Optimal Dynamic Allocation of Attention*

Yeon-Koo Che   Konrad Mierendorff

This draft: December 15, 2018
First draft: April 18, 2016

Abstract

We consider a decision maker (DM) who, before taking an action, seeks information by allocating her limited attention dynamically over different news sources that are biased toward alternative actions. Endogenous choice of information generates rich dynamics: The chosen news source either reinforces or weakens the prior, shaping subsequent attention choices, belief updating, and the final action. The DM adopts a learning strategy biased toward the current belief when the belief is extreme and against that belief when it is moderate. Applied to consumption of news media, observed behavior exhibits an "echo-chamber" effect for partisan voters and a novel "anti echo-chamber" effect for moderates.

KEYWORDS: Wald sequential decision problem, choice of information, own-biased and opposite-biased learning strategies, limited attention.

1 Introduction

Information is central to decision making. Individuals, firms, and government agencies often expend significant resources to evaluate their choices in consumption, investment, and public projects. This is particularly so in high-stakes deliberation; e.g., when a voter deliberates on alternative candidates, when a researcher investigates a hypothesis, or when a judge or a juror weighs a defendant's guilt in a criminal case. In such situations and others, decision makers have access to different "news" sources or diverse views on their

*We thank Martino Bardi, Guy Barles, Hector Chade, Martin Cripps, Mark Dean, Mira Frick, Duarte Goncalves, Johannes Hörner, Han Huynh, Philippe Jehiel, Ian Krajbich, Botond Kőszegi, Suehyun Kwon, George Mailath, Filip Matějka, Pietro Ortoleva, Andrea Prat, Ed Schlee, Sara Shahanaghi, Philipp Strack, Tomasz Strzalecki, Glen Weyl, Michael Woodford, Xingye Wu, Weijie Zhong, and audiences at numerous conferences and seminars for helpful comments; and Timur Abbyasov for excellent research assistance.
Che: Department of Economics, Columbia University, 420 W. 118th Street, 1029 IAB, New York, NY 10025, USA; email: yeonkooche@gmail.com; web: http://blogs.cuit.columbia.edu/yc2271/.
Mierendorff: Department of Economics, University College London, 30 Gordon Street, London WC1H 0AX, UK; email: k.mierendorff@ucl.ac.uk; web: http://www.homepages.ucl.ac.uk/∼uctpkmi/.

actions. For instance, voters may expose themselves to like-minded news channels or to opposite-minded ones. Jurors may hear adversarial lawyers advocating opposing views on the case. Individuals also make attention choices in their internal processes of deliberation and thinking. For instance, a researcher may expend efforts to either prove or disprove a hypothesis.

In this paper, we ask how a decision maker (DM) should allocate her limited attention across different news sources or different deliberation strategies dynamically over time, and how that process shapes her choice of action. An important aspect of this problem is how long the DM should search for information before stopping to take an action. This *stopping problem* has been studied extensively by many authors, starting with Wald (1947) and Arrow, Blackwell, and Girshick (1949). While sharing the premise that information is costly and takes time to arrive, the contribution of the current paper is to study how a DM allocates her attention over different *types* of news sources.

In our model the DM faces binary actions, r and ℓ, which are optimal in states R and L, respectively. The state is initially unknown, and the DM has a prior belief. At each point in time, the DM may stop and take an action which is irreversible, or she may acquire more information about the state. In the latter case, she incurs a flow cost and payoffs are discounted.

Information can be received from of two sources: an L-biased or an R-biased news source.[1] The L-biased news source always sends an L-signal in state L and sometimes also in state R. Otherwise, it sends an R-signal. Since the R-signal is sent only in state R, it fully reveals the state. Symmetrically, the R-biased news source is biased toward sending an R-signal, except that in state L it occasionally reveals the state to be L. At each point in time, the DM has a unit budget of attention to allocate between these two news sources, and she may "multi-home" by arbitrarily dividing her attention between the two sources. In our continuous time model, these two news sources reduce to two Poisson processes that each generate breakthrough news revealing one state. In the absence of a breakthrough, each source leads to continuous updating of the belief in the direction of the source's bias. We show that these Poisson processes can be justified as optimal within a class of experiments which encompass general non-conclusive Poisson signals.

The main trade-off in our model is the decision between news sources that are biased in favor of, or against one's current belief. We obtain a novel characterization of the optimal learning/attention strategy. 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) *own-biased learning*, and (iii) *opposite-biased learning*; and she never switches between these different modes of learning.

[1] We assume that the DM is Bayesian, so that a news source cannot systematically bias her belief. The term "bias" here refers to the frequency of a signal favoring one state, which will be clarified in detail later.

Immediate action is a simple strategy where the DM takes an optimal action given her prior without acquiring any information. **Own-biased learning** focuses attention on the news source that is biased toward the state that the DM finds *relatively likely*. An example is to focus on the R-biased news source when state R is relatively likely. Given this strategy, the DM will take action ℓ if breakthrough news reveals the state to be L. Otherwise, and more likely, the DM becomes more convinced of state R, which leads to further own-biased learning. Eventually, she becomes sufficiently certain that the state is R—her belief reaches a stopping boundary, and she chooses action r without fully learning the state. **Opposite-biased learning** focuses attention on the news source biased toward the state she finds *relatively unlikely*. An example is to focus initially on the L-biased news source when state R is relatively likely. When following this strategy, the DM becomes less confident about state R when no breakthrough news arrives. Eventually, she becomes so uncertain that she switches to a second phase where she divides her attention equally between the R-biased and the L-biased news source. She continues to acquire information and never stops until a breakthrough reveals the true state.

The optimality of the alternative learning strategies depends on the parameters as well as the DM's prior. In particular, the cost of information as measured by the flow cost and discounting is important. 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 own-biased learning when she is more uncertain. Finally, if information acquisition costs are low, immediate action is again optimal for extreme priors, and own-biased learning is optimal for less extreme priors. For very uncertain priors, however, opposite-biased 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 even without evidence, so further information acquisition is worth relatively little compared to the delay it causes. 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 belief space. This trade-off also explains which strategy is optimal inside the experimentation region. Opposite-biased 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, own-biased learning is likely to produce a quick decision, because when no breakthrough arrives, it takes only a finite period of time for the DM to reach the stopping boundary and take an action. The price of a quick decision is lower accuracy, since the DM sometimes takes an incorrect action when she reaches the stopping region without breakthrough news. When the DM is already quite certain, under own-biased learning the time needed to reach the stopping boundary is very short, and the higher accuracy of opposite-biased learning is less valuable. This explains why the DM chooses own-biased learning when she is more certain and opposite-biased learning only when she is more uncertain. An implication is that *a "skeptic" is more likely to make an accurate decision but with a longer delay than a "believer."* This prediction—particularly the dependence of decision accuracy on the prior beliefs—constitutes an important difference relative to the existing literature, as will be explained in Section 4.3.

Our model yields rich implications in terms of dynamic feedback between the DM's selective exposure to a news source and her belief updating. This feedback is particularly relevant for voters who consult media, say before an election. Our results imply that optimal media choice leads to an "echo-chamber effect," where voters with relatively extreme beliefs subscribe to own-biased media that are likely to reinforce their prior beliefs. With their beliefs reinforced, such voters further subscribe to own-biased news media, repeating the same process until they become sufficiently convinced. The resulting feedback loop results in polarization of beliefs. Interestingly, with sufficiently informative media, this effect is reversed for voters with moderate beliefs. They optimally seek opposite-biased outlets. As a result, they are likely to become more skeptical about their initial beliefs. With growing skepticism, such voters will eventually multi-home both types of media outlets until they receive conclusive news that leads them to make up their minds. Their behavior thus exhibits an "anti-echo chamber effect."

Relation to the Literature

Our model incorporates endogenous choice of information in an optimal stopping framework à la Wald (1947). Optimal stopping problems with exogenous information have been analyzed in a Poisson framework (see Peskir and Shiryaev, 2006, ch. VI), but economic applications have focused on drift-diffusion models (DDM) in which the signal follows a Brownian motion with a drift determined by the state. Moscarini and Smith (2001) extend the stopping problem by allowing for an endogenous choice of signal precision. Other applications of DDMs include Chan, Lizzeri, Suen, and Yariv (2018), Fudenberg, Strack, and Strzalecki (2017), Henry and Ottaviani (2017), and McClellan (2017). An exception is Nikandrova and Pancs (2018) who consider the problem of selectively learning about different investment projects in a Poisson framework. In their model, the payoffs of final actions are uncorrelated whereas our model assumes negatively correlated payoffs.[2]

Ultimately, whether a DDM or a Poisson model is appropriate depends on the specific application at hand.[3] From a theoretical point of view, Zhong (2017) justifies Poisson

[2]See also the recent paper by Mayskaya (2016), and Ke and Villas-Boas (2016), which is concurrent to our paper.

[3]DDMs, which lead to a continuous belief process, are more suitable for the problem of learning the properties of a data-generating process from a sequence of samples, as in clinical trials, or for the analysis of statistical information. Poisson models, which lead to discontinuous updating, are useful to model the discovery of individual pieces of information that are very informative, as is common in the political

learning as an optimal information choice in a model that also allows for learning from a diffusion process.[4]

Our model shares a common theme with the rational inattention model (henceforth, RI) introduced by Sims (2003) and further developed by Matejka and McKay (2015) and Steiner, Stewart, and Matějka (2017). Like our paper, the RI models explain individual choice as resulting from the optimal allocation of limited attention over diverse information. However, our model differs from the RI models in two respects. One feature of the RI model is that the same Blackwell experiment entails different costs for different beliefs. We do not allow for such belief-dependence of the information technology. Thus any dependence of information choice on beliefs arises from the DM's incentives, rather than the technology she faces. Second and more important, the RI model abstracts from the precise dynamic process of allocating one's attention. By contrast, our objective is to unpack the "black box" and explicitly characterize the DM's dynamic attention choice. Steiner, Stewart, and Matějka (2017) extend the RI model to a dynamic setting in which a DM chooses a sequence of multiple actions. However, their main objective is to characterize stochastic choice rather than the dynamic information acquisition, the main focus of our paper. As they acknowledge, the information acquisition strategy implementing the optimal stochastic choice is not uniquely pinned down (see p. 527 of Steiner, Stewart, and Matějka (2017)). Hebert and Woodford (2017) and Morris and Strack (2017) provide a link between RI and DDM models by showing that a class of static reduced-form cost functions, including RI, can be microfounded by dynamic DDM models.

Our model leads to rich predictions about the stochastic choice function that are not obtained in the DDM or RI framework. In the latter two, the accuracy of the DM's decision is independent of the prior (conditional on information acquisition). In our model, the accuracy varies with the prior.[5] There are two reasons for this difference. First, the endogenous choice of information sources leads to different modes of learning depending on the prior, which results in significant differences in the accuracy and delay. Second, in our Poisson model, the DM may reach a decision either following a breakthrough, or after the belief drifts to a stopping bound. In the DDM framework, all decisions are reached after drifting to a decision boundary.[6] (See Section 4.3 for further discussion.)

sphere, in criminal investigations, or when a scientist searches for a breakthrough insight that will prove or disprove a hypothesis.

[4]His model differs from ours in that it assumes a posterior separable cost function (see e.g. Caplin, Dean, and Leahy, 2018). Much like the rational inattention model described below, the cost of a Blackwell experiment depends on the DM's belief in his model. By contrast, feasible experiments and their costs do not depend on the DM's belief in our model. Due to this difference, our models are not directly comparable; his framework cannot be used to formulate our model (see also footnote 36 in Section 6.1).

[5]Note however, that this dichotomy is less clear in decision problems with a richer state-space such as Fudenberg, Strack, and Strzalecki (2017) and Ke and Villas-Boas (2016). A complete comparison of the implications of continuous versus discontinuous learning is beyond the current state of the literature.

[6]Note that we are comparing our model with discontinuous learning and endogenous information choice to a DDM model with continuous learning and exogenous information. In our two-state model, it is not possible to formulate a DDM with multiple information sources making it impossible to describe

The Poisson signal structure was introduced in bandit problems by Keller, Rady, and Cripps (2005). The negatively-correlated bandits model of Klein and Rady (2011) parallels the choice between two biased Poisson processes studied in the current paper. However, there are two fundamental differences. First, there is difference in timing. In our model, exploiting a payoff requires the DM to stop learning, whereas in bandit problems, she can exploit payoffs while learning. Second, in bandit models, information sources are linked to exploitation of specific arms. For these reasons, a distinct characterization emerges; for instance, there is no analogue to our "own-biased learning" in the bandit literature.[7]

Finally, the current paper is related to the media choice literature. It has been observed before that a Bayesian voter may find it optimal to consume news from a biased outlet (see Calvert (1985), Suen (2004), and Burke (2008)). In particular, Suen (2004) observes self-perpetuation and polarization of beliefs. As we elaborate later, however, these models consider very special cases that prevent nontrival dynamics from emerging.

The paper is organized as follows. Section 2 presents the model. Section 3 presents an example to illustrate the main results. Section 4 characterizes the optimal policy. Section 5 applies the model to media choice. Section 6 extends the model in several directions. Section 7 concludes. Proofs are defered to Appendix A and the Supplemental Material.

2 Model

States, Actions and Payoffs. A DM must choose from two *actions*, r or ℓ, whose payoffs depend on the unknown *state* $\omega \in \{R, L\}$. The *payoff* of taking action $x \in \{r, \ell\}$ 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, $u_r^R > \max\{0, u_\ell^R\}$ and $u_\ell^L > \max\{0, u_r^L\}$.[8] 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 $\rho \geq 0$. Either c or ρ may be zero, but not both.

The DM's *belief* is denoted by the probability $p \in [0, 1]$ that the state is R. 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 $U(p) := \max\{U_r(p), U_\ell(p)\}$, where $U_x(p) := pu_x^R + (1-p)u_x^L$ is the expected payoff of taking action x. $U(p)$ takes the piece-wise linear form depicted in Figure 4.2 on page 14.

---

separately the effect of discontinuous learning and endogenous information.

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

[8]Note that we allow for $u_x^R = u_x^L$ so that one action $x \in \{r, \ell\}$ can be a *safe action*. We rule out the trivial case in which $u_x^\omega \geq u_y^\omega$ for $x \neq y$, in both states $\omega = R, L$.

Information Acquisition and Attention. We model information acquisition in continuous time. At each point in time, the DM may allocate one unit of *attention* across an *L-biased* and an *R-biased* news source. The L-biased source sends Poisson breakthrough news only in state R with an arrival rate of $\lambda > 0$, and the R-biased source sends breakthrough news only in state L with an arrival rate of λ. Since, for a given source, a breakthrough arrives only in one state, it conclusively reveals the true state. In this sense, paying attention to the L-biased source say can be interpreted as "looking for R-evidence." Indeed, this interpretation describes some circumstances well and is hence useful to keep in mind. For instance, one can interpret this as a judge or juror focusing attention to evidence proving a certain state—guilt or innocence of a suspect, or a scientist seeking firm evidence that either proves or disproves a certain hypothesis. For the media choice application, however, interpreting news-sources according to their "biases" is more useful.

To get a better understanding of this interpretation, it is useful to study the *statistical experiments* induced by different attention choices. Suppose the DM pays full attention to the L-biased news source for a short duration $dt > 0$. Then, she receives one of two signals—breakthrough news ("signal σ_R^L"), or its absence ("signal σ_L^L"). Panel (a) of Figure 2.1 describes the probabilities of receiving these two signals in each state.

(a) L-biased experiment: σ^L

state/signal	σ_L^L	σ_R^L
L	1	0
R	$1-\lambda dt$	λdt

(b) R-biased experiment: σ^R

state/signal	σ_L^R	σ_R^R
L	λdt	$1-\lambda dt$
R	0	1

Figure 2.1: Experiments induced by two Poisson signals.

Experiment σ^L is "biased" toward L in the sense of sending the L-favoring signal σ_L^L excessively—always in state L but even in state R with some probability.[9] An implication of this is that σ_L^L is a relatively weak signal that moves the belief toward L but not by much. By contrast, the signal σ_R^L—the R-favoring signal from source σ^L—reveals conclusively that the state is R. Similarly, if the DM chooses the R-biased news source for a duration $dt > 0$, then this induces the experiment σ^R depicted in Panel (b). This experiment is biased toward R in the sense of sending the R-favoring signal excessively. While our model focuses on conclusive Poisson experiments, we show in Section 6.1 that this focus can be justified in a more general class of information technologies.

We assume that the DM may allocate any fraction $\alpha \in [0, 1]$ of her attention to L-biased news and the remaining fraction $\beta = 1 - \alpha$ to R-biased news. In this case, she receives R-evidence with arrival rate $\alpha\lambda$ in state R, and L-evidence with arrival rate $\beta\lambda$ in

[9] As mentioned in Footnote 1, we use the term "bias" only in this sense. This terminology is in keeping with the media choice literature as we will discuss in Section 6.1.

state L. An interior attention choice can be interpreted as "multi-homing"—or switching back and forth arbitrarily frequently—between the two news sources. We denote the DM's attention strategy by $(\alpha_t) = (\alpha_t)_{t\geq 0}$, and assume that α_t is a measurable function of t.

Suppose the DM uses the attention strategy (α_t). In the absence of a breakthrough, the DM's belief will evolve according to Bayes rule:[10]

$$\dot{p}_t = -\lambda(\alpha_t - \beta_t)p_t(1-p_t) = -\lambda(2\alpha_t - 1)p_t(1-p_t). \tag{2.1}$$

The bias of a news source corresponds to the direction of updating in the absence of breakthrough news. The higher α, the more L-biased is the mix of news sources that the DM pays attention to. For instance, full attention to L-biased news ($\alpha = 1$) makes the DM pessimistic about state R if no breakthrough arrives. Consequently, her belief drifts towards L. Finally, note that if the DM divides her attention equally between both news sources, she never updates her belief in case of no breakthrough news: $\dot{p}_t = 0$ if $\alpha_t = 1/2$.

3 Illustrative Examples

Before proceeding, we use simple examples to illustrate the main insights of our results. The examples will highlight the role of dynamics by considering one-period and two-period versions of our model.[11] For simplicity, assume the DM enjoys a payoff of 1 when matching the state and -1 otherwise ($u_r^R = u_\ell^L = 1$, and $u_\ell^R = u_r^L = -1$). Time is discrete with a period length of $dt = 1$, and the DM incurs a cost c per period if she acquires information. There is no discounting. In each period, the DM can either stop and take an irreversible action, or pay the cost c to observe one of the experiments σ^L and σ^R in Figure 2.1.

One-Period Problem. Suppose the DM can experiment only once before taking an action. For a prior p_0 sufficiently high or low, the DM will optimally choose to take an action immediately, since experimentation is costly. For less extreme priors p_0, she will experiment. The question is which news source she will pay attention to. We argue that it is optimal to pick the "own-biased news source"—namely the experiment biased toward one's prior: σ^R if $p_0 > 1/2$ and σ^L if $p_0 < 1/2$.

The reason is that own-biased news lowers the chance of mistakes compared to opposite-biased news. Specifically, suppose $p_0 > 1/2$, so that the prior indicates that action r is optimal, and assume that the DM picks the opposite-biased experiment, σ^L. There are

[10] Since the belief is a martingale, 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 (2.1).

[11] These examples are similar in spirit to Suen (2004) and Burke (2008), except that information is not costly in these models, which leads to a different characterization. In particular, dynamics has no significant effect in these models, in contrast to the point made here.

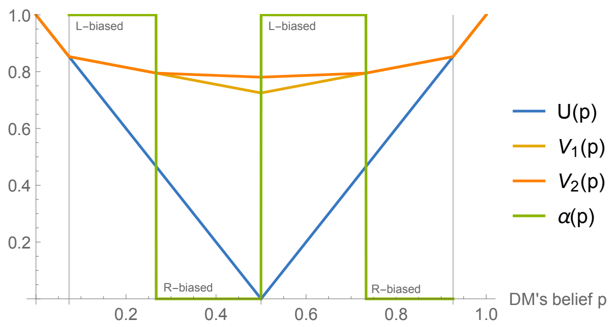


Figure 3.1: Optimal policy in two-period learning

Note: $\lambda = .85, c = .125$; $U(p)$ is the payoff from immediate action, $V_1(p)$ is the optimal payoff in the one-period problem, $V_2(p)$ is the optimal payoff in the two-period problem, and $\alpha(p)$ is the choice of experiment in the first (of two) periods, $\alpha = 1$ corresponds to σ^L and $\alpha = 0$ corresponds to σ^R.

two cases. Suppose first p_0 is very high. Then even when the DM receives the L-favoring signal σ^L_L, this will not move her belief enough to "change" her action to ℓ, meaning she will choose r regardless of the signal. Hence, σ^L is worthless to her. By contrast, a DM choosing σ^R will be "convinced" by an L-signal to change her action to ℓ, which makes σ^R valuable.

Next suppose that $p_0 > 1/2$ but lower, so that an L-signal from σ^L also leads the DM to choose ℓ. Why is σ^L still inferior to σ^R? The reason is that σ^L chooses ℓ as the "default action" which is taken unless conclusive R-evidence is observed. Since the prior $p_0 > 1/2$ favors state R, action r is a better default; so the own-biased experiment σ^R is better because it leads to action r unless conclusive L-evidence is observed.[12]

Two-Period Problem. Now suppose the DM can experiment for up to two periods. After observing the first signal, the DM may stop and take an action, or experiment for another period. The problem facing the DM in the second period is precisely the one-period problem above: Depending on her posterior belief after the first experiment, she will either take an immediate action, or choose an own-biased experiment. But what should she do in the first period?

[12] To see precisely why σ^R is more valuable than σ^L if $p_0 > 1/2$, we consider a (hypothetical) experiment S with three possible signals S_L, S_0, and S_R:

state/signal	S_L	S_0	S_R
L	λ	$1-\lambda$	0
R	0	$1-\lambda$	λ

Clearly, S is Blackwell more informative than σ^L and σ^R, respectively. σ^L can be obtained from S by observing only the partition $\{\{S_L, S_0\}, \{S_R\}\}$, and σ^R can be obtained by observing only the partition $\{\{S_L\}, \{S_0, S_R\}\}$. What is the optimal decision rule following experiment S? If $p_0 > 1/2$, the DM should choose action r if she observes a signal in $\{S_0, S_R\}$ and ℓ only if she observes S_L. Action r is the "default" chosen after the uninformative signal S_0. This corresponds to the decision implemented by the partition corresponding to σ^R. By contrast the partition corresponding to σ^L leads to action ℓ as the "wrong default," which reduces the value of this experiment compared to the value of σ^R (or of S).

Figure 3.1 illustrates a case where the optimal choice of experiment in the first period differs from that of the one-period problem. For extreme values of p_0, an immediate action is optimal as before. For less extreme priors ($p_0 \in [.07, .93]$), the DM chooses to experiment. When the belief is close to the boundaries of the experimentation region ($p_0 \in [.07, .27]$ or $p_0 \in [.73, .93]$), the DM chooses own-biased news in period one. Afterwards, she immediately takes an action without any further experimentation. The intuition is that when her belief is already quite extreme, the marginal value of increasing the accuracy for the decision is small, so she experiments only once. Conditional on experimenting once, she prefers own-biased news for the same reason as in the one-period problem.

If the belief is moderate ($p_0 \in [.27, .73]$), the DM picks an opposite-biased news source. For instance, for $p_0 = .7$, she chooses σ^L. Recall that this was not a good strategy in the one-period problem since an action had to be taken even after a "weak" L-signal. In the two-period problem, however, the DM can continue to experiment. This option value can make both opposite-biased and own-biased learning more attractive. However, for own-biased learning, the posterior after the first experiment becomes more extreme, which limits the option value so that it is optimal to stop after period one.[13] By contrast, opposite-biased learning leads to a second round of experimentation, so the option value is more significant in particular for moderate beliefs where the value of accuracy is high. This explains the optimality of the opposite-biased learning with longer delay for the moderate beliefs.

These examples suggest that dynamics matters for the optimal allocation of attention. Different learning strategies are optimal for different beliefs, and a learning strategy that is not myopically optimal becomes optimal when a DM can experiment for more than one period. The pattern of the optimal strategy as well as the insights illustrated in these examples generalize to our continuous time model, which we now turn to.

4 Analysis of the Optimal Strategy

We now analyze the DM's optimal strategy in the continuous time model introduced in Section 2.

4.1 Formulation of the Problem

The DM chooses an attention strategy $(\alpha_t) = (\alpha_t)_{t \geq 0}$, and a stopping time $T \in [0, \infty]$ at which a decision will be made if no breakthrough news is received by then.[14] Her problem

[13] It is in fact optimal to stop after period one for $p_0 > .65$. We focus on this range of beliefs because it illustrates the difference between a static and a dynamic model most clearly.

[14] Given the linearity of the arrival rates in α, the DM cannot benefit from randomization in the continuous time model. For this reason, we only consider deterministic strategies $(\alpha_t)_{t \in \mathbb{R}_+}$. Moreover, it suffices to consider strategies that specify α as a function of t since the attention choice at time t is only relevant if the DM has not received any conclusive signals until time t.

is thus given by

$$V^*(p_0) = \max_{(\alpha_t),T} \int_0^T e^{-\rho t} P_t \left[p_t \lambda \alpha_t u_r^R + (1-p_t)\lambda \beta_t u_\ell^L - c\right] dt + e^{-\rho T} P_T U(p_T), \tag{4.1}$$

where p_t satisfies (2.1), $\beta_t = 1 - \alpha_t$, and $P_t = p_0 e^{-\lambda \int_0^t \alpha_s ds} + (1-p_0)e^{-\lambda \int_0^t \beta_s ds}$ is the probability that no signal is received by time t given strategy (α_τ). The integrand in the objective function captures the payoffs from taking an action following discovery of evidence, and the flow cost incurred until the DM stops. Specifically, at each time t, conditional on no discovery so far (which occurs with probability P_t), the strategy α_t leads to discovery of R-evidence with probability $p_t \lambda \alpha_t$, and of L-evidence with probability $(1-p_t)\lambda\beta_t$, per unit time. The second term accounts for the payoff from the optimal decision in case of no discovery by T.[15]

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

$$c + \rho V(p) = \max_{\alpha \in [0,1]} \left\{ \begin{array}{c} \lambda \alpha p \left(u_r^R - V(p)\right) + \lambda(1-\alpha)(1-p)\left(u_\ell^L - V(p)\right) \\ -\lambda(2\alpha - 1)p(1-p)V'(p) \end{array} \right\}, \tag{4.2}$$

if $V(p) > U(p)$. If $V(p) = U(p)$, the LHS of (4.2) should be no less than the RHS—in this case $T(p) = 0$ and immediate action is optimal. The objective in (4.2) is linear in α, which implies that the optimal policy is a bang-bang solution $\alpha^*(p) \in \{0,1\}$, except when the derivative of the objective vanishes. While this observation narrows down our search, there is a large class of strategies consistent with bang-bang choices. Ultimately, one needs to characterize the attention choice for each belief, which we now turn to.

4.2 Learning Heuristics

We begin with several intuitive learning heuristics that the DM could employ. These heuristics form basic building blocks for the DM's optimal strategy. Specifically, it will be seen that for each prior, the optimal strategy employs the heuristic with the highest value and never switches even after the belief has moved away from the prior. The details of the formal construction are presented in Appendix A.

Immediate action (without learning). A simple strategy 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.

[15] For a given $(\alpha_t)_{t\in\mathbb{R}_+}$, conditional on no discovery, the posterior belief evolves according to a deterministic rule (2.1). Since stopping matters only when there is no discovery, it is without loss to focus on a *deterministic* stopping time T.

Own-Biased Learning. When the DM decides to experiment, one natural strategy is to focus attention on the news source that is biased toward the more likely state. Formally, own-biased learning prescribes

$$\alpha(p) = \begin{cases} 1 & \text{if } p \in \left(\underline{p}^*, \check{p}\right), \\ 0 & \text{if } p \in [\check{p}, \overline{p}^*), \end{cases} \tag{4.3}$$

for some reference belief $\check{p}$ and boundaries of the *experimentation region* $\left(\underline{p}^*, \overline{p}^*\right)$, which will each be chosen optimally.[16] For instance, if L is relatively likely, the DM chooses the L-biased news source, or equivalently looks for conclusive R-evidence. In the absence of such "contradictory" evidence, the DM's belief drifts toward L. The belief updating is illustrated by the direction of the arrows in Panel (a) of Figure 4.1. Eventually, the DM's belief will reach one of the boundary points $\underline{p}^*$ or $\overline{p}^*$, at which she is sufficiently certain to take an immediate action without conclusive evidence. Since contradictory evidence is unlikely, the DM adopting this strategy can be seen as seeking to gradually rule out the unlikely state. For example, a juror or a judge sympathetic to a defendant's innocence may try to rule out incriminating evidence by actively looking for it, or a mathematician convinced of her proof may try to rule out "errors" by actively searching for them.

