WINPEC Working Paper Series No. E2521
October 2025

The Model of Clickbait: Fact-Checking and Endogenous Information Acquisition

Masayuki Odora

Waseda INstitute of Political EConomy
Waseda University
Tokyo, Japan

The Model of Clickbait: Fact-Checking and Endogenous Information Acquisition*

Masayuki Odora†

September 22, 2025

Abstract

I study the role of fact-checking in a two-period strategic communication game between a decision maker and a media outlet. The decision maker relies on the outlet's article to take a binary action, while the outlet may exert costly effort to acquire information before publishing. The decision maker is uncertain about the outlet's motive: the outlet might be opportunistic, caring only about attracting clicks. Fact-checking probabilistically reveals the payoff-relevant state. I highlight a trade-off between the *diagnostic* effect and the *discipline* effect that arises when the probability of fact-checking successfully revealing the state increases. Consequently, introducing fact-checking or increasing its success probability may, in some parameter ranges, reduce the decision maker's welfare.

Keywords: Fact-checking; Information acquisition; Cheap-talk
JEL classification: C72; D83; L82

1 Introduction

The public relies on news articles to make decisions (e.g., which candidate to vote for or whether to get vaccinated). Before publishing, the media outlets exert costly effort to acquire information. Today, most news is distributed online, where readers typically must click a headline to access the full content.

---

*I am grateful to Yasushi Asako, Kohei Kawamura, Yukio Koriyama, Daiki Kishishita, Shotaro Shiba for their helpful discussions and comments. I also wish to thank the participants at WINPEC Microeconomics Workshop 2024, the Game Theory Workshop 2024, and the Japanese Society for Quantitative Political Science (JSQPS) 2023 summer meeting. An earlier version of this paper was circulated under the title "The Effect of Fact-checking: The Media's Endogenous Information Acquisition." This study was supported by JSPS Topic-Setting Program to Advance Cutting-Edge Humanities and Social Sciences Research Grant Number JPJS00123811919 and Japan Society for the Promotion of Science, Grant Number 25K23149.

†Global Education Center, Waseda University, 1-104 Totsukamachi, Shinjuku-ku, Tokyo 169-8050, Japan. Email: odora_masa@fuji.waseda.jp

Such informational environment has fueled the clickbait phenomenon, a growing concern in modern democracies. Clickbait refers to articles that use sensational or misleading headlines, which are often misaligned with the underlying content, to induce clicks. In many cases, clickbait producers invest little in acquiring information since their only motivation is attracting clicks; as a result, the content frequently includes misinformation or fake news. A prominent example is the fake news disseminated during the 2016 U.S. presidential election (Allcott and Gentzkow, 2017; Hughes and Waismel-Manor, 2021). Macedonian teenagers posted false stories on political news sites—for instance, "The Pope endorsed Donald Trump"—to earn advertising revenue from clicks rather than to endorse their preferred candidate (see also Subramanian, 2017). In many cases, they simply copied pro-Trump articles from other websites and republished them on their own ad-monetized sites (Subramanian, 2017).

In response to the rise of fake news, fact-checking organizations emerged. This raises two questions: How does fact-checking affect outlets' information acquisition and transmission? Do news consumers benefit from the introduction of fact-checking?

I developed a two-period strategic communication game between a media outlet and a decision maker. The media outlet publishes a news article to the decision maker. Before publication, the outlet acquires information by exerting costly effort. A news article comprises a *headline* and a *content*; the decision maker observes the headline, but must click it to access the content. The media outlet chooses its headline as a non-verifiable message (i.e., cheap-talk), whereas the content must be consistent with the acquired information. By clicking and observing the content, the decision maker gains access to the outlet's information as well. Fact-checking reveals the payoff-relevant state in period 1 with positive probability at the end of the first period. The decision maker is uncertain about the outlet's motive: the outlet is either a good type, sharing the preferences with the decision maker, or an opportunistic type, caring only about attracting clicks. The outlet has no reputational concern per se; rather, it is endogenously created by fact-checking (Morris, 2001).

