Speed, Accuracy, and the Optimal Timing of Choices[*]

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

This version: January 17, 2018, 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.

[†]Department of Economics, MIT
[‡]Department of Economics, University of California, Berkeley
[§]Department of Economics, Harvard University

1   Introduction

This paper studies the question of when early decisions are more likely to be correct in the classical setting of sequential information acquisition. We find that on average agents are more likely to be correct when they decide sooner, even though when agents are forced decide later they have more information. This mechanical effect is offset by two selection effects that come from the agent's choice of stopping time: First, in a fixed decision problem, when the agent receives precise information the value of acquiring further information is low, and in this case the agent is both quick and accurate. Second, the agent is more likely to receive precise information in easy problems. As we show, the first selection effect is not always sufficient to outweigh the mechanical effect of increased information, but the two selection effects combined always do so, leading to *decreasing accuracy* over time.

We believe that these effects are present in many economic choice problems, but we focus here on laboratory experiments where decision time can be carefully measured. In these experiments, individuals are repeatedly faced with the same choice set, and 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.

We restrict attention to the binary choice tasks that have been used in most neuroscience choice experiments, and suppose that the agent needs to choose between two items that we call left (l) and right (r), and can decide how long to think about the choice. In this setting, we can ask how the probability of the correct choice varies with the time taken to make the decision. Many such choice experiments find that 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

---

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

[2]The 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).

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 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 is uncertain about the utilities $\theta = (\theta^l, \theta^r)$ of the two choices; this allows her to learn not only which alternative is better, but also by how much. We further assume that the agent believes θ is bivariate normal.[4] 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 of a given agent 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 $1/t$.[5] Using the optimality of the stopping rule, we show that the optimal boundary is pointwise non-

---

[3]The 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).

[4]Typically problems of sequential information acquisition cannot be solved in full generality, and instead various sorts of functional forms and special assumptions are used. Many papers assume that there are only two possible states of the world and/or fully revealing signals.

[5]The 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. Time-dependent stopping thresholds also arise if the cost, signal structure, or utility functions are time-dependent or if there is a fixed terminal date, see e.g. Rapoport and Burkheimer (1971), Drugowitsch et al. (2012) and Frazier and Yu (2007).

increasing in the cost so that increasing cost decreases both accuracy and decision time. We also show that an analyst who is uncertain of each agent's true utilities and so aggregates data across many utility realizations from that agent 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 essentially constant. Our final theoretical result shows that when the flow cost can vary arbitrarily with time, optimal stopping imposes essentially no restrictions on the observed choice process.

To relate our results to the empirical literature, we show that the individual-level barrier of any DDM can be uniquely identified from the data, and show how to estimate the model using simulations of the distribution of stopping times. We also compare the fit of the optimal boundary with those of the contant and exponential functional forms.

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, cross-individual variation in speed and accuracy on cognitive reflection tasks (Frederick, 2005) has been widely used to classify individuals as "impulsive" or "reflective," and more recently Baron et al. (2015) and Johnson, Tubau and De Neys (2016) show that across individuals on a given task, people who respond more quickly are less likely to be correct. This pattern is predicted by our model if the impulsive agents have much higher sampling costs than the reflective ones.[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

---

[6]In 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).

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 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 in footnote 5, 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.

---

[7]Wald (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).

[8]Stone (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.

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 *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 $i \in A$.

We let $p(t)$ denote the probability of making the correct choice conditional on stopping at time t. In perceptual decision tasks, 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. Moreover, when we estimate our uncertain-difference model in Section 4, we will make the stronger assumption that these scores correspond to utility levels.[9] [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]

---

[9]In those experiments the rating task is done before the choice tasks. The literature finds that 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).

[10]In 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.

[11]The online appendix considers two additional, closely related, measures of how accuracy varies with time.

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 $2\alpha^2$:

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

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 $b : \mathbb{R}_+ \to \mathbb{R}$, 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

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

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

---

[12]Recall 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.

[13]There 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.

[14]Note 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.

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 correct decision. Endogenous stopping when the signal hits the boundary 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 proof of this theorem is in the Appendix, as the proofs of all of the following results and claims except where noted. 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

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

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

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

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. Section 4.3 shows how to use (4) to identify the boundary from the data.

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

---

[15]Fudenberg, 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.

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 of the same individual 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 $0 < t < t'$*

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

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.*

---

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

3 Optimal Stopping

3.1 Statement of the model

Both the simple DDM used to explain data from perception tasks and our uncertain-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, so that

$$dZ_t^i = \theta^i dt + \alpha dB_t^i \tag{5}$$

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 $\{Z_s^i\}_{0 \le s < t}$. We denote the agent's posterior belief about θ given this information by μ_t. Let $X_t^i = \mathbb{E}_{\mu_t} \theta^i = \mathbb{E}[\theta^i \mid \mathcal{F}_t]$ 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 $c > 0$.[18] (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.

---

[17] This 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.

[18] In our model only the expected difference in utility between the options matters, and not the absolute level. In many models the cost of delay takes the form of discounting. This is not a good assumption for the lab experiments we discuss, where rewards are received at the end of the session, and each decision takes only a few seconds. The discounting specification would make the problem harder, because the stopping boundary will depend on beliefs about the levels of the two options as well as on their difference. Our intuition is that this should lead to earlier stopping when the expected levels of both choices are high.

