Optimal Sequential Decision with Limited Attention*

Yeon-Koo Che   Konrad Mierendorff

This draft: May 3, 2017
First Draft: April 18, 2016

Abstract

We consider a dynamic model of information acquisition. Before taking an action, a decision maker may direct her limited attention to collecting different types of evidence that support alternative actions. The optimal policy combines three strategies: (i) *immediate action,* (ii) a *contradictory strategy* seeking to challenge the prior belief, and (iii) a *confirmatory strategy* seeking to confirm the prior. The model produces a rich dynamic stochastic choice pattern as well as predictions in settings such as jury deliberation and political media choice.

KEYWORDS: Wald sequential decision problem, choice of information, contradictory and confirmatory learning strategies, limited attention.

1 Introduction

In many situations, decision makers (DMs) must choose actions whose payoffs are initially unknown. Before taking an action, they can acquire information or spend time processing information already available to them. Consider for example a firm or government agency that contemplates investing in a project or hiring an employee. Other examples include a prosecutor who decides whether to indict a suspect; a judge or a jury who deliberates on the guilt or innocence of the accused; voters who decide which candidate to vote for; or a researcher who seeks to ascertain a hypothesis. Acquiring more information allows the DM to make better decisions but can be costly and may delay the action. The DM

must therefore decide how much information she gathers, and typically this is not a one time decision. During the process of information acquisition, she can decide sequentially to continue acquiring information or to stop and take an action, depending on what she has learned so far. Seminal papers by Wald (1947) and Arrow, Blackwell, and Girshick (1949) have analyzed this stopping problem.

In many problems, a critical aspect of sequential decision making is not just *how much* information the DM acquires, but often more importantly *what kind* of information she seeks. For instance, a firm may seek evidence in favor of, or against investing in a given project. A judge or jury may scrutinize incriminating or exculpatory evidence. A researcher may set out to prove a hypothesis by way of a mathematical proof or empirical evidence, or attempt to disprove it via a counter-example or contradictory evidence. How decision makers direct their resources and attention to alternative types of evidence influences the information they receive and ultimately the quality of decisions they make.

The present paper focuses on this aspect of sequential decision problems. We consider a model with binary actions, a and b, which are optimal in states A and B, respectively. The state is initially unknown, and the DM has a prior belief $p_0 \in (0, 1)$ about the probability that the state is A. In continuous time, the DM may decide to acquire information, or to stop and take a final and irreversible action. Information acquisition has a direct flow cost, and/or payoffs are discounted exponentially.

The DM may seek different types of evidence that would reveal alternative states in varying degrees of accuracy.[1] In the baseline model evidence is fully conclusive. The DM may seek A-evidence, and if the state is A, she will find such evidence at Poisson rate λ. If the state is B, she will not discover any A-evidence. Alternatively, she may seek B-evidence, which arrives at Poisson rate λ if the state is B, and she will not discover B-evidence if the state is A. Since evidence never arrives for the wrong state, it is conclusive. It is thus optimal to stop learning and take the optimal action immediately after discovery.

To capture the idea that decision makers often have limited resources (cognitive, financial, time, equipment, manpower) that can be devoted to acquiring or processing information, we assume that the DM has a unit budget of "attention," which she must allocate between seeking A-evidence and B-evidence. Allocating less attention to one type of evidence will proportionally reduce the corresponding arrival rate. Being Bayesian, the DM also updates her belief in the absence of discovery. For example, seeking A-evidence but failing to find it makes her more pessimistic that the state is A.

Given this model, we characterize the optimal strategy, and provide comparative statics with respect to the information cost, discount rate, and payoff parameters. Our model yields rich predictions which we explore in two real-world settings: deliberation by grand and trial juries, and media choice by voters. Going beyond the baseline model, we ana-

---

[1]Seeking evidence could also mean introspective deliberation, i.e., scrutinizing and revisiting information that is already available.

lyze a model with non-conclusive evidence. We extend our main characterization to this general model and explore the implications for stochastic choice and response time.

Before discussing these results in greater detail, we describe the main characterization and provide an intuition. While our model allows for general strategies, we show that for each prior belief, the DM optimally uses one of three simple heuristics: (i) *immediate action*; (ii) *contradictory learning*; and (iii) *confirmatory learning*; and never switches between these different modes of learning.

Immediate action is the simple strategy where the DM takes an optimal action without acquiring any information. **Contradictory learning** involves seeking evidence that would confirm the state the DM finds relatively *unlikely*. An example is to seek B-evidence when state A is relatively likely. Seeking B-evidence but not discovering one makes the DM even more certain that the state is A. Ultimately, if no B-evidence arrives, she becomes so certain about the state being A that she chooses action a without further learning. To the extent that this is a likely event, a DM is effectively trying to "disprove" or "rule out" the unlikely state by playing the "devil's advocate" to her own mind.

Confirmatory learning involves seeking evidence that would confirm the state the DM finds relatively *more likely*. An example is to seek A-evidence when state A is relatively *likely*. When following this strategy, the DM becomes less confident about the likely state when no discovery occurs. Eventually, she becomes so uncertain that she switches to a second phase where she divides her attention seeking both A-evidence and B-evidence until she observes a signal that reveals the true state.

We characterize which strategy is optimal for each prior belief, and show how the structure of the optimal policy depends on the cost of information. Not surprisingly, if information is very costly, the DM takes an immediate action for all beliefs. For moderate information acquisition costs, we show that the DM optimally takes an immediate action when she is extremely certain, while she employs contradictory learning when she is more uncertain. Finally, if information acquisition costs are low, immediate action is again optimal for extreme priors, and contradictory learning is optimal for less extreme priors. For uncertain priors, however, confirmatory learning becomes optimal.

The intuition behind the optimal policy is explained by a trade-off between *accuracy* and *delay*. With an extreme belief, a fairly accurate decision can be made, and further information acquisition has a smaller benefit than the cost of delaying the action. Conversely, with a less extreme belief, information acquisition is more valuable. This explains why the *experimentation region* contains moderate beliefs and the *stopping region* is located at the extreme ends of the the belief space. The trade-off also explains which strategy is optimal inside the experimentation region. Confirmatory learning will lead to a fully accurate decision because the DM never takes an action before learning the state, but this could take a long time. By contrast, contradictory learning seeks evidence only for a limited duration. The DM is more likely to make a quick decision, but one that is

not as accurate when the DM mistakenly "rules out" the unlikely state. Again, the more certain the DM is, the less valuable the evidence is. This explains why the DM chooses contradictory learning when she is more certain and confirmatory learning when she is more uncertain. An implication is that *a "skeptic" is more likely to make an accurate decision with a longer delay than a "believer."*

Our model yields rich implications which we explore in two settings. First, it predicts distinct ways in which grand juries and trial juries deliberate: Trial juries adopt a high evidentiary standard for conviction and focus attention on incriminating evidence, whereas grand juries adopt a lower evidentiary standard for an indictment and focus attention on exculpatory evidence. We also extend the model to analyze how the possibility of a "hung jury" affects jury decisions.

Second, we derive implications of our model for the choice of news media by a voter. We show that optimal media choice leads to an "echo-chamber effect," where voters subscribe to media that are likely to push them in the direction of their prior belief. Interestingly, with sufficiently informative media, this effect is reversed for voters with moderate beliefs. They now optimally seek opposite-biased outlets, creating an "anti-echo chamber effect". We extend the model to allow for a trade-off between bias and informativeness in media choice and show that voters with more extreme beliefs value informativeness less than moderates.

Finally, we formulate a generalized model that allows for *non-conclusive* evidence. We focus on the case where the optimal strategy of the DM has the *single experimentation property* (SEP)—that is, she finds it optimal to take an action as soon as she observes either A- or B-evidence. We show that this is the case if the cost of information is sufficiently high. If SEP is satisfied, the optimal policy is composed of immediate action, contradictory learning, and confirmatory learning strategies as in the baseline model.[2] While the structure of the optimal policy is preserved, the possibility of non-conclusive evidence leads to richer implications due to the imperfect learning and permits an interesting comparison with other models of stochastic choice.

First, we show that a DM with a more uncertain prior—a "skeptic"—ends up making more accurate decisions but with a longer delay, compared with a DM with a more extreme prior—a "believer." Generally, the stochastic choice function in our model depends on the prior belief, which is not the case in classic random utility models of stochastic choice such as logit. This prior dependence is a common feature of models of information acquisition such as the rational inattention (RI) model, or drift diffusion models (DDM), which we discuss below. Unlike our model, however, these models predict that the *accuracy* of the decision is *independent* of the prior belief.[3]

---

[2]In Section 6 we go beyond SEP and construct examples where the DM optimally engages in repeated experimentation. These examples show that the central conclusion of the baseline model are robust to the introduction of non-conclusive evidence.

[3]Fudenberg, Strack, and Strzalecki (2017) show that extending the DDM to a rich state-space leads

Second, our model generates an interesting prediction on how the accuracy of the decision relates to *the response time*—that is, the amount of time it takes to reach that decision. If the DM obtains a signal, say in favor of A, after a long duration of trying, then it is relatively weak evidence in favor of A, compared to obtaining the same signal only after a short duration of trying. This means that for a given prior belief, a longer deliberation is associated with a less accurate decision, thus capturing a sense of *speed-accuracy complementarity* documented in cognitive psychology experiments.[4]

Related Literature. Wald's sequential decision problem has been formulated in continuous time as a drift-diffusion model (DDM) as well as with Poisson signals (see Peskir and Shiryaev, 2006, ch. VI, for rigorous treatments and references). In DDMs, the decision maker observes a Brownian motion whose drift is correlated with the true state. Our Poisson model assumes a learning technology in which the DM can seek "lumpy" evidence, which is applicable in situations where signals arrive rarely but reveal a lot of information. In contrast, DDMs are more suitable for situations where small amounts of information are revealed frequently.[5]

The Poisson signal structure builds on the exponential bandit model of Keller, Rady, and Cripps (2005). Closest to our baseline model is the planner's problem for negatively-correlated bandits in Klein and Rady (2011) and Francetich (2014). Note, however, that the separation of experimentation and exploitation in our model leads to fundamentally different behavior. In particular, the bandit model has no analogue of stopping and taking an irreversible action, so a distinct characterization emerges. For instance, there is no analogue of our "contradictory strategy."[6]

A recent incarnation of Wald's problem is Fudenberg, Strack, and Strzalecki (2017), who introduce rich states in DDM with two actions and obtain speed-accuracy complementarity. Moscarini and Smith (2001) endogenize the signal precision through costly effort decisions.[7] By contrast we focus on endogenous *types* of information. Related to this Nikandrova and Pancs (2017) have considered the problem of selectively learning about different investment projects. This paper uses Poisson signals as in our baseline model, but the payoffs of final actions are uncorrelated whereas our model assumes neg-

---

to prior dependence in the accuracy.

[4]Subjects of perceptual or choice experiments often exhibit speed-accuracy complementarity (see Ratcliff and McKoon (2008) for a survey). While our model produces this prediction, the extent to which our model, particularly the Poisson signal, describes adequately the neuro-physiology underpinning subjects' behavior in these experiments is unclear (see footnote 49).

[5]While different applications call for different assumptions on the signal structure, an interesting question is which type of signal a DM prefers is she can acquire information *flexibly,* and incurs a cost that depends on the informativeness of the signal structure. Zhong (2017) shows that for a large class of (flow) cost functions, in the continuous time limit, it is optimal to obtain Poisson signals.

[6]This is also the case in Damiano, Li, and Suen (2017) who add learning to a Poisson bandit model.

[7]Chan, Lizzeri, Suen, and Yariv (2016) analyze a stopping decision by a heterogeneous committee. Henry and Ottaviani (2017) turn Wald's problem into a persuasion game by splitting authority over final decisions and stopping between two players.

atively correlated payoffs.[8] In addition to the different payoff structure and applications, these papers do not consider a general model with noisy signals or explore the stochastic choice implications.[9]

Our model shares a common theme with the rational inattention model introduced by Sims (2003) in explaining individual choice as resulting from one's optimal allocation of limited attention over diverse information sources. (Matejka and McKay (2015) provide a discrete choice formulation comparable to our framework.) While RI abstracts from the dynamic process of deliberation, we attempt to unpack the "black box" and explicitly model a dynamic learning process that gives rise to the predicted choice outcome.[10]

The rest of the paper is organized as follows. Section 2 presents the baseline model and Section 3 characterizes the optimal policy. Section 4 applies the model to jury deliberation and media choice. Section 5 generalizes the model under SEP and explores implications on stochastic choice and response time. Section 6 treats the case of repeated experimentation. Section 7 concludes.

2 Baseline Model

We consider a DM who must take an action with unknown payoff. In the exposition, we will refer to three canonical examples: (i) a voter subscribing to news media; (ii) a jury deliberating on a verdict; and (iii) a scientist performing experiments on a hypothesis.

States, Actions and Payoffs. The DM must choose from two *actions* a and b whose payoffs depend on the unknown *state* $\omega \in \{A, B\}$. The *payoff* of taking action x in state ω is denoted by $u_x^\omega \in \mathbb{R}$. We label states and actions such that it is optimal to "match the state," and assume that the optimal action yields a positive payoff—that is, $u_a^A > \max\left\{0, u_b^A\right\}$ and $u_b^B > \max\left\{0, u_a^B\right\}$.[11] The DM may delay her action and acquire information. In this case, she incurs a flow cost of $c \geq 0$ per unit of time. In addition, her payoffs (and the flow cost) are discounted exponentially at rate $r \geq 0$.[12]

---

[8]See also the recent papers by Mayskaya (2016) and Ke and Villas-Boas (2016) which are concurrent to our paper.

[9]In a binary action model, Fudenberg, Strack, and Strzalecki (2017) also consider the allocation of (limited) attention between two Brownian motions, each indicating the (unknown) payoff of one action, and show that it is optimal to devote equal attention to each process. Woodford (2016) considers a DM optimizing on the types of evidence she seeks subject to entropy-based capacity constraint but in an exogenously specified DDM framework with exogenous boundaries. Meyer (1991) analyzes the sequential choice of bias in information acquisition, but in contrast to the literature following Wald, there is no stopping decision but an exogenous finite horizon.

[10]The dynamic RI model of Steiner, Stewart, and Matejka (2015) considers a sequence of multiple actions. Applied to a Wald setup with a single irreversible action, the predictions of the model are similar to the static case.

[11]Note that we allow for $u_x^A = u_x^B$ so that one x can a be *safe action*, but we rule out the trivial case that for $x \neq y$, $u_x^\omega \geq u_y^\omega$ in both states $\omega = A, B$.

[12]One of c and r may be zero but not both.

In the voter example, the actions a and b may correspond to voting for "right-leaning" and "left-leaning" candidates, respectively, and the state captures which candidate has a better platform. In the jury example, the actions correspond to "convict" (a) or "acquit" (b), and the state corresponds to the defendant's guilt (A) or innocence (B).

The DM's *belief* is denoted by the probability $p \in [0, 1]$ that the state is A. Her *prior belief* at time $t = 0$ is denoted by p_0. If the DM chooses her action optimally without information acquisition, then given belief p, she will realize an expected payoff of $U(p) := \max\{U_a(p), U_b(p)\}$, where $U_x(p) := pu_x^A + (1 - p)u_x^B$ is the expected payoff of taking action x. $U(p)$ takes a piece-wise linear form as depicted in Figure 2 on page 15, and we denote the belief where $U(p)$ has a kink by $\hat{p}$. We denote the optimal action by $x^*(p) \in \arg\max_{x \in \{a,b\}} U_x(p)$, which is unique if $p \neq \hat{p}$.

Information Acquisition and Attention. We model information acquisition in continuous time. At each point in time, the DM may allocate one unit of *attention* to seeking one of two types of evidence: *A-evidence* which reveals state A conclusively, and *B-evidence* which reveals state B conclusively. If the DM allocates a fraction $\alpha \in [0, 1]$ of her attention to seeking A-evidence, and the remaining fraction $\beta = 1 - \alpha$ to seeking B-evidence, then she receives A-evidence at the Poisson arrival rate of $\alpha\lambda$ in state A, and B-evidence at the Poisson arrival rate of $\beta\lambda$ in state B. We denote the DM's attention strategy by (α_t), and assume that α_t is a measurable function of t.[13]

A more concrete interpretation of the signal structure can be given depending on the context. In the voter example, α corresponds to particular a news medium (e.g., MSNBC or Fox), which publishes evidence in favor of either candidate or platform. In the absence of evidence each medium publishes partisan rhetoric corresponding to its bias. A jury deliberates on the evidence presented by the prosecutor and defense attorney, and α corresponds to the attention the jury devotes to finding incriminating evidence (as opposed to exculpatory evidence). In the scientist example, α corresponds to the nature of the experiment the scientist designs—that is, proving or disproving a hypothesis.