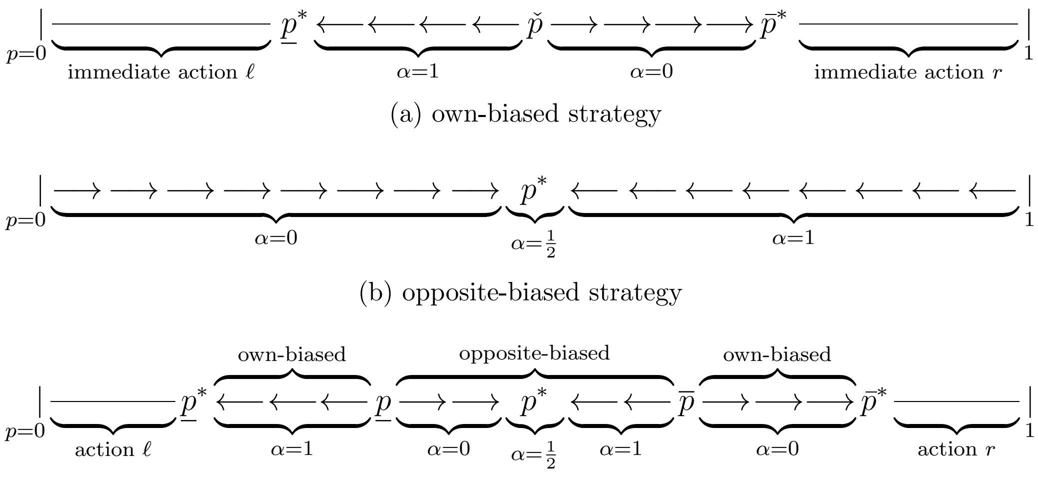
Opposite-Biased Learning. Alternatively, the DM could focus attention on a news source biased toward the state she finds relatively unlikely. Formally, opposite-biased learning prescribes:

$$\alpha(p) = \begin{cases} 0 & \text{if } p < p^*, \\ \frac{1}{2} & \text{if } p = p^*, \\ 1 & \text{if } p > p^*, \end{cases} \tag{4.4}$$

for some reference belief $p^* \in (0, 1)$, which will be chosen optimally.[17] At the optimal p^*, the DM is indifferent between all $\alpha \in [0, 1]$, so that the bang-bang result from the previous section does not apply. For beliefs below p^*, the DM subscribes to the R-biased news source, or equivalently, she seeks "confirmatory" evidence supporting the likely state L. An example of such a strategy could be that a mathematician tries to prove a promising hypothesis and to disprove unpromising one. In the absence of such affirmative proof, the DM becomes more uncertain in her belief, which thus drifts inward. The belief updating is illustrated in Panel (b) of Figure 4.1, with arrows indicating the direction of Bayesian updating. As is clear from equation (2.1) and (4.4), her belief drifts from both extremes towards the absorbing point p^*. Once p^* is reached, the DM divides her attention equally

[16] When payoffs are symmetric between ℓ and r, then $\check{p}$ equals 1/2. In general, $\check{p}$ may not equal 1/2.

[17] In our notation, an asterisk $(*)$ indicates an absorbing point, as in $p^*, \underline{p}^*, \overline{p}^*$. Cutoff beliefs without an asterisk indicate points were the belief diverges (e.g. $\check{p}$, $\underline{p}$, $\overline{p}$). An overline (underline) is used to denote cutoff beliefs to the right (left).



(a) own-biased strategy

(b) opposite-biased strategy

(c) combination of own-biased and opposite-biased strategies

Figure 4.1: Structure of Heuristic Strategies and Optimal Solution.

between both news sources. In this case, no further updating occurs, and she repeats the same strategy until she obtains evidence that reveals the true state.

4.3 Optimal Strategy

The structure of the optimal policy depends on the *cost of information* c. Intuitively, the higher the flow cost, the lower the net 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 theorem shows that there are three cases. If c is very high, immediate action is always optimal (case (a)). For intermediate values of c, the optimal strategy involves only own-biased learning (case (b)). For low values of c, both own-biased and opposite-biased learning occur (case (c)). For the theorem, we denote the optimal immediate action by $x^*(p) \in \arg\max_{x\in\{r,\ell\}} U_x(p)$, which is unique almost everywhere.

Theorem 1. *For given utilities u_x^ω, $\lambda > 0$, and $\rho \geq 0$, there exist $\overline{c} = \overline{c}(\rho, u_x^\omega, \lambda)$ and $\underline{c} = \underline{c}(\rho, u_x^\omega, \lambda)$, $\overline{c} \geq \underline{c} \geq 0$, with strict inequalities for ρ sufficiently small, such that the unique optimal strategy is characterized as follows:*[18]

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

*(b) (**Own-biased learning**) If $c \in [\underline{c}, \overline{c})$, there exist cutoffs $0 < \underline{p}^* < \check{p} < \overline{p}^* < 1$ such that for $p \in (\underline{p}^*, \overline{p}^*)$, the optimal policy $\alpha^*(p)$ is given by (4.3). If $p \notin (\underline{p}^*, \overline{p}^*)$, the DM takes action $x^*(p)$ without any information acquisition.*

*(c) (**Own-biased and Opposite-biased learning**) If $c < \underline{c}$, then there exist cutoffs $0 < \underline{p}^* < \underline{p} < p^* < \overline{p} < \overline{p}^* < 1$ such that for $p \in (\underline{p}^*, \overline{p}^*)$, the optimal policy is given*

[18] The strategy is unique up to tie breaking at the beliefs $\underline{p}^*, \overline{p}^*, \underline{p}, \overline{p}, \check{p}$. See (A.17) and (A.18) in Appendix A for explicit expressions for $\underline{c}$ and $\overline{c}$. See (A.9) and (A.10) for $\underline{p}^*$ and $\overline{p}^*$, and (A.13) for p^*.

by

$$\alpha^*(p) = \begin{cases} 1, & \textit{if } p \in \left(\underline{p}^*, \underline{p}\right), \\ 0, & \textit{if } p \in \left[\underline{p}, p^*\right), \\ \frac{1}{2} & \textit{if } p = p^*, \\ 1, & \textit{if } p \in \left(p^*, \overline{p}\right], \\ 0, & \textit{if } p \in \left(\overline{p}, \overline{p}^*\right). \end{cases} \tag{4.5}$$

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

In cases (b) and (c), 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; namely, an immediate action is chosen outside these boundaries.

The optimal policy in case (b) is depicted in Panel (a) of Figure 4.1 on page 13 above. In case (c) the pattern is more complex (see Panel (c) of Figure 4.1). Both own-biased and opposite-biased learning are optimal for some beliefs. Theorem 1 shows that the opposite-biased region $(\underline{p}, \overline{p})$ is always sandwiched between two regions where the own-biased learning strategy is employed. That is, near the boundaries of the experimentation region, the own-biased strategy is always optimal. Figure 4.2 shows the value of opposite-biased learning ($V_{opp}(p)$) and own-biased learning ($V_{own}(p)$) for high and low costs. It illustrates that the optimal policy picks the learning heuristic with the higher value.

The intuition behind the optimal strategy can be explained by a trade-off between speed and accuracy. The opposite-biased strategy leads to complete learning: The DM takes an action only if she receives conclusive evidence from breakthrough news. Therefore, she never makes a mistake. By contrast, own-biased learning may lead to mistakes if the DM's belief drifts to the stopping boundary ($\underline{p}^*$ or $\overline{p}^*$) before observing a breakthrough. In this case the DM stops without learning completely and may choose the wrong action. Therefore, opposite-biased learning has an accuracy advantage. At the

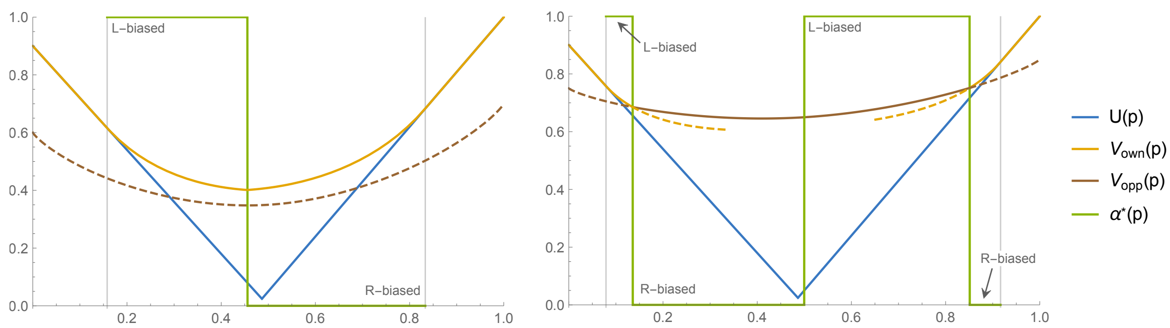


Figure 4.2: Value Function and Optimal Policy.

Note: The value function is the upper envelope of V_{own} and V_{opp} (solid). ($\lambda = 1$, $\rho = 0$, $u_r^R = 1$, $u_\ell^L = .9$, $u_\ell^R = u_r^L = -.9$)

same time, full learning under the opposite-biased strategy leads to a potentially long delay, because the DM has to wait for a breakthrough to arrive. By contrast, the delay in the own-biased strategy is limited by the time it takes the belief to reach the stopping boundary. Therefore own-biased learning has a speed advantage. This explains why the DM never uses the opposite-biased learning except when the cost of learning is sufficiently low, which makes the speed advantage less important.

The speed-accuracy tradeoff in the choice between the two strategies also explains why own-biased learning is always optimal near the stopping boundaries. Here, the speed advantage of own-biased learning is particularly large because it takes only a short period of time for the belief to reach the stopping boundary. At the same time, for beliefs near the stopping boundary, there is little uncertainty so that the value of full accuracy is relatively small. The DM therefore prefers speed over accuracy. For more uncertain beliefs, the value of learning the state is higher, making accuracy more important. At the same time, the speed advantage of own-biased learning is smaller because it takes longer for the belief to reach the stopping boundary. Therefore, the DM prefers the more accurate opposite-biased strategy for more uncertain beliefs.

Speed and Accuracy of Decisions under the Optimal Strategy. We have already noted intuitively that the speed and accuracy of the DM's decision depends on the mode of learning and her prior belief. Now we make some formal observations about the optimal strategy in line with this intuition.

Proposition 1 (Speed and Accuracy)**.** *Suppose that $c < \overline{c}$.*

(a) The average delay in the DM's decision is quasi-concave in the prior belief. If $c \geq \underline{c}$, the delay is maximal at $p_0 = \check{p}$. If $c < \underline{c}$, the delay is maximal at some p inside the opposite-biased learning region.

(b) The probability of a mistake is quasi-convex; it is zero when the initial belief is in the opposite-biased region. Within the own-biased region, the probability of a mistake is positive and increasing as p gets closer to a stopping boundary.

This proposition shows that initial beliefs matter greatly for the accuracy of a decision and its timing. A "skeptic"—a DM with uncertain initial belief—reaches a fully accurate decision but at the expense of a long delay. By contrast, a "believer"—a DM with a more extreme belief—sacrifices accuracy in favor of speed.[19] This feature stands in contrast to rational inattention and drift-diffusion models with non-shifting stopping boundaries. These latter models predict that the accuracy of a chosen action is invariant to the DM's initial beliefs, as long as they are within the experimentation region.[20]

[19]Around the boundary beliefs $\overline{p}$ and $\underline{p}$, the outcome in terms of accuracy and speed varies discontinuously, although the value is continuous.

[20]In a rational inattention model, an optimal experiment for a DM involves no more signals than the number of actions chosen; otherwise, she can lower her information cost by eliminating the "wasteful"

4.4 Comparative Statics

It is instructive to study how the optimal strategy varies with the parameters. We start by considering the experimentation region.

Proposition 2 (Comparative Statics: Boundaries of the Experimentation Region)**.**

(a) The experimentation region expands as ρ or c falls, and covers $(0,1)$ in the limit as $(\rho,c)\to(0,0)$.[21]

(b) The experimentation region expands as u_r^L or u_ℓ^R falls (so that "mistakes" become more costly), and covers $(0,1)$ in the limit as $(u_r^L,u_\ell^R)\to(-\infty,-\infty)$.[22]

(c) If $c<\overline{c}$, then the experimentation region shifts down as u_r^R increases and up as u_ℓ^L increases.

Parts (a) and (b) are quite intuitive. The DM acquires information for a wider range of beliefs if the cost of learning (ρ,c) falls, or if mistakes become more costly in the sense that (u_r^L,u_ℓ^R) falls. The intuition for (c) is twofold: an increase in u_r^R increases the value of a conclusive R-signal and thus causes $\underline{p}^*$ to fall. Further, if $\rho>0$, the immediate action r becomes more attractive, which causes $\overline{p}^*$ to shift down.

Next we consider how the payoffs from mistakes affect the cutoff beliefs that separate regions in which the DM employs different modes of learning.

Proposition 3 (Comparative Statics: Mode of Learning)**.**

(a) If $\underline{c}<c<\overline{c}$, then the cutoff $\check{p}$ in own-biased learning decreases if u_ℓ^R falls, and increases if u_r^L falls.

(b) For given c, the opposite-biased region appears ($\underline{c}>c$) for (u_r^L,u_ℓ^R) sufficiently small, and expands as (u_r^L,u_ℓ^R) falls. Given $\underline{c}>c$, $\underline{p}$ converges monotonically to zero as $u_\ell^R\to-\infty$, and $\overline{p}$ converges monotonically to one as $u_r^L\to-\infty$.

To get an intuition for Part (a), suppose $p<\check{p}$. In this case, own-biased learning may lead to taking action ℓ in the wrong state (state R). This mistake becomes more costly as u_ℓ^R becomes smaller and $\check{p}$ shifts down to avoid this mistake. In other words, the DM avoids the L-biased source when action ℓ becomes more risky.

Part (b) shows that the cost of mistakes also matters for the relative appeal of the *alternative* learning strategies: opposite-biased learning becomes more appealing when

signals (see Matejka and McKay (2015)). This property means any action chosen after performing an experiment must correspond to a unique posterior, and hence a unique level of accuracy. In drift-diffusion models, the drift-to-boundary structure of the optimal policy means that the accuracy of each decision is pinned down by a corresponding stopping boundary. In many such models, the stopping boundaries do not change over time, so the accuracy of each action is fixed. See Ratcliff and McKoon (2008) for a survey.

[21] 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 both boundaries of the region increase (decrease).

[22] For given c, if u_ℓ^R and u_r^L are sufficiently small, $\overline{c}>c$ so that the experimentation region is non-empty. Given $\overline{c}>c$, $\underline{p}^*$ converges monotonically to zero as $u_\ell^R\to-\infty$, and $\overline{p}^*$ converges monotonically to one as $u_r^L\to-\infty$.

mistakes become more costly. In the limit where mistakes become completely unacceptable, the opposite-biased strategy becomes optimal for all beliefs. One could imagine this limit behavior as that of a scientist who views conclusive evidence of either kind—*proving* or *disproving* a hypothesis—as the only acceptable way of advancing science. Such a scientist will rely solely on opposite-biased learning: she will initially strive to prove the hypothesis she conjectures to be true; 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 disproving it as well.[23]

Finally we show how the effect of a higher discount rate compares to the effect of a higher flow cost. Intuitively, one would interpret ρ as a cost of learning and would thus expect that c and ρ are substitutes in the sense that a higher discount rate requires a lower flow cost for the same structure to emerge. Formally, one would expect $\partial\overline{c}/\partial\rho < 0$ and $\partial\underline{c}/\partial\rho < 0$. The following proposition shows that this is indeed the case if at least one "mistake payoff" (u_ℓ^R or u_r^L) is not too small.

Proposition 4 (Discounting vs. Flow Cost)**.**

(a) Suppose $\overline{c} > 0$*. Then* $\partial\overline{c}/\partial\rho < 0$ *if and only if* $U(p) > 0$ *for all* $p \in [0, 1]$*.*

(b) Suppose $\underline{c} > 0$*. Then* $\partial\underline{c}/\partial\rho < 0$ *if both* $u_r^R > \left|u_\ell^R\right|$ *and* $u_\ell^L > \left|u_r^L\right|$*;* $\partial\underline{c}/\partial\rho > 0$ *if* $\min\left\{u_\ell^R, u_r^L\right\}$ *is sufficiently small.*

If the mistake payoffs u_ℓ^R and u_r^L are both negative and sufficiently large in absolute value, both $\partial\overline{c}/\partial\rho$ and $\partial\underline{c}/\partial\rho$ are positive. Namely, a higher discount rate calls for more experimentation in this case. This is intuitive since if losses are sufficiently large, the DM would prefer to delay their realization, when they are discounted more. This favors longer experimentation.

5 Application: Media Consumption

Media outlets differ in their partisan biases.[24] There is evidence that the consumption of biased news outlets affects the political leaning of voters and may change their voting decisions (see DellaVigna and Kaplan (2007) and Martin and Yurukoglu (2017) among others). While a significant fraction of people multi-home and have a news diet that contains outlets with different biases, consumers tend to consume news from outlets with a partisan bias similar to their own position (see for example Gentzkow and Shapiro

[23]Own-biased 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, and without finding one may declare it a correct proof.

[24]One method to measure the partisan bias of an outlet is to compare the language used by the outlet to the language used by members of congress whose partisanship is identified by their voting decisions. Gentzkow and Shapiro (2010) pioneered this method for daily newspapers; Martin and Yurukoglu (2017) use it to identify the bias of cable news channels.

(2011)). This partisan (or own-biased) selective exposure can lead to an "echo-chamber" effect—partisan voters become increasingly polarized (see Martin and Yurukoglu (2017)). Our model contributes to the literature on media choice by providing theoretical predictions about the optimal news diets for voters with different subjective beliefs, their dynamic evolutions, and the implications for polarization.[25]

We interpret our DM as a citizen who votes for one of two candidates, r or ℓ, possibly after consulting media outlets.[26] Candidate r has a right-wing platform, and candidate ℓ has a left-wing platform. The state $\omega \in \{R, L\}$ indicates the optimal platform for the voter, u_x^ω representing the voter's utility from voting for x in state ω. The voter incurs flow cost $c > 0$ for paying attention to the media.[27] Her belief about the state is captured by p. We say the voter is more right-leaning the higher p is.[28]

Naturally, we interpret the L-biased source as a *left-wing outlet* and the R-biased source as a *right-wing outlet*. For example, a right-wing outlet publishes information that supports the left-wing candidate only if it passes a high standard of accuracy—that is, if it constitutes conclusive evidence.[29] Of course this occurs rarely. Most of the time, the outlet instead reports R-favoring content, which Bayesian consumers perceive as having weak informational content. In the same vein, we interpret an interior "attention" choice $\alpha \in (0, 1)$ as multi-homing by the voter across the two outlets, where α represents the share of time she spends on the left-wing outlet.

Our interpretation of media bias accords well with the media literature (see Gentzkow, Shapiro, and Stone (2015) for a survey). A common model in this literature views media bias as arising from the manner in which an outlet "filters" raw news signals for its viewers.

[25]Our rational Bayesian framework to study media choice shares a theme with Calvert (1985), 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. Meyer (1991) makes a similar observation in a dynamic contest environment. Mullainathan and Shleifer (2005) assume a "behavioral" bias on the part of consumers to predict media slanting.

[26]An alternative interpretation is that the citizen has no decision to make but derives a non-instrumental payoff from reaching a certain opinion. The citizen subscribes to media and consults these on an issue of interest. After some time, she may "make up her mind" on the issue and stop acquiring further information. When making up her mind, she enjoys a payoff that increases in the precision of her belief, for instance $\max\{p_t, 1 - p_t\}$. This interpretation corresponds to our model with $u_r^R = u_\ell^L = 1$ and $u_\ell^R = u_r^L = 0$.

[27]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 $\rho = 0$. Our model does not capture the effect of a "deadline," which is clearly relevant for the election example. However, in line with the two-period example discussed in Section 3, we conjecture that the salient features of our characterization which are discussed in the context of this application carry over to a model with a sufficiently long deadline. See Section 6 for further discussion of the effect of a deadline.

[28]Alternatively one could model a voter's bias in terms of her payoffs. In this case u_x^ω could incorporate her "partisan" preference.

[29]The feature that a right-wing outlet could, albeit rarely, broadcast left-favoring news may appear unusual but is a consequence of our voter being a Bayesian who cannot be systematically misled. This feature is also consistent with empirical evidence. Chiang and Knight (2011) find that endorsements of presidential candidates by newspapers are only influential if they go against the bias of the newspaper, suggesting that consumers are, to some extent, able to correct for the bias of newspapers, as predicted by the Bayesian model.

Section 6.1 discusses how our information structure can be microfounded by such a model.

Given this interpretation, our theory, more specifically Theorem 1, provides a rich portrayal of voters' dynamic media choices and their effects. Voters with extreme beliefs $p \notin (\underline{p}^*, \overline{p}^*)$ always vote for their favorite candidates without consulting media. Those who consume media exhibit the following behavior:

- If *news media are moderately informative* so that the cost of information satisfies $\underline{c} \leq c < \overline{c}$, right-leaning voters with $p > \check{p}$ subscribe to right-wing outlets and left-leaning voters subscribe to left-wing outlets. Over time, in the absence of breakthrough news that goes against their initial beliefs, all voters stick to their initial media choice.
- If *media are highly informative*, so that the cost of information is low ($c < \underline{c}$), moderately left-wing ($p \in (\overline{p}, \overline{p}^*)$) or moderately right-wing voters ($p \in (\underline{p}^*, \underline{p})$) consume media that are biased against their beliefs, whereas partisan voters with more extreme beliefs consume media that are biased in favor of their beliefs. Over time, absent breakthrough news in favor of their initial beliefs, moderate voters ($p \in (\underline{p}, \overline{p})$) become increasingly undecided and when their beliefs reach p^*, they multi-home and divide their attention between both types of outlets; whereas partisan voters become more extreme and continue to subscribe to own-based media.

The choice of opposite-biased media by moderate voters may seem counterintuitive. Consider for example a moderate right-wing voter. This voter initially tries to find right-favoring evidence which she expects more likely to arise given their belief. Interestingly, she expects to find such evidence in left-wing outlets as they scrutinize right-favoring information more and apply a very high standard for reporting such information.

In order to understand how consumers' media choices interact with their beliefs, it is useful to perform a simple thought experiment. Suppose there is a unit mass of consumers with identical costs of information acquisition c; whose payoffs u_x^ω are identical and symmetric (so that $\hat{p} = 1/2$); and whose prior beliefs p_0 are distributed according to some distribution function F which is symmetric around 1/2 (i.e., $F(p) = 1 - F(1-p)$). Now, fix the state, say $\omega = L$, and study how the distribution of consumers' beliefs change over time due to their media choice. Of particular interest is the extent to which theirs beliefs become more or less polarized over time. While one can use a number of different measures of polarization,[30] we simply focus on the difference between the median belief for consumers with $p \geq 1/2$ and the median belief for consumers with $p \leq 1/2$.

Proposition 5. *Fix the true state to be $\omega = L$ and assume symmetric payoffs $u_\ell^L = u_r^R$ and $u_\ell^R = u_r^L$.*[31]

(a) The beliefs in the subpopulation of voters consuming own-biased outlets at time $t = 0$ become more polarized over time and converge to a distribution containing three mass points, $\{0, \underline{p}^*, \overline{p}^*\}$.

[30] See Esteban and Ray (2012) for a survey of polarization measures.

[31] Symmetric results hold if the true state is $\omega = R$.

(b) If $c \in (\underline{c}, \overline{c})$, the beliefs of all voters become more polarized over time.

(c) The beliefs of the voters consuming opposite-biased outlets at time $t = 0$ converge to the true state in the limit as $t \to \infty$.

Figure 5.1 shows snapshots of the evolution of beliefs taken at three different times, assuming that initial distribution F is uniform on $[0, 1]$. The colors represent the media choice by voters who are still subscribing to media.

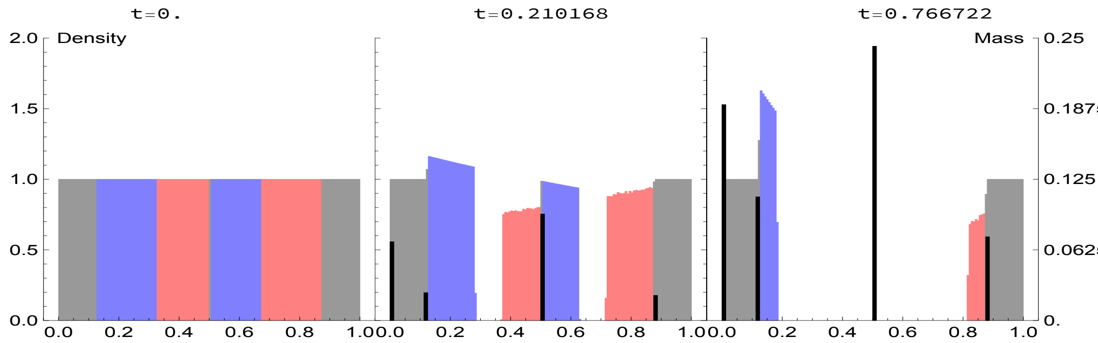


Figure 5.1: Evolution of media choice and beliefs when the true state is L.

Note: Shaded areas represent the density of beliefs (left axis). Colors indicate the choice of media with red indicating right-wing/conservative and blue indicating left-wing/liberal. Bold bars represent mass points of beliefs (right axis).

The figure shows that those with extreme beliefs consume own-biased outlets. Over time, in the absence of contradictory breakthrough news, these outlets feed such voters with what they believe in, leading them to become more extreme. Proposition 5.(a) and (b) states that the beliefs of these voters become more polarized over time. Consequently, our model generates self-reinforcing beliefs—sometimes called an "echo-chamber" effect—that persist until strong contradictory evidence arrives.

The evolution of beliefs is quite different for the voters with moderate beliefs. The figure shows that they consume opposite-biased outlets. Absent breakthrough news this leads to an "anti-echo chamber" effect. Over time, the anti-echo chamber effect makes voters more undecided. Ultimately, they multi-home both left-wing and right-wing outlets, and in the limit, learn the true state, as stated in Proposition 5.(c).[32]

In summary, our dynamic model of media choice predicts two different dynamics of belief evolution resulting from optimal media choice: the beliefs of 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 multi-homing of opposite media outlets. The "anti-echo chamber" effect is a novel prediction of our dynamic model and has no analogue in previous literature.

[32] Gentzkow and Shapiro (2011) present evidence that a significant number of consumers multi-home news channels with different slants. To the best of our knowledge, ours is the first theoretical model to predict the multi-homing of news outlets with conflicting slants.

6 Extensions

In this section, we provide a foundation for our model and link it to a model of filtering bias; and discuss several interesting extensions that correspond to various realistic features of information acquisition. The extensions suggest that our characterization of the optimal policy as well as the proof techniques are surprisingly robust. While the discussion is kept deliberately informal and intuitive, more detailed arguments can be found in Appendix C in the Supplemental Material.

6.1 Discrete-Time Foundation for Conclusive Poisson Model

Although we have only considered *conclusive* Poisson experiments, we show here that these experiments are justified as optimal within a more general class of experiments.

Consider a discrete time analogue of our model with an arbitrary period length $dt \in (0, 1/\lambda)$. The DM's problem is the same as before, except that she incurs a cost of cdt and discounts by the factor of $e^{-\rho dt}$ for each period of information acquisition. In each period, the DM may choose an experiment of the form described in Table 6.1.

state/signal	L-signal	R-signal
L	a	$1-a$
R	$1-b$	b
constraints: $a, b \in [0,1], 1 \leq a+b \leq 1+\lambda dt$		

Table 6.1: General binary-signal experiment

The total probability $a+b$ of "informative" signals is bounded above by $1+\lambda dt$. Note that the overall informativeness of the experiment, measured by λdt, is proportional to the length of a period, and vanishes as $dt \to 0$. This captures the idea that real information takes time to arrive. In the limit as $dt \to 0$, λ parameterizes the constraint for "flow" information.

General binary-signal experiments in discrete time encompass rich and flexible information structures. Setting $(a, b) = (1, \lambda dt)$ or $(a, b) = (\lambda dt, 1)$, we obtain the experiments in Table 2.1 that converge to our conclusive Poisson information structure as $dt \to 0$. More generally, if we set $a = \gamma dt$ and $1 - b = (\gamma - \lambda)dt$, for $\gamma > \lambda$, we obtain an inconclusive Poisson experiment in which breakthrough news arrives in both states but at a higher rate in state L than in state R. In this way, *any* posterior belief $\phi < p$ can be obtained from breakthrough news in the limit as $dt \to 0$, and a similar construction yields jumps to $\phi > p$. If we pick a posterior closer to the prior, breakthrough news arrives with a higher rate. This captures the intuitive idea that a less informative signal can be obtained more easily.[33] Further, a "mixing" of Poisson processes can be attained by switching across

[33]To see this, fix any posterior ϕ below the prior p. Consider the experiment with $a = \frac{p(1-\phi)}{p-\phi}\lambda dt$ and

different experiments within the class. For example, dividing attention with $\alpha = 1/2$ in the baseline model, is obtained by switching between $(a, b) = (1, \lambda dt)$ and $(a, b) = (\lambda dt, 1)$ over time.[34]

In summary, this class encompasses a range of both conclusive and inconclusive experiments.[35] Suppose the DM is free to choose from this rich class of experiments. Which experiments are optimal? Will she necessarily choose an accurate signal? The answer is not obvious, since the DM may find inaccurate signals appealing as they are easier to obtain. Nevertheless, we show that conclusive experiments are optimal, thus justifying our focus on them within this class of experiments.

