Working Paper Series 739
(ISSN 2788-0443)

Sequential Sampling Beyond Decisions? A Normative Model of Decision Confidence

Rastislav Rehák

CERGE-EI
Prague, November 2022

ISBN 978-80-7343-546-2 (Univerzita Karlova, Centrum pro ekonomický výzkum a doktorské studium)
ISBN 978-80-7344-657-4 (Národohospodářský ústav AV ČR, v. v. i.)

Sequential Sampling Beyond Decisions? A Normative Model of Decision Confidence*

Rastislav Rehák †

CERGE-EI‡

November 2022

Abstract

We study informational dissociations between decisions and decision confidence. We explore the consequences of a dual-system model: the decision system and confidence system have distinct goals, but share access to a source of noisy and costly information about a decision-relevant variable. The decision system aims to maximize utility while the confidence system monitors the decision system and aims to provide good feedback about the correctness of the decision. In line with existing experimental evidence showing the importance of post-decisional information in confidence formation, we allow the confidence system to accumulate information after the decision. We aim to base the post-decisional stage (used in descriptive models of confidence) in the optimal learning theory. However, we find that it is not always optimal to engage in the second stage, even for a given individual in a given decision environment. In particular, there is scope for post-decisional information acquisition only for relatively fast decisions. Hence, a strict distinction between one-stage and two-stage theories of decision confidence may be misleading because both may manifest themselves under one underlying mechanism in a non-trivial manner.

Keywords: decision, confidence, sequential sampling, optimal stopping

JEL codes: C11, C41, C44, D11, D83, D91

---

*I am very grateful to Filip Matějka for guidance and support, to Alex Bloedel, Tom Griffiths, Pavel Kocourek, Xiaosheng Mu, Pietro Ortoleva, Maxim Senkov, Milan Ščasný, João Thereze, Can Urgun, Ansgar Walther, and Weijie Zhong for very useful discussions, and to numerous other discussants for insights and comments. This project was supported by Charles University GAUK project No. 666420 and by the H2020-MSCA-RISE project GEMCLIME-2020 GA No. 681228. This paper is part of a project that has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 870245. This project has received funding from the European Research Council under the European Union's Horizon 2020 research and innovation programme (grant agreements No. 101002898 and No. 770652). Parts of this paper were developed and written during my research stays at Harvard University and Princeton University.

†E-mail address: Rastislav.Rehak@cerge-ei.cz

‡CERGE-EI, a joint workplace of Center for Economic Research and Graduate Education, Charles University and the Economics Institute of the Czech Academy of Sciences, Politických vězňů 7, P.O. Box 882, 111 21 Prague 1, Czech Republic.

1 Introduction

Decision confidence is a subjective assessment of one's own decision quality—a belief that a decision is correct. It is a manifestation of one's metacognitive abilities, which has been an area of great interest in cognitive science. However, it has received much less attention in economics even though it is relevant in many situations of economic interest.[1] Moreover, decision confidence provides cheap additional information about people's preferences and judgements, which relates to the proposal of Krajbich et al. (2014) to use previously neglected measures.

We jointly model decisions, decision times, confidence judgements, and interjudgement times.[2] Such data are commonly gathered in the two-alternative choice-followed-by-confidence experimental paradigm.[3] For example, a participant may be asked to choose which cloud on the screen (left or right) contains more dots and, after making the decision, to indicate on a slider between 0 and 1 her confidence in the decision.

In cognitive science, several descriptive[4] models have been developed to capture this kind of data. We relate directly to two prominent models in this literature: the Two-Stage Dynamic Signal Detection (2DSD) model of Pleskac and Busemeyer (2010) and the Collapsing Confidence Boundary (CCB) model of Moran et al. (2015). We adopt the idea of (potential) post-decisional evidence accumulation for confidence judgement formation (Yeung and Summerfield, 2012; Fleming and Daw, 2017). The primary role of these models is to provide a flexible framework

[1] We list some areas and literature that provide economic motivation for interest in decision confidence in Appendix F.

[2] *Confidence judgements* can be recorded on a slider between 0 and 1, with 1 corresponding to being certain about the decision and 0 to recognizing an error with certainty. *Decision times* (also called 'reaction times' or RT) measure the time between the start of the visual display of stimulus and the recording of the decision. *Interjudgement times* (also denoted RT2) measure the time between the recording of the decision and the recording of the confidence judgement.

[3] The best available source of data is the Confidence Database (Rahnev et al., 2020).

[4] We use the nomenclature of *normative* and *descriptive* models in the sense of Baron (2012). In particular, descriptive models "try to explain how people make judgments and decisions" (Baron, 2012, p. 1); they are data-driven. On the other hand, normative models serve as "standards for evaluation" and "[t]hey must be justified independently of observations of people's judgments and decisions, once we have observed enough to define what we are talking about." (Baron, 2012, p. 1). Hence, normative models capture our understanding of a situation, formulate a problem associated with that situation, and assume an optimal solution to the problem to derive behavior.

that is able to capture a multitude of empirical patterns and thus enable us to summarize and think about those patterns in a succinct way. However, these models are heuristic in the sense that they are not derived from first principles and the objects characterizing the behavior are not tied to the primitives of the decision environment. To be able to conduct counterfactual analyses and theorize about the drivers of decision confidence, normative models are needed.

At the normative spectrum of the literature, we follow closely the model of Fudenberg et al. (2018). However, this and similar normative models (Wald, 1947; Tajima et al., 2016) were developed to model primarily decisions, but not decision confidence. On one hand, they feature belief confidence,[5] so one can define decision confidence in these models as the belief confidence in the chosen option upon stopping. On the other hand, such decision confidence is incompatible with the observed ability of people to recognize their own errors even without explicit feedback (Yeung and Summerfield, 2012; Fleming and Daw, 2017) and the importance of post-decisional evidence in confidence formation (Moran et al., 2015). Hence, an extension of these models is needed to account for the whole mechanism behind decision confidence.

In this paper, we develop a normative dynamic model of decision confidence. At Marr's (1982) computational level, we posit that a decision maker employs two systems with distinct goals—a *decision system* and a *confidence system*. The decision system aims to choose the best option, i.e., it maximizes expected utility. The confidence system aims to monitor the decision system and provide good feedback on its performance. We assume that confidence acts as a substitute for explicit feedback (thus, it plays a key role in situations when explicit feedback is not (immediately) available); the confidence system minimizes the mean-squared error (MSE) of confidence relative to the perfect feedback indicator of (in)correctness of the decision. An implication of this assumption is that decision confidence is the posterior probability of being correct, in accordance with the Bayesian confidence hypothesis (Pouget et al., 2016). Finally, the two separate systems have access to a common source of costly and noisy evidence about the values of the options and resolve the speed-performance trade-off optimally. Naturally, we assume that the confidence system can continue evidence accumulation beyond decisions.

[5]To clarify the terminology, *belief confidence* (in an option) is accessible at any point of deliberation. However, in this paper, we are modeling *decision confidence*, which we define as a subjective assessment of one's own decision quality. In particular, this definition requires implicitly that (i) a decision is made (i.e., there is no decision confidence without a decision) and (ii) decision confidence is a committed judgement (i.e., the endpoint of deliberation, not a running variable). This distinction mirrors the distinction between "confidence" and "certainty" advocated by Pouget et al. (2016).

We are interested especially in dissociations between decision performance and metacognition. Specifically, we ask when decision and decision confidence should be based on the same evidence. Stated differently, we allow for two-stage confidence formation and we ask when it is optimal to use only one stage. This amounts to the comparison of the optimal stopping regions for evidence accumulation of the decision stopping problem and the unconstrained confidence stopping problem.[6]

Our main analytical result is a closed-form tight bound on the unconstrained confidence stopping time of the evidence process.[7] Together with the unboundedness of the decision stopping time, this implies that (relatively)[8] slow decisions will lead to so called *one-stage confidence*, which is a situation when the confidence is based on the same evidence as the decision. Therefore, there is scope for *two-stage confidence*, which is a situation when the confidence is based on more evidence than the decision, only for fast decisions. Consequently, there is space for error monitoring only for the fast decisions. However, we demonstrate numerically that under some parameters, not all fast decisions must lead to two-stage confidence. Surprisingly, it may happen that the very fastest decisions (together with the slow decisions) lead to one-stage confidence, while only intermediately fast decisions lead to two-stage confidence. Finally, an intuitive result is that it is only under low cost of time and/or strong preference for good confidence that there is room for two-stage confidence at all.

We contribute to two main strands of literature. First, we build on the analysis of Chernoff's (1961) problem of sequential testing of the sign of the normal drift of a Brownian motion (Zhitlukhin and Muravlev, 2013; Fudenberg et al., 2018). However, our confidence stopping problem is, to the best of our knowledge, a novel problem. Second, we contribute to the first long-term goal for the field of metacognition formulated by Rahnev et al. (2021, p. 6)—development of detailed models of visual metacognition. To the best of our knowledge, we are the first to propose a *dynamic normative model* of decision confidence.

[6]We express both stopping algorithms in the space of the evidence process (as opposed to the space of posterior expected values, for example).

[7]The original confidence stopping problem is constrained by the decision stopping time.

[8]In short, "relatively" means relative to a decision environment. Intuitively, choosing a house in one minute is fast, but choosing an apple for a snack in one minute is slow. We discuss in detail the meaning of "relatively" (fast/slow decisions) in Section 3.3.

Our main contribution to the discussion about one-stage vs. two-stage theories of decision confidence is that a strict distinction between them (Moran et al., 2015) might be misleading because both may manifest themselves under one underlying mechanism, even within one individual in a given controlled decision environment. Moreover, our approach allows us to predict how an individual might change the modes of decision confidence formation under different circumstances. Hence, our model speaks to both intra- and inter-individual differences in the formation of decision confidence.

In a complementary work, Fleming and Daw (2017) propose a Bayesian framework for grounding a discussion about a related aspect of metacognitive computation—whether decision and confidence are informed by the same signal or different but correlated signals. In their framework, our model falls into the category of "postdecisional" models (Pleskac and Busemeyer, 2010; Moran et al., 2015), in which a single process informs both decision and decision confidence. We are thus leaving aside their proposed "second-order" architecture that allows decision and decision confidence to be informed by distinct but correlated processes. However, our model has a feature reminiscent of the "second-order" architecture: the goals of the decision and confidence systems are distinct. Neverthless, Fleming and Daw (2017) do not postulate an explicit goal for the confidence system (for the decision system, it is implicitly expected utility maximization). Moreover, they work in a static environment and are concerned with a high-level structure of confidence computation, while we focus on *procedural* details of the *optimal* evidence accumulation.[9]

The question we ask—when is it optimal to gather additional evidence after the decision?—is similar to the question studied by the literature about metacognitive control (Schulz et al., 2021; Boldt et al., 2019; Desender et al., 2018). In a typical experimental paradigm in this literature, participants are asked to decide whether to obtain additional information after the first decision/stimulus presentation, but they are given an explicit motivation for doing so, e.g., a revision of the initial decision or a subsequent related decision. Moreover, the stimulus presentation is often not under full control of the participant. In contrast, we aim to contribute primarily to the literature about metacognitive monitoring, i.e., we are interested in how decision confidence arises rather than how it is used to control subsequent behavior. In particular, sampling beyond decisions in our setup leads to formation of decision confidence, while in the

[9] Fleming and Daw (2017) recognize the importance of modeling the dynamics in their footnote 1, p. 94.

control literature, the roles are reversed—decision confidence is used to decide about additional sampling for a specified goal.

2 Model

Our model consists of two separate systems—decision system and confidence system. The systems have access to a common evidence process and sampling is costly. Since the systems have different goals, they have to solve different optimal stopping problems.

2.1 Decision system

The model of the decision system follows closely the model of Fudenberg et al. (2018).[10] In the decision stage, the agent is choosing between options l and r, which can bring her utilities $\theta^l \in \mathbb{R}$ and $\theta^r \in \mathbb{R}$, respectively. The agent does not know the true utilities but has a prior about (θ^l, θ^r)

$$\mu_0 \sim N\left(\begin{pmatrix} X_0^l \\ X_0^r \end{pmatrix}, \begin{pmatrix} \sigma_0^2 & 0 \\ 0 & \sigma_0^2 \end{pmatrix}\right). \tag{1}$$

Moreover, she can learn about the utilities by observing continuously the signal $\{(Z_t^l, Z_t^r)\}_{t\geq 0}$ such that for $i \in \{l, r\}$

$$Z_t^i = \theta^i t + \alpha B_t^i, \; t \geq 0, \tag{2}$$

where $\{B_t^i\}_{t\geq 0}$ is a standard Brownian motion (the l and r versions are independent of each other and of (θ^l, θ^r)) and parameter $\alpha > 0$ captures the noisiness of the signal components. However, the agent pays constant flow cost $c > 0$ for accumulating evidence.

We denote by $\mathcal{F}_t$ the information up to time t (formally, on our probability space $(\Omega, \mathcal{F}, \mathrm{P})$, we have a filtration $\{\mathcal{F}_t\}_{t\geq 0}$ generated by process $\{(Z_t^l, Z_t^r)\}_{t\geq 0})$ and beliefs about (θ^l, θ^r) at time t by μ_t. We denote the posterior means for $i \in \{l, r\}$ by $X_t^i = \mathrm{E}\left[\theta^i | \mathcal{F}_t\right]$.