Bayesian Updating. Suppose the DM uses the attention strategy (α_t). Given her belief p_t, she observes signals confirming state A with Poisson rate $\alpha_t \lambda p_t$, and signals confirming state B with rate $\beta_t \lambda (1 - p_t)$. As long as she does not observe any signal, Bayes rule yields[14]

$$\dot{p}_t = -\lambda(\alpha_t - \beta_t)p_t(1 - p_t) = -\lambda(2\alpha_t - 1)p_t(1 - p_t). \tag{1}$$

---

[13]Given the linearity of the arrival rates, the DM cannot benefit from randomization. For this reason, we only consider deterministic strategies $(\alpha_t)_{t \in \mathbb{R}_+}$.

[14]By the martingale property we have $\lambda\alpha_t p_t dt + (1 - \lambda\alpha_t p_t dt - \lambda\beta_t (1 - p_t) dt) [p_t + \dot{p}_t dt] = p_t$. Dividing by dt and letting $dt \to 0$ yields (1).

For example, if the DM devotes all her attention to seeking A-evidence ($\alpha = 1$), she may not achieve discovery because the state is B, or because the state is A but no evidence has arrived. The longer she waits for a signal, the more convinced she becomes that the state is B. Finally, note that $\dot{p}_t = 0$ if $\alpha_t = 1/2$. That is, if the DM divides her attention between two types of evidence, she never updates her belief in case of no discovery.

The Decision Maker's Problem. The DM chooses an attention strategy (α_t) and a stopping time $T \geq 0$ at which a decision will be taken if no discovery has been made by then. Her problem is thus given by

$$
V(p_0) = \max_{(\alpha_\tau), T} \left\{ \begin{array}{c} \int_0^T e^{-rt} P_t(p_0, (\alpha_\tau)) \left[p_t \lambda \alpha_t u_a^A + (1-p_t) \lambda \beta_t u_b^B - c \right] dt \\ + e^{-rT} P_T(p_0, (\alpha_\tau)) U(p_T) \end{array} \right\}, \tag{2}
$$

where $\beta_t = 1 - \alpha_t$, and $P_t(p_0, (\alpha_\tau)) = p_0 e^{-\lambda \int_0^t \alpha_s ds} + (1-p_0) e^{-\lambda \int_0^t \beta_s ds}$ is the probability that no signal is received by time t given strategy (α_τ) and p_t satisfied (1). The first line of the objective function captures the flow cost which is incurred until the DM stops, and the payoffs from taking an action following discovery of evidence.[15] The second line accounts for the payoff from the optimal decision in case of no discovery by T.[16]

The Hamilton-Jacobi-Bellman (HJB) equation for this problem is

$$
c + rV(p) = \max_{\alpha \in [0,1]} \left\{ \begin{array}{c} \lambda \alpha p \left(u_a^A - V(p) \right) + \lambda(1-\alpha)(1-p) \left(u_b^B - V(p) \right) \\ - \lambda(2\alpha - 1)p(1-p)V'(p) \end{array} \right\}, \tag{3}
$$

for p such that $V(p) > U(p)$. If the LHS of (3) is larger than the RHS we must have $V(p) = U(p)$—that is, $T(p) = 0$ and immediate action is optimal. Since the problem is autonomous, the optimal policy and the stopping decision at time t depends only on the current belief p_t. Note also that the objective in (3) is linear in α, which implies that the optimal policy is a bang-bang solution: the optimal attention strategy must satisfy $\alpha^*(p) \in \{0, 1\}$, except when the derivative of the objective vanishes.

Outlook for Generalization. In Section 5 we generalize the model by relaxing the assumption that signals are perfectly revealing of the state. While this more general model yields richer predictions and implications, the qualitative nature of the optimal policy remains the same as long as the signal is sufficiently accurate. Hence, we focus the exposition primarily on the baseline model. In Section 5 we describe the general model in more detail and explore its implications for the optimal stochastic choice rule.

---

[15]Specifically, at each time t, conditional on no discovery so far, A-evidence is discovered at the instantaneous rate of $p_t \lambda \alpha_t$, B-evidence is discovered at the instantaneous rate of $(1-p_t)\lambda \beta_t$, and the flow cost c is always incurred.

[16]For a given $(\alpha_t)_{t \in \mathbb{R}_+}$, conditional on no discovery, posterior belief evolves according to a deterministic rule (1). Since stopping matters only when there is no discovery, it is without loss to focus on a *deterministic* stopping time T.

(a) confirmatory strategy

(b) contradictory strategy

(c) combination of contradictory and confirmatory
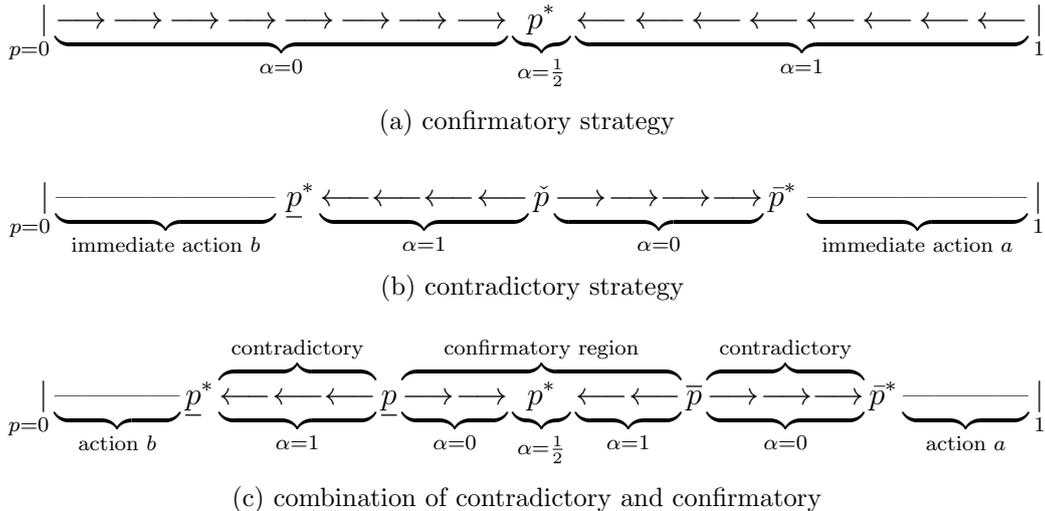
Figure 1: Structure of Heuristic Strategies and Optimal Solution

3 Optimal Strategy in the Baseline Model

We begin with a description of several intuitive learning heuristics that the DM could employ. These learning heuristics form basic building blocks for the DM's optimal strategy. The details of the formal construction are presented in Appendix A.

3.1 Learning Heuristics

Immediate action (without learning). A simple strategy the DM can choose is to take an immediate action and realize $U(p)$ without any information acquisition. Since information acquisition is costly, this can be optimal if the DM is sufficiently confident in her belief—that is, if p is either sufficiently high or sufficiently low.

Confirmatory learning. When the DM decides to experiment, one natural strategy is to seek evidence that "confirms" her current belief. Formally, a confirmatory learning strategy prescribes:

$$\alpha(p) = \begin{cases} 0 & \text{if } p < p^* \\ \frac{1}{2} & \text{if } p = p^* \\ 1 & \text{if } p > p^*, \end{cases}$$

for some the reference belief $p^* \in (0, 1)$, which will be chosen optimally. For beliefs above p^*, the DM seeks A-evidence. Receiving A-evidence confirms her belief that A is relatively likely and moves it to $p = 1$. For beliefs below p^*, she seeks B-evidence, which moves the belief to $p = 0$. In the absence of discovery, the DM becomes more uncertain relative to her current belief. As is clear from (1), her belief drifts from both directions towards the absorbing point p^*. This is illustrated in Panel (a) of Figure 1, where arrows indicate the direction of Bayesian updating.

9

Once the reference belief p^* is reached, the DM divides her attention equally between seeking both types of evidence. In this case, no further updating occurs, and she repeats the same strategy until she obtains evidence that confirms the true state of the world.

One clear advantage of the confirmatory strategy is that it will never lead to a mistake in the action chosen by the DM. On the other hand, since the DM always waits for evidence before taking an action, the confirmatory strategy involves a potentially long delay, which is costly.

Contradictory learning. Alternatively, the DM could try to obtain evidence that "contradicts" her current belief. Formally, contradictory learning prescribes

$$\alpha(p) = \begin{cases} 1 & \text{if } p \in \left(\underline{p}^*, \check{p}\right), \\ 0 & \text{if } p \in [\check{p}, \overline{p}^*), \end{cases} \tag{4}$$

for some reference belief $\check{p}$ and boundaries of the *experimentation region* $\left(\underline{p}^*, \overline{p}^*\right)$, which will each be chosen optimally. Within the experimentation region, the DM seeks B-evidence if she believes state A to be relatively likely, and A-evidence if state B is relatively likely. In the absence of a signal, the DM's belief drifts outward in the direction confirming her original belief. In effect, the DM is gradually ruling out the unlikely, if she seeks but fails to obtain contradictory evidence. Eventually, the DM's belief will reach one of the boundary points $\underline{p}^*$ or $\overline{p}^*$, at which she is sufficiently certain to justify an immediate action without conclusive evidence. The belief updating is illustrated in Panel (b) of Figure 1.

The contradictory strategy may seem counter-intuitive, since the DM is looking for evidence for the unlikely state, which is less likely to arrive. Its value is that it generates a "surprise" which changes the DM's optimal action. In the absence of a surprise the DM can "rule out" the unlikely, and will reach a quick decision for the optimal action in the likely state.

3.2 Optimal Strategy

The structure of the optimal policy depends on the *cost of information* c. Intuitively the higher the flow cost, the lower is the value of experimenting. As will be seen, the experimentation region expands as the cost of information falls. More interestingly, the type of learning strategy employed also changes in a nontrivial way. The following characterization shows that there are three cases. If c is very high, immediate action is always optimal. For intermediate values of c, the optimal strategy involves only contradictory learning, and for low values of c, both contradictory and confirmatory learning occur.

Theorem 1. *For given utilities u_x^ω, $\lambda > 0$, and $r \in [0, \infty)$, there exist $\overline{c} = \overline{c}(r, u_x^\omega, \lambda)$ and $\underline{c} = \underline{c}(r, u_x^\omega, \lambda)$, $\overline{c} \geq \underline{c} \geq 0$, with strict inequalities for r sufficiently small, such that the*

optimal strategy is characterized as follows:[17]

  (a) (**No learning**) *If* $c \geq \bar{c}$*, the DM takes action* $x^*(p)$ *without any information acquisition.*

  (b) (**Contradictory learning**) *If* $c \in [\underline{c}, \bar{c})$*, there exist* $0 < \underline{p}^* < \check{p} < \overline{p}^* < 1$ *such that for* $p \in (\underline{p}^*, \overline{p}^*)$*, the optimal* α *is given by* (4)*. If* $p \notin (\underline{p}^*, \overline{p}^*)$*, the DM takes action* $x^*(p)$ *without any information acquisition.*

  (c) (**Contradictory and Confirmatory learning**) *If* $c < \underline{c}$*, then there exist* $0 < \underline{p}^* < \underline{p} < p^* < \overline{p} < \overline{p}^* < 1$ *such that for* $p \in (\underline{p}^*, \overline{p}^*)$*, the optimal* α *is given by*

$$\alpha(p) = \begin{cases} 1, & \text{if } p \in \left(\underline{p}^*, \underline{p}\right), \\ 0, & \text{if } p \in \left(\underline{p}, p^*\right), \\ \frac{1}{2} & \text{if } p = p^*, \\ 1, & \text{if } p \in \left(p^*, \overline{p}\right), \\ 0, & \text{if } p \in \left(\overline{p}, \overline{p}^*\right). \end{cases} \tag{5}$$

If $p \notin (\underline{p}^*, \overline{p}^*)$ *the DM takes action* $x^*(p)$ *without any information acquisition.*

We sketch the main steps of the proof in Section 3.4 below. The formal proof can be found in Appendix A.

First, we describe the structure of the optimal strategy and discuss the intuition. The optimal policy in case (b) is depicted in Panel (b) of Figure 1. In case (c) the pattern is more complex (see Panel (c) of Figure 1). Both confirmatory and contradictory learning are optimal for some beliefs. Theorem 1 also states that the confirmatory region $(\underline{p}, \overline{p})$ is always sandwiched between two regions where the contradictory strategy is employed. That is, near the boundaries of the experimentation region, the contradictory strategy is always optimal.

The intuition for the optimal strategy is explained as follows. First, since learning is costly, there are levels of confidence, given by $\overline{p}^*$ and $\underline{p}^*$, that the DM finds sufficient for making decisions without any evidence. These beliefs constitute the boundaries of the experimentation region; more extreme beliefs result in an immediate action.

Within the experimentation region $(\underline{p}^*, \overline{p}^*)$, the DM's choice depends on the trade-off between the confirmatory and contradictory strategies. The confirmatory strategy has the advantage that the DM will eventually discover evidence, and will thus never make a mistake.[18,19] At the same time, full discovery of evidence could lead to a long delay.

---

[17]See (24) and (25) in Appendix A for explicit expressions for $\underline{c}$ and $\bar{c}$.

[18]In the generalized model with noise, mistakes are possible but we will see that the confirmatory strategy still leads to more accurate decisions.

[19]If the DM would attempt full discovery with the contradictory strategy by setting, e.g., $\overline{p}^* = 1$ when $p_0 > \hat{p}$, she would only look for evidence for the unlikely state B. Her belief would reach, $p_t = 1$ only as $t \to \infty$, so she would never take action a in finite time. She would avoid a mistake only at the cost of inaction.

This is particularly the case when the DM is fairly certain in her belief. In that case, the contradictory strategy becomes relatively more appealing. Suppose for instance the DM's belief $p < \overline{p}^*$ is very close to $\overline{p}^*$. In that case, the contradictory strategy either yields a "surprise" (B-evidence), leading to a perfectly accurate decision b, or (more likely) allows the DM to "rule out" the unlikely state and reach the desired level of confidence $\overline{p}^*$ for action a with very little delay. Of course, the DM could instead look for A-evidence only briefly and take an action if no information arrives, but obtaining A-evidence has no value since it does not change her action. The only sensible alternative is to "go all the way" toward full discovery—i.e., the confirmatory strategy, but it takes a long time. Hence, the contradictory strategy is optimal near the stopping boundaries. As the DM becomes less certain, however, the trade-off tilts toward the confirmatory strategy, as the contradictory strategy involves a longer delay to reach the stopping region.

3.3 Comparative Statics

It is instructive to study how the optimal strategy varies with the parameters. Of particular interest are the experimentation region $(\underline{p}^*, \overline{p}^*)$, and the confirmatory region $(\underline{p}, \overline{p})$. We say a region *expands* when a parameter change leads to a superset of the original region. (This includes the case that the region appears when it was empty before.) We say a region *shifts up (down)* when the boundaries of the region increase (decrease) strictly.

Proposition 1 (Comparative statics). *(a) The experimentation region expands as r or c falls, and covers $(0,1)$ in the limit as $(r,c) \to (0,0)$ or as $(u_a^B, u_b^A) \to (-\infty, -\infty)$.*

(b) The experimentation region expands as u_a^B or u_b^A falls, and covers $(0,1)$ in the limit as $(u_a^B, u_b^A) \to (-\infty, -\infty)$. Specifically, if $c < \overline{c}$, then $\underline{p}^* \searrow 0$ monotonically as $u_b^A \to -\infty$, and $\overline{p}^* \nearrow 1$ monotonically as $u_a^B \to -\infty$.

(c) If $c < \overline{c}$, then the experimentation region shifts down as u_a^A increases and up as u_b^B increases.

(d) If $\underline{c} < c < \overline{c}$, then $\check{p}$, the cutoff in the contradictory strategy, increases in u_b^A and decreases u_a^B.

(e) Suppose $c < \overline{c}$. The confirmatory region expands as (u_a^B, u_b^A) falls and covers $(0,1)$ in the limit as $(u_a^B, u_b^A) \to (-\infty, -\infty)$. Specifically, if $c < \underline{c}$, then $\underline{p} \searrow 0$ monotonically as $u_b^A \to -\infty$, and $\overline{p} \nearrow 1$ monotonically as $u_a^B \to -\infty$.

The proof can be found in Appendix in the Supplemental Material.