Main insights are as follows. First, period-2 outcome is pinned down by the posterior reputation that the outlet is good at the beginning of period 2. When this reputation is low, the decision maker neither clicks nor follows recommendations; both types then choose zero effort and the period-2 equilibrium is babbling. When the posterior reputation is high, a *clickbait equilibrium* arises: the opportunistic type always sends a *suspicious* headline (one that goes against the decision maker's ex-ante optimal action), yet the decision maker clicks it for instrumental reasons (cf. Suen, 2004; Wu and Nachbar, 2024). In such an equilibrium, the good type reports headlines truthfully and exerts positive effort. Both types prefer the clickbait equilibrium to the babbling one.

Second, reputation concerns alone do not induce period-1 effort. Indeed, absent fact-checking, the opportunistic type never exerts effort (the good type may do so to induce the desired current-period action, but not for reputational reasons).

Third, fact-checking creates incentives to invest in information acquisition. Suppose period 1 also features the clickbait structure. If the decision maker clicks a suspicious headline and the observed content *matches* the headline, the outlet's type remains unidentified. Here is where fact-checking bites: by revealing the true state, it reveals whether the article was correct and thereby *diagnoses* the outlet's motive. Namely, confirmation shifts the posterior reputation upward, yielding the clickbait equilibrium in period 2; disconfirmation shifts it downward, yielding babbling. This diagnostic channel may operate as a carrot-and-stick that can induce even the opportunistic type to exert effort; when it does, the decision maker's welfare rises by introducing fact-checking.

Finally, improving the effectiveness of fact-checking can harm the decision maker. As fact-checking becomes more likely to reveal the true state, both types of the outlet exert more efforts—the *discipline effect*—because carrot-and-stick incentive becomes more responsive to effort. However, this also makes it harder for the decision maker to infer the outlets' motive: as the success probability increases, the effort gap between types shrinks, so the likelihood of obtaining correct evidence becomes similar across types. Consequently, the posterior becomes less responsive to confirmation or disconfirmation by fact-checking; fact-checking no longer serves as a diagnostic device, and confirmation and disconfirmation no longer provide carrot-and-stick incentive. When success probability is sufficiently high, the diagnostic effect vanishes; the opportunistic type ceases to exert effort, and then so does the good type. Thus, increasing the effectiveness of fact-checking might deter information acquisition by both types and reduce the decision maker's welfare.

The rest of the paper is organized as follows: Below I summarize the related literature. Section 2 presents the model. Section 3 characterizes period-2 equilibria. Section 4.1 analyzes the equilibria of an entire game including period 1 in the absence of fact-checking; Section 4.2 does so with fact-checking. Section 4.2.3 performs comparative statics with respect to the probability of fact-checking successfully revealing the state and highlights the trade-off between the diagnostic and discipline effect. Section 5 compares the outcome across two settings—with and without fact-checking. Section 6 concludes. Proofs omitted from the text appear in Appendix A.

1.1 Related Literature

This paper relates to and contributes to several strands of research.

Lie Detection and Strategic Communication A recent literature studies how lie-detection technologies affect information transmission between an informed sender and an uninformed receiver/decision maker (Balbuzanov, 2019; Levkun, 2022; Tam and Sadakane, 2023).[1] This paper differs in two central respects. First, in much of this literature the sender is fully informed about the payoff-relevant state, while here the outlet endogenously acquires information by exerting costly effort, and studies how fact-checking affect the incentive to acquire information. The perfect-information assumption may be suitable for politicians (who often talk about

[1]Florian and Weicheng (2021) studies the effect of lie detection in a Bayesian persuasion framework.

themselves), but endogenizing acquisition process is more natural for analyzing the effect of fact-checking on the media outlets; in this sense, this paper complements the game-theoretic analysis on fact-checking. Second, the sender's objective differs: the opportunistic type in my model is click-driven (it cares only about attracting clicks), whereas senders in the above studies typically have ideological motives and wish to induce particular actions by the decision maker—e.g., voting for a particular candidate. Although providers of fake news can be driven by either pecuniary or ideological motives (Allcott and Gentzkow, 2017, among others), the interaction between fact-checking and *clickbait* has, to my knowledge, not been studied. This paper fills this gap by offering a formal analysis in which a click-driven outlet produces clickbait in equilibrium.

