

Speed, Accuracy, and Complexity

Duarte Gonçalves*

Abstract

This paper re-examines the validity of using response time to infer problem complexity. It revisits a canonical Wald model of optimal stopping, taking signal-to-noise ratio as a measure of problem complexity. While choice quality is monotone in problem complexity, expected stopping time is inverse U -shaped. Indeed decisions are fast in both very simple and very complex problems: in simple problems it is quick to understand which alternative is best, while in complex problems it would be too costly – an insight which extends to general costly information acquisition models. This non-monotonicity also underlies an ambiguous relationship between response time and ability, whereby higher ability entails slower decisions in very complex problems, but faster decisions in simple problems. Finally, this paper proposes a new method to correctly infer problem complexity based on the finding that choices react more to changes in incentives in more complex problems.

Keywords: Sequential Sampling; Complexity; Information Acquisition; Drift-Diffusion; Statistical Decision Theory.

JEL Classifications: D83, D84, D90.

* Department of Economics, University College London; duarte.goncalves@ucl.ac.uk.

I am indebted to Nick Netzer for many long conversations that indelibly shaped this project. I also thank César Barilla, Doruk Cetemen, Teresa Esteban-Casanelles, Philippe Jehiel, Deniz Kattwinkel, Pietro Ortoleva, Antonio Rangel, Ludvig Sinander, Ran Spiegler, William Zhang, and the participants at the Southeast Theory Festival (Oxford), UCL-LSE Theory Workshop, Junior Theory Workshop (Bonn), and University of British Columbia, University of Southern California, University of California - San Diego, University of California - Santa Barbara, and Caltech for valuable feedback.

First posted draft: 12 February 2024. *This draft:* 14 May 2024.

1. Introduction

It is almost definitional that more complex decision problems entail more mistakes. Individuals make more mistakes and choose dominated alternatives in more complex environments, both in experimental settings (e.g. [Rabin and Weizsäcker, 2009](#); [Martínez-Marquina et al., 2019](#); [Banovetz and Oprea, 2023](#); [Enke and Graeber, 2023](#); [Puri, 2023](#)) and in real-world decisions with significant financial implications.¹ This has important implications, since deploying theoretically optimal but complex incentive structures (tax schedules, allocation mechanisms, pricing schemes, etc) may lead to unreliable predictions and suboptimal outcomes — an observation at the root of the ongoing focus on simple mechanisms, both in theoretical and in empirical work.² If problem complexity is taken to be a leading explanation of suboptimal decision-making, understanding which problems are more complex becomes crucial for simplifying them.³

Response time is often used as an indicator of problem complexity in various settings, including lottery choice ([Wilcox, 1993](#)), loans ([Kalaycı and Serra-Garcia, 2016](#)), and portfolio selection ([Carvalho and Silverman, 2017](#)).⁴ Going beyond economics, a large literature in psychology and neuroscience has also used response time as a proxy for problem complexity ([Wright and Ayton, 1988](#); [Ratcliff and McKoon, 2008](#); [Shadlen and Shohamy, 2016](#); [Bogacz et al., 2010](#)). The underlying assumption is that in more complex problems, making similarly good choices requires greater cognitive effort, and response time can serve as a proxy for cognitive effort ([Caplin et al., 2020](#)). If decision-makers are able to make more accurate choices while expending less cognitive effort, as indicated by faster responses, then the decision problem is considered simpler. Therefore, a decision problem is deemed simpler if individuals make fewer mistakes (more accurate choices) and do so with less cognitive effort (faster choices).

The reliance on response time as a proxy for problem complexity is based on the observation that,

¹Examples include choosing health insurance plans ([Sinaiko and Hirth, 2011](#); [Bhargava et al., 2017](#)), pension plans ([Choi et al., 2011](#)), and mortgages ([Keys et al., 2016](#); [Agarwal et al., 2016](#)).

²For recent theoretical work concerned with simplicity in mechanism design see [Li \(2017\)](#), [Börgers and Li \(2019\)](#), and [Pycia and Troyan \(2023\)](#); an example of empirical research in the same domain is [Rees-Jones and Taubinsky \(2020\)](#).

³Or to obfuscate them further: a growing literature discusses how complexity can be leveraged and manipulated to the advantage of principals — cf. [Ellison and Wolitzky \(2012\)](#), [Jehiel \(2011\)](#), [Spiegler \(2016\)](#), and [Li and Dworzak \(2022\)](#).

⁴Many of these studies *show* a correlation between response time and problem complexity, rather than assuming it. Other recent examples are [Frydman and Nunnari \(2023\)](#), who use decision time together other behavioural patterns in support of specific coordination settings being more complex, and [Horenstein and Grabiszewski \(2022\)](#), who use it to rank extensive-form games in terms of complexity. More examples are surveyed in [Spiliopoulos and Ortmann \(2018\)](#) and [Gill and Prowse \(2023\)](#).

in many cases, faster choices tend to be more accurate. For example, research in cognitive sciences has indicated that perceptual problems that are harder to discern require greater cognitive effort, leading to slower and less accurate decisions (e.g. [Forstmann et al., 2016](#)). Additionally, studies have found that decision-makers take longer when they are closest to being indifferent, indicating the difficulty in distinguishing the best alternative ([Mosteller and Nogee, 1951](#); [Krajbich et al., 2012](#); [Clithero, 2018](#); [Alós-Ferrer and Garagnani, 2022](#)). This intuition aligns with sequential sampling models ([Fudenberg et al., 2018](#); [Gonçalves, 2023](#); [Alós-Ferrer et al., 2021](#)), where decisions take longer when individuals are closer to being indifferent, as the value of information is highest in such cases.⁵ In short, the justification behind the use of response times as a proxy for problem complexity is straightforward: a problem is considered easier if choices are both faster and more accurate, with faster choices often implying greater accuracy.

Whilst response time may be a compelling indicator of problem complexity, it may not always be a reliable measure when used on its own. Contributing to an long-standing discussion ([Hogarth, 1975](#); [Ford, 1975](#); [Onken et al., 1985](#); [Wright and Ayton, 1988](#)), a more recent literature documents a more ambiguous relationship between speed and accuracy, where in some cases faster decisions are better whereas in others slower decisions are more accurate (e.g. [Achtziger and Alós-Ferrer, 2013](#); [Moritz et al., 2014](#); [Caplin and Martin, 2016](#)). Introspection also cautions against positing a monotone relationship between speed, accuracy, and complexity: if problem complexity demands greater cognitive effort and longer decision time, then in excessively complex problems thinking for a long time may not be worthwhile.

In this paper, we reconcile this ambiguous intuition on the relation between speed-accuracy by revisiting a standard Wald problem of optimal stopping ([Wald, 1947](#); [Dvoretzky et al., 1953](#)). As is conventional in this literature, we take the drift-to-volatility or signal-to-noise ratio as a measure of complexity of the decision problem, capturing the rate of information per unit of time. Our main result is but a simple observation: although choice quality is monotone in problem complexity, expected stopping time is inverse U -shaped. This suggests a more nuanced relation between speed and accuracy: When choice problems are simple, people choose fast and well; if they become slightly more complicated, they choose slower and less well; if they become much more complicated, choice quality necessarily deteriorates, but one should instead expect response times to be *faster*. Extending our model to dynamic effort control, we uncover that this non-monotonicity also suggests restraint in using response times to infer ability of the decision-

⁵See also [Drugowitsch et al. \(2012\)](#).

makers, as high ability individuals always choose better, but do so faster in simple problems and slower in complex ones. Moreover, taking complexity affecting the cost of information acquired and assuming response time is proportional to it, we show that this non-monotonicity is a generic feature of models with costly information acquisition. Finally, we leverage our results to address the crucial question at hand of how to infer problem complexity. We find that distorting incentives is more effective in steering choices in more complex problems, allowing one to infer problem complexity from simple manipulations of incentives.

Our baseline setup is the most canonical sequential sampling model (Wald, 1947; Dvoretzky et al., 1953), where a decision-maker faces a binary choice problem with a constant flow cost per unit of time and can obtain information about a binary event or parameter by observing a Brownian motion with drift. We take the signal-to-noise ratio, i.e. the drift intensity relative to the instantaneous volatility, as a measure of problem complexity, as is common in existing literature (e.g. Callander, 2011; Fehr and Rangel, 2011). The nature of the exercise is then to examine how speed (expected stopping time) and accuracy (probability of choosing the optimal alternative) depend on problem complexity.

We first fully characterise the solution to the optimal stopping problem, which we then use to derive the stopping boundaries and the expected stopping time. We do so for general payoffs, obtaining a clear characterisation of the optimal stopping beliefs as a function of problem complexity.

We then show that, in the canonical setup, accuracy is monotone in problem complexity, whereas expected stopping time is non-monotone. While in some cases it may be more important to choose the best alternative than in others, our results show that, no matter relative preference intensity, the probability with which a decision-maker chooses the ex-post optimal alternative monotonically increases as the decision problem becomes simpler. Regarding the stopping time, there are two opposing forces that arise when the problem becomes simpler. As signal-to-noise ratio increases, on the one hand, the quality of information is higher, implying that the marginal benefit of learning for longer is greater. On the other hand, the decision-maker can also save on the cost of information by stopping earlier. In a standard drift-diffusion version of the model, the decision-maker stops whenever the accumulated evidence exceeds an exogenous threshold, leading to a monotone relationship between problem complexity, speed, and accuracy.⁶

⁶This could be interpreted as if the decision-maker is unable to recognise whether the problem is simpler or more complex, and so uses the same threshold.

We leverage closed-form characterisations of the optimal stopping boundaries and expected optimal stopping time to show that the former expand when the problem becomes simpler — entailing greater accuracy — but the latter is non-monotone and single-peaked. This then provides a rationale for the evidence documented by [Achtziger and Alós-Ferrer \(2013\)](#) that, in simple situations, quick choices are better, but in more complicated ones, faster decisions are associated with worse performance. Moreover, our results are robust to several variations of the baseline model, including assuming that the drift term is proportional to payoff scale — as in standard drift-diffusion models ([Fehr and Rangel, 2011](#)) — or to having discounting instead of a constant flow cost.

Several papers have examined the relation between response time and ability. For instance, it is often found that slower responses are indicative of higher ability or sophistication (e.g. [Rubinstein, 2016](#); [Agranov et al., 2015](#); [Recalde et al., 2018](#); [Esteban-Casanelles and Gonçalves, 2022](#)). However, other studies suggest that, similarly to accuracy, this relationship is ambiguous, with greater ability being associated with *faster* response times in simple problems ([Dodonova and Dodonov, 2013](#); [Goldhammer et al., 2014](#); [Kyllonen and Zu, 2016](#)). Anecdotal evidence again cautions against generic claims for monotone association: for instance, returning exams earlier typically indicates that a student will score significantly higher or significantly lower than their peers.

We show that this ambiguity is a natural consequence of the non-monotone relation between speed and accuracy. For this we extend our model to allow for the decision-maker to control instantaneous effort, which entails a cost per unit of time. In our model, higher effort scales up the signal-to-noise ratio,⁷ and the decision-maker will optimally choose a constant level of effort. This can then be interpreted as if the decision-maker faces an idiosyncratic problem with complexity level that depends on their cost of effort. We contrast the speed and accuracy of high-ability and low-ability decision-makers, where higher ability captures a lower cost of effort and show that high ability individuals always choose better, but do so faster than lower ability individuals in simple problems, and slower in more complex ones. Furthermore, this also cautions against pooling data on multiple individuals, since then response time will not in general be single-peaked.

We extend our results to more general models of costly information acquisition. Going to a more general setup, we consider general costly information acquisition models (e.g. [Caplin and Dean, 2015](#)), where the decision-maker can acquire a Blackwell experiment to learn about an unknown

⁷Our setup is akin to [Moscarini and Smith \(2001\)](#), but with no discounting, to keep the model as close as possible to our baseline.

state paying a cost that is increasing with respect to the Blackwell order. We assume that (1) problem complexity corresponds to a parameter that scales up the cost of information,⁸ and (2) that expected stopping time is proportional to the total cost of information. The first assumption is motivated by the fact that, in the case of the Wald problem, the square of noise-to-signal ratio corresponds to such parameter (Morris and Strack, 2019), whereas the second assumption not only exactly corresponds to our baseline setup, but is also consistent with evidence from Caplin et al. (2020). We then show that, in this general setup, expected stopping time is then non-monotone in problem complexity, and provide conditions for it to be single-peaked.⁹

Finally, we return to the fundamental question of how an analyst can infer which problem is more complex. Since response time is not always an appropriate measure of problem complexity, we consider other alternatives. Naturally, when the analyst knows which choice is optimal, they can infer problem complexity from accuracy directly, where the simpler problem is the one in which agents are more likely to choose the best alternative. In many cases, however, which alternative is best depends on unobserved parameter or idiosyncratic preferences. For instance, the decision problem itself may be meant to have people select different options depending on their preferences or some private type, in which case it is impossible for an analyst to assess if choices are optimal or not.