Parts (a) and (b) are quite intuitive. The DM acquires information for a wider range of beliefs if the cost of learning (r,c) falls, or if mistakes become more costly in the sense that (u_a^B, u_b^A) falls. A similar intuition holds for (c). The intuition for (d) is more subtle. For example, if $p < \check{p}$, the contradictory strategy may lead to taking action b in the wrong state (A). If this becomes more costly, $\check{p}$ shifts down to avoid this mistake.

Part (e) shows that the cost of mistakes also matters for the relative appeal of the *alternative* learning strategies: the confirmatory strategy becomes a more attractive mode of learning for the DM when mistakes become more costly. In the limit where mistakes become completely unacceptable, the confirmatory strategy becomes optimal for all beliefs. One could imagine this limit behavior as that of a scientist who views collecting evidence of either kind—proving or disproving a hypothesis—as the only acceptable way of advancing science. Such a scientist will rely solely on the confirmatory strategy: she will initially strive to prove a hypothesis for instance if she conjectures it to be true (believes it to be more likely true than not); after a series of unsuccessful attempts to prove the hypothesis, however, she begins to doubt her initial conjecture, and when the doubt reaches a "boiling point" (i.e., p^*), she begins to put some effort to disprove it.[20]

The Role of Discounting. Intuitively, one would interpret r as a cost of learning and would thus expect that c and r are substitutes in the sense that a higher discount rate requires a lower flow cost for the same structure to emerge. Formally, $\partial \overline{c}/\partial r < 0$ and $\partial \underline{c}/\partial r < 0$. The following proposition shows that this is indeed the case if at least one loss payoff (u_b^A or u_a^B) is not too small.

Proposition 2. *(a) $\partial \overline{c}/\partial r < 0$ iff $U(\hat{p}) > 0$, which is equivalent to $u_a^A u_b^B - u_b^A u_a^B > 0$.*
  (b) $\partial \underline{c}/\partial r < 0$ if both $u_a^A > \left| u_b^A \right|$ and $u_b^B > \left| u_a^B \right|$; $\partial \underline{c}/\partial r > 0$ if $\min \left\{ u_b^A, u_a^B \right\}$ is sufficiently small.

The proof can be found in Appendix A.3 in the Supplemental Material.

If both loss payoffs u_b^A and u_a^B are negative and sufficiently large in absolute value, we have $\partial \overline{c}/\partial r > 0$ and $\partial \underline{c}/\partial r > 0$. A higher discount rate calls for more experimentation in this case. Intuitively, if losses are sufficiently large, the DM would prefer to delay their realization, which favors longer experimentation. This explains Part (a) of Proposition 2. Moreover, large losses in case of a mistake increase the need for accuracy, favoring the confirmatory strategy. This explains Part (b) of Proposition 2.

3.4 Sketch of the Proof

We now sketch the main arguments leading to Theorem 1 in several steps. A less technically interested reader may want to skip this section.

First, we compute the value of alternative learning strategies. Taking an action immediately simply yields $U(p)$. The values of the confirmatory and contradictory strategies are given by particular solutions of two ODEs obtained by substituting $\alpha = 0$ and $\alpha = 1$

---

[20]Contradictory learning could also describe some aspect of scientific inquiry if a scientist is willing to accept a small margin of error. For instance, even a careful theorist may not verify thoroughly her "proof" if she believes it to be correct. Rather, she may look for a mistake in her argument—a contradictory strategy, and without finding one, may declare it as a correct proof.

in the HJB equation (3). For given boundary conditions $(\tilde{p}, W)$, where $\tilde{p} \in (0,1)$, the ODEs admit unique solutions $V_0(p; \tilde{p}, W)$ and $V_1(p; \tilde{p}, W)$, respectively.

To compute the value of the confirmatory strategy, we start with an arbitrary $p^* \in (0,1)$, which will then be determined optimally. At p^*, the strategy prescribes $\alpha = 1/2$ until evidence arrives, whereupon the DM takes an action according to the evidence. Let $\overline{U}(p)$ denote the value of this "stationary" strategy.[21] We can then use $V(p^*) = \overline{U}(p^*)$ as a boundary condition for the value function and obtain

$$V_{cf}(p) := \begin{cases} V_0(p; p^*, \overline{U}(p^*)), & \text{for } p \leq p^*, \\ V_1(p; p^*, \overline{U}(p^*)), & \text{for } p \geq p^*. \end{cases}$$

We show that $V_{cf}(p^*) = \overline{U}(p^*)$ implies that $V_{cf}(p)$ is differentiable at p^* (Lemma 2, all Lemmas can be found in Appendix A). To pin down p^*, we require that $V_{cf}(p) \geq \overline{U}(p)$ for all p in a neighborhood of p^*—an obvious necessary condition for the confirmatory strategy to be part of the optimal strategy. With differentiability at p^*, this yields $V'_{cf}(p^*) = \overline{U}'(p^*)$. Together with $V(p^*) = \overline{U}(p^*)$, this uniquely pins down the reference belief as $p^* = \left(ru_b^B + c\right) / \left(ru_a^A + ru_b^B + 2c\right)$.

The value of contradictory learning is computed in a similar way. Intuitively, $\overline{p}^*$ and $\underline{p}^*$ are the beliefs at which the DM is indifferent between an optimal immediate action and contradictory learning for an instant followed by an immediate action in case of no discovery. This yields the boundary conditions $V(\underline{p}^*) = U(\underline{p}^*)$ and $V(\overline{p}^*) = U(\overline{p}^*)$. Next, we postulate that the slope of the value is equal to $U'(p)$ at these values (smooth pasting). Combining these two conditions pins down the critical beliefs $\overline{p}^*$ and $\underline{p}^*$.[22] We can thus construct the *left branch* and *right branch* of the value function for the contradictory strategy:

$$\underline{V}_{ct}(p) := \begin{cases} U(p), & \text{if } p \leq \underline{p}^*, \\ V_1(p; \underline{p}^*, U(\underline{p}^*)), & \text{if } p > \underline{p}^*; \end{cases} \quad \overline{V}_{ct}(p) := \begin{cases} U(p), & \text{if } p \geq \overline{p}^*, \\ V_0(p; \overline{p}^*, U(\overline{p}^*)), & \text{for } p < \overline{p}^*. \end{cases}$$

We combine these branches and construct the value of the contradictory learning[23]

$$V_{ct}(p) := \max\left\{\underline{V}_{ct}(p), \overline{V}_{ct}(p)\right\}.$$

Second, we consider a candidate solution $V^*(p) := \max\{V_{ct}(p), V_{cf}(p)\}$ which is obtained if the DM chooses optimally between confirmatory and contradictory learning based on the prior belief, and never switches. We show that this yields the strategy stated in Theorem 1. Two observations are crucial. First, consider a hypothetical "full attention

---

[21]We have $\overline{U}(p) = \frac{\lambda}{2r+\lambda}\left(pu_a^A + (1-p)u_b^B\right) - \frac{2c}{2r+\lambda}$. Intuitively, without mistakes, the DM achieves the "first-best" payoff $pu_a^A + (1-p)u_b^B$, but discounted and net off the expected discounted flow cost.

[22]Explicit expressions can be found in Appendix A.1.2.

[23]We discuss in Appendix A.1.2 why it is appropriate take the maximum.

	U(p)
	$V_{ct}(p)$
	$V_{cf}(p)$
	$\alpha(p)$
	$\overline{U}(p)$

(a) $c = .3$ (only contradictory)　　(b) $c = .15$ (confirmatory and contradictory)
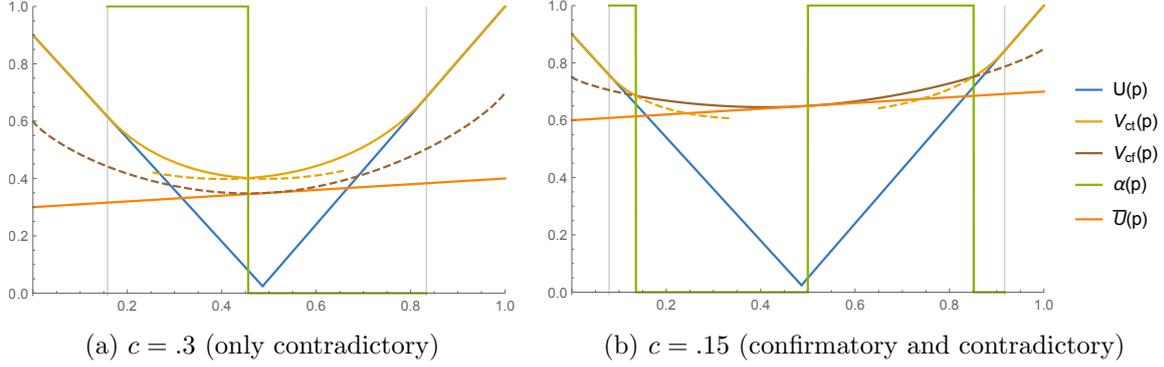
Figure 2: Value Function and Optimal Policy. The value function is the upper envelope of V_{ct} and V_{cf} (solid). ($\lambda = 1$, $r = 0$, $u_a^A = 1$, $u_b^B = .9$, $u_b^A = u_a^B = -.9$)

payoff"—the payoff the DM would attain if she could set $\alpha = \beta = 1$. We show that contradictory learning achieves this upper bound at the stopping boundaries $\overline{p}^*$ and $\underline{p}^*$, while confirmatory learning achieves a strictly lower value (Lemmas 4 and 6). This implies that contradictory learning is always part of the optimal strategy since it dominates confirmatory learning at the stopping boundaries $\overline{p}^*$ and $\underline{p}^*$. Second, we establish a *Crossing Property* (Lemma 2): a solution $V_1(p; x, W)$ can cross $V_0(p; x', W')$ only from above if they exceed $\overline{U}(p)$. This means that branches of the confirmatory and contradictory value function must intersect in the way illustrated in Figure 2 if they are undominated by $\overline{U}(p)$.

In particular, this implies that the choice between confirmatory and contradictory learning leads to one of the structures in Panels (b) and (c) of Figure 1.[24] Case (b) of Theorem 1 arises if $V_{ct}(p^*) \geq \overline{U}(p^*)$ and case (c) arises in the opposite case.

Finally, we show that no other strategies are part of the optimal strategy by verifying that it satisfies the HJB equation (3). Since $V^*(p)$ is not everywhere differentiable, we adopt a weaker solution concept and show that $V^*(p)$ is a *viscosity solution* of the HJB equation. By Theorem III.4.11 in Bardi and Capuzzo-Dolcetta (1997), this implies that it is the value function of the DM's problem.

4　Applications

The model of information acquisition with limited attention can be applied in various situations. We discuss two applications: jury deliberation and media choice by voters. For each we also discuss generalizations for that go beyond the baseline model: In the case of jury deliberation we consider the inclusion of a third action (a "hung jury") and the case of media choice, we consider news media with varying degrees of informativeness. The formal treatment of the extensions can be found in Appendices C and D in the Supplemental Material.

---

[24]Note that these structures indeed have the property that the DM never switches the mode of learning.

15

4.1 Jury Deliberation

In many states of the USA, when a prosecutor accuses an individual of a crime, a grand jury is impaneled to decide whether to indict the accused. Should that occur, a trial jury hears the case presented by the prosecutor and the defense attorney before returning its verdict. Both types of juries deliberate based on the evidence or lack thereof over a period of time. Our model can be used to consider their deliberation behavior and its implications on the outcome.

Suppose that the DM is a member of jury deciding whether to "indict/convict" (action a) or "not indict/acquit" (action b) a defendant.[25] The state is either "guilty" ($\omega = A$) or "innocent" ($\omega = B$). Suppose the jury has already heard all the evidence presented in the courtroom,[26] and has formed a preliminary opinion summarized by the prior belief p_0. The jury proceeds with deliberation which involves revisiting the evidence, testimonies, and arguments presented to them and scrutinizing some details jurors may have missed (i.e., juries often ask for transcripts of testimonies they wish to scrutinize). The attention decision α corresponds to the type of evidence/testimony that the jury may scrutinize in depth.

One salient difference between the two juries is the costs of making alternative errors: while a trial jury perceives a high cost of convicting an innocent (i.e., low u_a^B), a grand jury must be more concerned about failing to indict a guilty defendant (i.e., a low value of u_b^A). The evidentiary standards as well as the decision rule employed by US courts reflect this difference.[27] Proposition 1 predicts how this difference causes grand juries and trial juries to deliberate differently, and shows that the de facto evidentiary standards—reflected by the boundaries $\underline{p}^*$ and $\overline{p}^*$—differ for the two juries.

For ease of exposition we use examples to illustrate our findings which which hold generally. Panels (a) and (b) of Figure 3 depict examples where the parameters are chosen so that Part (b) of Theorem 1 obtains.

For the grand jury in panel (a), probable cause corresponds to $p > \overline{p}^* = .8$ in the example. If the prior falls in this range, the jury returns an indictment right away. Conversely, the accused is acquitted immediately only if $p_0 < \underline{p}^* = .1$. This shows that the stronger concern about acquitting a guilty person leads the grand jury to require a higher standard for this decision. If the grand jury does not return a decision immediately,

---

[25]Here, we view the DM as a single juror, or a coalition of jurors acting like a single individual. The conclusions would also apply for a single judge. In practice, the collective action aspect of jury deliberation brings another layer to the problem, which is beyond the scope of this paper. See Persico (2004), Stephenson (2010), and Chan, Lizzeri, Suen, and Yariv (2016). Our application to jury deliberation is a first step toward exploring information choice in the jury context.

[26]Only the prosecutor presents the case in grand jury proceedings, whereas in a trial, both prosecutor and defense attorney present the evidence.

[27]In a criminal case, conviction in a criminal court requires proving guilt "beyond the reasonable doubt" whereas the standard for grand jury indictment is "probable cause." The decision rule for trial jury is unanimity, whereas a grand jury indictment requires concurrence by 12 members out of 16-23 total members.

(a) Grand Jury ($u_b^A = -1$, $u_a^B = 0$) (b) Trial Jury ($u_b^A = 0$, $u_a^B = -1$) (c) Trial jury with hung jury ($\overline{u}_c = .7$)
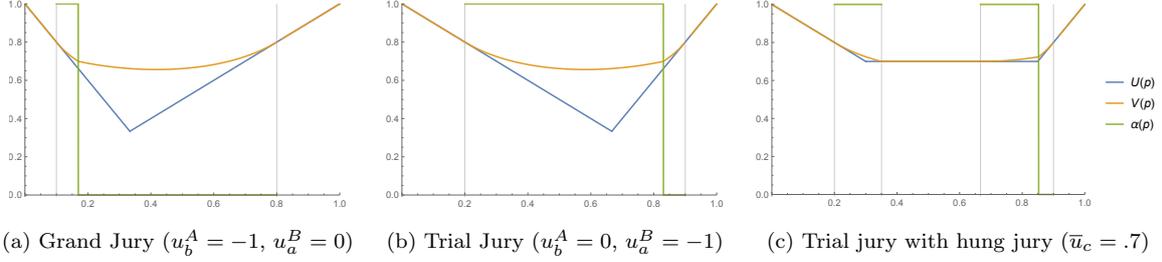
Figure 3: Deliberation of grand and trial juries. ($\lambda = 1$, $r = 0$, $c = .2$, $u_a^A = u_b^B = 1$.)

for most beliefs ($p \in (.17, .8)$ in the example), the jury will look for exculpatory evidence.[28] If no such evidence is found, the jury becomes more convinced that there is "probable cause," and ultimately returns an indictment.

The trial jury deliberates differently. The evidentiary standard for conviction is higher ($\overline{p}^* = .9$ in the example), whereas the standard for acquittal is lower ($\underline{p}^* = .2$), reflecting the greater concern about convicting an innocent person. Moreover, if no immediate verdict is returned, for most beliefs ($p \in (.2, .83)$ in the example), the jury scrutinizes incriminating evidence. Not finding such evidence pushes the jury's belief toward "reasonable doubt" and an acquittal.

In many cases, unanimity is required for a jury decision. If no consensus is reached, a *hung jury* arises. In the US this leads to a mistrial with the possibility to retry the case. Although a consequence of unanimity rule adopted by Common Law courts for criminal cases, in principle a mistrial need not be the only solution to a jury deadlock,[29] and its effect on jury deliberation is an important question. To analyze the effect, we introduce hung jury in "reduced form" as a third action c, which we assume to have a safe payoff of $u_c^A = u_c^B = \overline{u}_c$.[30] Panel (c) of Figure 3 depicts this case for a trial jury. The effects of introducing the option of a hung jury can be summarized as follows.

First, the possibility of a hung jury does not change the evidentiary standard for either verdict: they remain $\underline{p}^* = .2$ for acquittal and $\overline{p}^* = .9$ for conviction. The reason is that these standards are set by indifference between contradictory learning and immediate action, and the values of these strategies are not affected by the hung jury option.