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 $\tau : \Omega \to [0, +\infty]$ 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

$$\max_{\tau \in \mathcal{T}} \mathbb{E}\left[\max_{i=l,r} X_\tau^i - c\tau\right]^{19}. \tag{6}$$

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 is

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

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*

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

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

The proof is in the Online Appendix. 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 $\mathbb{E}\left[-\mathbf{1}_{\{X_\tau^l \geq X_\tau^r\}}(\theta^r - \theta^l)^+ - \mathbf{1}_{\{X_\tau^r > X_\tau^l\}}(\theta^l - \theta^r)^+\right]$.

---

[19]Following 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.

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_H, θ_L) and (θ_L, θ_H), where $\theta_H > \theta_L$. The agent receives payoff θ_H for choosing l in state (θ_H, θ_L) or r in state (θ', θ_H) and θ_L for choosing r in state (θ_H, θ_L) or r in state (θ_L, θ_H), so she knows that the magnitude of the utility difference between the two choices is $|\theta_H - \theta_L|$, but doesn't know which action is better. We let μ_0 denote the agent's prior probability of (θ_H, θ_L).

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:*

$$\hat{\tau} = \inf\{t \geq 0 \colon |Z_t| \geq b\}.\text{[20]}$$

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 agents 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.

---

[20]This 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.

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

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

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:

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

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

1. *k^* is well defined.*

2. *Let τ^* be the minimal optimal strategy in (6). Then*

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

3. *$k^*(t, c, \sigma_0, \alpha)$ is strictly positive, strictly decreasing in t, and $\lim_{t \to \infty} k^*(t, c, \sigma_0, \alpha) = 0$. 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*

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

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.

To gain intuition for why k^* is decreasing and asymptotes to 0, 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 $\left(\mathbb{E}_t \max_{i=l,r} X_{t+dt}^i - \max_{i=l,r} X_t^i \right) - c dt$. 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.[21] Part 5 says that k^*

---

[21]In the certain difference DDM, the agent believes she will stop in finite time with probability 1 even though the boundaries are constant, because the agent knows that the absolute value of the drift of Z_t is bounded away from 0.

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 $Z_t := Z_t^l - Z_t^r$.[22] 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 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 $\lambda > 0$. Bather (1962) shows that the stopping boundary converges to 0 and asymptotically 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. Drugowitsch et al. (2012) use numerical methods to argue that the stopping boundaries converge to 0, and to relate accuracy to the stopping boundary, which is an analog of our equation (4) for this model. In our model there are two reasons for the boundary to decline over time: as data accumulate additional data is less likely to lead to much change in beliefs, and if a decision hasn't been made after a long time the two choices are probably about as good. When the utility difference is known only the first of these reasons applies, so we expect the boundary to decrease more slowly; we explore this intuition in Fudenberg et al. (2017).

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 3 of the appendix. The conditions provide useful information about the identification of the parameters of the model,

---

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

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.

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.*[23]

Proposition 3 implies that the analyst will observe decreasing accuracy when a single fixed agent faces a series of decisions with utilities (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. $X_0^l \neq X_0^r$.[24] This implies that as long as the prior is correct, decreasing accuracy will hold for the average P for each agent 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, 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, though agents with very different costs should be expected to behave differently.

3.4 Approximations for large t and c

To gain more insight into the form of the optimal policy, we study an approximation $\bar{b}$ to the optimal boundary b^* that has a simple and tractable functional form. Using the results of Bather (1962) on the asymptotic behavior of the we show that $\bar{b}$ approximates the solution well for large t, and that the boundary is approximately constant when c is large. Let

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

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

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

---

[23]Note 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.

[24]Here P does not admit an average DDM representation as the upper and lower barrier are not symmetric.

Bather's result assumes that the agent has an improper prior with 0 prior precision. Our proof of Proposition 4 (in the Online Appendix) uses several of the rescaling arguments in Lemma O.2 (in the Online Appendix) to adapt his result to our setting. One useful implication of Proposition 4 is that b^* asymptotically declines to zero at rate $1/t$.

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 α, σ_0 there is a constant η independent of c such that*

$$\left| b^*(0, c, \sigma_0, \alpha) - \frac{1}{2\,c} \right| \le \frac{\eta}{c^{4/3}}$$

3.5   Endogenously Divided Attention

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

$$dZ_t^i = \left(\beta_t^i \right)^{1/2} \theta^i dt + dB_t^i.$$

We assume that attention levels are always positive $\beta_t^l, \beta_t^r \ge 0$, and that the total amount of attention is bounded by two, i.e. $\beta_t^l + \beta_t^r \le 2$.[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 $t \ge 0$.

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

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

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., $v_t := (\sigma_t^l)^2 + (\sigma_t^r)^2$.

---

[25] The 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.

Theorem 5 (Endogenous Attention).

1. *The posterior variance v_t at every point in time $t \geq 0$ is minimized by setting $\beta_s^r = \beta_s^l = 1$ at every point in time s.*

2. *It is optimal to set $\beta_t^l = \beta_t^r = 1$ 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 shows that the equal attention policy minimizes the posterior variance at every point in time and thus maximizes the speed of learning. A more general result in Section 2 of the Online Appendix shows that its conclusion extends as stated to the case where the signal technology is $dZ_t^i = (\beta_t^i)^{\gamma/2} \theta^i dt + dB_t^i$ for $0 < \gamma < 1$. The proof also shows that when $\gamma > 1$, the most efficient way for the agent to allocate equal attention to each signal is to oscillate or "chatter" very quickly between them; the limit policy is again equal attention to each signal over any interval of time.