The agent decides when to stop the accumulation of evidence and what option to choose upon stopping. Conditional on stopping at time t, she chooses an option to maximize her expected

[10] We also try to adopt the notation of Fudenberg et al. (2018) to the highest reasonable degree so that our new results about decision confidence can be better understood within their framework.

utility, so her (optimal) decision rule $d : \mathbb{R}^2 \to \{l, r\}$ is

$$d(X_t^l, X_t^r) = \begin{cases} l & \text{if } X_t^l \geq X_t^r, \\ r & \text{if } X_t^l < X_t^r \end{cases} \tag{3}$$

and she expects to get $\max\{X_t^l, X_t^r\}$. Hence, she faces a stopping problem

$$\sup_{\tau \in \mathcal{T}} \mathrm{E}\left[\max\{X_\tau^l, X_\tau^r\} - c\tau\right], \tag{4}$$

where τ is an $\mathcal{F}_t$-stopping time (formally, $\tau \colon \Omega \to [0, \infty]$ such that $\{\tau \leq t\} \in \mathcal{F}_t \ \forall t \geq 0$ and $\mathrm{P}(\tau < \infty) = 1$) and $\mathcal{T}$ is the set of all $\mathcal{F}_t$-stopping times.[11]

2.2 Confidence system

The agent can continue to accumulate signal $\{(Z_t^l, Z_t^r)\}_{t \geq 0}$ even after the end of the the decision stage τ in order to refine her beliefs for forming decision confidence. By accumulating the evidence in the confidence stage, the agent pays constant flow cost $\bar{c} > 0$.

In accordance with the Bayesian confidence hypothesis (Pouget et al., 2016), we define *decision confidence* as the probability of having made the optimal decision.[12] Hence, decision confidence at time $t \geq \tau$ is[13]

$$conf_t = \mathrm{P}\left(d(X_\tau^l, X_\tau^r) = \arg\max_{i \in \{l,r\}} \theta^i \,\middle|\, \mathcal{F}_t\right). \tag{5}$$

The agent decides when to stop the confidence stage in order to minimize the MSE loss from incorrectly assessing the (objective) (in)correctness of the decision and the additional cost of evidence accumulation in the confidence stage

$$\inf_{\tau_c \in \mathcal{T} \text{ s.t. } \tau_c \geq \tau} \mathrm{E}\left[\left(conf_{\tau_c} - \mathbb{1}\{d(X_\tau^l, X_\tau^r) = \arg\max_{i \in \{l,r\}} \theta^i\}\right)^2 + \bar{c}(\tau_c - \tau)\right]. \tag{6}$$

Note that flow cost $\bar{c}$ is likely to be different from c because it is expressed in different units: while c captures forgone utils during decision making, $\bar{c}$ captures forgone utils during confidence formation relative to lost utils due to imperfect confidence (confidence with higher MSE

[11] In cases of multiple optimal stopping times, we select the minimal optimal stopping time. This is assumed also for the confidence stopping problem.

[12] This definition of decision confidence can also be justified as the optimal assessment under the mean-squared error (MSE) loss from incorrectly assessing the (objective) (in)correctness of the decision.

[13] We implicitly exclude the event $\theta^l = \theta^r$, which is of zero probability (for any $\mathcal{F}_t$), so the arg max in the formula is a singleton.

loss). Formally, the most natural interpretation of $\bar{c}$ is that it is cost of time c relative to the importance of well-calibrated confidence denoted by γ, $\bar{c} = \frac{c}{\gamma}$. However, c and $\bar{c}$ can be manipulated even independently in experiments, e.g., a time-pressure-on-choice manipulation can be understood as an increase in c without a change in $\bar{c}$.

3 Analysis

Our research question leads to the comparison of the solutions of the optimal stopping problems of the decision and confidence system. Since the stopping problem of the decision system has been studied elsewhere (in particular, by Fudenberg et al. [2018]), we focus on the confidence stopping problem. The notorious difficulty of optimal stopping problems prevents us from going far analytically, so we first provide insights based on numerical solutions before presenting more general points.

3.1 Reformulation of the confidence problem

Since the decision of the agent depends only on the sign of the difference between the options, the object of interest is $\theta := \theta^l - \theta^r$.[14] The prior about θ is

$$N(X_0 := X_0^l - X_0^r, 2\sigma_0^2). \tag{7}$$

Hence, a sufficient statistic for the agent is the process

$$Z_t := Z_t^l - Z_t^r = \theta t + \alpha\sqrt{2}B_t,\ t \geq 0, \tag{8}$$

where $B_t = 1/\sqrt{2}(B_t^l - B_t^r)$ is a standard Brownian motion independent of θ. By Lemma A.1, the beliefs about θ at time t are normal

$$N(X_t, \sigma_t^2) \tag{9}$$

with mean

$$X_t = \frac{\sigma_0^{-2} X_0 + \alpha^{-2} Z_t}{\sigma_0^{-2} + \alpha^{-2} t} \tag{10}$$

and variance

$$\sigma_t^2 = \frac{2}{\sigma_0^{-2} + \alpha^{-2} t}. \tag{11}$$

[14]Note that Fudenberg et al. (2018) denote by θ something else—the pair (θ^l, θ^r).

Hence, we can express the confidence explicitly

$$conf_t = \Phi\left(\frac{X_t}{\sigma_t}\right)\mathbb{1}\{X_\tau \geq 0\} + \Phi\left(-\frac{X_t}{\sigma_t}\right)\mathbb{1}\{X_\tau < 0\}, \tag{12}$$

where Φ is the CDF of the standard normal distribution.

Since

$$\begin{aligned}
&\mathrm{E}\left[\left(conf_{\tau_c} - \mathbb{1}\{d(X_\tau^l, X_\tau^r) = \arg\max_{i\in\{l,r\}} \theta^i\}\right)^2\right] && (13)\\
&= \mathrm{E}\left[\mathrm{E}\left[\left(conf_{\tau_c} - \mathbb{1}\{d(X_\tau^l, X_\tau^r) = \arg\max_{i\in\{l,r\}} \theta^i\}\right)^2 \middle| \mathcal{F}_{\tau_c}\right]\right] && (14)\\
&= \mathrm{E}\left[\mathrm{var}\left(\mathbb{1}\{d(X_\tau^l, X_\tau^r) = \arg\max_{i\in\{l,r\}} \theta^i\} \middle| \mathcal{F}_{\tau_c}\right)\right] && (15)\\
&= \mathrm{E}\left[conf_{\tau_c}(1 - conf_{\tau_c})\right], && (16)
\end{aligned}$$

we can use (12) to rewrite the confidence-stage objective function as

$$\mathrm{E}\left[\Phi\left(\frac{X_{\tau_c}}{\sigma_{\tau_c}}\right)\Phi\left(-\frac{X_{\tau_c}}{\sigma_{\tau_c}}\right) + \bar{c}\tau_c\right] - \bar{c}\mathrm{E}\left[\tau\right]. \tag{17}$$

In the reformulation of the confidence objective (17), we can focus only on the first part because $\bar{c}\mathrm{E}[\tau]$ is irrelevant for the choice of τ_c. Moreover, the first part in (17) does not feature any τ elements, so we can simplify the analysis by studying the unconstrained stopping problem

$$\inf_{\tau'\in\mathcal{T}} \mathrm{E}\left[\Phi\left(\frac{\sigma_0^{-2}X_0 + \alpha^{-2}Z_{\tau'}}{\sqrt{2(\sigma_0^{-2} + \alpha^{-2}\tau')}}\right)\Phi\left(-\frac{\sigma_0^{-2}X_0 + \alpha^{-2}Z_{\tau'}}{\sqrt{2(\sigma_0^{-2} + \alpha^{-2}\tau')}}\right) + \bar{c}\tau'\right]. \tag{18}$$

Due to the Markovian structure of problem (18) (see Appendix A.2), the solution of this problem boils down to finding its continuation region in the (t, z)-space $(= [0, \infty) \times \mathbb{R})$. Specifically, we are looking for a set $C_C \subseteq [0, \infty) \times \mathbb{R}$ such that the complement of C_C is closed and the stopping time $\tau'^* = \inf\{t \geq 0 \colon (t, Z_t) \notin C_C\}$ is the (minimal) optimal stopping time in (18).

To simplify further notation, let us introduce the loss function $f\colon [0, \infty) \times \mathbb{R} \to \mathbb{R}$ defined for all $(r, x) \in [0, \infty) \times \mathbb{R}$ as follows

$$f(r, x) = \Phi\left(\frac{\sigma_0^{-2}X_0 + \alpha^{-2}x}{\sqrt{2(\sigma_0^{-2} + \alpha^{-2}r)}}\right)\Phi\left(-\frac{\sigma_0^{-2}X_0 + \alpha^{-2}x}{\sqrt{2(\sigma_0^{-2} + \alpha^{-2}r)}}\right) + \bar{c}r. \tag{19}$$

Moreover, we will denote the expected loss at time $t \in [0, \infty)$ when we start sampling with $s \in [0, \infty)$ amount of evidence and the evidence process at position $z \in \mathbb{R}$

$$u(t; s, z) = \mathrm{E}\left[f(s + t, Z_t^z)\right], \tag{20}$$

where $Z_t^z = z + \theta t + \alpha\sqrt{2}\widetilde{B}_t$, $t \geq 0$, with a new standard Brownian motion $\{\widetilde{B}_t\}_{t\geq 0}$ (see Appendix A.2). Finally, we will denote the value function

$$U(s, z) = \inf_{\tau \in \mathcal{T}} u(\tau; s, z). \tag{21}$$

With this notation, the confidence stopping problem can be formulated as the problem of characterizing the set

$$C_C = \{(s, z) \in [0, \infty) \times \mathbb{R} : U(s, z) < f(s, z)\}. \tag{22}$$

If we denote by C_D the continuation region of the decision stopping problem (4) in the (t, z)-space,[15] the continuation region of the original confidence stopping problem (6) will be $C_D \cup C_C$. We are interested especially in the analysis of the regions $C_C \backslash C_D$ and $C_D \backslash C_C$, which characterize when it is optimal to continue accumulating evidence beyond the decision stage and when it is optimal to stop immediately after the decision stage, respectively.

3.2 Qualitative insights from numerical solutions

In Figure 1, we depict the numerically computed decision and confidence stopping boundaries (∂C_D and ∂C_C, respectively) for an actual individual with estimated parameters $c = 0.02, \alpha = 2, \sigma_0 = 1.8, X_0 = 0$[16] and a hypothetical value of $\bar{c} = 0.007$. As is illustrated by this figure, the decision stopping boundary is a pair of barriers collapsing to zero at infinity and the confidence stopping boundary is a left-truncated ellipse.[17] These are the typical features of these regions (for another example of the boundaries with different parameters, see Figure 3 in Appendix G). One less typical feature of the boundaries depicted in Figure 1 is that they first expand before collapsing.[18] We chose this non-typical (but realistic) set of parameters to illustrate an important point developed later in this section.

The value of $\bar{c}$ used in Figure 1 is sufficiently low so that it is sometimes optimal to sample beyond decisions. In particular, if a realization of the evidence process Z_t is sufficiently strong,

[15] Fudenberg et al. (2018) characterize C_D in their Theorem 4 and footnote 22.

[16] See Subject 45 in Table 4 in the Online Appendix of Fudenberg et al. (2018). The value of $X_0 = 0$ was imposed in their work.

[17] For the characterization of the decision stopping boundary, see Fudenberg et al. (2018). The shape of the confidence stopping boundary is driven by region $\mathring{C}_C$ derived analytically in Appendix A.3.

[18] This can be seen from Figure 3 in the Online Appendix of Fudenberg et al. (2018). The estimated decision boundaries for most participants are monotonically collapsing.

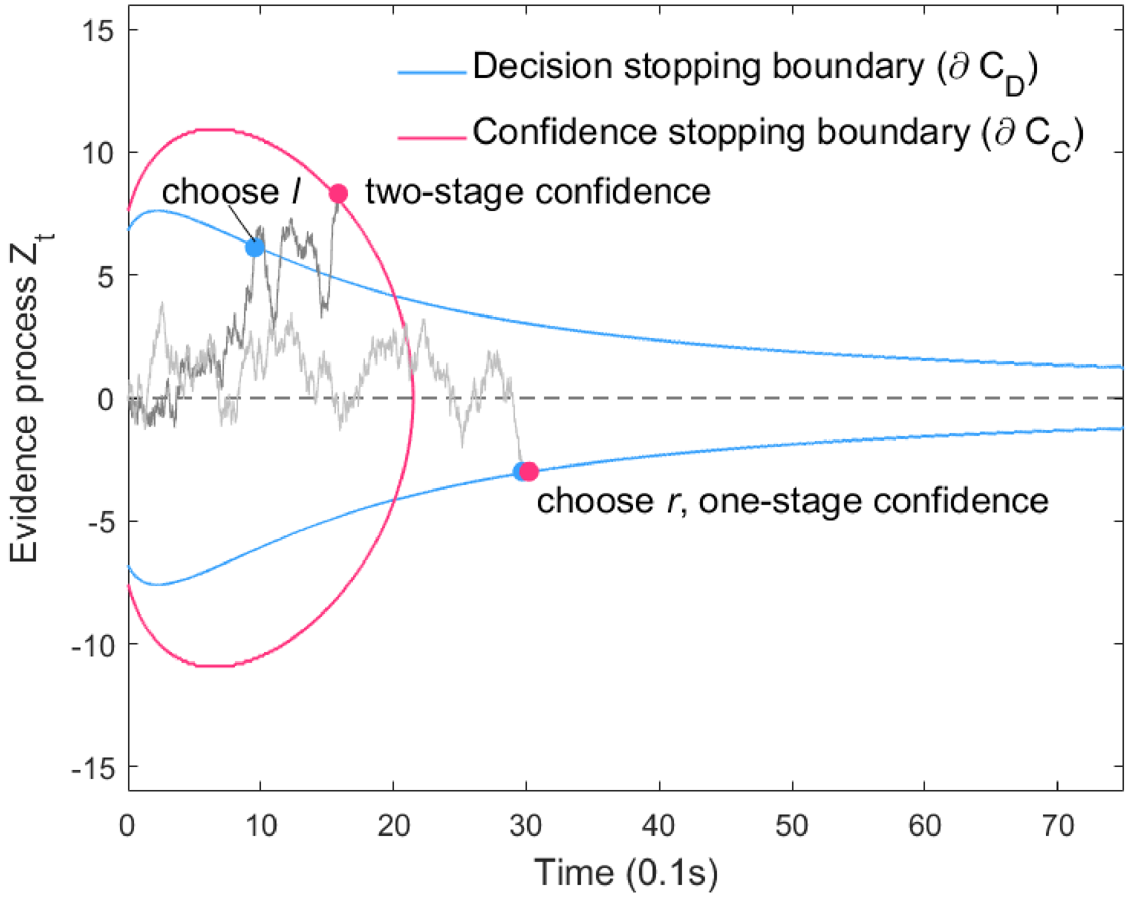

Figure 1: An illustration of optimal stopping of the evidence process for decision and confidence for a **low value of** $\bar{\mathbf{c}}$. The decision and confidence stopping boundaries are computed numerically (see Appendix D) for parameters $\bar{c} = 0.007, c = 0.02, \alpha = 2, \sigma_0 = 1.8, X_0 = 0$. The grey lines represent two possible realizations of the evidence process. The blue dots represent the moments of decision formation and the pink dots represent the moments of confidence formation.

e.g., as depicted by the dark grey line, then the decision is made inside the confidence continuation region C_C and it is optimal to sample beyond such decision to refine confidence. Hence, the confidence judgement will be based on more information than the decision. We call such cases of informational dissociation between decision and confidence *two-stage confidence*.