Proposition 6. *Consider the discrete-time problem with arbitrary period length $dt \in (0, 1/\lambda)$ and finite or infinite number of periods. In each period and for each belief, any binary experiment with $a + b \leq 1 + \lambda dt$ is weakly dominated by either σ^R or σ^L.*

Importantly, this proposition does not claim a Blackwell dominance relation. We show that at each history, one of the experiments σ^R or σ^L is optimal, but which one depends on the current belief and the continuation payoffs. This result follows from the convexity of the DM's continuation payoff in her beliefs, which holds because the payoff of any fixed strategy is linear in the prior belief.[36]

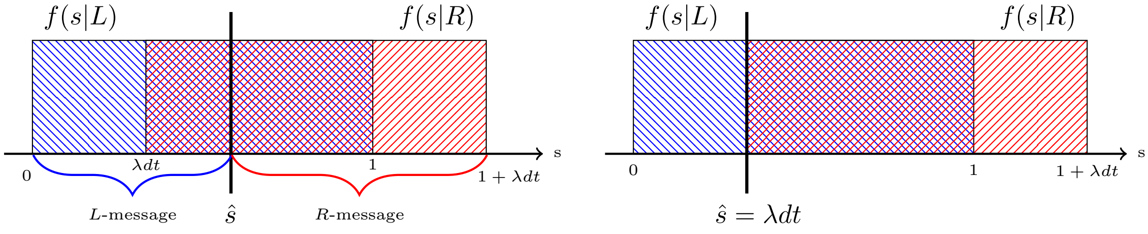
Filtering Interpretation. The class of experiments featured in Table 6.1 can be motivated via a "filtering" model commonly adopted in the media choice literature. According to this model, media "bias" or "slant" arises when an outlet filters rich raw information into coarse "messages" for the consumers (see Gentzkow, Shapiro, and Stone (2015) and Prat and Strömberg (2013)). Suppose that an outlet observes a signal $s \in \mathbb{R}$ where s is drawn uniformly from $[0, 1]$ in state L and uniformly from $[\lambda dt, 1 + \lambda dt]$ in state R. The outlet must "filter" the signal into coarse messages. This could reflect, for instance, limited publishing space or broadcasting time, or a limited capacity of consumers to process rich information. Suppose, the outlet sends binary messages to the consumer as depicted

$b = 1 - \frac{(1-p)\phi}{p-\phi}\lambda dt$. In the limit as $dt \to 0$, the experiment converges to a Poisson process which sends an L-breakthrough signal at rate $\frac{p(1-\phi)}{p-\phi}\lambda$ in state L, and at rate $\frac{(1-p)\phi}{p-\phi}\lambda$ in state R, so that upon receiving that signal the belief becomes exactly ϕ. The arrival rate increases and converges to ∞ as $\phi \nearrow p$.

[34]As usual, in the absence of news, the belief drifts in the direction implied by Bayes rule. One example of this is the stationary policy $\alpha(p) = 1/2$, in which no updating occurs in the absence of breakthrough news; this is obtained when two Poisson processes with jumps to zero and one are mixed equally.

[35]In discrete time, the class of experiments also admits a random walk, e.g. if $a = b = (1 + \lambda dt)/2$. However, this process becomes uninformative in the limit as $dt \to 0$. It converges to a diffusion process with identical drift in both states. In other words, an informative DDM cannot be obtained as a limit of the current class.

[36]Zhong (2017) demonstrates the optimality of a (non-conclusive) Poisson experiment when the DM incurs posterior separable cost that depends on the experiments as well as the current belief. While similar in spirit, our result is not an implication of his result. We adopt a class of feasible Blackwell experiments that is independent of the DM's belief. Our constraint $a + b < 1 + \lambda dt$ cannot be derived from a constraint on a posterior separable information cost function (details are available from the authors on request). Further, we prove optimality of conclusive experiments, which is not shown in his paper.



(a) choice of bias (b) right-wing outlet

Figure 6.1: Filtering Bias

in Figure 6.1. It sends an L-message if $s < \hat{s}$, and an R-message if $s > \hat{s}$, for a threshold $\hat{s}$ chosen by the media outlet. The threshold $\hat{s}$ characterizes the outlet's "ideological" orientation; the lower $\hat{s}$, the more R-biased it is. Each filtering-threshold $\hat{s}$ induces an experiment of the form described in Table 6.1 on page 21, with $a = \hat{s}$ and $b = 1+\lambda dt-\hat{s}$. The "left-wing" and "right-wing" outlets in Section 5 correspond to cutoffs $\hat{s} = 1$ and $\hat{s} = \lambda dt$, respectively. Compared with these outlets, the outlets choosing cutoffs $\hat{s} \in (\lambda dt, 1)$ can be interpreted as more moderate. In Section 5, our results hold unchanged if we expand the set of media outlets to contain all these moderate outlets since Proposition 6 shows that, even facing such a rich choice, consumers will still choose from two extreme media outlets.

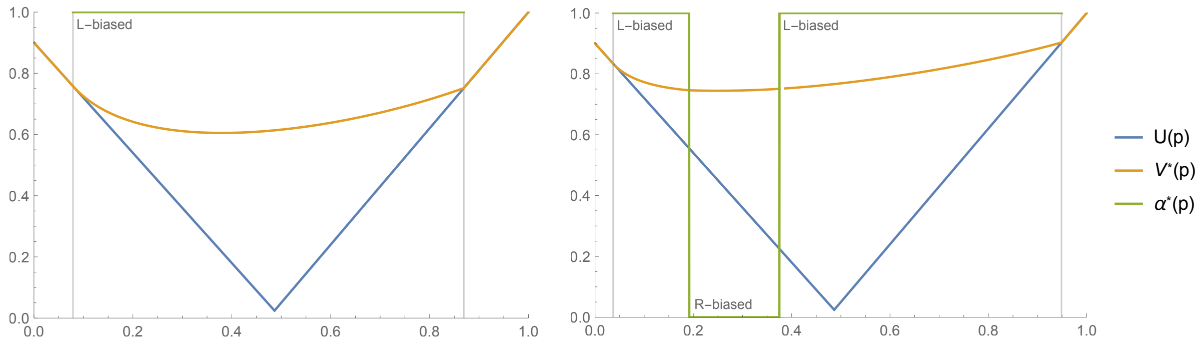
6.2 Non-Exclusivity of Attention

Our model does not allow for accidental discovery of evidence; i.e., a DM never receives evidence that she is not looking for. It is plausible, however, that an individual who looks for R-evidence may accidentally find the opposite and become convinced that the state is L. For example, a prosecutor seeking evidence that a suspect is guilty, may stumble upon evidence to the contrary.

This possibility can be easily accommodated within our model by assuming that the DM is limited to an interior attention choice $\alpha \in [\underline{\alpha}, \overline{\alpha}]$, where $0 < \underline{\alpha} < 1/2 < \overline{\alpha} < 1$. Consequently she will always be exposed to evidence of both types, and may find one type of evidence while looking only for the other. If we set $\underline{\alpha} = 1 - \overline{\alpha}$, Theorem 1 as well as Propositions 2, 3, and 4 remain qualitatively unchanged. Of course, the precise cutoffs that characterize the optimal policy change: The experimentation region shrinks and $\overline{c}$ becomes smaller as $\overline{\alpha}$ decreases. The intuition is that a restriction on the feasible information choices reduces the value of information acquisition. More interestingly, opposite-biased learning is part of the optimal policy for a larger range of cost parameters—the cutoff $\underline{c}$ increases as $\overline{\alpha}$ falls. Opposite-biased learning is less affected by the restriction since it calls the DM to ultimately divide attention once p^* is reached. Hence, at p^* the constraint on α is not binding, so that the value of opposite-biased learning is less sensitive to the restriction on the feasible choices for α than the value of own-biased learning.

6.3 Asymmetric Returns to Attention

We have also assumed that both states are equally easy to prove. The arrival rate for breakthrough news for each type of evidence is the same. We can easily relax this feature by introducing two different arrival rates, $\overline{\lambda}^R$ for R-evidence and $\overline{\lambda}^L$ for L-evidence. For a given attention choice (α, β), this means that in state R evidence arrives at rate $\alpha\overline{\lambda}^R$, and in state L evidence arrives at rate $\beta\overline{\lambda}^L$. To fix ideas, suppose $\overline{\lambda}^R > \overline{\lambda}^L$ so that state R is easier to prove. If $\overline{\lambda}^R - \overline{\lambda}^L$ is small, our characterization of the optimal policy in Theorem 1 carries over to this case: For low levels of the cost c, the optimal policy combines own-biased and opposite-biased learning, and for moderate costs only own-biased learning is optimal. If $\overline{\lambda}^R - \overline{\lambda}^L$ is large, the structure changes as illustrated in Figure 6.2.



(a) $c = .15$ (moderate cost) (b) $c = .07$ (low cost)

Figure 6.2: Value Function and Optimal Policy with asymmetric returns to attention.

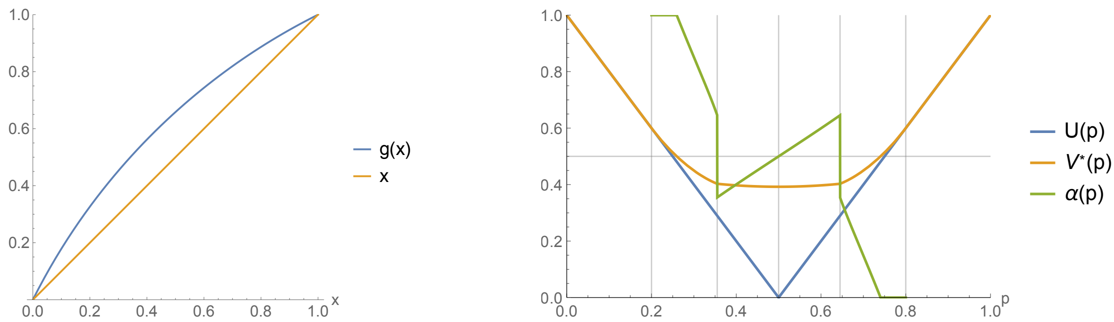
Note: $\overline{\lambda}^R = 1$, $\overline{\lambda}^L = .6$, $\rho = 0$, $u_r^R = u_\ell^L = 1$, $u_r^R = 1$, $u_\ell^L = .9$, $u_\ell^R = u_r^L = -.9$

In the case of moderate costs (Panel (a)), the DM never looks for L-evidence, meaning R-biased learning is not part of the optimal policy. In the case of low costs (Panel (b)), the opposite-biased learning strategy appears, but it is skewed toward L-biased learning (or R-evidence seeking), and the absorbing state p^* is less than 1/2 even when the payoffs are symmetric. As in Panel (a), for high beliefs near the stopping region, R-biased learning is not optimal, in contrast to the characterization in Theorem 1.

6.4 Diminishing Returns to Attention

In our model, the DM never splits her attention or multi-homes media outlets, except at the absorbing belief p^*. This feature is a consequence of the "linear" attention technology assumed in our model. The arrival rate of an R-breakthrough is $\lambda\alpha$ and the arrival rate of an L-breakthrough is $\lambda(1 - \alpha)$. This means that the marginal return to attention to a single news source is constant. In practice, however, a diminishing marginal return may be realistic in some contexts; namely, one may learn more efficiently from diverse news sources than from just one. For instance, one may obtain more information by reading the front pages of multiple newspapers, than by reading a single newspaper from front to back. In the Internet era, multi-homing is facilitated by *news aggregators* such as Google News or Yahoo News that curate diverse news sources or perspectives that complement one another. One can learn more efficiently from such aggregators than by focusing on a single news source.

Diminishing marginal returns to attention can be incorporated in our model by assuming that the arrival rate of R-evidence is given by $\lambda g(\alpha)$, and the arrival rate of L-evidence is given by $\lambda g(1-\alpha)$, where g is an increasing function that satisfies $g(0)=0$, $g(1)=1$. Our baseline model corresponds to $g(x)=x$ and diminishing returns obtain if $g(x)$ is strictly concave. Panel (a) of Figure 6.3 depicts these two cases.



(a) constant vs. diminishing returns

(b) optimal media consumption

Figure 6.3: Diminishing returns to attention.

Note: $\lambda=1$, $c=.4$, $\rho=0$, $u_\ell^L=u_r^R=1$, $u_r^L=u_\ell^R=0$, $g(x)=\sqrt{1+4x-x^2}-1$.

When attention has diminishing marginal returns, multi-homing, or interior choices of α, become optimal for a wide range of beliefs as depicted in panel (b) of Figure 6.3. Nevertheless, the basic structure of optimal policy resembles that of Theorem 1, if we call the attention choice L-biased if $\alpha>1/2$, so that absent breakthrough news the belief drifts toward L, and call the attention choice R-biased if $\alpha<1/2$. Specifically, the optimal policy is again characterized by own-biased and opposite-biased learning strategies. The former is optimal for extreme beliefs and the latter may be optimal for moderate beliefs.

The resulting characterization yields richer implications for the interplay between information choice and beliefs in the context of the media choice. For voters with extreme beliefs, the echo-chamber effect is reinforced. Not only do the beliefs evolve over time due to a biased news diet, the media choice itself evolves. Over time absent breakthrough news, partisan voters' beliefs become more extreme, and this in turn leads to a more biased new-diet. For example, a right-leaning voter consumes more and more right-biased news by decreasing α as her belief moves more to the right over time (see Panel (b) of Figure 6.3). A more formal result is derived for the case of symmetric payoffs in Appendix C.4 in the Supplemental Material.

6.5 A Deadline for Decision Making

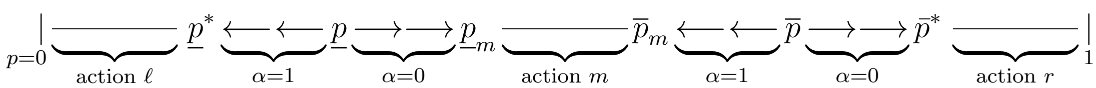
What happens if the DM faces a firm deadline for her decision? Our media application raises this issue since a voter must stop deliberating on the election date. While a general analysis incorporating a deadline is beyond the scope of the current paper, our two-period example discussed in Section 3 sheds some light both on the robustness of our characterization as well as the effect of a deadline.

First, both modes of learning identified in Theorem 1 can arise in the first period. In the example in Section 3, a combination of own-biased learning and opposite-biased learning is optimal. For other parameter values in the example, own-biased learning is optimal for all beliefs. Second, we find a clear deadline effect. In the second period, own-biased learning is always optimal. We conjecture that this pattern will hold more generally if one were to introduce a firm deadline in our continuous time model—namely, the DM will shift her attention increasingly toward own-biased news sources as the deadline approaches. This is indeed the pattern found by Stroud (2008) from her analysis of the 2004 National Annenberg Election Survey. She finds that selective exposure and partisan consumption of media outlets intensifies as the election date approaches (see Figures 1 and 2 therein).

6.6 Non-Binary States and Actions

Our model can be easily extended to include more than two actions. Suppose there is a third action m with payoffs $u_m^R \in (u_\ell^R, u_r^R)$ and $u_m^L \in (u_r^L, u_\ell^L)$. Then we can define a new learning heuristic, called "m-strategy," that takes the following form:

The two cutoffs $\underline{p}_m$ and $\overline{p}_m$ are chosen optimally. The optimal policy is a combination of own-biased learning, opposite-biased learning, and the m-strategy. It can take various forms. For example if $c \in (\underline{c}, \overline{c})$, and the value of the m-strategy is higher than the value of opposite-biased learning, the optimal policy takes the following form:



Along these lines, a finite number of actions can be added (see Appendix C.5 in the Supplemental Material).

Extending the model to more than two states raises several issues. First, the state-space of the DM's problem becomes multi-dimensional. Second, it is natural to allow for a larger set of news sources if there are more than two states. Within our Poisson framework, with two states, two sources are sufficient to allow for good news and bad news about each state. With $n > 2$ states, there could be n good-news sources and n bad-news sources, greatly increasing the complexity of the DM's attention allocation problem. While limited progress has been made in multi-state models with only two news sources,[37] we conjecture that a characterization in a general model will not be tractable.[38]

6.7 Non-Conclusive Evidence

So far, we have assumed that the DM can obtain conclusive evidence—that is, a signal that arrives only in one state. This can be relaxed by introducing "noise," or "false evidence." Suppose the DM looks for ω-evidence. With noise, this is received in state ω with a Poisson rate of $\overline{\lambda}$ but *also in state* $\omega' \neq \omega$ with a lower rate $\underline{\lambda} < \overline{\lambda}$. If $\underline{\lambda} > 0$, then an ω-signal is no longer conclusive evidence for state ω. In a previous version of this paper (Che and Mierendorff, 2017), we analyze this extension and show that, as long as $\underline{\lambda}$ is sufficiently small, the DM finds it optimal to take an action immediately after receiving a (noisy) breakthrough signal—a property we call *Single Experimentation Property* (SEP). Given SEP, our characterization applies without any qualitative changes.[39] Moreover, the main implications in terms of accuracy and delay reported in Proposition 1 continue to hold. A DM with a more uncertain belief, who chooses opposite-biased learning, ends up making a more accurate decision but with a greater delay than a DM with a more extreme belief who employs own-biased learning and as a result is more prone to mistakes.

This dependence of the stochastic choice function on the DM's prior belief was already present in the baseline model. With noise the stochastic choice becomes richer, and new phenomena appear. In particular, we show that, conditional on the prior belief, a decision maker who (by chance) received a breakthrough very quickly, makes a more accurate decision than a DM who had to wait a long time for a breakthrough. This finding is consistent with so-called "speed-accuracy complementarity"—a phenomenon that a more delayed decision tends to be less accurate.[40] The simple intuition is that noisy evidence is less convincing if the DM was more skeptical when receiving it. This means that if the DM is unlucky and waits for a longer time before receiving breakthrough news, the accuracy of her action will suffer.

[37] See the discussion of Nikandrova and Pancs (2018) and Mayskaya (2016) in the Introduction.

[38] As far as we know, even in the Wald stopping problem, tractable characterizations are not available for more than two states (Peskir and Shiryaev, 2006).

[39] A general characterization is difficult to obtain if SEP does not hold. In Che and Mierendorff (2017) we solve such a case. While our analytical method developed here continues to be useful, the characterization is much more complex, involving multiple jumps across different learning regions.

[40] This finding is often documented in perceptual judgment or consumption choice experiments conducted in cognitive psychology. See Ratcliff and McKoon (2008) for a survey, and Fudenberg, Strack, and Strzalecki (2017) for a recent economic theory.

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, own-biased learning*, and *opposite-biased learning* for different prior beliefs. We have used this characterization to obtain rich predictions about information acquisition and choices.

We envision several avenues of extending the current work. First, our model is relatively tractable (e.g., the value function and optimal policy are available in closed form). Therefore, we hope that this framework will be useful for integrating dynamic choice of information in applied theory models. This includes our application to media choice, which could be extended beyond our analysis in the present paper. Dynamic information choice might be also embedded in a model of a committee or a jury, or in 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 to allow for more gradual information acquisition. Our analysis about repeated experimentation in Che and Mierendorff (2017) 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 Proof of Theorem 1

In this appendix, we describe the construction of the heuristic strategies and state the main steps of the proof of Theorem 1. Omitted proofs can be found in Section B.1 in the Online Appendix. For mathematical details on optimal control problems that are used here, see Bardi and Capuzzo-Dolcetta (1997).

A.1 The DM's problem

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

$$V^*(p_0) = \max_{(\alpha_\tau),T} \int_0^T e^{-\rho t} P_t(p_0, (\alpha_\tau)) \left[\lambda \alpha_t p_t \, u_r^R + \lambda \beta_t (1 - p_t) \, u_\ell^L - c\right] dt + e^{-\rho T} P_T(p_0, (\alpha_\tau)) U(p_T), \tag{A.1}$$

s.t. $\dot{p}_t = -\lambda\left(\alpha_t - \beta_t\right) p_t \left(1 - p_t\right),$

where $P_t(p_0, (\alpha_\tau)) := p_0 e^{-\lambda\int_0^t \alpha_s ds} + (1-p_0)e^{-\lambda\int_0^t \beta_s ds}$, and $\beta_t = 1 - \alpha_t$. Given that the problem is autonomous, the optimal α_t only depends on the belief at time t, and can thus be written as a policy $\alpha(p)$. A policy $\alpha(p)$ is *admissible* if, together with (A.2) it defines a unique path p_t for any prior $p_0 \in [0,1]$. Similarly, the decision to stop and take an action only depends on p_t.

A.2 Two Benchmarks and a Condition for Experimentation

Two benchmark value functions prove useful for our analysis. The first is the value of the *stationary strategy*:

$$U^S(p) := p\frac{\frac{1}{2}\lambda\, u_r^R - c}{\rho + \frac{1}{2}\lambda} + (1-p)\frac{\frac{1}{2}\lambda\, u_\ell^L - c}{\rho + \frac{1}{2}\lambda} = \frac{\lambda\left(pu_r^R + (1-p)u_\ell^L\right)}{2\rho + \lambda} - \frac{2c}{2\rho+\lambda}, \tag{A.2}$$

which arises when the DM chooses $\alpha_t = \beta_t = 1/2$ for all t and takes an optimal action only after receiving the conclusive breakthrough news.

The second is what we call the *full-attention value*:

$$U^{FA}(p) := p\frac{\lambda\, u_r^R - c}{\rho + \lambda} + (1-p)\frac{\lambda\, u_\ell^L - c}{\rho + \lambda} = \frac{\lambda\left(pu_r^R + (1-p)u_\ell^L\right)}{\rho + \lambda} - \frac{c}{\rho+\lambda}, \tag{A.3}$$

which arises in a "hypothetical" scenario in which the DM chooses (infeasible) attention $\alpha_t = \beta_t = 1$ for all t and again takes an optimal action only after receiving breakthrough news. Since limited attention prevents the DM from achieving this value in our model, $U^{FA}(p)$ serves only as an analytical device for proofs.

Intuitively, $U^{FA}(p)$ is an upper bound for a payoff the DM can obtain from experimentation. Note further that $U^{FA}(\cdot)$ is linear and $U(\cdot)$ is piecewise-linear with a kink at $\hat{p}$. Hence, the condition[41]

$$U^{FA}(\hat{p}) > U(\hat{p}) \tag{EXP}$$

would be necessary for experimentation to be optimal. Indeed, if (EXP) does not hold, an immediate action is optimal for all $p \in [0,1]$:

Proposition 7. *For all $p \in [0,1]$, $U(p) \leq V^*(p) \leq \max\left\{U(p), U^{FA}(p)\right\}$. In particular, if (EXP) is violated, then $V^*(p) = U(p)$ for all p.*

A.3 The Bellman equation

In light of Proposition 7, in the sequel we only consider the case where (EXP) holds, and construct the value function for the range of beliefs where $U^{FA}(p) > U(p)$. The HJB

[41] It is easy to check that $U(p) \geq U^{FA}(p)$ for all p if (EXP) is violated.

equation for the DM's problem in (A.1) is the following variational inequality

$$\max\left\{-c-\rho V(p)+\max_{\alpha\in[0,1]} F_\alpha(p,V(p),V'(p)), U(p)-V(p)\right\}=0, \tag{A.4}$$

where

$$F_\alpha(p,V(p),V'(p)) := \left\{\begin{matrix}\alpha\lambda p\left(u_r^R-V(p)\right)+(1-\alpha)\lambda(1-p)\left(u_\ell^L-V(p)\right)\\ -\lambda\left(2\alpha-1\right)p(1-p)V'(p)\end{matrix}\right\}. \tag{A.5}$$

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

$$c+\rho V(p)=F(p,V(p),V'(p))\left(:=\max_{\alpha\in[0,1]} F_\alpha(p,V(p),V'(p))\right). \tag{A.6}$$

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

In the following, we will construct a candidate value function and show that it satisfies (A.4) for all points of differentiability. This would be sufficient if the candidate were differentiable everywhere. Since our candidate function has kinks, we show instead that it is a *viscosity solution* of (A.4), a necessary and sufficient condition for the value function according to the verification theorem we invoke (see Proposition 9 below).

Note that $F_\alpha(\cdot)$ is linear in α. Therefore, 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. With α set respectively to 0 and 1, we can define functions, V_0 and V_1, satisfying the ODEs:

$$c+\rho V_0(p)=F_0(p,V_0(p),V_0'(p))=\lambda(1-p)\left(u_\ell^L-V_0(p)\right)+\lambda p(1-p)V_0'(p), \tag{A.7}$$

$$c+\rho V_1(p)=F_1(p,V_1(p),V_1'(p))=\lambda p\left(u_r^R-V_1(p)\right)-\lambda p(1-p)V_1'(p). \tag{A.8}$$

Solutions to these ODEs with boundary condition $V(x)=W$ are well-defined if $x\in(0,1)$, and denoted by $V_0(p;x,W)$ and $V_1(p;x,W)$, respectively.[42]

A.4 Own-Biased Strategy

Recall the structure of the own-biased strategy given by (4.3). We will define its value, labeled $V_{own}(\cdot)$ to be an upper envelope of two value functions, $\underline{V}_{own}(\cdot)$ and $\overline{V}_{own}(\cdot)$, respectively its left- and right-branches. To this end, we first compute the boundary beliefs, $\underline{p}^*$ and $\overline{p}^*$, and then construct the two branches by solving the ODEs (A.7) and (A.8), using boundary conditions at $\underline{p}^*$ and $\overline{p}^*$. The particular construction will be ultimately justified later through our verification argument (Proposition 9).

[42] $V_0(p;x,W)$ and $V_1(p;x,W)$ are uniquely defined if $x\in(0,1)$ because (A.7) and (A.8) satisfy local Lipschitz continuity for all $p\in(0,1)$.

First, value matching and smooth pasting (relative the immediate action payoffs) pin down the boundary beliefs:[43]

$$\underline{p}^* := \frac{u_\ell^L \rho + c}{\rho\left(u_\ell^L - u_\ell^R\right) + \left(u_r^R - u_\ell^R\right)\lambda}, \tag{A.9}$$

$$\overline{p}^* = \frac{\left(u_\ell^L - u_r^L\right)\lambda - u_r^L \rho - c}{\rho\left(u_r^R - u_r^L\right) + \left(u_\ell^L - u_r^L\right)\lambda}, \tag{A.10}$$

Next, we define the value of the left branch as $\underline{V}_{own}(p) = U_\ell(p)$ for $p \leq \underline{p}^*$. For $p > \underline{p}^*$, we set $\underline{V}_{own}(p) = V_1(p; \underline{p}, U_\ell(\underline{p}^*))$ which yields:

$$\underline{V}_{own}(p) = -\frac{c}{\rho}(1-p) + \frac{u_r^R \lambda - c}{\lambda + \rho} p + \frac{\lambda\left(c + u_\ell^L \rho\right)}{\rho(\lambda+\rho)} \left(\frac{\underline{p}^*}{1-\underline{p}^*}\right)^{\frac{\rho}{\lambda}} \left(\frac{1-p}{p}\right)^{\frac{\rho}{\lambda}} (1-p). \tag{A.11}$$

Similarly, the value $\overline{V}_{own}(p)$ of the right branch equals $U_r(p)$ for $p \geq \overline{p}^*$. For $p < \overline{p}^*$, we set $\overline{V}_{own}(p) = V_0(p; \overline{p}^*, U_r(\overline{p}^*))$ which yields:

$$\overline{V}_{own}(p) := -\frac{c}{\rho} p + \frac{u_\ell^L \lambda - c}{\lambda + \rho}(1-p) + \frac{\lambda\left(c + u_r^R \rho\right)}{\rho(\lambda+\rho)} \left(\frac{1-\overline{p}^*}{\overline{p}^*}\right)^{\frac{\rho}{\lambda}} \left(\frac{p}{1-p}\right)^{\frac{\rho}{\lambda}} p. \tag{A.12}$$

Combining these functions, we define the value of the own-biased strategy as $V_{own}(p) := \max\left\{\underline{V}_{own}(p), \overline{V}_{own}(p)\right\}$. Without further analysis, it is not clear when $V_{own}(p)$ is the value of a strategy of the form (4.3). This will be clarified in Section A.6.

A.5 Opposite-Biased Strategy

Recall the structure of the opposite-biased strategy given by (4.4). The value of this strategy, denoted by $V_{opp}(p)$, and the reference belief p^* are defined as follows. First, we observe that the value must equal the stationary value $U^S(p^*)$ at p^*. Given this, we invoke value matching and smooth pasting to pin down[44]

$$p^* := \frac{\left(u_\ell^L \rho + c\right)}{\left(u_r^R \rho + c\right) + \left(u_\ell^L \rho + c\right)}. \tag{A.13}$$

For $p \leq p^*$ we have $V_{opp}(p) = V_0(p; p^*, U^S(p^*))$ which yields:

$$\underline{V}_{opp}(p) := -\frac{c}{\rho} p + \frac{u_\ell^L \lambda - c}{\lambda + \rho}(1-p) + \frac{\lambda}{\rho(2\rho+\lambda)} \frac{\lambda\left(u_r^R \rho + c\right)}{\lambda + \rho} \left(\frac{1-p^*}{p^*} \frac{p}{1-p}\right)^{\frac{\rho}{\lambda}} p. \tag{A.14}$$

[43] We show in Section A.6 that the boundary beliefs satisfy $0 < \underline{p}^* < \hat{p} < \overline{p}^* < 1$ if (EXP) is satisfied.

