing l in state (θ'', θ') or r in state (θ', θ'') and $\theta' < \theta''$ for choosing r in state (θ'', θ') or l in state (θ', θ'') , so she knows that the magnitude of the utility difference between the two choices is $|\theta'' - \theta'|$, but doesn't know which action is better. We let μ_0 denote the agent's prior probability of (θ'', θ') .

This model was first studied in discrete time by [Wald \(1947\)](#) (with a trade-off between type I and type II errors taking the place of utility maximization) and by [Arrow, Blackwell](#)

and Girshick (1949) in a standard dynamic programming setting. The solution is essentially the same in continuous time.

Theorem 3. (*Shiryaev (1969, 2007).*) *With a binomial prior, there is $k > 0$ such that the minimal optimal stopping time is $\hat{\tau} = \inf\{t \geq 0 : |l_t| = k\}$, where $l_t = \log\left(\frac{\mathbb{P}[\theta=\theta_l|\mathcal{F}_t]}{\mathbb{P}[\theta=\theta_r|\mathcal{F}_t]}\right)$. Moreover, when $\mu_0 = 0.5$, the optimal stopping time has a DDM representation with a constant boundary b :*

$$\hat{\tau} = \inf\{t \geq 0 : |Z_t| \geq b\}^{19}$$

The simple DDM misses an important feature, as the assumption that the agent knows the magnitude of the payoff difference rules out cases in which the agent is learning the intensity of her preference. Intuitively, one might expect that if Z_t is close to zero and t is large, the agent would infer that the utility difference is small and so stop. This inference is ruled out by the binomial prior, which says that the agent is sure that he is not indifferent. We now turn to a model with a Gaussian prior which makes such inferences possible.

3.3 Uncertain-difference DDM

In the uncertain-difference DDM, the agent's prior μ_0 is independent for each action and normally distributed, with prior means X_0^i and common variance, σ_0^2 . Given the specification of the signal process (5), the posterior μ_t is $N(X_t^i, \sigma_t^2)$ where

$$X_t^i = \frac{X_0^i \sigma_0^{-2} + Z_t^i \alpha^{-2}}{\sigma_0^{-2} + t \alpha^{-2}} \quad \text{and} \quad \sigma_t^2 = \frac{1}{\sigma_0^{-2} + t \alpha^{-2}}. \quad (7)$$

Moreover, these equations describe the agent's beliefs at time t conditional on any sequence of the signal process up to t , so in particular they describe the agent's beliefs conditional on not having stopped before t .

Note that the variance of the agent's beliefs decreases at rate $1/t$ regardless of the data she receives.

Define the continuation value V as the expected value an agent can achieve by using the optimal continuation strategy if her posterior means are (x^l, x^r) at time t , the initial variance of the prior is σ_0^2 , and the noisiness of the signal is α :

$$V(t, x^l, x^r, c, \sigma_0, \alpha) := \sup_{\tau \geq t} \mathbb{E}_{(t, x^l, x^r, \sigma_0, \alpha)} [\max\{X_\tau^l, X_\tau^r\} - c(\tau - t)].$$

¹⁹This is essentially Theorem 5, p. 185 of Shiryaev (2007). In his model the drift depends on the sign of the utility difference, but not on its magnitude; his proof extends straightforwardly to our case.

Lemma 2 in the appendix establishes a number of useful properties of V , including that it is increasing and Lipschitz continuous in x^l and x^r and non-increasing in t . This leads to the following theorem, which characterizes the agent's optimal stopping rule.

Theorem 4 (Characterization of the optimal stopping time). *Let $k^*(t, c, \sigma_0, \alpha) = \min\{x \in \mathbb{R} : 0 = V(t, -x, 0, c, \sigma_0, \alpha)\}$. Then*

1. k^* is well defined.
2. Let τ^* be the minimal optimal strategy in (6). Then

$$\tau^* = \inf\{t \geq 0 : |X_t^l - X_t^r| \geq k^*(t, c, \sigma_0, \alpha)\}.$$

3. $k^*(t, c, \sigma_0, \alpha)$ is strictly positive, strictly decreasing in t , and $\lim_{t \rightarrow \infty} k^*(t, c, \sigma_0, \alpha) = 0$. Moreover, it is Lipschitz continuous with constant at most $2\alpha^{-2}\sigma_t^2 k^*$.
4. If $X_0^l = X_0^r$, then for $b^*(t, c, \sigma_0, \alpha) = \alpha^2 \sigma_t^{-2} k^*(t, c, \sigma_0, \alpha)$ we have

$$\tau^* = \inf\{t \geq 0 : |Z_t^l - Z_t^r| \geq b^*(t, c, \sigma_0, \alpha)\}.$$

5. $k^*(t, c, \sigma_0, \alpha)$ and $b^*(t, c, \sigma_0, \alpha)$ are pointwise non-increasing in c

Corollary 1. 1. *As c increases the agent decides earlier, in the sense of first-order stochastic dominance.*

2. *The probability of making the correct choice is non-increasing in c conditional on stopping at any time t for all t, θ^l, θ^r .*

The boundary k^* is given by the smallest difference in posterior means that makes the continuation value equal to the expected value of stopping, which is the higher of the two posterior means at the time the agent stops; this higher value can be set to 0 due to the shift invariance of the normally-distributed posterior beliefs. Part (2) of the theorem describes the optimal strategy τ^* in terms of stopping regions for posterior means $X_t^l - X_t^r$: It is optimal for the agent to stop once the expected utility difference exceeds the threshold $k^*(t, c, \sigma_0, \alpha)$. Intuitively, if the difference in means is sufficiently high it becomes unlikely that future signals will change the optimal action, and thus it is optimal to make a decision immediately and not incur additional cost. The proof of this follows from the principle of optimality for continuous time processes.

Note that the optimal strategy depends only on the difference in means and not on their absolute levels; however, it also depends on other parameters, in particular the prior variance. For example, if l and r are two houses with a given utility difference $\delta = \theta^l - \theta^r$, we expect the agent to spend on average more time here than on a problem where l and r are two lunch items with the same utility difference δ . This is because we expect the prior belief of the agent to be domain specific and in particular, the variance of the prior, σ_0^2 , to be higher for houses than for lunch items.

Part (3) says that k^* is decreasing and asymptotes to 0. To gain intuition for this result, consider the agent at time t deciding whether to stop now or to wait dt more seconds and then stop. The utility of stopping now is $\max_{i=l,r} X_t^i$. If the agent waits, she will have a more accurate belief and so be able to make a more informed decision, but she will pay an additional cost, leading to an expected change in utility of $(\mathbb{E}_t \max_{i=l,r} X_{t+dt}^i - \max_{i=l,r} X_t^i) - cdt$. Because belief updating slows down as shown in equation (6), the value of the additional information gained per unit time is decreasing in t , which leads the stopping boundaries to shrink over time; the boundaries shrink all the way to 0 because otherwise the agent would have a positive subjective probability of never stopping and incurring an infinite cost.²⁰ Part 5 says that k^* is pointwise non-increasing in c ; this is because k^* is defined with reference to the agent's value function, and the value function is non-increasing in c . This directly implies Corollary 1.1; Corollary 1.2 follows from Theorem 1.

Part 4 of the theorem describes the optimal strategy τ^* in terms of stopping regions for the signal process $Z_t := Z_t^l - Z_t^r$.²¹ This facilitates comparisons with the certain-difference DDM, where the process of beliefs lives in a different space and is not directly comparable. One way to understand the difference between these two models is to consider the agent's posterior beliefs when $Z_t \approx 0$ for large t . In the certain difference model, the agent interprets the signal as noise, since according to her prior the utilities of the the two alternatives are a fixed distance apart, so the agent disregards the signal and essentially starts from anew. This is why the optimal boundaries are constant in this model. On the other hand, in the uncertain difference model the agent's interpretation of $Z_t \approx 0$ for large t is that the two alternatives are nearly indifferent, which prompts the agent to stop the costly information gathering process and make a decision right away. This is why the optimal boundaries are

²⁰In contrast, in the certain difference DDM, the agent believes she will stop in finite time with probability 1 even though the boundaries are constant. This is because the agent knows that the absolute value of the drift of Z_t is bounded away from 0, while in the uncertain-difference model the agent believes it might be arbitrarily small.

²¹ When $X_0^l \neq X_0^r$, the optimal strategy can be described in terms of asymmetric boundaries for the signal process: $\underline{b}_t = \alpha^2 [-k(t)\sigma_t^{-2} - (X_0^l - X_0^r)\sigma_0^{-2}]$ and $\bar{b}(t) = \alpha^2 [k(t)\sigma_t^{-2} - (X_0^l - X_0^r)\sigma_0^{-2}]$.

decreasing in this model.

Bather (1962, example *i*) and Drugowitsch et al. (2012) study a similar problem where the agent knows the utility difference (that is $\delta = 1$ or $\delta = -1$) but is uncertain about the signal intensity (that is the drift is $\lambda\delta$ for some unknown λ). Drugowitsch et al. (2012) solve the problem numerically, and announce that the boundaries collapse to 0; however, they do not discuss the robustness or numerical stability of their results. Bather (1962) shows this convergence analytically and proves that the boundary decreases at the rate $1/\sqrt{t}$. Here too the boundaries collapse to zero because when $Z_t \approx 0$ for large t the agent thinks he is unlikely to learn more in the future.

To obtain a sharper characterization of the optimal boundary, we use space/time change arguments and basic facts about optimization problems to show that the functions k^* and b^* have to satisfy the conditions stated in Lemma 2 and Lemma 4 of the appendix. The conditions provide useful information about the identification of the parameters of the model, and about how the predictions of the model change as the parameters are varied. Moreover, they are an important underpinning for the rest of the results in this section; in particular, they are also used to show that k^* is Lipschitz continuous, which simplifies the analysis of the boundary value problem, and that it declines with time at rate at least $1/\sqrt{t}$, which is at the heart of the proof of the next theorem.

Proposition 3. *The optimal stopping boundary for the uncertain DDM has $b^*(t) \cdot \sigma_t$ non-increasing in t , so by Theorem 2, the average DDM with prior μ_0 and boundary function b^* has non-increasing accuracy.²²*

Proposition 3 implies that the analyst will observe decreasing accuracy when the agent faces a series of decisions with states (θ^r, θ^l) that are distributed according to the agent's prior. Moreover, the proof shows that decreasing accuracy even holds when the agent's prior is not symmetric, i.e. $X_0^l \neq X_0^r$.²³ This implies that as long as the prior is correct, decreasing accuracy will hold for the average P in a given experiment. In addition, we expect that decreasing accuracy should hold at least approximately if the agent's beliefs are approximately correct, but we have not shown this formally. Moreover, decreasing accuracy can hold even across experiments as long as the distributions of the states are close enough. That is, while we expect decreasing accuracy to hold within a given class of decision problems,

²²Note that if b^* is decreasing in t , then by Theorem 1, any for each realization of δ the induced choice probabilities $P(\delta, \alpha, b^*)$ display decreasing accuracy.

²³As mentioned in footnote 21, in this case P does not admit an average DDM representation as the upper and lower barrier are not symmetric.

it need not hold across classes with different prior variances. Similarly, decreasing accuracy can hold across subjects as long as their boundaries are not too different.

3.4 Approximations for large t and c

To gain more insight into the form of the optimal policy, we study approximations \bar{k}, \bar{b} to the optimal boundaries k^*, b^* that have simple and tractable functional forms. Using the results of Bather (1962) on the model we then show that \bar{b} approximates the solution well for large t , and that the boundary is approximately constant when c is large. Let

$$\bar{k}(t, c, \sigma_0, \alpha) = \frac{1}{2c\alpha^2(\sigma_0^{-2} + \alpha^{-2}t)^2} \quad \text{and} \quad \bar{b}(t, c, \sigma_0, \alpha) = \frac{1}{2c(\sigma_0^{-2} + \alpha^{-2}t)}.$$

Proposition 4. *There are constants $\beta, T > 0$ such that for all $t > T$*

$$|\bar{b}(t, c, \sigma_0, \alpha) - b^*(t, c, \sigma_0, \alpha)| \leq \frac{\beta}{(\sigma_0^{-2} + \alpha^{-2}t)^{5/2}}.$$

Bather’s result assumes that the agent has an improper prior with 0 prior precision; the proof of this proposition uses several of the rescaling arguments in Lemma 3 to adapt his result to our setting. One useful implication of Proposition 4 is that b^* asymptotically declines to zero at rate $1/t$. Moreover, when computing the solution to the optimal stopping problem numerically by working backwards from a fixed terminal date, we have found that \bar{k} and \bar{b} are extremely good numerical approximations to the optimal boundaries k^* and b^* .²⁴

We show now that for large c the initial portion of the boundary b^* is approximately $\frac{1}{2c}$, so that the agent decides immediately if $|X_0^l - X_0^r| > \sigma_0^2/2c\alpha^2$.