On the other hand, if a realization of the evidence process Z_t is sufficiently weak, e.g., as depicted by the light grey line, then the decision is made outside the confidence continuation region C_C and it is optimal to stop as soon as the decision is made. In fact, from the perspective of confidence, it would be optimal to stop even sooner in this case—upon hitting ∂C_C; however, the confidence stopping time is chosen only from a restricted set of stopping times that come after the optimal decision stopping time. Hence, the confidence judgement will be based on the same information as the decision. We call such cases of informational congruence between decision and confidence *one-stage confidence*.

As one might expect, the higher is $\bar{c}$, the smaller is the confidence continuation region C_C, thus the less likely confidence will be two-stage. Moreover, from the shape of the decision and confidence stopping boundaries, one might expect that the two-stage confidence would prevail only for the fastest decisions. However, as suggested by the numerical solution in the next paragraph, this intuition might not hold. This might be an important point speaking to the architecture underlying confidence formation that would be missed by heuristic models of decision confidence (Moran et al., 2015; Pleskac and Busemeyer, 2010).

In Figure 2, we depict the numerically computed decision and confidence stopping boundaries for an individual with $\bar{c} = 0.012$ and the same remaining parameters as in Figure 1, $c = 0.02, \alpha = 2, \sigma_0 = 1.8, X_0 = 0$. The value of $\bar{c}$ is sufficiently low so that there is space for two-stage confidence. However, for the fastest decisions, C_C is contained in C_D, so confidence will be one-stage for these decisions—for example, see the strong realization of the evidence process depicted in dark grey. On the other hand, the weak realization of the evidence process depicted in light grey will lead to a relatively fast (but not the fastest) decision and two-stage confidence.[19] Finally, the weakest realizations of the evidence process (as the yellow one) will again lead to one-stage confidence. This illustrates a potential "non-monotonicity" of occurrence of two-stage confidence.

In summary, the numerical solutions suggest several qualitative insights about the informational underpinning of confidence judgements relative to decisions:

- The slowest decisions are associated with one-stage confidence, i.e., decisions and decision confidence are based on the same evidence.
- If the cost of time relative to the importance of good confidence $\bar{c}$ is sufficiently low, (intermediately) fast decisions are associated with two-stage confidence, i.e., decision confidence is based on more evidence than decisions.

[19] Notice that even though the decision and confidence boundaries ∂C_D and ∂C_C are not very distant in the region for intermediately fast decisions, the resulting timing of decision and confidence formation may be very distinct as illustrated by the light grey realization of the evidence process.

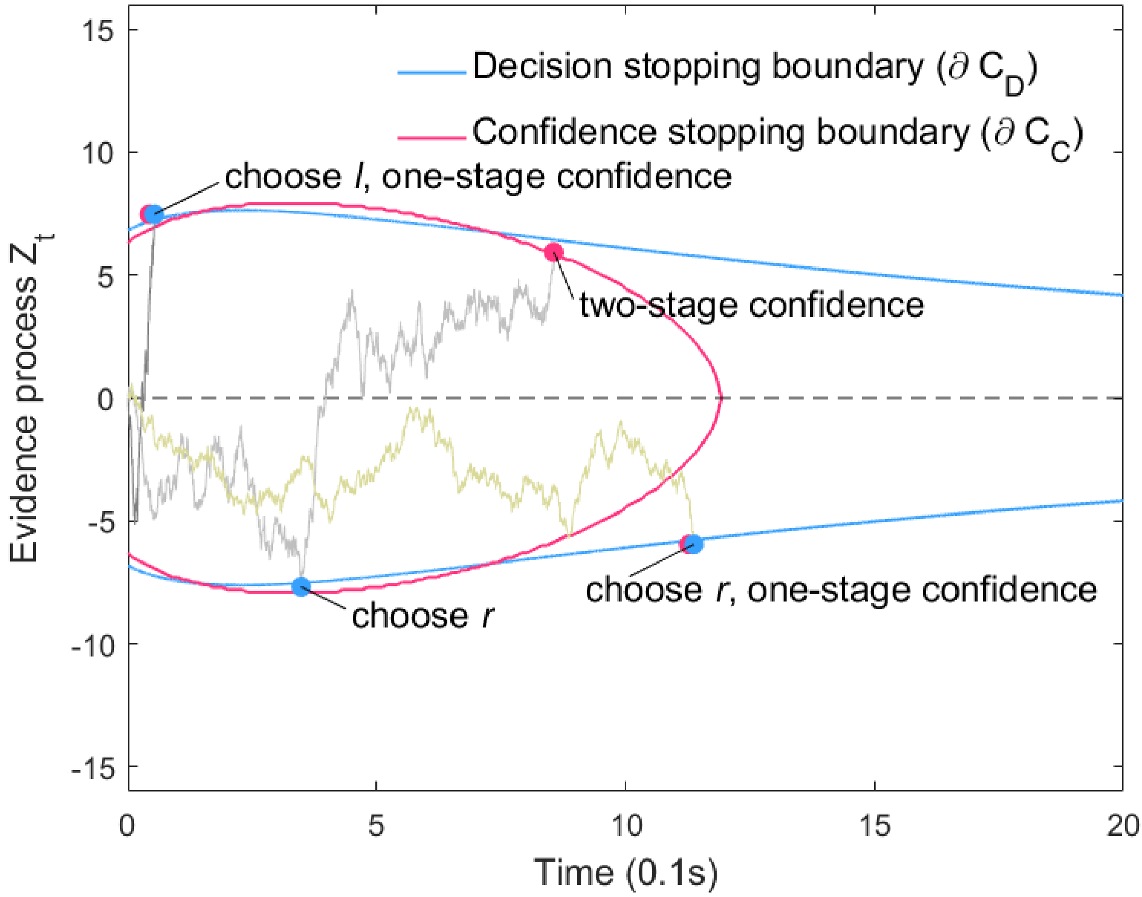

Figure 2: An illustration of optimal stopping of the evidence process for decision and confidence for a **moderately high value of** $\bar{\mathbf{c}}$. The decision and confidence stopping boundaries are computed numerically (see Appendix D) for parameters $\bar{c} = 0.012, c = 0.02, \alpha = 2, \sigma_0 = 1.8, X_0 = 0$. The grey and yellow lines represent three possible realizations of the evidence process. The blue dots represent the moments of decision formation and the pink dots represent the moments of confidence formation.

- If the cost of time relative to the importance of good confidence $\bar{c}$ is sufficiently low, the fastest decisions might be associated with two-stage confidence.

The meaning of "fast" and "slow" decisions in these statements is relative to the decision environment. It is supposed to capture a qualitative distinction between decisions in terms of timing for a given decision environment, but it does not mean there are some fixed time thresholds that could help to classify decisions as fast or slow universally. In the next section, we provide a threshold as a function of the parameters of a decision environment that brings more meaning to this distinction between fast and slow decisions and corroborates the qualitative insights just described.

3.3 Bound for two-stage confidence

Our main analytical result is the explicit tight bound on the optimal stopping time in the unconstrained confidence stopping problem (18).

Proposition 1. *Any optimal stopping time in the unconstrained confidence stopping problem* (18) *is bounded almost surely by* $\max\{\alpha^2[(2\pi\alpha^2\bar{c})^{-1} - \sigma_0^{-2}], 0\}$.

Proof. The proof is in Appendix A.4. □

To gain intuition for the overall shape and boundedness of the confidence stopping region, assume $X_0 = 0$ for simplicity and fix a time $t > \tau$. Conditional on t, the MSE loss, which we want to minimize, can be rewritten as in (13). In particular, it can be noticed that as a function of the evidence particle Z_t (denoted $MSE_t(Z_t)$), it is bell-shaped and smooth around zero. Hence, when deciding at time t whether to continue to $t + \mathrm{d}t$, we can see from the bell shape of $MSE_t(Z_t)$ that some scope for a decrease in the loss can be expected only around zero due to concavity, i.e., we can expect the evidence process to diffuse almost equally likely to parts with much lower loss than to parts with only slightly higher loss.

The boundedness stems from the smoothness of $MSE_t(Z_t)$ around zero. In particular, as time t increases, $MSE_t(Z_t)$ becomes flatter around zero,[20] so the expected drop in loss near zero becomes smaller as time passes. However, there is no kink at zero as in the case of the decision problem that would prevent the expected drop in loss to be dominated by the cost of sampling time at a finite point of time.

The bound in Proposition 1 is the threshold distinguishing "slow" and "fast" decisions alluded to in the previous section. For decisions made after this point, the decision confidence will certainly be one-stage. For decisions made slightly before this point, the decision confidence will still be one-stage, which comes from the fact that the decision stopping boundary is bounded away from zero at any specific time while the confidence boundary collapses to zero at a finite time.

[20]As time increases, low values of Z_t become more consistent with the value of the utility difference θ around zero, for which the most cautious confidence judgement is 0.5. Moreover, as time increases, a larger interval of values of Z_t around zero becomes more consistent with the values of θ around zero because larger θ would have likely drifted the process far away by that time.

Next, three distinct cases may occur:

1. The confidence region C_C is entirely contained in the decision region C_D. Confidence will always be one-stage in this case.
2. The confidence region C_C "peaks out" above the decision region C_D in an interval of decision times $[\underline{t}, \bar{t}]$, $0 < \underline{t} < \bar{t} < \infty$, but is contained in C_D near zero as in Figure 2. Confidence will be mostly one-stage (for fast and slow decisions) except for the interval of decision times $[\underline{t}, \bar{t}]$, where it will be two-stage.
3. The confidence region C_C "embraces" the decision region C_D in an interval of decision times $[0, \bar{t}]$, $\bar{t} > 0$ (as in Figure 1), in which confidence will be two-stage, while it will be one-stage beyond this interval.

Unfortunately, a full analytical characterization of cases of one-stage and two-stage decision confidence, which requires a nontrivial analysis of the relative position of objects C_C and C_D, is a difficult problem.[21]

The bound in Proposition 1 varies intuitively with the parameters of the decision environment: it increases with lower cost of time relative to importance of precise confidence $\bar{c}$, lower noisiness of the evidence process α, and higher prior variance σ_0^2. In particular, prior variance is higher for a choice of a house than for a choice of a snack, for example. Hence, "fast" decisions distinguished by this threshold bear naturally different meaning in different contexts. Finally, the explicit expression for the bound also enables us to say that for cost of time $\bar{c}$ larger than $\frac{\sigma_0^2}{2\pi\alpha^2}$ there will be certainly no scope for two-stage confidence.

[21]Besides the numerical solution (see Appendix D), we can also derive analytically the first-order condition for deterministic stopping, see Appendix B, which is necessary but not sufficient. The interest of studying the deterministic stopping lies in the following implication: if immediate stopping is not optimal in the set of deterministic stopping times, then it is not optimal in the set of all stopping times $\mathcal{T}$ either. Nevertheless, based on numerical solutions, the reverse implication does not hold, so the connection between the unconstrained and deterministic optimal stopping is not tight even from this perspective.

4 Discussion

In this section, we evaluate our model (referred to as Two-Stage Sequential Sampling or 2SS) on the basis of empirical patterns of Moran et al. (2015) (which extend the list of patterns of Pleskac and Busemeyer [2010]). Moreover, we contrast the 2SS model with the most related models.