Endogenous Information Acquisition and Transmission This paper is also related to the literature on strategic communication with endogenous information acquisition (Argenziano, Severinov and Squintani, 2016; Luo and Rozenas, 2025; Pei, 2015). Argenziano, Severinov and Squintani (2016) and Pei (2015) introduce an information acquisition stage into the Crawford-Sobel framework (with different information acquisition technologies). In particular, Pei (2015) shows that information obtained at a cost is transmitted truthfully: if a sender preferred to hide certain information, she would not have paid to acquire it in the first place. While the acquisition technology in this paper differs, the good outlet, whose preference coincides with the decision maker, faces similar incentives: she exerts effort to acquire information only if the decision maker obeys the information. By contrast, the opportunistic outlet, who cares nothing about the decision maker's action, may still misreport even after acquiring information at cost when fact-checking is present, because the driving force of acquisition in my model is the reputational concern.

Argenziano, Severinov and Squintani (2016) compare delegation to communication, whereas I compare the decision maker's welfare with and without fact-checking. Luo and Rozenas (2025) examines endogenous acquisition prior to the transmission of a non-verifiable message and allows for lie detection. They model the acquisition process as an information design (Bergemann and Morris, 2019; Kamenica and Gentzkow, 2011). Hence, there is no direct cost, which plays a crucial role in this study.

2 Model

Consider a two-period advice game involving an uninformed decision maker (he) and a media outlet (she). In each period $t \in \{0, 1\}$, an unknown state $\omega_t \in \{0, 1\}$ is drawn independently across periods, with each state occurring with equal probability. In each period, the media outlet exerts costly effort to acquire information and publishes a news article addressed to the decision maker, (henceforth the DM). At the end of each period, the DM chooses an action $a_t \in \{0, 1\}$.

Type of the Media: The media outlet is either good or opportunistic, and its type is the outlet's private information. The prior probability that the outlet is good type is $\lambda_1 \in (0,1)$, which can be interpreted as its prior reputation. Throughout, I abbreviate the event that the outlet is of the good (opportunistic) type by G (O), and write Pr(G) and Pr(O) for Pr(Good) and Pr(Opportunistic), respectively.

Information Acquisition: In each period, the outlet obtains the private signal about the state by exerting costly effort. Specifically, she first chooses an effort level $\mu^t \in [0, 1/2]$ and then receives a signal $\sigma_t \in C := \{0, 1\}$ with distribution

$$\Pr(\sigma_t = 1 \mid \omega_t = 1) \;=\; \Pr(\sigma_t = 0 \mid \omega_t = 0) \;=\; \tfrac{1}{2} + \mu^t.$$

If the outlet exerts effort μ, she incurs the cost μ^2.

The DM cannot observe the outlet's chosen effort level. In general, the accuracy of the outlet's evidence is neither required to be disclosed nor inferable from the article. It is therefore natural to assume the effort level is unobservable to the DM. The effort level chosen by the good and the opportunistic type are denoted μ^t_{G} and μ^t_{O}, respectively.

Publication of the News Article: In general, the media outlet publishes a news article, whose *content* is summarized by a *headline*. The outlet may freely choose the headline, including one that may appear inconsistent with the eventual content. By contrast, the content must provide some evidentiary basis for the article's conclusion, as it is substantially longer than the headline.

The consumer of the news article can observe the headline without clicking, but must *click* the headline to access the content. Upon clicking, he obtains full access to the article and can examine its content.

The model in this study captures these features as follows. After observing her signal, the outlet publishes a non-verifiable (cheap-talk) headline $m_t \in \mathcal{M} := \{0, 1\}$, which is publicly observed by the DM. The DM then decides whether to click on the headline m_t. If he clicks, he observes the outlet's private signal. Clicking entails an arbitrarily small cost $c > 0$.

Thus, the outlet may choose any headline at her will. By contrast, this paper implicitly assumes that the content is verifiable and the outlet is forced to disclose the evidence, which is only observable by clicking.

Payoffs: The DM's payoff in each period U^t_{DM} is given by

$$U^t_{\mathrm{DM}} = u^t_{\mathrm{DM}}(a_t, \omega_t) - c \cdot \mathbf{1}_{\mathrm{click}},$$

where

$$u^t_{\text{DM}}(a_t, \omega_t) = \begin{cases} 0 & \text{if } a_t = 0 \\ -q & \text{if } a_t = 1,\ \omega_t = 0 \\ 1-q & \text{if } a_t = 1,\ \omega_t = 1 \end{cases},$$

and $\mathbf{1}_{\text{click}}$ is an indicator function equal to one if the DM clicks the headline. The DM's payoff in an entire game is $U^1_{\text{DM}} + U^2_{\text{DM}}$.