[44] Namely, we insert $V_0(p^*) = U^S(p^*)$ and $V_0'(p^*) = U^{S\prime}(p^*)$ in (A.7). It turns out that this yields the same value for p^* as the one obtained by inserting $V_1(p^*) = U^S(p^*)$ and $V_1'(p^*) = U^{S\prime}(p^*)$ in (A.8).

Likewise, for $p \geq p^*$, we have $V_{opp}(p) = V_1(p; p^*, U^S(p^*))$ which yields:

$$\overline{V}_{opp}(p) := -\frac{c}{\rho}(1-p) + \frac{u_r^R\lambda - c}{\lambda+\rho}p + \frac{\lambda}{\rho(2\rho+\lambda)}\frac{\lambda\left(u_\ell^L\rho + c\right)}{\lambda+\rho}\left(\frac{p^*}{1-p^*}\frac{1-p}{p}\right)^{\frac{\rho}{\lambda}}(1-p). \quad \text{(A.15)}$$

A.6 Solution Candidate

Again we assume (EXP) is satisfied. We define our solution candidate as the upper envelope of $V_{own}(p)$ and $V_{opp}(p)$, denoted by $V_{Env}(p) := \max\{V_{own}(p), V_{opp}(p)\}$. This function is characterized as follows.

Proposition 8 (Structure of V_{Env}). *(a) If (EXP) holds and $V_{own}(p^*) \geq V_{opp}(p^*)$, then there exists a unique $\check{p} \in \left(\underline{p}^*, \overline{p}^*\right)$ such that $\underline{V}_{own}(\check{p}) = \overline{V}_{own}(\check{p})$ and*

$$V_{Env}(p) = V_{own}(p) = \begin{cases} \underline{V}_{own}(p), & \text{if } p < \check{p}, \\ \overline{V}_{own}(p), & \text{if } p \geq \check{p}. \end{cases}$$

(b) If (EXP) holds and $V_{own}(p^*) < V_{opp}(p^*)$, then $p^* \in (\underline{p}^*, \overline{p}^*)$, and there exist a unique $\underline{p} \in (\underline{p}^*, p^*)$ such that $V_{own}(\underline{p}) = V_{opp}(\underline{p})$, and a unique $\overline{p} \in (p^*, \overline{p}^*)$ such that $V_{own}(\overline{p}) = V_{opp}(\overline{p})$ and

$$V_{Env}(p) = \begin{cases} \underline{V}_{own}(p), & \text{if } p < \overline{p}, \\ V_{opp}(p), & \text{if } p \in [\underline{p}, \overline{p}], \\ \overline{V}_{own}(p), & \text{if } p > \overline{p}. \end{cases}$$

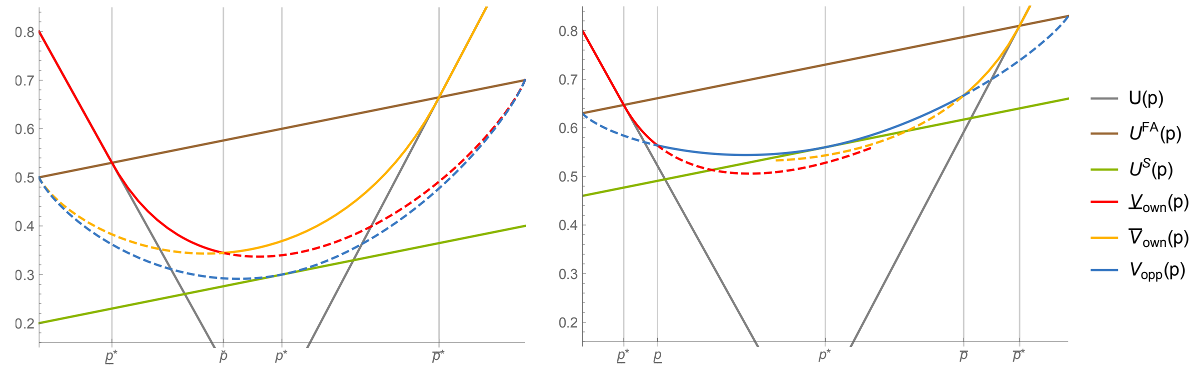
To understand how we derive the structure of the solution candidate it is useful to make several geometric observations, which are depicted for illustrative purpose in Figure A.1. Note first that $\underline{V}_{own}(p)$ is strictly convex on $[\underline{p}^*, 1]$, $\overline{V}_{own}(p)$ is strictly convex on $[0, \overline{p}^*]$, and $V_{opp}(p)$ is strictly convex on $[0, 1]$. This can be seen directly from (A.11)–(A.12) and (A.14)–(A.15).[45] Figure A.1 shows that these three functions coincide with $U^{FA}(p)$ at the endpoints of the respective intervals above. When the endpoints are $p = 0$ and $p = 1$ this can be seen by comparing the explicit expression and (A.3). Our first crucial Lemma shows that the value of own-biased learning coincides with U^{FA} at the boundary points $\underline{p}^*$ and $\overline{p}^*$.

Lemma 1. *The boundary points $\underline{p}^*$ and $\overline{p}^*$ satisfy*

$$U_\ell(\underline{p}^*) = U^{FA}(\underline{p}^*) \quad \text{and} \quad U_r(\overline{p}^*) = U^{FA}(\overline{p}^*). \quad \text{(A.16)}$$

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

[45] For $V_{opp}(p)$, its convexity on the whole interval $[0, 1]$ follows from the convexity of the two branches $\underline{V}_{opp}(p)$ and $\overline{V}_{opp}(p)$ and smooth pasting at p^*.



(a) only own-biased ($c = .3$) (b) opposite-biased and own-biased ($c = .13$)

Figure A.1: Branches of the value function and solution candidate.

Note: Dashed lines indicated segments of the branches that are not part of V_{Env}. (Parameters: $\lambda = 1$, $\rho = 0$, $u_r^R = 1$, $u_\ell^L = .8$, $u_\ell^R = -1$, $u_r^L = -.8$)

Equipped with these preliminary observations, we can turn to the characterization of the upper envelope in Proposition 8. The following crucial "crossing lemma" characterizes how a function solving (A.7) (such as $\overline{V}_{own}$ or $\underline{V}_{opp}$) intersects a function that solves (A.8) (such as $\underline{V}_{own}$ or $\overline{V}_{opp}$).

Lemma 2 (Crossing Lemma). *Let V_0 satisfy (A.7) and V_1 satisfy (A.8). If $V_0(p) = V_1(p) > (=)U^S(p)$ at some at $p \in (0,1)$, then $V_0'(p) > (=)V_1'(p)$.*

The Lemma implies that $\underline{V}_{own}$ must cross $\overline{V}_{own}$ (and likewise $\underline{V}_{own}$ and $\overline{V}_{opp}$ must cross $\underline{V}_{opp}$ and $\overline{V}_{own}$ respectively from above), when an intersection occurs above the stationary value function.

Suppose first that $\max\{\underline{V}_{own}(p), \overline{V}_{own}(p)\} \geq U^S(p)$ for all $p \in [0,1]$. Our preliminary observations imply that $\underline{V}_{own}$ and $\overline{V}_{own}$ must cross each other at some $p \in (\underline{p}^*, \overline{p}^*)$.[46] Since their upper envelope exceeds $U^S(p)$, the Crossing Lemma 2 implies that $\underline{V}_{own}(p)$ intersects $\overline{V}_{own}(p)$ from above. Consequently the intersection point $\check{p}$ must be unique, as depicted in Panel (a) of Figure A.1. This summarizes the crucial step for part (a) of Proposition 8.[47]

Suppose next that $\max\{\underline{V}_{own}(p^*), \overline{V}_{own}(p^*)\} < U^S(p^*)$, as in Part (b) of Proposition 8, a case depicted in Panel (b) of Figure A.1. Consider the interval $(\underline{p}^*, p^*)$. From Lemma 1 and our preliminary observations, we have $V_{opp}(\underline{p}^*) < U^{FA}(\underline{p}^*) = \underline{V}_{own}(\underline{p}^*)$. Since $V_{opp}(p^*) > \underline{V}_{own}(p^*)$, there must be an intersection between $V_{opp}(p)$ and $\underline{V}_{own}(p)$ at some $p \in (\underline{p}^*, p^*)$. Since $V_{opp}(p) = \underline{V}_{opp}(p) > U^S(p)$ for all $p < p^*$, the Crossing Lemma 2

[46]Since $\underline{V}_{own}$ is strictly convex and equal to the linear function $U^{FA}(p)$ for $p \in \{\underline{p}^*, 1\}$, we have $\underline{V}_{own}(\overline{p}^*) < U^{FA}(\overline{p}^*) = \overline{V}_{own}(\overline{p}^*)$. Similarly, we obtain $\overline{V}_{own}(\underline{p}^*) < \underline{V}_{own}(\underline{p}^*)$. Hence there must be an intersection.

[47]See the complete proof in Appendix B.1 for the remaining steps: (i) $\max\{\underline{V}_{own}(p^*), \overline{V}_{own}(p^*)\} \geq U^S(p^*)$ implies the stronger condition $\max\{\underline{V}_{own}(p), \overline{V}_{own}(p)\} \geq U^S(p)$ for all $p \in [0,1]$ that we used here, and (ii) $V_{opp}(p) \leq V_{own}(p)$ for all $p \in (0,1)$ if $\max\{\underline{V}_{own}(p^*), \overline{V}_{own}(p^*)\} \geq U^S(p^*)$.

implies that $\underline{V}_{own}(p)$ intersects $V_{opp}(p)$ from above and hence there is a unique intersection $\underline{p} \in (\underline{p}^*, p^*)$. The characterization of the upper envelope for $p < p^*$ is completed by noting that $\overline{V}_{own}(p) < V_{opp}(p)$ for all $p < p^*$.[48] A symmetric argument is used to characterize $V_{Env}(p)$ for $p > p^*$.

A.7 Verification of the Candidate

We now show that V_{Env} is the value function of the DM's problem in (A.1).

Proposition 9. *$V^*(p) = V_{Env}(p)$ for all $p \in [0,1]$. Up to tie-breaking at $\check{p}$, $\underline{p}$, $\overline{p}$, $\underline{p}^*$, and $\overline{p}^*$, the optimal policy is unique.*

For $p \notin (\underline{p}^*, \overline{p}^*)$, $V_{Env}(p) = U(p)$ which is equal to $V^*(p)$ by Proposition 7. To show optimality for beliefs inside the experimentation region, the following Lemma is crucial.

Lemma 3 (Unimprovability). *(a) If V_0 satisfies (A.7) and $V_0(p) \geq U^S(p)$ for some $p \in [0,1]$, then V_0 satisfies (A.6) at p, and $\alpha = 0$ is a maximizer.*
(b) If V_1 satisfies (A.8) and $V_1(p) \geq U^S(p)$ for some $p \in [0,1]$, then V_1 satisfies (A.6) at p, and $\alpha = 1$ is a maximizer.
The maximizers are unique if $V_0(p), V_1(p) > U^S(p)$.

As we have argued in the previous section, $V_{opp}(p) \geq U^S(p)$ for all $p \in [0,1]$, which implies $V_{Env}(p) \geq U^S(p)$. Remember that $V_{Env}(p)$ is constructed of functions that satisfy (A.7) or (A.8), respectively. Therefore, Lemma 3 shows that $V_{Env}(p)$ satisfies the HJB equation for all points where it is differentiable. We have thus verified optimality for all p where $V_{Env}(p)$ is differentiable. For verification at points where $V_{Env}(p)$ is not differentiable, we show that it is a viscosity solution of the HJB equation (see the proof of Proposition 9 in Appendix B in the Supplemental Material).

A.8 Proof of Theorem 1

We show that Theorem 1 holds with the cutoffs:

$$\bar{c} := 0 \vee \frac{\lambda\left(u_r^R - u_\ell^R\right)\left(u_\ell^L - u_r^L\right) - \rho\left(u_r^R u_\ell^L - u_\ell^R u_r^L\right)}{\left(u_r^R - u_\ell^R\right) + \left(u_\ell^L - u_r^L\right)}, \tag{A.17}$$

$$\underline{c} := 0 \vee \begin{cases} \bar{c} \wedge \min\left\{ \frac{(\rho+\lambda)\left(u_r^R - u_\ell^R\right)}{1+\left(\frac{2\rho+\lambda}{\lambda}\right)^{\lambda/\rho}} - \rho u_r^R, \frac{(\rho+\lambda)\left(u_\ell^L - u_r^L\right)}{1+\left(\frac{2\rho+\lambda}{\lambda}\right)^{\lambda/\rho}} - \rho u_\ell^L \right\} & \text{if } \rho > 0, \\ \bar{c} \wedge \frac{\lambda}{1+e^2} \min\left\{\left(u_r^R - u_\ell^R\right), \left(u_\ell^L - u_r^L\right)\right\} & \text{if } \rho = 0, \end{cases} \tag{A.18}$$

where $x \vee y = \max\{x, y\}$ and $x \wedge y = \min\{x, y\}$.

[48] This follows from the fact that $\overline{V}_{own}(p^*) < \overline{V}_{opp}(p^*) = \underline{V}_{opp}(p^*)$. Since both $\overline{V}_{own}$ and $\underline{V}_{opp}$ satisfy (A.7), the former stays below the latter for all $p < p^*$.

Proof of Theorem 1. Straightforward algebra shows that (EXP) is equivalent to

$$c\left(\left(u_r^R-u_\ell^R\right)+\left(u_\ell^L-u_r^L\right)\right)+\rho\left(u_r^R u_\ell^L-u_\ell^R u_r^L\right)<\lambda\left(u_r^R-u_\ell^R\right)\left(u_\ell^L-u_r^L\right).$$

By Propositions 7, immediate action is optimal for all $p\in[0,1]$, if (EXP) is violated, which holds if and only if $c\geq\overline{c}$, where $\overline{c}$ is given by (A.17). This proves part (a).

Conversely, if $c\leq\overline{c}$, then (EXP) is satisfied, and by Propositions 8 and 9 experimentation is optimal for some beliefs. We show that

$$c\geq\underline{c}\qquad\Longleftrightarrow\qquad\max\left\{\underline{V}_{own}(p^*),\overline{V}_{own}(p^*)\right\}\geq U^S(p^*).\tag{A.19}$$

By Propositions 8 and 9, and the Unimprovability Lemma 3, this implies that the policies stated in Parts (b) and (c) of Theorem 1 are optimal.

We first assume $\rho>0$. The closed-form solutions for $\underline{V}_{own}(p^*)$ and $\overline{V}_{own}(p^*)$ in (A.11) and (A.12) can be used to show that

$$\begin{aligned}&\max\left\{\underline{V}_{own}(p^*),\overline{V}_{own}(p^*)\right\}\geq U^S(p^*)\\ \Longleftrightarrow\;&\left(\max\left\{\frac{c+\rho u_r^R}{\lambda u_r^R-(\rho+\lambda)u_\ell^R-c},\frac{c+\rho u_\ell^L}{\lambda u_\ell^L-(\rho+\lambda)u_r^L-c}\right\}\right)^{\frac{\rho}{\lambda}}\geq\frac{\lambda}{2\rho+\lambda}\;\Longleftrightarrow\;c\geq\underline{c}.\end{aligned}$$

This proves (A.19) for $\rho>0$. Taking the limit $r\to 0$ yields the result for $r=0$.[49]

By definition we have $\overline{c}\geq\underline{c}\geq 0$. It remains to show that the inequalities are strict for ρ sufficiently small. For $\rho=0$, $\underline{c}>0$ and

$$\begin{aligned}\overline{c}&=\lambda\min\left\{\left(u_r^R-u_\ell^R\right),\left(u_\ell^L-u_r^L\right)\right\}\frac{\max\left\{\left(u_r^R-u_\ell^R\right),\left(u_\ell^L-u_r^L\right)\right\}}{\left(u_r^R-u_\ell^R\right)+\left(u_\ell^L-u_r^L\right)}\\&>\lambda\min\left\{\left(u_r^R-u_\ell^R\right),\left(u_\ell^L-u_r^L\right)\right\}\frac{1}{2}>\underline{c}.\end{aligned}$$

Since both cutoffs are continuous in ρ, $\overline{c}>\underline{c}>0$ for ρ in a neighborhood of zero. □

References

Arrow, K. J., D. Blackwell, and M. A. Girshick (1949): "Bayes and Minimax Solutions of Sequential Decision Problems," *Econometrica*, pp. 213–244.

Bardi, M., and I. Capuzzo-Dolcetta (1997): *Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations*. Birhäuser.

Bronshtein, I. N., K. A. Semendyayev, G. Musiol, and H. Muehlig (2007): *Handbook of Mathematics*. Springer, 5th edn.

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

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

Calvert, R. L. (1985): "The Value of Biased Information: A Rational Choice Model of Political Advice," *Journal of Politics*, 47(2), 530–555.

Caplin, A., M. Dean, and J. Leahy (2018): "Rationally Inattentive Behavior: Characterizing and Generalizing Shannon Entropy," working paper.

Chan, J., A. Lizzeri, W. Suen, and L. Yariv (2018): "Deliberating Collective Decisions," *Review of Economic Studies*, 85(2), 929–963.

Che, Y.-K., and K. Mierendorff (2017): "Optimal Sequential Decision with Limited Attention," http://www.homepages.ucl.ac.uk/~uctpkmi/papers/CM-2017-07-02.pdf.

Chiang, C.-F., and B. Knight (2011): "Maedia Bias and Influence: Evidence from Newspaper Endorsements," *Review of Economic Studies*, 78, 795–820.

Damiano, E., H. Li, and W. Suen (2017): "Learning while Experimenting," unpublished.

DellaVigna, S., and E. Kaplan (2007): "The Fox News Effect: Media Bias and Voting," *Quarterly Journal of Economics*, 122(3), 1187–1234.

Esteban, J., and D. Ray (2012): "Comparing Polarization Measures," in *The Oxford Handbook of the Economics of Peace and Conflict*, ed. by M. R. Garfinkel, and S. Skaperdas. Oxford University Press.

Fudenberg, D., P. Strack, and T. Strzalecki (2017): "Stochastic Choice and Optimal Sequential Sampling," Available at SSRN 2602927.

Gentzkow, M., and J. Shapiro (2011): "Ideological Segregation Online and Offline," *Quarterly Journal of Economics*, 126, 1799–1839.

Gentzkow, M., and J. M. Shapiro (2010): "What drives Media Slant? Evidence from U.S. Daily Newspapers," *Econometrica*, 78(1), 35–71.

Gentzkow, M., J. M. Shapiro, and D. F. Stone (2015): *Bias in the Marketplace: Theory*chap. 14, pp. 623–645, Handbook of Media Economics 1B. Elsevier.

Hebert, B., and M. Woodford (2017): "Rational Inattention with Sequential Information Sampling," mimeo, Columbia University and Stanford University.

Henry, E., and M. Ottaviani (2017): "Research and Approval Process: The Organization of Persuasion," *Economic Journal*, forthcoming.

Ke, T. T., and J. M. Villas-Boas (2016): "Optimal Learning before Choice," unpublished manuscript, MIT.

Keller, G., S. Rady, and M. Cripps (2005): "Strategic Experimentation with Exponential Bandits," *Econometrica*, 73(1), 39–68.

Klein, N., and S. Rady (2011): "Negatively Correlated Bandits," *The Review of Economic Studies*, 78(2), 693–732.

Martin, G. J., and A. Yurukoglu (2017): "Bias in Cable News: Persuasion and Polarization," *American Economic Review*, 107(9), 2565–2599.

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.

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

McClellan, A. (2017): "Experimentation and Approval Mechanisms," unpublished working paper, New York University.

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

Morris, S., and P. Strack (2017): "The Wald Problem and the Equivalence of Sequential Sampling and Static Information Costs," Available at SSRN 2991567.

Moscarini, G., and L. Smith (2001): "The Optimal Level of Experimentation," *Econometrica*, 69(6), 1629–1644.

Mullainathan, S., and A. Shleifer (2005): "The Market for News," *American Economic Review*, 95(4), 1031–1053.

Nikandrova, A., and R. Pancs (2018): "Dynamic Project Selection," *Theoretical Economics*, 13(1), 115–144.

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

Peskir, G., and A. Shiryaev (2006): *Optimal Stopping and Free-Boundary Problems.* Birkhäuser Verlag.

Prat, A., and D. Strömberg (2013): "The Political Economy of Mass Media," in *Advances in Economics and Econometrics*, ed. by D. Acemoglu, M. Arellano, and E. Dekel, vol. 2 of *Econometric Society Monographs*, pp. 135–187.

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

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

Steiner, J., C. Stewart, and F. Matějka (2017): "Rational Inattention Dynamics: Inertia and Delay in Decision-Making," *Econometrica*, 85(2), 521–553.

Stroud, N. J. (2008): "Media Use and Political Predispositions: Revisiting the Concept of Selective Exposure," *Political Behavior*, 30, 341–366.

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

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

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

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

Supplemental Material (for online publication)

B Remaining Proofs from Section 4

B.1 Omitted Proofs from Appendix A.

B.1.1 Proof of Proposition 7

Proof. Since the DM can stop immediately, we have $V^*(p) \geq U(p)$. For the second inequality, consider the problem of a decision maker who can choose $\alpha_t \in [0,1]$ and $\beta_t \in [0,1]$ without the constraint that $\alpha_t + \beta_t = 1$. Clearly the value of this problem exceeds $V^*(p)$ for all p. The value function of the unconstrained problem is $\max\left\{U(p), U^{FA}(p)\right\}$. To see this, it is optimal to choose $\alpha_t = \beta_t = 1$. Given this policy, the belief does not change over time if no breakthrough occurs. The optimal policy is therefore either to stop immediately or to wait without deadline until a breakthrough occurs. Hence the value of the unconstrained problem is $\max\left\{U(p), U^{FA}(p)\right\}$. Therefore $V^*(p) = U(p) = \max\left\{U(p), U^{FA}(p)\right\}$ if (EXP) is violated. □

B.1.2 Proof of Lemma 2

Proof. Suppose $V_0(p) = V_1(p) = V(p)$ for some $p \in (0,1)$. Solving (A.7) and (A.8) for $V_0'(p)$ and $V_1'(p)$ and some algebra yields

$$V_0'(p) - V_1'(p) = \frac{\lambda + 2\rho}{\lambda p(1-p)} \left(V(p) - U^S(p)\right).$$

Therefore $\operatorname{sgn}(V_0'(p) - V_1'(p)) = \operatorname{sgn}\left(V(p) - U^S(p)\right)$. □

B.1.3 Proof of Lemma 3

Proof. Consider first the case that $V_0(p)$ satisfies (A.7). With $V = V_0(p)$, and substituting $V' = V_0'(p)$ from (A.7), we have

$$\frac{\partial F_\alpha(p, V_0(p), V_0'(p))}{\partial \alpha} = \frac{2\rho + \lambda}{\lambda} \left(U^S(p) - V_0(p)\right).$$

This implies that $\alpha = 0$ is a maximizer if $V_0(p) \geq U^S(p)$, and the unique maximizer if the inequality is strict. This proves Part (a). The proof of Part (b) follows from a similar argument. □

B.1.4 Proof of Proposition 8

The following three lemmas establish properties of the function U^S, V^{FA}, V_{own} and V_{opp} that are used in the proof of Proposition 8. Some of these properties were already established in Appendix A and are repeated here for convenience.

Lemma 4 (Properties of $U^S(p)$ and $U^{FA}(p)$)**.** *(a) $U^S(p) < U^{FA}(p)$ for all $p \in [0,1]$.*

(b) $U^S(p)$ and $U^{FA}(p)$ are linear in p.

If $U^S(p) \geq U(p)$ for some $p \in [0,1]$, then $U_\ell'(p) < U^{S\prime}(p) < U_r'(p)$ for all $p \in [0,1]$.

If $U^{FA}(p) \geq U(p)$ for some $p \in [0,1]$, then $U_\ell'(p) < U^{FA\prime}(p) < U_r'(p)$ for all $p \in [0,1]$.

(c) $U^S(p), U^{FA}(p) < U(p)$ at $p \in \{0,1\}$; and for all $p \in [0,1]$, $U^S(p)$ and $U^{FA}(p)$ are strictly decreasing without bound in c.

Proof. (a) $U^S(p) < U^{FA}(p)$ is immediate from the expressions in (A.2) and (A.3).

(b) Linearity is obvious. Suppose $U^S(p) \geq U(p)$ for some $p \in [0,1]$. To show $U_\ell'(p) < U^{S\prime}(p)$ for all p, suppose by contradiction that $U^{S\prime}(p) \leq U_\ell'(p)$ for some p. Note that $U^S(0) = \frac{u_\ell^L \lambda - 2c}{\lambda + 2\rho} < u_\ell^L = U_\ell(0)$. Hence, $U^{S\prime}(p) \leq U_\ell'(p)$ and the linearity of these functions imply $U^S(p) < U_\ell(p) \leq U(p)$ for all p, which is a contradiction. The other inequalities are proven similarly.

Part (c) is obtained from straightforward algebra. $\square$

The following lemma summarizes the properties of the own-biased strategy:

Lemma 5. *(a) $\underline{V}_{own}(p)$ and $\overline{V}_{own}(p)$ are continuously differentiable and convex on $(0,1)$;*

(b) $\underline{V}_{own}(p)$ is strictly convex and $\underline{V}_{own}(p) > U_\ell(p)$ on $(\underline{p}^*, 1]$, and $\overline{V}_{own}(p)$ is strictly convex and $\overline{V}_{own}(p) > U_r(p)$ on $[0, \overline{p}^*)$. $V_{own}(p) > U(p)$ for $p \in (\underline{p}^*, \overline{p}^*)$.

(c) If $\underline{p}^*, \overline{p}^* \in (0,1)$, they satisfy

$$U_\ell(\underline{p}^*) = U^{FA}(\underline{p}^*), \quad \text{and} \quad U_r(\overline{p}^*) = U^{FA}(\overline{p}^*). \tag{B.1}$$

(d) Suppose (EXP) holds. Then, $0 < \underline{p}^* < \overline{p}^* < 1$, $\underline{V}_{own}(p) < U^{FA}(p)$ for $p \in (\underline{p}^*, 1)$, $\overline{V}_{own}(p) < U^{FA}(p)$ for $p \in (0, \overline{p}^*)$, and $V_{own}(p) = U(p) > U^{FA}(p)$ for $p \notin [\underline{p}^*, \overline{p}^*]$.

(e) If (EXP) is violated, then $V_{own}(p) = U(p)$ for all $p \in [0,1]$.

Proof. Parts (a)-(c) follow from straightforward algebra. For part (d), note that (EXP) together with part (c) and Lemma 4.(b) imply $0 < \underline{p}^* < \hat{p} < \overline{p}^* < 1$ and $U^{FA}(p) < U(p)$ for $p \notin [\underline{p}^*, \overline{p}^*]$. This implies $V_{own}(p) = U(p) > U^{FA}(p)$ for $p \notin [\underline{p}^*, \overline{p}^*]$. To show that $\underline{V}_{own}(p) < U^{FA}(p)$ for $p \in (\underline{p}^*, 1)$, note that $\underline{V}_{own}(\underline{p}^*) = U_\ell(\underline{p}^*) = U^{FA}(\underline{p}^*)$ from part (c), and $\underline{V}_{own}(1) = U^{FA}(1)$ from (A.11). Since $U^{FA}(p)$ is linear by Lemma 4 and $\underline{V}_{own}(p)$ is strictly convex $(\underline{p}^*, 1]$ by part (b), this implies implies that $\underline{V}_{own}(p) < U^{FA}(p)$ for $p \in (\underline{p}^*, 1)$. $\overline{V}_{own}(p) < U^{FA}(p)$ for $p \in (0, \overline{p}^*)$ is proven similarly.

Part (e) holds because by part (c), $\underline{p}^* > \overline{p}^*$ if (EXP) is violated. $\square$

We next observe several properties of $V_{opp}(p)$.

Lemma 6. *(a) $V_{opp}(p)$ is continuously differentiable and strictly convex on $(0,1)$, and $V_{opp}(p) \geq U^S(p)$ for all $p \in [0,1]$ with strict inequality for $p \neq p^*$.*

(b) Then, $V_{opp}(p) \leq U^{FA}(p)$ for all $p \in [0,1]$, with equality if and only if $p \in \{0,1\}$.