4.1 Empirical evaluation and contrast to heuristic models

Using a small simulation study, we report in Table 1 in Appendix E that 2SS matches all empirical patterns of Moran et al. (2015, Table 1, p. 102) except for the positive correlation of decision and interjudgement times (pattern 8.4). Intuitively, there is scope for non-zero interjudgement times in 2SS only for relatively fast decisions.[22]

The 2SS model can be seen as an attempt to microfound the heuristic model of Moran et al. (2015) (referred to as Collapsing Confidence Boundary or CCB). CCB matches all patterns in Table 1. In particular, it captures the positive correlation of decision and interjudgement times, unlike 2SS. The driving force behind this feature is non-stationarity of the confidence boundary and selection: in CCB, the confidence boundary is decision-triggered, i.e., it is placed only at the moment of decision. Hence, the distance the evidence particle has to travel to reach a given point on the confidence boundary is independent of the decision stopping point (unlike in the 2SS model). Therefore, the selection effect operates: slow decisions are more likely to be generated by a signal with low drift (in absolute value) and this carries over to the confidence stage, leading to slower confidence judgements more likely as well.

The decision-triggered confidence boundary in CCB reflects the assumption that confidence is always two-stage. In contrast, in this paper we do not impose this assumption; instead, we allow for the second stage and derive when it should actually occur. Thus, our paper is an attempt to give a statistical foundation for the two-stage confidence architecture. Our analysis reveals that this uniform assumption may not be well justified and a strict distinction between one-stage and two-stage theories of decision confidence (Moran et al., 2015) might be misleading. Both modes of confidence formation may manifest themselves under one underlying mechanism. Moreover,

[22] 65% and 49% trials lead to two-stage confidence (non-zero interjudgement times) in 2SS under the two parametrizations used in the simulations specified in Table 1.

they may manifest themselves in a surprising manner as in Figure 2, i.e., one-stage confidence for the fastest and slow decisions while two-stage for intermediately fast ones.[23]

It is worth noting that even the empirical results of Moran et al. (2015) "may indicate that there are important individual differences with respect to the number of information collection stages" (p. 111). Our approach enables us to study more rigorously not only these individual differences (e.g., different values of mental noise α [Enke and Graeber, 2022]), but also differences for a given individual across decision environments (e.g., different opportunity cost c), and even differences in the use of one-stage vs. two-stage mode for a given individual in a given decision environment (for fixed parameters).

4.2 Comparison with the model of Fudenberg et al. (2018)

Since the 2SS model is an extension of the uncertain-difference model of Fudenberg et al. (2018) (referred to as FSS), it is natural to discuss predictions of FSS about decision confidence. As outlined in the Introduction, the natural way to define decision confidence in the FSS model is by (12) at the time of decision τ. However, such model is incompatible with the observed ability of people to recognize their own errors even without explicit feedback (Yeung and Summerfield, 2012; Fleming and Daw, 2017). Moreover, Moran et al. (2015) conduct experiments where post-decisional evidence availability has a causal effect on confidence resolution.[24] These experiments put forward the interjudgement time (RT2) as an important variable to guide the modeling of decision confidence—a variable that FSS is silent about.[25]

In Table 1, we report that FSS does not capture the increased confidence resolution under time pressure (higher c). Intuitively, decisions and confidence are tied in FSS, so time pressure affects both negatively. On the other hand, in two-stage models, time pressure makes the job easier for the confidence system because mistakes are easier to recognize.

[23]Moran et al. (2015) build on the work of Pleskac and Busemeyer (2010) who also propose heuristic models of decision confidence capturing the two-stage idea. Importantly, these insights provided by 2SS in relation to CCB apply to the models of Pleskac and Busemeyer (2010) too.

[24]Confidence resolution is the correlation between confidence and decision correctness.

[25]The inability to capture error monitoring and the effect of post-decisional evidence availability is shared by all single-stage models, e.g., the Wald model.

4.3 Simpler version: two-stage Wald model

One may wonder whether our extension to the model of Fudenberg et al. (2018) would bring something interesting in the simpler Wald model. We develop this idea in Appendix C; call it *two-stage Wald model*. We find that the decision maker either (i) never wants to sample beyond decisions (one-stage confidence always) and there is a single confidence level achieved, or (ii) always wants to sample beyond decisions (two-stage confidence always) and there are two possible confidence levels achieved—one above $\frac{1}{2}$ for correctness confirmation and one below $\frac{1}{2}$ for error recognition.[26,27]

People are likely to report more than two levels of confidence, so to account for that we can introduce a noisy readout of confidence, i.e., the confidence levels predicted by the two-stage Wald model would be reported with some independent noise. Nevertheless, there are several empirical regularities that such model cannot explain. For example, there will be no correlation between decision times and confidence because the belief process is a time-homogeneous diffusion. Intuitively, conditional on stopping for decision, the belief process restarts at the same point regardless of the decision time due to the constant decision boundary. Independent readout noise does not change this relationship. This is inconsistent with the empirical regularity no. 4 from Moran et al. (2015) (negative correlation between decision time and confidence). Moreover, the regularities tied to the manipulation of stimulus discriminability will pose problems for the two-stage Wald model too, e.g., pattern no. 3 of negative correlation between confidence and difficulty.[28]

[26]The use of mode (i) or (ii) depends on the parameters, especially $\bar{c}$.

[27]In fact, these predictions would be obtained with an even simpler static model of a rationally inattentive decision maker. However, the static version would be silent about the reaction times.

[28]Nevertheless, one has to be cautious about the use of these empirical regularities because manipulating the objective difference between the two options across trials does not fit the setup of the Wald model. In the Wald model, the decision maker is assumed to know the difference between the options a priori; she is uncertain only about which option is the better one. Hence, the prior of the decision maker is not correct in this setup and there is a mismatch between the experimenter and decision maker.

5 Conclusion

We develop a normative dynamic model of decision confidence. This model captures the dual-system view of mind with the decision system aiming to maximize expected utility and the confidence system aiming to monitor the decision system and provide good feedback on the decision system's performance. We focus on studying optimal informational dissociations between decisions and decision confidence, i.e., we ask when we should expect the decision and confidence to be based on the same evidence and when not. We question the following assumption adopted by models from cognitive science: decision confidence is always (i.e., uniformly across all decisions) based on post-decisional evidence accumulation. We find that this assumption is not justified by our normative model; our model suggests that there is scope for post-decisional evidence accumulation only for relatively fast decisions. Moreover, a nontrivial pattern may emerge in some situations—confidence based on the same information as decisions for very fast and slow decisions and post-decisional evidence accumulation for intermediately fast decisions.

Our findings contribute to the theory of metacognitive monitoring. In particular, we provide a normative foundation for post-decisional evidence accumulation and derive how one- and two-stage modes of decision confidence may arise under a single mechanism. Moreover, an empirical evaluation of our model indicates that it is consistent with the patterns of Moran et al. (2015), except for the positive correlation of decision times and interjudgement times. This and similar patterns may guide future theorizing about the mechanisms underlying decision confidence, which will lead to a better understanding of its functional role. Moreover, it will be interesting to contrast and merge our approach to theorizing about mechanisms behind decision confidence with the complementary direction of "second-order" models (Fleming and Daw, 2017), in which distinct but related processes inform decision and decision confidence. Finally, decision confidence is involved in the control of future mental processes and behavior. Different mechanisms underlying decision confidence may lead to different predictions about the relationships between confidence and future behavior (Schulz et al., 2021). Hence, analyzing the implications of our model of decision confidence for the control of future behavior is an exciting avenue for future research.

Appendix

A Technical details and proofs

A.1 Beliefs

Lemma A.1. *Let* $\theta \sim N(X_0, 2\sigma_0^2)$ *and* $Z_t = \theta t + \alpha\sqrt{2}B_t,\ t \geq 0,$ *where* $\alpha > 0$ *is a parameter and* B_t *is a standard Brownian motion independent of* θ*. Then the posterior distribution of* θ *after observing* $Z_s, s \leq t$*, is normal with mean*

$$\frac{\sigma_0^{-2} X_0 + \alpha^{-2} Z_t}{\sigma_0^{-2} + \alpha^{-2} t} \tag{23}$$

and variance

$$\frac{2}{\sigma_0^{-2} + \alpha^{-2} t}. \tag{24}$$

Proof. As outlined by Chernoff (1961, p. 81) (and developed in more detail by Shiryaev[29]), the conditional distribution of θ is determined by

$$\mathrm{P}\left(\theta \leq y | \mathcal{F}_t\right) = \frac{\int_{-\infty}^{y} \frac{\mathrm{dP}\left(Z_0^t | \theta = \xi\right)}{\mathrm{dP}\left(Z_0^t | \theta = 0\right)} \mathrm{dP}_\theta\left(\xi\right)}{\int_{-\infty}^{\infty} \frac{\mathrm{dP}\left(Z_0^t | \theta = \xi\right)}{\mathrm{dP}\left(Z_0^t | \theta = 0\right)} \mathrm{dP}_\theta\left(\xi\right)}, \tag{25}$$

where

$$\frac{\mathrm{dP}\left(Z_0^t | \theta = \xi\right)}{\mathrm{dP}\left(Z_0^t | \theta = 0\right)} \tag{26}$$

is the Radon-Nikodym derivative of the measure of the process $Z_0^t = \{Z_s\}_{s=0}^t$ with $\theta = \xi$ with respect to the measure of the process $Z_0^t = \{Z_s\}_{s=0}^t$ with $\theta = 0$. The Radon-Nikodym derivative can be calculated explicitly as[30]

$$\frac{\mathrm{dP}\left(Z_0^t | \theta = \xi\right)}{\mathrm{dP}\left(Z_0^t | \theta = 0\right)} = e^{\frac{1}{2}\left(\frac{\xi}{\alpha^2} Z_t - \frac{1}{2}\frac{\xi^2}{\alpha^2} t\right)}. \tag{27}$$

Hence, the conditional density of θ is

$$p(y; t, Z_t) = \frac{\mathrm{dP}\left(\theta \leq y | \mathcal{F}_t\right)}{\mathrm{d}y} = \frac{e^{\frac{1}{2}\left(\frac{y}{\alpha^2} Z_t - \frac{1}{2}\frac{y^2}{\alpha^2} t\right)} p(y)}{\int_{-\infty}^{\infty} e^{\frac{1}{2}\left(\frac{\xi}{\alpha^2} Z_t - \frac{1}{2}\frac{\xi^2}{\alpha^2} t\right)} p(\xi)\, \mathrm{d}\xi}, \tag{28}$$

[29] See p. 8–9 at https://www.uni-ulm.de/fileadmin/website_uni_ulm/mawi.inst.110/lehre/ws13/Workshop_Probab_Anal_Geom/Shiryaev.pdf.

[30] For example, see the Girsanov theorem, specifically Theorem 8.6.4 in Øksendal (2003).

where p is the prior density of θ. The numerator is

$$
\begin{aligned}
&\frac{1}{\sqrt{4\pi\sigma_0^2}} e^{\frac{1}{2}\left(\frac{y}{\alpha^2}Z_t - \frac{1}{2}\frac{y^2}{\alpha^2}t - \frac{1}{2}\frac{(y-X_0)^2}{\sigma_0^2}\right)} \\
&= \frac{1}{\sqrt{4\pi\sigma_0^2}} e^{-\frac{1}{4\sigma_0^2}(1+\sigma_0^2\alpha^{-2}t)\left(y^2 - 2y\frac{\sigma_0^2\alpha^{-2}Z_t+X_0}{1+\sigma_0^2\alpha^{-2}t} + \frac{X_0^2}{1+\sigma_0^2\alpha^{-2}t} + \left[\left(\frac{\sigma_0^2\alpha^{-2}Z_t+X_0}{1+\sigma_0^2\alpha^{-2}t}\right)^2 - \left(\frac{\sigma_0^2\alpha^{-2}Z_t+X_0}{1+\sigma_0^2\alpha^{-2}t}\right)^2\right]\right)} \\
&= \frac{1}{\sqrt{4\pi\sigma_0^2}} e^{-\frac{1}{4\sigma_0^2}(1+\sigma_0^2\alpha^{-2}t)\left(y - \frac{\sigma_0^2\alpha^{-2}Z_t+X_0}{1+\sigma_0^2\alpha^{-2}t}\right)^2} e^{\frac{-(1+\sigma_0^2\alpha^{-2}t)}{4\sigma_0^2}\left[\frac{X_0^2}{1+\sigma_0^2\alpha^{-2}t} - \left(\frac{\sigma_0^2\alpha^{-2}Z_t+X_0}{1+\sigma_0^2\alpha^{-2}t}\right)^2\right]}.
\end{aligned} \tag{29}
$$

Therefore, by exploiting the density of normal distribution with mean

$$
\frac{X_0 + \sigma_0^2\alpha^{-2}Z_t}{1+\sigma_0^2\alpha^{-2}t} \tag{30}
$$

and variance

$$
\frac{2\sigma_0^2}{1+\sigma_0^2\alpha^{-2}t}, \tag{31}
$$

the denominator in (28) is

$$
e^{\frac{-(1+\sigma_0^2\alpha^{-2}t)}{4\sigma_0^2}\left[\frac{X_0^2}{1+\sigma_0^2\alpha^{-2}t} - \left(\frac{\sigma_0^2\alpha^{-2}Z_t+X_0}{1+\sigma_0^2\alpha^{-2}t}\right)^2\right]} \frac{1}{\sqrt{1+\sigma_0^2\alpha^{-2}t}}. \tag{32}
$$

Hence, by putting (28), (29), and (32) together, we can see that the conditional density is

$$
p(y; t, Z_t) = \frac{1}{\sqrt{2\pi\frac{2\sigma_0^2}{1+\sigma_0^2\alpha^{-2}t}}} e^{-\frac{1}{4\sigma_0^2}(1+\sigma_0^2\alpha^{-2}t)\left(y - \frac{\sigma_0^2\alpha^{-2}Z_t+X_0}{1+\sigma_0^2\alpha^{-2}t}\right)^2}, \tag{33}
$$

which is the density of the normal distribution with mean

$$
\frac{X_0 + \sigma_0^2\alpha^{-2}Z_t}{1+\sigma_0^2\alpha^{-2}t} = \frac{\sigma_0^{-2}X_0 + \alpha^{-2}Z_t}{\sigma_0^{-2} + \alpha^{-2}t} \tag{34}
$$

and variance

$$
\frac{2\sigma_0^2}{1+\sigma_0^2\alpha^{-2}t} = \frac{2}{\sigma_0^{-2} + \alpha^{-2}t}. \tag{35}
$$

□

A.2 Structure of the unrestricted confidence stopping problem

Let us recall the full unrestricted confidence stopping problem (18)

$$
\inf_{\tau' \in \mathcal{T}} \mathrm{E}\left[\Phi\left(\frac{\sigma_0^{-2}X_0 + \alpha^{-2}Z_{\tau'}}{\sqrt{2(\sigma_0^{-2} + \alpha^{-2}\tau')}}\right)\Phi\left(-\frac{\sigma_0^{-2}X_0 + \alpha^{-2}Z_{\tau'}}{\sqrt{2(\sigma_0^{-2} + \alpha^{-2}\tau')}}\right) + \bar{c}\tau'\right]. \tag{36}
$$

Here, we discuss the structure of this problem.