Second, the hung jury option affects the deliberation strategy in favor of seeking incriminating evidence (B-evidence). The value of seeking incriminating evidence increases since the jury faces a more appealing option, namely to "settle" for mistrial instead of deliberating until they reach the reasonable doubt. This means that for some pre-deliberation belief ($p_0 \in [.83, .85]$), the jury switches attention from exculpatory evidence

---

[28]If $u_a^A = u_b^B$ and $u_a^B > u_b^A$, then $\underline{p}^* < 1 - \overline{p}^*$ by part (b), and $\check{p} < 1/2$ by part (d) of Proposition 1.

[29]For instance, in England and New Zealand, the unanimity requirement is relaxed in the event of a jury deadlock, allowing a judge to require supermajority instead. Hung jury is not possible in a Scottish court where a simple majority is required for a verdict. See https://en.wikipedia.org/wiki/Hung_jury.

[30]An interpretation is that this is a continuation value jurors assign to a mistrial. We formally analyze the model with a third action in Appendix C in the Supplemental Material.

to incriminating evidence. The flip side of this behavior is that for a large range of pre-deliberation beliefs ($p_0 \in [.35, .67]$) mistrial is declared.

Third, the change in deliberation behavior also affects verdicts. For $p_0 \in [.67, .85)$, the probability of guilty verdict falls.[31] Remarkably, the probability of "not-guilty" verdict also falls. In fact, with the mistrial option, the jury never returns a "not-guilty" verdict for any pre-deliberation belief $p_0 > 0.35$ above the lower bound of the mistrial region, whereas, without that option, the jury would have returned "not guilty" verdict with positive probability after a long deliberation (and reached $\underline{p}^*$).

These findings suggest that the possibility of hung jury has complex and nontrivial effects on jury deliberation and verdicts and also point to the richness of the prediction that can be brought out by use of a dynamic model such as the current one.[32]

4.2 Choice of News Media by Voters

Our second application is the choice of news media by a voter.[33] Suppose a citizen votes for one of two candidates a and b. Candidate a has a right-wing platform, and candidate b has a left-wing platform.[34] The state $\omega \in \{A, B\}$ captures which platform serves the interest of the voter better, and u_x^ω represents the voter's utility from voting for x in state ω. Her belief about the state is captured by p. We say the voter is more right-leaning the higher p is.[35]

Before voting, the voter may obtain information from a range of "news media" or "news channels," at a flow cost $c > 0$.[36] We now interpret each choice $\alpha \in [0, 1]$ as a

---

[31]Fix any belief $p_0 \in [0.83, 0.85]$ in the switching region. Without hung jury, the probability of a guilty verdict exceeds p_0 since the jury never finds the guilty innocent whereas it may find the innocent guilty (in case the jury reaches $\overline{p}^*$ without finding exculpatory evidence). With hung jury, the probability of guilty verdict falls strictly below p_0 by an analogous argument. For $p_0 \in [0.67, 0.83]$, mistrial is declared in any event in which, without hung jury, the jury would have found incriminating evidence after p fell below 0.67. Hence, the probability of a guilty verdict is reduced commensurately.

[32]The analysis so far focused on the situation in which, without hung jury, jury deliberation involves only contradictory learning. The case with both confirmatory and contradictory learning is analogous: introduction of mistrial increases the value of confirmatory learning, which reduces the probability of the guilty verdict for high p (around $\overline{p}$) and that of the not guilty verdict for a low p (around $\underline{p}$), at the expense of increased probability of mistrial.

[33]This application shares a theme with Suen (2004), Burke (2008), Oliveros (2015), and Yoon (2016)—particularly the optimality of consuming a biased medium for a Bayesian agent. These papers are largely static, unlike the current model which is fully dynamic. The specific signal structure differs also. Mullainathan and Shleifer (2005) assumes "behavioral" bias on the part of consumers to predict media slanting. Gentzkow and Shapiro (2008) and Perego and Yuksel (2016) study media competition, which we take as given.

[34]In the context of a referendum, one could interpret $x = a$ as voting for the status quo and $x = b$ for an alternative, or vice versa.

[35]Alternatively one could model a voter's bias in terms of her payoffs. In this case u_x^ω could incorporate her "partisan" preference. This approach is left for future research.

[36]Since the time at which the payoff is realized (the election or implementation of a policy) is independent of when the voter makes up her mind, we assume $r = 0$. Our model does not capture the effect of a "deadline," which is clearly relevant for the election example. We conjecture that the salient features of our characterization which are discussed in the context of this application carry over to a model with

|	state A	rhetoric (A-favoring)		state A	rhetoric (B-favoring)	facts (A-favoring)
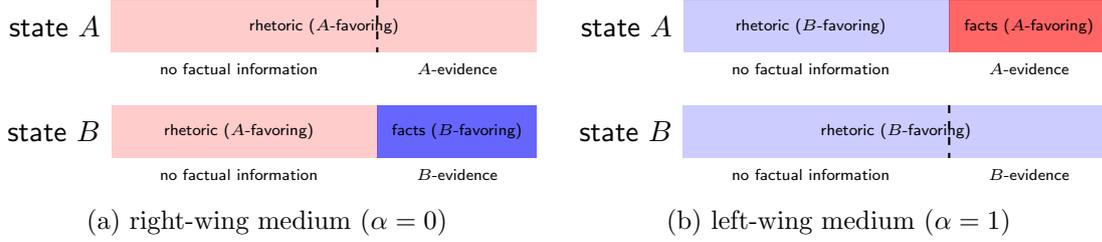
Figure 4: Media bias: Labels below the bars describe the information the medium has learned. Colors and labels inside the bars describe informational content (the message) published in each case.

(a) right-wing medium ($\alpha = 0$)   (b) left-wing medium ($\alpha = 1$)

particular news medium the voter may subscribe to.[37] Medium α publishes three possible types of content: (i) factual information in favor of A, (ii) factual information in favor of B, (iii) partisan rhetoric that corresponds to the bias of a medium. Over a time interval $dt > 0$, these messages are sent with probabilities (i) $p\alpha\lambda dt$, (ii) $(1-p)(1-\alpha)\lambda dt$, and (iii) $1 - [p\alpha + (1-p)(1-\alpha)]\lambda dt$, respectively.

The informational content of partisan rhetoric—and thus the partisan bias—depends on the medium α. To illustrate this, consider the most extreme media. As depicted in Panel (a) of Figure 4, medium $\alpha = 0$ sends B messages in state B only when it learns facts supporting B, which happens with *some* probability. Otherwise it sends partisan rhetoric favoring the right-wing candidate a both in state B and in state A.[38] We therefore call medium $\alpha = 0$ *right-wing*. Intuitively, a right-wing medium will advocate the left-wing candidate b only if it possesses hard facts supporting state B. On the other hand, it is willing to advocate the right-wing candidate even in the absence of factual information. Therefore, even if it possesses factual information in favor of a, voters will not distinguish this from partisan rhetoric.

As depicted in Panel (b) medium $\alpha = 1$ has the opposite bias. It always publishes (left-wing) rhetoric in state B. In state A it also publishes left-wing rhetoric, except when it learns factual evidence in favor of candidate a. Corresponding to the bias of the rhetoric we call such a medium *left-wing*. More generally a medium $\alpha \in (0, 1/2)$ is *moderately right-wing*, a medium $\alpha \in (1/2, 1)$ is *moderately left-wing*, and the medium $\alpha = 1/2$ is called *unbiased*.[39]

We treat the behavior of media as exogenous to the model and focus on optimal subscription choices by voters for a given set of media. With this interpretation of the model, Theorem 1 has the following implications for the choice of news media.

---

sufficiently long deadline, but leave the formal analysis of this case for future work.

[37]The notation suggests that voters subscribe to a single medium. Given that the arrival rates of evidence are linear in α, this is without loss—a "multi-homing" voter cannot achieve a higher payoff.

[38]Note that the partisan bias of the rhetoric is consistent with Bayesian updating. Since rhetoric is more frequent in state A, it must be in favor of the right-wing candidate.

[39]This interpretation of media bias is more precisely micro-founded in Appendix B of the Supplemental Material.
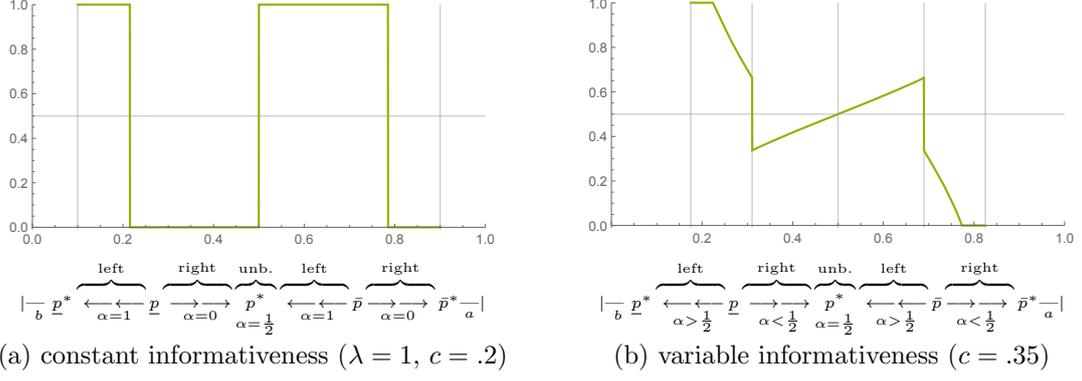
Figure 5: Optimal media choice. $(r = 0, u_a^A = u_b^B = 1, u_a^B = u_b^A = -1)$

Corollary 1. *Voters with extreme beliefs $p \notin \left(\underline{p}^*, \overline{p}^*\right)$ always vote for their favorite candidates without consulting media. Those subscribing to media exhibit the following behavior.*

*(a) (**Moderately informative media**) Suppose $\underline{c} \leq c < \overline{c}$:*

- *All voters to the right of $\check{p}$ subscribe to right-wing media and all voters to the left of $\check{p}$ subscribe to left-wing media.*
- *Over time absent breakthrough news, all voters become progressively more extreme and polarized.*

*(b) (**Highly informative media**) Suppose $c < \underline{c}$:*

- *(i) Right-leaning voters ($p \in (\overline{p}, \overline{p}^*)$) and moderate left-leaning voters ($p \in (\underline{p}, \underline{p}^*)$) subscribe to right-wing media. (ii) Left-leaning voters ($p \in (\underline{p}, \underline{p}^*)$) and moderate right-leaning voters ($p \in (p^*, \overline{p})$) subscribe to left-wing media. (iii) Undecided voters ($p = p^*$) subscribe to unbiased media.[40]*
- *Over time absent breakthrough news, moderate voters ($p \in \left(\underline{p}, \overline{p}\right)$) become increasingly undecided and more extreme voters become increasingly more extreme and polarized.*

Panel (a) of Figure 5 shows the choice of media α by voters with different beliefs. Figure 6 shows how the distribution of beliefs evolves over time (measured at three different times), where colors represent the media choice by voters who are still subscribing to media.

Among those subscribing to news media, voters on the far right choose right-wing media. Such a medium is valuable to them since it mostly publishes a content reinforcing their belief, and publishes left-favoring information *only if* it is accurate enough (in fact, fully conclusive in the baseline model) to actually convince them to change their votes. Over time, in the absence of convincing contradictory news, the right-wing medium feeds such voters with what they believe, leading them to become more extreme in their beliefs. Hence, applied to dynamic media choice with fully Bayesian voters, our model generates

---

[40]As is clear, the voters may equivalently divide their attention between right-wing and left-wing media.
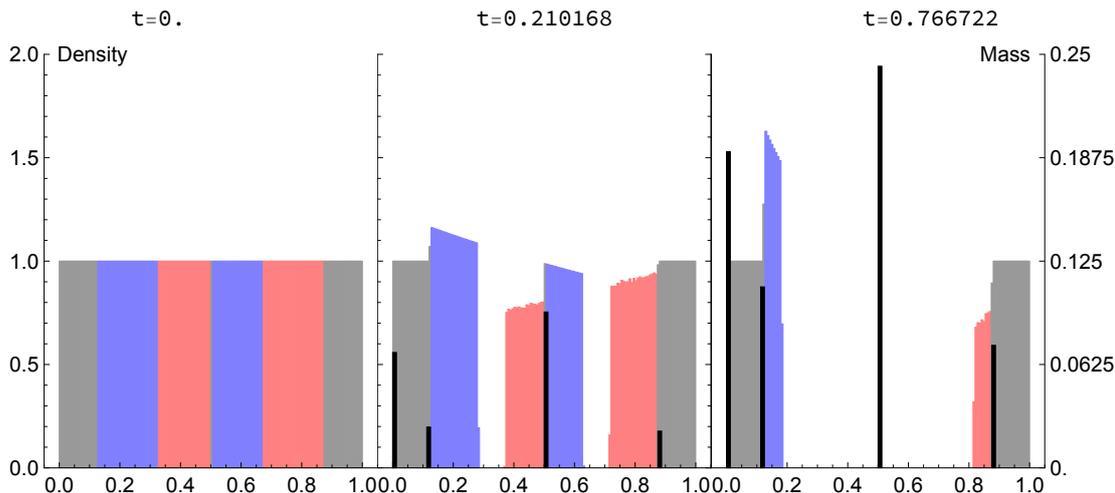
Figure 6: Evolution of media choice and beliefs when the true state is B. Shaded areas represent the density of beliefs (left axis). Bold bars represent mass points of beliefs (right axis).

self reinforcing beliefs—sometimes called an "echo-chamber" effect—that persists until strong contradictory evidence arrives.

The moderately right voter's media choice is quite different. Their moderate beliefs cause them to seek accurate evidence (of either kind) for voting. Initially they look for right-favoring evidence which they expect more likely to arise given their beliefs. Interestingly, they expect to find such evidence in left-wing media, since these media scrutinize the right-favoring information more and apply a very high standard for reporting such information. The moderate's media choice thus differs from the extreme voter's choice. The prediction of an "anti-echo chamber" effect is novel and has no analogue in previous literature. Over time, absent breakthrough news, their anti-echo chamber choice leads voters to be undecided. Ultimately, they switch to the unbiased medium.

In sum, our dynamic model of media choice predicts two different dynamics of belief evolution resulting from optimal media choice: the beliefs for those who are sufficiently extreme become more polarized, and the beliefs of those who are sufficiently moderate converge toward the middle and result in the subscription of unbiased media.

Extreme Beliefs and Preference for Skewed Media. One prediction of the baseline model is that voters only choose from three media, right-wing, left-wing, and unbiased. This is a consequence of the assumption that all media have access to information with identical arrival rates, and differ only in the rhetoric they publish. Formally, each news medium is characterized by a pair of arrival rates (λ^A, λ^B) for A- and B-evidence, respectively. The sum of arrival rates $\lambda^A + \lambda^B$ can be viewed as a measure of the *informativeness* of a medium. In the baseline model, if we normalize $\lambda = 1$, we have $\lambda^A = \alpha$ and $\lambda^B = \beta = 1 - \alpha$, so that $\lambda^A + \lambda^B = 1$ for all media. The analysis of this case

suggests that there is no demand for moderate media and the market will be dominated by extremely-biased outlets, as long as all media are identical in their informativeness.

We relax this assumption and introduce a trade-off between bias and informativeness by assuming less biased media to be more informative. This generates predictions about which voters have a stronger preference for informativeness versus bias.

Formally we assume that $\lambda^B = \Gamma(\lambda^A)$, where $\Gamma(\lambda^A)$ is a decreasing concave function, and set $\alpha = \lambda^A / (\lambda^B + \Gamma(\lambda^B))$. The interpretation is that news outlets that are more balanced have access to more factual information, for example because they employ journalists focused more on hard evidence.[41] In Appendix D in the Supplemental Material, we extend our characterization to this more general framework and show that it leads to choices of moderately biased media. Here, we illustrate the optimal choice in Panel (b) of Figure 5 with Γ given by: $(\lambda + \Gamma) + 3\sqrt{\lambda^2 + \Gamma^2} = 4$.

The overall structure of the optimal policy is similar, with extreme voters choosing media corresponding to their own bias and moderates doing the opposite. Regarding the trade-off between informativeness and bias, the model predicts that more moderate voters have a stronger preference for informativeness, whereas voters with more extreme beliefs, prefer more biased outlets. This reinforces the echo-chamber effect as extreme voters turn to more and more biased media as they become more polarized over time. Moderate voters who becomes less and less certain over time shift toward more and more balanced outlets. Overall, absent breakthrough news demand for outlets with moderate biases falls over time, as extreme voters become more extreme, switching to more extreme media; and moderate voters become more moderate and eventually subscribe to balanced media.