Proof. Part (a) follows from straightforward algebra. For part (b), again by straightforward algebra we get $U^{FA}(0) = \underline{V}_{opp}(0) = V_{opp}(0)$ and $U^{FA}(1) = \overline{V}_{opp}(0) = V_{opp}(1)$. Since $U^{FA}(p)$ is linear and V_{opp} is strictly convex, this implies $V_{opp}(p) < U^{FA}(p)$ for all $p \in [0,1]$. □

We are now ready to prove Proposition 8. For the reader's convenience, we restate the proposition.

Proposition (Structure of V_{Env}). (a) If (EXP) holds and $V_{own}(p^*) \geq V_{opp}(p^*)$, then there exists a unique $\check{p} \in \left(\underline{p}^*, \overline{p}^*\right)$ such that $\underline{V}_{own}(\check{p}) = \overline{V}_{own}(\check{p})$ and

$$V_{Env}(p) = V_{own}(p) = \begin{cases} \underline{V}_{own}(p), & \text{if } p < \check{p}, \\ \overline{V}_{own}(p), & \text{if } p \geq \check{p}. \end{cases}$$

(b) If (EXP) holds and $V_{own}(p^*) < V_{opp}(p^*)$, then $p^* \in (\underline{p}^*, \overline{p}^*)$, and there exist a unique $\underline{p} \in (\underline{p}^*, p^*)$ such that $V_{own}(\underline{p}) = V_{opp}(\underline{p})$, and a unique $\overline{p} \in (p^*, \overline{p}^*)$ such that $V_{own}(\overline{p}) = V_{opp}(\overline{p})$ and

$$V_{Env}(p) = \begin{cases} \underline{V}_{own}(p), & \text{if } p < \overline{p}, \\ V_{opp}(p), & \text{if } p \in [\underline{p}, \overline{p}], \\ \overline{V}_{own}(p), & \text{if } p > \overline{p}. \end{cases}$$

Proof. **Part (a):** We first prove that $V_{own}(p) \geq V_{opp}(p)$ for all $p \in [0,1]$. Since $V_{own}(p) \geq U^{FA}(p) > V_{opp}(p)$ for $p \notin [\underline{p}^*, \overline{p}^*]$, it suffices to show $V_{own}(p) \geq V_{opp}(p)$ for $p \in [\underline{p}^*, \overline{p}^*]$. To this end, suppose first $p^* > \underline{p}^*$ and consider $p \in [\underline{p}^*, p^*]$ so that $V_{opp}(p) = \underline{V}_{opp}(p)$. Recall from Lemmas 5 and 6 that $\underline{V}_{own}(\underline{p}^*) = U^{FA}(\underline{p}^*) > V_{opp}(\underline{p}^*)$. Since $V_{opp}(\cdot) \geq U^S(\cdot)$, by the Crossing Lemma 2, $\underline{V}_{own}$ can cross $V_{opp} = \underline{V}_{opp}(p)$ only from above on $[\underline{p}^*, p^*)$. If $\underline{V}_{own}(p^*) \geq \underline{V}_{opp}(p)(p^*)$, by the Crossing Lemma 2, $\underline{V}_{opp}(p) < \underline{V}_{own}(p) \leq V_{own}(p)$ for all $p \in [\underline{p}^*, p^*]$. If $\underline{V}_{own}(p^*) < \underline{V}_{opp}(p^*)$, then $\overline{V}_{own}(p^*) = V_{own}(p^*) \geq \underline{V}_{opp}(p^*)$. Since both $\overline{V}_{own}(p)$ and $\underline{V}_{opp}(p^*)$ satisfy (A.7), we must have $\underline{V}_{opp}(p) \leq \overline{V}_{own}(p) \leq V_{own}(p)$ for all $p \in [\underline{p}^*, p^*]$. Either way, we have proven that $V_{opp}(p) = \underline{V}_{opp}(p) \leq V_{own}(p)$ for all $p \in [\underline{p}^*, p^*]$. A symmetric argument proves that $V_{opp}(p) \leq V_{own}(p)$ for all $p \in [p^*, \overline{p}^*]$ in case $p^* < \overline{p}^*$.

We have now proven that $V_{own}(p) \geq V_{opp}(p)$ for all $p \in [0,1]$. Recall from Lemma 5 that $\underline{V}_{own}(\underline{p}^*) = U^{FA}(\underline{p}^*) > \overline{V}_{own}(\underline{p}^*)$ and $\overline{V}_{own}(\overline{p}^*) = U^{FA}(\overline{p}^*) > \underline{V}_{own}(\overline{p}^*)$. By the intermediate value theorem, there exists $\check{p} \in \left(\underline{p}^*, \overline{p}^*\right)$ where $\underline{V}_{own}(\check{p}) = \overline{V}_{own}(\check{p})$. For any p we have $V_{own}(p) \geq V_{opp}(p)$ and $V_{opp}(p) \geq U^S(p)$ and hence $V_{own}(\check{p}) \geq U^S(\check{p})$. The Crossing

Lemma 2 then implies that $\underline{V}_{own}$ cannot cross $\overline{V}_{own}$ from below at $\check{p}$.[50] This means that the intersection point $\check{p}$ is unique and the structure stated in part (a) obtains.

Part (b): We first prove that $p^* \in \left(\underline{p}^*, \overline{p}^*\right)$. By Lemma 5, $V_{own}(p) \geq U(p)$ for all $p \in [0,1]$. This implies $V_{opp}(p^*) > U(p^*)$, and since $V_{opp}(p^*) = U^S(p^*) < U^{FA}(p^*)$, and since by Lemma 5.(d) $U^{FA}(p) \leq U(p)$ for $p \notin \left(\underline{p}^*, \overline{p}^*\right)$, we must have $p^* \in \left(\underline{p}^*, \overline{p}^*\right)$. Next, by Lemma 6.(b), $V_{opp}(\underline{p}^*) < U^{FA}(\underline{p}^*) = \underline{V}_{own}(\underline{p}^*)$. Therefore, $V_{opp}(p)$ and $\underline{V}_{own}(p)$ intersect at some $\underline{p} \in \left(\underline{p}^*, p^*\right)$ and by the Crossing Lemma 2, the intersection is unique since $V_{opp}(\underline{p}) > U^S(\underline{p})$ for $\underline{p} \in \left(\underline{p}^*, p^*\right)$ by Lemma 6.(a). Moreover, for $p < p^*$, we have $V_{opp}(p) > \overline{V}_{own}(p)$ since both satisfy (A.7), and hence $\overline{V}_{own}(p) < \underline{V}_{own}(p)$ for all $p \in \left(\underline{p}^*, \underline{p}\right)$. This proves the result for $p \leq p^*$. For $p > p^*$ the arguments are symmetric. □

B.1.5 Proof of Proposition 9

Proof. If (EXP) is violated, $V_{Env}(p) = U(p)$ since $\underline{p}^* > \overline{p}^*$ by Proposition 7. Moreover Proposition 7 shows that $V^*(p) = U(p) = V_{Env}(p)$ in this case. Similarly, if (EXP) is satisfied, by Lemma 1 and Proposition 8, $V_{Env}(p) = U(p)$ for all $p \notin (\underline{p}^*, \overline{p}^*)$ and Proposition 7 shows that $V^*(p) = U(p) = V_{Env}(p)$ for $p \notin (\underline{p}^*, \overline{p}^*)$.

It remains to verify $V^*(p) = V_{Env}(p)$ for $p \in (\underline{p}^*, \overline{p}^*)$ when EXP is satisfied. In the remainder of this proof we write $V(p) = V_{Env}(p)$. 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 (A.1) as the (unique) viscosity solution of the HJB equation.[51] For all $p \in (0,1)$ where $V(p)$ is differentiable, this requires that $V(p)$ satisfy (A.4).

Consider points of differentiability $p \in \left(\underline{p}^*, \overline{p}^*\right)$. From (A.11) and (A.12), we obtain that $\underline{V}_{own}$ and $\overline{V}_{own}$ are strictly convex on $(\underline{p}^*, \overline{p}^*)$. Smooth pasting at $\underline{p}^*$ and $\overline{p}^*$, respectively, implies that $\underline{V}_{own}(p) > U_\ell(p)$ and $\overline{V}_{own}(p) > U_r(p)$, and therefore $V_{own}(p) > U(p)$ for $p \in (\underline{p}^*, \overline{p}^*)$. This implies that (A.4) is equivalent to (A.6) for all $p \in \left(\underline{p}^*, \overline{p}^*\right)$. Since $V(p)$ satisfies (A.7) or (A.8) at points of differentiability, and $V(p) \geq V_{opp}(p) \geq U^S(p)$, the Unimprovability Lemma 3 implies that $V(p)$ satisfies (A.6). Since V_{opp} is strictly convex (see the discussion after Lemma 2), $V_{opp}(p) > U^S(p)$, and hence Lemma 3 implies that the optimal policy is unique at all points where $V(p)$ is differentiable except p^*. At p^*, the HJB equation is satisfied for any $\alpha \in [0,1]$ but $\alpha = 1/2$ is the only maximizer that defines an admissible policy.

We have shown that $V(p)$ satisfies (A.4) for all points of differentiability. For $V(p)$ to

[50] $\underline{V}_{own}$ and $\overline{V}_{own}$ could be equal to U^S at $\check{p}$ which means that two branches are tangent. However, the convexity of both branches and the fact that $V_{own}(p) \geq U^S(p)$ for all p, means that $\underline{V}_{own}$ cannot cross $\overline{V}_{own}$ from below at any point of intersection. Therefore $\check{p}$ is unique.

[51] To formally apply their theorem, we have to use 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 $V(p)$.

be a viscosity solution it remains to show that for all points of non-differentiability,

$$\max\{-c-\rho V(p)+F(p,V(p),\rho),U(p)-V(p)\}\leq 0, \tag{B.2}$$

for all $\rho\in\left[V'_-(p),V'_+(p)\right]$; and the opposite inequality holds for all $\rho\in\left[V'_+(p),V'_-(p)\right]$, where $V'_-(p)$ denotes the left derivative at p, and $V'_+(p)$ denotes the right derivative at p. By Proposition 8, non-differentiability at $\check{p}$ if (EXP) holds and $V_{own}(p^*)\geq V_{opp}(p^*)$; and at $\underline{p}$ and $\overline{p}$ if (EXP) holds and $V_{own}(p^*)<V_{opp}(p^*)$. Since $V(p)\geq U^S(p)$, the Crossing Lemma 2 implies that $V(p)$ has convex kinks at all these points so that $V'_-(p)\leq V'_+(p)$. Therefore it suffices to check (B.2) for all $\rho\in\left[V'_-(p),V'_+(p)\right]$. F_α is linear in α (see (A.5)), so it suffices to consider $\alpha\in\{0,1\}$. For $\alpha=1$ we have $F_1(p,V(p),\rho)\leq F_1(p,V(p),V'_-(p))$ and for $\alpha=0$ we have $F_0(p,V(p),\rho)\leq F_0(p,V(p),V'_+(p))$. Therefore if $U(p)\leq V(p)$, which holds for our candidate solution by construction,

$$c+\rho V(p)\geq \max\left\{F_1(p,V(p),V'_-(p)),F_0(p,V(p),V'_+(p))\right\} \tag{B.3}$$

implies that (B.2) holds for all for $\rho\in\left[V'_-(p),V'_+(p)\right]$. We distinguish two cases.

Case A: (EXP) *is satisfied and* $V_{own}(p^*)\geq V_{opp}(p^*)$. Consider $p=\check{p}$. (B.3) becomes

$$c+\rho\underline{V}_{own}(\check{p})=c+\rho\overline{V}_{own}(\check{p})\geq \max\left\{F_1(\check{p},\underline{V}_{own}(\check{p}),\underline{V}'_{own}(\check{p})),F_0(\check{p},\overline{V}_{own}(\check{p}),\overline{V}'_{own}(\check{p}))\right\}.$$

By the Unimprovability Lemma 3, this holds with equality since $\underline{V}_{own}(p)$ satisfies (A.8) and $\overline{V}_{own}(p)$ satisfies (A.7) at $\check{p}$. As we have argued earlier, $V_{own}(p)>U(p)$ for all $p\in\left(\underline{p}^*,\overline{p}^*\right)$ and hence $V(\check{p})>U(\check{p})$. (B.2) is thus satisfied at $\check{p}$.

Case B: (EXP) *is satisfied and* $V_{own}(p^*)<V_{opp}(p^*)$. The proof is similar to Case B.

We have thus shown that $V(p)$ is a viscosity solution of (A.4) which is sufficient for $V(p)$ to be the value function of problem (A.1). □

B.2 Proof of Proposition 1

Proof. (a) Denote the expected delay until the DM takes an action by $\tau(p)$. At p^* the DM uses $\alpha=1/2$. Hence the arrival rate of a signal is $\lambda/2$ and the expected delay is given by the expectation of the exponential distribution:

$$\tau(p^*)=\frac{2}{\lambda}.$$

For $p_0\in(p^*,\overline{p})$, the expected delay must satisfy a recursive equation with respect to any t:

$$\tau(p_0)=\int_0^t s(p_0\lambda e^{-\lambda s})ds+(p_0e^{-\lambda t})(\tau(p_t)+t).$$

Differentiating both sides by t yields

$$0 = (p_0 e^{-\lambda t})(\tau'(p_t)\dot{p}_t + 1) - (\lambda p e^{-\lambda t})\tau(p_t),$$

which, upon setting $t = 0$, reduces to:

$$\tau'(p) = \frac{1 - (\lambda p)\tau(p)}{p(1-p)\lambda}.$$

Solving this differential equation with boundary condition $\tau(p^*) = \frac{2}{\lambda}$ and some algebra yields $\tau''(p) < 0$. Moreover the right derivative of τ at p^* is given by

$$\tau'(p^*_+) = \frac{1 - 2p^*}{p^*(1-p^*)\lambda}.$$

Using similar steps for $p_0 \in (\underline{p}, p^*)$ we have $\tau'(p_t) = \frac{\lambda(1-p)\tau(p_t)-1}{p(1-p)\lambda}$ and $\tau''(p) < 0$. The left derivative of τ at p^* is given by $\tau'(p^*_-) = \tau'(p^*_+)$. Since τ is concave on $\left(\underline{p}, p^*\right)$ and on $(p^*, \overline{p})$ and $\tau'(p^*_-) = \tau'(p^*_+)$, we conclude that τ is concave on $\left(\underline{p}, \overline{p}\right)$.

To show that τ is quasi-concave, it remains to show that τ is decreasing on $[\overline{p}, \overline{p}^*]$ and increasing on $\left[\underline{p}^*, \underline{p}\right]$. Since the argument is essentially the same for both cases, we consider $[\overline{p}, \overline{p}^*]$. The expected delay implied by the own-biased strategy is

$$\tau(p) = \int_0^{\overline{T}^*(p)} s((1-p)\lambda e^{-\lambda s})ds + (1-p)e^{-\lambda \overline{T}^*(p)}\overline{T}^*(p).$$

where $\overline{T}^*(p)$ is the time it takes for the belief to reach $\overline{p}^*$ in the absence of a signal if the DM follows the own-biased strategy (i.e., seeks L-signals). Since $\overline{T}^*(p)$ is decreasing in p we have

$$\tau'(p) = (1-p)e^{-\lambda \overline{T}^*}\overline{T}^{*\prime}(p) < 0.$$

Therefore it remains to show that $\tau(\overline{p}_-) \geq \tau(\overline{p}_+)$.

Suppose $r = 0$. If at $\overline{p}$, the DM follows the own-biased strategy, she enjoys the payoff of

$$\left[\overline{p}u_r^R + (1-\overline{p})u_\ell^L\right] - \overline{p}\left[u_r^R - u_\ell^R\right] - c\int_0^{\overline{T}^*(\overline{p})} (1 - H(t)))dt. \tag{B.4}$$

where H is the distribution of the time at which the DM makes a decision.

Suppose instead that the DM follows the opposite-biased strategy. In this case her expected payoff (for $r = 0$) is given by

$$\left[\overline{p}u_r^R + (1-\overline{p})u_\ell^L\right] - c\int_0^{\infty} (1 - G(t)))dt. \tag{B.5}$$

where G is the distribution of time at which the DM makes a decision.

Since at $\overline{p}$, the DM is indifferent between both strategies we must have

$$\int_0^{\overline{T}^*(\overline{p})} (1 - H(t)))dt < \int_0^{\infty} (1 - G(t)))dt,$$

i.e., the DM will take a longer time for decision if she chooses a opposite-biased strategy instead. This proves part (a) of the Proposition for $r = 0$ if $c < \underline{c}$. By continuity the result extends to r in a neighborhood of zero. For the case that $c \in [\underline{c}, \overline{c})$, it suffices to invoke the argument used for $(\overline{p}, \overline{p}^*)$ for the whole interval $(\check{p}, \overline{p}^*)$.

(b) Consider $p > \check{p}$ so that the DM uses $\alpha = 0$ according to the opposite-biased strategy. Inserting this in (2.1) and integrating we get

$$p_t = \frac{e^{t\lambda} p_0}{1 + (e^{t\lambda} - 1)\, p_0}.$$

Setting $p_{\overline{T}^*} = \overline{p}^*$ and solving for $\overline{T}^*$ we get

$$\overline{T}^* = \frac{1}{\lambda} \log \left(\frac{\overline{p}^*}{\overline{p}^* - \overline{p}^*} \frac{1 - p_0}{p_0} \right).$$

The probability of a mistake is therefore

$$(1 - p) \left(1 - e^{-\lambda \overline{T}^*} \right) = \frac{(1 - \overline{p}^*)\,(\overline{p}^* - p_0)}{\overline{p}^*(1 - p_0)}.$$

Differentiating this with respect to p_0, we get

$$-\frac{\overline{p}^* - p_0}{\overline{p}^*(1 - p_0)} < 0.$$

This proves that the probability of a mistake decreases in the distance to $\overline{p}^*$ for high p. For low p the argument is symmetric. □

B.3 Proof of Proposition 2

Proof. (a) By (B.1), $\underline{p}^*$ and $\overline{p}^*$ are given by the intersections of $U(p)$ and $U^{FA}(p)$. Since $U(p)$ is independent of r and c, and $U^{FA}(p)$ is strictly decreasing in both parameters, the experimentation region expands as r or c fall. As $(r, c) \to (0, 0)$, we have $U^{FA}(p) \to U(p)$ for $p \in \{0, 1\}$, hence the experimentation region converges to $(0, 1)$.

(b) The dependence of $\underline{p}^*$ and $\overline{p}^*$ on u_ℓ^R and u_r^L is straightforward from the expressions for the cutoffs in (A.9) and (A.10). (c) By (B.1), $\underline{p}^*$ is the intersection between $U_\ell(p)$ and $U^{FA}(p)$. The former is independent of u_r^R and the latter is increasing in u_r^R. Hence

$\partial \underline{p}^*/\partial u_r^R < 0$. Also by (B.1), $\overline{p}^*$ is the intersection between $U_r(p)$ and $U^{FA}(p)$. We have

$$\frac{\partial U_r(p)}{\partial u_r^R} = p > \frac{\lambda}{r+\lambda}p = \frac{\partial U^{FA}(p)}{\partial u_r^R}.$$

This implies that $\partial \overline{p}^*/\partial u_r^R < 0$. The comparative statics with respect to u_ℓ^L is derived similarly. □

B.4 Proof of Proposition 3

Proof. (a) We prove $\partial \check{p}/\partial u_\ell^R > 0$; the other case follows from a symmetric argument. Consider $\overline{V}_{own}(p)$. Since the right branch of the own-biased value function is obtained from a strategy that takes action ℓ only if a signal has been received, its value is independent of u_ℓ^R, as can be seen from (A.12). On the other hand we have $\partial \underline{V}_{own}(p)/\partial u_\ell^R > 0$ from (A.11). Therefore the point of intersection of $\underline{V}_{own}$ and $\overline{V}_{own}$ is increasing in u_ℓ^R.

(b) It is clear from (A.18) that $\underline{c}$ is decreasing in u_ℓ^R and u_r^L. Therefore, it suffices to consider the case that $c < \underline{c}$. We prove that $\underline{p} \to 0$ monotonically as $u_\ell^R \to -\infty$. If a opposite-biased region exists, $\underline{p} \in (\underline{p}^*, p^*)$ is defined as the unique intersection between $\underline{V}_{opp}(p)$ and $\underline{V}_{own}(p)$. Note that $\underline{V}_{opp}(p)$ is independent of u_ℓ^R since the opposite-biased strategy never leads to a mistake. As in (d) we have $\partial \underline{V}_{own}(p)/\partial u_\ell^R > 0$. Moreover, Lemma 2 shows that $\underline{V}_{own}(p)$ crosses $\underline{V}_{opp}(p)$ from above at $\underline{p}$. Since V_{opp} is independent of u_ℓ^R this implies that of $\underline{p}$ is monotonically increasing in u_ℓ^R.

Since $\underline{p}$ is bounded from below, there exists $q = \lim_{u_\ell^R \to -\infty} \underline{p} < p^*$. Suppose by contradiction that $q > 0$. Notice that, for each $p \in [q, p^*]$, as $u_\ell^R \to -\infty$.

$$\underline{V}_{own}(p) \to \frac{\lambda u_r^R pr - \lambda c(1-p) - cr}{(r+\lambda)\, r} =: \underline{V}_{own}^\circ(p),$$

where we used the fact that $\underline{p}^*/(1-\underline{p}^*) \to 0$ as $u_\ell^R \to -\infty$.

Note that the convergence is uniform on $[q, p^*]$ since $q > 0$.[52] Simple algebra yields

$$\underline{V}_{own}^\circ(p^*) \leq U^S(p^*),$$
$$\text{and} \quad \underline{V}_{own}^{\circ\prime}(p) > U^{S\prime}(p).$$

Since $\underline{V}_{own}^\circ(p)$ is linear in p, this implies that $\underline{V}_{opp}(q) \geq U^S(q) > \underline{V}_{own}^\circ(q)$ which is a contradiction and we must have $q = 0$. The proof for $\overline{p}$ is essentially the same. □

[52] Recall from (A.11) that $\underline{V}_{own}(p) \to \infty$ as $p \to 0$, hence the condition $q > 0$ is necessary here.

B.5 Proof of Proposition 4

Proof. (a) We have

$$\frac{\partial \overline{c}}{\partial r} = -\frac{u_r^R u_\ell^L - u_\ell^R u_r^L}{(u_r^R + u_\ell^L) - (u_\ell^R + u_r^L)},$$

and hence $\operatorname{sgn}(\partial \overline{c}/\partial r) = \operatorname{sgn}\left(u_\ell^R u_r^L - u_r^R u_\ell^L\right)$. It is straightforward to verify that $U(\hat{p}) > 0$ if and only if $u_r^R u_\ell^L - u_\ell^R u_r^L > 0$.

(b) Denoting $Z(\rho) := (\rho+1)/\left(1 + (2\rho+1)^{\frac{1}{\rho}}\right)$, we have

$$\frac{\partial \underline{c}}{\partial r} = \begin{cases} Z'(r/\lambda)\left(u_r^R - u_\ell^R\right) - r u_r^R & \text{if } (\lambda Z(r/\lambda) - r)\left(u_r^R - u_\ell^L\right) - \lambda Z(r/\lambda)\left(u_\ell^R - u_r^L\right) < 0, \\ Z'(r/\lambda)\left(u_\ell^L - u_r^L\right) - r u_\ell^L & \text{if } (\lambda Z(r/\lambda) - r)\left(u_r^R - u_\ell^L\right) - \lambda Z(r/\lambda)\left(u_\ell^R - u_r^L\right) > 0. \end{cases}$$

Consider the first case. Since $Z'(\rho) \in \left[(1+3e^2)/(1+e^2)^2, 1/2\right]$,

$$Z'(r/\lambda)\left(u_r^R - u_\ell^R\right) - u_r^R < \frac{1}{2}\left(u_r^R - u_\ell^R\right) - u_r^R = -\frac{1}{2}\left(u_r^R + u_\ell^R\right),$$

which is negative if $u_r^R > \left|u_\ell^R\right|$. Conversely, if u_ℓ^R is sufficiently negative $Z'(r/\lambda)\left(u_r^R - u_\ell^R\right) - u_r^R > 0$. The argument for the second case is similar. □

B.6 Proof of Proposition 5

Proof. Let $F_t(p)$ be the distribution function of beliefs in the whole population at time t. Denote the density, whenever it exists by $f_t(p)$. Denote by $\delta_t(p) = F_t(p) - F_t^-(p)$ the mass at p if there is a mass point. For $t = 0$ we have the uniform distribution $F_0(p) = p$.

(a) For part (a) we consider the subpopulation of voters with prior beliefs in $\mathcal{P}_{own} = [\underline{p}^*, \underline{p}] \cup [\overline{p}, \overline{p}^*]$. Initially, these voters consume own-biased news. If we consider the same subpopulation at later points $t > 0$, then their beliefs either remain inside $\mathcal{P}_{own}$, voters who have received an L-breakthrough, however, have a belief $p_t = 0$. Therefore, for $t > 0$ we consider the subpopulation of voters with beliefs in $\mathcal{P}_{own}^0 = \mathcal{P}_{own} \cup \{0\}$. Within $\mathcal{P}_{own}^0$ we consider the median belief for voters with $p_t > 1/2$, denoted m_t^r and the median belief for voters with $p_t < 1/2$, denoted by m_t^ℓ.

We first consider m_t^r which is given by

$$m_t^r = F_t^{-1}\left(F_t(\overline{p}) + \frac{F_t(\overline{p}^*) - F_t(\overline{p})}{2}\right)$$

We show that this is increasing in t whenever $m_t^r < \overline{p}^*$. All individuals in $\mathcal{P}_{own}^0 \cup (1/2, 1]$ consume R-biased news. This leads to two possible changes in their beliefs that effects the median. First, for voters who receive breakthrough news the belief becomes 0 so that they leave the set $\mathcal{P}^0_{own} \cup (1/2, 1]$. Note that conditional on the state being L all individuals who acquire information recieve L-breakthroughs at rate λ. If $m^r_t < \overline{p}^*$, all voters in $\mathcal{P}^0_{own} \cup (1/2, 1]$ below the median still acquire information but some voters above the median have already stoped. Therefore, more voters below the median recieve breakthrough than above the median. This increases the median.

Second absent a breakthrough the belief of a voter in $\mathcal{P}^0_{own} \cup (1/2, 1]$ drifts upwards. The upward drift also increses the median. Hence, if $m^r_t < \overline{p}^*$, m^r_t is increasing over time. If $m^r_t = \overline{p}^*$, it remains constant for all $t' > t$.

Next consider the subpopulation of individuals with beliefs in $\mathcal{P}^0_{own} \cup [0, 1/2)$. This subpopulation is composed of (i) the voters who initially consume own-biased news and have a prior $p_0 < 1/2$, and (ii) voters who initially consume own-biased news and have a prior of $p > 1/2$, but received breakthrough news at some time $t' \leq t$. The median belief at time t of individuals with beliefs below 1/2 in this subset is given by

$$m^\ell_r = \begin{cases} \underline{p}^*, & \text{if } F_t(0) \geq F_t(\underline{p}) - F_t(\underline{p}^*) \\ F_t^{-1}\left(F_t(\underline{p}) - \frac{\delta_t(0) + F_t(\underline{p}) - F_t^-(\underline{p}*)}{2}\right), & \text{otherwise.} \end{cases}$$

m^ℓ_t is moved by two forces. First, individuals with $p > 1/2$ who receive breakthroughs enter the population with $p < 1/2$, and since they have a belief $p = 0$ after the breakthrough this reduces the median. Second, indivduals with beliefs $p < 1/2$ who consume own-biased news never receive breakthroughs if the true state is L. Therefore their beliefs drift downwards which further decreases m^ℓ_t.

In summary we have shown that $m^r_t - m^\ell_t$ is increasing which concludes the proof of part (a).

Part (b) follows from similar arguments since all voters who consume any news choose own-biased news by assumption. Therefore their beleif dynamics as as in case (a). The reminaing voters do not consume any news so that their beliefs remain constant and leave the median in the subpopulations above and below 1/2 unaffected.

The proof of part (c) is immediate from the definition of the opposite-biased strategy.

□

C Extensions

C.1 Discrete Time Foundation

Proof of Proposition 6. If the DM chooses an experiment with parameters a and $b = 1 + \lambda dt - a$, then the posteriors are $q^R := p\,(\lambda dt + 1 - a)\,/\,(p\lambda dt + (1 - a))$ when the R-signal is received, and $q^L := p\,(a - \lambda dt)\,/\,(a - p\lambda dt)$ when an L-signal is received. The unconditional probabilities of the signals are Prob$\,[R\text{-signal}] = p\lambda dt + (1 - a)$, and

Prob [L-signal] $= a - p\lambda dt$. Hence the DM maximizes

$$\max_{a\in[\lambda dt,1]} (p\lambda dt + (1-a))\,\widetilde{V}(q^R) + (a - p\lambda dt)\,\widetilde{V}(q^L) \tag{C.1}$$

where $\widetilde{V}(q) = \max\left\{U(q), e^{-\rho dt}V(q) - c\right\}$ and $V(p)$ is the optimal value function. We note that $V(p)$ is weakly convex.[53] Therefore the *continuation value* $\widetilde{V}(p)$ is also weakly convex.