We can restart process $\{Z_t\}_{t\geq 0}$ after time s

$$Z_{s+t} = \theta(s+t) + \alpha\sqrt{2}B_{s+t} = Z_s + \theta t + \alpha\sqrt{2}(B_{s+t} - B_s), \tag{37}$$

where we denote the new starting point $z := Z_s$ and the new standard Brownian motion $\widetilde{B}_t := B_{s+t} - B_s$ for $t \geq 0$. This motivates us to introduce process $\{Z_t^z\}_{t\geq 0}$ with the same differential as $\{Z_t\}_{t\geq 0}$, but starting at z

$$Z_t^z = z + \theta t + \alpha\sqrt{2}\widetilde{B}_t,\ t \geq 0. \tag{38}$$

The innovation representation (Liptser and Shiryaev, 2000, section 7.4) of $\{Z_t^z\}_{t\geq 0}$ starts with a different initial information set, $\widetilde{\mathcal{F}}_t = \mathcal{F}_{s+t}$ for $t \geq 0$, thus the prior is $N(X_s, \sigma_s^2)$. Hence, its innovation representation is

$$Z_t^z = z + \int_0^t \widetilde{X}_r \,\mathrm{d}r + \alpha\sqrt{2}\overline{\widetilde{B}}_t, \tag{39}$$

where

$$\widetilde{X}_r = \mathrm{E}\left[\theta|\widetilde{\mathcal{F}}_r\right] = \frac{\sigma_0^{-2}X_0 + \alpha^{-2}Z_r^z}{\sigma_0^{-2} + \alpha^{-2}(s+r)} \tag{40}$$

and

$$\overline{\widetilde{B}}_t = \frac{1}{\alpha\sqrt{2}}\left(\theta t + \alpha\sqrt{2}\widetilde{B}_t - \int_0^t \widetilde{X}_r \,\mathrm{d}r\right) \tag{41}$$

is a standard Brownian motion.

We can introduce process $\{Y_t^{(s,z)}\}_{t\geq 0}$ such that $Y_t^{(s,z)} = (s+t, Z_t^z)'$ for $t \geq 0$ with the differential

$$\mathrm{d}Y_t^{(s,z)} = \begin{pmatrix} 1 \\ \frac{\sigma_0^{-2}X_0 + \alpha^{-2}Z_t^z}{\sigma_0^{-2} + \alpha^{-2}(s+t)} \end{pmatrix} \mathrm{d}t + \begin{pmatrix} 0 \\ \alpha\sqrt{2} \end{pmatrix} \mathrm{d}\overline{\widetilde{B}}_t,\ t \geq 0,\ Y_0^{(s,z)} = (s,z). \tag{42}$$

Since this process is Markovian (it always restarts anew), the unrestricted confidence stopping problem (18) has the structure of the problem (2.2.2) of Peskir and Shiryaev (2006). However, their condition (2.2.1) is not satisfied in our problem because of the potentially unbounded sampling cost $\bar{c}t$. Nevertheless, their comments on p. 27 and 2 indicate that this may not be a problem, especially if we restrict our search for optimal stopping times to $\tau' \in \mathcal{T}$ such that $\mathrm{E}[\tau'] < \infty$ for which the expectation of the cost upon stopping is well defined. Restriction to such stopping times is innocuous in our problem (18) because, by boundedness of the function $x \mapsto \Phi(x)\Phi(-x)$, stopping times that are expected to be infinite are dominated by immediate stopping due to unbounded sampling cost.

A.3 Subcontinuation region for the unrestricted confidence stopping problem

Lemma A.2. *Let*

$$\mathring{C}_C = \left\{ (r, x) \in [0, \infty) \times \mathbb{R} \colon \frac{\alpha^{-2}}{\sigma_0^{-2} + \alpha^{-2} r} \varphi^2 \left(\frac{\sigma_0^{-2} X_0 + \alpha^{-2} x}{\sqrt{2(\sigma_0^{-2} + \alpha^{-2} r)}} \right) > \bar{c} \right\}. \tag{43}$$

It is not optimal to stop sampling for confidence in this region, i.e., $\mathring{C}_C \subseteq C_C$.

Proof. Consider process $\{Y_t^{(s,z)}\}_{t \geq 0}$ introduced in (42) and function f introduced in (19). By applying the Itô formula, we obtain

$$f(Y_t^{(s,z)}) = f(s, z) + \int_0^t Af(s + r, Z_r^z) \,\mathrm{d}r + \int_0^t \alpha\sqrt{2} \frac{\partial f}{\partial x}(s + r, Z_r^z) \,\mathrm{d}\bar{\tilde{B}}_r, \tag{44}$$

where A is a differential operator acting on f such that, when evaluated at (r, x),

$$Af(r, x) = \frac{\partial f}{\partial r}(r, x) + \frac{\sigma_0^{-2} X_0 + \alpha^{-2} x}{\sigma_0^{-2} + \alpha^{-2} r} \frac{\partial f}{\partial x}(r, x) + \alpha^2 \frac{\partial^2 f}{\partial x^2}(r, x). \tag{45}$$

Similar to Lemma 7.3.2 of Øksendal (2003), it can be proved that for a stopping time τ' such that $\mathrm{E}[\tau'] < \infty$,

$$\mathrm{E}\left[\int_0^{\tau'} \alpha\sqrt{2} \frac{\partial f}{\partial x}(s + r, Z_r^z) \,\mathrm{d}\bar{\tilde{B}}_r \right] = 0, \tag{46}$$

thus, from (44),

$$\mathrm{E}\left[f(Y_{\tau'}^{(s,z)}) \right] = f(s, z) + \mathrm{E}\left[\int_0^{\tau'} Af(s + r, Z_r^z) \,\mathrm{d}r \right]. \tag{47}$$

Consider set

$$\mathring{C}_C = \{(r, x) \in [0, \infty) \times \mathbb{R} \colon Af(r, x) < 0\}. \tag{48}$$

For $(s, z) \in \mathring{C}_C$ and a bounded open set V such that $(s, z) \in V \subset \mathring{C}_C$, consider stopping time

$$\tau_V = \inf\{t \geq 0 \colon Y_t^{(s,z)} \notin V\}. \tag{49}$$

Then, we can see from (47) that

$$\mathrm{E}\left[f(Y_{\tau_V}^{(s,z)}) \right] < f(s, z). \tag{50}$$

Hence, as long as we are in $\mathring{C}_C$, it is not optimal to stop sampling for confidence because we expect the total confidence cost f to decrease at least until we exit from $\mathring{C}_C$. Therefore, $\mathring{C}_C \subseteq C_C$ and that is why we call $\mathring{C}_C$ a *subcontinuation region* for the unrestricted confidence stopping problem (18).[31]

[31] In fact, $\mathring{C}_C \neq C_C$ in general. As Øksendal (2003, p. 205) writes, this is “the typical situation” in these problems.

To express region $\mathring{C}_C$ explicitly, denote

$$M(r,x) = \frac{\sigma_0^{-2} X_0 + \alpha^{-2} x}{\sqrt{2(\sigma_0^{-2} + \alpha^{-2} r)}}. \tag{51}$$

First, we calculate

$$M_r(r,x) = \frac{\partial M(r,x)}{\partial r} = -\frac{\alpha^{-2}}{2(\sigma_0^{-2} + \alpha^{-2} r)} M(r,x), \tag{52}$$

$$M_x(r,x) = \frac{\partial M(r,x)}{\partial x} = \frac{\alpha^{-2}}{\sqrt{2(\sigma_0^{-2} + \alpha^{-2} r)}}, \tag{53}$$

$$M_{xx}(r,x) = \frac{\partial^2 M(r,x)}{\partial x^2} = 0. \tag{54}$$

Second, using (54) and the following properties of the standard normal pdf φ

$$\varphi'(x) = -x\varphi(x), \tag{55}$$

$$\varphi(-x) = \varphi(x), \tag{56}$$

we obtain

$$\frac{\partial f(r,x)}{\partial r} = \varphi(M(r,x)) M_r(r,x)(1 - 2\Phi(M(r,x))) + \bar{c}, \tag{57}$$

$$\frac{\partial f(r,x)}{\partial x} = \varphi(M(r,x)) M_x(r,x)(1 - 2\Phi(M(r,x))), \tag{58}$$

$$\frac{\partial^2 f(r,x)}{\partial x^2} = \varphi(M(r,x)) M_x^2(r,x)[-M(r,x)(1 - 2\Phi(M(r,x))) - 2\varphi(M(r,x))]. \tag{59}$$

Finally, plugging expressions (52)–(53) and (57)–(59) into (45) and simplifying yields

$$Af(r,x) = \bar{c} - \frac{\alpha^{-2}}{\sigma_0^{-2} + \alpha^{-2} r} \varphi^2(M(r,x)). \tag{60}$$

Hence, the subcontinuation region is

$$\mathring{C}_C = \left\{ (r,x) \in [0,\infty) \times \mathbb{R} \colon \frac{\alpha^{-2}}{\sigma_0^{-2} + \alpha^{-2} r} \varphi^2 \left(\frac{\sigma_0^{-2} X_0 + \alpha^{-2} x}{\sqrt{2(\sigma_0^{-2} + \alpha^{-2} r)}} \right) > \bar{c} \right\}. \tag{61}$$

□

A.4 Boundedness of the unrestricted confidence stopping time

Intuitively, the bound is the rightmost point of $\mathring{C}_C$.

Proof of Proposition 1. Denote $T_c := \alpha^2[(2\pi\alpha^2\bar{c})^{-1} - \sigma_0^{-2}]$ and suppose $T_c > 0$.[32] Toward contradiction, suppose we have an optimal stopping time of the unconstrained confidence stopping

[32] T_c is the (unique) solution of equation $Af(r, -\alpha^2\sigma_0^{-2}X_0) = 0$, where $Af(r,x)$ is given by (60) (the motivation for this choice will become obvious in the proof). Notice that we are looking for the solution $r \in \mathbb{R}$, so it can also be negative.

problem (18), σ, that ends beyond T_c with strictly positive probability, i.e., $\mathrm{P}(\sigma > T_c) > 0$. Nevertheless, we still assume that $\mathrm{E}[\sigma] < \infty$ because stopping times that are expected to be infinite cannot be optimal as we already argued at the end of section A.2.[33]

With the use of notation from section A.3 and by using (47), the expected loss in problem (18) for σ can be written as

$$\mathrm{E}\left[f(Y_\sigma^{(0,0)})\right] = f(0,0) + \mathrm{E}\left[\int_0^\sigma Af(r, Z_r)\,\mathrm{d}r\right]. \tag{62}$$

Similarly, we can write for stopping time $\sigma \wedge T_c$

$$\mathrm{E}\left[f(Y_{\sigma\wedge T_c}^{(0,0)})\right] = f(0,0) + \mathrm{E}\left[\int_0^{\sigma\wedge T_c} Af(r, Z_r)\,\mathrm{d}r\right]. \tag{63}$$

Hence,

$$\mathrm{E}\left[f(Y_\sigma^{(0,0)})\right] - \mathrm{E}\left[f(Y_{\sigma\wedge T_c}^{(0,0)})\right] = \mathrm{E}\left[\int_0^\sigma Af(r, Z_r)\,\mathrm{d}r - \int_0^{\sigma\wedge T_c} Af(r, Z_r)\,\mathrm{d}r\right]. \tag{64}$$

By the Law of Iterated Expectations, this can be simplified to

$$\mathrm{E}\left[\int_{T_c}^\sigma Af(r, Z_r)\,\mathrm{d}r \middle| \sigma > T_c\right] \mathrm{P}(\sigma > T_c). \tag{65}$$

But from (60) and the definition of T_c, we can see that for any $r > T_c$, $Af(r, x) > 0\ \forall x \in \mathbb{R}$. Therefore, stopping time $\sigma \wedge T_c$ (which differs from σ with strictly positive probability) achieves strictly lower expected loss, which is a contradiction to σ being optimal.

Finally, when $T_c \leq 0$ (which corresponds to high cost of time $\bar{c}$), we can see from (47) and (60) that it is never beneficial to wait, so any optimal stopping time is trivially bounded by 0 almost surely. □

[33] As a consequence, we also have $\mathrm{P}(\sigma < \infty) = 1$.

B Deterministic stopping times

Finding the optimal deterministic stopping time

$$\inf_{t\in[0,\infty)} u(t;s,z) \tag{66}$$

amounts to analyzing the derivative of function u defined in (20).