Proposition 5 (The initial position of the boundary). *For any $\alpha, \tilde{\sigma}_0$ there is a constant β independent of c such that*

$$\left| b^*(0, c, \tilde{\sigma}_0, \alpha) - \frac{1}{2c} \right| \leq \frac{\beta}{c^{4/3}}$$

3.5 Endogenously Divided Attention

We now consider a simple model of endogenous attention, where the agent can pay variable amounts of attention to each signal, and costlessly change these weights at any time. Specif-

²⁴The approximations for $t \rightarrow 0$ obtained by Bather (1962) show that $b^* \neq \bar{b}$; however our simulations and the bounds from Proposition 4 indicate that \bar{b} approximates the solution quite accurately for moderate t , so it may be useful for estimation purposes.

ically, at every point in time t the agent can choose attention levels (β_t^l, β_t^r) which influence the signals Z_t^1, Z_t^2 given by

$$dZ_t^i = \beta_t^i \theta^i dt + dB_t^i.$$

We assume that attention levels are always positive $\beta_t^l, \beta_t^r \geq 0$ and that the total amount of attention is bounded by two, i.e. $\beta_t^l + \beta_t^r \leq 2$.²⁵ Note that this is identical to the model analyzed before (with $\alpha = 1$) if the agent pays equal attention to the two signals, i.e. $\beta_t^l = \beta_t^r = 1$ for all $t \geq 0$.

The posterior distribution the agent assigns to the utility θ^i of alternative i at time t is Normal with mean X_t^i and variance $(\sigma_t^i)^2$ where

$$X_t^i = \frac{(\sigma_0^i)^{-2} X_0^i + \int_0^t (\beta_s^i)^2 dZ_s^i}{(\sigma_0^i)^{-2} + \int_0^t (\beta_s^i)^2 ds} \quad \text{and} \quad (\sigma_t^i)^2 = \frac{1}{(\sigma_0^i)^{-2} + \int_0^t (\beta_s^i)^2 ds}.$$

The variance of the posterior belief about the difference $\theta^l - \theta^r$ is given by the sum of the variances of θ^l and θ^r , i.e., $v_t := (\sigma_t^l)^2 + (\sigma_t^r)^2$.

Theorem 5.

1. *The posterior variance v_t at every point in time $t \geq 0$ is minimized by paying equal attention to the two signals $\beta_s^r = \beta_s^l = 1$ at every point in time s .*
2. *It is optimal to pay equal attention to both signals $\beta_t^l = \beta_t^r = 1$ at every point in time t .*

Note that the theorem does not say that it is optimal to give equal attention to both signals if the agents has paid unequal attention in the past, as then the agent might want to pay more attention to the signal with the higher posterior variance. Instead, the proof uses the comparison principle for differential equations to show that the equal-attention policy dominates any other.

A more realistic model might add an adjustment cost for changing attention, and to fit eye-tracking data such as that of [Krajbich, Armel and Rangel \(2010\)](#) one might also assume that the agent can only pay attention to one signal at a time. This results in a more complex control problem with an additional state variable. We have not solved that model but conjecture that the as the adjustment cost shrinks to 0 the optimal policy is to quickly oscillate or “chatter” to approximate the equal-attention solution derived above. As

²⁵The normalization of the attention budget to two allows us to relate the attention model to the previous model. All our results generalize to arbitrary attention budgets.

quick oscillation of attention is not observed in laboratory experiments this highlights the importance role switching costs might play when modeling attention.

Che and Mierendorff (2016) study a certain difference model with perfectly revealing Poisson signals, where as in our model the agent can split a fixed attention budget in arbitrary ways between the alternatives. They show that it is optimal for the agent to pay full attention to only one alternative at a time and switch attention back and forth depending on the posterior. Ke and Villas-Boas (2016) study endogenous attention in a multi-object choice stopping problem when each choice has two possible values and the agent can only pay attention to one signal at a time .

In Woodford (2014) and Hebert and Woodford (2016) the agent directly chooses the informativeness of a signal process. Woodford (2014) assumes the agent can optimize the dependence of the process Z_t on θ subject to a Shannon capacity constraint and that the stopping rule is constrained to have constant boundaries. Hebert and Woodford (2016) combine optimal stopping and optimal information gathering when the agent can choose from a very general class of signals; they characterize the cost functions that have a form of “sequential prior invariance” and show that the sequential evidence accumulation problem can be thought of as a static rational inattention problem with a particular cost function.

3.6 Non-Linear Cost

In deriving the DDM representation from optimal stopping, we have so far assumed that the cost of continuing per unit time is constant. We have seen that in the uncertain-difference model, the optimal boundary decreases due to the fact there is less to learn as time goes on. One would expect that the boundary could increase if costs decrease sufficiently quickly. This raises the question of which DDM representations can be derived as a solution to an optimal stopping problem when the cost is allowed to vary arbitrarily over time. The next result shows that for any boundary there exists a cost function such that the boundary is optimal in the learning problem with normal or binomial priors. Thus optimal stopping on its own imposes essentially no restrictions on the observed choice process; the force of the model derives from its joint assumptions about the evolution of beliefs and the cost function.

Theorem 6. *Consider either the Certain or the Uncertain-Difference DDM. For any finite boundary b and any finite set $G \subseteq \mathbb{R}_+$ there exists a cost function $d : \mathbb{R}_+ \rightarrow \mathbb{R}$ such that b is*

optimal in the set of stopping times T that stop in G with probability one

$$\inf\{t \in G: |Z_t| \geq b(t)\} \in \arg \max_{\tau \in \mathcal{T}} \mathbb{E} [\max\{X_\tau^1, X_\tau^2\} - d(\tau)] .$$

In particular, there is a cost function such that the exponentially decreasing boundaries in Milosavljevic et al. (2010) are optimal, and a cost function that leads to constant accuracy.

Intuitively, the reason this result obtains is that the optimal stopping rule always takes the form of a cut-off: If the agent stops at time t when $X_t = x$, she stops at time t whenever $|X_t| > x$. This allows us to recursively construct a cost function that rationalizes the given boundary by setting the cost at time t equal to the expected future gains from learning. To avoid some technical issues having to do with the solutions to PDE's, we consider a discrete-time finite-horizon formulation of the problem, where the agent is only allowed to stop at times in a finite set G . This lets us construct the associated cost function period by period instead of using smoothness conditions and stochastic calculus.²⁶

4 Empirical Analysis

In this section we report three different ways of relating our theoretical results to the data that inspired us: an aggregate analysis that groups together people and decision problems to get more power, a test of whether in individual-level data the boundary is constant, and finally the results of fitting the exact model to data on each individual using a numerical computation of the optimal stopping boundary. In all three sections we use data from Krajbich, Armel and Rangel (2010); we thank them for sharing the data with us.

In this experiment, subjects were asked to refrain from eating for 3 hours before the experiment started. They were told they would be making a number of binary choices between food items, and that at the end of the experiment they would be required to stay in the room with the experimenter for 30 minutes while eating the food item that they chose in a randomly selected trial.²⁷ Before making their choices, subjects entered liking ratings

²⁶Our proof relies on a result on implementable stopping times from Kruse and Strack (2015). In another paper Kruse and Strack (2014) generalize this result to continuous time, but as the absolute value is not covered by their result we can not use it here. Nevertheless, we conjecture that the methods used in that paper can be extended to prove the result in continuous time directly. Drugowitsch et al. (2012) propose a related result, namely a way to construct a cost function so that the associated boundary is optima for any observed stopping times. However their construction implicitly requires that the constructed cost function be absolutely continuous, and they do not discuss what restrictions on the data are needed for this.

²⁷Food items that received a negative rating in the rating phase of the experiment were excluded from the choice phase but subjects were not informed of this. The items shown in each trial were chosen pseudo-

for 70 different foods on a scale from -10 to +10.

4.1 Aggregate Analysis

In this subsection we group together all subjects, and group decision problems by the difference in ratings of the left-hand and right-hand item, ranging from -5 to 5 by unit increments, cf Figure 5 and 6 in [Krajbich, Armel and Rangel \(2010\)](#).

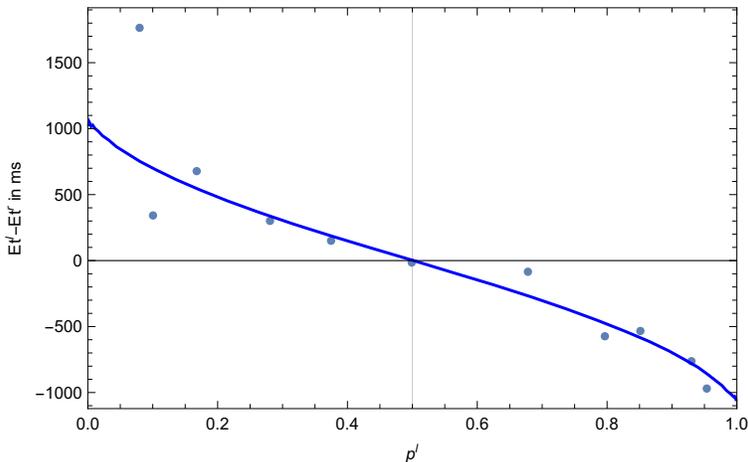


Figure 1: The dots display the data from [Krajbich, Armel and Rangel \(2010\)](#), grouping pairs by difference in self-reported item rankings. The blue line displays simulations of the approximately optimal barrier \bar{b} for the uncertain-difference DDM (derived in Section 3.3) with parameters calibrated to $c = 0.05, \sigma_0 = \alpha = 1$; each simulation uses $5 \cdot 10^6$ draws and time is discretized with $dt = 0.01$. The expected time to choose l is lower than the expected time to choose r for pairs of alternatives where l is the modal choice; it is higher for pairs where r is modal.

Figure 1 shows that in the aggregate there is decreasing accuracy: the average time to choose l is lower than the expected time to choose r when l is more highly rated, and conversely the time for l is higher when it is lower rated. The blue line shows that our model tracks the qualitative properties of the data. Note that certain-difference DDM cannot reproduce the decreasing accuracy observed in the data, as it implies that the average decision time and the choice probability are independent.

randomly according to the following rules: (i) no item was used in more than 6 trials; (ii) the difference in liking ratings between the two items was constrained to be 5 or less; (iii) if at some point in the experiment (i) and (ii) could no longer both be satisfied, then the difference in allowable liking ratings was expanded to 7, but these trials occurred for only 5 subjects and so were discarded from the analyses.

However, the aggregate data leaves open the possibility that the decreasing accuracy comes from aggregating over different subjects as opposed to occurring at the individual level. For this reason we now proceed to analyze the data of each individual subject.

4.2 Individual Level Analysis of the Slope of the Boundary

Here we study the slope of the boundary using the functional form $\tilde{b}(t) = \frac{1}{g+ht}$ for some $g > 0$ and $h \geq 0$. This functional form nests the approximate boundary \bar{b} of Proposition 4 (by setting $g := 2c\sigma_0^{-2}$ and $h := 2c\alpha^{-2}$) and also nests the simple-DDM case of a constant boundary (by setting $h := 0$), which the exact boundary does not. We show that for all but 3 out of 39 subjects h is significantly different from zero at the 3 percent significance level. By Theorem 1, this implies that except for those outliers, all subjects display decreasing choice accuracy over time.

More specifically, the stopping time distribution depends only on (δ, α, g, h) , where α is the volatility of Z_t and we assume that the drift δ equals the difference in reported ratings and (g, h) determine the boundary. We computed the likelihoods for all combinations of (δ, α, g, h) with $\delta \in \{-5, -4, \dots, 4, 5\}$, $\alpha \in \{1.0, 1.2, \dots, 5.0\}$, $g \in \{0.01, 0.02, \dots, 0.50\}$, and $h \in \{0.000, 0.005, \dots, 0.500\}$.²⁸ We used maximum likelihood to estimate both the unrestricted parameters (α^*, g^*, h^*) and the restricted parameters $(\alpha^\dagger, g^\dagger)$ when h is set equal to 0. The results are shown in Table 1 in the online appendix.