5 Generalized Model with Non-Conclusive Evidence

In the baseline model, we have assumed that the DM can access fully revealing signals. We now generalize the model to allow for the signals to be noisy. A-signals now arrive at rate $\overline{\lambda}$ in state A, and at rate $\underline{\lambda}$ in state B, where we assume $\overline{\lambda} > \underline{\lambda} \geq 0$. Importantly, the DM does not observe the state giving rise to these signals. Similarly, B-signals arrive at rate $\overline{\lambda}$ in state B, and at rate $\underline{\lambda}$ in state A.[42] If $\underline{\lambda} = 0$, the signals are fully revealing, and we obtain the baseline model as a special case.

With noisy signals, the DM may find it optimal to wait for more than one signal before taking an action. Our characterization from the baseline model generalizes naturally, if the DM finds it optimal *not* to wait for multiple signals. In this case, we say the model

---

[41]One could consider the opposite case where more biased outlets have access to more factual information which would reinforce the demand for skewed media that is already present in the baseline model. Formally, a convex function Γ will have no effect on the optimal choices, as long as the DM can divide attention across different news media in linear fashion.

[42]Allowing the DM to choose any signal with state contingent arrival rates $(\lambda_A, \lambda_B) \in [\underline{\lambda}, \overline{\lambda}]^2$ does not change the analysis; the DM will optimally choose only between $(\lambda_A, \lambda_B) = (\underline{\lambda}, \overline{\lambda})$ and $(\lambda_A, \lambda_B) = (\overline{\lambda}, \underline{\lambda})$.

satisfies the *single experimentation property* (SEP). We provide a necessary and sufficient condition for SEP in terms of a critical cost level and characterize the optimal strategy under SEP. In Section 6, we illustrate what the optimal strategy looks like when SEP does not hold.

5.1 Optimal Strategy under SEP

As in the baseline model, the DM's attention strategy determines which signals are actually observed. If she allocates a fraction $\alpha \in [0, 1]$ of her attention to seeking A-signals and $\beta = 1 - \alpha$ to seeking B-signals, she receives an A-signal with arrival rate $\alpha \lambda^A(p)$, where $\lambda^A(p) := p\overline{\lambda} + (1-p)\underline{\lambda}$. An A-signal leads to a posterior $q^A(p) := \overline{\lambda}p / \left(\overline{\lambda}p + \underline{\lambda}(1-p) \right) > p$. Similarly B-signals arrive at rate $(1 - \alpha)\lambda^B(p)$ where $\lambda^B(p) := p\underline{\lambda} + (1-p)\overline{\lambda}$, and lead to posterior $q^B(p) := \underline{\lambda}p / \left(\overline{\lambda}p + \underline{\lambda}(1 - p) \right) < p$. If the DM does not receive any signal, her belief evolves according to $\dot{p}_t = -\left(2\alpha_t - 1\right)\delta p_t(1 - p_t)$, where $\delta := \overline{\lambda} - \underline{\lambda}$ denotes the difference in arrival rates between signal and noise. We now state the optimal strategy, with the detailed analysis and explicit definitions relegated to Appendix A. As before, the optimal policy is a combination of immediate action, contradictory learning and confirmatory learning.

Theorem 2. *For any $(r, u_x^\omega, \overline{\lambda}, \underline{\lambda})$, there exists a critical cost-level $\underline{c}^{SEP} = \underline{c}^{SEP}(r, u_x^\omega, \overline{\lambda}, \underline{\lambda}) \geq 0$ such that SEP holds if and only if $c \geq \underline{c}^{SEP}$.[43] Given $c \geq \underline{c}^{SEP}$, there exists $\overline{c} = \overline{c}(r, u_x^\omega, \overline{\lambda}, \underline{\lambda}) > 0$ such that the optimal strategy is characterized as follows:*

(a) If $c \geq \overline{c}$, then the DM takes an optimal action $x^*(p)$ without information acquisition.

(b.i) If $c \in \left[\underline{c}^{SEP}, \overline{c}\right)$, for $p \in (\underline{p}^*, \overline{p}^*)$, the optimal strategy takes either the contradictory form (4) or the mixed contradictory and confirmatory form (5). For $p \notin (\underline{p}^*, \overline{p}^*)$, it is optimal to take an immediate action.

(b.ii) For $c \in \left[\underline{c}^{SEP}, \overline{c}\right)$ in a neighborhood of $\overline{c}$, (4) is optimal, and for r and $\underline{\lambda}$ both sufficiently low, (5) is optimal for $c \in \left[\underline{c}^{SEP}, \overline{c}\right)$ in a neighborhood of $\underline{c}^{SEP}$.

5.2 Implications for Stochastic Choice and Response Time

The general model exhibits rich implications for the choice process. We explore them and compare them to some well-known benchmarks. In classic models such as Luce's logit model, stochastic choice can be interpreted as arising from unobserved random utility. In contrast, in models of information acquisition such as Wald, RI, and DDM, as well as our model, choice is stochastic because the DM learns the payoff relevant state with noise, so that the choice probabilities depend on the DM's prior. This is not the case in random-utility models. There is a particular structure to this dependence and we first explore the predictions for the accuracy and delay of decisions across DMs with different

---

[43]We have $\underline{c}^{SEP} = 0$ if $r > 0$ and $\underline{\lambda}$ is sufficiently close to zero.

(a) Accuracy (dotted: RI, solid: our model)   (b) Expected Delay
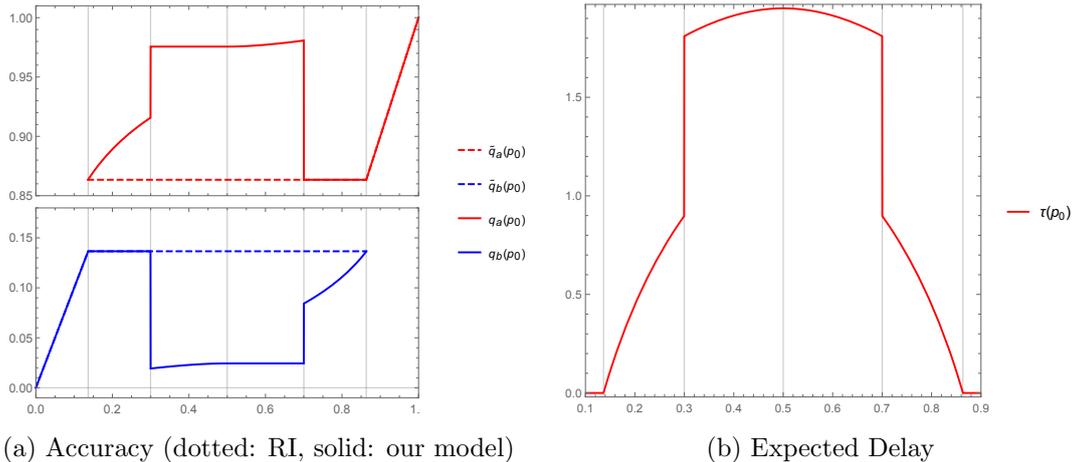
Figure 7: Parameters: $\overline{\lambda} = 1$, $\underline{\lambda} = .025$, $r = 0$, $c = 0.115$, $u_a^A = u_b^B = 1$, $u_b^A = u_a^B = 0$. RI-weight of cost function $\mu = 1.025$. Grid lines at: $\underline{p}^*$, $\underline{p}$, p^*, $\overline{p}$, $\overline{p}^*$ (left to right).

prior beliefs. Second, we explore the predicted time pattern of choice in a single DM with fixed prior.

Accuracy of decisions. A measure of the accuracy of the decision is the average posterior belief that the DM holds when taking a particular action. In the RI model, the accuracy does not vary with the DM's initial belief (Matejka and McKay, 2015). This property is also shared by the DDM in which the DM makes a decision if and only if her belief drifts to one of the stopping boundaries that are constant over time.[44] The accuracy, or the belief, at the boundary therefore determines the accuracy of the decision.

In our model, by contrast, the decision accuracy varies with the initial belief. The reason is that decisions are not only taken when the belief has reached the bound of the experimentation region. For both the contradictory and the confirmatory region, decisions are also taken after signals are observed, in which case the belief will land in the interior of the stopping region. Moreover, since our optimal strategy is a combination of the contradictory strategy, which has a drift to bound structure, and the confirmatory strategy, where actions are only taken if signals are observed, the accuracy depends on the initial belief in a complex way, as depicted in Panel (a) of Figure 7, which also reports the accuracy in the RI-model (the constant dashed lines).

Most striking is the difference between the contradictory region and the confirmatory region. The higher initial uncertainty in the latter is associated with a significantly higher accuracy of the final action. To make this precise, we say that a DM with belief p makes a **uniformly more accurate decision** than a DM with belief p' if the following holds: *For any two histories h and h', resulting from p and p', respectively, where the DM takes*

---

[44]This holds in our framework with binary states and actions (Peskir and Shiryaev, 2006). Below we will discuss the DDM with a continuum of states considered by Fudenberg, Strack, and Strzalecki (2017) which generates prior-dependence.

the same action $x \in \{a, b\}$, the posterior belief at h must be more extreme in favor of action x than the posterior belief at h'.

Proposition 3 ("Skepticism fosters accuracy"). *Suppose $c \in [\underline{c}^{SEP}, \overline{c})$ and $(\underline{p}, \overline{p}) \neq \emptyset$. Then, a DM with any prior belief $p \in (\underline{p}, \overline{p})$ makes a uniformly more accurate decision than a DM with any prior belief $p' \in (\underline{p}^*, \underline{p}) \cup (\overline{p}, \overline{p}^*)$.*

Proposition 3 makes precise the intuition why the DM employs the confirmatory strategy when she is uncertain and the contradictory strategy when she is more certain. As the proposition shows, the former strategy involves more information acquisition than the latter. Since the value of information is higher when the DM is initially more uncertain, she employs a strategy that is more effective in fuller learning (i.e., confirmatory strategy). Of course this comes at a cost: more accurate information requires longer deliberation, as we show next.

Delay. Let $\tau(p)$ denote the *expected delay* of taking an action, where p is the prior belief of the DM. We show that $\tau(p)$ is single-peaked as illustrated in Panel (b) of Figure 7.

Proposition 4 ("Skepticism entails delay"). *Suppose $c \in [\underline{c}^{SEP}, \overline{c})$. Then $\tau(p)$ is quasi-concave with maximum at $p' \in (\underline{p}^*, \overline{p}^*)$ if either of the following holds:*
  (a) *The contradictory strategy is optimal. In this case $p' = \check{p}$.*
  (b) *The optimal strategy combines contradictory and confirmatory evidence, r is sufficiently close to zero, and $p^* \in (\underline{\lambda}/(\underline{\lambda} + \overline{\lambda}), \overline{\lambda}/(\underline{\lambda} + \overline{\lambda}))$. In this case $p' \in [\underline{p}, \overline{p}]$.*

Clearly, in the contradictory region, the delay increases in the distance between the prior belief and the boundary of the experimentation region. For the confirmatory region, we show that the delay is concave in p. If payoffs are symmetric, the longest expected delay arises at $p^* = 1/2$. Intuitively, at $p \neq 1/2$ a decision is made faster since the DM focuses her attention on the signal that is more likely to arrive.[45] Finally, we show using a revealed preference argument that at $\underline{p}$ and $\overline{p}$, the delay in the contradictory strategy must be shorter to compensate for the less accurate decision.

Similar results hold in the RI model and DDM. The difference is that in our model, the two modes of learning lead to discontinuities in the average delay when the posterior moves from the confirmatory to the contradictory region.[46]

Speed-Accuracy Complementarity. So far, we have analyzed properties of the stochastic choice as a function of the prior belief. Next, we obtain a prediction on how, for a given prior, the accuracy varies as a function of decision time.

---

[45]In the asymmetric case, p^* differs from $1/2$ so that the DM sometimes focuses attention on the less likely signals. τ is still concave in this case but the peak may differ from both p^* and $1/2$.

[46]In the static RI model, we can interpret the (average) cost incurred as delay and the result follows from concavity of mutual information. In the DDM, the result follows straightforwardly from the convexity of the value function (see Peskir and Shiryaev, 2006, Theorem 21.1).

Proposition 5. *Suppose $c \in [\underline{c}^{SEP}, \overline{c})$. For a DM with belief in $p_0 \in [\underline{p}^*, \underline{p}] \cup [p^*, \overline{p}]$, conditional on taking a, a later decision results in lower accuracy. For a DM with $p_0 \in [\underline{p}, p^*] \cup [\overline{p}, \overline{p}^*]$, conditional on choosing b, a later decision results in lower accuracy.*

We obtain *speed-accuracy complementarity*, meaning later decisions tend to be less accurate. The reason is that (regardless of the prior), the DM becomes less and less convinced of the state she is seeking evidence for. Hence, if she discovers, say A-evidence, after a long period of waiting, she updates to a lower posterior than if she discovers A-evidence quickly (formally, $q^A(p_t) < q^A(p_0)$).

Whether a long delay produces a more or less accurate decision has been of considerable interest in cognitive psychology. In a typical experiment in this area, subjects make choices over time—often of perceptual judgment type (see Ratcliff and McKoon, 2008, for a survey). What emerges as a pattern, is that decisions with a longer response time tend to be less accurate.

This phenomenon, known as *speed-accuracy complementarity*, is difficult to explain via the DDM often used to fit the data, unless the stopping boundaries collapse over time.[47] Fudenberg, Strack, and Strzalecki (2017) justify collapsing boundaries by introducing a rich uncertainty in terms of payoff differences between two actions, and produce a less accurate decision over time.[48] We obtain a similar prediction in a Poisson model but the mechanism behind our result is different.[49]

6 Repeated Experimentation

Our paper has focused on the situation in which the SEP holds, which requires the cost of information acquisition to be sufficiently high (Theorem 2). We now discuss cases where the SEP fails. In the examples we solve, the optimal solution retains the same structure

---

[47]The DDM assumes that a subject's perception follows a diffusion process with drift reflecting the underlying stimulus. In DDMs with constant boundaries, the threshold beliefs for decisions do not change over time, so a decision is made with the same accuracy over time, and *speed-accuracy independence* follows. If the boundaries collapse over time, then a later decision is made with less extreme or more uncertain beliefs, so speed-accuracy complementarity can be explained. See Drugowitsch, Moreno-Bote, Churchland, Shadlen, and Pouget (2012) and Hawkins, Forstmann, Wagenmakers, Ratcliff, and Brown (2015) for discussions of the DDM with collapsing boundaries and the related evidence.

[48]Given the uncertainty about payoff differences, a long passage of time without observing strong evidence in favor of either action indicates that payoff differences are likely to be small. This reduces the incentive to acquire further information, causing the DM to stop with lower accuracy.

[49]The extent to which Poisson signals accurately describe decision making in these experiments, or the underlying neurophysiology, is not clear. Whether brain processes stimuli through a gradual process described by DDM or a discrete process described by a Poisson model is subject of active research in neuroscience (see Latimer, Yates, Meister, Huk, and Pillow (2015) for instance). Within a Poisson model, our reliance on SEP further limits the applicability of our model in explaining these experiments. Nevertheless, our finding points to the potential usefulness of a Poisson-based approach; we leave for future research extension of the current model toward this direction.
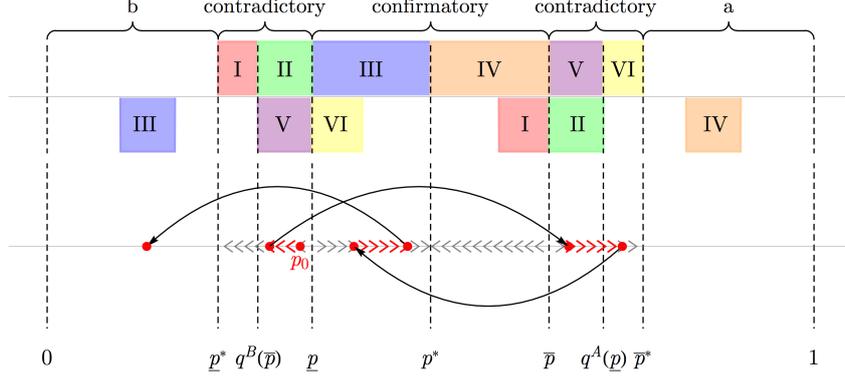
Figure 8: Repeated Experimentation. $\overline{\lambda} = 1$, $\underline{\lambda} = .2$, $r = 0$, $c = .1$, $u_a^A = u_b^B = 1$, $u_b^A = u_a^B = 0$.

as in Theorems 1 and 2.[50] In particular, the optimal policy combines immediate action, confirmatory and contradictory evidence. At the boundaries of the experimentation region, it is still optimal to seek contradictory signals, and the confirmatory region appears if the cost of information acquisition is sufficiently low.