Lemma B.1.

$$\begin{aligned}\frac{\partial u(t;s,z)}{\partial t} &= \bar{c} - \frac{\alpha^{-2}}{\sigma_0^{-2}+\alpha^{-2}(s+t)}\sqrt{\frac{\sigma_0^{-2}+\alpha^{-2}s}{\sigma_0^{-2}+\alpha^{-2}s+2\alpha^{-2}t}} \\ &\quad \times \frac{1}{2\pi}\exp\left(-\frac{1}{2}\frac{\sigma_0^{-2}+\alpha^{-2}(s+t)}{\sigma_0^{-2}+\alpha^{-2}s+2\alpha^{-2}t}\frac{(\sigma_0^{-2}X_0+\alpha^{-2}z)^2}{\sigma_0^{-2}+\alpha^{-2}s}\right).\end{aligned} \tag{67}$$

Proof. As noticed by Øksendal (2003) in (8.1.1), we can see from (47) that

$$\frac{\partial u(t;s,z)}{\partial t} = \mathrm{E}\left[Af(s+t, Z_t^z)\right], \tag{68}$$

where $Af(r,x)$ takes the form (60). Hence, we need to calculate

$$\mathrm{E}\left[\varphi^2\left(\frac{\sigma_0^{-2}X_0+\alpha^{-2}Z_t^z}{\sqrt{2(\sigma_0^{-2}+\alpha^{-2}(s+t))}}\right)\right]. \tag{69}$$

By the Law of Iterated Expectations

$$\begin{aligned}&\mathrm{E}\left[\varphi^2\left(\frac{\sigma_0^{-2}X_0+\alpha^{-2}Z_t^z}{\sqrt{2(\sigma_0^{-2}+\alpha^{-2}(s+t))}}\right)\right] \\ &= \mathrm{E}\left[\mathrm{E}\left[\varphi^2\left(\frac{\sigma_0^{-2}X_0+\alpha^{-2}Z_t^z}{\sqrt{2(\sigma_0^{-2}+\alpha^{-2}(s+t))}}\right)\middle|\theta\right]\right].\end{aligned} \tag{70}$$

Since

$$\left.\frac{\sigma_0^{-2}X_0+\alpha^{-2}Z_t^z}{\sqrt{2(\sigma_0^{-2}+\alpha^{-2}(s+t))}}\right|\theta \sim N(\mu,\sigma^2) \tag{71}$$

with

$$\mu = \frac{\sigma_0^{-2}X_0+\alpha^{-2}(z+\theta t)}{\sqrt{2(\sigma_0^{-2}+\alpha^{-2}(s+t))}}, \tag{72}$$

$$\sigma^2 = \frac{\alpha^{-2}t}{\sigma_0^{-2}+\alpha^{-2}(s+t)}, \tag{73}$$

the interior expectation in (70) is

$$\begin{aligned}\mathrm{E}\left[\varphi^2\left(\frac{\sigma_0^{-2}X_0+\alpha^{-2}Z_t^z}{\sqrt{2(\sigma_0^{-2}+\alpha^{-2}(s+t))}}\right)\middle|\theta\right] &= \int_{-\infty}^{\infty}\varphi^2(x)\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}\mathrm{d}x \\ &= (2\pi)^{-\frac{3}{2}}\sigma^{-1}\int_{-\infty}^{\infty}e^{-x^2-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}\mathrm{d}x.\end{aligned} \tag{74}$$

We can write

$$-x^2 - \frac{1}{2}\frac{(x-\mu)^2}{\sigma^2} = -\frac{1}{2\sigma^2}((1+2\sigma^2)x^2 - 2x\mu + \mu^2)$$
$$= -\frac{1+2\sigma^2}{2\sigma^2}\left(x^2 - 2x\frac{\mu}{1+2\sigma^2} + \frac{\mu^2}{(1+2\sigma^2)^2}\right)$$
$$+ \frac{\mu^2}{2\sigma^2(1+2\sigma^2)} - \frac{\mu^2}{2\sigma^2}. \tag{75}$$

Plugging (75) to (74) yields

$$\mathrm{E}\left[\varphi^2\left(\frac{\sigma_0^{-2}X_0 + \alpha^{-2}Z_t^z}{\sqrt{2(\sigma_0^{-2}+\alpha^{-2}(s+t))}}\right)\middle|\theta\right]$$
$$= (2\pi)^{-\frac{3}{2}}\sigma^{-1}\int_{-\infty}^{\infty} e^{-\frac{1}{2}\frac{1+2\sigma^2}{\sigma^2}\left(x-\frac{\mu}{1+2\sigma^2}\right)^2}\mathrm{d}x \cdot e^{\frac{\mu^2}{2\sigma^2(1+2\sigma^2)}-\frac{\mu^2}{2\sigma^2}}$$
$$= (2\pi)^{-\frac{3}{2}}\sigma^{-1}\sqrt{2\pi\frac{\sigma^2}{1+2\sigma^2}}\underbrace{\frac{1}{\sqrt{2\pi\frac{\sigma^2}{1+2\sigma^2}}}\int_{-\infty}^{\infty} e^{-\frac{1}{2}\frac{1+2\sigma^2}{\sigma^2}\left(x-\frac{\mu}{1+2\sigma^2}\right)^2}\mathrm{d}x}_{=1} \cdot e^{-\frac{\mu^2}{1+2\sigma^2}}$$
$$= \frac{1}{2\pi\sqrt{1+2\sigma^2}}e^{-\frac{\mu^2}{1+2\sigma^2}}. \tag{76}$$

By Lemma A.1, the beliefs about θ at (s, z) are

$$N(X_s, \sigma_s^2) \tag{77}$$

with

$$X_s = \frac{\sigma_0^{-2}X_0 + \alpha^{-2}z}{\sigma_0^{-2} + \alpha^{-2}s}, \tag{78}$$
$$\sigma_s^2 = \frac{2}{\sigma_0^{-2} + \alpha^{-2}s}. \tag{79}$$

Hence, we can insert (76) back to (70) and calculate the outer expectation explicitly. Let us focus on the part in (76) that depends on θ

$$\mathrm{E}\left[e^{-\frac{\mu^2}{1+2\sigma^2}}\right] = \frac{1}{\sqrt{2\pi\sigma_s^2}}\int_{-\infty}^{\infty} e^{-\left(\frac{\mu^2}{1+2\sigma^2}+\frac{(\theta-X_s)^2}{2\sigma_s^2}\right)}\mathrm{d}\theta. \tag{80}$$

We can write

$$\frac{\mu^2}{1+2\sigma^2} + \frac{(\theta-X_s)^2}{2\sigma_s^2} = (2\sigma_s^2(1+2\sigma^2))^{-1}[(\theta^2 - 2\theta X_s + X_s^2)(1+2\sigma^2) + 2\sigma_s^2\mu^2]. \tag{81}$$

By expanding μ according to (72) and denoting

$$\sigma_{s+t}^2 = \frac{2}{\sigma_0^{-2} + \alpha^{-2}(s+t)}, \tag{82}$$

we can continue rewriting (81) as

$$
\begin{aligned}
(2\sigma_s^2(1+2\sigma^2))^{-1}\Big[\theta^2(1+2\sigma^2+\frac{1}{2}\sigma_s^2\sigma_{s+t}^2\alpha^{-4}t^2)&\\
-2\theta(X_s(1+2\sigma^2)-\frac{1}{2}\sigma_s^2\sigma_{s+t}^2(\sigma_0^{-2}X_0+\alpha^{-2}z)\alpha^{-2}t)&\\
+X_s^2(1+2\sigma^2)+\frac{1}{2}\sigma_s^2\sigma_{s+t}^2(\sigma_0^{-2}X_0+\alpha^{-2}z)^2\Big].&
\end{aligned} \tag{83}
$$

Using $\sigma^2 = \sigma_{s+t}^2\alpha^{-2}t/2$ and $X_s = \sigma_s^2(\sigma_0^{-2}X_0+\alpha^{-2}z)/2$, we can simplify the terms in (83) and introduce term K to simplify further calculations

$$
\begin{aligned}
K &:= 1+2\sigma^2+\frac{1}{2}\sigma_s^2\sigma_{s+t}^2\alpha^{-4}t^2 = 1+\sigma_{s+t}^2\alpha^{-2}t(1+\frac{1}{2}\sigma_s^2\alpha^{-2}t)\\
&= 1+\frac{2\alpha^{-2}t}{\sigma_0^{-2}+\alpha^{-2}(s+t)}\left(1+\frac{\alpha^{-2}t}{\sigma_0^{-2}+\alpha^{-2}s}\right) = \frac{\sigma_0^{-2}+\alpha^{-2}s+2\alpha^{-2}t}{\sigma_0^{-2}+\alpha^{-2}s},
\end{aligned} \tag{84}
$$

$$
X_s(1+2\sigma^2)-\frac{1}{2}\sigma_s^2\sigma_{s+t}^2(\sigma_0^{-2}X_0+\alpha^{-2}z)\alpha^{-2}t = X_s, \tag{85}
$$

$$
\begin{aligned}
&X_s^2(1+2\sigma^2)+\frac{1}{2}\sigma_s^2\sigma_{s+t}^2(\sigma_0^{-2}X_0+\alpha^{-2}z)^2\\
&\quad = X_s^2(1+\sigma_{s+t}^2\alpha^{-2}t+2\sigma_s^{-2}\sigma_{s+t}^2) = X_s^2[1+\sigma_{s+t}^2\underbrace{(\alpha^{-2}t+2\sigma_s^{-2})}_{2\sigma_{s+t}^{-2}}] = 3X_s^2.
\end{aligned} \tag{86}
$$

Then, we can continue with (83)

$$
\begin{aligned}
&(2\sigma_s^2(1+2\sigma^2))^{-1}\left[K\theta^2-2\theta X_s+3X_s^2\right]\\
&\quad = (2\sigma_s^2(1+2\sigma^2))^{-1}K\left[\theta^2-2\theta\frac{X_s}{K}+\frac{X_s^2}{K^2}-\frac{X_s^2}{K^2}+\frac{3X_s^2}{K}\right]\\
&\quad = \frac{K}{2\sigma_s^2(1+2\sigma^2)}\left[\left(\theta-\frac{X_s}{K}\right)^2-\frac{X_s^2}{K^2}+\frac{3X_s^2}{K}\right].
\end{aligned} \tag{87}
$$

Plugging this back to (80) yields

$$
\begin{aligned}
\mathrm{E}\left[e^{-\frac{\mu^2}{1+2\sigma^2}}\right] &= \exp\left(-\frac{K}{2\sigma_s^2(1+2\sigma^2)}\left(-\frac{X_s^2}{K^2}+\frac{3X_s^2}{K}\right)\right)\\
&\quad\times\frac{1}{\sqrt{2\pi\sigma_s^2}}\int_{-\infty}^{\infty}\exp\left(-\frac{1}{2\frac{\sigma_s^2(1+2\sigma^2)}{K}}\left(\theta-\frac{X_s}{K}\right)^2\right)\mathrm{d}\theta\\
&= \exp\left(\frac{1}{2\sigma_s^2(1+2\sigma^2)}\left(\frac{X_s^2}{K}-3X_s^2\right)\right)\sqrt{\frac{1+2\sigma^2}{K}}\\
&\quad\times\underbrace{\frac{1}{\sqrt{2\pi\sigma_s^2\frac{1+2\sigma^2}{K}}}\int_{-\infty}^{\infty}\exp\left(-\frac{1}{2\frac{\sigma_s^2(1+2\sigma^2)}{K}}\left(\theta-\frac{X_s}{K}\right)^2\right)\mathrm{d}\theta}_{=1}\\
&= \exp\left(\frac{1}{2\sigma_s^2(1+2\sigma^2)}\left(\frac{X_s^2}{K}-3X_s^2\right)\right)\sqrt{\frac{1+2\sigma^2}{K}}.
\end{aligned} \tag{88}
$$

The following steps lead to a simplification of expression (88):

$$
\begin{aligned}
2\sigma_s^2(1+2\sigma^2) &= 4\frac{\sigma_0^{-2}+\alpha^{-2}s+3\alpha^{-2}t}{(\sigma_0^{-2}+\alpha^{-2}s)(\sigma_0^{-2}+\alpha^{-2}(s+t))}, \\
1-3K &= \frac{-2(\sigma_0^{-2}+\alpha^{-2}s+3\alpha^{-2}t)}{\sigma_0^{-2}+\alpha^{-2}s}, \\
\frac{1-3K}{K} &= \frac{-2(\sigma_0^{-2}+\alpha^{-2}s+3\alpha^{-2}t)}{\sigma_0^{-2}+\alpha^{-2}s+2\alpha^{-2}t}, \\
\frac{1}{2\sigma_s^2(1+2\sigma^2)}\frac{1-3K}{K}X_s^2 &= -\frac{1}{2}\frac{\sigma_0^{-2}+\alpha^{-2}s+\alpha^{-2}t}{\sigma_0^{-2}+\alpha^{-2}s+2\alpha^{-2}t}\frac{(\sigma_0^{-2}X_0+\alpha^{-2}z)^2}{\sigma_0^{-2}+\alpha^{-2}s}.
\end{aligned} \tag{89}
$$

Therefore, by putting together (60), (68), (70), (76), (84), (88), and (89), we obtain

$$
\begin{aligned}
\frac{\partial u(t;s,z)}{\partial t} = \bar{c} &- \frac{\alpha^{-2}}{\sigma_0^{-2}+\alpha^{-2}(s+t)}\sqrt{\frac{\sigma_0^{-2}+\alpha^{-2}s}{\sigma_0^{-2}+\alpha^{-2}s+2\alpha^{-2}t}} \\
&\times \frac{1}{2\pi}\exp\left(-\frac{1}{2}\frac{\sigma_0^{-2}+\alpha^{-2}(s+t)}{\sigma_0^{-2}+\alpha^{-2}s+2\alpha^{-2}t}\frac{(\sigma_0^{-2}X_0+\alpha^{-2}z)^2}{\sigma_0^{-2}+\alpha^{-2}s}\right).
\end{aligned} \tag{90}
$$

□

C Confidence in the Wald model

Fudenberg et al. (2018) also study a variant of the model with a priori known difference between the options $\theta_d := |\theta_l - \theta_r|$ with the only uncertainty about which option is the best. This corresponds to the binomial prior about θ:

$$\theta = \begin{cases} \theta_d & \text{with probability } \mu_0 \in (0,1), \\ -\theta_d & \text{with probability } 1-\mu_0. \end{cases} \tag{91}$$

This model is the continuous-time version of the model of Wald (1947). As Fudenberg et al. (2018) write, Shiryaev (2007) provides a solution to this problem.