4.3 Individual Level Analysis using the Exact Boundary

Finally, we estimated the exact model at the individual level using a numerically computed b function. Here we find that there is substantial heterogeneity between the subjects. Figure 2 presents the histograms of the marginal distributions of α , c , and σ_0 across all subjects.

Figure 2 also plots the estimated cost for each subject along with that subject's average stopping time. As it shows, the dispersion in costs explains a substantial fraction of the variation in times.

While Proposition 3 proves that the accuracy of the agent is decreasing on average whenever the agent makes optimal choices, the accuracy of the agents choice need not be decreasing. An example of such a situation is the optimal boundary for the parameters $\sigma = 1.8, c = 0.02, \alpha = 2$ which is first increasing and then decreasing (c.f. Figure 3, subject 45). However, our estimation suggests that for all but two subjects the optimal boundary

²⁸The supplementary appendix has a detailed description of our numerical methods.

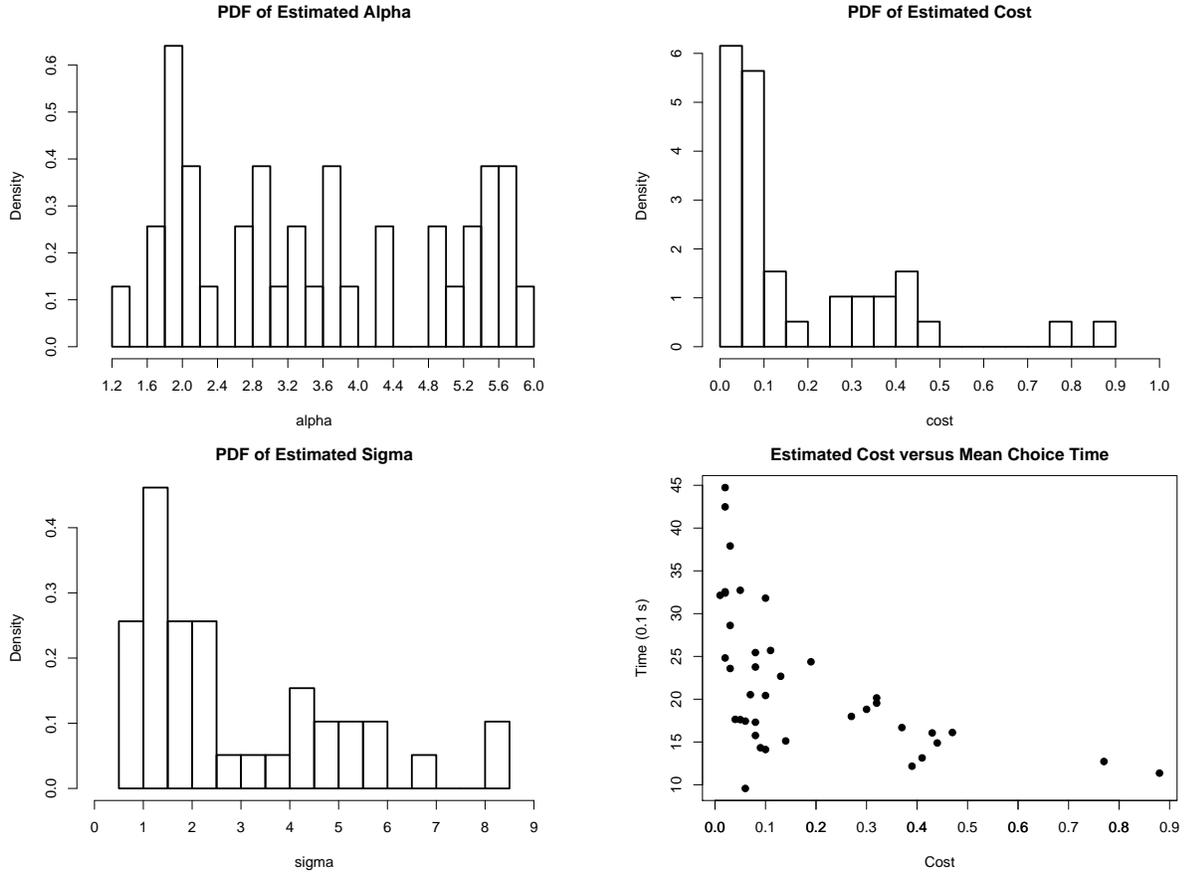


Figure 2: Marginal distributions of α , c , and σ_0 and the correlation between the average stopping time for each subject and their estimated cost c .

which best fits their choice is monotone decreasing (see Figure 3). Theorem 1 then implies that those subjects display decreasing accuracy.

5 Conclusion

The recent literature in economics and cognitive science uses drift-diffusion models with time-dependent boundaries. This is helpful in matching observed properties of reaction times, notably their correlation with chosen actions, and in particular a phenomenon that we call speed-accuracy complementarity, where earlier decisions are better than later ones. In Section 2 we showed that the monotonicity properties of the boundary characterize whether the observed choice process displays speed-accuracy complementarity, or the opposite pattern of a speed-accuracy tradeoff. This ties an observable property of behavior (the correlation

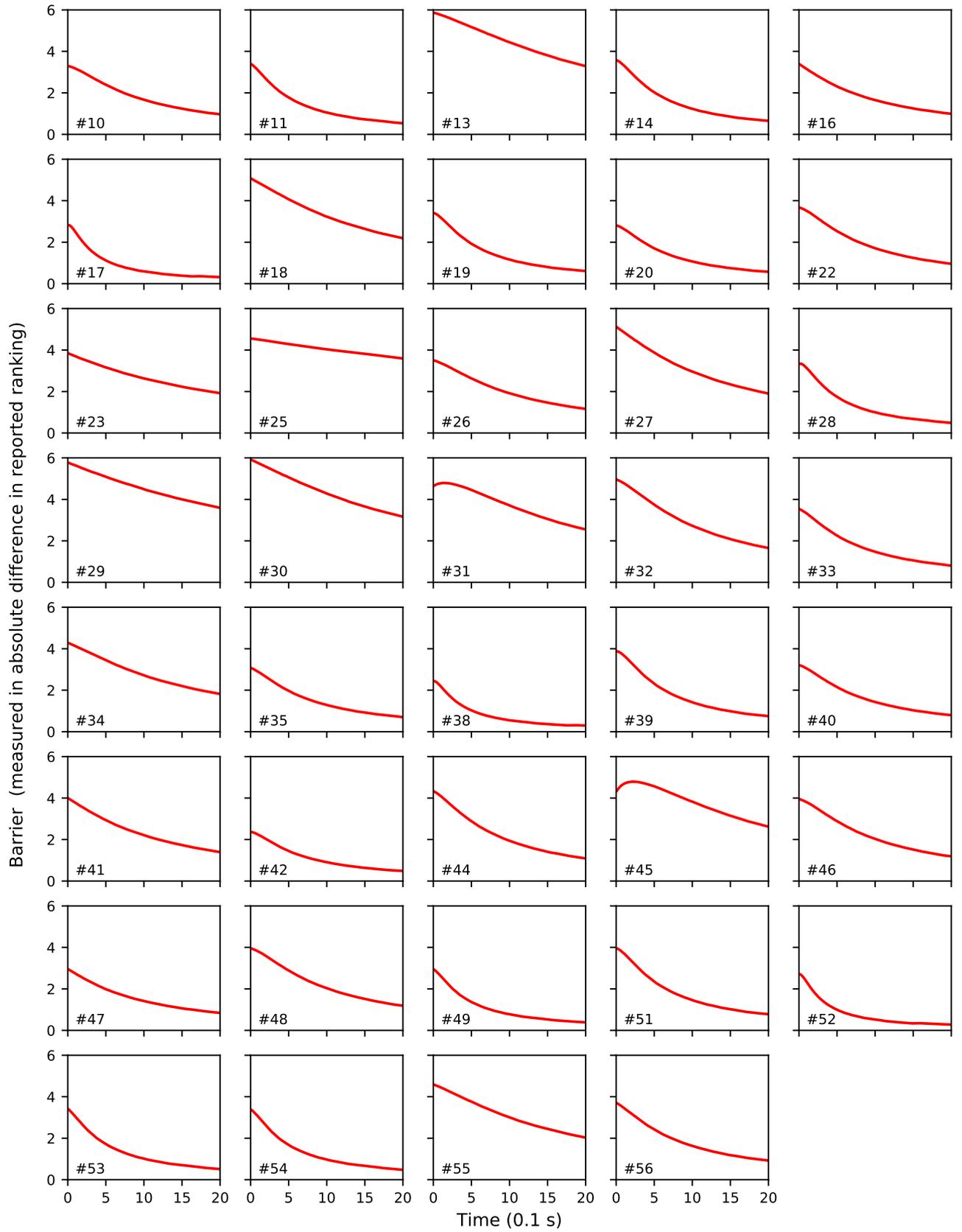


Figure 3: Estimated optimal boundaries for different subjects.

between reaction times and decisions) to an unobservable construct of the model (the boundary). This connection is helpful for understanding the qualitative properties of DDMs; it may also serve as a useful point of departure for future quantitative exploration of the connection between the rate of decline of the boundary and the strength of correlation between reaction times and actions.

In Section 3 we investigated the DDM as a solution to the optimal sequential sampling problem, where the agent is unsure about the utility of each action and is learning about it as the time passes, optimally deciding when to stop. We studied the dependence of the solution on the nature of the learning problem and also on the cost structure. In particular, we proposed a model in which the agent is learning not only about which option is better, but also by how much. We showed that the boundary in this model asymptotically declines to zero at the rate $1/t$, and that it is pointwise decreasing in c . We also showed that any boundary could be optimal if the agent is facing a nonlinear cost of time.

Our analysis provides a precise foundation for DDMs with time-varying boundaries, and establishes a set of useful connections between various parameters of the model and predicted behavior, thus enhancing the theoretical understanding of the model as well as making precise its empirical content. We hope these results will be a helpful stepping stone for further work. We expect the forces identified in this paper to be present in other decisions involving uncertainty: not just in tasks used in controlled laboratory experiments, but also in decisions involving longer time scales, such as choosing an apartment rental, or deciding which papers to publish as a journal editor.

A Appendix: Proofs

A.1 Proof of Theorem 1

Let $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ be the density of the distribution of stopping times, and $g : \mathbb{R}_+ \times \mathbb{R} \rightarrow \mathbb{R}_+$ be the density of Z_t , i.e.,

$$\begin{aligned} g(t, y) &= \frac{\partial}{\partial y} \mathbb{P} [Z_t \leq y | \delta, \alpha] = \frac{\partial}{\partial y} \mathbb{P} [\delta t + \alpha \sqrt{2} B_t \leq y | \delta, \alpha] \\ &= \frac{\partial}{\partial y} \mathbb{P} \left[\frac{B_t}{\sqrt{t}} \leq \frac{y - \delta t}{\alpha \sqrt{2} \sqrt{t}} \right] = \phi \left(\frac{y - \delta t}{\alpha \sqrt{2t}} \right) \end{aligned}$$

where $\phi(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$ is the density of the standard normal. By Bayes rule:

$$\begin{aligned} p^l(t) &= \mathbb{P}[Z_t = b(t) | \tau = t, \delta, \alpha] = \frac{g(t, b(t))\mathbb{P}[\tau = t | Z_t = b(t), \delta, \alpha]}{f(t)} \\ p^r(t) &= \mathbb{P}[Z_t = -b(t) | \tau = t, \delta, \alpha] = \frac{g(t, -b(t))\mathbb{P}[\tau = t | Z_t = -b(t), \delta, \alpha]}{f(t)} \end{aligned}$$

It follows from $Z_0 = 0$ and the symmetry of the upper and the lower barrier that

$$\mathbb{P}[\tau = t | Z_t = b(t), \delta, \alpha] = \mathbb{P}[\tau = t | Z_t = -b(t), -\delta, \alpha], \quad (8)$$

because for any path of Z that ends at $b(t)$ and crosses any boundary before t , the reflection of this path ends at $-b(t)$ and crosses some boundary at the same time.