A more realistic model might add an adjustment cost for changing attention. In some cases it is also more natural to assume that the agent can only pay attention to one signal at a time, as in a number of recent papers on related optimal stopping problems. Liang, Mu and Syrgkanis (2017) study the choice between signals when only one-signal can be observed at a time, allowing for multiple dimensions of normally-distributed uncertainty and correlated normal signals. They characterize the relationship between the dynamically optimal attention policy and the myopic one that always selects the signal that maximizes the expected payoff from making a choice immediately.

Che and Mierendorff (2016) study a model with a known utility difference between the two choices and perfectly revealing Poisson signals. 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. Hebert and Woodford (2016) combine optimal stopping and optimal information gathering when the agent can choose from a very general class of signals, and relate the solution to that of a static rational-inattention problem.[26]

---

[26]Woodford (2014) is also related, but instead of allowing optimal stopping constrains the stopping rule to have constant boundaries.

3.6 Non-Linear Cost

In deriving the DDM representation, 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. 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, whose proof is in the Online Appendix, 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.[27] 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 $G \subseteq R_+$ there exists a cost function $d : R_+ \to R$ such that b is optimal in the set of stopping times T that stop in G with probability one*

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

4   Empirical Analysis

In this section we relate our theoretical results to data from Krajbich, Armel and Rangel (2010); we thank them for sharing it with us. In this experiment, 39 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.[28] Before making their choices, subjects entered liking ratings for 70 different foods on a scale from -10 to +10.

---

[27]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. The proof of this result is based on Kruse and Strack (2015, 2014).

[28]There were 100 trials per subject. Food 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-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.

First we report a test of whether in individual-level data the boundary is constant, and then the results of fitting the exact model to data on each individual using a numerical computation of the optimal stopping boundary. We also compare the relative fit of an approximately optimal boundary and the exponential boundary used in Milosavljevic et al. (2010). Section 4.3 shows that the parameters are point identified.

4.1 Individual Level Analysis of the Slope of the Boundary

Here we show how to estimate the slope of subjects boundary under different functional form assumptions. First, we look at the functional form $\tilde{b}(t) = \frac{1}{g+h\,t}$ 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 used Monte Carlo simulations with 1 million random paths to compute the distribution of hitting times, and then combined this with formula (4) for the choice probabilities as a function of stopping time to compute the likelihood function and found the maximum using a gradient descent algorithm, both for 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.[29]

The same methodology can be applied to any functional form of the boundary. For example, we look at the two-parameter exponentially decreasing boundary $\check{b}(t) = g\exp(-ht)$, which was used by Milosavljevic et al. (2010) to fit to data from 8 subjects. We fit their model to the data of Krajbich, Armel and Rangel (2010), and found that $\tilde{b}$ fits better (had higher likelihood) for 30 of the 39 subjects. Fitting these two functional forms has exactly the same computational cost; see online appendix for the details.

---

[29]The simulations took about a week to run in parallel on a cluster. The supplementary appendix has a detailed description of our numerical methods.
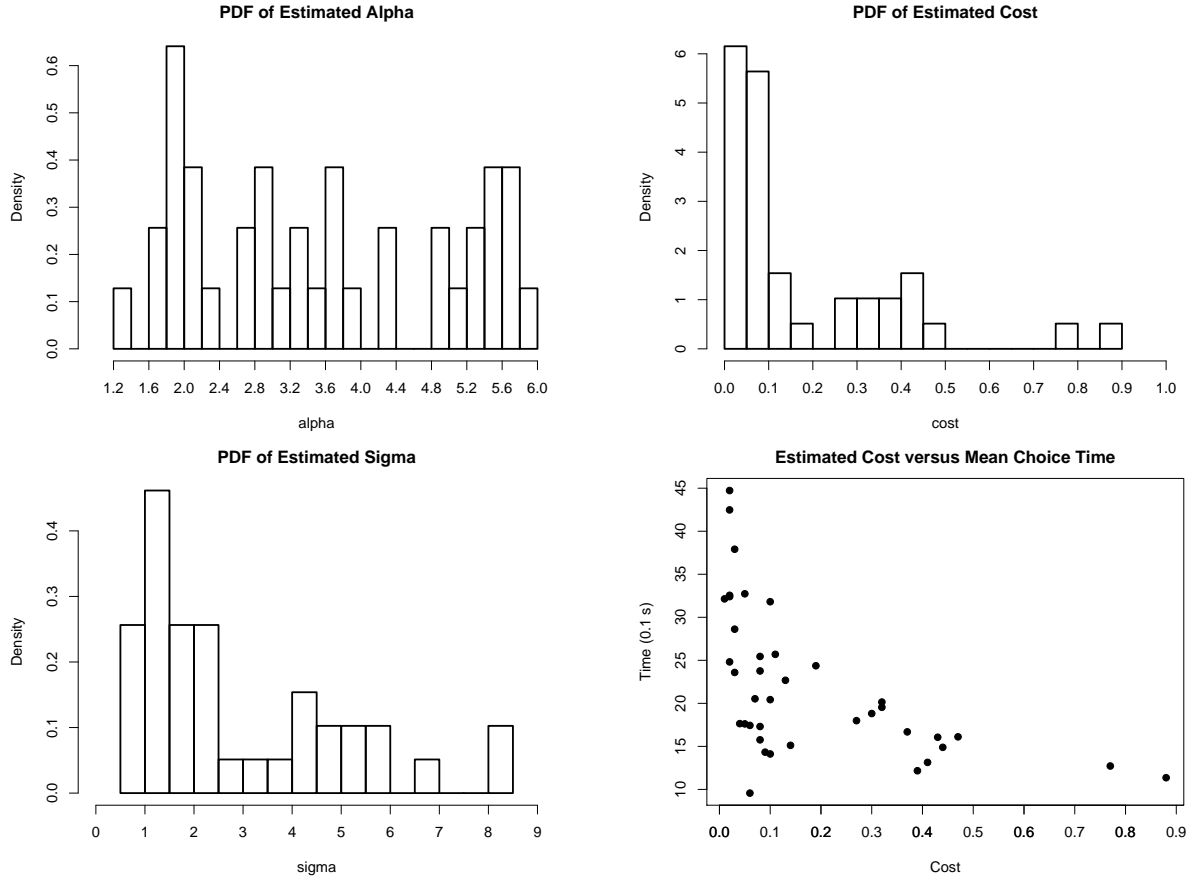
Figure 1: Marginal distributions of α, c, and σ_0 and the correlation between the average stopping time for each subject and their estimated cost c.