Let $\pi := \Pr(\omega_t = 1 \mid \Omega)$ denote the DM's posterior belief that $\omega_t = 1$, where Ω is the available information. The expected payoff from choosing $a_t = 1$ is $\pi - q$, while that from $a_t = 0$ is 0. Hence, his optimal action is $a_t = 1$ if and only if $\pi \geq q$. Hereafter assume $q \in (1/2, 1)$; with a uniform prior over states, the ex-ante optimal action is $a_t = 0$.

For the outlet, the good type's period-t payoff depends on: (i) whether the DM takes the correct action, (ii) the cost of information acquisition, and (iii) an arbitrarily small truth-telling benefit $k_{\text{G}} > 0$ (equivalently, a negligible disutility from lying). Specifically, the payoff of the good type is

$$U^t_{\text{G}} = u^t_{\text{DM}} - (\mu^t_G)^2 + k_{\text{G}} \cdot \mathbf{1}_{\{m_t = \sigma_t\}}.$$

If the good type exerts effort μ^t_{G}, her posterior belief that $\omega_t = 1$ conditional on σ_t is

$$\Pr(\omega_t = 1 \mid \sigma_t) = \begin{cases} \frac{1}{2} + \mu^t_{\text{G}} & \text{if } \sigma_t = 1, \\ \frac{1}{2} - \mu^t_{\text{G}} & \text{if } \sigma_t = 0. \end{cases}$$

Hence, upon receiving $\sigma_t = 0$, she strictly prefers the DM to choose $a_t = 0$ regardless of her effort μ^t_{G}; Upon receiving $\sigma_t = 1$, she prefers $a_t = 1$ if and only if

$$\underbrace{\tfrac{1}{2} + \mu^t_{\text{G}} - q}_{\text{payoff from } a_t=1} \geq \underbrace{0}_{\text{payoff from } a_t=0}.$$

The opportunistic type's period-t payoff depends on : (i) whether her headline is clicked, (ii) her information acquisition cost, and (iii) truth-telling benefit $k_{\text{O}} > 0$. Specifically, it is given by

$$U^t_{\text{O}} = \mathbf{1}_{\text{clicked}} - (\mu^t_{\text{O}})^2 + k_{\text{O}} \cdot \mathbf{1}_{\{m_t = \sigma_t\}}.$$

The payoffs over two periods are $U^1_{\text{G}} + U^2_{\text{G}}$ for the good type and $U^1_{\text{O}} + U^2_{\text{O}}$ for the opportunistic type.

Fact-Checking: At the end of period 1, a fact-checker reveals the true state ω_1 with probability $\beta \in (0, 1]$; with probability $1 - \beta$, the check fails. Let $\gamma \in \{0, 1, \emptyset\}$ denote the outcome of the fact-checking, where $\emptyset$ indicates failure. Then

$$\Pr(\gamma = 1 \mid \omega_1 = 1) = \Pr(\gamma = 0 \mid \omega_1 = 0) = \beta,$$
$$\Pr(\gamma = \emptyset \mid \omega_1 = 1) = \Pr(\gamma = \emptyset \mid \omega_1 = 0) = 1 - \beta.$$

I compare the equilibrium outcome with and without fact-checking; the no-fact-checking benchmark corresponds to $\beta = 0$.

Timing of the Game and Equilibrium Concept. The game unfolds over two periods. At the start of period 1, Nature draws a state $\omega_1 \in \{0, 1\}$, which is unobserved by either player. The media outlet then chooses an effort level, receives a private signal σ_1, and publishes a headline m_1. Upon observing m_1, the DM decides whether to click and subsequently chooses an action a_1. If fact-checking is available, the true state ω_1 is publicly revealed with probability $\beta \in (0, 1]$; otherwise, no additional information is disclosed.