The induced probability measure over paths conditional on $Z_t = b(t)$ is the same as the probability of the Brownian Bridge.²⁹ The Brownian Bridge is the solution to the SDE $dZ_s = -\frac{b(t)-Z}{t-s}ds + \alpha dB_s$ and notably does not depend on the drift δ , which implies that

$$\mathbb{P}[\tau = t | Z_t = -b(t), -\delta, \alpha] = \mathbb{P}[\tau = t | Z_t = -b(t), \delta, \alpha] \quad (9)$$

Thus, by (8) and (9) we have that

$$\frac{p^l(t)}{p^r(t)} = \frac{g(t, b(t))}{g(t, -b(t))} = \exp\left(\frac{\delta b(t)}{\alpha^2}\right).$$

Wlog $m = l$ and $\delta > 0$; the above expression decreases with t if and only if $b(t)$ decreases with t . \square

A.2 Proof of Proposition 1

In this proof we use the notation introduced in Section 3.3. Note that $Z_t = Z_t^l - Z_t^r$ and $\delta = \theta^l - \theta^r$. The outside observer knows that if the agent stopped at time t then the absolute value of the difference in posterior means, $|X_t^l - X_t^r|$, is equal to $k(t)$. Conditional on this information $|\delta| = |\theta^l - \theta^r|$ is folded Normal distributed, with mean $k(t)$ and variance $2\sigma_t^2$, i.e.

$$\begin{aligned} \mathbb{P}[|\delta| < d | \tau = t] &= \mathbb{P}\left[\theta^l - \theta^r \in (-d, d) | \tau = t, |X_\tau^l - X_\tau^r| = k(\tau)\right] \\ &= \frac{1}{\sqrt{2\pi}} \int_{\frac{k(t)-d}{\sqrt{2\sigma_t}}}{\frac{k(t)+d}{\sqrt{2\sigma_t}}} e^{-y^2/2} dy = \frac{1}{\sqrt{2\pi}} \int_{\beta(t) - \frac{d}{\sqrt{2\sigma_t}}}{\beta(t) + \frac{d}{\sqrt{2\sigma_t}}} e^{-y^2/2} dy \end{aligned} \quad (10)$$

where $\beta(t) := k(t)/(\sqrt{2}\sigma_t)$. As $b(t)\sigma_t$ is non-increasing $\beta(t) = k(t)/(\sqrt{2}\sigma_t)$ is non-increasing. As $\beta(t)$ and σ_t are non-increasing it suffices to show that (10) decreases in β and σ to prove that (10)

²⁹See, e.g., Proposition 12.3.2 of Dudley (2002) or Exercise 3.16, p. 41 of Revuz and Yor (1999).

increases in t :

$$\begin{aligned} \frac{\partial}{\partial \beta} \left[\frac{1}{\sqrt{2\pi}} \int_{\beta - \frac{d}{\sqrt{2\sigma}}}^{\beta + \frac{d}{\sqrt{2\sigma}}} e^{-y^2/2} dy \right] &= \frac{1}{\sqrt{2\pi}} \left[\exp \left\{ -\frac{1}{2} \left(\beta + \frac{d}{\sqrt{2\sigma}} \right)^2 \right\} - \exp \left\{ -\frac{1}{2} \left(\beta - \frac{d}{\sqrt{2\sigma}} \right)^2 \right\} \right] < 0 \\ \frac{\partial}{\partial \sigma} \left[\frac{1}{\sqrt{2\pi}} \int_{\beta - \frac{d}{\sqrt{2\sigma}}}^{\beta + \frac{d}{\sqrt{2\sigma}}} e^{-y^2/2} dy \right] &= -\frac{d}{2\sqrt{\pi}\sigma^2} \left[\exp \left\{ -\frac{1}{2} \left(\beta + \frac{d}{\sqrt{2\sigma}} \right)^2 \right\} + \exp \left\{ -\frac{1}{2} \left(\beta - \frac{d}{\sqrt{2\sigma}} \right)^2 \right\} \right] \\ &< 0. \quad \square \end{aligned}$$

A.3 Proof of Theorem 2

In this proof we use the notation introduced in Section 3.3, i.e. $X_t = \mathbb{E}[\delta \mid (Z_s)_{s \leq t}]$, $\sigma_t^2 = \frac{1}{2} \mathbb{E}[(\delta - X_t)^2 \mid (Z_s)_{s \leq t}]$. As we show in Section 3.3, the posterior beliefs about δ conditional on stopping at time t when $Z_t = b(t)$ are normal with mean X_t and variance $2\sigma_t^2$.

If the agent stops at time t when $Z_t^l - Z_t^r = b(t)$, then the conditional means X_t^l, X_t^r satisfy $X_t^l - X_t^r = b(t)\alpha^{-2}\sigma_t^2 \equiv k(t)$, and the probability that the agent picks l when r is optimal is

$$\begin{aligned} \mathbb{P}[\theta^l < \theta^r \mid X_t^l - X_t^r = k(t)] &= \mathbb{P} \left[(\theta^l - \theta^r) - (X_t^l - X_t^r) \leq -k(t) \mid X_t^l - X_t^r = k(t) \right] \\ &= \mathbb{P} \left[\frac{(\theta^l - \theta^r) - (X_t^l - X_t^r)}{\sqrt{2}\sigma_t} \leq -\frac{k(t)}{\sqrt{2}\sigma_t} \mid X_t^l - X_t^r = k(t) \right] = \Phi \left(-\frac{1}{\sqrt{2}} k(t) \sigma_t^{-1} \right). \end{aligned}$$

Note that since the beliefs depend only on the endpoint and not on the whole path, we have

$$\mathbb{P}[\theta^l < \theta^r \mid X_t^l - X_t^r = k(t)] = \mathbb{P}[\theta^l < \theta^r \mid X_t^l - X_t^r = k(t) \text{ and } |X_s^l - X_s^r| < k(s) \text{ for all } s < t]$$

By the symmetry of the problem, the probability of mistakenly picking r instead of l is the same. \square

A.4 The Value Function

We use the following representation of the posterior process in the uncertain-difference model.

Lemma 1. *For any $t > 0$*

$$X_t^i = X_0^i + \int_0^t \frac{\alpha^{-1}}{\sigma_0^{-2} + s\alpha^{-2}} dW_s^i$$

where W_s^i is a Brownian motion with respect to the filtration information of the agent.

Proof: This follows from Theorem 10.1 and equation 10.52 of Liptser and Shiryaev (2001) by setting $a = b = 0$ and $A = 1, B = \alpha$. \square

Define the continuation value as the expected value an agent can achieve by using the optimal continuation strategy if she believes the posterior means to be (x^l, x^r) at time t and the variance of his prior equaled σ_0^2 at time 0 and the noisiness of the signal is α .

$$V(t, x^l, x^r, c, \sigma_0, \alpha) := \sup_{\tau \geq t} \mathbb{E}_{(t, x^l, x^r, \sigma_0, \alpha)} \left[\max\{X_\tau^l, X_\tau^r\} - c(\tau - t) \right].$$

Lemma 2. *The continuation value V has the following properties:*

1. $\mathbb{E}_{(t, x^l, x^r, \sigma_0, \alpha)}[\max\{\theta^l, \theta^r\}] \geq V(t, x^l, x^r, c, \sigma_0, \alpha) \geq \max\{x^l, x^r\}$.
2. $V(t, x^l, x^r, c, \sigma_0, \alpha) - \beta = V(t, x^l - \beta, x^r - \beta, c, \sigma_0, \alpha)$ for every $\beta \in \mathbb{R}$.
3. The function $V(t, x^l, x^r, c, \sigma_0, \alpha) - x^i$ is decreasing in x^i for $i \in \{l, r\}$.
4. $V(t, x^l, x^r, c, \sigma_0, \alpha)$ is increasing in x^l and x^r .
5. $V(t, x^l, x^r, c, \sigma_0, \alpha)$ is Lipschitz continuous in x^l and x^r .
6. $V(t, x^l, x^r, c, \sigma_0, \alpha)$ is non-increasing in t .
7. $V(t, x^l, x^r, c, \sigma_0, \alpha) = V(0, x^l, x^r, c, \sigma_t, \alpha)$ for all $t > 0$.

Proof of Lemma 2

In this proof we equivalently represent a continuation strategy by a pair of stopping times (τ^l, τ^r) , one for each alternative.

Proof of 1: For the lower bound, the agent can always stop immediately and get x^l or x^r . For the upper bound, the agent can't do better than receiving a fully informative signal right away and pick the better item immediately.

Proof of 2: Intuitively, this comes from the translation invariance of the belief process X . To prove the result formally, fix a continuation strategy (τ^l, τ^r) . Recall that Lemma 1 we can represent X_t as $X_t^i = X_0^i + \int_0^t \frac{\alpha^{-1}}{\sigma_0^{-2} + s\alpha^{-2}} dW_s^i$ and thus (τ^l, τ^r) can be interpreted as a mapping from paths of the Brownian motion (W^l, W^r) into stopping times. As such we can define the strategy as a function of the Brownian motion W without explicit reliance on the belief process X and its starting value. The expected payoff when using the strategy (τ^l, τ^r) equals

$$\begin{aligned}
& \mathbb{E} [\mathbf{1}_{\{\tau^l \leq \tau^r\}} X_{\tau^l}^l + \mathbf{1}_{\{\tau^l > \tau^r\}} X_{\tau^r}^r - c(\min\{\tau^l, \tau^r\} - t) \mid X_t^l = x^l - k, X_t^r = x^r - k] \\
&= \mathbb{E} \left[\mathbf{1}_{\{\tau^l \leq \tau^r\}} \left(\int_t^{\tau^l} \frac{\alpha^{-1}}{\sigma_0^{-2} + s\alpha^{-2}} dW_s^l + x^l - k \right) + \mathbf{1}_{\{\tau^l > \tau^r\}} \left(\int_t^{\tau^r} \frac{\alpha^{-1}}{\sigma_0^{-2} + s\alpha^{-2}} dW_s^r + x^r - k \right) \right. \\
&\quad \left. - c(\min\{\tau^l, \tau^r\} - t) \right] \\
&= \mathbb{E} [\mathbf{1}_{\{\tau^l \leq \tau^r\}} X_{\tau^l}^l + \mathbf{1}_{\{\tau^l > \tau^r\}} X_{\tau^r}^r - c(\min\{\tau^l, \tau^r\} - t) \mid X_t^l = x^l, X_t^r = x^r] - k.
\end{aligned}$$

Since V is defined as the supremum over all continuation strategies (τ^l, τ^r) the result follows.

Proof of 3: The expected difference between stopping at time t with option l and using the continuation strategy (τ^l, τ^r) is

$$\begin{aligned} & \mathbb{E} [\mathbf{1}_{\{\tau^l \leq \tau^r\}} X_{\tau^l}^l + \mathbf{1}_{\{\tau^l > \tau^r\}} X_{\tau^r}^r - c(\min\{\tau^l, \tau^r\} - t) \mid X_t^l = x^l, X_t^r = x^r] - x^l \\ &= \mathbb{E} [\mathbf{1}_{\{\tau^l \leq \tau^r\}} (X_{\tau^l}^l - x^l) + \mathbf{1}_{\{\tau^l > \tau^r\}} (X_{\tau^r}^r - x^l) - c(\min\{\tau^l, \tau^r\} - t) \mid X_t^l = x^l, X_t^r = x^r] \\ &= \mathbb{E} \left[\mathbf{1}_{\{\tau^l \leq \tau^r\}} \int_t^{\tau^l} \frac{\alpha^{-1}}{\sigma_0^{-2} + s\alpha^{-2}} dW_s^l + \mathbf{1}_{\{\tau^l > \tau^r\}} (X_{\tau^r}^r - x^l) - c(\min\{\tau^l, \tau^r\} - t) \mid X_t^l = x^l, X_t^r = x^r \right] \end{aligned}$$

Note that the first part is independent of x^l , and $(X_{\tau^r}^r - x^l)$ is weakly decreasing in x^l . As for every fixed strategy (τ^l, τ^r) the value of waiting is decreasing, the supremum over all continuation strategies is also weakly decreasing in x^l . Thus it follows that the difference between continuation value $V(t, x^l, x^r, c, \sigma_0, \alpha)$ and value of stopping immediately on the first arm x^l is decreasing in x^l for every t and every x^r . Intuitively, because the valuations of the objects are independent, increasing the belief about one arm has no effect on the expected value of the other. If there were only one choice, then $V(x) - x$ would be constant and equal to 0; because the agent might take the other arm the impact of higher signals is “damped” and so $V - x$ is decreasing in x .