4.2 Individual Level Analysis using the Exact Boundary

Finally, we estimated the exact model at the individual level using the exact boundary. The additional computational step here is numerically solving for the optimal boundary as a function of the parameters. But since the simulations need to be done many times for each parameter configuration, and the computation only needs to be done for a single parameter configuration (as shown by Lemma O.4 in the online appendix, which follows from Lemma 3 in the appendix), the difference in computational cost was negligible.[30] Here we find that there is substantial heterogeneity between the subjects. Figure 1 presents the histograms of

---

[30]We computed the boundary by imposing a large finite terminal time, discretizing time and space on a fine grid, and solving backwards. Computations in Matlab and Python yielded numerically identical results. This computation takes less than two hours on a (vintage 2014) laptop, which is a negligible computation cost compared to the Monte Carlo simulations of the distribution of stopping times.

the marginal distributions of α, c, and σ_0 across all subjects.

Figure 1 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 agen's 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 2, subject 45). However, our estimation suggests that for all but two subjects the optimal boundary which best fits their choice is monotone decreasing (see Figure 2). Theorem 1 then implies that those subjects display decreasing accuracy.

4.3 Identification

For the estimation results to have meaning, the model must be identified. The following proposition shows that this is the case for a general boundary and in particular for the three kinds of boundaries we estimated.

Proposition 6. *Fix δ and suppose that the analyst observes a DDM choice process P. There exists unique $\alpha > 0$ and a unique function b such that $P = P(\delta, \alpha, b)$. In addition,*

1. *If $b(t) = \tilde{b}(t; g, h) = \frac{1}{g+ht}$ for some $g, h \geq 0$, then g and h are unique.*

2. *If $b(t) = \hat{b}(t; g, h) = g \exp(-ht)$ for some $g, h \geq 0$, then g and h are unique.*

3. *If $b(t) = b^*(t; \sigma_0, c, \alpha)$ for some $c, \sigma_0 \geq 0$, then c and σ_0 are unique.*