In the following, we fix an arbitrary weakly convex continuation value $\widetilde{V}$ and belief $p \in (0,1)$. We show that (C.1) is maximized by $\alpha = \lambda dt$ or $\alpha = 1$. To do this, we rewrite the objective in (C.1) for an arbitrary choice $\hat{a} \in [\lambda dt, 1]$ in a way that can be bounded by the value for $\alpha = \lambda dt$ or $\alpha = 1$.

So we fix any $\hat{a} \in [\lambda dt, 1]$ and denote the implied posteriors by $\hat{q}^R$ and $\hat{q}^L$. To rewrite the objective in (C.1), we construct alternative payoff parameters $\hat{u}_x^\omega$ so that the resulting stopping payoffs satisfy $\widehat{U}_\ell(\hat{q}^L) = \widetilde{V}(\hat{q}^L)$ and $\widehat{U}'_\ell(\hat{q}^L) = \widetilde{V}'(\hat{q}^L)$, as well as $\widehat{U}_\rho(\hat{q}^R) = \widetilde{V}(\hat{q}^R)$ and $\widehat{U}'_r(\hat{q}^R) = \widetilde{V}'(\hat{q}^R)$.[54] Theses conditions yields:

$$\begin{aligned}
\hat{u}_r^R &:= \widetilde{V}(\hat{q}^R) + (1-\hat{q}^R)\widetilde{V}'(\hat{q}^R), & \hat{u}_\ell^R &:= \widetilde{V}(\hat{q}^R) - \hat{q}^R\widetilde{V}'(\hat{q}^R),\\
\hat{u}_r^L &:= \widetilde{V}(\hat{q}^L) + (1-\hat{q}^L)\widetilde{V}'(\hat{q}^L), & \hat{u}_\ell^R &:= \widetilde{V}(\hat{q}^L) - \hat{q}^L\widetilde{V}'(\hat{q}^L).
\end{aligned}$$

By definition, $\widehat{U}(p)$ is tangent to $\tilde{V}(p)$ at $p = \hat{q}^L$ and at $p = \hat{q}^R$, and is everywhere weakly below $\widetilde{V}(p)$, given the convexity of $\widetilde{V}(p)$.

The objective in (C.1) for $\hat{a}$ can be rearranged and bounded as follows:

$$\begin{aligned}
&(p\lambda dt + (1-\hat{a}))\,\widetilde{V}(\hat{q}^R) + (\hat{a} - p\lambda dt)\,\widetilde{V}(\hat{q}^L)\\
=&\,(p\lambda dt + (1-\hat{a}))\,\widehat{U}(\hat{q}^R) + (\hat{a} - p\lambda dt)\,\widehat{U}(\hat{q}^L)\\
=&\,p(1+\lambda dt-\hat{a})\hat{u}_r^R + (1-p)(1-\hat{a})\hat{u}_r^L + p(a-\lambda dt)\hat{u}_\ell^R + (1-p)\hat{a}\hat{u}_\ell^L\\
\le&\max_{a\in\{\lambda dt,1\}} p(1+\lambda dt-a)\hat{u}_r^R + (1-p)(1-a)\hat{u}_r^L + p(a-\lambda dt)\hat{u}_\ell^R + (1-p)a\hat{u}_\ell^L\\
=&\,(p\lambda dt + (1-\hat{a}^*))\,\widehat{U}(\hat{q}^{R*}) + (\hat{a}^* - p\lambda dt)\,\widehat{U}(\hat{q}^{L^*})\\
\le&\,(p\lambda dt + (1-\hat{a}^*))\,\widetilde{V}(\hat{q}^{R*}) + (\hat{a}^* - p\lambda dt)\,\widetilde{V}(\hat{q}^{L^*}).
\end{aligned}$$

In the second line, we have replaced $\widetilde{V}$ by $\widehat{U}$. Writing this out in the third line, we see that the expression is linear in a. Therefore, maximizing over $a \in \{\lambda dt, 1\}$, we get a weakly higher value. In the fifth line $\hat{a}^*$ denotes a maximizer from the forth line and $\hat{q}^{\omega*}$ denotes the corresponding posterior beliefs. The last inequality follows from the fact that $\widetilde{V}$ is

[53] To see this, note that the expected value of a fixed strategy (i.e. a mapping that specifies the attention choice and action for each history) is linear in the prior belief. The value function is therefore the upper envelope of a family of linear functions, which implies convexity.

[54] We use the notation $\widehat{U}_\ell(p) := p\hat{u}_\ell^R + (1-p)\hat{u}_\ell^R$, $\widehat{U}_r(p) := p\hat{u}_r^R + (1-p)\hat{u}_r^R$, and $\widehat{U}(p) := \max\{\widehat{U}_\ell(p), \widehat{U}_r(p)\}$.

weakly above $\widehat{U}$. This shows that the optimal a can be found in $\{\lambda dt, 1\}$. $\square$

C.2 Non-Exclusivity of Attention

The proofs of our main results only require minor modifications. One important change is that the full attention strategy has to be defined using $\alpha = \beta = \overline{\alpha}$. Without this modification, Lemmas 1 and 4–6 are no longer valid. We also have to replace V_0 and V_1 by solutions to $c + \rho V(p) = F_\alpha(p, V(p), V'(p))$ for $\alpha = \underline{\alpha}$ and $\alpha = \overline{\alpha}$, respectively. The value of the stationary strategy $U^S(p)$ is unchanged as discussed in the main text. The crucial Lemmas 2 and 3 continue to hold without modification.

Explicit expressions for the boundaries of the experimentation region and the absorbing point p^* are now given by

$$\begin{aligned}
\underline{p}^* &= \frac{u_\ell^L \rho + c}{\rho\left(u_\ell^L - u_\ell^R\right) + \left(u_r^R - u_\ell^R\right)\lambda\overline{\alpha}}, \\
\overline{p}^* &= \frac{\left(u_\ell^L - u_r^L\right)\lambda\overline{\alpha} - u_r^L \rho - c}{\rho\left(u_r^R - u_r^L\right) + \left(u_\ell^L - u_r^L\right)\lambda\overline{\alpha}}, \\
p^* &= \frac{\left(u_\ell^L \rho + c\right)}{\left(u_r^R \rho + c\right) + \left(u_\ell^L \rho + c\right)}.
\end{aligned}$$

One can see from the first two expressions that $\underline{p}^*$ increases and $\overline{p}^*$ decreases if we decrease the upper bound $\overline{\alpha}$. This confirms the claim that the experimentation region shrinks if the constraint on α is tightened.

The cutoffs $\overline{c}$, $\underline{c}$ are given by:

$$\overline{c} := 0 \vee \frac{\lambda\overline{\alpha}\left(u_r^R - u_\ell^R\right)\left(u_\ell^L - u_r^L\right) - \rho\left(u_r^R u_\ell^L - u_\ell^R u_r^L\right)}{\left(u_r^R - u_\ell^R\right) + \left(u_\ell^L - u_r^L\right)}, \tag{C.2}$$

$$\underline{c} := 0 \vee \begin{cases} \overline{c} \wedge \frac{\lambda}{1+e^2} \min\left\{\left(u_r^R - u_\ell^R\right), \left(u_\ell^L - u_r^L\right)\right\} & \text{if } \rho = 1 - \overline{\alpha} = 0, \\ \overline{c} \wedge \min\left\{ \frac{(\rho+\lambda\overline{\alpha})\left(u_r^R - u_\ell^R\right)}{1+\left(\frac{2\rho+\lambda}{(2\overline{\alpha}-1)\lambda}\right)^{\frac{2\overline{\alpha}-1}{(1-\overline{\alpha})+\frac{\rho}{\lambda}}}} - \rho u_r^R, \frac{(\rho+\lambda\overline{\alpha})\left(u_\ell^L - u_r^L\right)}{1+\left(\frac{2\rho+\lambda}{(2\overline{\alpha}-1)\lambda}\right)^{\frac{2\overline{\alpha}-1}{(1-\overline{\alpha})+\frac{\rho}{\lambda}}}} - \rho u_\ell^L \right\} & \text{otherwise.} \end{cases} \tag{C.3}$$

From the first expression it is immediately clear that $\overline{c}$ decreases if we reduce the upper bound $\overline{\alpha}$. It is less obvious that $\underline{c}$ increases at the same time. To see this, remember from the proof of Theorem 1 that $c > \underline{c}$ is equivalent to

$$\max\left\{\underline{V}_{own}(p^*), \overline{V}_{own}(p^*)\right\} > U^S(p^*).$$

The right-hand side of this inequality is independent of $\overline{\alpha}$. The left-hand however, is decreasing in $\overline{\alpha}$.

C.3 Asymmetric Returns to Attention

In this section we revisit three crucial results that are used to prove Theorem 1, and outline how they are changed if $\overline{\lambda}^R \neq \overline{\lambda}^L$. Throughout we assume that $\overline{\lambda}^R \geq \overline{\lambda}^L$. Up to relabeling this is without loss of generality. The three crucial results are:

(a) The Crossing Lemma 2 and the Unimprovability Lemma 3. In Appendix A, we considered solutions V_0 and V_1 to the HJB equation where we set $\alpha = 0$ or $\alpha = 1$, respectively. If we generalize the HJB equation to allow for $\overline{\lambda}^R > \overline{\lambda}^L$, we can obtain similar solutions V_0 and V_1. Lemma 2 also uses the value of the stationary strategy as a benchmark. The definition of the stationary strategy has to be modified if $\overline{\lambda}^R > \overline{\lambda}^L$. The Bayesian updating formula in the absence of a signal is now given by:

$$\dot{p}_t = -\left(\overline{\lambda}^R \alpha_t - \overline{\lambda}^L \beta_t\right) p_t \left(1 - p_t\right), \tag{C.4}$$

Hence the stationary attention strategy is given by

$$\alpha^S = \frac{\overline{\lambda}^L}{\overline{\lambda}^R + \overline{\lambda}^L}.$$

Note that this coincides with the definition of our main model where $\alpha^S = 1/2$ if $\overline{\lambda}^R = \overline{\lambda}^L$. The value of the stationary strategy is now

$$U^S(p) := p \frac{\alpha^S \overline{\lambda}^R u_r^R - c}{\rho + \alpha^S \overline{\lambda}^R} + (1-p) \frac{\beta^S \overline{\lambda}^L u_\ell^L - c}{\rho + \beta^S \overline{\lambda}^L}.$$

With this definition, Lemmas 2 and 3 continue to hold.

(b) Properties of the own-biased strategy in Lemma 5:[55] In Appendix A we have constructed the own-biased strategy by first obtaining the boundary points $\underline{p}^*$ and $\overline{p}^*$ from value matching and smooth pasting. Following the same steps while allowing for $\overline{\lambda}^R > \overline{\lambda}^L$ we get

$$\underline{p}^* = \frac{u_\ell^L \rho + c}{\rho \left(u_\ell^L - u_\ell^R\right) + \left(u_r^R - u_\ell^R\right) \overline{\lambda}^R}, \tag{C.5}$$

$$\overline{p}^* = \frac{\left(u_\ell^L - u_r^L\right) \overline{\lambda}^L - u_r^L \rho - c}{\rho \left(u_r^R - u_r^L\right) + \left(u_\ell^L - u_r^L\right) \overline{\lambda}^L}. \tag{C.6}$$

The branches of the own-biased solution are then given by particular solutions V_0 and V_1 that satisfy the boundary conditions $V_0(\overline{p}^*) = U_r(\overline{p}^*)$ and $V_1(\underline{p}^*) = U_\ell(\underline{p}^*)$. Lemma 5.(a)-(b) hold unchanged if $\overline{\lambda}^R > \overline{\lambda}^L$. For the other results in Lemma 5, we need to modify the definition of the full-attention strategy. We compute separately

[55] Lemma 5 repeats the statements of Lemma 1 so we do not discuss Lemma 1 separately.

the value of full attention if the DM can obtain both types of evidence at rate $\overline{\lambda}^R$:

$$U_R^{FA}(p) := p\frac{\overline{\lambda}^R u_r^R - c}{\rho + \overline{\lambda}^R} + (1-p)\frac{\overline{\lambda}^R u_\ell^L - c}{\rho + \overline{\lambda}^R} = \frac{\overline{\lambda}^R \left(pu_r^R + (1-p)u_\ell^L\right) - c}{\rho + \overline{\lambda}^R},$$

and at rate $\overline{\lambda}^L$:

$$U_L^{FA}(p) := p\frac{\overline{\lambda}^L u_r^R - c}{\rho + \overline{\lambda}^L} + (1-p)\frac{\overline{\lambda}^L u_\ell^L - c}{\rho + \overline{\lambda}^L} = \frac{\overline{\lambda}^L \left(pu_r^R + (1-p)u_\ell^L\right) - c}{\rho + \overline{\lambda}^L}.$$

Generalizing Lemma 5.(c) we obtain now obtain:

$$U_\ell(\underline{p}^*) = U_R^{FA}(\underline{p}^*), \qquad \text{and} \qquad U_r(\overline{p}^*) = U_r^{FA}(\overline{p}^*).$$

Lemma 5.(d) refers to the condition (EXP). If $\overline{\lambda}^R > \overline{\lambda}^L$, we need to define two separate conditions

$$U_R^{FA}(\hat{p}) > U(\hat{p}), \tag{EXP$_R$}$$
$$U_L^{FA}(\hat{p}) > U(\hat{p}). \tag{EXP$_L$}$$

With this Lemma 5.(d) generalizes as follows: If (EXP$_R$) holds, $0 < \underline{p}^* < \hat{p}$ and $\underline{V}_{own}(p) < U_R^{FA}(p)$ for all $p \in (\underline{p}^*, 1)$, $\underline{V}_{own}(p) > U_R^{FA}(p)$ for $p < \underline{p}^*$ and $\underline{V}_{own}(p) = U_R^{FA}(p)$ if $p \in \left\{\underline{p}^*, 1\right\}$. If (EXP$_L$) holds, $0 < \underline{p}^* < \overline{p}^* < \hat{p}$ and $\overline{V}_{own}(p) < U_L^{FA}(p)$ for all $p \in (0, \overline{p}^*)$, $\overline{V}_{own}(p) > U_L^{FA}(p)$ for $p > \overline{p}^*$ and $\overline{V}_{own}(p) = U_L^{FA}(p)$ if $p \in \{0, \overline{p}^*\}$. Lemma (5).(e) generalizes as follows: If (EXP$_L$) is violated, then $\overline{V}_{own} = U(p)$ for all $p \in [\hat{p}, 1]$. If (EXP$_R$) is violated, then $V_{own} = U(p)$ for all $p \in [0, \hat{p}]$.

(c) Properties of the opposite-biased solution in Lemma 6: As in the main model, we can use smooth pasting and value matching with U^S to obtain p^* as follows:

$$p^* = \frac{\left(u_\ell^L \rho + c\right)\overline{\lambda}^L}{\left(u_r^R \rho + c\right)\overline{\lambda}^R + \left(u_\ell^L \rho + c\right)\overline{\lambda}^L}. \tag{C.7}$$

As before we obtain the branches of the opposite-biased strategy as particular solutions V_0 and V_1 with the boundary condition $V_0(p^*), V_1(p^*) = U^S(p^*)$, and set

$$V_{own}(p) := \begin{cases} \underline{V}_{own}(p), & \text{if } p < p^*, \\ \overline{V}_{own}(p), & \text{if } p \geq p^*. \end{cases}$$

With this definition Lemma 6.(a) holds unchanged. Part (b) of the Lemma has to be modified: $V_{opp}(p) = \underline{V}_{opp}(p) \leq U_L^{FA}(p)$ for all $p \in [0, p^*]$ with strict inequality for $p \neq p^*$, and $V_{opp}(p) = \overline{V}_{own}(p) \leq U_R^{FA}(p)$ for all $p \in [p^*, 1]$ with strict inequality for $p \neq p^*$.

The fact that the important Lemmas 2 and 3 continue to hold and we still have $V_{opp}(p) > U^S(p)$ for all $p \neq p^*$ from Lemma 6.(a), together imply that many of the structural properties of $V_{Env}(p) = \max\{V_{own}(p), V_{opp}(p)\}$ are preserved and the structure of the optimal policy is similar to the main model with $\overline{\lambda}^R = \overline{\lambda}^L$.

However, there is one crucial difference. It is now possible that $\overline{V}_{own}(p)$ is dominated by $\underline{V}_{own}(p)$ or by $V_{opp}(p)$ for all $p < \overline{p}^*$. We only consider the case that $V_{own}(p^*) > U^S(p^*)$. In this case we can use similar steps as in the proof of Proposition 8.(a) to show that $V_{Env}(p) = V_{own}(p)$ for all $p \in [0,1]$, i.e., opposite-biased learning is never optimal. However, it is no longer guaranteed that there exists a point of intersection $\check{p} \in (\underline{p}^*, \overline{p}^*)$ between $\underline{V}_{own}(p)$ and $\overline{V}_{own}(p)$. This is most easily seen by considering the generalization of Lemma 5.(c) outlined above. It is easy to see that $U^{FA}_R(p) > U^{FA}_L(p)$ for all $p \in [0,1]$ since $\overline{\lambda}^R > \overline{\lambda}^L$. Since both functions are strictly decreasing in c, we can find levels of c for which $U^{FA}_L(p) < U(p)$ for all $p \in [0,1]$ but $U^{FA}_R(p) > U(p)$ for some p. In this case

$$V_{own}(p) = \max\left\{\underline{V}_{own}(p), \overline{V}_{own}(p)\right\} = \max\left\{\underline{V}_{own}(p), U_r(p)\right\}, \quad \forall p \in [0,1]. \tag{C.8}$$

More generally, (C.8) may also hold if $U^{FA}_L(p) > U(p)$ for some $p \in [0,1]$. Before we could rule out this case since for $\overline{\lambda}^R = \overline{\lambda}^L$, $\underline{V}_{own}(p) < U^{FA}(p)$ for all $p \in (\underline{p}^*, 1)$ and $\overline{V}_{own}(\overline{p}^*) = U^{FA}(\overline{p}^*)$. This is not longer true if $\overline{\lambda}^R > \overline{\lambda}^L$. The example in Panel (a) of Figure 6.2 depicts such a case. Moreover we can argue that, as claimed in Section 6.3, this case only arises if $\overline{\lambda}^R - \overline{\lambda}^L$ is sufficiently large. To see this fix $\overline{\lambda}^R$ such that $\underline{V}_{own}(p) > U(p)$ for some p. Note that $\underline{V}_{own}(p)$ is independent of $\overline{\lambda}^L$, since $\underline{p}^*$ does not depend on $\overline{\lambda}^L$ and $\underline{V}_{own}(p)$ is the value of seeking R-evidence. Moreover, we have $\underline{V}_{own}(\underline{p}^*) > U_r(\underline{p}^*)$ and one can easily verify that $\underline{V}_{own}(1) = U^{FA}_R(1) < U_r(p)$. Since $U_r(p)$ is linear in p and $\underline{V}_{own}(p)$ can be verified to be strictly convex, there exists a unique $\overline{q}$ such that $\overline{V}_{own}(\underline{q}) = U_r(\overline{q})$. If $\overline{q} \geq \overline{p}^*$, the Crossing Lemma 2 implies that there exists no intersection of $\underline{V}_{own}(p)$ and $\overline{V}_{own}(p)$ between $\underline{p}^*$ and $\overline{p}^*$. In this case (C.8) holds. Conversely, if $\overline{q} < \overline{p}^*$ an intersection point $\check{p} \in (\underline{p}^*, \overline{p}^*)$ exists and $V_{own}(p)$ has the same structure as in our main model. It remains to argue that $\overline{q} \geq \overline{p}^*$ only if $\overline{\lambda}^R - \overline{\lambda}^L$ is sufficiently large. For $\overline{\lambda}^R - \overline{\lambda}^L = 0$, $\underline{V}_{own}(\overline{p}^*) < U^{FA}_R(\overline{p}^*) = U^{FA}_L(\overline{p}^*) = \overline{V}_{own}(\overline{p}^*) = U_r(\overline{p}^*)$. Therefore, $\overline{q} < \overline{p}^*$. Decreasing λ^L while holding λ^R fixed does not change $\overline{q}$ but decreases $\overline{p}^*$. Hence, there exists a cutoff for λ^L below which (for given λ^R), $\overline{q} \geq \overline{p}^*$.

C.4 Diminishing Returns to Attention

As an extension to the main model in Section 2, we show that the general structure of the solution is preserved if the arrival rate of breakthroughs from a given news source does not increase linearly in the amount of attention allocated to the source. For the proofs we adopt a different notation than in Section 6.4. Note that each choice of attention gives rise to a pair of arrival rates (λ^R, λ^L). For given $g(x)$ a pair (λ^R, λ^L) is feasible if there exists $\alpha \in [0,1]$ such that $\lambda^R \leq \lambda g(\alpha)$ and $\lambda^L \leq \lambda g(1-\alpha)$. Instead of working with the function $g(x)$ we introduce a function $\Gamma(\lambda^R)$ that characterizes the upper bound of the set of feasible pairs (λ^R, λ^L) as follows:[56]

$$\left\{(\lambda^R, \lambda^L) \in \mathbb{R}_+ \,\middle|\, \lambda^L \leq \Gamma(\lambda^R)\right\}.$$

Remember from Section 6.4 that $\lambda^R = \lambda g(\alpha)$ and $\lambda^L = \lambda g(1-\alpha)$. If we normalize $\lambda = 1$, we can derive $\Gamma(\lambda)$ from the function $g(x)$:[57]

$$\Gamma(\lambda^R) = g(1 - g^{-1}(\lambda^R)).$$

Clearly, the DM will only chose pairs of arrival rates on the upper bound, i.e., $(\lambda^R, \Gamma(\lambda^R))$, so we can describe her choice by λ^R. To simplify the notation we omit the superscript and write λ instead of λ^R. Moreover we assume that $\lambda \in [0,1]$. We maintain the following assumptions about the function Γ.

Assumption 1. $\Gamma : [0,1] \to [0,1]$ *is twice continuously differentiable, strictly decreasing, strictly convex, and satisfies* $\Gamma(0) = 1$, $\Gamma(1) = 0$ *and* $\Gamma'(\gamma) = -1$, *where* γ *is the unique fixed point of* Γ.

Note that $\Gamma'(\gamma) = -1$ is always fulfilled if Γ is derived from a differentiable function g since $\Gamma(\Gamma(x)) = x$ in this case which implies that the graph of Γ is symmetric with respect to the 45-degree line.

Example 1. A parametric example is obtained by setting $g(x) = \sqrt{1 + 4x + x2} - 2$. The inverse is $g^{-1}(x) = 2\sqrt{4 - 2x - x^2}$ and we obtain

$$\Gamma(\lambda^R) = g(1 - g^{-1}(\lambda^R)) = \sqrt{6\sqrt{4 - 2\lambda^R - (\lambda^R)^2} + \lambda^R(2 + \lambda^R) - 8} - 1.$$

This is the example used in Figure 6.3 in Section 6.4.

[56] The feasible set of arrival rates can also be derived from a model with many news sources but constant returns to attention. In this model, a news source is now characterized by two parameters (λ^R, λ^L). If an amount of attention α_i is directed to a news-source given by $(\lambda_i^R, \lambda_i^L)$, the DM will receive a signal from that source that confirms state R with Poisson arrival rate $\lambda_i^R \alpha_i$ if the state is indeed R and she will receive a signal that confirms state L with Poisson arrival rate $\lambda_i^L \alpha_i$ if the state is L. Hence, when allocating her attention over two news sources with parameters $(\lambda_i^R, \lambda_i^L)$ and $(\lambda_j^R, \lambda_j^L)$ with attention levels α_i and $\alpha_j = 1 - \alpha_i$, the DM will receive a signal that confirms R with Poisson rate $\lambda^R = \alpha_i \lambda_i^R + (1 - \alpha_j)\lambda_j^R$, and a signal that confirms L with Poisson rate $\lambda^L = \alpha_i \lambda_i^L + (1 - \alpha_i)\lambda_j^L$. The set of feasible arrival rates (λ^R, λ^L) is thus a weakly convex subset of $\mathbb{R}_+$. We denote the upper bound of this set $\Gamma(\lambda^R)$ and note that weak convexity of the set implies weak concavity $\Gamma(\lambda^R)$. In the main model studied before we had $\Gamma(\lambda^R) = 1 - \lambda^R$, which is the linear boundary that is spanned by the two primitive news sources given by $(1,0)$ and $(0,1)$.

[57] The normalization of the upper bound is without loss of generality since only the ratios ρ/λ and c/λ matter.

C.4.1 The Decision Maker's Problem

The DM's posterior evolves according to

$$\dot{p}_t = -p_t(1-p_t)\left(\lambda_t - \Gamma(\lambda_t)\right), \tag{C.9}$$

The objective is given by

$$J\left((\lambda_t)_{t\geq 0}, T; p_0\right) := \begin{Bmatrix} \int_0^T e^{-\rho t} P_t(p_0, (\lambda_\tau))\left(p_t \lambda_t u_r^R + (1-p_t)\Gamma(\lambda_t) u_\ell^L\right) dt \\ +e^{-\rho T} P_T(p_0, (\lambda_\tau)) U(p_T) \end{Bmatrix},$$
$$\text{where} \quad P_t(p_0, (\lambda_\tau)) := p_0 e^{-\int_0^t \lambda_s ds} + (1-p_0) e^{-\int_0^t \Gamma(\lambda_s) ds}.$$

The DM solves the problem ($\mathcal{P}^\Gamma$) given by:

$$V(p_0) := \sup_{\left((\lambda_t)_{t\geq 0}, T\right)} J\left((\lambda_t)_{t\geq 0}, T; p_0\right) \quad \text{s.t. (C.9), and } \lambda_t \in [0,1]. \tag{$\mathcal{P}^\Gamma$}$$

We define

$$H(p, V(p), V'(p), \lambda) := \begin{Bmatrix} \lambda p \left(u_r^R - V(p)\right) + \Gamma(\lambda)(1-p)\left(u_\ell^L - V(p)\right) \\ -p(1-p)(\lambda - \Gamma(\lambda)) V'(p) \end{Bmatrix}.$$

The HJB equation for ($\mathcal{P}^\Gamma$) is

$$\max\left\{-c - \rho V(p) + \max_{\lambda \in [0,1]} H(p, V(p), V'(p), \lambda), U(p) - V(p)\right\} = 0. \tag{C.10}$$

If $V(p) > U(p)$ this simplifies to

$$c + \rho V(p) = \max_{\lambda \in [0,1]} H(p, V(p), V'(p), \lambda). \tag{C.11}$$

The first-order condition is given by

$$\frac{\partial H(p, V(p), V'(p), \lambda)}{\partial \lambda} = \begin{Bmatrix} p\left(u_r^R - V(p)\right) + \Gamma'(\lambda)(1-p)\left(u_\ell^L - V(p)\right) \\ -p(1-p)(1 - \Gamma'(\lambda)) V'(p) \end{Bmatrix} = 0. \tag{C.12}$$

For a given policy $\lambda(p)$, we obtain the differential equation

$$c + \rho V(p) = H(p, V(p), V'(p), \lambda(p)) \tag{C.13}$$
$$\iff c + \rho V(p) = \begin{Bmatrix} \lambda(p) p\left(u_r^R - V(p)\right) + \Gamma(\lambda(p))(1-p)\left(u_\ell^L - V(p)\right) \\ -p(1-p)(\lambda(p) - \Gamma(\lambda(p))) V'(p) \end{Bmatrix}.$$

As in our original model, we will define two candidate value functions. For this purpose, we state the HJB equation for problems in which the DM is either restricted to choose

$\lambda \geq \gamma$,

$$c + \rho V_+(p) \quad = \max\nolimits_{\lambda \in [\gamma,1]} H(p, V_+(p), V_+'(p), \lambda), \tag{C.14}$$

or $\lambda \leq \gamma$:

$$c + \rho V_-(p) \quad = \max\nolimits_{\lambda \in [0,\gamma]} H(p, V_-(p), V_-'(p), \lambda). \tag{C.15}$$

we denote policies corresponding to solution to (C.14) and (C.15) by $\lambda_+(p)$ and $\lambda_-(p)$, respectively.

C.4.2 Preliminary results

We first revisit some definitions made for the original model. The stationary strategy is now given by choosing $\lambda = \gamma$ until a signal arrives and then taking an optimal action according to the signal. The value of this strategy is now given by

$$U^S(p) = \frac{\gamma}{\rho + \gamma} U^*(p) - \frac{c}{\rho + \gamma},$$

where

$$U^*(p) = p u_r^R + (1-p) u_\ell^L$$

is the first best value that is achieved if the DM can learn the state without any delay.

As in the original model, we obtain a crossing condition for functions that satisfy (C.14) and (C.15) and a condition under which solutions to (C.14) and (C.15) satisfy (C.11).