Proof of 4: The expected value of using the continuation strategy (τ^l, τ^r) equals

$$\begin{aligned} & \mathbb{E} [\mathbf{1}_{\{\tau^l \leq \tau^r\}} X_{\tau^l}^l + \mathbf{1}_{\{\tau^l > \tau^r\}} X_{\tau^r}^r - c(\min\{\tau^l, \tau^r\} - t) \mid X_t^l = x^l, X_t^r = x^r] \\ &= \mathbb{E} \left[\mathbf{1}_{\{\tau^l \leq \tau^r\}} \left(\int_t^{\tau^l} \frac{\alpha^{-1}}{\sigma_0^{-2} + s\alpha^{-2}} dW_s^l + x^l \right) + \mathbf{1}_{\{\tau^l > \tau^r\}} X_{\tau^r}^r - c(\min\{\tau^l, \tau^r\} - t) \mid X_t^l = x^l, X_t^r = x^r \right] \\ &= \mathbb{E} \left[\mathbf{1}_{\{\tau^l \leq \tau^r\}} \int_t^{\tau^l} \frac{\alpha^{-1}}{\sigma_0^{-2} + s\alpha^{-2}} dW_s^l + \mathbf{1}_{\{\tau^l > \tau^r\}} X_{\tau^r}^r - c(\min\{\tau^l, \tau^r\} - t) \mid X_t^l = x^l, X_t^r = x^r \right] \\ & \quad + x^l \mathbb{E} [\mathbf{1}_{\{\tau^l \leq \tau^r\}} \mid X_t^l = x^l, X_t^r = x^r], \end{aligned}$$

which is weakly increasing in x^l . Consequently, the supremum over all continuation strategies (τ^l, τ^r) is weakly increasing in x^l . By the same argument it follows that $V(t, x^l, x^r, c, \sigma_0, \alpha)$ is increasing in x^r .

Proof of 5 : To see that the value function is Lipschitz continuous in x^l and x^r with constant 1, note that changing the initial beliefs moves the posterior beliefs at any fixed time linearly and has no effect on the cost of stopping at that time. Thus, the supremum over all stopping times can at most be linearly affected by a change in initial belief. To see this explicitly, observe that

$$\begin{aligned} & |V(0, x^l, x^r, c, \sigma_0, \alpha) - V(0, y^l, x^r, c, \sigma_0, \alpha)| \\ &= \left| \sup_{\tau} \mathbb{E} \left[\max\left\{ x^l + \int_0^{\tau} \frac{\alpha^{-1}}{\sigma_0^{-2} + s\alpha^{-2}} dW_s^1, x^r + \int_0^{\tau} \frac{\alpha^{-1}}{\sigma_0^{-2} + s\alpha^{-2}} dW_s^2 \right\} - c(\tau) \right] \right. \\ & \quad \left. - \sup_{\tau} \mathbb{E} \left[\max\left\{ y^l + \int_0^{\tau} \frac{\alpha^{-1}}{\sigma_0^{-2} + s\alpha^{-2}} dW_s^1, x^r + \int_0^{\tau} \frac{\alpha^{-1}}{\sigma_0^{-2} + s\alpha^{-2}} dW_s^2 \right\} - c(\tau) \right] \right| \\ &\leq \left| \sup_{\tau} \mathbb{E} \left[\max\left\{ x^l + \int_0^{\tau} \frac{\alpha^{-1}}{\sigma_0^{-2} + s\alpha^{-2}} dW_s^1, x^r + \int_0^{\tau} \frac{\alpha^{-1}}{\sigma_0^{-2} + s\alpha^{-2}} dW_s^2 \right\} \right. \right. \\ & \quad \left. \left. - \max\left\{ y^l + \int_0^{\tau} \frac{\alpha^{-1}}{\sigma_0^{-2} + s\alpha^{-2}} dW_s^1, x^r + \int_0^{\tau} \frac{\alpha^{-1}}{\sigma_0^{-2} + s\alpha^{-2}} dW_s^2 \right\} \right] \right| \leq |y^l - x^l|. \end{aligned}$$

Proof of 6: We show that $V(t, x^l, x^r, c, \sigma_0, \alpha)$ is decreasing in t . Note that by Doob's optional sampling theorem for every fixed stopping strategy τ

$$\begin{aligned}\mathbb{E} \left[\max\{X_\tau^l, X_\tau^r\} - c\tau \mid X_t = (x^l, x^r) \right] &= \mathbb{E} \left[\max\{X_\tau^l - X_\tau^r, 0\} + X_\tau^r - c\tau \mid X_t = (x^l, x^r) \right] \\ &= \mathbb{E} \left[\max\{X_\tau^l - X_\tau^r, 0\} - c\tau \mid X_t = (x^l, x^r) \right] + x_t^r.\end{aligned}$$

Define the process $X_t := X_t^l - X_t^r$, and note that

$$X_t = X_t^l - X_t^r = X_0^l - X_0^r + \int_0^t \frac{\alpha^{-1}}{\sigma_0^{-2} + s\alpha^{-2}} (dW_s^l - dW_s^r) = X_0^l - X_0^r + \int_0^t \frac{\sqrt{2}\alpha^{-1}}{\sigma_0^{-2} + s\alpha^{-2}} d\tilde{W}_s, \quad (11)$$

where \tilde{W} is a Brownian motion. Define a time change as follows: Let $q(k)$ solve $k = \int_0^{q(k)} \left(\frac{\sqrt{2}\alpha^{-1}}{\sigma_0^{-2} + s\alpha^{-2}} \right)^2 ds$. This implies that $q(k) = \frac{k\alpha^2\sigma_0^{-2}}{2\sigma_0^2 - k}$. Define $\psi(t) = \frac{2\sigma_0^2 t}{\alpha^2\sigma_0^{-2} + t}$. By (Theorem 1.6, chapter V of [Revuz and Yor, 1999](#)) $W_s := (X_{q(s)})_{s \in [0, 2\sigma_0^{-2}]}$ is a Brownian motion and thus we can rewrite the problem as

$$\begin{aligned}V(t, x^l, x^r, c, \sigma_0, \alpha) &= \sup_{\tau \geq \psi(t)} \mathbb{E} \left[\max\{W_\tau, 0\} - c \left(q(\tau) - q(\psi(t)) \right) \mid W_{\psi(t)} = x^l - x^r \right] + x^r \\ &= \sup_{\tau \geq \psi(t)} \mathbb{E} \left[\max\{W_\tau, 0\} - c \left(\int_{\psi(t)}^\tau q'(s) ds \right) \mid W_{\psi(t)} = x^l - x^r \right] + x^r, \\ &= \sup_{\tau \geq \psi(t)} \mathbb{E} \left[\max\{W_\tau, 0\} - c \left(\int_{\psi(t)}^\tau \frac{2\alpha^2}{(2\sigma_0^2 - s)^2} ds \right) \mid W_{\psi(t)} = x^l - x^r \right] + x^r.\end{aligned}$$

Next, we remove the conditional expectation in the Brownian motion by adding the initial value

$$V(t, x^l, x^r, c, \sigma_0, \alpha) = \sup_{\tau \geq \psi(t)} \mathbb{E} \left[\max\{W_\tau + (x^l - x^r), 0\} - c \int_{\psi(t)}^\tau \frac{2\alpha^2}{(2\sigma_0^2 - s)^2} ds \right] + x^r.$$

Define $\hat{\tau} = \tau - \psi$ and let wlog $x^l < x^r$, then

$$V(t, x^l, x^r, c, \sigma_0, \alpha) = \sup_{\hat{\tau} \geq 0} \mathbb{E} \left[\max\{W_\tau - |x^l - x^r|, 0\} - c \int_{\psi(t)}^{\psi(t) + \hat{\tau}} \frac{2\alpha^2}{(2\sigma_0^2 - s)^2} ds \right] + \max\{x^l, x^r\};$$

because the current state is a sufficient statistic for Brownian motion we have

$$V(t, x^l, x^r, c, \sigma_0, \alpha) = \sup_{\hat{\tau} \geq 0} \mathbb{E} \left[\max\{W_\tau - |x^l - x^r|, 0\} - c \int_0^{\hat{\tau}} \frac{2\alpha^2}{(2\sigma_0^2 - s - \psi(t))^2} ds \right] + \max\{x^l, x^r\}.$$

Note that for every fixed strategy τ the cost term is increasing in t and $\psi(t)$ and thus $V(t, x^l, x^r, c, \sigma_0, \alpha) - \max\{x^l, x^r\}$ is non-increasing.

Proof of 7: Note that Lemma 1 implies that for any $t < t'$

$$X_{t'}^i = X_t^i + \int_0^{t'-t} \frac{\alpha^{-1}}{(\sigma_0^{-2} + \alpha^{-2}t) + \alpha^{-2}s} dW_{t+s}^i$$

where W_s^i is a Brownian motion with respect to the filtration information of the agent. Thus, if the agent starts with a prior at time 0 equal to $\mathcal{N}(X_0, \sigma_0^2)$, then her belief at time t' is exactly the same as if she started with a prior at t equal to $\mathcal{N}(X_t, \sigma_t^2)$ where $\sigma_t^{-2} = \sigma_0^{-2} + \alpha^{-2}t$. Thus, $V(t, x^l, x^r, c, \sigma_0, \alpha) = V(0, x^l, x^r, c, \sigma_t, \alpha)$.

A.5 Proof of Theorem 4

Part 4 of the Theorem follows from simple algebra. We prove the remaining parts below. Note that part 3d is proved using Lemma 4, which relies only on parts 1–3c of Theorem 4.

1. k^* is well defined

Define the function k implicitly by $k^*(t) := \min\{x \in \mathbb{R} : 0 = V(t, -x, 0, c, \sigma_0, \alpha)\}$. To see that the set above is nonempty for all t , suppose toward contradiction that there is some t for which $V(t, -x, 0, c, \sigma_0, \alpha) > 0$ for all $x > 0$. As V nonincreasing by Lemma 2, it follows that $V(t', -x, 0, c, \sigma_0, \alpha) > 0$ for all $t' < t$.³⁰ Fix $t' < t$; this implies that the agent never stops between t' and t , which implies that he incurs a sure cost of $(t - t')c$. An upper bound for his value of continuing at t is given by part 1 of Lemma 2. But $\lim_{x \rightarrow \infty} \mathbb{E}_{(t', -x, 0, \sigma_0, \alpha)} \max\{\theta^l, \theta^r\} = 0$, a contradiction. Since V is continuous in x by part 5 of Lemma 2, the minimum is attained.

2. Characterization of the optimum by k^*

Note that due to the symmetry of the problem $V(t, x^l, x^r, c, \sigma_0, \alpha) = V(t, x^r, x^l, c, \sigma_0, \alpha)$. Without loss of generality suppose $x^l \leq x^r$. As X_t is a Markov process, the principle of optimality³¹ implies that the agent's problem admits a solution of the form $\tau = \inf\{t \geq 0 : \max_{i=l,r} X_t^i \geq V(t, X_t^l, X_t^r, c, \sigma_0, \alpha)\}$. Thus, it is optimal to stop if and only if

$$0 = V(t, x^l, x^r, c, \sigma_0, \alpha) - \max\{x^l, x^r\} = V(t, x^l, x^r, c, \sigma_0, \alpha) - x^r = V(t, x^l - x^r, 0, c, \sigma_0, \alpha).$$

As $x^l - x^r \leq 0$, V is monotone increasing in the second argument (by Lemma 2, part 4), and $V(t, x^l - x^r, 0, c, \sigma_0, \alpha) \geq 0$ we have

$$\{0 = V(t, x^l - x^r, 0, c, \sigma_0, \alpha)\} = \{x^l - x^r \leq -k^*(t)\} = \{|x^l - x^r| \geq k^*(t)\}.$$

Hence the optimal strategy equals $\tau^* = \inf\{t \geq 0 : |X_t^l - X_t^r| \geq k^*(t)\}$.

³⁰If $t = 0$, then use part 7 of Lemma 2 to shift time.

³¹Our model does not satisfy condition (2.1.1) of Peskir and Shiryaev (2006) because for some stopping times the expected payoff is $-\infty$, but as they indicate on p. 2 the proof can be extended to our case.