The proof is in the Online Appendix. In outline, the ratio of choice probabilities identifies $\frac{b(t)}{\alpha^2}$, which lets us use the expected stopping time to identify α (closed form expressions for α and b can be found in the proof). This makes the proofs of clauses 1 and 2 relatively straightforward; we use our asymptotic characterization of the optimal boundary to prove clause 3. In our estimation and the above identification result we assume that the analyst knows δ, and in particular that it equals the difference in the numerical ratings. Including an additional parameter that multiplies that difference leads to a model that is not identified.
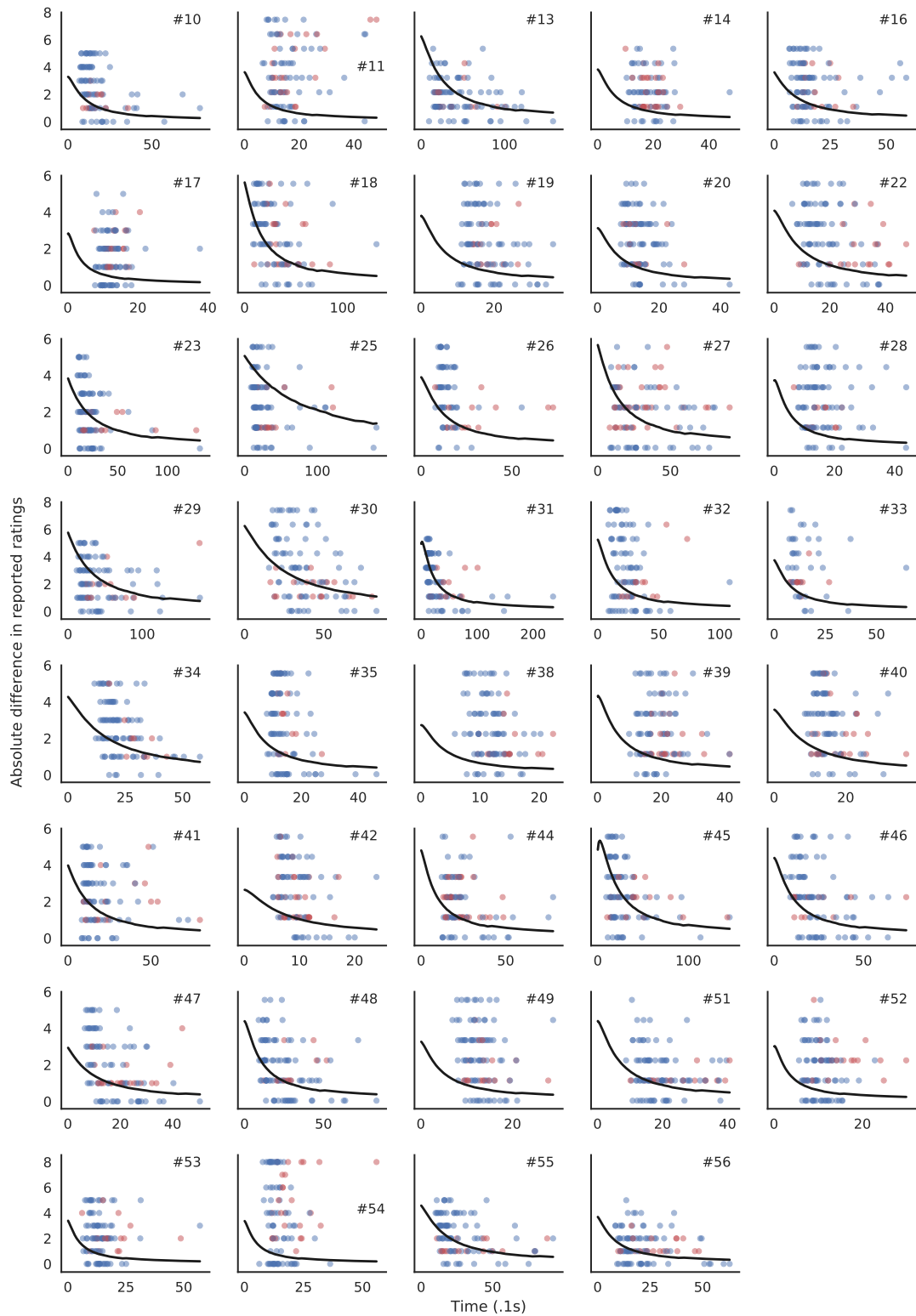
22

Figure 2: Estimated optimal boundaries for different subjects. Blue dots are correct decisions, red dots are incorrect ones.

23

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 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. Indeed, subsequent to our paper, Bhui (2017) ran an experiment more closely designed to fit our model. In his Experiment 2, each subject faces a screen with some dots moving to the right, some moving to the left, and some moving randomly. The subject has to decide whether more are moving to the right or to the left. The payoff is proportional to the number of dots moving in the chosen direction (the coherence), so here the right choice is observable by the analyst, as is the payoff difference, in contrast to the data we use, where we have to assume that reported scores are the same as utility. Consistently with our model, he finds that quicker decisions are more accurate. He also finds that our model fits his data better than does the classic model with constant boundaries.

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

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

$$= \frac{\partial}{\partial y}\mathbb{P}\left[\frac{B_t}{\sqrt{t}} \leq \frac{y - \delta t}{\alpha\sqrt{2}\sqrt{t}}\right] = \phi\left(\frac{y - \delta t}{\alpha\sqrt{2t}}\right)$$

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

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

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

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

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

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.[31] 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

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

Thus, by (8) and (9) we have that $\frac{p^l(t)}{p^r(t)} = \frac{g(t,b(t))}{g(t,-b(t))} = \exp\left(\frac{\delta b(t)}{\alpha^2}\right)$.

---

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

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.

$$\mathbb{P}\left[|\delta| < d \mid \tau = t\right] = \mathbb{P}\left[\theta^l - \theta^r \in (-d, d) \mid \tau = t, |X_\tau^l - X_\tau^r| = k(\tau)\right]$$

$$= \frac{1}{\sqrt{2\pi}} \int_{\frac{k(t)-d}{\sqrt{2}\sigma_t}}^{\frac{k(t)+d}{\sqrt{2}\sigma_t}} e^{-y^2/2} dy = \frac{1}{\sqrt{2\pi}} \int_{\beta(t)-\frac{d}{\sqrt{2}\sigma_t}}^{\beta(t)+\frac{d}{\sqrt{2}\sigma_t}} e^{-y^2/2} dy \qquad (10)$$

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) increases in t:

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

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

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

$$\mathbb{P}[\theta^l < \theta^r \mid X_t^l - X_t^r = k(t)] = \mathbb{P}\left[(\theta^l - \theta^r) - (X_t^l - X_t^r) \leq -k(t) \,\Big|\, X_t^l - X_t^r = k^*(t)\right]$$

$$= \mathbb{P}\left[\frac{(\theta^l - \theta^r) - (X_t^l - X_t^r)}{\sqrt{2}\sigma_t} \leq -\frac{k(t)}{\sqrt{2}\sigma_t} \,\Big|\, X_t^l - X_t^r = k(t)\right] = \Phi\left(-\frac{1}{\sqrt{2}}k(t)\sigma_t^{-1}\right).$$