Theorem C.1. *For the model with the binomial prior, there exists $k > 0$ such that the minimal optimal stopping time is $\hat{\tau} = \inf\{t \geq 0 : |l_t| = k\}$, where $l_t = \log\left(\frac{p_t}{1-p_t}\right)$ and $p_t = \mathrm{P}\left(\theta = \theta_d | \mathcal{F}_t\right)$.*

Since p_t can be seen as belief confidence (in the left option), this solution is often interpreted as a decision maker committing to a desired level of confidence ex ante. Since there is no further deliberation in this model, the predicted (one-stage) decision confidence is constant $\frac{e^k}{1+e^k}$.

Given the interest of this paper, we can ask whether there is a scope for two-stage confidence in this model. Motivated by calculations (13), we will analyze the unconstrained stopping problem

$$\inf_{\tau' \in \mathcal{T}} \mathrm{E}\left[p_{\tau'}(1-p_{\tau'}) + \bar{c}\tau'\right]. \tag{92}$$

The belief process $\{p_t\}_{t\geq 0}$ solves the stochastic differential equation

$$\mathrm{d}p_t = p_t(1-p_t)\frac{\sqrt{2}}{\alpha}\mathrm{d}W_t, \tag{93}$$

where $\{W_t\}_{t\geq 0}$ is a Brownian motion with respect to $\mathcal{F}_t$ (Morris and Strack, 2019, Lemma 1). By applying the differential operator of the process $\{(s+t, p_t^p)\}_{t\geq 0}$ to the function $(t,p) \mapsto p(1-p) + \bar{c}t$ and inspecting where it is negative, we obtain a candidate stopping region in the (t,p)-space

$$\mathring{C}_W = \left\{(t,p) \in [0,\infty) \times [0,1] : p(1-p) > \alpha\sqrt{\frac{\bar{c}}{2}}\right\}. \tag{94}$$

If $\bar{c} > \frac{1}{8\alpha^2}$, this region is empty and it is always optimal to stop immediately. Otherwise this region implicitly prescribes constant boundaries for p_t, $0 \leq p_* \leq \frac{1}{2} \leq p^* \leq 1$.

We will show that the optimal continuation region takes the same form as $\mathring{C}_W$, i.e.,

$$C_W := \left\{(t,p) \in [0,\infty) \times [0,1] : \underline{p} < p < \bar{p}\right\} \tag{95}$$

with the boundaries $\underline{p} \leq p_* \leq \frac{1}{2} \leq p^* \leq \bar{p}$ (since $\mathring{C}_W \subseteq C_W$). To do that, we follow a similar approach as Øksendal (2003, p. 210).

Denote the value function

$$V(s, p) = \inf_{\tau \in \mathcal{T}} \mathrm{E}\left[p_\tau^p(1 - p_\tau^p) + \bar{c}(\tau + s)\right], \tag{96}$$

where

$$p_t^p = p + \int_0^t p_r(1 - p_r)\frac{\sqrt{2}}{\alpha}\mathrm{d}W_r. \tag{97}$$

Since

$$\begin{aligned} V(s - t_0, p) &= \inf_{\tau \in \mathcal{T}} \mathrm{E}\left[p_\tau^p(1 - p_\tau^p) + \bar{c}(\tau + s - t_0)\right] \\ &= \inf_{\tau \in \mathcal{T}} \mathrm{E}\left[p_\tau^p(1 - p_\tau^p) + \bar{c}(\tau + s)\right] - \bar{c}t_0 \\ &= V(s, p) - \bar{c}t_0, \end{aligned}$$

we can show that C_W is invariant in time:

$$\begin{aligned} C_W + (t_0, 0) &= \{(t_0 + t, p) : (t, p) \in C_W\} \\ &= \{(s, p) : (s - t_0, p) \in C_W\} \\ &= \{(s, p) : V(s - t_0, p) < p(1 - p) + \bar{c}(s - t_0)\} \\ &= \{(s, p) : V(s, p) - \bar{c}t_0 < p(1 - p) + \bar{c}(s - t_0)\} \\ &= \{(s, p) : V(s, p) < p(1 - p) + \bar{c}s\} = C_W. \end{aligned}$$

Hence, C_W must take the form (95).

This analysis reveals that if the parameters $\bar{c}, c, \alpha, \mu_0$ are such that $\frac{e^k}{1+e^k} \geq \bar{p}$, confidence will always be one-stage with the resulting confidence level $\frac{e^k}{1+e^k}$; otherwise confidence will always be two-stage with the two potential levels $\underline{p}$ (for error recognition) and $\bar{p}$ (for correctness confirmation).

D Numerical solution

In this section, we describe the numerical methods for finding the optimal decision and confidence stopping boundaries ∂C_D and ∂C_C, respectively.

The decision boundary is found by building on the code and descriptions of Fudenberg et al. (2018) from their paper, Online Appendix, and replication folder accessible at OPENICPSR. In particular, script `h_patching.m`[34] imports all csv files[35] from the subfolder 'solver/mat' of their replication folder and patches them together as described in Section 4.1.2 of their Online Appendix. Moreover, script `h_patching.m` also (i) reduces the size of the resulting vector by dropping unnecessary points[36] and (ii) rectifies monotonicity as mentioned in Footnote 2 of their Online Appendix.[37] The resulting vector representing their h function is saved and used by function `opt_dec_bound.m` to calculate the optimal decision boundary, which Fudenberg et al. (2018) denote by b^* and define in Theorem 4 (page 3661), with the generalization to non-symmetric prior means in their Footnote 22 (page 3662).[38] Function `opt_dec_bound.m` thus takes as input characteristics/parameters of an individual (c, σ_0, α, X_0) and the h vector, and outputs upper and lower portions of the decision stopping boundary.

The confidence boundary is computed by function `conf_bound.m`, which takes as input characteristics/parameters of an individual ($\bar{c}$, σ_0, α, X_0) and outputs a list of points in the (t, z)-space. The computation is performed by backward induction. Denote by $\mathcal{U}$ the numerical approximation of the value function U defined in (21).

$\mathcal{U}$ is computed as follows. Fix a space grid by choosing sufficiently large (in absolute value) $\bar{z} > 0$ and $\underline{z} < 0$ and a sufficiently small $\Delta_z > 0$: $\zeta = \{\underline{z}, \underline{z} + \Delta_z, \ldots, \bar{z} - \Delta_z, \bar{z}\}$.[39] Start at T

[34] The codes are available upon request.

[35] Except for `h-30-100.csv` because they write in Section 5 of `readme.pdf` that it is just a mistake.

[36] The original resulting vector has more than 95% of useless constant parts, which we drop and use linear interpolation instead to speed up further computations.

[37] 9 points are redefined.

[38] See also Lemma O.4 of their Online Appendix for details of these calculations.

[39] Δ_z and Δ_t are set to 0.1 (unless the time bound from Proposition 1 is smaller than 5; then a smaller Δ_t is chosen). Initially, $\bar{z}$ and $\underline{z}$ are set ad hoc to $50 - \alpha^2 \sigma_0^{-2} X_0$ and $-50 - \alpha^2 \sigma_0^{-2} X_0$, respectively. During the backward induction computation, if these initial $\bar{z}$ and $\underline{z}$ become insufficient (the computed continuation region approaches $0.8\bar{z}$ from below or $0.8\underline{z}$ from above), the backward induction repeatedly restarts anew with a larger range of bounds until it is sufficient.

slightly above the bound from Proposition 1 where it is surely optimal to stop for any $z \in \mathbb{R}$, i.e., set $\mathcal{U}(T,z) = f(T,z)$. Going backwards in time by Δ_t, suppose we already computed $\mathcal{U}(s,z),\ z \in \zeta$, for all $s \in \{t+\Delta_t, \dots, T\}$. At time point t for space point $z \in \zeta$, we want to compare $f(t,z)$ and

$$\mathrm{E}_{(t,z)}\left[U(t+\Delta_t, Z^z_{\Delta_t})\right] = \int_{-\infty}^{\infty} U(t+\Delta_t, x) g_{(t,z)}(x)\,\mathrm{d}x, \tag{98}$$

where $g_{(t,z)}$ is the pdf of the normal distribution

$$N\left(z + \frac{\sigma_0^{-2}X_0 + \alpha^{-2}z}{\sigma_0^{-2} + \alpha^{-2}t}\Delta_t, 2\alpha^2\Delta_t + \frac{2}{\sigma_0^{-2} + \alpha^{-2}t}\Delta_t^2\right). \tag{99}$$

Denote the cdf of this distribution $G_{(t,z)}$. To approximate the expected value (98), use a constant approximation of $U(t+\Delta_t, x)$ below $\underline{z}$ and above $\bar{z}$[40] and the trapezoidal rule on ζ

$$\begin{aligned}
\int_{-\infty}^{\infty} & U(t+\Delta_t, x) g_{(t,z)}(x)\,\mathrm{d}x \approx U(t+\Delta_t, \underline{z}) G_{(t,z)}(\underline{z}) \\
&+ \sum_{x \in \{\underline{z}, \underline{z}+\Delta_z, \dots, \bar{z}-\Delta_z\}} \frac{1}{2}[U(t+\Delta_t, x) g_{(t,z)}(x) + U(t+\Delta_t, x+\Delta_z) g_{(t,z)}(x+\Delta_z)]\Delta_z \\
&+ U(t+\Delta_t, \bar{z})(1 - G_{(t,z)}(\bar{z})) \\
&=: \mathcal{E}_{(t,z)}.
\end{aligned} \tag{100}$$

Finally, set

$$\mathcal{U}(t,z) = \min\{f(t,z), \mathcal{E}_{(t,z)}\}. \tag{101}$$

The numerically computed continuation region is

$$\{(t,z) : \mathcal{U}(t,z) < f(t,z)\}. \tag{102}$$

More precisely, for a given $t \in \{0, \Delta_t, \dots, T\}$, we find the extreme points $z \in \zeta$ for which $\mathcal{U}(t,z) < f(t,z)$ and consider them to be on the boundary of the confidence continuation region.

[40]Below $\underline{z}$ and above $\bar{z}$, the MSE part of the loss function is almost zero, so $U(t+\Delta_t, x)$ flattens out to approximately $\bar{c}(t+\Delta_t)$ beyond these points.

E Empirical evaluation of models

Table 1: Evaluation of models using empirical patterns of Moran et al. (2015)

.

	Empirical pattern	Explanation	FSS	2SS
1.	Speed-accuracy trade-off	Higher error rate under time pressure	✓	✓
2.	Slow/fast errors	Error choices can be slower or faster than correct choices	Slow errors	Slow errors
3.	Negative correlation of confidence and difficulty	Positive correlation of confidence and stimulus discriminability	✓	✓
4.	Negative correlation of decision time and confidence		✓	✓
5.	Lower confidence under time pressure		✓	✓
6.	Positive confidence resolution	Higher confidence in correct decisions	✓	✓
7.	Increased confidence resolution under time pressure	Difference in confidence between correct and error choices is higher under time pressure	×	✓
8.1.	Positive correlation of RT2 and difficulty	Negative correlation of RT2 and stimulus discriminability	–	✓
8.2.	Lower RT2 in correct choices		–	✓
8.3.	Negative correlation of RT2 and confidence		–	✓
8.4.	Positive correlation of RT2 and RT		–	×
9.	Decreased confidence resolution for difficult decisions	Difference in confidence between correct and error choices is lower under lower stimulus discriminability	✓	✓
10.	Higher RT2 for correct choices and lower RT2 for errors under higher difficulty		–	✓

FSS is the model of Fudenberg et al. (2018) with confidence determined at decision stopping. 2SS is the model proposed in this paper. We tried the following parameters: (i) $c = 0.06, \sigma_0 = 1, \alpha = 2, X_0 = 0$ with $\bar{c} = 0.02$ for 2SS model ("typical" subject as in Figure 3), (ii) $c = 0.02, \sigma_0 = 1.8, \alpha = 2, X_0 = 0$ with $\bar{c} = 0.012$ for 2SS model ("atypical" subject as in Figure 2). The results are based on 100,000 simulated trials for each model and set of parameters. We tried θ distributed according to (a) the agent's prior $N(X_0, 2\sigma_0^2)$, (b) $N(0, 1)$. Discriminability is measured by $|\theta|$. RT is response time, RT2 is interjudgement time. Time pressure effect was analyzed by decreasing c to 0.05 in (i) and 0.01 in (ii).