Unlike when SEP holds, however, the DM may optimally switch between different modes of learning, as discovery of evidence while following one strategy may lead to a posterior for which a different learning strategy is prescribed. Figure 8 depicts an optimal policy that exhibits one such pattern.[51]

The colored bars indicate originating beliefs on top, and on the bottom, the corresponding regions where the posterior is located after receiving a signal. For instance, discovery made in top region I will lead to a posterior in bottom region I. In this example, discovery made in a contradictory region (top regions I-II and V-VI) leads to a posterior in another part of the contradictory region or in the confirmatory region, that is, it will never lead to an immediate action.[52]

The bottom graph of Figure 8 illustrates a possible evolution of the belief under the optimal strategy. Suppose the DM has a prior belief p_0 in top II region. She then looks for A-evidence, as prescribed by contradictory learning. Her belief drifts down until she discovers A-evidence. Suppose she does so at p_1, still within region II. Her belief then jumps up to p_2 (in bottom II, or top V, region). Now she looks for B-evidence, as prescribed by contradictory learning for that belief. Her belief drifts up until she finds B-evidence. Suppose she does so at p_3 in top VI region. Her belief then jumps down to p_4 in bottom VI, or top III, region, a confirmatory region where she again looks for

---

[50]Providing a general characterization is difficult since the number of possible strategies increases. Our analysis of examples studied in Appendix E in the Supplemental Material reveals both the difficulties and the potential ways to overcome them.

[51]A plot of the value function and posterior as well as other cases can be found in Appendix E of Supplemental Material.

[52]As in the case where SEP holds, a discovery made in the confirmatory region (top regions III-IV) leads to a posterior in the stopping region and an immediate action.

B-evidence. Suppose she discovers B-evidence at p_5. This time her belief jumps down to p_6, in bottom region III, which is outside the experimentation region. She now picks action b, completing the journey.

7   Conclusion

We have studied a model in which a decision maker may allocate her limited attention to collecting different types of evidence that support alternative actions. Assuming Poisson arrival of evidence, we have shown that the optimal policy combines *immediate action, contradictory learning*, and *confirmatory learning* for different prior beliefs. We have then used this characterization to yield rich predictions about information acquisition and choices several in real-world settings.

We envision several avenues of extending the current work. First, our model yields a relatively tractable characterization (e.g., the value function and the optimal policy are available in closed form). Therefore, we hope that this framework will be useful to integrate dynamic choice of information in applied theory models. This includes our applications of jury and media choice, which could be extended beyond our analysis in the present paper. Our dynamic information choice problem might be also embedded into a principal-agent setup in which a principal tries to induce an agent to acquire information or a strategic setup such as R&D competition where different firms choose from alternative innovation approaches over time.

Second, one may relax the "lumpiness" of information arrival to allow for more gradual information acquisition. The previous section dealing with repeated experimentation already points to an extension in this direction, and suggests that our characterization is robust to such a generalization. A complete analysis of this case will be useful for applications in which decision makers learn gradually, such as a researcher who makes day-to-day decisions about the next steps in a project, as opposed to a manager who decides only based on reports that are made once a month. We leave this for future research.

A   Proofs

A.1   Proofs for the General Model

All parts of the proofs that are common to the baseline are stated for the general model. The results specialize to the baseline model by setting $\overline{\lambda} = \lambda$ and $\underline{\lambda} = 0$. We start by formally stating the DM's problem for the general model in Section 5. We analyze the problem assuming that the DM is forced to take an action after receiving the first signal. We will then show this is true iff $c > \underline{c}^{SEP}$

As in the baseline model, the DM chooses an attention strategy (α_t) and a time $T \in [0, \infty]$ at which she will stop acquiring information if she does not observe any signal up to time T. Her problem is thus given by

$$V(p_0) = \max_{(\alpha_\tau), T} \left\{ \int_0^T e^{-rt} P_t(p_0, (\alpha_\tau)) \left[\alpha_t \lambda^A(p) U(q^A(p)) + (1 - \alpha_t) \lambda^B(p) U(q^B(p)) - c \right] dt \right.$$

$$\left. + e^{-rT} P_\tau(p_0, (\alpha_\tau)) U(p_T) \right\} \tag{6}$$

where $P_t(p_0, (\alpha_\tau)) := p_0 e^{-\left(\overline{\lambda} \int_0^t \alpha_s ds + \underline{\lambda} \int_0^t \beta_s ds \right)} + (1 - p_0) e^{-\left(\underline{\lambda} \int_0^t \alpha_s ds + \overline{\lambda} \int_0^t \beta_s ds \right)}$. Given that the problem is autonomous, the optimal $\alpha(p_t)$ and $T(p_t)$ only depend on the belief at time t. Similarly, the decision to stop and take an action only depends on p_t.

The HJB equation for this problem is a variational inequality

$$\max \left\{ -c - rV(p) + F(p, V(p), V'(p)), U(p) - V(p) \right\} = 0, \tag{7}$$

where

$$F(p, V(p), V'(p)) = \max_{\alpha \in [0,1]} \left\{ \begin{array}{c} \alpha \lambda^A(p) \left(U(q^A(p)) - V(p) \right) + (1 - \alpha) \lambda^B(p) \left(U(q^B(p)) - V(p) \right) \\ - (2\alpha_t - 1) \delta p (1 - p) V'(p) \end{array} \right\}. \tag{8}$$

In the "experimentation region" where $V(p) > U(p)$ the HJB equation reduces to

$$c + rV(p) = F(p, V(p), V'(p)). \tag{9}$$

If $V(p) = U(p)$, then $T(p) = 0$ is optimal and we must have $c + rV(p) \geq F(p, V(p), V'(p))$.

A.1.1 Preliminary Lemmas

For the stationary strategy described in the text, $\alpha_t = 1/2$ for all t implies that the posterior is constant for all t. The value achieved by this strategy is obtained from (6):

$$\overline{U}(p) = \frac{\frac{1}{2}\lambda^A(p) U_a \left(q^A(p) \right) + \frac{1}{2}\lambda^B(p) U_b \left(q^B(p) \right) - c}{r + \frac{1}{2}\overline{\lambda} + \frac{1}{2}\underline{\lambda}}. \tag{10}$$

We will also use the value of a hypothetical "full attention strategy" where $\alpha = \beta = 1$, i.e., the DM devotes full attention to both types of evidence. In our model, limited attention prevents the DM from doing so. Therefore, this only a benchmark used in the proofs. We denote the value of this strategy by $\widehat{U}(p)$ and we have

$$\widehat{U}(p) = \frac{\lambda^A(p) U_a(q^A(p)) + \lambda^B(p) U_b(q^B(p)) - c}{r + \overline{\lambda} + \underline{\lambda}}. \tag{11}$$

For $\overline{U}(p)$ and $\widehat{U}(p)$ we have assumed that the DM will always take action a after an A-signal, and action b after a B-signal. This will be convenient in the following. We also note that $\overline{U}(p)$ and $\widehat{U}(p)$ are linear in p. Lemma 9 in Appendix A.4 shows that $\max\left\{U(p),\widehat{U}(p)\right\}$ is an upper bound for the value function if SEP holds. Therefore, the following straightforward Lemma will be important for our analysis.

Lemma 1. $\widehat{U}(\hat{p}) \geq U(\hat{p})$ *is equivalent to*

$$\frac{c}{\delta}\left(\Delta^A - \Delta^B\right) + \frac{r}{\delta}\left(u_a^A u_b^B - u_b^A u_a^B\right) \leq -\Delta^A\Delta^B. \tag{EXP}$$

We now turn to the characterization of the value function in (6). $V(p)$ must satisfy the HJB equation (9) for all p where it is differentiable and $V(p) > U(p)$. Since the objective in (8) is linear in α, the optimal policy is a bang-bang solution and we have $\alpha^*(p) \in \{0, 1\}$ except for posteriors where the derivative of the objective vanishes. We therefore consider solutions to the ODEs given by $\alpha = 0$ and $\alpha = 1$:

$$c + rV_0(p) = \lambda^B(p)\left(U(q^B(p)) - V_0(p)\right) + \delta p(1-p)V_0'(p), \tag{12}$$

$$c + rV_1(p) = \lambda^A(p)\left(U(q^A(p)) - V_1(p)\right) + \delta p(1-p)V_1'(p). \tag{13}$$

Solutions to these equations with boundary condition $V(x) = W$ are denoted by $V_0(p; W, x)$ and $V_1(p; W, x)$, respectively.[53]

We start by showing two basic properties of solutions to (12) and (13) that hold whenever these solutions are greater or equal than the value of the stationary strategy $\overline{U}(p)$. The first Lemma shows that V_1 can cross V_0 only from above.

Lemma 2. *Let V_0 and V_1 differentiable and satisfy (12) and (13), respectively, at $p \in (0,1)$. If $V_0(p) = V_1(p)$ then*

$$V_0'(p) > (=)V_1'(p) \iff V_0(p) = V_1(p) > (=)\overline{U}(p).$$

Proof. Suppose $V_0(p) = V_1(p) = V(p)$ for some $p \in (0,1)$. Solving (12) and (13) for $V_0'(p)$ and $V_1'(p)$ and some algebra yields $p(1-p)\delta\left(V_0'(p) - V_1'(p)\right)/\left(2r + \overline{\lambda} + \underline{\lambda}\right) = V(p) - \overline{U}(p)$. Therefore $\mathrm{sgn}(V_0'(p) - V_1'(p)) = \mathrm{sgn}\left(V(p) - \overline{U}(p)\right)$. $\qquad\square$

The next Lemma shows that solutions to (12) and (13) that exceed $\overline{U}(p)$ are "locally unimprovable", that is, they satisfy the HJB equation.

Lemma 3. *Let $V(p)$ be differentiable at $p \in (0,1)$ and $V(p) \geq \overline{U}(p)$.*
  (a) *If V satisfies (12) at p, then V satisfies (9) at p, and $\alpha = 0$ is a maximizer.*
  (b) *If V satisfies (13) at p, then V satisfies (9) at p, and $\alpha = 1$ is a maximizer.*

---

[53]$V_0(p; W, x)$ and $V_1(p; W, x)$ are uniquely defined if $x \in (0,1)$ because (12) and (13) satisfy local Lipschitz continuity for all $p \in (0,1)$.

The maximizers are unique if $V(p) > \overline{U}(p)$.

Proof. Consider first the case that $V(p)$ satisfies (12). Denote the objective on the RHS of (9) by $X(\alpha, p, V, V')$. With $V = V_0(p)$ and $V' = V_0'(p)$ substituted from (12) we have

$$\frac{\partial X(\alpha, p, V_0(p), V_0'(p))}{\partial \alpha} = \left(2r + \overline{\lambda} + \underline{\lambda}\right)\left(\overline{U}(p) - V_0(p)\right).$$

This implies that $\alpha = 0$ is a maximizer if $V(p) \geq \overline{U}(p)$ and $\alpha = 0$ is a unique maximizer if the inequality is strict. This proves part 1. The proof of part 2 follows from a similar argument. $\qquad\square$

A.1.2 Contradictory Evidence

We start the construction of the contradictory strategy by defining

$$\underline{p}^* := \inf\left\{p \in [0, \hat{p}) \,\big|\, U(p) \geq 0 \text{ and } c + rU_b(p) \leq \lambda^A(p)\left(U_a(q^A(p)) - U_b(p)\right) - \delta p(1 - p)U_a'(p)\right\}. \tag{14}$$

$$\overline{p}^* := \sup\left\{p \in (\hat{p}, 1] \,\big|\, U(p) \geq 0 \text{ and } c + rU_a(p) \leq \lambda^B(p)\left(U_b(q^B(p)) - U_a(p)\right) + \delta p(1 - p)U_a'(p)\right\}, \tag{15}$$

In case the inf or sup is taken over the empty set, we define $\underline{p}^* := 1$ and $\overline{p}^* := 0$, respectively. If these bounds are non-trivial, we can characterize them in terms of $\widehat{U}(p)$.

Lemma 4. *If* (EXP) *is satisfied,* $0 < \underline{p}^* < \hat{p} < \overline{p}^* < 1$ *and*

$$\widehat{U}(\underline{p}^*) = U(\underline{p}^*) \quad \text{and} \quad \widehat{U}(\overline{p}^*) = U(\overline{p}^*)$$

If (EXP) *is violated,* $\underline{p}^* = 1$ *and* $\overline{p}^* = 0$.

Proof. We show the result for $\underline{p}^*$. Some algebra shows that

$$c + rU_b(p) = \lambda^A(p)\left(U_a(q^A(p)) - U_b(p)\right) - \delta p(1 - p)U_b'(p) \quad \Longleftrightarrow \quad \widehat{U}(p) = U_b(p).$$

The solution to this defines a candidate for $\underline{p}^*$. It is straightforward to verify that

$$c + rU_b(p) \leq \lambda^A(p)\left(U_a(q^A(p)) - U_b(p)\right) - \delta p(1 - p)U_b'(p) \quad \Longleftrightarrow \quad p \geq \underline{p}^*.$$

Therefore (EXP) implies that $\underline{p}^*$ is given by $\widehat{U}(\underline{p}^*) = U(\underline{p}^*)$. On the other hand, if (EXP) is violated, the candidate is greater than $\hat{p}$. Therefore the infimum in (14) is taken over the empty set and we have $\underline{p}^* = 1$. The proof for $\overline{p}^*$ is similar. $\qquad\square$

The bounds are given by

$$\underline{p}^* = \frac{\left(u_b^B - u_a^B\right)\underline{\lambda} + u_b^B r + c}{r\left(u_b^B - u_a^A\right) + \left(u_a^A - u_b^A\right)\overline{\lambda} + \left(u_b^B - u_a^B\right)\underline{\lambda}}, \tag{16}$$

$$\overline{p}^* = \frac{\left(u_b^B - u_a^B\right)\overline{\lambda} - u_a^B r - c}{r\left(u_a^A - u_a^B\right) + \left(u_b^B - u_a^B\right)\overline{\lambda} + \left(u_a^A - u_b^A\right)\underline{\lambda}}. \tag{17}$$

We define a left branch $\underline{V}_{ct} : [0,1] \to \mathbb{R}$ and the right branch $\overline{V}_{ct} : [0,1] \to \mathbb{R}$ of the value function for the contradictory strategy as follows:[54]

$$\underline{V}_{ct}(p) := \begin{cases} U(p), & \text{if } p \leq \underline{p}^*, \\ V_1(p; \underline{p}^*, U(\underline{p}^*)), & \text{if } p > \underline{p}^*, \end{cases}$$

$$\overline{V}_{ct}(p) := \begin{cases} U(p), & \text{if } p \geq \overline{p}^*, \\ V_0(p; U(\overline{p}^*, \overline{p}^*)), & \text{if } p < \overline{p}^*. \end{cases}$$

Note that $\underline{V}_{ct}(p) = \overline{V}_{ct}(p) = U(p)$ for all p if (EXP) is violated, so that $\underline{p}^* = 1$ and $\overline{p}^* = 0$. A simple observation is the following. Note in particular that this Lemma shows smooth pasting at $\underline{p}^*$ and $\overline{p}^*$.

Lemma 5. $\underline{V}_{ct}(p)$ and $\overline{V}_{ct}(p)$ are continuously differentiable and convex on $(0,1)$. $\underline{V}_{ct}(p)$ is strictly convex on $[\underline{p}^*, 1]$, and $\overline{V}_{ct}(p)$ is strictly convex on $[1, \overline{p}^*]$.

Proof. If $\underline{p}^* = 1$, the statement is trivial. Otherwise, smooth pasting at $\underline{p}^*$ holds since $\underline{V}_{ct}(p)$ satisfies (13), $\underline{V}_{ct}(\underline{p}^*) = U_b(\underline{p}^*)$, and the second condition in (14) holds with equality. The argument is similar for $\overline{V}_{ct}(p)$. Verifying $\underline{V}_{ct}''(p), \overline{V}_{ct}''(p) > 0$ yields strict convexity for each branch. Convexity extends to the whole interval $[0,1]$ by smooth pasting. $\square$

We combine the branches and define $V_{ct}(p) := \max\left\{\underline{V}_{ct}(p), \overline{V}_{ct}(p)\right\}$. Next we want to identify a condition under which $\underline{V}_{ct}(p)$ and $\overline{V}_{ct}(p)$ have a unique point of intersection $\check{p} \in \left(\underline{p}^*, \overline{p}^*\right)$. For this result we need an additional Lemma

Lemma 6. Let $V(p)$ be differentiable at $p \in (0,1)$, and $V(p) = \widehat{U}(p)$.
  (a) If V satisfies (13) at p and $q^B(p) < \underline{p}^*$, then $V'(p) < \widehat{U}'(p)$.
  (b) If V satisfies (12) at p and $q^A(p) > \underline{p}^*$, then $V'(p) > \widehat{U}'(p)$.