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 $t > 0$*

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

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 $A = 1, B = \alpha$. □

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

$$\mathbb{E}\left[\mathbf{1}_{\{\tau^l \leq \tau^r\}} X_{\tau^l}^l + \mathbf{1}_{\{\tau^l > \tau^r\}} X_{\tau^r}^r - c\left(\min\{\tau^l, \tau^r\} - t\right) \mid X_t^l = x^l - k, X_t^r = x^r - k\right]$$

$$= \mathbb{E}\left[\mathbf{1}_{\{\tau^l \leq \tau^r\}} \left(\int_t^{\tau^l} \frac{\alpha^{-1}}{\sigma_0^{-2} + s\alpha^{-2}} dW_s^l + x^l - k\right) + \mathbf{1}_{\{\tau^l > \tau^r\}} \left(\int_t^{\tau^r} \frac{\alpha^{-1}}{\sigma_0^{-2} + s\alpha^{-2}} dW_s^r + x^r - k\right)\right.$$

$$\left. - c\left(\min\{\tau^l, \tau^r\} - t\right)\right]$$

$$= \mathbb{E}\left[\mathbf{1}_{\{\tau^l \leq \tau^r\}} X_{\tau^l}^l + \mathbf{1}_{\{\tau^l > \tau^r\}} X_{\tau^r}^r - c\left(\min\{\tau^l, \tau^r\} - t\right) \mid X_t^l = x^l, X_t^r = x^r\right] - k.$$

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

$$\mathbb{E}\left[\mathbf{1}_{\{\tau^l \leq \tau^r\}} X_{\tau^l}^l + \mathbf{1}_{\{\tau^l > \tau^r\}} X_{\tau^r}^r - c\left(\min\{\tau^l, \tau^r\} - t\right) \mid X_t^l = x^l, X_t^r = x^r\right] - x^l$$

$$= \mathbb{E}\left[\mathbf{1}_{\{\tau^l \leq \tau^r\}} (X_{\tau^l}^l - x^l) + \mathbf{1}_{\{\tau^l > \tau^r\}} (X_{\tau^r}^r - x^l) - c\left(\min\{\tau^l, \tau^r\} - t\right) \mid X_t^l = x^l, X_t^r = x^r\right]$$

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

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

$$\mathbb{E}\left[\mathbf{1}_{\{\tau^l \le \tau^r\}} X^l_{\tau^l} + \mathbf{1}_{\{\tau^l > \tau^r\}} X^r_{\tau^r} - c\left(\min\{\tau^l, \tau^r\} - t\right) \mid X^l_t = x^l, X^r_t = x^r\right]$$

$$= \mathbb{E}\left[\mathbf{1}_{\{\tau^l \le \tau^r\}} \left(\int_t^{\tau^l} \frac{\alpha^{-1}}{\sigma_0^{-2} + s\alpha^{-2}} dW^l_s + x^l\right) + \mathbf{1}_{\{\tau^l > \tau^r\}} X^r_{\tau^r} - c\left(\min\{\tau^l, \tau^r\} - t\right) \mid X^l_t = x^l, X^r_t = x^r\right]$$

$$= \mathbb{E}\left[\mathbf{1}_{\{\tau^l \le \tau^r\}} \int_t^{\tau^l} \frac{\alpha^{-1}}{\sigma_0^{-2} + s\alpha^{-2}} dW^l_s + \mathbf{1}_{\{\tau^l > \tau^r\}} X^r_{\tau^r} - c\left(\min\{\tau^l, \tau^r\} - t\right) \mid X^l_t = x^l, X^r_t = x^r\right]$$

$$+ x^l \mathbb{E}\left[\mathbf{1}_{\{\tau^l \le \tau^r\}} \mid X^l_t = x^l, X^r_t = x^r\right],$$

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

$$|V(0, x^l, x^r, c, \sigma_0, \alpha) - V(0, y^l, x^r, c, \sigma_0, \alpha)|$$

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

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

$$\le \left| \sup_\tau \mathbb{E}\left[\max\{x^l + \int_0^\tau \frac{\alpha^{-1}}{\sigma_0^{-2} + s\alpha^{-2}} dW^1_s, x^r + \int_0^\tau \frac{\alpha^{-1}}{\sigma_0^{-2} + s\alpha^{-2}} dW^2_s\}\right.\right.$$

$$\left.\left. - \max\{y^l + \int_0^\tau \frac{\alpha^{-1}}{\sigma_0^{-2} + s\alpha^{-2}} dW^1_s, x^r + \int_0^\tau \frac{\alpha^{-1}}{\sigma_0^{-2} + s\alpha^{-2}} dW^2_s\}\right]\right| \le |y^l - x^l|.$$

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

$$\mathbb{E}\left[\max\{X^l_\tau, X^r_\tau\} - c\tau \mid X_t = (x^l, x^r)\right] = \mathbb{E}\left[\max\{X^l_\tau - X^r_\tau, 0\} + X^r_\tau - c\tau \mid X_t = (x^l, x^r)\right]$$

$$= \mathbb{E}\left[\max\{X^l_\tau - X^r_\tau, 0\} - c\tau \mid X_t = (x^l, x^r)\right] + x^r_t.$$