We propose a simple alternative: to infer problem complexity, the analyst need only to manipulate incentives and observe how choices change. Specifically, we consider the case in which one picks one of the two alternatives in both problems and simply pays decision-makers one extra cent if they choose such alternative. As in any reasonable model, decision-makers would now choose this alternative more often and do so faster, and choose the other less often and slower.¹⁰ The analyst can then infer problem complexity from *how much more often* the incentivised alternatives are now chosen. Specifically, the extra cent has a greater effect on choice probabilities in more complex problems than in simpler ones. In other words, when the problem is simpler, the decision-maker is surer about which alternative is better and so small changes in relative payoffs have a smaller effect on choice probabilities.

⁸As is standard in costly information acquisition or rational inattention models with entropy costs (Matějka and McKay, 2015, e.g.).

⁹While we only consider models in which the cost information is continuous and convex, recent literature has also indicated that in binary decision problem with one-sided Poisson process, the expected decision time could be non-monotone in the arrival rate (Bobtcheff and Levy, 2017; Halac et al., 2016). This non-monotonicity, however, does seem to depend on modelling details — see Bobtcheff and Levy (2017, Proposition 1).

¹⁰An observation that holds in for general dynamic costly information acquisition models — see Gonçalves (2023, Theorem 3).

2. Setup

The understanding of the relationship between speed and accuracy has been indelibly shaped by the wide-spread use of sequential sampling models in psychology, neuroscience, and economics. In this section we introduce a simple version of the drift-diffusion model, a particularly influential stylised model of decision-making that models the accumulation of evidence in favour of one of two alternatives as a Brownian motion with drift.

Alternatives and Payoffs. A decision-maker is choosing from a set $A = \{a, b\}$, with generic element $\alpha \in A$. Let $(\Omega, \mathcal{F}, \mathbb{P})$ denote a probability space, where $\omega \in \Omega$ denotes the state of the world, \mathcal{F} a Borel σ -algebra, and \mathbb{P} a probability measure. Payoffs to each alternative depend on a parameter or event $\theta \in \Theta := \{a, b\}$, and are given by $u : A \times \Theta \rightarrow \mathbb{R}$. We assume there are not weakly dominant alternatives so that the choice problem is non-trivial and, without loss of generality, θ specifies which alternative is best, i.e. $u(\theta, \theta) > u(\theta', \theta)$, $\theta \neq \theta'$.

We denote the prior probability that $\theta = a$ by $p_0 := \mathbb{P}(\omega : \theta(\omega) = a)$ and we extend the utility function to distributions over Θ , writing $u(\alpha, p_0) := \mathbb{E}[u(\alpha, \theta)]$ and $v(p_0) := \max_{\alpha} u(\alpha, p_0)$ where the dependence on p_0 is implicit.¹¹ For simplicity, we assume each option is equally likely – this simplifies the presentation and the proofs, but does not affect the substance of the results.

Sequential Sampling. Prior to choosing, the decision-maker can acquire information about θ , paying a flow cost per unit of time $c > 0$. The decision-maker observes a continuous-time process given by a Brownian motion with drift, $X_t := \mu s(\theta)t + \sigma B_t$, where B_t is a standard Brownian motion defined on (Ω, \mathcal{F}, P) , $s(\theta) = 1$ if $\theta = a$ and $= -1$ if $\theta = b$,¹² $\mu > 0$ corresponds to the drift rate, and $\sigma > 0$ scales the instantaneous volatility.

The model is typically interpreted to be a stylised representation of the process of deliberation or reasoning in decision-making, with the state of the process X_t meant to represent accumulated evidence in favour of or against any of the alternatives. While in psychology and neuroscience it became a benchmark in modelling the neural basis of decision-making,¹³ in economics it has

¹¹Instead of understanding varying p_0 as changing the measure \mathbb{P} , it can be equivalently understood as varying the random variable $\theta : \Omega \rightarrow \Theta$.

¹²Note that, given $s(a) \neq s(b)$, insofar as s is known, having the drift being $\pm\mu$ is without loss of generality since the process is observationally equivalent.

¹³For instance, it has been the main modelling paradigm for the recalling past experiences and sampling from memory (Ratcliff, 1978; Shadlen and Shohamy, 2016), the sensory perception of physical characteristics (Gold and Shadlen,

been used to model a wide range of situations and behavioural patterns, including randomness in consumer choice (Clithero, 2018; Alós-Ferrer et al., 2021), the relation between misperception and small-stakes risk aversion (Frydman and Jin, 2022; Khaw et al., 2021), consumer’s search for product information (Branco et al., 2012), choosing actions in games (Schotter and Trevino, 2021), information gathering and deliberation in groups (Reshidi et al., 2024), and brand recognition (Alós-Ferrer, 2018).

Problem Complexity. As the signal-to-noise ratio μ/σ captures the rate of information flow per unit of time, we will take this as a measure of problem complexity – with higher values of μ/σ denoting easier problems. This is a common interpretation in economics, cognitive sciences, and psychology (e.g. Callander, 2011; Fehr and Rangel, 2011; Forstmann et al., 2016). In essence, it captures the intuition that simple problems (high μ/σ) are those in which a decision-maker quickly realizes which alternative or course of action is best, whereas in complex problems (low μ/σ) it would take longer to be as confident. A decision problem may be more complex owing to the intrinsic properties of the items involved – e.g. deciding between savings products is arguably easier and faster than understanding which structured financial product is best for oneself; solving easy exam questions is faster than harder ones. Alternatively, how complex a problem is may have to do with how information is presented or how easy it is to obtain information, making it easier or harder to compare across alternatives.

Optimal Stopping. The decision-maker then faces a trade-off between taking longer to have a better understanding of which alternative is better, and the effort involved in doing so. Formally, the decision-maker chooses a stopping time τ adapted to the natural filtration induced by the stochastic process $(X_t)_t$ ¹⁴ in order to maximise their expected payoff, $\mathbb{E}[v(p_\tau) - c\tau] = \mathbb{E}[\max_\alpha \mathbb{E}[u(\alpha, \theta) | \mathcal{F}_\tau] - c\tau]$, where $p_t := \mathbb{E}[1_{\theta=\alpha} | X_t]$ denotes the posterior belief about θ given X_t . That is,

$$V(p_0) = \sup_{\tau} \mathbb{E}[v(p_\tau) - c\tau],$$

where again the dependence on p_0 is implicit.

2001; Ratcliff et al., 2016), and brain activation in decision-making (Bogacz et al., 2010; Shushruth et al., 2022). Recent variations extend the model to interpret eye gaze, e.g. in purchasing decisions (Krajbich et al., 2012).

¹⁴That is, the decision-maker chooses a random variable taking values in \mathbb{R}_+ such that $\{\omega \in \Omega : \tau(\omega) \leq t\} \in \mathcal{F}_t^X$, where \mathcal{F}_t^X is the smallest σ -field under which X_t is measurable.

3. Speed, Accuracy, and Optimal Sequential Sampling

The main focus of this section is to provide comparative statics relating, on the one hand, problem complexity, μ/σ , and, on the other hand, both choice accuracy – the probability of choosing the best alternative, $\mathbb{P}(\alpha_{\tau^*} = \theta)$ – and how long it takes to make a choice, that is, the expected stopping time, $\mathbb{E}[\tau^*]$.

Intuitively, simpler problems, with higher signal-to-noise ratio, imply that the decision-maker can more easily distinguish between the two alternatives, leading to more accurate choices. In what regards time, however, the relation is less clear. When making a problem simpler, there is a greater marginal informational benefit of engaging for longer, suggesting longer expected stopping time. However, the decision-maker can also save on the flow cost of information by stopping earlier, leading to shorter expected stopping times. We show that which of the two forces dominates is ambiguous. To do so, we first start by characterising the optimal stopping time and the optimal stopping boundaries, and then proceed to examine how speed and accuracy depend on problem complexity. All proofs are provided in [Appendix A](#).

3.1. Characterising Optimal Stopping

Auxiliary Payoff Functions. Let us introduce the following auxiliary payoff functions: $u_a(\alpha, p) := 1_{\alpha=a} \cdot (p - \tilde{p})$ and $u_b(\alpha, p) = 1_{\alpha=b}(\tilde{p} - p)$, where \tilde{p} denotes the belief that renders the decision-maker indifferent between both alternatives, i.e. $\tilde{p} : u(a, \tilde{p}) = u(b, \tilde{p})$. The interpretation of these payoff functions is simple: for $\alpha \in A$, u_α is such that α is the risky alternative and the other is the safe one, delivering a payoff normalised to zero. This then implies that the relation between beliefs and preference is retained: $u(a, p) \geq u(b, p) \iff u_a(a, p) \geq u_a(b, p) \iff u_b(a, p) \geq u_b(b, p) \iff p \geq \tilde{p}$.

We can then define maximised expected payoffs and value functions for both these auxiliary payoff functions: For $\alpha \in \{a, b\}$, define $v_\alpha(p) := \max_{\hat{\alpha}} u_\alpha(\hat{\alpha}, p)$ and $V_\alpha(p) := \sup_{\tau} \mathbb{E}[v_\alpha(p_\tau) - (c/\delta)\tau]$, where $p = \mathbb{P}(\theta = a)$.

These will turn out to be quite useful, as the optimal stopping time of the original problem can be characterised in terms of the optimal stopping times of these auxiliary problems:

Lemma 1. *For any $p \in [0, 1]$, $V(p) = \delta V_a(p) + u(b, p) = \delta V_b(p) + u(a, p)$, and*

$$\arg\max_{\tau} \mathbb{E}[v(p_\tau) - c\tau] = \arg\max_{\tau} \mathbb{E}[v_\alpha(p_\tau) - (c/\delta)\tau] = \arg\max_{\tau} \mathbb{E}[v_b(p_\tau) - (c/\delta)\tau],$$

where $\delta := (u(a, a) - u(b, a)) + (u(b, b) - u(a, b))$.

The above fact follows almost directly from the simple observation that one can obtain both u_a and u_b as positive affine transformations of u : $u = \delta u_a + u(b, \cdot) = \delta u_b + u(a, \cdot)$, where $\delta := (u(a, a) - u(b, a)) + (u(b, b) - u(a, b))$ can be interpreted as the overall incentive level or stakes of the problem.

Optimal Stopping Boundaries. Relying on these auxiliary problems, it is then straightforward to show that the decision-maker's optimal stopping time is pinned-down by two constant belief thresholds, \underline{p} and \bar{p} , such that the decision-maker stops as soon as their belief about θ reaches either of the two thresholds.

The argument is based on the following lemma:

Lemma 2. V_a is convex and increasing in p_0 and V_b is convex and decreasing in p_0 .

This is intuitive, given that v_a is increasing in p and v_b is decreasing in p . Convexity, a property that will be useful later on, follows from linearity in the measure.

Given this, we can then show that

Lemma 3. For any prior $p_0 \in (0, 1)$, the earliest optimal stopping time τ^* is given by $\tau^* := \inf\{t \geq 0 | V(p_t) \leq v(p_t)\} = \inf\{t \geq 0 : p_t \notin (\underline{p}, \bar{p})\}$, where $0 < \underline{p} < \tilde{p} < \bar{p} < 1$.

In particular, we have that $\underline{p} := \max\{p \in [0, 1] | V_a(p) \leq 0\}$ and $\bar{p} := \min\{p \in [0, 1] | V_b(p) \leq 0\}$. When $p \leq \underline{p}$, monotonicity of V_a implies that $V_a(p) = 0 = u_a(b, p)$, and so the decision-maker must stop and choose b ; and this must be the case in both the auxiliary problems. Similarly, when $p \geq \bar{p}$, $V_b(p) = 0 = u_b(a, p)$, and so the decision-maker must stop and choose a . Hence, the decision-maker has not stopped yet if and only if $p \in (\underline{p}, \bar{p})$. **Lemma 1** then ties in the argument and yields that the optimal stopping time is characterised by these two thresholds, $\tau^* = \inf\{t \geq 0 | p_t \notin (\underline{p}, \bar{p})\}$.

Figure 1 illustrates these lemmas. In panel (a) we have the original maximised expected payoff, v , as well as the two auxiliary versions v_a and v_b , all of which piece-wise linear functions with a kink at the indifference point \tilde{p} . The value functions, V , V_a , and V_b , are then given on the right panel (b), coinciding with v , v_a , and v_b , respectively, outside the continuation region (\underline{p}, \bar{p}) .