3a. Monotonicity

Recall that by Lemma 2 the value function V is non-increasing in t . Suppose that $t < t'$; then

$$0 = V(t, -k^*(t, c, \sigma_0, \alpha), 0, c, \sigma_0, \alpha) \geq V(t', -k^*(t, c, \sigma_0, \alpha), 0, c, \sigma_0, \alpha).$$

By Lemma 2, $V(t', -k^*(t, c, \sigma_0, \alpha), 0, c, \sigma_0, \alpha) \geq 0$, so $0 = V(t', -k^*(t, c, \sigma_0, \alpha), 0, c, \sigma_0, \alpha)$. Hence,

$$k^*(t, c, \sigma_0, \alpha) \geq \inf\{x \in \mathbb{R} : 0 = V(t', -x, 0, c, \sigma_0, \alpha)\} = k^*(t', c, \sigma_0, \alpha).$$

3b. Positivity

The payoff of the optimal decision rule is at least as high as the payoff from using the strategy that stops at time Δ for sure. Because the information gained over a short time period Δ is of order $\epsilon^{\frac{1}{2}}$ and the cost is linear, it is always worth buying some information when the expected utility of both options is the same. To see this formally, suppose that $x^l = x^r = x$, and note that

$$\begin{aligned} V(t, x, x, c, \sigma_0, \alpha) - x &= \sup_{\tau} \mathbb{E} \left[\max\{W_{\tau}, 0\} - \int_0^{\tau} \frac{2c\alpha^2}{(2\sigma_0^2 - s - \psi(t))^2} ds \right] \\ &\geq \mathbb{E} \left[\max\{W_{\epsilon}, 0\} - \int_0^{\epsilon} \frac{2c\alpha^2}{(2\sigma_0^2 - s - \psi(t))^2} ds \right] \\ &= \int_0^{\infty} z \frac{1}{\sqrt{2\pi\epsilon}} e^{-\frac{z^2}{2\epsilon}} dz - \int_0^{\epsilon} \frac{2c\alpha^2}{(2\sigma_0^2 - s - \psi(t))^2} ds \\ &\geq \sqrt{\frac{\epsilon}{2\pi}} - \frac{2c\alpha^2\epsilon}{(2\sigma_0^2 - \psi(t) - \epsilon)^2} \\ &\geq \sqrt{\frac{\epsilon}{2\pi}} - \frac{2c\alpha^2\epsilon}{(2\sigma_0^2 - \psi(t) - \tilde{\epsilon})^2} \end{aligned}$$

for all fixed $\tilde{\epsilon} \in [\epsilon, 2\sigma_0^2 - \psi(t)]$. As the first term goes to zero with the speed of square root while the second term shrinks linearly we get that $V(t, x, x, c, \sigma_0, \alpha) - \max\{x, x\} > 0$ for some small $\epsilon > 0$ and thus the agent does not stop when her posterior mean is the same on both options.

3c. Zero limit

Let $k^*(s, c, \sigma_0, \alpha) \geq K^* > 0$ for all $s \geq t$. Consider the time t history where $X_t^l = X_t^r$. The probability that the agent never stops (and thus pays infinity costs) is bounded from below by the probability that the process $X^l - X^r$ stays in the interval $[-K^*, K^*]$,

$$\mathbb{P} \left[\sup_{s \in [t, \infty)} |X_s^l - X_s^r| < k(s, c, \sigma_0, \alpha) \mid X_t^l = X_t^r \right] \geq \mathbb{P} \left[\sup_{s \in [t, \infty)} |X_s^l - X_s^r| < K^* \mid X_t^l = X_t^r \right].$$

By the time change argument used in Section A.5 this equals the probability that a Brownian motion $(W_t)_{t \in \mathbb{R}_+}$ leaves the interval $[-K, K]$ in the time from $\psi(t)$ to $2\sigma_0^2$,

$$\mathbb{P} \left[\sup_{s \in [t, \infty)} |X_s^l - X_s^r| < K^* \mid X_t^l = X_t^r \right] = \mathbb{P} \left[\sup_{s \in [\psi(t), 2\sigma_0^2]} |W_s| < K^*(s) \right].$$

This probability is non-zero. Thus, there is a positive probability the agent incurs infinite cost. Because the expected gain is bounded by the full information payoff, this is a contradiction.

3d. Lipschitz continuity of k^* in t

Let $\lambda_\epsilon = (1 + \epsilon\alpha^{-2}\sigma_0^2)^{-1/2} < 1$ and note that by definition $\lambda_\epsilon \sigma_0 = \sigma(\epsilon)$. We can thus use equations (13), (14) and (16) to get

$$k^*(\epsilon, c, \sigma_0, \alpha) = k^*(0, c, \sigma(\epsilon), \alpha) = \lambda_\epsilon k^*(0, c\lambda_\epsilon^{-3}, \sigma_0) \geq \lambda_\epsilon^4 k^*(0, c, \sigma_0, \alpha).$$

As a consequence we can bound the difference between the value of the barrier at time zero and at time ϵ from below

$$k^*(\epsilon, c, \sigma_0, \alpha) - k^*(0, c, \sigma_0, \alpha) \geq \left[(1 + \epsilon\alpha^{-2}\sigma_0^2)^{-2} - 1 \right] k^*(0, c, \sigma_0, \alpha) \geq -2\alpha^{-2}\sigma_0^2\epsilon k^*(0, c, \sigma_0, \alpha), \quad (12)$$

where the last inequality follows from convexity of the function $\epsilon \mapsto (1 + \epsilon\alpha^{-2}\sigma_0^2)^{-2} - 1$. Since k^* is nonincreasing in t , the upper bound is zero. Thus, by equation (13), inequality (12), and then equation (13) again, we have:

$$\begin{aligned} 0 &\geq k^*(t + \epsilon, c, \sigma_0, \alpha) - k^*(t, c, \sigma_0, \alpha) = k^*(\epsilon, c, \sigma_t, \alpha) - k^*(0, c, \sigma_t, \alpha) \geq -2\alpha^{-2}\sigma_t^2\epsilon k^*(0, c, \sigma_t, \alpha), \\ &= -2\alpha^{-2}\sigma_t^2\epsilon k^*(t, c, \sigma_0, \alpha) \geq -2\alpha^{-2}\sigma_0^2\epsilon k^*(t, c, \sigma_0, \alpha), \end{aligned}$$

where the last inequality follows since σ_t is decreasing in t . Hence, the function is Lipschitz with constant at most $2\alpha^{-2}\sigma_0^2 k^*(0, c, \sigma_0, \alpha)$. \square

5. k^* and b^* pointwise decreasing

$V(t, -x, 0, c, \sigma_0, \alpha)$ is non-increasing in c as the decision maker with a lower c can always use the same strategy to guarantee himself a strictly higher value. Thus

$$k^*(t, c, \sigma_0, \alpha) = \min\{x \in \mathbb{R} : 0 = V(t, -x, 0, c, \sigma_0, \alpha)\}$$

is pointwise non-increasing in c . Since $b^*(t, c, \sigma_0, \alpha) = \alpha^2\sigma_t^{-2}k^*(t, c, \sigma_0, \alpha)$ it follows that b^* is non-increasing in c . \square

A.5.1 Corollary 1: increasing c decreases stopping time and accuracy

Since b falls at every t , each sample path stops at least as soon when cost increases. Moreover, as we argued in the proof of Theorem 1 the ratio of the probability of picking l divided by the probability of picking r conditional on making a decision at time t is given by $\frac{p^l(t)}{p^r(t)} = \exp\left(\frac{4[\theta^l - \theta^r]b(t)}{\alpha}\right)$. Hence,

as $b^*(t)$ is decreasing in c , so is the probability of making the correct choice conditional on stopping at time t . \square

A.6 Additional Lemmas

Lemma 3.

$$\begin{aligned} V(0, x^l, x^r, c\lambda, \sigma_0, \alpha) &= \\ &= \lambda^{-1} \sup_{\tau'} \mathbb{E} \left[\max \left\{ \lambda x^l + \int_0^{\tau'} \frac{\alpha^{-1}}{\sigma_0^{-2} + s\lambda^{-2}\alpha^{-2}} dM_s^l, \lambda x^r + \int_0^{\tau'} \frac{\alpha^{-1}}{\sigma_0^{-2} + s\lambda^{-2}\alpha^{-2}} dM_s^r \right\} - c\tau' \right], \end{aligned}$$

where M_t^i are Brownian motions.

Proof: By definition, $V(0, x^l, x^r, c\lambda, \sigma_0, \alpha)$ equals

$$\sup_{\tau} \mathbb{E} \left[\max \left\{ x^l + \int_0^{\tau} \frac{\alpha^{-1}}{\sigma_0^{-2} + s\alpha^{-2}} dW_s^l, x^r + \int_0^{\tau} \frac{\alpha^{-1}}{\sigma_0^{-2} + s\alpha^{-2}} dW_s^r \right\} - c\lambda\tau \right]$$

and by simple algebra this equals

$$\lambda^{-1} \sup_{\tau} \mathbb{E} \left[\max \left\{ \lambda x^l + \int_0^{\tau} \frac{\lambda\alpha^{-1}}{\sigma_0^{-2} + s\alpha^{-2}} dW_{(s\lambda^2)\lambda^{-2}}^l, \lambda x^r + \int_0^{\tau} \frac{\lambda\alpha^{-1}}{\sigma_0^{-2} + s\alpha^{-2}} dW_{(s\lambda^2)\lambda^{-2}}^r \right\} - c\lambda^2\tau \right].$$

We now change the speed of time and apply Proposition 1.4 of Chapter V of [Revuz and Yor \(1999\)](#) with $C_s := s\lambda^{-2}$ and $H_s := \frac{\alpha^{-1}\lambda}{\sigma_0^{-2} + \alpha^{-2}\lambda^{-2}s}$ (pathwise to the integrals with limits τ and $\tau\lambda^2$) to get

$$\lambda^{-1} \sup_{\tau} \mathbb{E} \left[\max \left\{ \lambda x^l + \int_0^{\tau\lambda^2} \frac{\lambda\alpha^{-1}}{\sigma_0^{-2} + s\lambda^{-2}\alpha^{-2}} dW_{s\lambda^{-2}}^l, \lambda x^r + \int_0^{\tau\lambda^2} \frac{\lambda\alpha^{-1}}{\sigma_0^{-2} + s\lambda^{-2}\alpha^{-2}} dW_{s\lambda^{-2}}^r \right\} - c\lambda^2\tau \right].$$

In the next step we apply the time rescaling argument to conclude that $M_r^i := \lambda W_{r\lambda^{-2}}^i$ is a Brownian motion, and $\tau' = \tau\lambda^2$ is a stopping time measurable in the natural filtration generated by M . This yields

$$\lambda^{-1} \sup_{\tau'} \mathbb{E} \left[\max \left\{ \lambda x^l + \int_0^{\tau'} \frac{\alpha^{-1}}{\sigma_0^{-2} + s\lambda^{-2}\alpha^{-2}} dM_s^l, \lambda x^r + \int_0^{\tau'} \frac{\alpha^{-1}}{\sigma_0^{-2} + s\lambda^{-2}\alpha^{-2}} dM_s^r \right\} - c\tau' \right]$$

\square

Lemma 4. *The optimal solution $k^*(t, c, \sigma_0, \alpha)$ to problem (6) satisfies:*

$$k^*(t, c, \sigma_0, \alpha) = k^*(0, c, \sigma_t, \alpha) \text{ for all } t \geq 0 \quad (13)$$

$$k^*(0, c, \lambda\sigma_0, \alpha) = \lambda k^*(0, c\lambda^{-3}, \sigma_0, \alpha) \text{ for all } \lambda > 0 \quad (14)$$

$$k^*(t, c, \sigma_0, \lambda\alpha) = \lambda k^*(t, \lambda^{-1}c, \lambda^{-1}\sigma_0, \alpha) \text{ for all } t, \lambda > 0 \quad (15)$$

$$k^*(0, \lambda c, \sigma_0, \alpha) \geq \lambda^{-1} k^*(0, c, \sigma_0, \alpha) \text{ for all } \lambda > 0. \quad (16)$$

Proof: Equations (13) and (14) are a simple consequence of Lemma 3. Equation (15) simply follows from dividing V by α . Equation (16) follows because having more information is always better. The details of the proof can be found in the online appendix.

Lemma 5. \bar{k} is the only function that satisfies (13) – (16) with equality, and \bar{b} is the associated boundary in the signal space.