Define the process $X_t := X^l_t - X^r_t$, and note that

$$X_t = X^l_t - X^r_t = X^l_0 - X^r_0 + \int_0^t \frac{\alpha^{-1}}{\sigma_0^{-2} + s\alpha^{-2}} \left(dW^l_s - dW^r_s\right) = X^l_0 - X^r_0 + \int_0^t \frac{\sqrt{2}\alpha^{-1}}{\sigma_0^{-2} + s\alpha^{-2}} d\tilde{W}_s, \qquad (11)$$

where $\tilde{W}$ is a Brownian motion. Define a time change as follows: Let $q(k)$ solve $k = \int_0^{q(k)} \left(\frac{\sqrt{2}\alpha^{-1}}{\sigma_0^{-2}+s\alpha^{-2}}\right)^2 \mathrm{d}s$.
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

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

$$= \sup_{\tau\geq\psi(t)} \mathbb{E}\left[\max\{W_\tau,0\} - c\left(\int_{\psi(t)}^\tau q'(s)\mathrm{d}s\right) \mid W_{\psi(t)} = x^l - x^r\right] + x^r,$$

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

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

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

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

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

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

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

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 $t < t'$

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

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 3, which relies only on parts 1–3c of Theorem 4.

1. k^* is well defined

Define the function k implicitly by $k^*(t) := \min\{x \in \mathbb{R} \colon 0 = V(t, -x, 0, c, \sigma_0, \alpha)\}$. 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 $t' < t$.[32] 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 $\lim_{x \to \infty} \mathbb{E}_{(t', -x, 0, \sigma_0, \alpha)} \max\{\theta^l, \theta^r\} = 0$, 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 optimality[33] implies that the agent's problem admits a solution of the form $\tau = \inf\{t \geq 0 \colon \max_{i=l,r} X_t^i \geq V(t, X_t^l, X_t^r, c, \sigma_0, \alpha)\}$. 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

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

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

3a. Monotonicity

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

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

---

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

[33]Our 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.

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

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

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 $\lambda_\epsilon = \left(1 + \epsilon \alpha^{-2} \sigma_0^2\right)^{-1/2} < 1$ 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

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

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

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

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

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

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

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. $\qquad \square$

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. $\qquad \square$

A.6   Additional Lemma

Lemma 3. *The optimal solution $k^*(t, c, \sigma_0, \alpha)$ to problem* (6) *satisfies:*

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

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

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

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

Proof: Equations (13) and (14) are a simple consequence of the time-rescaling Lemma O.2 in the Online Appendix. Equation (15) simply follows from dividing V by α. Equation (16) follows because having more information is always better. The details can be found in the online appendix. $\square$

A.7   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 $\frac{\partial}{\partial\sigma_t}\left[k^*(0, c, \sigma_t, \alpha)\sigma_t^{-1}\right] = k_\sigma^*(0, c, \sigma_t, \alpha)\sigma_t^{-1} - k^*(0, c, \sigma_t, \alpha)\sigma_t^{-2}$. We will now show that this is equal to $-3\,c\,k_c^*(0, \sigma_t, c, \alpha)\sigma_t^2$, which is nonnegative. To see that, we show that $k_\sigma^*(0, c, \sigma_0, \alpha)\sigma_0 = -3\,c\,k_c^*(0, c, \sigma_0, \alpha) + k^*(0, c, \sigma_0, \alpha)$. Set $\beta_\epsilon\sigma_0 = \sigma_0 + \epsilon \Rightarrow \beta_\epsilon = 1 + \frac{\epsilon}{\sigma_0}$. Inserting in equation (14) gives

$$k^*(0, c, \sigma_0\beta_\epsilon, \alpha) = k^*(0, c, \sigma_0 + \epsilon, \alpha) = \beta_\epsilon k^*(0, c\beta_\epsilon^{-3}, \sigma_0, \alpha)$$

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

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

$$k_\sigma^*(0, c, \sigma_0, \alpha) = k_c^*(0, c, \sigma_0, \alpha)c\left[\lim_{\epsilon \to 0}\frac{\beta_\epsilon^{-3} - 1}{\epsilon}\right] + \frac{1}{\sigma_0}k^*(0, c, \sigma_0, \alpha)$$

$$= k_c^*(0, c, \sigma_0, \alpha)c\,[-3]\frac{\partial\beta_\epsilon}{\partial\epsilon} + \frac{1}{\sigma_0}k^*(0, c, \sigma_0, \alpha)$$

$$= -k_c^*(0, c, \sigma_0, \alpha)c\frac{3}{\sigma_0} + \frac{1}{\sigma_0}k^*(0, c, \sigma_0, \alpha)$$

$$\Leftrightarrow k_\sigma^*(0, c, \sigma_0, \alpha)\sigma_0 = -3\,c\,k_c^*(0, c, \sigma_0, \alpha) + k^*(0, c, \sigma_0, \alpha).$$

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.8 Proof of Theorem 5

The posterior variance on the difference is $v_t = \left(\sigma_t^l\right)^2 + \left(\sigma_t^r\right)^2 = \frac{1}{\left(\sigma_0^i\right)^{-2} + y_t^l} + \frac{1}{\left(\sigma_0^i\right)^{-2} + y_t^r}$, where $y_t^i = \int_0^t \left(\beta_s^i\right) \mathrm{d}s$ measures the total attention the agent has spent on alternative $i \in \{l, r\}$.