Lemma 7 (Crossing Lemma)**.** *Suppose $V_+(p)$ is $\mathcal{C}^1$ at p and satisfies* (C.14) *and $V_-(p)$ is $\mathcal{C}^1$ at p and satisfies* (C.15)*. If $V_+(p) = V_-(p) \geq U^S(p)$, then $V_+'(p) \leq V_-'(p)$. If $V_+(p) = V_-(p) > U^S(p)$, then $V_+'(p) < V_-'(p)$.*

Proof of Lemma 7. Suppose $V(p) := V_+(p) = V_-(p) \geq U^S(p)$ at p and denote the maximizers in (C.14) and (C.15) by $\lambda_+(p)$ and $\lambda_-(p)$ respectively.

From (C.14) and (C.15), we obtain

$$\begin{aligned} & p(1-p)(\Gamma(\lambda_-(p)) - \lambda_-(p))(\lambda_+(p) - \Gamma(\lambda_+(p)))(V_-'(p) - V_+'(p)) \\ &= (\delta(p)\rho + \Delta(p)) \left[V(p) - \frac{\frac{\Delta(p)}{\delta(p)}}{\frac{\Delta(p)}{\delta(p)} + \rho} U^*(p) + \frac{1}{\frac{\Delta(p)}{\delta(p)} + \rho} c \right] \\ &\geq (\delta(p)\rho + \Delta(p)) \left[V(p) - \frac{\gamma}{\gamma + \rho} U^*(p) + \frac{1}{\gamma + \rho} c \right] \\ &= (\delta(p)\rho + \Delta(p)) \left[V(p) - U^S(p) \right], \end{aligned}$$

where

$$\delta(p) :=\Gamma(\lambda_-(p)) - \lambda_-(p) + \lambda_+(p) - \Gamma(\lambda_+(p)) > 0,$$
$$\Delta(p) :=\lambda_+(p)\Gamma(\lambda_-(p)) - \lambda_-(p)\Gamma(\lambda_+(p)) > 0,$$

since $\lambda_+(p) > \gamma > \lambda_-(p)$. The inequality can be seen as follows. First, one can verify that $(\Delta(p)/\delta(p), \Delta(p)/\delta(p))$ is the point of intersection between the forty-five degrees line and the line segment between two points, $(\lambda_-(p), \Gamma(\lambda_-(p)))$ and $(\lambda_+(p), \Gamma(\lambda_+(p)))$. Since Γ is concave, we must have $\Delta(p)/\delta(p) < \gamma$. Since $\delta(p), \Delta(p) \geq 0$, if $V(p) \geq U^S(p)$, the last expression is non-negative, and if $V(p) > U^S(p)$, it is strictly positive. □

Lemma 8 (Unimprovability). *(a) Suppose $V_+(p)$ is $\mathcal{C}^1$ at p and satisfies (C.14). If $V_+(p) \geq \max\{U^S(p), U(p)\}$, then $V_+(p)$ satisfies (C.11) at p.*

(b) Suppose $V_-(p)$ is $\mathcal{C}^1$ at p and satisfies (C.15). If $V_-(p) \geq \max\{U^S(p), U(p)\}$, then $V_-(p)$ satisfies (C.11) at p.

Proof of Lemma 8. We prove the first statement; the second follows symmetrically. Suppose the optimal policy satisfies $\lambda_+(p) > \gamma$. By the condition, it is not improvable by an immediate action or by any $\lambda \geq \gamma$. Hence, it suffices to show that it is not improvable by any $\lambda_- < \gamma$.

Substituting $V_+'(p)$ from (C.14) and rearranging we get

$$\begin{aligned}
&H(p, V_+(p), V_+'(p), \lambda_+(p)) - H(p, V_+(p), V_+'(p), \lambda_-)\\
=&\frac{\hat{\delta}(p)\rho + \hat{\Delta}(p)}{\lambda_+(p)) - \Gamma(\lambda_+(p)))}\left[V_+(p) - \frac{\frac{\hat{\Delta}(p)}{\hat{\delta}(p)}}{\frac{\hat{\Delta}(p)}{\hat{\delta}(p)} + \rho}U^S + \frac{1}{\frac{\hat{\Delta}(p)}{\hat{\delta}(p)} + \rho}c\right]\\
\geq&\frac{\hat{\delta}(p)\rho + \hat{\Delta}(p)}{\lambda_+(p)) - \Gamma(\lambda_+(p)))}\left[V_+(p) - U^S(p)\right],
\end{aligned}$$

where

$$\hat{\delta}(p) := \Gamma(\lambda_-) - \lambda_- + \lambda_+(p) - \Gamma(\lambda_+(p)) \text{ and } \hat{\Delta}(p) := \lambda_+(p)\Gamma(\lambda_-) - \lambda_-\Gamma(\lambda_+(p)).$$

The inequality follows from the same observation as in the proof of Lemma (7). □

Before constructing the value function for $(\mathcal{P}^\Gamma)$, we make one general observation about the boundaries of the experimentation region and the value opposite-biased signals at the boundaries.

For this purpose we consider a model in which the DM has full attention. In this case we have $\lambda^R = 1 = \lambda^L$ and the DM only chooses when to stop. Note that Assumption 1 precludes the DM from choosing $\lambda^R = 1 = \lambda^L$ so the full attention model only serves as a hypothetical benchmark.

The value of this stopping problem is given by

$$\widehat{V}(p) := \max\left\{U(p), U^{FA}(p)\right\},$$

where

$$U^{FA}(p) = \frac{1}{\rho+1}U^*(p) - \frac{c}{\rho+1}.$$

Moreover, we note that by Assumption 1, $(\lambda, \Gamma(\lambda)) \leq (1,1)$ for all $\lambda \in (0,1)$. Therefore, $\widehat{V}(p)$ is an upper bound for the value function of the problem ($\mathcal{P}^\Gamma$).

Remember that in our original model, the boundaries of the experimentation region are given by the points of intersection between $U^{FA}(p)$ and $U(p)$:

$$U^{FA}(\overline{p}^*) = U_r(\overline{p}^*). \tag{C.16}$$

$$U^{FA}(\underline{p}^*) = U_\ell(\underline{p}^*). \tag{C.17}$$

If (EXP) is satisfied, we have $\underline{p}^* < \overline{p}^*$. We now show that the value of ($\mathcal{P}^\Gamma$) is equal to $\widehat{V}$ at these boundaries. This immediately shows that $\underline{p}^*$ and $\overline{p}^*$ are the boundaries of the experimentation region in ($\mathcal{P}^\Gamma$). Moreover, we show that under Assumption 1, at these boundaries, the DM does not benefit from interior choices $\lambda \in (0,1)$.

Proposition 10. *Suppose (EXP) is satisfied. Then $\underline{p}^*$ and $\overline{p}^*$ given by (C.16) and (C.17) are the boundaries of the the experimentation region for the optimal solution to ($\mathcal{P}^\Gamma$). At $\underline{p}^*$ and $\overline{p}^*$, the value of ($\mathcal{P}^\Gamma$) coincides with the value of our original model and equals $U^{FA}(p)$ The loss of restricting the DM to chose $\lambda \in \{0,1\}$ vanishes as $p \downarrow \underline{p}^*$ and $p \uparrow \overline{p}^*$.*

Proof. If the DM is restricted to chose $\lambda \in \{0,1\}$, her optimal strategy coincides with the optimal strategy in our original model. The value in our original model is a lower bound for the value of ($\mathcal{P}^\Gamma$). Since at $\underline{p}^*$ and $\overline{p}^*$ the value in our original model coincides with the upper bound $U^{FA}(p)$, it must also coincide with the value of ($\mathcal{P}^\Gamma$). □

Note that while Assumption 1 requires $\Gamma(\lambda) < 1$ for $\lambda > 0$, it does not rule out an Inada condition like $\lim_{\lambda\to 0}\Gamma'(\lambda) = 0$. This shows that at the boundaries of the experimentation region, the value of a opposite-biased signal is zero even if it is cost-less to obtain. We will see below when we characterize the value function that without an Inada condition, there exist neighborhoods of $\underline{p}^*$ and $\overline{p}^*$ such that the DM does not suffer any loss if in these neighborhoods she uses $\lambda = 1$ and $\lambda = 0$, respectively.

C.4.3 Construction of Solutions to the HJB equation

For the remainder of this section, we will focus on the cases that the payoffs are symmetric. This simplifies the derivations and is sufficient to understand the main features of the optimal solution in the extension. Formally we impose:

Assumption 2. *$u_r^R = u_\ell^L = U^S$ and $u_\ell^R = u_r^L = \underline{u}$ for some $\overline{u} > \underline{u} > 0$.*

In contrast to our original model, it may now be optimal to choose $\lambda \in (0,1)$ for beliefs $p \in (\underline{p}^*, \overline{p}^*)$, i.e., in the interior of the experimentation region. For an interval where this is the case, we will obtain a differential equation for $\lambda(p)$ and furthermore an equation that expresses $V(p)$ as a function of $\lambda(p)$. We begin with the latter. To state the result in concise form we define

$$A(\lambda) := \frac{\Gamma(\lambda) - \Gamma'(\lambda)\,\lambda}{\Gamma(\lambda) - \Gamma'(\lambda)\,\lambda + \rho\,(1 - \Gamma'(\lambda))}, \quad \text{and} \quad B(\lambda) := \frac{1 - \Gamma'(\lambda)}{\Gamma(\lambda) - \Gamma'(\lambda)\,\lambda + \rho\,(1 - \Gamma'(\lambda))}.$$

A basic observation that we will use at several points is that these two functions are (inverse) U-shaped with (maximum) minimum at $\lambda = \gamma$.

Lemma 9. *If Assumption 1 is satisfied,*

$$A'(\lambda) > (<) 0 \iff B'(\lambda) < (>) 0 \iff \lambda > (<) \gamma.$$

Proof. The Lemma follows from straightforward algebra which we omit here. □

Lemma 10. *Suppose Assumptions 1 and 2 are satisfied. If $p \in (0,1)$, $V(p)$ is continuously differentiable at p and satisfies (C.11) with maximizer $\lambda(p) \neq \gamma$, then*

$$V(p) \geq A(\lambda(p))\overline{u} - B(\lambda(p))\,c \geq U^S(p) \tag{C.18}$$

If λ satisfies (C.12) at p, then the first inequality binds. The statement continues to hold if we replace V, λ, and (C.11), by V_+, λ_+ and (C.14), or V_-, λ_- and (C.15).

Proof of Lemma 10. We define the LHS of (C.12) as

$$X := (p + (1-p)\Gamma'(\lambda))\,(\overline{u} - V(p)) - p(1-p)(1 - \Gamma'(\lambda))V'(p). \tag{C.19}$$

Eliminating $V'(p)$ from (C.13) and (C.19) we obtain an expression for $V(p)$ in terms of $\lambda(p)$ and X:

$$V(p) = A(\lambda(p))\overline{u} - B(\lambda(p))\,c + \frac{X\,(\lambda - \Gamma(\lambda(p)))}{\Gamma(\lambda) - \Gamma'(\lambda)\,\lambda + \rho\,(1 - \Gamma'(\lambda))}.$$

If $\lambda(p)$ is a maximizer in (C.11), we must have

$$X \begin{cases} \geq 0 & \text{if } \lambda = 1, \\ = 0 & \text{if } \lambda \in (0,1), \\ \leq 0 & \text{if } \lambda = 0. \end{cases}$$

Since $\lambda = 1$ implies $\lambda - \Gamma(\lambda(p)) > 0$ and $\lambda = 0$ implies $\lambda - \Gamma(\lambda(p)) < 0$ we have

$$V(p) \geq A(\lambda(p))\overline{u} - B(\lambda(p))\, c,$$

and the inequality holds with equality if $X = 0$ which is equivalent to λ satisfying (C.12). This proves the first inequality and the first statement.

The second inequality follows from Lemma 9 and $A(\gamma)\overline{u} - B(\gamma)\, c = U^S(p)$, which is obtained from straightforward algebra. It is straightforward to adapt the proofs to V_+ and V_-. □

Using Lemma 10 we can obtain an ODE for λ that holds whenever the optimal policy is interior, i.e., it satisfies (C.12).

Lemma 11. *Suppose Assumptions 1 and 2 are satisfied. If $p \in (0, 1)$, V is continuously differentiable at p and satisfies (C.13) and the maximizer is $\lambda(p) \neq \gamma$ and satisfies (C.12) at p, then*

$$\lambda'(p) = \frac{[p + (1-p)\Gamma'(\lambda(p))]\,[\Gamma(\lambda(p)) - \Gamma'(\lambda(p))\,\lambda(p) + \rho\,(1 - \Gamma'(\lambda(p)))]}{p(1-p)\,(\Gamma(\lambda(p)) - \lambda(p))\,\Gamma''(\lambda(p))}. \tag{C.20}$$

The statement continues to hold if we replace V and λ, by V_+ and λ_+, or V_- and λ_-.

Proof of Lemma 11. If $\lambda(p) \neq \gamma$ satisfies (C.12), then by Lemma 10

$$\begin{aligned} V(p) &= A(\lambda(p))\overline{u} - B(\lambda(p))\, c, \\ \text{and} \quad V'(p) &= A'(\lambda(p))\lambda'(p)\overline{u} - B'(\lambda(p))\lambda'(p)\, c. \end{aligned}$$

Inserting these two equations in (C.13) and solving for $\lambda'(p)$ we get (C.20) □

Next, we state a Lemma that identifies conditions under which the solution to (C.20) remains bounded away from $\lambda = 0$ or $\lambda = 1$.

Lemma 12. *Suppose Assumptions 1 and 2 are satisfied. Then there exists function $p^1(x) > 1/2$ for $x > \gamma$ and $p^0(x) < 1/2$ for $x < \gamma$ such that*

$$\begin{aligned} \lambda(p) = \lambda_+ > \gamma \quad &\Rightarrow \quad \left\{\lambda'(p) < 0 \iff p < p^1(\lambda_+)\right\}, \\ \lambda(p) = \lambda_- < \gamma \quad &\Rightarrow \quad \left\{\lambda'(p) > 0 \iff p > p^0(\lambda_-)\right\}. \end{aligned}$$

Proof. Inserting $\lambda(p) = \lambda_+ > \gamma$ in (C.20) yields

$$\begin{aligned} &\lambda'(p) < 0 \\ \iff [p + (1-p)\Gamma'(\lambda_+)]\, &\frac{\Gamma(\lambda_+) - \Gamma'(\lambda_+)\,\lambda_+ + \rho\,(1 - \Gamma'(\lambda_+))}{p(1-p)\,(\Gamma(\lambda(p)) - \lambda(p))\,\Gamma''(\lambda(p))} < 0 \end{aligned}$$

$$\iff p + (1-p)\Gamma'(\lambda_+) < 0$$
$$\iff p < p^1(\lambda_+) = \frac{|\Gamma'(\lambda_+)|}{1 + |\Gamma'(\lambda_+)|}$$

Since $|\Gamma'(\lambda_+)| > 1$ $p^1 > 1/2$. The proof for $\lambda(p) = \lambda_- < \gamma$ is similar. □

Next, we show the following property that relates sufficiency of the FOC (C.12) to convexity of the value function.

Lemma 13. *Suppose Assumptions 1 and 2 are satisfied.*

(a) Let $W : [0,1] \to \mathbb{R}$ be weakly convex and satisfy $W(p) = U(p)$ in neighborhoods of 0 and 1. Then $H(p, W(p), W'(p), \lambda)$ is weakly concave in λ for all p and strictly concave whenever $W(p) > U(p)$.

(b) Let $\lambda(p)$ be a solution to (C.20) such that $\lambda(p) \in (0,1)$ at some p. Let

$$\pi(\ell) = \frac{(\rho + \ell)\,\Gamma'(\ell)}{(\rho + \ell)\,\Gamma'(\ell) - (\rho + \Gamma(\ell))}.$$

Then

$$\frac{\partial^2 \left[A(\lambda(p))\overline{u} - B(\lambda(p))c\right]}{\partial p^2} \geq 0 \quad \text{if} \quad \begin{cases} & \lambda(p) > \gamma \text{ and } p \leq \pi(\lambda(p)), \\ \text{or} & \lambda(p) < \gamma \text{ and } p \geq \pi(\lambda(p)). \end{cases}$$

$\pi(\ell) > 1/2$ if $\ell > \gamma$, and $\pi(\ell) < 1/2$ if $\ell < \gamma$.

Proof. (a) Some algebra yields

$$\frac{\partial^2 H(p, W(p), W'(p), \lambda)}{\partial \lambda^2} \leq 0 \quad \iff W(p) - pW'(p) \leq U^S.$$

The latter inequality is satisfied under the assumptions on W and both are strict if $W(p) > U(p)$.

(b) Differentiating $A(\lambda(p))\overline{u} - B(\lambda(p))c$ with respect to p, substituting $\lambda'(p)$ from (C.20) and differentiating again yields (after some algebra):

$$\frac{\partial^2 \left[A(\lambda(p))\overline{u} - B(\lambda(p))c\right]}{\partial p^2} < 0$$
$$\iff -\frac{\left(p^2 - (1-p)^2\Gamma'(\lambda(p))\right)\left(\rho + \Gamma(\lambda(p)) - (\rho + \lambda(p))\Gamma'(\lambda(p))\right)}{p(1-p)\left(\rho + p\lambda(p) + (1-p)\Gamma(\lambda(p))\right)\Gamma''(\lambda(p))} > \lambda'(p).$$

Substituting $\lambda'(p)$ from (C.20) in the last line and rearranging we get

$$\left(\lambda(p) - \Gamma(\lambda(p))\right)\left(p\left[\rho + \Gamma(\lambda(p))\right] + (1-p)\left[\rho + \lambda(p)\right]\Gamma'(\lambda(p))\right) < 0.$$

Solving for p this yields an upper bound if $\lambda(p) > \gamma$ so that the first term is positive and a lower bound if $\lambda(p) < \gamma$. The bound is $\pi(\lambda(p))$ in both cases. If $\ell > \gamma > \Gamma(\ell)$ we have

$$\begin{aligned}
\pi(\ell) &= \frac{(\rho+\ell)\,|\Gamma'(\ell)|}{(\rho+\ell)\,|\Gamma'(\ell)| + (\rho+\Gamma(\ell))} \\
&> \frac{(\rho+\ell)\,|\Gamma'(\ell)|}{(\rho+\ell)\,|\Gamma'(\ell)| + (\rho+\ell)} \\
&= \frac{|\Gamma'(\ell)|}{|\Gamma'(\ell)|+1} \\
&> 1/2.
\end{aligned}$$

where the last step follows because Assumption 1 implies that $|\Gamma'(\ell)| > 1$ if $\ell > \gamma$. Similarly we obtain $\pi(\ell) < 1/2$ if $\ell < \gamma$. □

C.4.4 Solution Candidates

Own-Biased Learning The first candidate is obtained by assuming that the DM uses an own-biased attention strategy. In contrast to our original model, where we choose $\lambda \in \{0,1\}$, we will now also use interior values for λ. In an own-biased strategy, the DM may now receive breakthrough news for both states but with a higher likelihood in the state that she find relatively unlikely. For instance, for low posterior beliefs p, the own-biased strategy involves $\lambda > \gamma$. At the same time, the belief moves in the same direction as the initial bias if now breakthrough arrives: $\dot{p}_t < 0$ if $\lambda > \gamma$. We have already identified the boundaries of the experimentation region.

Lemma 14. *Suppose* (EXP) *is satisfied. Then* $\underline{p}^*$ *and* $\overline{p}^*$ *satisfy*

$$\underline{p}^* = \inf\left\{ p \in [0,\hat{p}) \,\middle|\, c + \rho U_\ell(p) \le \max_{\lambda\in[\gamma,1]} \left\{ \begin{array}{c} (\lambda p + \Gamma(\lambda)(1-p))\,(\overline{u} - U_\ell(p)) \\ -p(1-p)(\lambda-\Gamma(\lambda))U_\ell'(p) \end{array} \right\} \right\}, \tag{C.21}$$

$$\overline{p}^* = \sup\left\{ p \in (\hat{p},1] \,\middle|\, c + \rho U_r(p) \le \max_{\lambda\in[0,\gamma]} \left\{ \begin{array}{c} (\lambda p + \Gamma(\lambda)(1-p))\,(\overline{u} - U_r(p)) \\ -p(1-p)(\lambda-\Gamma(\lambda))U_r'(p) \end{array} \right\} \right\}, \tag{C.22}$$

and the maximizers on the right-hand side are given by $\lambda = 1$ *and* $\lambda = 0$*, respectively. Moreover,*

$$\begin{aligned}
& U_\ell(\underline{p}^*) \ge A(1)\overline{u} - B(1)c, \\
\textit{and} \quad & U_r(\overline{p}^*) \ge A(1)\overline{u} - B(1)c.
\end{aligned}$$

The first inequality is strict if and only if $\Gamma'(1)$ *is finite. The second is strict if and only if* $\Gamma'(0) < 0$.

Proof of Lemma 14. We only give the proof for $\underline{p}^*$, the other case is symmetric. Consider the maximization problem in (C.21). The derivative of the objective function simplifies to $p\,(\overline{u} - \underline{u})$. Therefore we can set $\lambda = 1$ and (C.21) reduces to the definition via smooth pasting and value matching as in our original model.

The first inequality is equivalent to

$$\frac{1}{(1+\rho)\Gamma'(1) - \rho} \leq 0,$$

which holds under Assumption 1. The inequality is strict if and only if $\Gamma'(1)$ is finite. The second inequality is equivalent to

$$\frac{\Gamma'(0)}{1+\rho-\rho\Gamma'(0)} \leq 0,$$

which is strict if and only if $\Gamma'(0) < 0$. □

We are now ready to define the opposite-biased strategy. Given that we impose Assumption 1, we only describe the construction for the left branch which is used for $p \leq 1/2$. There are up to four intervals where the opposite-biased strategy takes a different form. First, for $p \leq \underline{p}^*$, the DM takes immediate action. Then there is an interval $(\underline{p}^*, \underline{q}^b]$ where the DM uses the own-biased strategy from our original model. $\underline{q}^b$ is given by

$$\frac{\partial H(\underline{q}^b, \underline{V}_{own}(\underline{q}^b), \underline{V}'_{own}(\underline{q}^b), 1)}{\partial \lambda} = 0.$$

Rearranging this we get

$$\frac{(1+\rho)\Gamma'(\underline{q}^b)}{\rho - (1+\rho)\Gamma'(\underline{q}^b)} + \underline{q}^b + (1-\underline{q}^b)\left(\frac{1-\underline{q}^b}{\underline{q}^b}\frac{\underline{p}^*}{1-\underline{p}^*}\right)^{\rho} = 0,$$

which is equivalent to

$$\underline{V}_{opp}(\underline{q}^b) = A(1)\overline{u} - B(1)c.$$

By Lemma 14, $\underline{q}^b = \underline{p}^*$ if $\Gamma'(1)$ is infinite and otherwise $\underline{q}^b > \underline{p}^*$. If $\underline{q}^b \geq 1/2$ we define the own-biased strategy as in our original model. If $\underline{q}^b < 1/2$, Lemma 12 implies that $\lambda'(\underline{q}^b) < 0$ if we impose the boundary condition $\lambda(\underline{q}^b) = 1$. Denote the unique solution for $p \geq \underline{q}^b$ to (C.20) with $\lambda(\underline{q}^b) = 1$ by $\lambda(p; \underline{q}^b, 1)$. Since by Lemma 12, $\lambda'(p; p, 1) < 0$ for all $p \leq 1/2$, we have $\lambda(p; \underline{q}^b, 1) < 1$ for $p \in (\underline{q}^b, 1/2)$. Finally we need to take care of the possibility that there exists $\underline{q}^s \in (\underline{q}^b, 1/2]$ such that $\lambda(p; \underline{q}^b, 1) = \gamma$. If no such $\underline{q}^s$ exists we set $\underline{q}^s = 1/2$. If Assumption 2 is satisfied, a symmetric construction can be used for the right branch with cutoffs $\overline{q}^b = 1 - \underline{q}^b$ and $\overline{q}^s = 1 - \underline{q}^s$.

We thus define the opposite biased strategy as follows. For $p \notin (\underline{p}^*, \overline{p}^*)$: take the optimal immediate action. For $p \in (\underline{p}^*, \overline{p}^*)$, experiment according to the following attention strategy:

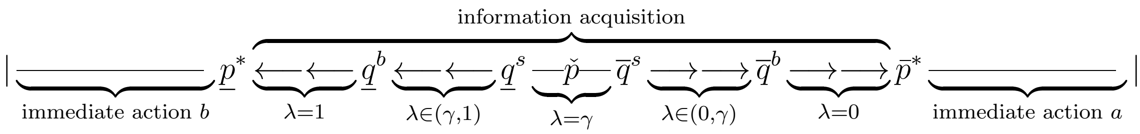
$$\lambda_{own}^{\Gamma}(p) = \begin{cases} 1, & \text{if } p \in (\underline{p}^*, \underline{q}^b], \\ \lambda(p; \underline{q}^b, 1), & \text{if } p \in (\underline{q}^b, \underline{q}^s], \\ \gamma, & \text{if } p \in (\underline{q}^s, \overline{q}^s), \\ \lambda(p; \overline{q}^b, 0), & \text{if } p \in [\overline{q}^s, q^b), \\ 0, & \text{if } p \in [q^b, \overline{p}^*), \end{cases}$$

and take an action corresponding to the signal if one is received.[58] Note that by Lemma 12, $\lambda_{own}^{\Gamma}(p)$ is strictly decreasing if $p \in (\underline{q}^b, \underline{q}^s] \cup [\overline{q}^s, q^b)$. The value of this strategy is given by

$$V_{own}^{\Gamma}(p) = \begin{cases} V_{own}(p), & \text{if } p \leq \underline{q}^b, \\ A\left(\lambda(p; \underline{q}^b, 1)\right)\overline{u} - B\left(\lambda(p; \underline{q}^b, 1)\right)c, & \text{if } p \in (\underline{q}^b, \underline{q}^s], \\ U^S(p), & \text{if } p \in (\underline{q}^s, \overline{q}^s), \\ A\left(\lambda(p; \overline{q}^b, 0)\right)\overline{u} - B\left(\lambda(p; \overline{q}^b, 0)\right)c, & \text{if } p \in [\overline{q}^s, q^b), \\ V_{own}(p), & \text{if } p \geq q^b, \end{cases}$$

where $V_{own}(p)$ denotes the value of the opposite-biased strategy from our original model. Note that since we focus attention on the symmetric case (Assumption 2), the belief that separates the "left branch" and the "right branch" of the opposite-biased solution is given by $\check{p}$. Note also, that in contrast to our original model, we defined the own-biased strategy in a way that it is always weakly greater than $U^S(p)$.

The implied dynamics of the posterior as well as the attention strategy are summarized by the following diagram:



Lemma 15. *Suppose Assumptions 1 and 2 are satisfied. Then V_{own}^{Γ} is continuously differentiable and convex on $\left[0, \underline{q}^s\right)$ and on $(\overline{q}^s, 1]$, respectively, and satisfies (C.11) on $\left[\underline{p}^*, \underline{q}^s\right)$ and on $(\overline{q}^s, \overline{p}^*]$, respectively.*

Proof. We show the Lemma for $p \leq 1/2$. The remaining results follow from a symmetric argument.

We need to show that V_{own}^{Γ} is continuously differentiable at $\underline{q}^b$. For $r > 0$, some algebra yields for $p \leq 1/2$[59]

$$V_{own}^{\Gamma} = A(1)\overline{u} - B(1)c$$

[58]If $\underline{q}^s = \overline{q}^s$, $\lambda_{own}^{\Gamma}(\overline{q}^s) \in \left\{\lambda(\underline{q}^s; \underline{q}^b, 1), \lambda(\underline{q}^s; \overline{q}^b, 0)\right\}$ with an arbitrary tie-breaking rule.

[59]The derivation for $r = 0$ is similar.

$$\iff \left(\frac{\underline{p}^*}{1-\underline{p}^*}\frac{1-p}{p}\right)^{\rho} = 1 - \frac{r}{(1-p)\left(\rho-(1+\rho)\Gamma'[1]\right)}.$$

Substituting this expression in $V_{own}^{\Gamma\prime}(p)$ yields

$$\begin{aligned} V_{own}^{\Gamma\prime}(p)\big|_{V_{own}(p)=A(1)\overline{u}-B(1)c} &= \frac{\left(c+\rho U^S\right)\left(p+(1-p)\Gamma'[1]\right)}{(1-p)\,p\left(\rho-(1+\rho)\Gamma'[1]\right)} \\ &= \frac{\partial\left[A(\lambda(p))\overline{u}-B(\lambda(p))c\right]}{\partial\lambda}\bigg|_{\lambda(p)=1}. \end{aligned}$$

Convexity on $\left[\underline{p}^*, \underline{q}^s\right]$ follows from strict convexity of V_{own} (Lemma 5) and strict convexity of $A(\lambda(p))\overline{u} - B(\lambda(p))c$ (Lemma 13.(b)) and continuous differentiability.