Period 2 proceeds analogously. A new state ω_2 is independently drawn, after which the outlet chooses an effort level, observes a signal σ_2, and publishes a headline m_2. The DM again decides whether to click and then chooses an action a_2. The game then terminates and payoffs are realized. For simplicity, future payoffs are not discounted.

Throughout, I confine attention to Perfect Bayesian equilibria satisfying in which the good type reports headlines truthfully in both periods.[2]

3 Equilibrium in Period 2

Let $\lambda_2 \in [0, 1]$ denote the posterior probability that the outlet is of the good type. Although λ_2 is endogenously determined in the equilibrium, in this section it is taken as given and the period-2 equilibrium is characterized for a given λ_2. Since this section focuses exclusively on period 2, the time superscript is omitted, and the effort levels are simply denoted by μ_{G} and μ_{O}.

First observation is that the opportunistic type exerts no effort in the last period. This holds true in any period-2 equilibrium.

Proposition 1. *Fix any $\lambda_2 \in [0, 1]$. In any equilibrium in period 2, the opportunistic type exerts no effort: $\mu_{\mathrm{O}}^* = 0$.*

The conclusion follows from the fact that the DM cannot observe the effort level. Suppose by contradiction that the opportunistic outlet chooses a positive effort level. Because effort is unobservable to the DM, whether or not being clicked is independent of the outlet's effort. The opportunistic type can therefore profitably deviate to zero effort—leaving the probability of being clicked unchanged while strictly reducing the cost of effort—and thereby raise her

[2] Of course, players are allowed to deviate to any other strategy.

payoff. Hence, no period-2 equilibrium can sustain strictly positive effort by the opportunistic type.

Definition 1 (Informativeness of the period-2 equilibrium). A period-2 equilibrium is *uninformative* if both types choose zero effort, i.e., $(\mu_G^*, \mu_O^*) = (0, 0)$. Conversely, the equilibrium is *informative* if at least one type exerts positive effort in period 2.

3.1 Informative Equilibrium in Period 2

I begin by characterizing the good type's effort level in an informative period-2 equilibrium, denoted by μ_G^*, which is positive by definition of informativeness. In any such equilibrium, the good type's effort μ_G^* must be a best response at the moment it is chosen: no deviation from μ_G^* can raise her continuation payoff when all other strategies—including her headline strategy—remain fixed.

The continuation payoff from choosing μ_G is obtained as follows. Conditional on observing $\sigma_2 = 1$ (which occurs with probability $1/2$), the good type obtains $1/2 + \mu_G - q$ if the DM chooses $a_2 = 1$, and zero otherwise. Conditional on $\sigma_2 = 0$, she obtains $1/2 - \mu_G - q$ if $a_2 = 1$, and zero otherwise.

Hence μ_G^* solves

$$\max_{\mu_G} \Bigg[\overbrace{\frac{1}{2}\Pr(a_2 = 1 \mid \sigma_2 = 1)\left(\tfrac{1}{2} + \mu_G - q\right) + \frac{1}{2}\Pr(a_2 = 1 \mid \sigma_2 = 0)\left(\tfrac{1}{2} - \mu_G - q\right)}^{\text{action-related payoff}}$$
$$\underbrace{-(\mu_G)^2}_{\text{cost of effort}} + \underbrace{\tfrac{k_G}{2}\left\{\Pr(m_2 = 1 \mid \sigma_2 = 1) + \Pr(m_2 = 0 \mid \sigma_2 = 0)\right\}}_{\text{truth-telling benefit}} \Bigg],$$

where $\Pr(m_2 \mid \sigma_2)$ denotes the good type's headline strategy, and $\Pr(a_2 \mid \sigma_2)$ denotes the probability that the DM chooses a_2 from the point of view of the good type with σ_2. Note that the first line of the objective function captures the action-related payoff, and neither $\Pr(m_2 \mid \sigma_2)$ nor $\Pr(a_2 \mid \sigma_2)$ depends on μ_G.

The first-order condition yields

$$\mu_G^* = \frac{1}{4}\left(\Pr(a_2 = 1 \mid \sigma_2 = 1) - \Pr(a_2 = 1 \mid \sigma_2 = 0)\right). \tag{1}$$

Hence, in any informative equilibrium, it must be that $\Pr(a_2 = 1 \mid \sigma_2 = 1) > 0$ and $\Pr(a_2 = 0 \mid \sigma_2 = 0) > 0$; otherwise the good type exerts no effort. Here $\Pr(a_2 \mid \sigma_2)$ denotes the probability of the DM choosing a_2 from the point of view of the good type with σ_2. Note that this is pinned down by both the DM's strategy and the good type's headline strategy.