In showing that the continuation region is characterised by two constant thresholds, **Lemma 3** extends this standard result from [Shiryayev \(2007\)](#) to our slightly more general setting in which

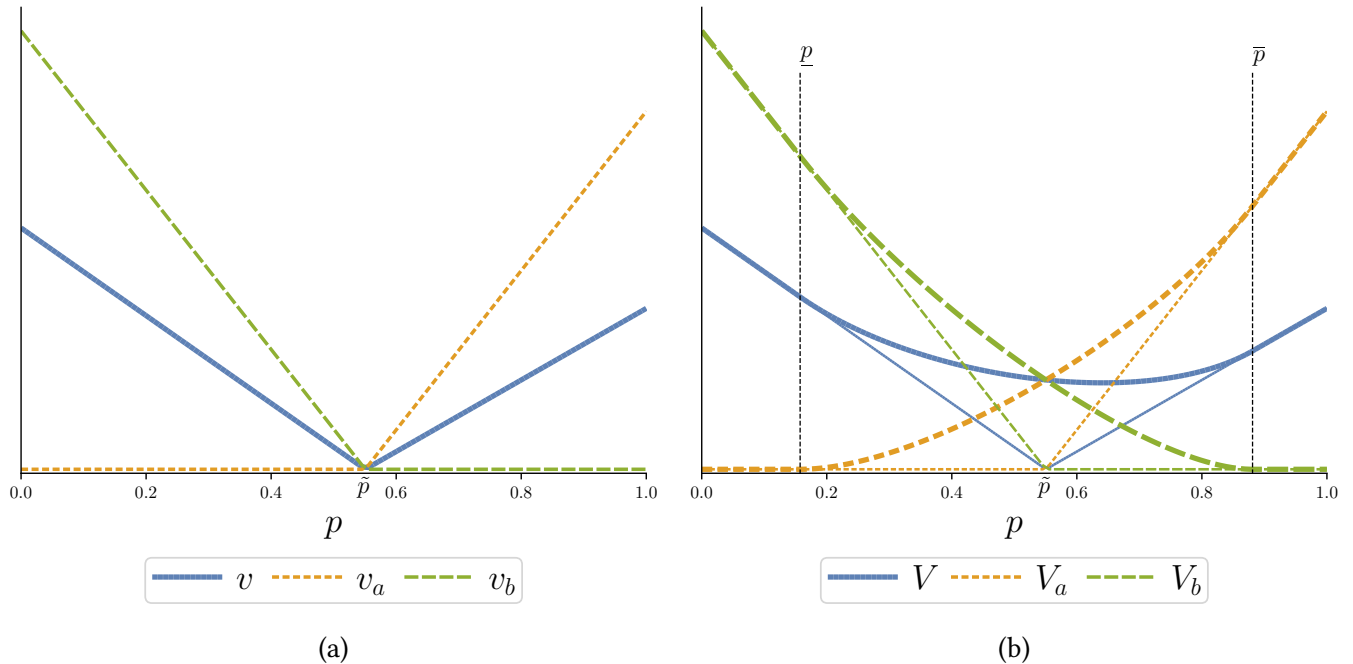


Figure 1: Payoff and Value Functions

the payoff difference between the two alternatives is not certain as in [Fudenberg et al. \(2018\)](#). In other words, it is not the fact that $|u(a, \theta) - u(b, \theta)|$ is constant with respect to θ that underlies the boundary being constant, but rather that $u(a, \theta) - u(b, \theta)$ takes only two possible values.

This observation is valuable as it makes transparent the connection between our model and the drift-diffusion model. In the standard use of the drift-diffusion model, the decision-maker stops as soon as the accumulated evidence reaches a pre-specified threshold ([Forstmann et al., 2016](#)), which is equivalent to the decision-maker stopping as soon as their belief about θ reaches a pre-specified threshold. While these thresholds and the model parameters are typically estimated from choice data, it is understood that the thresholds are endogenous: “[t]he assumption is that payoffs or instructions induce the subject to adjust the positions of the (...) boundaries and so to adjust the amount of information required for a decision” ([Ratcliff, 1978](#)). Our focus is exactly on how problem complexity, μ/σ , affects the optimal stopping boundaries, and how this in turn affects the speed and accuracy of choices.

In order to precisely determine the optimal stopping thresholds, we rely on the differential characterisation of the value function V :

Lemma 4. *The value function V is given by the unique viscosity solution to the free-boundary problem*

$$\begin{aligned} 0 &= -c + 2(\mu/\sigma)^2 (p(1-p))^2 V''(p) && \text{on } p \in (\underline{p}, \bar{p}) && \text{(HJB)} \\ 0 &= V(p) - v(p) && \text{on } p \notin (\underline{p}, \bar{p}). && \text{(BC)} \end{aligned}$$

Note that V by definition is not a classical solution to the boundary problem since it will necessarily fail to be \mathcal{C}^2 .¹⁵

We then obtain a sharp characterisation of the optimal stopping boundaries. To state it, let the log-odds of p be denoted by $\ell(p) := \ln(p/(1-p))$. The log-odds representation will be useful, since the log-odds of the posterior beliefs at time t is linear in X_t :

$$\ell_t = \ln \left(\frac{\mathbb{P}(\theta = a | X_t)}{\mathbb{P}(\theta = b | X_t)} \right) = \ln \left(\frac{\phi((X_t - \mu t)/\sqrt{t}\sigma)}{\phi((X_t + \mu t)/\sqrt{t}\sigma)} \right) = 2(\mu/\sigma^2)X_t.$$

We then have:

Proposition 1. *The optimal stopping time τ^* is given by $\tau^* = \inf\{t \geq 0 : p_t \notin (\underline{p}, \bar{p})\}$, where \underline{p} and \bar{p} are the unique solution to*

$$\frac{(\frac{\mu}{\sigma})^2}{\tilde{c}} = \left[\exp(\bar{\ell}) - \exp(-\bar{\ell}) + 2\bar{\ell} \right] - \left[\exp(\underline{\ell}) - \exp(-\underline{\ell}) + 2\underline{\ell} \right] \quad (1)$$

$$0 = \left[\bar{\ell} + \exp(\bar{\ell}) - \exp(\tilde{\ell})(\bar{\ell} - \exp(-\bar{\ell})) \right] - \left[\underline{\ell} + \exp(\underline{\ell}) - \exp(\tilde{\ell})(\underline{\ell} - \exp(-\underline{\ell})) \right] \quad (2)$$

where $\bar{\ell} := \ell(\bar{p})$, $\underline{\ell} := \ell(\underline{p})$, $\tilde{\ell} := \ell(\tilde{p})$, and $\tilde{c} := c/(2\delta)$.

The proof of **Proposition 1** is but a corollary of **Lemma 4**: by explicitly solving the free-boundary problem, we obtain the optimal stopping boundaries as the unique solution to (1) and (2).

3.2. Speed and Accuracy

As indicated, our main result for this section is the following:

Theorem 1. *Accuracy $\mathbb{P}(\alpha_{\tau^*} = \theta)$ is increasing in μ/σ , and expected optimal stopping time $\mathbb{E}[\tau^*]$ is non-monotone and quasi-concave in μ/σ , and strictly so for all μ/σ such that $\tau^* \neq 0$.*

Given the stopping boundaries in log-odds beliefs, $\bar{\ell}, \underline{\ell}$, one can translate these to stopping boundaries for the Wiener process, as $\ell_t = 2(\mu/\sigma^2)X_t$. From there, it is well-known how to obtain closed-form expressions for the expected stopping time and the probability of choosing the best

¹⁵Note that $V''(\bar{p}^-), V''(\bar{p}^+) > 0 = v''(\bar{p}^+), v''(\underline{p}^-)$.

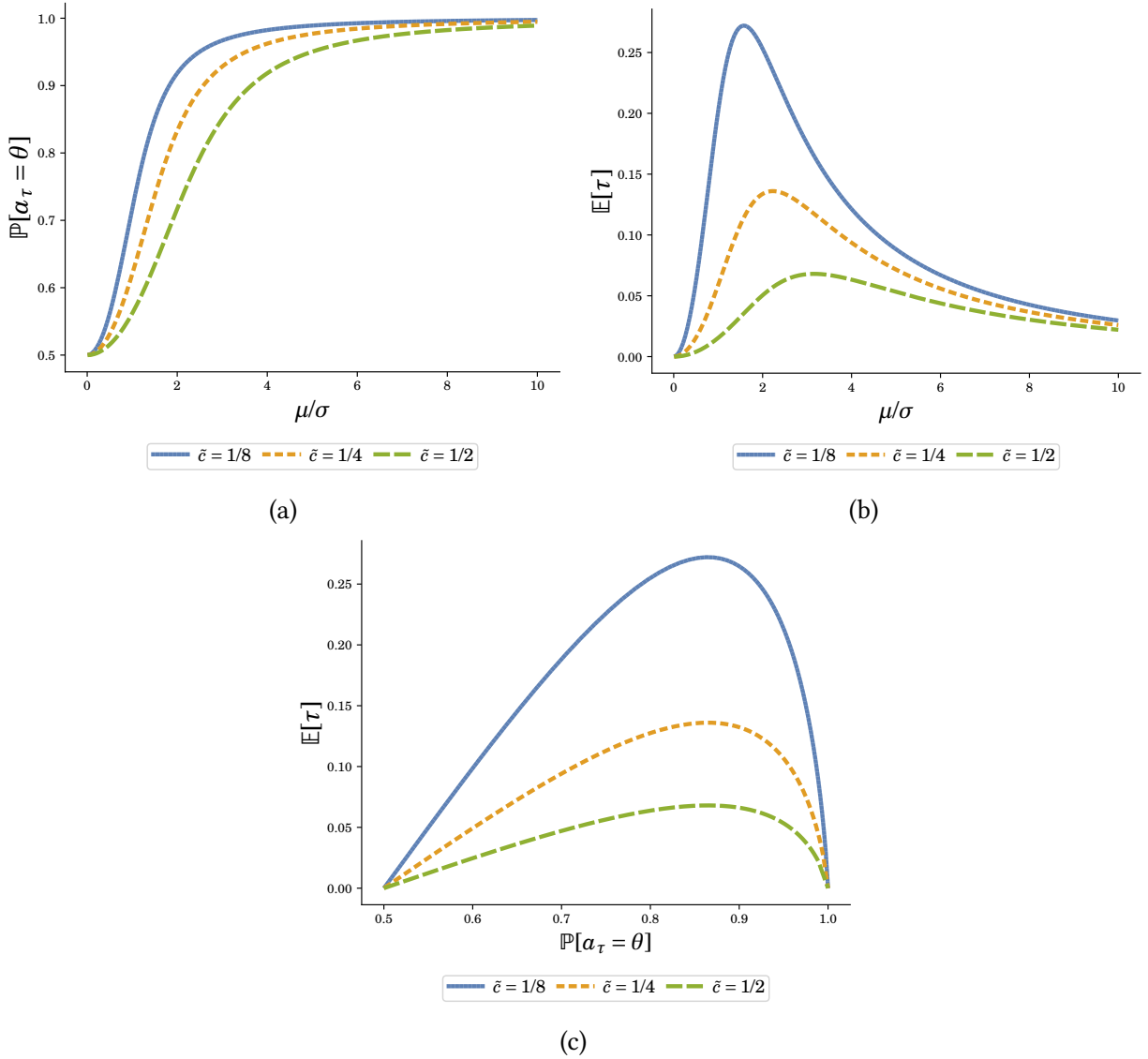


Figure 2: Speed and Accuracy

alternative as a function of μ/σ . Then, the comparative statics in [Theorem 1](#) arise from a simple, even if cumbersome, application of the implicit function theorem using (1) and (2). [Figure 2](#) exhibits how speed and accuracy depend on problem complexity and their implied inverse U-shaped relationship.

While, at first glance, that accuracy is monotone in problem complexity may not seem particularly surprising, it is important to note that the result in [Theorem 1](#) is independent of ex-post relative preference intensity. In particular, it might be more important for the decision-maker to choose $\alpha = \theta$ when $\theta = a$ than when $\theta = b$. For instance, when alternative a is better it is much

better than b , but when it is worse it is only mildly worse – something that would be captured by a small value of the indifference parameter \tilde{p} . Our results show that such considerations regarding relative preference intensity do not matter: the probability with which the decision-maker chooses the ex-post optimal alternative is monotone in problem complexity.

Theorem 1 means that easier problems translate into more accurate choices, which, as the proof shows, is due to wider stopping thresholds for beliefs. While wider thresholds entail longer expected stopping times, a higher signal-to-noise ratio entail faster choices. The former effect dominates for more difficult problems, when the signal-to-noise ratio is low, whereas the latter dominates when simpler problems are made simpler.

Theorem 1 also implies that easier problems translate into more accurate choices, which, as the proof shows, is due to wider stopping thresholds for beliefs. While wider thresholds entail longer expected stopping times, a higher signal-to-noise ratio entail faster choices. The former effect dominates for more difficult problems, when the signal-to-noise ratio is low, whereas the latter dominates when simpler problems are made simpler. This result is consistent with existing evidence that, in simple situations, faster choices are more accurate, but in more complicated ones, the opposite is true (e.g. [Onken et al., 1985](#); [Achtziger and Alós-Ferrer, 2013](#)).

We want to highlight two common variants of the model that will not affect this observation. In many instances, the drift is proportional to the payoff difference (e.g. [Fehr and Rangel, 2011](#)). In our model, this would be observationally equivalent to having a drift rate of $\mu \delta$ rather than just μ . **Theorem 1** would hold just the same in this case. A second variant of the model would feature exponential discounting instead of a flow cost of time; in [Appendix B](#) we show that our results are robust: the relation between accuracy and expected stopping time remains non-monotone.

One important implicit assumption in our analysis is that the decision-maker can to some extent recognise the complexity of the problem they face. In this, we differ from previous literature, e.g. [Fudenberg et al. \(2018\)](#), who preclude this possibility. While it may often be the case that one learns about how complicated a particular problem is while trying to solve it, there may be many instances in which there are apparent differences in complexity across problems, and the decision-maker can realise this. In such cases, without further assumptions, one should not in general expect the speed of decision to indicate response accuracy or problem complexity.