Proof. For (a) we solve (13) for $V'(p)$. Some algebra then shows that $V'(p) > \widehat{U}'(p)$ is equivalent to $q^B(p) < \underline{p}^*$. (b) is similar. $\square$

This Lemma allows us to be more precise about the structure of the contradictory solution. For the proof of the following Lemma, we use the following condition:

$$q^A(\underline{p}^*) > \overline{p}^*. \tag{NSEC}$$

---

[54]Explicit expressions are given in (1) and (2) in the Supplemental Material.

This condition requires that the DM's posterior after an A-signal is outside the experimentation region for all beliefs in the experimentation region. Notice, that $q^A(q^B(p)) = p$ for all $p \in (0, 1)$. Since $q^A(p)$ is strictly increasing, (NSEC) is equivalent to $q^B(\overline{p}^*) < \underline{p}^*$. Hence the condition also implies that the DM's posterior after a B-signal is outside the experimentation region. Clearly, (NSEC) is a necessary condition for the single experimentation property. We will see later that it is also sufficient. For the moment, however, we use this property to derive the solution to the problem in (6) where the DM is forced to take an action after receiving the first signal.

Lemma 7. *Suppose* (EXP) *and* (NSEC) *are satisfied. If* $V_{ct}(p) \geq \overline{U}(p)$ *for all* p, *then there exists a unique* $\check{p} \in (\underline{p}^*, \overline{p}^*)$ *such that* $\underline{V}_{ct}(\check{p}) = \overline{V}_{ct}(\check{p})$ *and* $\underline{V}_{ct}(p)$ *crosses* $\overline{V}_{ct}(p)$ *from above at* $\check{p}$. *Moreover* $V_{ct}(p)$ *is convex and strictly convex on* $(\underline{p}^*, \overline{p}^*)$.

Proof. Since (EXP) is satisfied, Lemma 4 implies $\underline{p}^* < \overline{p}^*$, $\underline{V}_{ct}(\underline{p}^*) = \widehat{U}(\underline{p}^*)$, and $\overline{V}_{ct}(\overline{p}^*) = \widehat{U}(\overline{p}^*)$. Since $\underline{V}'_{ct}(\underline{p}^*) = U'_b(\underline{p}^*) < \widehat{U}'(\underline{p}^*)$ and $\overline{V}'_{ct}(\overline{p}^*) = U'_a(\overline{p}^*) > \widehat{U}'(\overline{p}^*)$, Lemma 6 thus implies $\underline{V}_{ct}(\overline{p}^*) < \overline{V}_{ct}(\overline{p}^*)$ and $\overline{V}_{ct}(\underline{p}^*) < \underline{V}_{ct}(\underline{p}^*)$. By the intermediate value theorem, there exists $\check{p} \in (\underline{p}^*, \overline{p}^*)$ such that $\underline{V}_{ct}(\check{p}) = \overline{V}_{ct}(\check{p})$. By Lemma 2 and $V_{ct}(p) \geq \overline{U}(p)$ for all $p \in (\underline{p}^*, \overline{p}^*)$, $\check{p}$ is unique and $\underline{V}_{ct}(p)$ crosses $\overline{V}_{ct}(p)$ from above at $\check{p}$. Since $V_{ct}(p)$ is strictly convex on $(\underline{p}^*, \check{p})$ and $(\check{p}, \overline{p}^*)$ and has a convex kink at $\check{p}$, it is strictly convex on $(\underline{p}^*, \overline{p}^*)$. $\square$

In the case where $V_{ct}(p) < \overline{U}(p)$ for some p, there may be multiple intersections of $\underline{V}_{ct}(p)$ and $\overline{V}_{ct}(p)$. We will see, however, that the contradictory strategy will not be part of the optimal strategy for beliefs between the lowest and the highest intersection in this case. Therefore, we do not refine the definition of $V_{ct}(p)$ for that case.

A.1.3 Confirmatory Evidence

For given p^*, the value of the confirmatory strategy is given by[55]

$$V_{cf}(p) := \begin{cases} \underline{V}_{cf}(p) := V_0(p; p^*, \overline{U}(p^*)), & \text{if } p \leq p^*, \\ \overline{V}_{cf}(p) := V_1(p; p^*, \overline{U}(p^*)), & \text{if } p > p^*. \end{cases} \tag{18}$$

Since

$$V(p^*) = \underline{V}_{cf}(p^*) = \overline{V}_{cf}(p^*) = \overline{U}(p^*), \tag{19}$$

Lemma 2 implies smooth pasting: $\underline{V}'_{cf}(p^*) = \overline{V}'_{cf}(p^*)$ and $V_{cf}(p)$ is differentiable at p^*. Moreover, we choose p^* such that the value function is tangent to $\overline{U}(p)$

$$V'(p^*) = \overline{U}'(p^*). \tag{20}$$

---

[55]Explicit expressions can be found in (3) and (4) in the Supplemental Material.

Otherwise the confirmatory strategy is dominated by the stationary attention strategy for some p. Inserting (19) and (20) in (12) or (13) yields a unique solution for p^*:

$$p^* = \frac{\left(u_b^B - u_a^B\right)\overline{\lambda}\underline{\lambda} + r\left(u_b^B\overline{\lambda} - u_a^B\underline{\lambda}\right) + c\left(\overline{\lambda} - \underline{\lambda}\right)}{\left(u_a^A - u_b^A\right)\overline{\lambda}\underline{\lambda} + r\left(u_a^A\overline{\lambda} - u_b^A\underline{\lambda}\right) + \left(u_b^B - u_a^B\right)\underline{\lambda}\overline{\lambda} + r\left(u_b^B\overline{\lambda} - u_a^B\underline{\lambda}\right) + 2c\left(\overline{\lambda} - \underline{\lambda}\right)}. \tag{21}$$

Given our assumptions, it is straightforward to verify that $p^* \in (0,1)$.

Lemma 8. *Let $V_{cf}(p)$ be defined by (18) and (21).*
  (a) *$V_{cf}(p)$ is continuously differentiable and strictly convex on $(0,1)$, and $V_{cf}(p) \geq \overline{U}(p)$ for all p with strict inequality for $p \neq p^*$.*
  (b) *If (EXP) and (NSEC) are satisfied, $V_{cf}(p) < V_{ct}(p)$ for $p \in \{\underline{p}^*, \overline{p}^*\}$*

Proof. (a) Continuous differentiability follows from Lemma 2 and (19). $\underline{V}''_{cf}(p), \overline{V}''_{cf}(p) > 0$ are equivalent to $p^* \in (0,1)$. We omit the algebra. Since $\underline{V}'_{cf}(p^*) = \overline{V}'_{cf}(p^*)$, V_{cf} is strictly convex on $[0,1]$. Since $V_{cf}(p^*) = \overline{U}(p^*)$, $\overline{U}(p)$ is linear in p, and $V'_{cf}(p^*) = \overline{U}'(p^*)$, strict convexity implies $V_{cf}(p) > \overline{U}(p)$ for all $p \neq p^*$.

(b) Since $\overline{V}_{cf}(p^*) = \overline{U}(p^*) < \widehat{U}(p^*)$ and $\overline{V}_{cf}(p)$ satisfies (13), Lemmas 4 and 6 imply $\overline{V}_{cf}(\overline{p}^*) < \widehat{U}(\overline{p}^*) = V_{ct}(\overline{p}^*)$. The proof for $\underline{p}^*$ follows from a symmetric argument. $\qquad\square$

A.1.4 Characterization of the value function

To prove the main theorems, we first characterize the properties of the upper envelope of V_{ct} and V_{cf}.

Proposition 6. *Suppose (EXP) and (NSEC) hold and let $V(p) = \max\{V_{ct}(p), V_{cf}(p)\}$. Then $V(p) \geq \overline{U}(p)$ for all $p \in [0,1]$.*
  (a) *If $V_{ct}(p^*) \geq V_{cf}(p^*)$, then $V(p) = V_{ct}(p)$ for all $p \in [0,1]$.*
  (b) *If $V_{ct}(p^*) < V_{cf}(p^*)$, then there exist $\underline{p} \in (\underline{p}^*, p^*)$ such that $\underline{V}_{ct}(\underline{p}) = V_{cf}(\underline{p})$ and $\overline{p} \in (p^*, \overline{p}^*)$ such that $\overline{V}_{ct}(\overline{p}) = V_{cf}(\overline{p})$ and*

$$V(p) = \begin{cases} \underline{V}_{ct}(p), & \text{if } p \leq \underline{p}, \\ V_{cf}(p), & \text{if } p \in [\underline{p}, \overline{p}], \\ \overline{V}_{ct}(p), & \text{if } p \geq \overline{p}. \end{cases}$$

Proof. By Lemma 8, $V_{cf}(p) \geq \overline{U}(p)$ for all p. Therefore $V(p) \geq \overline{U}(p)$ for all p.

(a) Suppose that $\underline{V}_{ct}(p^*) \geq \overline{V}_{ct}(p^*)$. The proof of the opposite case follows from a symmetric argument. It is easy to check that $V'_{cf}(p) > U'_b(p)$ if $p \leq \underline{p}^*$ and $V'_{cf}(p) < U'_a(p)$ for $p \geq \overline{p}^*$. Since $V_{ct}(p) = U_b(p)$ for $p \leq \underline{p}^*$ and $V_{ct}(p) = U_b(p)$ for $p \geq \overline{p}^*$, it therefore suffices to show that $V_{cf}(p) \leq V_{ct}(p)$ for $p \in [\underline{p}^*, \overline{p}^*]$. We distinguish two cases.

Case A: $p^* \geq \underline{p}^*$. Since $\underline{V}_{ct}(p^*) \geq \overline{V}_{ct}(p^*)$, Lemma 7 implies that $p^* < \overline{p}^*$. For $p \in [p^*, \overline{p}^*]$, $\underline{V}_{ct}(p)$ and $V_{cf}(p)$ satisfy (13). $V_{ct}(p^*) \geq V_{cf}(p^*)$ therefore implies $\underline{V}_{ct}(p) \geq V_{cf}(p)$

for $p \in [p^*, \overline{p}^*]$. For $p \in [\underline{p}^*, p^*]$, $V_{cf}(p)$ satisfies (12). Lemma 2 thus implies that it cannot cross $\underline{V}_{ct}(p)$ at any $p \in [\underline{p}^*, p^*]$ and hence $\underline{V}_{ct}(p) \geq \underline{V}_{cf}(p)$ for all $p \in [\underline{p}^*, \overline{p}^*]$. This proves the result for Case A.

Case B: $p^* < \underline{p}^*$. Since $V_{cf}(\underline{p}^*) = \overline{U}(\underline{p}^*)$ we have $V_{cf}(\underline{p}^*) < \widehat{U}(\underline{p}^*)$. For $p > p^*$, $V_{cf}(p)$ satisfies (13). By Lemma 6, we therefore have $V_{cf}(\underline{p}^*) < \widehat{U}(\underline{p}^*)$. By Lemma 4 $\widehat{U}(\underline{p}^*) = \underline{V}_{ct}(\underline{p}^*)$ and hence $V_{cf}(\underline{p}^*) < \underline{V}_{ct}(\underline{p}^*)$. Since $\underline{V}_{ct}(p)$ also satisfies (13) for all $p \geq \underline{p}^*$, we must therefore have $\underline{V}_{ct}(p) > V_{cf}(p)$ for all $p \geq \underline{p}^*$. This proves the result.

(b) Since $V_{cf}(p^*) = \overline{U}(p^*) < \widehat{U}(p^*)$ and $V_{ct}(p) \geq \widehat{U}(p)$ for all $p \notin (\underline{p}^*, \overline{p}^*)$, we have $p^* \in (\underline{p}^*, \overline{p}^*)$. By Lemma 6, $V_{cf}(\underline{p}^*) < \widehat{U}(\underline{p}^*) = \underline{V}_{ct}(\underline{p}^*)$. Therefore $V_{cf}(p)$ and $\underline{V}_{ct}(p)$ intersect at some $\underline{p} \in (\underline{p}^*, p^*)$ and by Lemma 2, the intersection is unique. Moreover, for $p < p^*$, we have $V_{cf}(p) > \overline{V}_{ct}(p)$ and hence $\overline{V}_{ct}(p) < \underline{V}_{ct}(p)$ for all $p \in (\underline{p}^*, \underline{p})$. This proves the result for $p \leq p^*$. For $p > p^*$ the arguments are symmetric. $\qquad\square$

We are now ready characterize the value function of the DM's problem in (6).

Proposition 7. *Suppose* (NSEC) *holds. Then* $V(p) = \max\{V_{ct}(p), V_{cf}(p)\}$ *is the value function in* (6).

Proof of Theorem 7. Theorem III.4.11 in Bardi and Capuzzo-Dolcetta (1997) characterizes the value function of a dynamic programming problem with an optimal stopping decision as in (6) as the (unique) viscosity solution of the corresponding HJB equation.[56] For our problem this requires that for all $p \in (0,1)$ where $V(p)$ is differentiable, it satisfies (7). We start by considering points of differentiability where $V(p) = U(p)$, i.e., $p \notin (\underline{p}^*, \overline{p}^*)$.[57] If $p < \hat{p}$, the maximand in (8) with $\alpha = 1$ is less than or equal than $c + rU(p)$ since otherwise $\underline{p}^* < p$. Also, if $p < \hat{p}$, the maximand in (8) with $\alpha = 0$ is $\lambda^B(p)\left(U_b(q^B(p)) - U_b(p)\right) + \delta p(1-p)U_b'(p) = 0$. Hence (7) holds for all $p < \hat{p}$, $p \notin (\underline{p}^*, \overline{p}^*)$. A similar arguments shows that (7) holds for all $p > \hat{p}$, $p \notin (\underline{p}^*, \overline{p}^*)$. We have shown that $V(p)$ satisfies (7) for all $p \notin (\underline{p}^*, \overline{p}^*)$ where it is differentiable. In particular, if (EXP) is violated (7) holds for all $p \neq \hat{p}$.

Next, we consider points of differentiability $p \in (\underline{p}^*, \overline{p}^*)$. $(\underline{p}^*, \overline{p}^*) \neq \emptyset$ requires that (EXP) is satisfied. Therefore, by Proposition 6, we have $V(p) \geq \overline{U}(p)$ for all $p \in (\underline{p}^*, \overline{p}^*)$. In this case (7) is equivalent to (9). Since $V(p)$ satisfies (12) or (13) at points of differentiability, Lemma 3 implies that $V(p)$ satisfies (9).

So far we have shown that $V(p)$ satisfies (7) for all points of differentiability. For $V(p)$ to be a viscosity solution of (7) we must show that for all points of non-differentiability, we have

$$\max\left\{-c - rV(p) + F(p, V(p), \rho), U(p) - V(p)\right\} \leq 0, \tag{22}$$

---

[56]To formally apply their theorem, we have to us P_t as a second state-variable and define a value function $v(p, P) = PV(p)$. Since v is continuously differentiable in P, it is straightforward to apply the result directly to $V(p)$.

[57]If $\underline{p}^* = 1$ and $\overline{p}^* = 0$ we set $(\underline{p}^*, \overline{p}^*) = \emptyset$ so that $p \in [0,1]$.

for all $\rho \in \left[V'_-(p), V'_+(p)\right]$; and the opposite inequality must hold for all $\rho \in \left[V'_+(p), V'_-(p)\right]$. Discontinuities arise at $\hat{p}$ if (EXP) is violated; at $\check{p}$ if (EXP) is satisfied and $V_{ct}(p^*) \geq V_{cf}(p^*)$; and at $\underline{p}$ and $\overline{p}$ if (EXP) is satisfied and $V_{ct}(p^*) < V_{cf}(p^*)$. By Lemma 2, $V(p)$ has a convex kink at all these points so that $V'_-(p) \leq V'_+(p)$. Therefore it suffices to check (22) for all $\rho \in \left[V'_-(p), V'_+(p)\right]$. The maximand in F is linear in α (see (8)), so it suffices to consider $\alpha \in \{0,1\}$. For $\alpha = 1$ we have

$$\lambda^A(p) \left(U(q^A(p)) - V(p)\right) - \delta p(1-p)\rho \leq \lambda^A(p) \left(U(q^A(p)) - V(p)\right) - \delta p(1-p)V'_-(p),$$

and for $\alpha = 0$ we have

$$\lambda^B(p) \left(U(q^B(p)) - V(p)\right) + \delta p(1-p)\rho \leq \lambda^B(p) \left(U(q^B(p)) - V(p)\right) + \delta p(1-p)\rho V'_+(p).$$

Therefore if $U(p) \leq V(p)$,

$$c + rV(p) \geq \max \begin{cases} \lambda^A(p) \left(U(q^A(p)) - V(p)\right) - \delta p(1-p)V'_-(p) \\ \lambda^B(p) \left(U(q^B(p)) - V(p)\right) + \delta p(1-p)V'_+(p) \end{cases} \tag{23}$$

implies that (22) holds for all for $\rho \in \left\{V'_-(p), V'_+(p)\right\}$. We distinguish three cases.