Equation (1) implies the following requirement on the DM's equilibrium strategy: from the perspective of the good type with signal $\sigma_2 \in \{0, 1\}$, there must exist at least one headline that induces the DM to choose action $a_2 = \sigma_2$. For example, consider the following DM's strategy:

$$\begin{cases} \text{if } m_2 = 1: & \text{click; upon observing } \sigma_2, \text{ choose } a_2 = \sigma_2, \\ \text{if } m_2 = 0: & \text{do not click; choose } a_2 = 0. \end{cases}$$

Under this strategy, a good outlet with $\sigma_2 = 1$ can secure $a_2 = 1$ only by sending $m_2 = 1$, whereas when $\sigma_2 = 0$ both headlines ultimately lead to $a_2 = 0$.

I construct an informative equilibrium in which the DM follows the above strategy. I now derive the equilibrium headline strategies for each type. The opportunistic type must choose $m_2 = 1$, since only $m_2 = 1$ is clicked. For the good type, the case $\sigma_2 = 0$ is straightforward: she must choose $m_2 = 0$, since she prefers the DM to take action $a_2 = 0$, and the truthful headline (i.e., $m_2 = 0$) achieves this. When $\sigma_2 = 1$, truthful reporting ($m_2 = 1$) induces $a_2 = 1$ and yields payoff

$$\frac{1}{2} + \mu_{\mathrm{G}}^* - q + k_{\mathrm{G}},$$

whereas $m_2 = 0$ yields zero. In this class of equilibria, $1/2 + \mu_{\mathrm{G}}^* - q + k_{\mathrm{G}}$ must be positive, so she chooses truthful headline $m_2 = 1$; if it were negative, she would choose $m_2 = 0$ in this equilibrium. However, given such a headline strategy, the DM would no longer choose $a_2 = 1$ upon observing $\sigma_2 = 1$ after $m_2 = 1$, leading to a contradiction; if it were zero, her action-related payoff would be zero, making costly effort suboptimal and contradicting the informativeness of the equilibrium. Hence, the good type reports truthfully. Given the good type's and the DM's strategies, equation (1) pins down $\mu_{\mathrm{G}}^* = 1/4$. Thus, the equilibrium strategies have been fully specified.

In this equilibrium, for the DM to choose $a_2 = 1$ upon clicking $m_2 = 1$ and observing $\sigma_2 = 1$, his posterior that $\omega_2 = 1$ must satisfy

$$\Pr(\omega_2 = 1 \mid m_2 = 1, \sigma_2 = 1) \geq q,$$

equivalently,

$$\frac{1}{2} + \Pr(\mathrm{G} \mid m_2 = 1, \sigma_2 = 1)\, \mu_{\mathrm{G}}^* + \Pr(\mathrm{O} \mid m_2 = 1, \sigma_2 = 1)\, \mu_{\mathrm{O}}^* \geq q.$$

With $(\mu_{\mathrm{G}}^*, \mu_{\mathrm{O}}^*) = (1/4, 0)$, this reduces to $1/2 + \lambda_2/4 \geq q$, i.e., $\lambda_2 \geq \overline{\lambda} := 2(2q - 1)$. In fact, the equilibrium requires $\lambda_2 > \overline{\lambda}$; if $\lambda_2 = \overline{\lambda}$, the DM would have no reason to click given a positive cost, see Appendix A.1.

Note that in this equilibrium the opportunistic type always selects $m_2 = 1$, a headline that runs against the DM's ex-ante optimal action $a_2 = 0$. Thus, the opportunistic type disseminates an article that runs counter to public opinion. She does so without incurring any information-acquisition cost, so the resulting article is entirely uninformative (the underlying signal is babbling). The good type sometimes also sends a headline that opposes the predisposition—namely when $\sigma_2 = 1$—which triggers the DM's scrutiny and induces him to click. He clicks on such a "suspicious" headline only when the realized content can affect his action; this requires high reputation, so that when $(m_2, \sigma_2) = (1, 1)$ the DM follows the content (chooses $a_2 = 1$). I refer to this period-2 equilibrium as the *clickbait equilibrium*.