4. Ability and Effort

We extend our model to speak to effort and ability. Our intent is twofold. First, we wish to acknowledge that time spent is not necessarily a good measure of the effort spent in a task. For instance, a decision-maker might spend a lot of time on a task, but not exert much effort because it is a simple problem, or might spend little time on a hard task but exert a lot of effort. We therefore allow the decision-maker to control how much effort they will exert, positing that more effort is costly but increases how much progress they make per unit of time.

Second, we wish to speak to the ambiguity in the literature regarding the relation between ability and response time. In many contexts slower response times are taken to be indicative of greater ability or sophistication. For instance, more financially literate individuals will by definition be at a better position for understanding information about alternative financial assets (e.g. [Darriet et al., 2020](#)). Also in strategic settings, it is often found that individuals who commit more time to a problem tend to be more successful or sophisticated (e.g. [Rubinstein, 2016](#); [Agranov et al., 2015](#); [Alós-Ferrer and Bruckenmaier, 2021](#); [Esteban-Casanelles and Gonçalves, 2022](#); [Gill and Prowse, 2023](#)).

While the studies that relate response time to ability rely on measuring ability from performance independently, one may be tempted to infer ability from response time directly when performance outcomes are ambiguous or not available. Anecdotes and introspection again caution against this. For example, students who return their exam papers early would believe that spending more time and effort would not benefit them, but this could indicate either strong performance or a lack of progress. Our results corroborate this intuitive ambiguity.

We revisit the model in [Section 2](#) and extend it to allow for the decision-maker to control instantaneous effort $e_t \geq 0$, borrowing from [Moscarini and Smith \(2001\)](#). Greater effort translates to more informative signals, formally

$$dX_t = \sqrt{e_t} \mu s(\theta) dt + \sigma dB_t.$$

where $\mu, \sigma > 0$ are the same as before, B_t is a standard Brownian motion, and e_t is the instantaneous effort exerted by the decision-maker.

Exerting effort is costly, entailing a cost per unit of time of $c(e/\lambda) \geq 0$, with $c', c'' > 0$. The term $\lambda > 0$ captures the decision-maker's ability, with higher values of λ capturing lower costs of exerting effort and thus higher ability.

The decision-maker of ability λ then chooses a stopping time τ and a path of effort (e_t) , adapted to the natural filtration induced by X_t , to maximise their expected payoff,

$$\mathbb{E} \left[v(p_\tau) - \int_0^\tau c(e_t/\lambda) dt \right]$$

with $p_t = \mathbb{E}[1_{\theta=\alpha}|X_t]$ denoting the decision-maker's posterior belief, $dX_t = \sqrt{e_t}\mu s(\theta)dt + \sigma dB_t$ and $X_0 = 0$.

We find that decision time alone is not a good measure of ability. Specifically, we show that:

Proposition 2. *Let $\lambda_H > \lambda_L > 0$. Then, (1) accuracy is always higher for the high ability type, $\mathbb{P}(\alpha_H^* = \theta) > \mathbb{P}(\alpha_L^* = \theta)$, and (2) there are $\bar{k}, \underline{k} > 0$ such that, for all problems which are sufficiently complex, $\mu/\sigma \leq \underline{k}$, the expected stopping time is higher for the higher ability type $\mathbb{E}[\tau_H^*] \geq \mathbb{E}[\tau_L^*]$, and for all simple enough problems, $\mu/\sigma \geq \bar{k}$, the reverse is true. Furthermore, the inequalities are strict whenever $\tau_H^* \neq 0$.*

The proof of [Proposition 2](#) proceeds by observing that, from the HJB equation, optimal effort is constant and, for the optimal effort of a decision-maker with ability λ , e_λ^* , is given by $e_\lambda^* = \lambda e_1^*$. Hence, the higher the ability, the greater the effort exerted. This then translates to an effectively higher signal-to-noise ratio, $\sqrt{e_\lambda^*}\mu/\sigma$, and [Proposition 2](#) then follows from [Theorem 1](#).

We note that [Proposition 2](#) is consistent with recent research in psychology that highlights the importance of taking into account task complexity when interpreting the relation between ability and response time – see e.g. [Dodonova and Dodonov \(2013\)](#), [Goldhammer et al. \(2014\)](#), [Kyllonen and Zu \(2016\)](#), and [Junghaenel et al. \(2023\)](#). In short, these studies find that, in simple tasks, faster responses are indicative of greater ability, but in complex tasks, the opposite is true. Our simple model can rationalise this nuanced prediction.

5. General Costs of Information Acquisition

We turn to the question of how general this non-monotonic relation between speed and accuracy is when considering a general class of models with costly information acquisition.

Specifically, we consider a decision-maker faces uncertainty about an unknown parameter or event $\theta \in \Theta$ and can acquire a Blackwell experiment π to learn about θ , which, following existing literature (e.g. [Kamenica and Gentzkow, 2011](#); [Caplin et al., 2022](#); [Lipnowski and Ravid, 2023](#)), we model as a distribution over posterior beliefs, i.e. $\pi \in \Pi := \Delta(\Delta(\Theta))$ such that $\mathbb{E}_\pi[p] = p_0$. We denote by $C : \Pi \rightarrow \bar{\mathbb{R}}$ the cost of the experiment, which we assume is strictly increasing with

respect to the Blackwell order \geq_B , continuous with respect to the Lévy-Prokhorov metric, strictly convex, and normalised such that a fully uninformative experiment bears zero cost, $C(\delta_{p_0}) = 0$. We model the benefit of information as a function $B : \Pi \rightarrow \mathbb{R}$ that is bounded, linear, and strictly increasing with respect to the Blackwell order, which arises as the decision-maker's expected payoff from choosing an action $a \in A$ in order to maximise their expected payoff given posterior beliefs $p \sim \pi$, i.e. $B(\pi) = \mathbb{E}_\pi[\max_{a \in A} \mathbb{E}_p[u(a, \theta)]]$.

The decision-maker then chooses an experiment in order to maximise their value of information, $v(\pi; \kappa) := B(\pi) - \kappa C(\pi)$, where $\kappa \geq 0$ is a scaling parameter. We denote the value function, optimal experiment, and cost of the optimal experiment as $V(\kappa) := \sup_{\pi \in \Pi} v(\pi; \kappa)$, $\pi^*(\kappa) := \arg \max_{\pi \in \Pi} v(\pi; \kappa)$, and $C^*(\kappa) := C(\pi^*(\kappa))$.¹⁶

In order to connect this framework to time, we assume that the decision-maker's expected stopping time is proportional to the total cost of information, i.e. $T^*(\kappa) := \mathbb{E}[\tau^*] \propto \kappa C^*(\kappa)$. The scale κ is then interpreted as a measure of problem complexity, scaling the marginal cost of information. In the case of the Wald problem, this is simply reciprocal of the square of the signal-to-noise ratio, $\kappa = (\mu/\sigma)^{-2}$, which extends to a setting with many states (Morris and Strack, 2019).¹⁷

We then obtain the following result from a simple application of Berge's maximum theorem:

Proposition 3. *$T^*(\kappa)$ is non-monotone in κ .*

Single-peakedness in the general case is more difficult to establish, requiring that $-\kappa V''(\kappa)/V'(\kappa)$ crosses at most once.¹⁸ Going beyond this loose characterisation requires additional structure.

Remark 1. Returning to the binary decision problems introduced in Section 2, we suppose $A = \Theta = \{a, b\}$ and $\bar{p} = 1/2$. If one assumes costs of information are posterior-separable (Caplin et al., 2022), i.e. given by $C(\pi) = \mathbb{E}_\pi[c(p)]$ for some $c : [0, 1] \rightarrow \mathbb{R}_+$ satisfying $c', c'' > 0$, and $c(p_0) = 0$, then $T^*(\kappa)$ is single-peaked if and only if $c(p)/c'(p)$ is quasiconcave on $[1/2, 1]$.

The argument is simple. First, we recall the well-known result that additively separable information costs, it is without loss of optimality to choose information structures with at most $|A|$ signals, in this case \bar{p} and \underline{p} . Symmetry of the problem ($\bar{p} = p_0 = 1/2$) implies at an optimum

¹⁶Note that, by Berge maximum theorem, V is continuous, π^* is single-valued and upper hemicontinuous, and so C^* is continuous.

¹⁷See Pomatto et al. (2023) for an axiomatic characterisation of such cost function.

¹⁸Note that, since V is convex, by Alexandrov's theorem, V' and V'' exist a.e., and by an application of the envelope theorem, $\kappa C^*(\kappa) = -\kappa V'(\kappa)$. Since $\kappa C^*(\kappa)$ is non-monotone, $-\kappa V''(\kappa)/V'(\kappa)$ must cross one at least once; crossing it exactly once is a sufficient and necessary condition for it to be single-peaked.

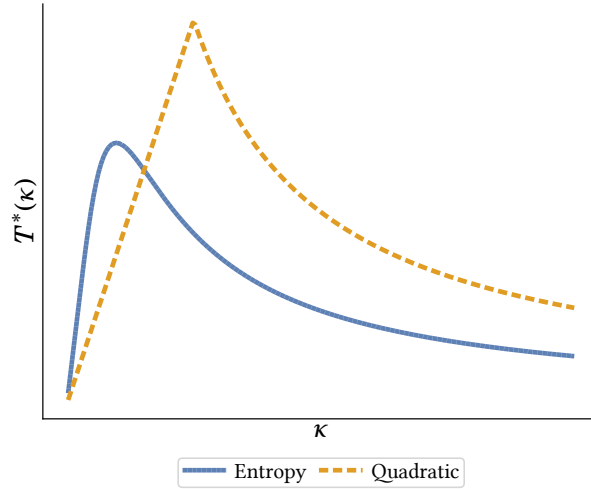


Figure 3: Other Information Costs

$\bar{p} = 1 - \underline{p}$. The observation follows from the first-order condition $2 - \kappa c'(p) = 0$.

While this implies that single-peakedness of $T^*(\kappa)$ is not generic, the most common posterior-based information cost functions do have that $c(p)/c'(p)$ is quasiconcave on $[1/2, 1]$, e.g. entropy costs, $c(p) = p \ln(p) + (1 - p) \ln(1 - p) - \ln(1/2)$, and quadratic or variance costs, $c(p) = (p - p_0)^2$ — see [Figure 3](#).

6. Identifying Complexity

Finally, we turn to the question that often motivates the use of response times in the first place: identifying complexity. As complexity is almost tantamount to mistakes and suboptimal choices, being able to say when a problem is more complex than another is a necessary first step in understanding what makes a problem complex and how to make it simpler (or more complex, depending on the goal). This is well exemplified by many everyday-life situations, going from choosing insurance plans to comparing phone features.

In this section we take the stance of an analyst considering how to assess whether one problem is more or less complex than another. If the analyst knows which choices are optimal, then they can use the decision-maker's accuracy to infer the complexity of the problem, since more complex problems are associated with lower accuracy. This is the approach often taken in experimental work, where the analyst knows the true state of the world, and the decision-maker's accuracy is then used to infer the complexity of the problem.

However, in many situations, the analyst does not know which choices are optimal. For instance, suppose that one is interested in assessing which features of a particular mechanism that elicits private information make it more complex. In this case, which choices are optimal will typically depend exactly on the private information the designer is trying to elicit. Furthermore, our earlier results highlight the limitations of using expected response time as a measure of problem complexity.¹⁹ Hence, we would like some behavioural measure that is monotone in complexity that can be easily observed, aggregated across heterogeneous populations, and used to infer the complexity of a problem.

We propose an alternative way forward, which relies on distorting incentives. To be more specific, suppose that our decision-maker in problem A is choosing between alternatives a and b and in problem \hat{A} is choosing between alternatives \hat{a} and \hat{b} . In the context of our model, we would say that, all else equal, problem A is more complex than problem \hat{A} if the latter has a higher signal-to-noise ratio, $\hat{\mu}/\hat{\sigma}$, than the former.

Suppose that one could simply pay 1 cent when the decision-maker chooses alternative b in A and \hat{b} in \hat{A} . This would marginally increase the decision-maker's indifference point, \tilde{p} , without significantly affecting anything else, and so one would expect b and \hat{b} to be now chosen more often in both problems. What our main result in this section shows is that this small incentive distortion has a greater effect the more complex the problem is.

Theorem 2. *Suppose $\mathbb{P}(\alpha_\tau = b) \in (0, 1)$. Then $\frac{\partial^2}{\partial \tilde{p} \partial (\mu/\sigma)} \mathbb{P}(\alpha_\tau = b) < 0$.*

Figure 4 clearly illustrates the result: the higher the indifference point \tilde{p} , the more likely the decision-maker is to choose b . The effect of problem complexity is then to affect how much more the decision-maker is likely to choose b when the indifference point is increased. It is immediate that the more complex the problem is, insofar as the the choice probability is interior, the stronger the dependence of the choices on the indifference point is, and the steeper the lines are. This is captured by the cross-partial derivative $\partial^2 \mathbb{P}(\alpha_\tau = b) / (\partial \tilde{p} \partial (\mu/\sigma))$ being positive. The proof of this fact follows by obtaining a closed-form characterisation of this derivative using implicit differentiation. The important point is that this result is robust to aggregating over heterogenous agents with different preferences, insofar as these are distributed the same across the two problems.

¹⁹We found that this non-monotonicity is present even if we consider other moments of the distribution, such as variance, even if the population is homogeneous in preferences and ability.