We first consider the auxiliary problem of minimizing the posterior variance at some *fixed time t*. At each point in time s the agent optimally exhausts the total attention budget of two, $\beta_t^s + \beta_s^r = 2$. We claim that the agent can minimize the posterior variance v_t by paying equal attention to the two signals $\beta_t^l = \beta_t^r = 1$. To see this, note that, $y_t^l + y_t^r$ is bounded by $2t$

$$y_t^l + y_t^r \leq \int_0^t \max_{\beta^l, \beta^r, \beta^l + \beta^r \leq 2} \beta^l + \beta^r \mathrm{d}s = \int_0^t 2\mathrm{d}s = 2t.$$

Hence, we consider the the problem

$$\min_{(y_t^l, y_t^r)} \frac{1}{\left(\sigma_0^i\right)^{-2} + y_t^l} + \frac{1}{\left(\sigma_0^i\right)^{-2} + y_t^r}$$

$$\text{s.t.} \qquad y_t^l + y_t^r = 2\,t.$$

As the objective function is concave in y_t^l and y_t^r it follows that the solution to the above problem satisfies $y_t^l = y_t^r = t$. As this is achievable by $\beta_t^l = \beta_t^r = 1$ this means that the policy which pays equal attention minimizes the posterior variance at time any time t simultaneously. Denote the resulting time path of the posterior variance by $\tilde{v}_t$.

The last step is to show that minimizing the posterior variance at each time is optimal. We are thus left to argue that for any stopping strategy τ according to which the agent takes a decision between the two objects it is better always if the posterior variance is lower. The remainder of the proof argues that this is indeed true and learning quicker is beneficial to the agent.] Fix an attention strategy β and denote by $E^\beta[\cdot]$ the associated expectation operator, and by $E^{\beta^\star}[\cdot]$ the expectation operator associated with the limiting case where the agent switches attention instantaneously between the two objects. The optimal stopping policy τ is a solution to

$$\sup_\tau \mathbb{E}^\beta \left[\max\{X_\tau^l, X_\tau^r\} - c\tau\right] = \sup_\tau \mathbb{E}^\beta \left[\max\{X_\tau^l - X_\tau^r, 0\} - c\tau\right] + X_0^r \qquad (17)$$

By the Dambis, Dubins, Schwarz Theorem (see, e.g., Theorem 1.6, chapter V of Revuz and Yor, 1999) there exists a Brownian motion $(B_\nu)_{\nu \in [0, \sigma_0^2]}$ such that $X_t^l - X_t^r = B_{\sigma_0^2 - v_t}$; this a time change where the new scale is the posterior variance. Furthermore, we can define the stochastic process $\phi_\nu := \inf\{t : \sigma_0^2 - v_t \geq \nu\}$. By eq. (17) the value of the agent is given by

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

As the posterior variance v_t is greater than the minimal posterior variance $\tilde{v}_t$ we have that $\phi_\nu \geq \tilde{\phi}_\nu := \inf\{t \colon \sigma_0^2 - \tilde{v}_t \geq \nu\}$. It follows from $\phi_r \geq \tilde{\phi}_r$ that the value when using the attention strategy β is smaller that the value achieved in the limit when the agent constantly switches attention between the two signals

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

$$= \sup_\tau \mathbb{E}^{\beta^\star} \left[\max\{X_\tau^l, X_\tau^r\} - c\tau \right] . \quad \square$$

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.

Baron, Jonathan, Sydney Scott, Katrina Fincher, and S Emlen Metz. 2015. "Why does the cognitive reflection test (sometimes) predict utilitarian moral judgment (and other things)?" *Journal of Applied Research in Memory and Cognition*, 4(3): 265–284.

Bather, John A. 1962. "Bayes procedures for deciding the sign of a normal mean." Vol. 58, 599–620, Cambridge Univ Press.

Bhui, Rahul. 2017. "A Statistical Test for the Optimality of Deliberative Time Allocation." *mimeo.*

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.

Busemeyer, Jerome R, and James T Townsend. 1992. "Fundamental derivations from decision field theory." *Mathematical Social Sciences*, 23(3): 255–282.

Busemeyer, 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.

Busemeyer, 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 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 prepraration.*

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 choicerecent advances." *The Journal of Economic Perspectives*, 25(4): 3–30.

Frazier, Peter, and Angela Yu. 2007. "Sequential hypothesis testing under stochastic deadlines." 465–472.

Frederick, Shane. 2005. "Cognitive reflection and decision making." *The Journal of Economic*

Perspectives, 19(4): 25–42.

Fudenberg, D., K. K. Li, P. Strack, and T. Strzalecki. 2017. "Speed and accuracy in perception and choice." *mimeo.*

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 D, Elisabet Tubau, and Wim De Neys. 2016. "The Doubting System 1: Evidence for automatic substitution sensitivity." *Acta psychologica*, 164: 56–64.

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.

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.

Liang, A., X. Mu, and V. Syrgkanis. 2017. "Dynamic Information Acquisition from Multiple Sources." *mimeo.*

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.

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.

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.

Shiryaev, Albert N. 1969. *Optimal stopping rules.* Russian Edition published by Nauka.

Shiryaev, 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." NBER.