Proof: Notice that by equations (13), (15), (14) and (16) applied in that order, it follows that

$$\begin{aligned}\bar{k}(t, c, \sigma_0, \alpha) &= \bar{k}(0, c, \sigma_t, \alpha) = \alpha \bar{k}(0, \alpha^{-1}c, \alpha^{-1}\sigma_t, 1) = \sigma_t \bar{k}(0, \alpha^2 c \sigma_t^{-3}, 1, 1) \\ &= \alpha^{-2} c^{-1} \sigma_t^4 \bar{k}(0, 1, 1, 1) = \frac{\kappa}{c \alpha^2 (\sigma_0^{-2} + \alpha^{-2}t)^2},\end{aligned}$$

where $\kappa = \bar{k}(0, 1, 1, 1)$. Since $\bar{b}(t, c, \sigma_0, \alpha) = \alpha^2 \bar{k}(t, c, \sigma_0, \alpha) \sigma_t^{-2}$, it follows that $\bar{b}(t, c, \sigma_0, \alpha) = \frac{\kappa}{c(\sigma_0^{-2} + \alpha^{-2}t)}$. The fact that $\kappa = \frac{1}{2}$ follows from the proof of Proposition 4, as any other constant would result in a contradiction as $t \rightarrow \infty$. \square

Remark 1. Proposition 4 says that \bar{k} is a good approximation of k^* for large t . An intuition for why this is true is as follows: Inequality (16) becomes an equality if additional information does not have value, which is the case when the agent already learned a lot, which is the case when t is large. Thus, for large t , k^* almost satisfies (16) with equality, i.e., it is almost equal to \bar{k} .

A.7 Proof of Proposition 2

Let $\kappa := \mathbb{E}[\max\{\theta^l, \theta^r\}]$ and fix a stopping time τ . To show that

$$\mathbb{E}\left[-\mathbf{1}_{\{X_\tau^l \geq X_\tau^r\}}(\theta^r - \theta^l)^+ - \mathbf{1}_{\{X_\tau^r > X_\tau^l\}}(\theta^l - \theta^r)^+ - c\tau\right] = \mathbb{E}\left[\max\{X_\tau^l, X_\tau^r\} - c\tau\right] - \kappa,$$

the cost terms can be dropped. Let D be the difference between the expected payoff from the optimal decision and the expected payoff from choosing the correct action, $D := \mathbb{E}[\max\{X_\tau^l, X_\tau^r\}] - \mathbb{E}[\max\{\theta^l, \theta^r\}]$. By decomposing the expectation into two events,

$$D = \mathbb{E}\left[\mathbf{1}_{\{X_\tau^l \leq X_\tau^r\}}(X_\tau^l - \max\{\theta^l, \theta^r\}) + \mathbf{1}_{\{X_\tau^r < X_\tau^l\}}(X_\tau^r - \max\{\theta^l, \theta^r\})\right].$$

Plugging in the definition of X_τ^i and using the law of iterated expectations, this equals

$$\begin{aligned}&\mathbb{E}\left[\mathbf{1}_{\{X_\tau^l \leq X_\tau^r\}}(\mathbb{E}[\theta^l | \mathcal{F}_\tau] - \max\{\theta^l, \theta^r\}) + \mathbf{1}_{\{X_\tau^r < X_\tau^l\}}(\mathbb{E}[\theta^r | \mathcal{F}_\tau] - \max\{\theta^l, \theta^r\})\right] \\ &= \mathbb{E}\left[\mathbf{1}_{\{X_\tau^l \leq X_\tau^r\}}(\mathbb{E}[\theta^l | \mathcal{F}_\tau] - \mathbb{E}[\max\{\theta^l, \theta^r\} | \mathcal{F}_\tau]) + \mathbf{1}_{\{X_\tau^r < X_\tau^l\}}(\mathbb{E}[\theta^r | \mathcal{F}_\tau] - \mathbb{E}[\max\{\theta^l, \theta^r\} | \mathcal{F}_\tau])\right] \\ &= \mathbb{E}\left[\mathbf{1}_{\{X_\tau^l \leq X_\tau^r\}}\mathbb{E}[-(\theta^r - \theta^l)^+ | \mathcal{F}_\tau] + \mathbf{1}_{\{X_\tau^r < X_\tau^l\}}\mathbb{E}[-(\theta^l - \theta^r)^+ | \mathcal{F}_\tau]\right] \\ &= \mathbb{E}\left[-\mathbf{1}_{\{X_\tau^l \leq X_\tau^r\}}(\theta^r - \theta^l)^+ - \mathbf{1}_{\{X_\tau^r < X_\tau^l\}}(\theta^l - \theta^r)^+\right].\end{aligned}$$

A.8 Proof of Proposition 3

When $X_0^l = X_0^r$, $\delta =: \theta^l - \theta^r$ has a normal distribution with mean 0, so by Theorem 2 it suffices to show that $k^*(t, c, \sigma_0, \alpha)\sigma_t^{-1} = k^*(0, c, \sigma_t, \alpha)\sigma_t^{-1}$ is decreasing in t . From equation 13 we have that k^* is strictly monotone in σ_0 so the partial derivative exists almost everywhere and at the points of differentiability

$$\frac{\partial}{\partial \sigma_t} [k^*(0, c, \sigma_t, \alpha)\sigma_t^{-1}] = k_{\sigma}^*(0, c, \sigma_t, \alpha)\sigma_t^{-1} - k^*(0, c, \sigma_t, \alpha)\sigma_t^{-2}.$$

We will now show that this is equal to $-3ck_c^*(0, \sigma_t, c, \alpha)\sigma_t^2$, which is nonnegative. To see that, we show that $k_{\sigma}^*(0, c, \sigma_0, \alpha)\sigma_0 = -3ck_c^*(0, c, \sigma_0, \alpha) + k^*(0, c, \sigma_0, \alpha)$. Set

$$\beta_{\epsilon}\sigma_0 = \sigma_0 + \epsilon \Rightarrow \beta_{\epsilon} = 1 + \frac{\epsilon}{\sigma_0}.$$

Inserting in equation (14) gives

$$\begin{aligned} k^*(0, c, \sigma_0\beta_{\epsilon}, \alpha) &= k^*(0, c, \sigma_0 + \epsilon, \alpha) = \beta_{\epsilon}k^*(0, c\beta_{\epsilon}^{-3}, \sigma_0, \alpha) \\ \Leftrightarrow k^*(0, c, \sigma_0 + \epsilon, \alpha) - k^*(0, c, \sigma_0, \alpha) &= k^*(0, c\beta_{\epsilon}^{-3}, \sigma_0, \alpha) - k^*(0, c, \sigma_0, \alpha) + \frac{\epsilon}{\sigma_0}k^*(0, c\beta_{\epsilon}^{-3}, \sigma_0, \alpha). \end{aligned}$$

Dividing by ϵ and taking the limit $\epsilon \rightarrow 0$ yields

$$\begin{aligned} k_{\sigma}^*(0, c, \sigma_0, \alpha) &= k_c^*(0, c, \sigma_0, \alpha)c \left[\lim_{\epsilon \rightarrow 0} \frac{\beta_{\epsilon}^{-3} - 1}{\epsilon} \right] + \frac{1}{\sigma_0}k^*(0, c, \sigma_0, \alpha) \\ &= k_c^*(0, c, \sigma_0, \alpha)c[-3] \frac{\partial \beta_{\epsilon}}{\partial \epsilon} + \frac{1}{\sigma_0}k^*(0, c, \sigma_0, \alpha) \\ &= -k_c^*(0, c, \sigma_0, \alpha)c \frac{3}{\sigma_0} + \frac{1}{\sigma_0}k^*(0, c, \sigma_0, \alpha) \\ \Leftrightarrow k_{\sigma}^*(0, c, \sigma_0, \alpha)\sigma_0 &= -3ck_c^*(0, c, \sigma_0, \alpha) + k^*(0, c, \sigma_0, \alpha). \end{aligned}$$

Note that even $X_0^l \neq X_0^r$, the agent will still use the same boundary in belief space, which implies that his conditional probability of making the correct choice will be the same.

A.9 Proof of Theorem 5

The change in posterior variance satisfies the ordinary differential equation $\frac{d}{dt}(\sigma_t^i)^2 = -(\sigma_t^i)^4(\beta_t^i)^2$. To prove part 1, note that by the previous equation the dynamics of the posterior variance are given by

$$\frac{d}{dt}v_t = \frac{d}{dt}(\sigma_t^l)^2 + \frac{d}{dt}(\sigma_t^r)^2 = -\left\{(\sigma_t^l)^4(\beta_t^l)^2 + (\sigma_t^r)^4(\beta_t^r)^2\right\}. \quad (17)$$

We have that for any fixed (β^l, β^r) we can bound the right-hand side of (17) by using the myopically optimal policy

$$(\sigma^l)^4(\beta^l)^2 + (\sigma^r)^4(\beta^r)^2 \leq \max_{\beta \in [0, 2]} (\sigma^l)^4(\beta)^2 + (\sigma^r)^4(2 - \beta)^2 = 4 \frac{(\sigma^l)^4(\sigma^r)^8 + (\sigma^l)^8(\sigma^r)^4}{((\sigma^l)^4 + (\sigma^r)^4)^2}. \quad (18)$$

Next we argue that for any given value of v the instantaneous decrease in posterior variance can be bounded by the instantaneous decrease when the posterior variances for both options is the same. The right hand side of (18) is maximized when $v/2 = (\sigma^l)^2 = (\sigma^r)^2$

$$\begin{aligned} 4 \frac{(\sigma^l)^4(\sigma^r)^8 + (\sigma^l)^8(\sigma^r)^4}{((\sigma^l)^4 + (\sigma^r)^4)^2} &\leq \max_{(\sigma^l)^2 + (\sigma^r)^2 \leq v} 4 \frac{(\sigma^l)^4(\sigma^r)^8 + (\sigma^l)^8(\sigma^r)^4}{((\sigma^l)^4 + (\sigma^r)^4)^2} \\ &= 4 \frac{(v/2)^2(v/2)^4 + (v/2)^4(v/2)^2}{((v/2)^2 + (v/2)^2)^2} = 4 \frac{2(v/2)^6}{(2(v/2)^2)^2} = \frac{1}{2}v^2. \end{aligned}$$

We thus have that the change in posterior variance satisfies

$$\frac{d}{dt}v_t \geq - \max_{(\sigma_t^l)^2 + (\sigma_t^r)^2 \leq v_t} \max_{\beta_t^l + \beta_t^r \leq 1} \left\{ (\sigma_t^l)^4(\beta_t^l)^2 + (\sigma_t^r)^4(\beta_t^r)^2 \right\} = -\frac{1}{2}v_t^2.$$

It follows from the comparison principle for ODEs that the variance v_t is bounded from below by \tilde{v}_t for all t , where \tilde{v}_t is the solution to the ODE $\frac{d}{dt}\tilde{v}_t = -\frac{1}{2}\tilde{v}_t^2$, $\tilde{v}_0 = 2\sigma_0^2$. As this is the solution to (17) when the agent pays equal attention $\beta_t^l = \beta_t^r = 1$ and the variances equal $v_t/2 = (\sigma_t^l)^2 = (\sigma_t^r)^2$ at every point in time t it follows that the posterior variance about the difference in utilities is minimized uniformly over time by paying equal attention.