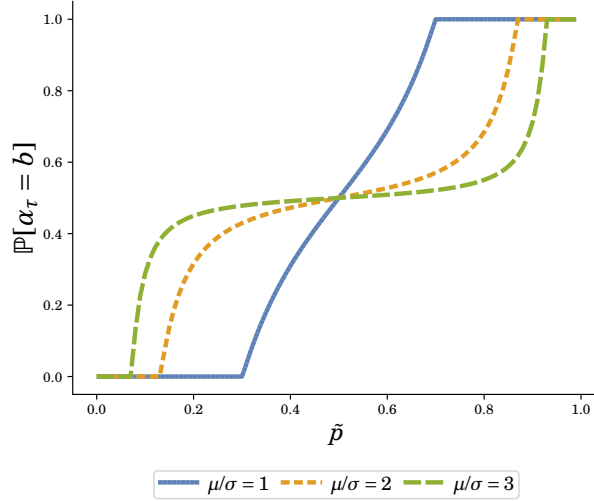


Figure 4: Choice, Problem Complexity, and Payoff Manipulation

7. Final Remarks

We believe that our results have important implications for the use of response times in the social sciences. While response times provide an important source of information that is able to shed light on the cognitive processes underlying decision-making, our results support existing evidence that, on their own, they are not a reliable indicator of problem complexity or ability. Instead, we propose to use small incentive distortions to infer problem complexity. These are easy to implement in real-life settings and could be used to study which features make specific problems more or less complex, and evaluate if particular policies improve them. In current ongoing, [Gonçalves et al. \(2024\)](#) not only corroborate the results in [Theorem 1](#) — already evidenced in the above-referenced literature — but also demonstrate that incentive manipulation can indeed be used to reveal problem complexity.

Our results also highlight the importance of recognising how complex a problem is, as well as one’s own ability in tackling problems of a particular nature. Several questions emerge naturally. Can we infer problem complexity via incentive manipulation also when the decision-maker is not immediately aware of the problem’s complexity? Would this kind of uncertainty, leading to individuals exerting too much or too little effort, be able to rationalise phenomena akin to over- and underconfidence? Are agents with different abilities able to learn to self-sort into tackling problems (activities or jobs) regarding which they hold a comparative advantage? We find these to be promising avenues for future research.

8. References

- Achtziger, A., and C. Alós-Ferrer.** 2013. “Fast or Rational? A Response-Times Study of Bayesian Updating.” *Management Science* 60 (4): 923–938. 10.1287/mnsc.2013.1793.
- Agarwal, S., R. J. Rosen, and V. Yao.** 2016. “Why Do Borrowers Make Mortgage Refinancing Mistakes?” *Management Science* 62 (12): 3494–3509. 10.1287/mnsc.2015.2272.
- Agranov, M., A. Caplin, and C. Tergiman.** 2015. “Naive play and the process of choice in guessing games.” *Journal of the Economic Science Association* 1 146–157. 10.1007/s40881-015-0003-5.
- Alós-Ferrer, C.** 2018. “A Dual-Process Diffusion Model.” *Journal of Behavioral Decision Making* 31 203–218. 10.1002/bdm.1960.
- Alós-Ferrer, C., and J. Bruckenmaier.** 2021. “Cognitive sophistication and deliberation times.” *Experimental Economics* 24 558–592. 10.1007/s10683-020-09672-w.
- Alós-Ferrer, C., E. Fehr, and N. Netzer.** 2021. “Time Will Tell: Recovering Preferences When Choices Are Noisy.” *Journal of Political Economy* 129 (6): 1828–1877. 10.1086/713732.
- Alós-Ferrer, C., and M. Garagnani.** 2022. “Strength of preference and decisions under risk.” *Journal of Risk and Uncertainty* 64 309–329. 10.1007/s11166-022-09381-0.
- Banovetz, J., and R. Oprea.** 2023. “Complexity and Procedural Choice.” *American Economic Journal: Microeconomics* 15 (2): 384–413. 10.1257/mic.20210032.
- Bhargava, S., G. Loewenstein, and J. Sydnor.** 2017. “Choose to Lose: Health Plan Choices from a Menu with Dominated Option.” *Quarterly Journal of Economics* 132 (3): 1319–1372. 10.1093/qje/qjx011.
- Bobtcheff, C., and R. Levy.** 2017. “More Haste, Less Speed? Signaling through Investment Timing.” *American Economic Journal: Microeconomics* 9 (3): 148–186. 10.1257/mic.20160200.
- Bogacz, R., E.-J. Wagenmakers, B. Forstmann, and S. Nieuwenhuis.** 2010. “The neural basis of the speed-accuracy tradeoff.” *Trends in Neurosciences* 33 (1): 10–16. 10.1016/j.tins.2009.09.002.
- Börgers, T., and J. Li.** 2019. “Strategically Simple Mechanisms.” *Econometrica* 87 (6): 2003–2035. 10.3982/ECTA15897.
- Borodin, A. N., and P. Salminen.** 2002. *Handbook of Brownian Motion Facts and Formulae*. Springer, 2nd edition.
- Branco, F., M. Sun, and J. Villas-Boas.** 2012. “Optimal Search for Product Information.” *Management Science* 58 (11): 2037–2056. 10.1287/mnsc.1120.1535.
- Callander, S.** 2011. “Searching and Learning by Trial and Error.” *American Economic Review* 111 (6): 2277–2308. 10.1257/aer.101.6.2277.

- Caplin, A., D. Csaba, J. Leahy, and O. Nov.** 2020. “Rational Inattention, Competitive Supply, and Psychometrics.” *Quarterly Journal of Economics* 135 (3): 1681–1724. 10.1093/qje/qjaa011.
- Caplin, A., and M. Dean.** 2015. “Revealed Preference, Rational Inattention, and Costly Information Acquisition.” *American Economic Review* 105 (7): 2183–2203. 10.1257/aer.20140117.
- Caplin, A., M. Dean, and J. Leahy.** 2022. “Rationally Inattentive Behavior: Characterizing and Generalizing Shannon Entropy.” *Journal of Political Economy* 130 (6): 1–40. 10.1086/719276.
- Caplin, A., and D. Martin.** 2016. “Decision Times Reveal Private Information in Strategic Settings: Evidence from Bargaining Experiments.” *Economic Inquiry* 54 (2): 1274–1282. 10.1111/ecin.12294.
- Carvalho, L., and D. Silverman.** 2017. “Complexity and Sophistication.” *Working Paper* 1–47.
- Choi, J. J., D. Laibson, and B. C. Madrian.** 2011. “\$100 Bills on the Sidewalk: Suboptimal Investment in 401(k) Plans.” *Review of Economics and Statistics* 93 (3): 748–763. 10.1162/REST_a_00100.
- Clithero, J.** 2018. “Improving out-of-sample predictions using response times and a model of the decision process.” *Journal of Economic Behavior and Organization* 148 344–375. 10.1016/j.jebo.2018.02.007.
- Crandall, M. G., H. Ishii, and P.-L. Lions.** 1992. “User’s Guide to Viscosity Solutions of Second Order Partial Differential Equations.” *Bulletin of the American Mathematical Society* 27 1–67. 10.1090/S0273-0979-1992-00266-5.
- Darriet, E., M. Guille, J.-C. Vergnaud, and M. Shimizu.** 2020. “Money illusion, financial literacy and numeracy: Experimental evidence.” *Journal of Economic Psychology* 76 102211. 10.1016/j.joep.2019.102211.
- Dodonova, Y. A., and Y. S. Dodonov.** 2013. “Faster on easy items, more accurate on difficult ones: Cognitive ability and performance on a task of varying difficulty.” *Intelligence* 41 (1): 1–10. 10.1016/j.intell.2012.10.003.
- Drugowitsch, J., R. Moreno-Bote, A. K. Churchland, M. N. Shadlen, and A. Pouget.** 2012. “The Cost of Accumulating Evidence in Perceptual Decision Making.” *The Journal of Neuroscience* 32 (11): 3612–28. 10.1523/JNEUROSCI.4010-11.2012.
- Dvoretzky, A., J. Kiefer, and J. Wolfowitz.** 1953. “Sequential Decision Problems for Processes with Continuous time Parameter. Testing Hypotheses.” *Annals of Mathematical Statistics* 24 (2): 254–264. 10.1214/aoms/1177729031.
- Ellison, G., and A. Wolitzky.** 2012. “A Search Cost Model of Obfuscation.” *RAND Journal of Economics* 43 (3): 417–441. 10.1111/j.1756-2171.2012.00180.x.

- Enke, B., and T. Graeber.** 2023. “Cognitive Uncertainty.” *Quarterly Journal of Economics* 138 (4): 2021–2067. 10.1093/qje/qjad025.
- Esteban-Casanelles, T., and D. Gonçalves.** 2022. “The Effect of Incentives on Choices and Beliefs in Games: An Experiment.” *Working Paper*.
- Ethier, S. N., and T. G. Kurtz.** 2005. *Markov Processes: Characterization and Convergence*. John Wiley & Sons, Inc..
- Fehr, E., and A. Rangel.** 2011. “Neuroeconomic Foundations of Economic Choice—Recent Advances.” *Journal of Economic Perspectives* 25 (4): 3–30. 10.1257/jep.25.4.3.
- Ford, D. L.** 1975. “Decision Time as a Function of Task Complexity: Comments on Hogarth’s Paper.” In *Utility, Probability, and Human Decision Making*, edited by Wendt, D., and C. Vlek 339–343, Dordrecht: Springer, . 10.1007/978-94-010-1834-0_19.
- Forstmann, B., R. Ratcliff, and E.-J. Wagenmakers.** 2016. “Sequential Sampling Models in Cognitive Neuroscience: Advantages, Applications, and Extensions.” *Annual Review of Psychology* 67 (1): 641–666. 10.1146/annurev-psych-122414-033645.
- Frydman, C., and L. Jin.** 2022. “Efficient Coding and Risky Choice.” *Quarterly Journal of Economics* 137 (1): 161–213. 10.1093/qje/qjab031.
- Frydman, C., and S. Nunnari.** 2023. “Coordination with Cognitive Noise.” *Working Paper*.
- Fudenberg, D., P. Strack, and T. Strzalecki.** 2018. “Speed, Accuracy, and the Optimal Timing of Choices.” *American Economic Review* 108 (12): 3651–3684. 10.1257/aer.20150742.
- Gill, D., and V. Prowse.** 2023. “Strategic Complexity and the Value of Thinking.” *The Economic Journal* 133 (650): 761–786. 10.1093/ej/ueac070.
- Gold, J., and M. Shadlen.** 2001. “Neural computations that underlie decisions about sensory stimuli.” *Trends in Cognitive Sciences* 5 (1): 10–16. 10.1016/s1364-6613(00)01567-9.
- Goldhammer, F., J. Naumann, A. Stelter, K. Tóth, H. Rölke, and E. Klieme.** 2014. “The time on task effect in reading and problem solving is moderated by task difficulty and skill: Insights from a computer-based large-scale assessment.” *Journal of Educational Psychology* 106 (3): 608–626. 10.1037/a0034716.
- Gonçalves, D.** 2023. “Sequential Sampling Equilibrium.” *Working Paper* 1–57. 10.48550/arXiv.2212.07725.
- Gonçalves, D., S. Nunnari, and J. Zarate-Pina.** 2024. “Revealing Complexity.” *Work in Progress*.
- Halac, M., N. Kartik, and Q. Liu.** 2016. “Optimal Contracts for Experimentation.” *Review of Economic Studies* 83 (3): 1040–1091. 10.1093/restud/rdw013.
- Hogarth, R. M.** 1975. “Decision Time as a Function of Task Complexity.” In *Utility, Probability,*

and Human Decision Making, edited by Wendt, D., and C. Vlek 321–338, Dordrecht: Springer, .
10.1007/978-94-010-1834-0_19.

Horenstein, A., and K. Grabiszewski. 2022. “Measuring tree complexity with response times.” *Journal of Behavioral and Experimental Economics* 98 101876. 10.1016/j.socec.2022.101876.

Jehiel, P. 2011. “Manipulative auction design.” *Theoretical Economics* 6 185–217. 10.3982/TE687.

Junghaenel, D. U., S. Schneider, B. Orriens, H. Jin, P.-J. Lee, A. Kapteyn, E. Meijer, E. Zelinski, R. Hernandez, and A. A. Stone. 2023. “Inferring Cognitive Abilities from Response Times to Web-Administered Survey Items in a Population-Representative Sample.” *Journal of Intelligence* 11 (1): 1–25. 10.3390/jintelligence11010003.

Kalaycı, K., and M. Serra-Garcia. 2016. “Complexity and Biases.” *Experimental Economics* 19 31–50. 10.1007/s10683-015-9434-3.

Kamenica, E., and M. Gentzkow. 2011. “Bayesian Persuasion.” *American Economic Review* 101 (6): 2590–2615. 10.1257/aer.101.6.2590.

Keys, B. J., D. G. Pope, and J. C. Pope. 2016. “Failure to Refinance.” *Journal of Financial Economics* 122 (3): 482–499. 10.1016/j.jfineco.2016.01.031.