F Literature: economic motivation for decision confidence

- Enke and Graeber (2022) demonstrate the economic relevance of decision confidence in predicting behavioral biases.
- Enke et al. (2022) show that confidence modulates attenuation of behavioral biases in aggregate through self-selection into institutions.
- Folke et al. (2016) show that “an explicit representation of confidence is harnessed for subsequent changes of mind.” This may be relevant for markets with the possibility of changing one’s mind, e.g., return policies in online shopping or cancellation insurance in airlines or races.
- Van den Berg et al. (2016) show that people exploit confidence in previous choices to adjust their termination mechanism in subsequent decisions. This may be relevant for projects, which are typically composed of several subsequent decisions directed towards an overarching goal. Moreover, it is well known that demand for many products is derived from demand for other products; however, the demand for the “later” products may also be affected by the mere confidence in the “earlier” products, e.g., demand for electricity may be affected by one’s decision confidence in a purchase of an appliance if one cares about its energy efficiency.
- Purcell and Kiani (2016) show that people use confidence to modulate the adjustment of their strategy when facing negative feedback, i.e., confidence disambiguates between two sources of errors: bad strategy vs. bad information. This suggests an important role of confidence in reinforcement learning and the Credit Assignment Problem in particular.
- Yin et al. (2016) document that consumers exhibit confirmation bias in evaluating online reviews and this bias is amplified by higher confidence in their initial beliefs. Hence, confidence plays a role in information acquisition (Schulz et al., 2021).
- Confidence is often communicated to others and it has a strong influence on their decisions (Vullioud et al., 2017; Shea et al., 2014; Sah et al., 2013; Brewer and Burke, 2002).
- Artiga González et al. (2022) show that voters change their policy preferences (in line with the supported candidate) after merely expressing their support. We suspect that confidence may play a role in this process of formation of policy preferences and polarization.

G Additional figures

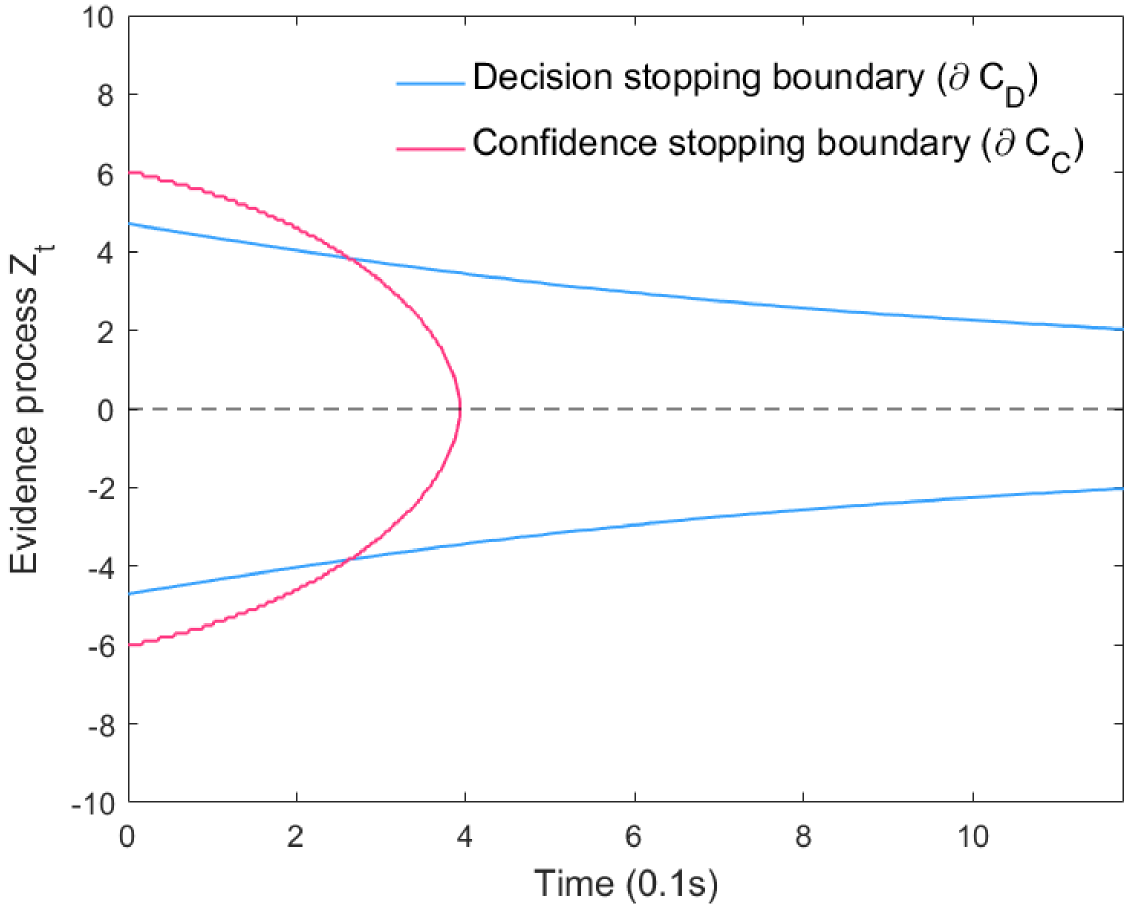

Figure 3: An example of decision and confidence stopping boundaries computed numerically (see Appendix D) for parameters $c = 0.06, \alpha = 2, \sigma_0 = 1, X_0 = 0$ (corresponding to Subject 47 in Table 4 in the Online Appendix of Fudenberg et al. (2018)) and $\bar{c} = 0.02$.

References

Artiga González, T., Capozza, F., and Granic, G. D. (2022). Political support, cognitive dissonance and political preferences. CESifo Working Paper No. 9549.

Baron, J. (2012). The point of normative models in judgment and decision making. *Frontiers in Psychology*, 3:577.

Boldt, A., Blundell, C., and De Martino, B. (2019). Confidence modulates exploration and exploitation in value-based learning. *Neuroscience of consciousness*, 2019(1):niz004.

Brewer, N. and Burke, A. (2002). Effects of testimonial inconsistencies and eyewitness confidence on mock-juror judgments. *Law and Human Behavior*, 26(3):353–364.

Chernoff, H. (1961). Sequential tests for the mean of a normal distribution. In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, volume 1, pages 79–91. University of California Press.

Desender, K., Boldt, A., and Yeung, N. (2018). Subjective confidence predicts information seeking in decision making. *Psychological science*, 29(5):761–778.

Enke, B. and Graeber, T. (2022). Cognitive uncertainty. Unpublished manuscript.

Enke, B., Graeber, T., and Oprea, R. (2022). Confidence, self-selection, and bias in the aggregate. Unpublished manuscript.

Fleming, S. M. and Daw, N. D. (2017). Self-evaluation of decision-making: A general Bayesian framework for metacognitive computation. *Psychological Review*, 124(1):91–114.

Folke, T., Jacobsen, C., Fleming, S. M., and De Martino, B. (2016). Explicit representation of confidence informs future value-based decisions. *Nature Human Behaviour*, 1(1):1–8.

Fudenberg, D., Strack, P., and Strzalecki, T. (2018). Speed, accuracy, and the optimal timing of choices. *American Economic Review*, 108(12):3651–84.

Krajbich, I., Oud, B., and Fehr, E. (2014). Benefits of neuroeconomic modeling: New policy interventions and predictors of preference. *American Economic Review*, 104(5):501–06.

Liptser, R. and Shiryaev, A. (2000). *Statistics of Random Processes, 2nd edn. I and II*. Springer, Berlin.

Marr, D. (1982). *Vision*. New York, NY: WH Freeman.

Moran, R., Teodorescu, A. R., and Usher, M. (2015). Post choice information integration as a causal determinant of confidence: Novel data and a computational account. *Cognitive psychology*, 78:99–147.

Morris, S. and Strack, P. (2019). The Wald problem and the relation of sequential sampling and ex-ante information costs. Working paper.

Øksendal, B. (2003). *Stochastic differential equations: an introduction with applications*. Springer.

Peskir, G. and Shiryaev, A. (2006). *Optimal stopping and free-boundary problems*. Springer.

Pleskac, T. J. and Busemeyer, J. R. (2010). Two-stage dynamic signal detection: a theory of choice, decision time, and confidence. *Psychological Review*, 117(3):864–901.

Pouget, A., Drugowitsch, J., and Kepecs, A. (2016). Confidence and certainty: distinct probabilistic quantities for different goals. *Nature neuroscience*, 19(3):366–374.

Purcell, B. A. and Kiani, R. (2016). Hierarchical decision processes that operate over distinct timescales underlie choice and changes in strategy. *Proceedings of the National Academy of Sciences*, 113(31):E4531–E4540.

Rahnev, D., Balsdon, T., Charles, L., De Gardelle, V., Denison, R., Desender, K., Faivre, N., Filevich, E., Fleming, S., Jehee, J., et al. (2021). Consensus goals for the field of visual metacognition. https://doi.org/10.31234/osf.io/z8v5x.

Rahnev, D., Desender, K., Lee, A. L., Adler, W. T., Aguilar-Lleyda, D., Akdoğan, B., Arbuzova, P., Atlas, L. Y., Balcı, F., Bang, J. W., et al. (2020). The confidence database. *Nature human behaviour*, 4(3):317–325.

Sah, S., Moore, D. A., and MacCoun, R. J. (2013). Cheap talk and credibility: The consequences of confidence and accuracy on advisor credibility and persuasiveness. *Organizational Behavior and Human Decision Processes*, 121(2):246–255.

Schulz, L., Fleming, S. M., and Dayan, P. (2021). Metacognitive computations for information search: Confidence in control. *bioRxiv*.

Shea, N., Boldt, A., Bang, D., Yeung, N., Heyes, C., and Frith, C. D. (2014). Supra-personal cognitive control and metacognition. *Trends in cognitive sciences*, 18(4):186–193.

Shiryaev, A. N. (2007). *Optimal rules.* 2nd ed. Translated by A.B. Aries. New York: Springer. Originally published as *Optimal'nye pravila ostanovski.* (Moscow: Nauka, 1969).

Tajima, S., Drugowitsch, J., and Pouget, A. (2016). Optimal policy for value-based decision-making. *Nature communications*, 7(1):1–12.

Van den Berg, R., Zylberberg, A., Kiani, R., Shadlen, M. N., and Wolpert, D. M. (2016). Confidence is the bridge between multi-stage decisions. *Current Biology*, 26(23):3157–3168.

Vullioud, C., Clément, F., Scott-Phillips, T., and Mercier, H. (2017). Confidence as an expression of commitment: Why misplaced expressions of confidence backfire. *Evolution and Human Behavior*, 38(1):9–17.

Wald, A. (1947). *Sequential analysis.* New York: John Wiley & Sons.

Yeung, N. and Summerfield, C. (2012). Metacognition in human decision-making: confidence and error monitoring. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1594):1310–1321.

Yin, D., Mitra, S., and Zhang, H. (2016). Research note—When do consumers value positive vs. negative reviews? An empirical investigation of confirmation bias in online word of mouth. *Information Systems Research*, 27(1):131–144.

Zhitlukhin, M. V. and Muravlev, A. A. (2013). On Chernoff's hypotheses testing problem for the drift of a Brownian motion. *Theory of Probability & Its Applications*, 57(4):708–717.

Abstrakt

Studujeme informační disociace mezi rozhodnutími a důvěrou v rozhodnutí. Zkoumáme důsledky modelu duálního systému myšlení: rozhodovací systém a systém důvěry mají odlišné cíle, ale sdílejí přístup ke zdroji nepřesných a nákladných informací o proměnné, která je důležitá pro rozhodování. Rozhodovací systém má za cíl maximalizovat užitek, zatímco systém důvěry monitoruje rozhodovací systém a má za cíl poskytnout kvalitní zpětnou vazbu o správnosti rozhodnutí. V souladu se stávajícími experimentálními výsledky ukazujícími důležitost informací, které následují po rozhodnutí při vytváření důvěry, také dovolujeme systému důvěry, aby shromažďoval informace po rozhodnutí. Naším cílem je podložit proces získávaní informací po rozhodnutí (používaný v deskriptivních modelech důvěry) teorií optimálního učení. Nicméně zjišťujeme, že není vždy optimální pokračovat v získávání informací po rozhodnutí, a to ani pro daného jedince v daném rozhodovacím prostředí. Zejména, prostor pro získávání informací po rozhodnutí existuje jenom pro relativně rychlá rozhodnutí. Striktní rozlišování mezi jednostupňovou a dvoustupňovou teorií důvěry v rozhodnutí může být tedy zavádějící, protože oba způsoby formace důvěry v rozhodnutí se mohou projevit pod jedním mechanismem netriviálním způsobem.

Klíčová slova: rozhodování, důvěra, sekvenční výběr, optimální zastavení

Working Paper Series
ISSN 2788-0443

Individual researchers, as well as the on-line version of the CERGE-EI Working Papers (including their dissemination) were supported from institutional support RVO 67985998 from Economics Institute of the CAS, v. v. i.

Specific research support and/or other grants the researchers/publications benefited from are acknowledged at the beginning of the Paper.

Published by
Charles University, Center for Economic Research and Graduate Education (CERGE)
and
Economics Institute of the CAS, v. v. i. (EI)
CERGE-EI, Politických vězňů 7, 111 21 Prague 1, tel.: +420 224 005 153, Czech Republic.
Phone: + 420 224 005 153
Email: office@cerge-ei.cz
Web: https://www.cerge-ei.cz/

Editor: Byeongju Jeong

The paper is available online at https://www.cerge-ei.cz/working-papers/.

ISBN 978-80-7343-546-2 (Univerzita Karlova, Centrum pro ekonomický výzkum a doktorské studium)
ISBN 978-80-7344-657-4 (Národohospodářský ústav AV ČR, v. v. i.)