To prove part 2, wlog let $X_0^l \geq X_0^r$. Fix an attention strategy β . The optimal stopping policy τ is a solution to

$$\sup_{\tau} \mathbb{E} \left[\max\{X_{\tau}^l, X_{\tau}^r\} - c\tau \right] = \sup_{\tau} \mathbb{E} \left[\max\{X_{\tau}^l - X_{\tau}^r, 0\} - c\tau \right] + X_0^r \quad (19)$$

Define the process $y_t = -\{(\sigma_t^l)^4(\beta_t^l)^2 + (\sigma_t^r)^4(\beta_t^r)^2\}$. By the Dambis, Dubins, Schwarz Theorem (see, e.g., Theorem 1.6, chapter V of Revuz and Yor, 1999) there exists a Brownian motion $(B_{\nu})_{r \in [0, \sigma_0^2]}$ such that $X_t^l - X_t^r = B_{v_0^2 - v_t^2}$; this a time change where the new scale is the posterior variance. Furthermore, we can define the stochastic process $\phi_r := \inf\{t: v_0^2 - v_t^2 \geq r\}$. By eq. (19) the value of the agent is given by a similar argument to that in the Proof of Lemma 2.

$$\sup_{\tau} \mathbb{E} \left[\max\{X_{\tau}^l - X_{\tau}^r, 0\} - c\tau \right] + X_0^r = \sup_{\nu} \mathbb{E} \left[\max\{B_{\nu}, 0\} - c\phi_{\nu} \right] + X_0^r.$$

Denote by $(\tilde{v}_t)_{t \geq 0}$ the posterior variance process if the agent pays equal attention to the signals. As the posterior variance v_t is greater than the posterior variance if the agent pays equal attention \tilde{v}_t we have that $\phi_r \geq \tilde{\phi}_r := \inf\{t: \tilde{v}_0^2 - \tilde{v}_t^2 \geq r\}$. It follows from $\phi_r \geq \tilde{\phi}_r$ that the value when using a give attention strategy is smaller than the value achieved by paying equal attention to the signals

$$\sup_r \mathbb{E} [B_r - c\phi_r] + X_0^r \leq \sup_{\nu} \mathbb{E} [\max\{B_{\nu}, 0\} - c\phi_{\nu}] + X_0^r. \quad \square$$

References

Arrow, Kenneth, David Blackwell, and M. Girshick. 1949. "Bayes and minimax solutions of sequential decision problems." *Econometrica*, 17: 213–244.

- Ballinger, T Parker, and Nathaniel T Wilcox.** 1997. “Decisions, error and heterogeneity.” *The Economic Journal*, 107(443): 1090–1105.
- Bather, John A.** 1962. “Bayes procedures for deciding the sign of a normal mean.” Vol. 58, 599–620, Cambridge Univ Press.
- Bogacz, Rafal, Eric Brown, Jeff Moehlis, Philip Holmes, and Jonathan D Cohen.** 2006. “The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks.” *Psychological review*, 113(4): 700.
- Branco, Fernando, Monic Sun, and J Miguel Villas-Boas.** 2012. “Optimal search for product information.” *Management Science*, 58(11): 2037–2056.
- Brocas, Isabelle, Juan D Carrillo, Stephanie W Wang, and Colin F Camerer.** 2014. “Imperfect choice or imperfect attention? Understanding strategic thinking in private information games.” *The Review of Economic Studies*, rdu001.
- Bussemeyer, Jerome R, and James T Townsend.** 1992. “Fundamental derivations from decision field theory.” *Mathematical Social Sciences*, 23(3): 255–282.
- Bussemeyer, Jerome R, and James T Townsend.** 1993. “Decision field theory: a dynamic-cognitive approach to decision making in an uncertain environment.” *Psychological review*, 100(3): 432.
- Bussemeyer, Jerome R, and Joseph G Johnson.** 2004. “Computational models of decision making.” *Blackwell handbook of judgment and decision making*, 133–154.
- Camerer, Colin F.** 1989. “An experimental test of several generalized utility theories.” *Journal of Risk and uncertainty*, 2(1): 61–104.
- Caplin, Andrew, and Daniel Martin.** 2014. “The Dual-Process Drift Diffusion Model: Evidence from Response Times.”
- Caplin, Andrew, and Mark Dean.** 2011. “Search, choice, and revealed preference.” *Theoretical Economics*, 6(1): 19–48.
- Cheremukhin, Anton, Anna Popova, and Antonella Tutino.** 2011. “Experimental evidence on rational inattention.” *Federal Reserve Bank of Dallas Working Paper*, 1112.
- Chernoff, Herman.** 1961. “Sequential tests for the mean of a normal distribution.” Vol. 1, 79–91, Univ of California Press.
- Che, Yeon-Koo, and Konrad Mierendorff.** 2016. “Optimal Sequential Decision with Limited Attention.” *in preparation*.
- Churchland, Anne K, Roozbeh Kiani, and Michael N Shadlen.** 2008. “Decision-making with multiple alternatives.” *Nature neuroscience*, 11(6): 693–702.
- 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.” *mimeo*.
- Costa-Gomes, Miguel, Vincent P Crawford, and Bruno Broseta.** 2001. “Cognition and behavior in normal-form games: An experimental study.” *Econometrica*, 69(5): 1193–1235.

- Ditterich, Jochen.** 2006. “Stochastic models of decisions about motion direction: behavior and physiology.” *Neural Networks*, 19(8): 981–1012.
- Drugowitsch, Jan, Rubén Moreno-Bote, Anne K Churchland, Michael N Shadlen, and Alexandre Pouget.** 2012. “The cost of accumulating evidence in perceptual decision making.” *The Journal of Neuroscience*, 32(11): 3612–3628.
- Dudley, Richard M.** 2002. *Real analysis and probability*. Vol. 74, Cambridge University Press.
- Edwards, Ward.** 1965. “Optimal strategies for seeking information: Models for statistics, choice reaction times, and human information processing.” *Journal of Mathematical Psychology*, 2(2): 312–329.
- Fehr, Ernst, and Antonio Rangel.** 2011. “Neuroeconomic foundations of economic choice: recent advances.” *The Journal of Economic Perspectives*, 25(4): 3–30.
- Frazier, Peter, and Angela Yu.** 2007. “Sequential hypothesis testing under stochastic deadlines.” 465–472.
- Fudenberg, Drew, Philipp Strack, and Tomasz Strzalecki.** 2015. “Stochastic Choice and Optimal Sequential Sampling.” *mimeo*.
- Fudenberg, Drew, Ryota Iijima, and Tomasz Strzalecki.** 2015. “Stochastic choice and revealed perturbed utility.” *Econometrica*, 83(6): 2371–2409.
- Gabaix, Xavier, and David Laibson.** 2005. “Bounded rationality and directed cognition.” *Harvard University*.
- Hanes, Doug P, and Jeffrey D Schall.** 1996. “Neural control of voluntary movement initiation.” *Science*, 274(5286): 427–430.
- Hebert, Benjamin, and Michael Woodford.** 2016. “Rational Inattention with Sequential Information Sampling.” *mimeo*.
- Hey, John D.** 1995. “Experimental investigations of errors in decision making under risk.” *European Economic Review*, 39(3): 633–640.
- Hey, John D.** 2001. “Does repetition improve consistency?” *Experimental economics*, 4(1): 5–54.
- Johnson, Eric J, Colin Camerer, Sankar Sen, and Talia Rymon.** 2002. “Detecting failures of backward induction: Monitoring information search in sequential bargaining.” *Journal of Economic Theory*, 104(1): 16–47.
- Jones, Matt, and Ehtibar N Dzhafarov.** 2014. “Unfalsifiability and mutual translatability of major modeling schemes for choice reaction time.” *Psychological review*, 121(1): 1.
- Ke, Tony, and Miguel Villas-Boas.** 2016. “Optimal Learning before Choice.” *mimeo*.
- Ke, T, Zuo-Jun Max Shen, and J Miguel Villas-Boas.** 2013. “Search for Information on Multiple Products.” Working paper, University of California, Berkeley.
- Krajbich, Ian, and Antonio Rangel.** 2011. “Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions.” *Proceedings of the National Academy of Sciences*, 108(33): 13852–13857.

- Krajbich, Ian, Carrie Armel, and Antonio Rangel.** 2010. “Visual fixations and the computation and comparison of value in simple choice.” *Nature neuroscience*, 13(10): 1292–1298.
- Krajbich, Ian, Dingchao Lu, Colin Camerer, and Antonio Rangel.** 2012. “The attentional drift-diffusion model extends to simple purchasing decisions.” *Frontiers in psychology*, 3: 193.
- Krajbich, I, B Bartling, T Hare, and E Fehr.** 2015. “Rethinking fast and slow based on a critique of reaction-time reverse inference.” *Nature Communications*, 6(7455): 700.
- Kruse, Thomas, and Philipp Strack.** 2014. “Inverse Optimal Stopping.” *arXiv preprint arXiv:1406.0209*.
- Kruse, Thomas, and Philipp Strack.** 2015. “Optimal stopping with private information.” *forthcoming in Journal of Economic Theory*.
- Laming, Donald Richard John.** 1968. *Information theory of choice-reaction times*. Academic Press.
- Link, SW, and RA Heath.** 1975. “A sequential theory of psychological discrimination.” *Psychometrika*, 40(1): 77–105.
- Liptser, R Sh, and AN Shiryaev.** 2001. “Statistics of stochastic processes.” *Springer-Verlag*, 1.
- Luce, Robert Duncan.** 1986. *Response times*. Oxford University Press.
- Lu, Jay.** 2016. “Random choice and private information.” *Econometrica*, 84(6): 1983–2027.
- Milosavljevic, Milica, Jonathan Malmaud, Alexander Huth, Christof Koch, and Antonio Rangel.** 2010. “The drift diffusion model can account for value-based choice response times under high and low time pressure.” *Judgement & Decision Making*, 5: 437–449.
- Natenzon, Paulo.** 2013. “Random choice and learning.” *Working paper, University of Washington*.
- Peskir, Goran, and Albert Shiryaev.** 2006. *Optimal stopping and free-boundary problems*. Springer.
- Rand, David G, Joshua D Greene, and Martin A Nowak.** 2012. “Spontaneous giving and calculated greed.” *Nature*, 489(7416): 427–430.
- Rapoport, Amnon, and Graham J Burkheimer.** 1971. “Models for deferred decision making.” *Journal of Mathematical Psychology*, 8(4): 508–538.
- Ratcliff, Roger.** 1978. “A theory of memory retrieval.” *Psychological review*, 85(2): 59.
- Ratcliff, Roger, and Gail McKoon.** 2008. “The diffusion decision model: Theory and data for two-choice decision tasks.” *Neural computation*, 20(4): 873–922.
- Ratcliff, Roger, and Philip L Smith.** 2004. “A comparison of sequential sampling models for two-choice reaction time.” *Psychological review*, 111(2): 333.
- Ratcliff, Roger, Anil Cherian, and Mark Segraves.** 2003. “A comparison of macaque behavior and superior colliculus neuronal activity to predictions from models of two-choice decisions.” *Journal of neurophysiology*, 90(3): 1392–1407.

- Reutskaja, Elena, Rosemarie Nagel, Colin F Camerer, and Antonio Rangel.** 2011. “Search dynamics in consumer choice under time pressure: An eye-tracking study.” *The American Economic Review*, 101(2): 900–926.
- Revuz, Daniel, and Marc Yor.** 1999. *Continuous martingales and Brownian motion*. Vol. 293, Springer.
- Roe, Robert M, Jermone R Busemeyer, and James T Townsend.** 2001. “Multialternative decision field theory: A dynamic connectionist model of decision making.” *Psychological review*, 108(2): 370.
- Rubinstein, Ariel.** 2007. “Instinctive and cognitive reasoning: A study of response times.” *The Economic Journal*, 117(523): 1243–1259.
- Shadlen, Michael N, and Roozbeh Kiani.** 2013. “Decision Making as a Window on Cognition.” *Neuron*, 80(3): 791–806.
- Shadlen, Michael N, Timothy D Hanks, Anne K Churchland, Roozbeh Kiani, and Tianming Yang.** 2006. “The speed and accuracy of a simple perceptual decision: a mathematical primer.” *Bayesian brain: Probabilistic approaches to neural coding*, 209–37.
- Shiryayev, Albert N.** 1969. *Optimal stopping rules*. Russian Edition published by Nauka.
- Shiryayev, Albert N.** 2007. *Optimal stopping rules*. Springer.
- Stone, Mervyn.** 1960. “Models for choice-reaction time.” *Psychometrika*, 25(3): 251–260.
- Swensson, Richard G.** 1972. “The elusive tradeoff: Speed vs accuracy in visual discrimination tasks.” *Perception & Psychophysics*, 12(1): 16–32.
- Vickers, D.** 1970. “Evidence for an accumulator model of psychophysical discrimination.” *Ergonomics*, 13(1): 37–58.
- Vul, Edward, Noah Goodman, Thomas L Griffiths, and Joshua B Tenenbaum.** 2014. “One and done? Optimal decisions from very few samples.” *Cognitive science*, 38(4): 599–637.
- Wald, Abraham.** 1947. *Sequential analysis*. John Wiley & Sons.
- Webb, Ryan.** 2013. “Dynamic constraints on the distribution of stochastic choice: Drift Diffusion implies Random Utility.” *Available at SSRN 2226018*.
- Woodford, Michael.** 2014. “An Optimizing Neuroeconomic Model of Discrete Choice.” National Bureau of Economic Research.