Khaw, M., Z. Li, and M. Woodford. 2021. “Cognitive Imprecision and Small-Stakes Risk Aversion.” *Review of Economic Studies* 88 (4): 1979–2013. 10.1093/restud/rdaa044.

Krajbich, I., D. Lu, C. Camerer, and A. Rangel. 2012. “The attentional drift-diffusion model extends to simple purchasing decisions.” *Frontiers in Psychology* 3 1–18. 10.3389/fpsyg.2012.00193.

Kyllonen, P. C., and J. Zu. 2016. “Use of Response Time for Measuring Cognitive Ability.” *Journal of Intelligence* 4 (4): 1–29. 10.3390/jintelligence4040014.

Li, J., and P. Dworzak. 2022. “Are Simple Mechanisms Optimal when Agents are Unsophisticated?” *Working Paper* 1–51.

Li, S. 2017. “Obviously Strategy-Proof Mechanisms.” *American Economic Review* 107 (11): 3257–3287. 10.1257/aer.20160425.

Lipnowski, E., and D. Ravid. 2023. “Predicting Choice from Information Costs.” *Working Paper*.

Martínez-Marquina, A., M. Niederle, and E. Vespa. 2019. “Failures in Contingent Reasoning: The Role of Uncertainty.” *American Economic Review* 109 (10): 3437–3474. 10.1257/aer.20171764.

Matějka, F., and A. McKay. 2015. “Rational Inattention to Discrete Choices: A New Foundation for the Multinomial Logit Model.” *American Economic Review* 105 (1): 272–298. 10.1257/aer.20130047.

Moritz, B., E. Siemsen, and M. Kremer. 2014. “Judgmental Forecasting: Cognitive Reflection and Decision Speed.” *Production and Operations Management* 23 (7): 1146–1160. 10.1111/poms.

12105.

- Morris, S., and P. Strack.** 2019. “The Wald Problem and the Equivalence of Sequential Sampling and Static Information Costs.” *Working Paper*. 10.2139/ssrn.2991567.
- Moscarini, G., and L. Smith.** 2001. “The Optimal Level of Experimentation.” *Econometrica* 69 (6): 1629–1644. 10.1111/1468-0262.00259.
- Mosteller, F., and P. Noguee.** 1951. “An Experimental Measurement of Utility.” *Journal of Political Economy* 59 (5): 371–404.
- Øksendal, B., and K. Reikvam.** 1998. “Viscosity solutions of optimal stopping problems.” *Stochastics and Stochastic Reports* 62 285–301. 10.1080/17442509808834137.
- Onken, J., R. Hastie, and W. Revelle.** 1985. “Individual differences in the use of simplification strategies in a complex decision-making task.” *Journal of Experimental Psychology: Human Perception and Performance* 11 (1): 14–27. 10.1037/0096-1523.11.1.14.
- Pomatto, L., P. Strack, and O. Tamuz.** 2023. “The Cost of Information: The Case of Constant Marginal Costs.” *American Economic Review* 113 (5): 1360–1393. 10.1257/aer.20190185.
- Puri, I.** 2023. “Simplicity and Risk.” *Working Paper* 1–91.
- Pycia, M., and P. Troyan.** 2023. “A Theory of Simplicity in Games and Mechanism Design.” *Econometrica* 91 (4): 1495–1526. 10.3982/ECTA16310.
- Rabin, M., and G. Weizsäcker.** 2009. “Narrow Bracketing and Dominated Choices.” *American Economic Review* 99 (4): 1508–1534. 10.1257/aer.99.4.1508.
- Ratcliff, R.** 1978. “A theory of memory retrieval.” *Psychological Review* 85 (2): 59–108. 10.1037/0033-295X.85.2.59.
- Ratcliff, R., and G. McKoon.** 2008. “The Diffusion Decision Model: Theory and Data for Two-Choice Decision Tasks.” *Neural Computation* 20 873–922.
- Ratcliff, R., P. Smith, S. Brown, and G. McKoon.** 2016. “Diffusion Decision Model: Current Issues and History.” *Trends in Cognitive Sciences* 20 (4): 260–281. 10.1016/j.tics.2016.01.007.
- Recalde, M. P., A. Riedl, and L. Vesterlund.** 2018. “Error-prone inference from response time: The case of intuitive generosity in public-good games.” *Journal of Public Economics* 160 132–147. 10.1016/j.jpubeco.2018.02.010.
- Rees-Jones, A., and D. Taubinsky.** 2020. “Measuring “Schmeduling”.” *Review of Economic Studies* 87 (5): 2399–2438. 10.1093/restud/rdz045.
- Reshidi, P., A. Lizzeri, L. Yariv, J. Chan, and W. Suen.** 2024. “Sequential Sampling by Individuals and Groups: An Experimental Study.” *Working Paper* 1–20.
- Rubinstein, A.** 2016. “A Typology of Players: Between Instinctive and Contemplative.” *Quarterly*

Journal of Economics 117 (523): 859–890. 10.1093/qje/qjw008.

Schotter, A., and I. Trevino. 2021. “Is response time predictive of choice? An experimental study of threshold strategies.” *Experimental Economics* 24 87–117. 10.1007/s10683-020-09651-1.

Shadlen, M., and D. Shohamy. 2016. “Decision Making and Sequential Sampling from Memory.” *Neuron* 90 (5): 927–939. 10.1016/j.neuron.2016.04.036.

Shiryaev, A. N. 2007. In *Optimal Stopping Rules*, 2nd edition, Springer.

Shushruth, S., A. Zylberberg, and M. Shadlen. 2022. “Sequential sampling from memory underlies action selection during abstract decision-making.” *Current Biology* 32 (9): 1949–1960.e5. 10.1016/j.cub.2022.03.014.

Sinaiko, A. D., and R. A. Hirth. 2011. “Consumers, health insurance and dominated choices.” *Journal of Health Economics* 30 (2): 450–457. 10.1016/j.jhealeco.2010.12.008.

Spiegler, R. 2016. “Choice Complexity and Market Competition.” *Annual Review of Economics* 8 1–25. 10.1146/annurev-economics-070615-115216.

Spiliopoulos, L., and A. Ortmann. 2018. “The BCD of response time analysis in experimental economics.” *Experimental Economics* 21 383–433. 10.1007/s10683-017-9528-1.

Wald, A. 1947. “Foundations of a General Theory of Sequential Decision Functions.” *Econometrica* 15 (4): 279–313. 10.2307/1905331.

Wilcox, N. 1993. “Lottery choice: Incentives complexity and decision time.” *Economic Journal* 103 (421): 1397–1417. 10.2307/2234473.

Wright, G., and P. Ayton. 1988. “Decision time, subjective probability, and task difficulty.” *Memory & Cognition* 16 (2): 176–185.

Appendix A. Omitted Proofs

A.1. Proof of Lemma 1

Proof. We focus on the $V(p) = \delta V_a(p) + u(b, p)$ and $\arg\max_{\tau} \mathbb{E}[v(p_{\tau}) - c\tau] = \arg\max_{\tau} \mathbb{E}[v_a(p_{\tau}) - (c/\delta)\tau]$, since the other case is analogous. Note that $u(\alpha, p) = \delta u_a(\alpha, p) + u(b, p)$, for $\alpha = a, b$. For any stopping time τ which is finite almost surely, $\mathbb{E}[v(p_{\tau}) - c\tau] = \delta \mathbb{E}[(p_{\tau} - \tilde{p})^+] - (c/\delta)\tau + \mathbb{E}[u(b, p_{\tau})] = \delta \mathbb{E}[v_a(p_{\tau}) - (c/\delta)\tau] + \mathbb{E}[u(b, p_{\tau})] = \delta \mathbb{E}[v_a(p_{\tau}) - (c/\delta)\tau] + u(b, p_0)$, where we used optional stopping theorem with the fact that beliefs are a martingale and lie in a compact set. Any optimal stopping time must be finite almost surely, as otherwise the expected value is negative infinity. Hence, $\arg\max_{\tau} \mathbb{E}[v(p_{\tau}) - c\tau] = \arg\max_{\tau} \delta \mathbb{E}[v_a(p_{\tau}) - (c/\delta)\tau] + u(b, p_0) = \arg\max_{\tau} \mathbb{E}[v_a(p_{\tau}) -$

$(c/\delta)\tau]$ and $\max_{\tau} \mathbb{E}[v(p_{\tau}) - c\tau] = \delta \max_{\tau} \mathbb{E}[v_{\alpha}(p_{\tau}) - (c/\delta)\tau] + u(b, p_0) = \delta V_{\alpha}(p_0) + u(b, p_0)$. \square

A.2. Proof of Lemma 2

Proof. We focus on V_{α} ; the proof for V_b is analogous. Let $g = (g_t)$ denote a stochastic process adapted to the natural filtration induced by $(X_t)_t$. Note that $V_{\alpha}(p_0) = \sup_{g, \tau} \mathbb{E}[u_{\alpha}(g_{\tau}, p_{\tau}) - c\tau]$. Letting $W_{\alpha}(p_0, \tau, g) := \mathbb{E}[u_{\alpha}(g_{\tau}, p_{\tau}) - c\tau] = \mathbb{E}[u_{\alpha}(g_{\tau}, \theta) - c\tau]$, we have $V_{\alpha}(p_0) = \sup_{g, \tau} W_{\alpha}(p_0, \tau, g)$. Since W_{α} is linear in p_0 , V_{α} is convex in p_0 .

Since $V_{\alpha} \geq v_{\alpha}$, both V_{α} and v_{α} are convex, v_{α} is increasing, and $V_{\alpha}(1) = v_{\alpha}(1)$ and $V_{\alpha}(0) = v_{\alpha}(0)$, we must have that V_{α} is also increasing. \square

A.3. Proof of Lemma 3

In text.

A.4. Proof of Lemma 4

Proof. From Lemma 3, there are $0 < \underline{p} \leq \bar{p} < 1$ such that $\tau^* := \inf\{t \geq 0 : p_t \notin (\underline{p}, \bar{p})\}$. Given this, the fact that V is the unique viscosity solution of the free-boundary problem follows from Øksendal and Reikvam (1998, Proposition 3.1). \square

A.5. Proof of Proposition 1

Proof. Given Lemma 1, the free-boundary problem can be equivalently given by

$$0 = -\tilde{c} + (\mu/\sigma)^2 (p(1-p))^2 V_{\alpha}''(p) \quad \text{on } p \in (\underline{p}, \bar{p}) \quad (\text{HJBa})$$

$$0 = V_{\alpha}(p) - v_{\alpha}(p) \quad \text{on } p \in \{\underline{p}, \bar{p}\}, \quad (\text{BCa})$$

where $\tilde{c} = c/(2\delta)$ and acknowledging that $V_{\alpha}'' = V''/\delta$.

As V is a viscosity solution of the problem in Lemma 4, on (\underline{p}, \bar{p}) , by definition of the boundaries, $V_{\alpha} > v_{\alpha}$. Moreover, by Lemma 2, V_{α} is convex, and so, by Alexandrov's theorem, V_{α}'' exists a.e. and left and right derivatives exist everywhere.

Suppose that V_{α} has a kink at $p \in (\underline{p}, \bar{p})$, i.e. $V_{\alpha}'(p^+) \neq V_{\alpha}'(p^-)$. Then, by convexity, $V_{\alpha}'(p^+) > V_{\alpha}'(p^-)$, and, for any $k \in (V_{\alpha}'(p^-), V_{\alpha}'(p^+))$ and significantly high $M > 0$, $V_{\alpha}(p') \geq V_{\alpha}(p) + k(p' - p) + (M/2)(p' - p)^2 =: \psi(p')$ for p' in a neighborhood of p . Then, we would obtain that $\tilde{c} - (\mu/\sigma)^2 (p(1-p))^2 M < 0$ for sufficiently high M , contradicting the fact that V_{α} is a viscosity supersolution.

Hence, V_a is \mathcal{C}^1 except possibly at $\{\underline{p}, \bar{p}\}$.

Since V_a'' exists a.e., from Crandall et al. (1992, p. 15), the HJB will hold with equality whenever V_a is twice differentiable, i.e. almost everywhere on (\underline{p}, \bar{p}) . Furthermore, $0 \leq V_a'' \leq (\mu/\sigma)^{-2} \tilde{c} \max\{(p(1-p))^{-2}, (\bar{p}(1-\bar{p}))^{-2}\}$, and so V_a'' is bounded whenever it exists. Take any sequence $(p_n) \subset (\underline{p}, \bar{p})$ such that $p_n \rightarrow p^* \in (\underline{p}, \bar{p})$. Then, there is a subsequence (p_m) such that $V_a'(p_m) \rightarrow V_{a,\infty}'$, $V_a''(p_m) \rightarrow V_{a,\infty}''$, and, by continuity, $0 = -\tilde{c} + (\mu/\sigma)^2 (p^*(1-p^*))^2 V_{a,\infty}''$. Hence $(V_{a,\infty}', V_{a,\infty}'')$ belongs to both the super- and subjet of V_a at p^* , implying that V_a is twice continuously differentiable everywhere except possibly at $\{\underline{p}, \bar{p}\}$.