Note that by Lemma 8, it suffices to show that V_{own}^{Γ} satisfies (C.14) for all $\left[\underline{p}^*, \underline{q}^s\right)$ since $V_{own}^{\Gamma}(p) > U^S(p)$ for $p < \underline{q}^s$. We have derived V_{own}^{Γ} from the first order-condition (C.12) and the respective Kuhn-Tucker condition of $p < \underline{q}^b$. Therefore it suffices to show that the maximization problem in the HJB equation is concave. By Lemma 13.(a), this is the case since we have shown that V_{own}^{Γ} is weakly convex. □

Opposite-Biased Learning The second candidate for the value function is obtained by assuming that the DM uses an opposite-biased attention strategy. Specifically, we define a "reference belief" p^* such that the DM chooses $\lambda < \gamma$ for lower beliefs $p < p^*$ and $\lambda > \gamma$ for higher beliefs $p > p^*$. The implied dynamics of the posterior as well as the attention strategy are summarized by the following diagram:

$$\Big|\underbrace{\longrightarrow\longrightarrow\longrightarrow\longrightarrow}_{\lambda\in[0,\gamma)} p^* \underbrace{\longleftarrow\longleftarrow\longleftarrow\longleftarrow}_{\lambda\in(\gamma,1]}\Big|$$

The reference belief is absorbing and we assume that once p^* is reached, the DM adopts the stationary attention strategy $\lambda = \gamma$. Under Assumption 2, we have $p^* = 1/2$. This can also be derived from value matching

$$V(p^*) = U^S(p^*)(= U^S), \tag{C.23}$$

and the tangency condition

$$V'(p^*) = U^{S\prime}(p^*)(= 0). \tag{C.24}$$

Substituting these two conditions together with $\lambda = \gamma$ in (C.12) yields $p^* = 1/2$.[60]

We would now like to construct the opposite-biased strategy in a similar fashion as the own-biased solution, that is, we will identify two types of regions. If $\lambda \in \{0, 1\}$, we will use solutions to (A.7) or (A.8) (with $\overline{\lambda} = 1$, $\underline{\lambda} = 1$ and α replace by λ.) On the other

[60]Note that in contrast to the linear model, we cannot use the HJB equation because for $\lambda = \gamma$, $V'(p)$ vanishes so that substituting (C.24) has no bite.

hand, if $\lambda \in (0,1)$ we will use solutions to (C.20) with a suitable boundary condition. A problem arises since we want to impose the boundary condition $\lambda(p^*) = \gamma$. Note that this implies $\lambda'(p^*) = 0/0$. We therefore begin by identifying a solution to (C.20) that satisfies $\lambda(p^*) = \gamma$ as well as $\lambda'(p^*) > 0$.

Lemma 16. *Suppose Assumptions 1 and 2 are satisfied. Then there exists a unique continuously differentiable function $\hat{\lambda}_{opp}(p)$ which satisfies (C.20) for all p in a neighborhood of $p^* = 1/2$, such that $\lambda(p^*) = \gamma$ and $\lambda'(p^*) > 0$. The derivative at p^* is given by*

$$\hat{\lambda}'_{opp}(p^*) = -(\rho+\gamma) + \sqrt{(\rho+\gamma)^2 - \frac{8(\rho+\gamma)}{\Gamma''(\gamma)}}.$$

Proof of Lemma 16. The ODE (C.20) can be written as

$$\lambda'(p) = \frac{P(p,\lambda(p))}{Q(p,\lambda(p))},$$

where

$$\begin{aligned} P(p,\lambda) &= [\Gamma(\lambda) - \Gamma'(\lambda)\,\lambda + \rho\,(1-\Gamma'(\lambda))] \times [p + \Gamma'(\lambda)\,(1-p)]\,, \\ Q(p,\lambda) &= p(1-p)\Gamma''(\lambda)\,[\Gamma(\lambda) - \lambda]\,. \end{aligned}$$

Since P and Q are both continuous and have continuous partial derivatives, the behavior of solutions that go through points in a neighborhood of (p^*, γ) is, under some conditions (see below), the same as for[61]

$$\lambda'(p) = \frac{a\,(p-p^*) + b\,(\lambda(p)-\gamma)}{c\,(p-p^*) + d\,(\lambda(p)-\gamma)}, \tag{C.25}$$

where

$$\begin{aligned} a &= \partial_p P(p^*,\gamma) = 4\,(\rho+\gamma) > 0, \\ b &= \partial_\lambda P(p^*,\gamma) = (\rho+\gamma)\,\Gamma''(\gamma) < 0, \\ c &= \partial_p Q(p^*,\gamma) = 0, \\ d &= \partial_\lambda Q(p^*,\gamma) = -\frac{1}{2}\Gamma''(\gamma) > 0. \end{aligned}$$

The characteristic equation is

$$x^2 - bx - ad = 0.$$

Since $ad > 0$, the characteristic equation has two reals roots of opposite sign. This implies

[61] See e.g. Bronshtein, Semendyayev, Musiol, and Muehlig (2007).

that (p^*, γ) is a saddle point and there are two continuously differentiable solutions $\lambda(p)$ that pass through (p^*, γ). In the case of a saddle point, the behavior of the solutions of (C.25) in a neighborhood of (p^*, γ) corresponds to the behavior of the solutions to (C.20). Hence there exist two continuously differentiable solutions $\lambda(p)$ that satisfy the boundary condition $\lambda(p^*) = \gamma$.

Next we want to obtain $\lambda'(p^*)$ for these solutions, and show that only one of them has a positive derivative. We have

$$\begin{aligned}\lambda'(p^*) = \lim_{p \to p^*} \lambda'(p) &= \lim_{p \to p^*} \frac{P(p, \lambda(p))}{Q(p, \lambda(p))} \\ &= \lim_{p \to p^*} \frac{\partial_p P(p, \lambda(p)) + \partial_\lambda P(p, \lambda(p))\lambda'(p)}{\partial_p Q(p, \lambda(p)) + \partial_\lambda Q(p, \lambda(p))\lambda'(p)} \\ &= \frac{a + b\lambda'(p^*)}{d\lambda'(p^*)}.\end{aligned}$$

Hence $\lambda'(p^*)$ solves

$$\begin{aligned}x^2 - \frac{b}{d}x - \frac{a}{d} = 0,& \\ \lambda'(p^*) = \frac{b}{2d} \pm \sqrt{\left(\frac{b}{2d}\right)^2 + \frac{a}{d}}.&\end{aligned}$$

Since $a/d > 0$, there is one positive and one negative solution. For the opposite-biased solution, we are interested in a solution that satisfies $\lambda'(p^*) > 0$. Hence we have

$$\begin{aligned}\lambda'(p^*) &= \frac{b}{2d} + \sqrt{\left(\frac{b}{2d}\right)^2 + \frac{4(\rho + \gamma)}{d}} \\ &= -(\rho + \gamma) + \sqrt{(\rho + \gamma)^2 - \frac{8(\rho + \gamma)}{\Gamma''(\gamma)}}.\end{aligned}$$

□

Lemma 16 provides the solution $\hat{\lambda}_{opp}$ which together with $V(p) = A(\hat{\lambda}_{opp}(p))\overline{u}$ defines V_{opp} in a neighborhood of p^*. To extend this definition to $(0, 1)$ we first extend $\hat{\lambda}_{opp}$ to the maximal interval $(\underline{q}, \overline{q})$ where $\hat{\lambda}_{opp}(p) \in (0, 1) \setminus \{\gamma\}$ unless $p = p^*$.

Lemma 17. *Suppose Assumptions 1 and 2 are satisfied. There exist two points $0 \leq \underline{q} < p^* < \overline{q} \leq 1$ such that*

(a) $\hat{\lambda}_{opp}(p)$ is well defined as the unique $\mathcal{C}^1$-solution to (C.20) that satisfies the properties in Lemma 16

(b) $\hat{\lambda}_{opp}(p) > \gamma$ if $p > p^*$ and $\hat{\lambda}_{opp}(p) < \gamma$ if $p < p^*$.

(c) Either $\underline{q} = 0$ or $\hat{\lambda}_{opp}(\underline{q}) = 0$.

(d) Either $\overline{q} = 1$ or $\hat{\lambda}_{opp}(\overline{q}) = 1$.

Note that Properties (c) and (d) mean that the interval $(\underline{q}, \overline{q})$ is the maximal interval where $\hat{\lambda}_{opp}(p) \in (0,1)$.

Proof of Lemma 17. Consider the interval $(\underline{q}, p^*)$. $\hat{\lambda}_{opp}(p) \in (0,\gamma)$ in a neighborhood of p^*. Moreover, (C.20) satisfies local Lipschitz continuity if $p \in (0,p^*)$ and $\lambda \neq \gamma$. Hence, if there exists a $\mathcal{C}^1$ solution to (C.20) with initial condition $\hat{\lambda}_{opp}(p^*-\varepsilon) \in (0,\gamma)$ that satisfies $\hat{\lambda}_{opp}(p) \in (0,\gamma)$ for all $p \in (\underline{q}, p^*)$, then it is the unique such solution. We first show that by extending the interval from a neighborhood of p^* to $(\underline{q}, p^*)$, we do not violate $\hat{\lambda}_{opp}(p) < \gamma$. Suppose by contradiction that there exists $p' < p^*$ such that $\lim_{p \searrow p'} \hat{\lambda}_{opp}(p) \nearrow \gamma$. Note that

$$p' + \Gamma'(\gamma)(1-p') < p^* + \Gamma'(\gamma)(1-p^*) = 0.$$

Hence, since $\Gamma'' < 0$,$\lim_{p \searrow p'} \hat{\lambda}'_{opp}(p) \to \infty$ which contradicts $\lim_{p \searrow p'} \hat{\lambda}_{opp}(p) \nearrow \gamma$. Therefore we can extend the domain of $\hat{\lambda}_{opp}(p)$ to the left until either $p = 0$ or $\hat{\lambda}_{opp}(p) = 0$. This completes the proof for $p < p^*$ and the argument for $p > p^*$ is similar. □

If $\underline{q} > 0$ and $\overline{q} < 1$, respectively, then we further extend $\lambda_{opp}(p)$ to $(0,1)$ by setting $\lambda = 0$ for $p < \underline{q}$ and $\lambda = 1$ for $p > \overline{q}$. We define

$$\lambda^{\Gamma}{}_{opp}(p) := \begin{cases} 0, & \text{if } p \leq \underline{q}, \\ \hat{\lambda}_{opp}(p), & \text{if } p \in (\underline{q}, \overline{q}), \\ 1, & \text{if } p \geq \overline{q}. \end{cases}$$

The value of this strategy is given by

$$V^{\Gamma}_{opp}(p) := \begin{cases} V_0\left(p; \underline{q}, A(0)\overline{u} - B(0)c\right) & \text{if } p \leq \underline{q}, \\ A(\lambda_{opp}(p))\overline{u} & \text{if } p \in (\underline{q}, \overline{q}), \\ V_1\left(p; \overline{q}, A(1)\overline{u} - B(1)c\right) & \text{if } p \geq \overline{q}. \end{cases}$$

Lemma 18. *Suppose Assumptions 1 and 2 are satisfied. Then $V^{\Gamma}_{opp}(p)$ is a $\mathcal{C}^1$ solution to (C.11) and $V^{\Gamma}_{opp}(p)$ is strictly convex on $(\underline{q}, \overline{q})$.*

Proof. The proof has several steps. We give arguments for $p \geq 1/2$. The Lemma then follows by symmetry (Assumption 2) and the fact that $V^{\Gamma}_{opp}(p)$ is constructed in a way that is continuously differentiable at p^* (see (C.24)). Suppose in the following that $p > 1/2$.

First we note that $V^{\Gamma}_{opp}(p)$ is continuously differentiable. This holds by construction for $p \neq \overline{q}$ and at $\overline{q}$ it follows by the same argument as in the proof of Lemma 15.

Second, we shows that $V^{\Gamma}_{opp}(p)$ is strictly convex. For $p > 1/2$, $\lambda^{\Gamma}_{opp}(p) > \gamma$. Therefore, by Lemma 13, strict convexity on $(p^*, \overline{q})$ follows if $p < \pi(\lambda^{\Gamma}_{opp}(p))$ for all $p \in (p^*, \overline{q})$. Note that $\pi(\lambda^{\Gamma}_{opp}(p^*)) = \pi(\gamma) = 1/2$. We show that whenever $p = \pi(\lambda^{\Gamma}_{opp}(p))$, then

$\pi'(\lambda^{\Gamma}_{opp}(p))\lambda^{\Gamma\prime}_{opp}(p) > 1$. This implies that $p < \pi(\lambda^{\Gamma}_{opp}(p))$ for all $p \in (p^*, \overline{q})$. We have

$$\begin{aligned} &\pi'(\lambda^{\Gamma}_{opp}(p^*))\lambda^{\Gamma\prime}_{opp}(p^*) > 1 \\ \iff \frac{2-(\rho+\gamma)\Gamma''(\gamma)}{4(\rho+\gamma)} &\left(\sqrt{(r+\gamma)^2 - \frac{8\,(\rho+\gamma)}{\Gamma''(\gamma)}} - (\rho+\gamma)\right) > 1 \\ &\iff \Gamma''(\gamma) < 0. \end{aligned}$$

for $p > p^*$, we substitute $p = \pi(\lambda^{\Gamma}_{opp}(p))$ in (C.20), which yields (after some algebra)

$$\pi'(\lambda^{\Gamma}_{opp}(p^*))\lambda^{\Gamma\prime}_{opp}(p^*) = 1 + \frac{\Gamma'(\lambda^{\Gamma}_{opp}(p))\left(\rho + \Gamma(\lambda^{\Gamma}_{opp}(p)) - \left(\rho + \lambda^{\Gamma}_{opp}(p)\right)\Gamma'(\lambda^{\Gamma}_{opp}(p))\right)}{\left(\rho + \lambda^{\Gamma}_{opp}(p)\right)\left(\rho + \Gamma(\lambda^{\Gamma}_{opp}(p))\right)\Gamma''(\gamma)} > 1.$$

This completes the proof of convexity on $(p^*, \overline{q})$. For $p > \overline{q}$, convexity has been shown in Lemma 6. Since $V^{\Gamma}_{opp}(p)$ is continuously differentiable at $p = \overline{q}$, $V^{\Gamma}_{opp}(p)$ is strictly convex on $[0, 1]$.

Third, by Lemma 13.(a), convexity implies that the maximization problem in (C.14) is concave so that the first-order condition is sufficient. Therefore, $V^{\Gamma}_{opp}(p)$ satisfies (C.14) or for $p > p^*$.

Finally, convexity, together with (C.23) and (C.24) implies that $V^{\Gamma}_{opp}(p) \geq U^S(p)$ for $p \geq p^*$. Lemma 8 then implies that $V^{\Gamma}_{opp}(p)$ satisfies (C.11). □

Finally we show that $\lambda^{\Gamma}_{opp}(p)$ is strictly increasing.

Lemma 19. *Suppose Assumptions 1 and 2 are satisfied and let $\underline{q}, \overline{q}$ be given as in Lemma 17. Then $\lambda^{\Gamma}_{opp}(p)$ is strictly increasing on $(\underline{q}, \overline{q})$.*

Proof. For $p \in (\underline{q}, \overline{q})$, $V^{\Gamma}_{opp}(p) = A(\lambda_{opp}(p))\overline{u}$. Differentiating with respect to p we get

$$V^{\Gamma\prime}_{opp}(p) = A'(\lambda_{opp}(p))\lambda'_{opp}(p)\overline{u}.$$

Hence if $\lambda'(p) = 0$ for $p \neq 1/2$, we must have $V^{\Gamma\prime}_{opp}(p) = 0$. Since $V^{\Gamma\prime}_{opp}(1/2) = 0$, this violates strict convexity of $V^{\Gamma}_{opp}(p)$. Therefore $\lambda'(p) \neq 0$ for all $p \in (\underline{q}, \overline{q})$. Since $\lambda'(1/2) > 0$, this implies that $\lambda'(p) > 0$ if $p \in (\underline{q}, \overline{q})$. □

C.4.5 Optimal Solution

As in our original model we show that the value function V^{Γ} is the upper envelope of the two solution candidates. In contrast to our original model, the optimal policy is not a bang-bang solution. We show that inside the own-biased region, $\alpha(p) = g^{-1}(\lambda(p))$ is decreasing whenever it is not a corner-solution. This means that more extreme beliefs lead to a more own-biased news-diet. In the opposite-biased region, $\alpha(p)$ is strictly increasing. This implies that more moderate beliefs lead to a more balanced news-diet.

Theorem 2. *Suppose Assumptions 1 and 2 are satisfied.*

(a) If (EXP) is violated then $V^\Gamma(p) = U(p)$ for all $p \in [0,1]$.

(b) If (EXP) is satisfied and $V^\Gamma_{own}(p) > U^S(p)$ for all $p \neq 1/2$, then $V^\Gamma(p) = V^\Gamma_{own}(p)$ for all $p \in [0,1]$, and $\alpha(p) = g^{-1}(\lambda(p))$ is strictly decreasing if $V^\Gamma(p) > U(p)$ and $\alpha(p) = g^{-1}(\lambda(p)) \in (0,1)$.

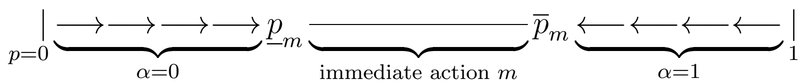
(c) If (EXP) is satisfied and $V^\Gamma_{own}(p) = U^S(p)$ for some $p \neq 1/2$, then $V^\Gamma(p) = \max\left\{V^\Gamma_{own}(p), V^\Gamma_{opp}(p)\right\}$, and $\alpha(p) = g^{-1}(\lambda(p))$ is strictly decreasing if $V^\Gamma(p) = V^\Gamma_{own}(p) > U(p)$ and $\lambda(p) \in (0,1)$, and strictly increasing if $V^\Gamma(p) = V^\Gamma_{opp}(p)$.

Proof of Theorem 2. Follows from the same arguments as the proof of Theorem 1. □

C.5 Multiple Actions

In this Appendix, we extend the model in Section 2 to include a third action $x = m$ which yields u^R_m and u^L_m in states R and L. Up to relabeling of the actions it is without loss to assume that $u^R_m \in \left(u^R_\ell, u^R_r\right)$. Further we assume $u^L_m < u^L_\ell$ which guarantees that action m does not dominate action ℓ for all beliefs.

The optimal policy will be affected by the availability of action m if it is optimal to take this action for some beliefs. To identify when this is the case, we define a strategy that specifies a stopping region $[\underline{p}_m, \overline{p}_m]$ in which action m is taken immediately. For $p > \overline{p}_m$, the strategy prescribes attention to the L-biased news source ($\alpha = 1$) and for $p < \underline{p}_m$, the strategy prescribes attention to the R-biased news source ($\alpha = 0$). We call this strategy the "m-strategy." It has the following structure:



If this strategy is part of the optimal solution (for some range of belief), the boundary points $\underline{p}_m$ and $\overline{p}_m$ must satisfy value-matching and smooth-pasting conditions that resemble those used to define $\underline{p}^*$ and $\overline{p}^*$. We will define $\overline{p}_m$ by imposing smooth pasting and value matching with $U_m(p)$ in (A.8):

$$c + \rho U_m(p) = \lambda p \left(u^R_r - U_m(p)\right) - \lambda p(1-p)U'_m(p). \tag{C.26}$$

Similarly we will define $\underline{p}_m$ by imposing smooth pasting and value matching with $U_m(p)$ in (A.7):

$$c + \rho U_m(p) = \lambda(1-p)\left(u^L_\ell - U_m(p)\right) + \lambda p(1-p)U'_m(p). \tag{C.27}$$

The following lemma identifies when solutions to (C.26) and (C.27) exist, and when these solutions can be used to define the cutoffs $\underline{p}_m$ and $\overline{p}_m$ in a way the m-strategy only prescribes information acquisition if it is not dominated by immediate action m or by the stationary strategy.

Lemma 20.

(a) Let $u_m^R \geq U^{FA}(1)$ or $c + \rho u_m^L \leq 0$. If $q_1 \in (0,1)$ is a solution to (C.26), then $V_1(p; q_1, U_m(q_1)) \leq U_m(p)$ for all $p \in [q_1, 1]$.

(b) If $u_m^R < U^{FA}(1)$ and $c + \rho u_m^L > 0$, then there exists a unique solution $q_1 \in (0,1)$ to (C.26) given by

$$q_1 = \frac{u_m^L \rho + c}{\rho\left(u_m^L - u_m^R\right) + \left(u_r^R - u_m^R\right)\lambda}. \tag{C.28}$$

and $V_1(p; q_1, U_m(q_1))$ is strictly convex on $[q_1, 1]$.

(c) Let $u_m^L \geq U^{FA}(0)$ or $c + \rho u_m^R \leq 0$. If $q_2 \in (0,1)$ is a solution to (C.27), then $V_0(p; q_2, U_m(q_2)) \leq U_m(p)$ for all $p \in [0, q_2]$.

(d) If $u_m^L < U^{FA}(0)$ and $c + \rho u_m^R > 0$, then there exists a unique solution $q_2 \in (0,1)$ to (C.27) given by

$$q_2 = \frac{\left(u_\ell^L - u_m^L\right)\lambda - u_m^L \rho - c}{\rho\left(u_m^R - u_m^L\right) + \left(u_\ell^L - u_m^L\right)\lambda} \tag{C.29}$$

and $V_0(p; q_2, U_m(q_2))$ is strictly convex on $[0, q_2]$.

(e) Suppose $u_m^R < U^{FA}(1)$ and $c + \rho u_m^L > 0$, and $u_m^L < U^{FA}(0)$ and $c + \rho u_m^R > 0$. If $U_m(q_1) \geq U^S(q_1)$ and $U_m(q_2) \geq U^S(q_2)$, then $q_1 \geq q_2$.

Proof. For (a) and (b) we note that the general solution to (A.8) is given by

$$V_1(p) = \underbrace{\frac{p\rho u_r^R \lambda - c\left(\rho + (1-p)\lambda\right)}{\rho(\rho+\lambda)}}_{=:z(p)} + \left(\frac{1-p}{p}\right)^{\frac{\rho}{\lambda}} (1-p)\, C,$$

there C is the constant of integration. Clearly, the sign of C determines whether the solution is convex or concave since

$$\frac{d^2}{dp^2}\left(\left(\frac{1-p}{p}\right)^{\frac{\rho}{\lambda}} (1-p)\right) > 0$$

Moreover, we note the $V_1(1) = U^{FA}(1)$ regardless of the value of the constant C.

For the proof of (a) we distinguish several cases: Case 1: If $u_m^R = U^{FA}(1)$ and $c + \rho u_m^L = 0$. In this case, $U_m(p) = z(p)$ for all p. Hence any $q_1 \in (0,1)$ satisfies (C.26) and smooth pasting but $V(p; q_1, U_m(q_1)) = U_m(p)$ for all $p \in [0,1]$.

Case 2: $u_m^R > U^{FA}(1)$. We first show that if (C.26) and smooth pasting is satisfied for $p' \in (0,1)$, then $U_m(p') < z(p)$. Suppose by contradiction that $U_m(p') \geq z(p')$. If (C.26) is satisfies at p', then $U_m(p)$ is tangent to $V_1(p; p', U_m(p')$at p' and since $V_1(p; p', U_m(p')) \geq z(p)$, $V_1(p; p', U_m(p'))$ is weakly convex as a function of p. But this implies that $V_1(1; p', U_m(p')) \geq U_m(1) = u_m^R > U^{FA}(1)$. This is a contradiction since we argued above that any solution to (A.8) satisfies $V_1(1) = U^{FA}(1)$. Hence (C.26) or smooth pasting is vioalated at p' if $U_m(p') \geq z(p')$. If $U_m(p') \leq z(p')$, (C.26), and smooth pasting is satisfied for $p' \in (0,1)$, then $V_1(p; p', U_m(p'))$ is strictly concave as a function of p and tangent to $U_m(p)$ at p'. Hence $V_1(p; p', U_m(p')) < U_m(p)$ for all $p > p'$.

Case 3: $u_m^R < U^{FA}(1)$. If $c + \rho u_m^L \leq 0$, then $U_m(0) < z(0)$ and since $z(1) = U^{FA}(1)$ we have $U_m(p) < z(p)$ for all p. As in case 2, if $p' \in (0,1)$ satisfies (C.26) and smooth pasting, then $V_1(p; p', U_m(p')) < U_m(p)$ for all $p > p'$ which contradicts $V_1(1; p', U_m(p')) = U^{FA}(1)$. Hence there is no solution to (C.26) that satisdies smooth pasting. This concludes the proof of (a).

For (b), note that if $u_m^R < U^{FA}(1)$ and $c + \rho u_m^L > 0$, then $U_m(p)$ crosses $z(p)$ from above. As in case 3 in the proof for part (a), $z(p') > U_m(p)$ implies that (C.26) and smooth pasting cannot be both satisfied. Next we identify a solution q_1 to (C.26) for which $V'(q_1; q_1, U_m(q_1)) = U_m'(q_1)$. If $U_m(q_1) = z(q_1)$ then $V'(q_1; q_1, U_m(q_1)) = z'(q_1) = U_m'(q_1)$. On the hand $\lim_{q_1 \to 0} V'(q_1; q_1, U_m(q_1)) = -\infty$. Therefore, the intermediate value theorem implies that there exists $q_1 \in (0,1)$ such that $V'(q_1; q_1, U_m(q_1)) = U_m'(q_1)$ and simple algebra shows that is is given by (C.28).

The proofs of (c) and (d) follow from a similar argument. For part (e) suppose by contradiction that $q_1 < q2$. Since both $V_1(p; q_1, U_m(q_1))$ and $V_0(p; q_0, U_m(q_0))$ are strictly convex on $[q_1, q_2]$ and coincide with $U_m(p)$ at q_1 and q_2, respectively, there exists $p' \in (q_1, q_2)$ such that $V_1(p'; q_1, U_m(q_1)) = V_0(p'; q_0, U_m(q_0)) > U_m(p')$ and $V_1'(p'; q_1, U_m(q_1)) > V_0'(p'; q_0, U_m(q_0))$. Since $U_m(p) \geq U^S(p)$ for $p \in q_1, q_2$ and both U_m and U^S are linear, we have $V_1(p'; q_1, U_m(q_1)) = V_0(p'; q_0, U_m(q_0)) > U_S(p')$. By Lemma 2 this implies $V_1'(p'; q_1, U_m(q_1)) < V_0'(p'; q_0, U_m(q_0))$ which is a contradiction. Therefore we must have $q_1 \geq q_2$. □

Based on the results of this lemma, we define $\underline{p}_m$ and $\overline{p}_m$ as follows:

$$\overline{p}_m = \begin{cases} q_1, & \text{if } u_m^R < U^{FA}(1),\ c + \rho u_m^L > 0, \text{ and } U_m(q_1) \geq U^S(q_1) \\ 1, & \text{otherwise.} \end{cases}$$

$$\underline{p}_m = \begin{cases} q_2, & \text{if } u_m^L < U^{FA}(0),\ c + \rho u_m^R > 0, \text{ and } U_m(q_2) \geq U^S(q_2) \\ 0, & \text{otherwise.} \end{cases}$$

Consider $\overline{p}_m$. By Lemma 20.(a)-(b) $u_m^R < U^{FA}(1)$ together with $c + \rho u_m^L > 0$ is a necessary and sufficient condition for the existence of a solution in $(0,1)$ to (C.26) that satisfies smooth pasting and is not dominated by immediate action m. Hence if the necessary and sufficient condition is violated we set $\overline{p}_m = 1$. Similarly, Lemma 20.(c)-(d) motivates the definition of $\underline{p}_m = 0$ if $u_m^B \geq U^{FA}(0)$ and $c + \rho u_m^R \leq 0$.

The requirements that $U_m(q_1) \geq U^S(q_1)$ in the definition of $\overline{p}_m$ and $U_m(q_2) \geq U^S(q_2)$ in the definition of $\underline{p}_m$, guarantee, respectively, that the m-strategy always has the structure depicted in the diagram above because it avoids defining $\overline{p}_m = q_1$ and $\underline{p}_m = q_2$ when $q_2 > q_1$.

The value of the m-strategy is

$$V_m(p) := \begin{cases} V_0(p; \underline{p}_m, U_m(\underline{p}_m)), & \text{for } p < \underline{p}_m, \\ U_m(p), & \text{for } p \in \left[\underline{p}_m, \overline{p}_m\right], \\ V_1(p; \overline{p}_m, U_m(\overline{p}_m)), & \text{for } p > \overline{p}_m. \end{cases}$$

The Lemmas leading to the upper envelope characterization of the value function in Proposition 9 depend on the properties of branches defined by particular solutions to (A.7) and (A.8). Therefore the same steps can be applied in this extension and we obtain that the value function of the extended problem is given by:

$$V(p) = \max \{V_{own}(p), V_{opp}(p), V_m(p)\} .$$

It is straightforward to extend this to more than three actions. Suppose we have actions ℓ, r as well as additional actions $m_1, m_2, \ldots$, where for all $i = 1, 2, \ldots$, $(u^R_{m_i}, u^A_{m_i})$ satisfy the conditions formulated for action m at the beginning of this section. In this case we define an m_i-strategy for each of the actions in the same way as above. Denote the value of strategy m_i by $V_{m_i}(p)$. The value function of the DM's problem is then given by

$$V(p) = \max \{V_{own}(p), V_{opp}(p), V_{m_1}(p), V_{m_2}(p), \ldots\} .$$