Case A: (EXP) is violated. In this case we only have to consider $p = \hat{p}$. We have $V(\hat{p}) = U_a(\hat{p}) = U_b(\hat{p})$, $V'_-(\hat{p}) = U'_b(\hat{p})$ and $V'_+(p) = U'_a(\hat{p})$. Now suppose by contradiction that (23) is violated at $\hat{p}$. Suppose that (the other case is similar)

$$c + rU_b(\hat{p}) < \lambda^A(\hat{p}) \left(U_a(q^A(\hat{p})) - U_b(\hat{p})\right) - \delta\hat{p}(1-\hat{p})U'_b(\hat{p}).$$

This implies $\underline{p}^* < 1$ which is a contradiction. Therefore $V(p)$ is a viscosity solution of (7) if (EXP) is violated.

Case B: (EXP) is satisfied and $V_{ct}(p^*) \geq V_{cf}(p^*)$. We have to consider $p = \check{p}$. (23) becomes

$$c + r\underline{V}_{ct}(\check{p}) = c + r\overline{V}_{ct}(\check{p}) \geq \max \begin{cases} \lambda^A(\check{p}) \left(U(q^A(\check{p})) - \underline{V}_{ct}(\check{p})\right) - \delta\check{p}(1-\check{p})\underline{V}'_{ct}(\check{p}) \\ \lambda^B(\check{p}) \left(U(q^B(\check{p})) - \overline{V}_{ct}(\check{p})\right) + \delta\check{p}(1-\check{p})\overline{V}'_{ct}(\check{p}) \end{cases}$$

This holds with equality since $\underline{V}_{ct}(p)$ satisfies (13) and $\overline{V}_{ct}(p)$ satisfies (12) at $\check{p}$. Since $V(\check{p}) > U(\check{p})$, (22) is thus satisfied at $\check{p}$.

Case C: (EXP) is satisfied and $V_{ct}(p^*) < V_{cf}(p^*)$. The proof is similar as in Case B.

We have thus shown that $V(p)$ is a viscosity solution of (7) which is sufficient for $V(p)$ to be the value function of problem (6). $\qquad\square$

A.2 Proof of Theorem 1

The main step of the proof of Theorem 1 is Proposition 7 in the previous section. It remains to show the existence of cutoffs $\underline{c}(r)$ and $\overline{c}(r)$ that determine the structure of the optimal solution. To simplify the notation we normalize c and r by λ and set $\zeta = c/\lambda$ and $\theta = r/\lambda$. We show that Theorem 1 holds with the cutoffs

$$\frac{\underline{c}(r)}{\lambda} = \underline{\zeta}(\theta) = \begin{cases} \min\left\{ \frac{(1+\theta)}{1+(1+2\theta)^{\frac{1}{\theta}}}\left(u_a^A - u_b^A\right) - \theta u_a^A, \frac{(1+\theta)}{1+(1+2\theta)^{\frac{1}{\theta}}}\left(u_b^B - u_a^B\right) - \theta u_b^B \right\} & \text{if } \theta > 0, \\ \min\left\{ \left(u_a^A - u_b^A\right), \left(u_b^B - u_a^B\right) \right\} / (1 + e^2) & \text{if } \theta = 0. \end{cases}$$

(24)

$$\frac{\overline{c}(r)}{\lambda} = \overline{\zeta}(\theta) = \frac{\left(u_a^A - u_b^A\right)\left(u_b^B - u_a^B\right) - \theta\left(u_a^A u_b^B - u_b^A u_a^B\right)}{\left(u_a^A + u_b^B\right) - \left(u_b^A + u_a^B\right)}$$

(25)

Proof of Theorem 1. By Proposition 7, no experimentation is optimal if (EXP) is violated. In the baseline model, this simplifies to

$$\zeta\left(\left(u_a^A + u_b^B\right) - \left(u_b^A + u_a^B\right)\right) + \theta\left(u_a^A u_b^B - u_b^A u_a^B\right) \geq \left(u_a^A - u_b^A\right)\left(u_b^B - u_a^B\right).$$

This is equivalent to $\zeta \leq \overline{\zeta}(r)$ which proves part (a).

Conversely if $\zeta \geq \overline{\zeta}(r)$, (EXP) is satisfied. Next we show that

$$\zeta \geq \underline{\zeta}(\theta) \quad \iff \quad \max\left\{\underline{V}_{ct}(p^*), \overline{V}_{ct}(p^*)\right\} \geq \overline{U}(p^*).$$

(26)

By Proposition 6 and Lemma 3, this implies that the policies stated in Parts (b) and (c) of Theorem 1 are optimal.

We first assume $\theta > 0$. Some algebra yields the following expressions

$$\underline{V}_{ct}(p^*) = \frac{1}{\theta(1+\theta)}\left(\frac{\left(\zeta + u_b^B\theta\right)\left(\zeta + u_a^A\theta\right)}{\left(\zeta + u_b^B\theta\right) + \left(\zeta + u_a^A\theta\right)}\left(\frac{\zeta + \theta u_a^A}{u_a^A - (1+\theta)u_b^A - \zeta}\right)^{\theta} + \frac{u_a^A u_b^B\theta^2 - \zeta^2}{\left(\zeta + u_b^B\theta\right) + \left(\zeta + u_a^A\theta\right)} - \zeta\theta\right)$$

$$\overline{V}_{ct}(p^*) = \frac{1}{\theta(1+\theta)}\left(\frac{\left(\zeta + u_b^B\theta\right)\left(\zeta + u_a^A\theta\right)}{\left(\zeta + u_b^B\theta\right) + \left(\zeta + u_a^A\theta\right)}\left(\frac{\zeta + \theta u_b^B}{u_b^B - (1+\theta)u_a^B - \zeta}\right)^{\theta} + \frac{u_a^A u_b^B\theta^2 - \zeta^2}{\left(\zeta + u_b^B\theta\right) + \left(\zeta + u_a^A\theta\right)} - \zeta\theta\right)$$

With some further algebra we obtain

$$\max\left\{\underline{V}_{ct}(p^*), \overline{V}_{ct}(p^*)\right\} \geq \overline{U}(p^*)$$

$$\iff \quad \left(\max\left\{\frac{\zeta + \theta u_a^A}{u_a^A - (1+\theta)u_b^A - \zeta}, \frac{\zeta + \theta u_b^B}{u_b^B - (1+\theta)u_a^B - \zeta}\right\}\right)^{\theta} \geq \frac{1}{1 + 2\theta}$$

$$\iff \quad \min\left\{\frac{(1+\theta)\left(u_a^A - u_b^A\right)}{1 + (1+2\theta)^{\frac{1}{\theta}}} - \theta u_a^A, \frac{(1+\theta)\left(u_b^B - u_a^B\right)}{1 + (1+2\theta)^{\frac{1}{\theta}}} - \theta u_b^B\right\} \leq \zeta.$$

Since the LHS in the last line is equal to $\underline{\zeta}(\theta)$, this proves (26) for $\theta > 0$. Taking the limit

$\theta \to 0$ yields the result for $\theta = 0$.[58]

Next we show that $\overline{\zeta}(\theta) > \underline{\zeta}(\theta)$. It is straightforward to verify that $\widehat{U}(p) > \overline{U}(p)$ for all p. Note that both functions are decreasing without bound in c. This means that for any u_x^ω, $r \geq 0$ and $\lambda > 0$, there exists $c > 0$ such that $\widehat{U}(\hat{p}) > 0$ and $\overline{U}(p) < 0$ for all p. In this case, the confirmatory solution cannot be part of the optimal strategy since it is dominated by immediate action. Conversely, since (EXP) is satisfied $V(p) \neq U(p)$ for some p. This contradicts $\overline{\zeta}(\theta) \leq \underline{\zeta}(\theta)$ and hence we have $\overline{\zeta}(\theta) > \underline{\zeta}(\theta)$. $\qquad\square$

A.3   Proof of Theorem 2

Proof of Theorem 2. Suppose SEP holds. (EXP) yields defines a cutoff $\overline{c}$ such that $V(p) = U(p)$ for all p if and only if $c > \overline{c}$ by the same argument as in the proof of Theorem 1. Proposition 7 shows that the the optimal strategy must take one of the forms 4 and 5. As in the proof of Theorem 1 we have that for c close to $\overline{c}$, $\overline{U}(p^*) < U(p^*)$ and $\widehat{U}(\hat{p}) > U(\hat{p})$ so that $V(p) = V_{ct}(p)$ and the optimal policy is of the form 4. Finally, as $\underline{\lambda} \to 0$, V_{cf} and V_{ct} converge to the values in the baseline model and if r is sufficiently small, $\underline{c}^{SEP} < \underline{c}$ so that the optimal policy is of the form 5.

Simple Algebra yields that (NSEC) is equivalent to

$$C_0 + C_1 c - \delta c^2 \leq 0, \tag{27}$$

where

$$C_0 = \left(-\overline{\lambda} u_a^A u_b^B + \underline{\lambda} u_b^A u_a^B\right) r^2 - 2(u_a^A u_b^B - u_b^A u_a^B)\overline{\lambda}\underline{\lambda}r - \delta\overline{\lambda}\underline{\lambda}\Delta^A\Delta^B,$$
$$C_1 = -r\left[\left(u_a^A + u_b^B\right)\overline{\lambda} + \left(u_a^B + u_b^A\right)\underline{\lambda}\right] - 2(\Delta^A - \Delta^B)\overline{\lambda}\underline{\lambda} < 0.$$

Since the LHS is decreasing in c, there exists a cutoff $\underline{c}^{SEP}$ such that (NSEC) is equivalent to $c > \underline{c}^{SEP}$. If $\underline{\lambda} = 0$, C_0 simplifies to $-r\overline{\lambda}u_a^A u_b^B < 0$ so $\underline{c}^{SEP} = 0$ for $\underline{\lambda}$ in a neighborhood of zero. Proposition 8 in Appendix A.4 in the Supplemental Material shows that SEP holds if and only if (NSEC) is satisfied. $\qquad\square$

References

ARROW, K. J., D. BLACKWELL, AND M. A. GIRSHICK (1949): "Bayes and Minimax Solutions of Sequential Decision Problems," *Econometrica*, pp. 213–244. 2

BARDI, M., AND I. CAPUZZO-DOLCETTA (1997): *Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations.* Birhäuser. 15, 35

---

[58]This can also be obtained directly by solving the ODEs for $\theta = 0$ to obtain $\underline{V}_{ct}(p^*)$ and $\overline{V}_{ct}(p^*)$.

BURKE, J. (2008): "Primetime Spin: Media Bias and Belief Confirming Information," *Journal of Economics and Management Strategy*, 17(3), 633–665. 18

CHAN, J., A. LIZZERI, W. SUEN, AND L. YARIV (2016): "Deliberating Collective Decisions," *Review of Economic Studies*, forthcoming. 5, 16

DAMIANO, E., H. LI, AND W. SUEN (2017): "Learning while Experimenting," unpublished. 5

DRUGOWITSCH, J., R. MORENO-BOTE, A. CHURCHLAND, M. SHADLEN, AND A. POUGET (2012): "The Cost of Accumulating Evidence in Perceptual Decision Making," *Journal of Neuroscience*, 32, 3612?–3628. 26

FRANCETICH, A. (2014): "Managing Multiple Research Projects," unpublished manuscript, Department of Decision Sciences and IGIER, Bocconi University. 5

FUDENBERG, D., P. STRACK, AND T. STRZALECKI (2017): "Speed, Accuracy, and the Optimal Timing of Choices," Available at SSRN 2602927. 4, 5, 6, 24, 26

GENTZKOW, M., AND J. SHAPIRO (2008): "Competition and Truth in the Market for News," *Journal of Economic Perspective*, 22, 133–154. 18

HAWKINS, G., B. FORSTMANN, E.-J. WAGENMAKERS, R. RATCLIFF, AND S. BROWN (2015): "Revisiting the Evidence for Collapsing Boundaries and Urgency Signals in Perceptual Decision-Making," *Journal of Neuroscience*, 35, 2476–2484. 26

HENRY, E., AND M. OTTAVIANI (2017): "Research and Approval Process: The Organization of Persuasion," *Economic Journal*, forthcoming. 5

KE, T. T., AND J. M. VILLAS-BOAS (2016): "Optimal Learning before Choice," unpublished manuscript, MIT. 6

KELLER, G., S. RADY, AND M. CRIPPS (2005): "Strategic Experimentation with Exponential Bandits," *Econometrica*, 73(1), 39–68. 5

KLEIN, N., AND S. RADY (2011): "Negatively Correlated Bandits," *The Review of Economic Studies*, 78(2), 693–732. 5

LATIMER, K., J. YATES, M. MEISTER, A. HUK, AND J. PILLOW (2015): "Single-trial spike trains in parietal cortex reveal discrete steps during decision-making," *Science*, 349, 184–187. 26

MATEJKA, F., AND A. MCKAY (2015): "Rational Inattention to Discrete Choices: A New Foundation for the Multinomial Logit Model," *The American Economic Review*, 105(1), 272–298. 6, 24

MAYSKAYA, T. (2016): "Dynamic Choice of Information Sources," unpublished manuscript, California Institute of Technology. 6

MEYER, M. A. (1991): "Learning from Coarse Information: Biased Contests and Career Profiles," *Review of Economic Studies*, 58(1), 15–41. 6

MOSCARINI, G., AND L. SMITH (2001): "The Optimal Level of Experimentation," *Econometrica*, 69(6), 1629–1644. 5

MULLAINATHAN, S., AND A. SHLEIFER (2005): "The Market for News," *American Economic Review*, 95(4), 1031–1053. 18

NIKANDROVA, A., AND R. PANCS (2017): "Dynamic Project Selection," *Theoretical Economics*, forthcoming. 5

OKSENDAL, B., AND A. SULEM (2009): *Applied Stochastic Control of Jump Diffusions.* Springer, 3rd edition edn.

OLIVEROS, S. (2015): "Demand for Slant: How Abstention Shapes Voters' Choice of News Media," *Economic Journal*, 125(587), 1327–1368. 18

PEREGO, J., AND S. YUKSEL (2016): "Media Competition and the Source of Disagreement," unpublished manuscript, NYU and UCSB. 18

PERSICO, N. (2004): "Committee Design with Endogenous Information," *Review of Economic Studies*, 71, 165–191. 16

PESKIR, G., AND A. SHIRYAEV (2006): *Optimal Stopping and Free-Boundary Problems.* Birkhäuser Verlag. 5, 24, 25

RATCLIFF, R., AND G. MCKOON (2008): "The Diffusion Decision Model: Theory and Data for Two-choice Decision Tasks," *Neural computation*, 20(4), 873–922. 5, 26

SIMS, C. A. (2003): "Implications of Rational Inattention," *Journal of Monetary Economics*, 50(3), 665–690. 6

SONER, H. M. (1986): "Optimal Control with State Space Constraint II," *SIAM Journal of Control and Optimization*, 24(6), 1110–1122.

STEINER, J., C. STEWART, AND F. MATEJKA (2015): "Rational Inattention Dynamics: Inertia and Delay in Decision-Making," Discussion paper, CEPR Discussion Papers. 6

STEPHENSON, M. C. (2010): "Information Acquisition and Institutional Design," *Harvard Law Review*, 124, 1422–1483. 16

SUEN, W. (2004): "The Self-Perpetuation of Biased Beliefs," *The Economic Journal*, 114(495), 337–396. 18

WALD, A. (1947): "Foundations of a General Theory of Sequential Decision Functions," *Econometrica*, pp. 279–313. 2

WOODFORD, M. (2016): "Optimal Evidence Accumulation and Stochastic Choice," unpublished manuscript, Columbia University. 6

YOON, Y. (2016): "Why Are Information Senders Biased? Indirect Demand for Information Bias," unpublished manuscript, Nanjing Audit University. 18

ZHONG, W. (2017): "Optimal Dynamic Information Acquisition," unpublished manuscript, Columbia University. 5