Then, for some $\beta_0, \beta_1 \in \mathbb{R}$ to be determined

$$V_a(p) = \frac{\tilde{c}}{(\mu/\sigma)^2} \frac{\exp(\ell(p)) - 1}{\exp(\ell(p)) + 1} \ell(p) + \frac{\exp(\ell(p))}{\exp(\ell(p)) + 1} \beta_1 + \beta_0,$$

$$V_a'(p) = \frac{\tilde{c}}{(\mu/\sigma)^2} (\exp(\ell(p)) - \exp(-\ell(p)) + 2\ell(p)) + \beta_1.$$

As the optimal stopping time is finite almost surely, $0 < \underline{p} \leq \bar{p} < 1$ and so $V_a''(\underline{p}^+), V_a''(\bar{p}^-) > 0 = v_a''(p) \forall p \neq \bar{p}$, which implies that the boundary condition is binding at the boundaries and therefore $V_a \in \mathcal{C}^0([0, 1])$ and convex.

We guess and verify that $V \in \mathcal{C}^1([0, 1])$. From $V_a'(\underline{p}^+) = v_a'(\underline{p}^-) = 0$ and $V_a'(\bar{p}^-) = v_a'(\bar{p}^+) = 1$, one obtains $-\beta_1 = \frac{\tilde{c}}{(\mu/\sigma)^2} (\exp(\underline{\ell}) - \exp(-\underline{\ell}) + 2\underline{\ell}) = \frac{\tilde{c}}{(\mu/\sigma)^2} (\exp(\bar{\ell}) - \exp(-\bar{\ell}) + 2\bar{\ell}) - 1$, from which (1) follows. Then,

$$0 = V_a(\underline{p}) - v_a(\underline{p}) = \frac{\tilde{c}}{(\mu/\sigma)^2} ((1 - \exp(\underline{\ell})) - \underline{\ell}) + \beta_0,$$

$$0 = V_a(\bar{p}) - v_a(\bar{p}) = \frac{\tilde{c}}{(\mu/\sigma)^2} ((1 - \exp(\bar{\ell})) - \bar{\ell}) + \beta_0 + \tilde{p},$$

which yields

$$((1 - \exp(\underline{\ell})) - \underline{\ell}) - ((1 - \exp(\bar{\ell})) - \bar{\ell}) = \frac{(\mu/\sigma)^2}{\tilde{c}} \tilde{p}.$$

Replacing $\frac{(\mu/\sigma)^2}{\tilde{c}}$ with (1) and rearranging the expression delivers (2).

Given $\underline{\ell} \leq \bar{\ell}$, (2) implies $\underline{\ell} \leq \tilde{\ell} \leq \bar{\ell}$. Then, both (1) and (2) implicitly define $\underline{\ell}$ as a function of $\bar{\ell}$, which we denote by $\underline{\ell}_1(\cdot)$ and $\underline{\ell}_2(\cdot)$, respectively. From the implicit function theorem, we obtain that $\underline{\ell}_1$ is strictly increasing and $\underline{\ell}_2$ strictly decreasing. Moreover, $\underline{\ell}_2(\bar{\ell}) = \bar{\ell}$, whereas since the right-hand side of (1) is decreasing in $\underline{\ell}$ and that it is zero for $\underline{\ell} = \bar{\ell}$, we obtain that $\underline{\ell}_1(\bar{\ell}) < \bar{\ell}$. Additionally, straightforward manipulation of (1) and (2) shows that for some large enough $\bar{\ell}$, $\underline{\ell}_1(\bar{\ell}) > 0 > \underline{\ell}_2(\bar{\ell})$, hence there is a unique $\bar{\ell}$ such that $\underline{\ell}_1(\bar{\ell}) = \underline{\ell}_2(\bar{\ell})$. This shows that the system

formed by (1) and (2) does have a (unique) solution $(\underline{\ell}, \bar{\ell})$ such that $\underline{\ell} \leq \tilde{\ell} \leq \bar{\ell}$, thus verifying the conjecture that $V_\alpha \in \mathcal{C}^1([0, 1])$ and concluding the proof. \square

A.6. Proof of Theorem 1

Proof. We focus on the case in which $\tilde{p}, p_0 \in (\underline{p}, \bar{p})$, since if this were not the case, the decision-maker with prior p_0 would prefer to stop immediately.

Letting $\ell(p) := \ln(p/(1-p))$, we first note that the information process X_t induces a process for log-odds beliefs given by $\ell_t := 2\frac{\mu}{\sigma^2}X_t$. Second, we observe that, given θ , $\exp(-(2s(\theta)\mu/\sigma^2)X_t)$ is a martingale and as X_τ is therefore bounded, we can apply the optional stopping theorem (Ethier and Kurtz, 2005, Ch. 2, Theorem 2.3) to obtain $1 = \exp(-(2s(\theta)\mu/\sigma^2)X_0) = \mathbb{E}[\exp(-(2s(\theta)\mu/\sigma^2)X_\tau) | \theta] = \mathbb{E}[\exp(-s(\theta)\ell_\tau) | \theta]$.

Recall that $\bar{\ell} := \ell(\bar{p})$, $\underline{\ell} := \ell(\underline{p})$, and $\tilde{\ell} := \ell(\tilde{p})$. As $\mathbb{P}(\alpha_\tau = a | \theta) = \mathbb{P}(\ell_\tau = \bar{\ell} | \theta) = 1 - \mathbb{P}(\ell_\tau = \underline{\ell} | \theta)$ and $\mathbb{E}[\exp(-s(\theta)\ell_\tau) | \theta] = \mathbb{P}(\ell_\tau = \bar{\ell} | \theta)\exp(-s(\theta)\bar{\ell}) + \mathbb{P}(\ell_\tau = \underline{\ell} | \theta)\exp(-s(\theta)\underline{\ell})$, we have $\mathbb{P}(\alpha_\tau = \theta) = \frac{1}{2}(\mathbb{P}(\alpha_\tau = a | \theta = a) + \mathbb{P}(\alpha_\tau = b | \theta = b)) = \frac{1}{2} \left(\exp(\bar{\ell}) \frac{1 - \exp(\underline{\ell})}{\exp(\bar{\ell}) - \exp(\underline{\ell})} + \frac{\exp(\bar{\ell}) - 1}{\exp(\bar{\ell}) - \exp(\underline{\ell})} \right)$. $\mathbb{P}(\alpha_\tau = \theta)$ is strictly increasing in $\bar{\ell}$ and decreasing in $\underline{\ell}$. A straightforward application of the implicit function theorem to the system formed by (1) and (2) reveals that $(\bar{\ell}, -\underline{\ell})$ is increasing in μ/σ , implying that accuracy, $\mathbb{P}(\alpha_\tau = \theta)$, is increasing in μ/σ .

We then use Wald's identity to obtain $\mathbb{E}[\ell_\tau | \theta] = 2 \left(\frac{\mu}{\sigma}\right)^2 s(\theta)\mathbb{E}[\tau | \theta]$. As $\mathbb{E}[\ell_\tau | \theta] = \mathbb{P}(\ell_\tau = \bar{\ell} | \theta)\bar{\ell} + \mathbb{P}(\ell_\tau = \underline{\ell} | \theta)\underline{\ell}$, we then obtain

$$\begin{aligned} \mathbb{E}[\tau] &= \frac{1}{4(\mu/\sigma)^2} \left(\mathbb{P}(\ell_\tau = \bar{\ell} | \theta = a) - \mathbb{P}(\ell_\tau = \bar{\ell} | \theta = b) \right) (\bar{\ell} - \underline{\ell}) \\ &= \frac{1}{4(\mu/\sigma)^2} \frac{(1 - \exp(\underline{\ell}))(\exp(\bar{\ell}) - 1)}{\exp(\bar{\ell}) - \exp(\underline{\ell})} (\bar{\ell} - \underline{\ell}). \end{aligned} \quad (3)$$

Replacing $(\mu/\sigma)^2$ in (3) with (1) and taking logs we obtain $T(\bar{\ell}, \underline{\ell})$, a strictly concave function of $(\bar{\ell}, \underline{\ell}) \in \mathbb{R}_+ \times \mathbb{R}_-$, satisfying $T(\bar{\ell}, \underline{\ell}) = T(-\underline{\ell}, -\bar{\ell})$. T therefore encodes the dependence of the stopping thresholds on μ/σ , but not the optimal relation between the upper and lower stopping thresholds, dictated by the indifference point \tilde{p} , in its log-odds form $\tilde{\ell}$ in (2).

Since, $\bar{\ell} \downarrow 0$ or $\bar{\ell} \uparrow +\infty$ (resp. $\underline{\ell} \uparrow 0$ or $\underline{\ell} \downarrow -\infty$) implies $\exp(T(\bar{\ell}, \underline{\ell})) \downarrow 0$ for any $\underline{\ell} < 0$ (resp. $\bar{\ell} > 0$), and as $\exp(T(\bar{\ell}, \underline{\ell})) > 0$ for some $\bar{\ell} > 0 > \underline{\ell}$, we have that the expected optimal stopping time is non-monotone.

Without loss of generality, we assume $\tilde{p} \geq 1/2$ (as the other case is symmetric), which implies

$\bar{\ell} > \tilde{\ell} > \underline{\ell}$. Then, from (2), it is immediate that $\bar{\ell} \geq \tilde{\ell} - \underline{\ell} \geq \tilde{\ell} \geq \ell_0 = 0 \geq \underline{\ell}$.

We then rearrange (2) so as to obtain

$$\exp(\tilde{\ell}) = \frac{(\bar{\ell} - \underline{\ell}) + (\exp(\bar{\ell}) - \exp(\underline{\ell}))}{(\bar{\ell} - \underline{\ell}) + (\exp(\bar{\ell}) - \exp(\underline{\ell})) \exp(-\bar{\ell} - \underline{\ell})}, \quad (4)$$

implicitly defining $\underline{\ell}$ as a function of $\bar{\ell}$. Let $L : \bar{\ell} \mapsto L(\bar{\ell}) = (\bar{\ell}, \underline{\ell}(\bar{\ell}))$.

Then,

$$\frac{d^2}{(d\bar{\ell})^2} T(\bar{\ell}, \underline{\ell}(\bar{\ell})) = (D_{\bar{\ell}} L(\bar{\ell}))^T \cdot D_{(\bar{\ell}, \underline{\ell})}^2 T(\bar{\ell}, \underline{\ell}) \cdot D_{\bar{\ell}} L(\bar{\ell}) + D_{\bar{\ell}}^2 L(\bar{\ell}) \cdot D_{(\bar{\ell}, \underline{\ell})} T(\bar{\ell}, \underline{\ell}) \Big|_{\underline{\ell}=L(\bar{\ell})},$$

where $D_x f$ and $D_x^2 f$ denote the first- and second-order derivatives of f with respect to x . Note that $D_{\bar{\ell}} L(\bar{\ell})$ and $D_{\bar{\ell}}^2 L(\bar{\ell})$ are obtained via the implicit function theorem from (4). A straightforward but tedious exercise in examining $\frac{d^2}{(d\bar{\ell})^2} T(\bar{\ell}, \underline{\ell}(\bar{\ell}))$ reveals that it is strictly negative for all $\bar{\ell}, \underline{\ell} : \bar{\ell} > 0 > \underline{\ell} > -\bar{\ell}$, thus showing that the expected optimal stopping time, taken as a function of $\bar{\ell}$, is strictly log-concave in $\bar{\ell} \geq 0$.

We recall that, by the implicit function theorem, using (1) and (2), $(\bar{\ell}, -\underline{\ell})$ is increasing in μ/σ . Hence, we conclude that the expected optimal stopping time, $\mathbb{E}[\tau^*]$, is non-monotone and quasi-concave in μ/σ , and strictly so for all μ/σ such that $\tau^* \neq 0$. \square

A.7. Proof of Proposition 2

Proof. We first note that the optimal effort for agent i , e_i^* , is constant over time and equal to $e_i^* = \lambda_i e^*$, where e^* is such that $e^* c'(e^*) = c(e^*)$. Write $c_i(e_i) := c(e_i/\lambda_i)$. The argument (already in Moscarini and Smith (2001)) is that, on the continuation region, the HJB equation is given by

$$0 = \sup_{e_i > 0} \{-c_i(e_i) + e_i 2(\mu/\sigma)^2 (p(1-p))^2 V''(p)\}.$$

The first-order condition yields $e_i c'_i(e_i) = e_i 2(\mu/\sigma)^2 (p(1-p))^2 V''(p)$, whereas the HJB equation implies $c_i(e_i) = e_i 2(\mu/\sigma)^2 (p(1-p))^2 V''(p)$. Observe that $e_i c'_i(e_i) = c_i(e_i)$ has unique solution and so effort is constant over time (while not stopping).

We then have $e_i c'_i(e_i) = c_i(e_i) \iff e_i/\lambda_i c'(e_i/\lambda_i) = c(e_i/\lambda_i)$, which implies $e_i = \lambda_i e^*$ and $c_i(e_i) = c(e_i/\lambda_i) = c(e^*)$.

Hence, it is as if the decision-maker faces a problem with a constant drift term $\mu_i \equiv \sqrt{\lambda_i} e^* \mu$ and a constant flow cost $c(e^*)$, and so will have constant stopping thresholds as before.

As accuracy is increasing in the signal-to-noise ratio, we have that the higher ability type will

always choose better, as they have a higher effective signal-to-noise ratio, $\mu_H/\sigma = \sqrt{\lambda_H e^*} \mu/\sigma$. In contrast, from **Theorem 1**, the expected stopping time is single-peaked; let its maximum be attained at k^* . Defining $\bar{k} := k^*(\lambda_L e^*)^{-1/2}$ and $\underline{k} := k^*(\lambda_H e^*)^{-1/2}$, the result follows immediately. The restriction $\tau_H^* \neq 0$ guarantees that we are in the part of the domain of λ_H and thus signal-to-noise ratio in which the expected stopping time is strictly quasi-concave (ensuring that the stopping thresholds of the high type are bounded away from 0 and the indifference point). \square

A.8. Proof of Theorem 2

Proof. From the proof of **Theorem 1**, we have that $\mathbb{P}(\alpha_\tau = a \mid \theta) = \mathbb{P}(\ell_\tau = \bar{\ell} \mid \theta)$ and $\mathbb{E}[\exp(-s(\theta)\ell_\tau) \mid \theta] = \mathbb{P}(\ell_\tau = \bar{\ell} \mid \theta) \exp(-s(\theta)\bar{\ell}) + \mathbb{P}(\ell_\tau = \underline{\ell} \mid \theta) \exp(-s(\theta)\underline{\ell})$.

Consequently, $\mathbb{P}(\alpha_\tau = a) = \frac{(1+\exp(\bar{\ell}))(1-\exp(\underline{\ell}))}{2(\exp(\bar{\ell})-\exp(\underline{\ell}))} =: q(\bar{\ell}, \underline{\ell})$.

From (1) and (4), using the implicit function theorem, we obtain $\frac{\partial}{\partial \exp(\bar{\ell})} q(\bar{\ell}, \underline{\ell})$ as a function $q_{\bar{\ell}}$ of $\bar{\ell}$ and $\underline{\ell}$, where

$$\begin{aligned} q_{\bar{\ell}}(\bar{\ell}, \underline{\ell}) &:= \frac{\partial}{\partial \exp(\bar{\ell})} q(\bar{\ell}, \underline{\ell}) = \frac{\partial}{\partial \bar{\ell}} q(\bar{\ell}, \underline{\ell}) \frac{\partial}{\partial \exp(\bar{\ell})} \bar{\ell} + \frac{\partial}{\partial \underline{\ell}} q(\bar{\ell}, \underline{\ell}) \frac{\partial}{\partial \exp(\bar{\ell})} \underline{\ell} \\ &= \frac{\exp(-5\bar{\ell}/2 - 2\underline{\ell})(1 + \exp(\bar{\ell})) \left(\exp(\bar{\ell}) - \exp(\underline{\ell}) + \exp(\bar{\ell} + \underline{\ell})(\bar{\ell} - \underline{\ell}) \right)^2 \operatorname{csch}((\bar{\ell} - \underline{\ell})/2) \operatorname{sech}(\underline{\ell}/2) \sinh(\bar{\ell})}{16(1 - \cosh(\bar{\ell} - \underline{\ell}))(\bar{\ell} - \underline{\ell} + \sinh(\bar{\ell}) - \sinh(\underline{\ell}))} \\ &+ \frac{\exp(-5\underline{\ell}/2 - 2\bar{\ell})(1 + \exp(\underline{\ell})) \left(\exp(\bar{\ell}) - \exp(\underline{\ell}) + \exp(\bar{\ell} + \underline{\ell})(\bar{\ell} - \underline{\ell}) \right)^2 \operatorname{csch}((\bar{\ell} - \underline{\ell})/2)^3 \operatorname{sech}(\bar{\ell}/2) \sinh(\underline{\ell})}{32(\bar{\ell} - \underline{\ell} + \sinh(\bar{\ell}) - \sinh(\underline{\ell}))}, \end{aligned}$$

which, as expected, can be shown to be strictly negative for all $\bar{\ell} > 0 > \underline{\ell}$. Relying again on implicit differentiation, we obtain an explicit characterisation of the cross-partial derivative, $q_{\bar{\ell}, \mu/\sigma} := \frac{\partial^2}{\partial \exp(\bar{\ell}) \partial ((\mu/\sigma)^2)} q = \frac{\partial}{\partial \bar{\ell}} q_{\bar{\ell}} \frac{\partial}{\partial ((\mu/\sigma)^2)} \bar{\ell} + \frac{\partial}{\partial \underline{\ell}} q_{\bar{\ell}} \frac{\partial}{\partial ((\mu/\sigma)^2)} \underline{\ell}$, involving exclusively $\bar{\ell}, \underline{\ell}$. With a careful examination of the resulting function and extremely tedious algebra it can be shown that $q_{\bar{\ell}, \mu/\sigma} > 0$ for all $\bar{\ell} > 0 > \underline{\ell}$.

As $\mu/\sigma > 0$, $\exp(\bar{\ell})$ and $(\mu/\sigma)^2$ are strictly increasing functions of μ/σ and \bar{p} , respectively, and so this implies that $\frac{\partial^2}{\partial \bar{p} \partial (\mu/\sigma)} \mathbb{P}(\alpha_\tau = a) > 0 \iff \frac{\partial^2}{\partial \bar{p} \partial (\mu/\sigma)} \mathbb{P}(\alpha_\tau = b) < 0$. \square

A.9. Proof of Proposition 3

Proof. Let π_I denote a fully informative experiment — i.e. assigning probability $p_0(\theta)$ to posterior belief 1_θ — and π_U a fully uninformative experiment, assigning probability 1 to the prior p_0 .

Take any $\bar{C} > 0$ such that $\exists \bar{\kappa}$ such that $\bar{\kappa} C^*(\bar{\kappa}) > \bar{C}$. Then, there are $\kappa < \bar{\kappa}$ and $\kappa' > \bar{\kappa}$ such that $\kappa C^*(\kappa) \vee \kappa' C^*(\kappa') < \bar{C}$. To see this note that, by Berge's maximum theorem, $\kappa C^*(\kappa)$ is continuous in κ . Taking $\kappa_n \downarrow 0$, $V(\kappa_n) \uparrow V(0) = B(\pi_I)$, and so $V(0) - V(\kappa_n) = \kappa_n C^*(\kappa_n) \downarrow 0$. Take now $\kappa_n \uparrow \infty$. Note that $0 \leq \lim_{n \rightarrow \infty} V(\kappa_n)/\kappa_n = \lim_{n \rightarrow \infty} B(\pi^*(\kappa_n))/\kappa_n - C^*(\kappa_n) = \lim_{n \rightarrow \infty} -C^*(\kappa_n) \leq 0$, and so it must be that $\pi^*(\kappa_n) \rightarrow \pi_U$. Consequently, $\kappa_n C^*(\kappa_n) \leq B(\pi^*(\kappa_n)) \rightarrow 0$, a contradiction. \square

Appendix B. Discounting

In this appendix, we consider a variation of the problem in [Section 2](#) in which the decision-maker, instead of bearing an explicit cost of time, exponentially discounts payoffs at a rate $r > 0$. We make the simplifying assumption that $u(\alpha, \theta) = 1_{\alpha=\theta}$.

Formally, the decision-maker chooses a stopping time τ and a stochastic process (α_t) taking values in A , both adapted to the natural filtration induced by the stochastic process $(X_t)_t$ in order to maximise their expected payoff, $\mathbb{E}[\exp(-r\tau)(1_{\alpha_\tau=\theta})]$, where p_t denotes the posterior belief about θ given X_t , and $v(p) := \max\{p, 1-p\}$. Denote their value function by

$$V(p_0) = \sup_{\tau} \mathbb{E}[\exp(-r\tau)(1_{\alpha_\tau=\theta})].$$

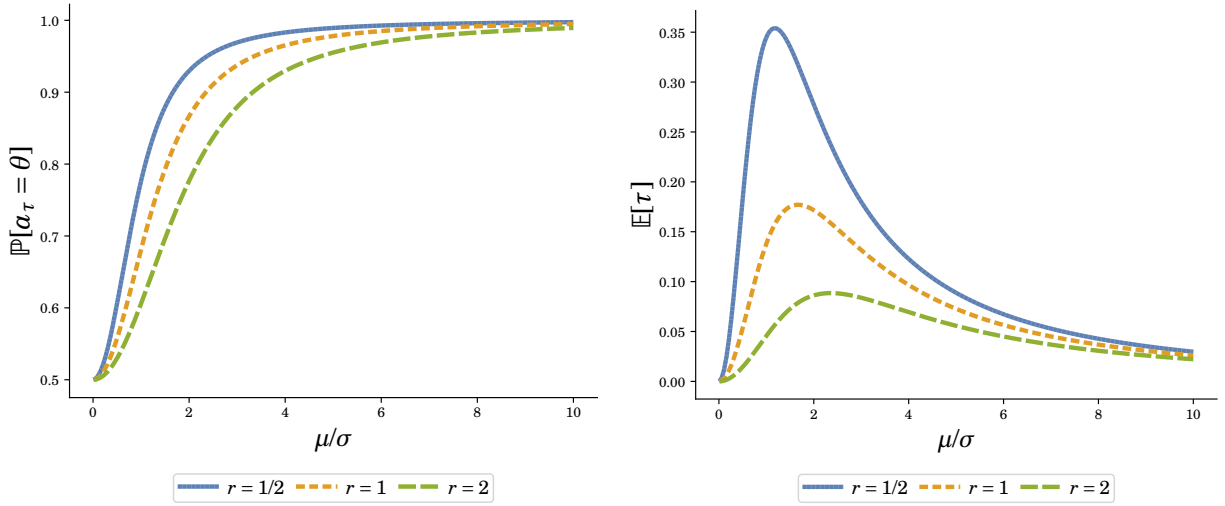
The earliest optimal stopping time is then given by $\tau^* := \inf\{t \geq 0 \mid V(p_t) \leq v(p_t)\}$. Let $D := \{p \in [0, 1] : V(p) > v(p)\}$. Since v, V are convex in p and $V(p) = v(p)$ for $p \in \{0, 1\}$, and v is piecewise-linear, D must be connected. Letting $\underline{p} := \inf D$ and $\bar{p} := \sup D$, we have that $V(p) = v(p)$ for $p \leq \underline{p}$ and $V(p) = v(p)$ for $p \geq \bar{p}$.

Furthermore, since payoffs are symmetric (i.e. $u(a, a) - u(b, a) = u(b, b) - u(a, b)$ and $u(a, b) = u(b, a)$), the stopping thresholds will also be symmetric around $1/2$, i.e. $\underline{p} = 1 - \bar{p}$. Then,

$$\mathbb{E}[\exp(-r\tau^*)(1_{\alpha_{\tau^*}=\theta})] = \mathbb{E}[\exp(-r\tau^*)(1_{\alpha_{\tau^*}=\theta}) \mid \theta = a] = \frac{1}{2} \exp\left(\frac{\bar{\ell}}{2}\right) \operatorname{sech}\left(\frac{\bar{\ell}}{2} \kappa\right)$$

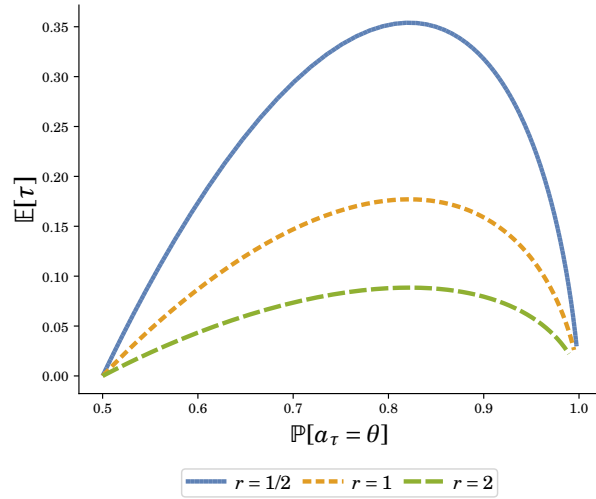
where the last equality follows from [Borodin and Salminen \(2002, p. 309, 3.0.5\(b\)\)](#) and $\kappa := (2r(\sigma/\mu)^2 + 1)^{1/2}$.

Taking first-order conditions, we get that the optimal stopping boundary is given by $\bar{\ell}^* = 2 \tanh^{-1}(1/\kappa)/\kappa$. It is then immediate that accuracy, $\mathbb{P}(\alpha_{\tau^*} = \theta) = (1 + \exp(-\bar{\ell}^*))^{-1}$, is increasing in μ/σ . The expected stopping time fixing $\bar{\ell} = -\underline{\ell}$ is known and given by (3), that is, $\mathbb{E}[\tau] = \frac{\bar{\ell} \tanh(\bar{\ell}/2)}{2(\mu/\sigma)^2}$. Substituting for the optimal stopping threshold $\bar{\ell}^*$, we have $\mathbb{E}[\tau^*] = \frac{\kappa^2 - 1}{\kappa r} \coth^{-1}(\kappa) \tanh(\coth^{-1}(\kappa)/\kappa)$, which is inverse U -shaped in μ/σ , with $\lim_{\mu/\sigma \downarrow 0} \mathbb{E}[\tau^*] = \lim_{\mu/\sigma \uparrow \infty} \mathbb{E}[\tau^*] = 0$ – see [Figure 5](#).



(a)

(b)



(c)

Figure 5: Speed and Accuracy: Discounting