Definition 2 (Clickbait equilibrium). Assume $q > 1/2$, so the DM's ex-ante optimal action is $a_2 = 0$. An informative period-2 equilibrium with the following strategies is said to be the *clickbait equilibrium*:

- **Good type:** truthful reporting, $m_2 = \sigma_2$.
- **Opportunistic type:** always sends $m_2 = 1$ (against the DM's ex-ante optimal action).
- **DM:** upon $m_2 = 0$, do not click and chooses $a_2 = 0$; upon $m_2 = 1$, clicks, observes σ_2, and chooses $a_2 = \sigma_2$.

To ensure the existence of the period-2 equilibrium in which the good type exerts positive effort, I henceforth assume $q < 5/8$:[3]

Assumption 1. $q < 5/8$.

Next proposition argues that an informative equilibrium in period 2 exists only if the reputation λ_2 is high enough. Moreover, when the reputation is high enough, there are only two types of informative equilibria in which the good type chooses headline truthfully—one of which is the clickbait equilibrium.

Proposition 2 (Informative equilibria in period 2). *Suppose that the game enters into the second period with the common posterior belief $\lambda_2 \in [0, 1]$. Define $\overline{\lambda} := 2(2q - 1)$.*

(i) *If $\lambda_2 < \overline{\lambda}$, no informative equilibrium exists in period 2.*

(ii) *Suppose $\lambda_2 \in (\overline{\lambda}, 1)$. There exists $\overline{c} > 0$ such that the clickbait equilibrium exists whenever $c \leq \overline{c}$. The equilibrium payoffs are:*

$$\begin{aligned} \textit{DM:} \quad & \frac{1}{2}\left(\frac{1}{2} + \frac{\lambda_2}{4} - q\right) - \left(1 - \frac{\lambda_2}{2}\right) c, \\ \textit{Good type:} \quad & \frac{5}{16} - \frac{q}{2} + k_{\mathrm{G}}, \\ \textit{Opportunistic type:} \quad & 1 + \frac{k_{\mathrm{O}}}{2}. \end{aligned}$$

(iii) *Suppose $\lambda_2 \in [\overline{\lambda}, 1)$. For all $c > 0$, the following constitutes an informative second-period equilibrium:*

[3] If $q \geq 5/8$, the good type never exerts positive effort in any equilibrium in period 2. When she exerts effort, her maximal action-related payoff is $1/2(1/2 + 1/4 - q) - 1/16$, attained at $\mu_{\mathrm{G}} = 1/4$. By contrast, choosing zero effort and inducing $a_2 = 0$ yields 0. Thus, $1/2 \cdot (1/2 + 1/4 - q) - 1/16 > 0 \Leftrightarrow 5/8 > q$ is necessary for an informative period-2 equilibrium.

- *Both types report headlines truthfully. The good type exerts* $\mu_G^* = 1/4$. *The DM never clicks any headlines and chooses an action according to the headline, i.e.,* $a_2 = m_2$.

The payoffs are:

$$\begin{aligned} \textit{DM:} \quad & \frac{1}{2}\left(\frac{1}{2}+\frac{\lambda_2}{4}-q\right), \\ \textit{Good type:} \quad & \frac{5}{16}-\frac{q}{2}+k_G, \\ \textit{Opportunistic type:} \quad & k_O. \end{aligned}$$

(iv) *Suppose* $\lambda_2 = 1$. *The only informative second-period equilibrium in which the good type reports headlines truthfully is:*

- *The good type exerts* $\mu_G^* = 1/4$ *and reports truthfully; the DM never clicks and chooses* $a_2 = m_2$.
- *The equilibrium payoffs coincides with those in (iii).*

(v) *Suppose* $\lambda_2 \in [\bar{\lambda}, 1)$. *For sufficiently small* $c > 0$, *the only informative equilibrium in which the good type reports truthfully are those described in (ii) and (iii).*

Proof. See Appendix A.1. □

Proposition 2 establishes that in period 2, an informative equilibrium exists only if the outlet's reputation is not too low, i.e., $\lambda_2 > \bar{\lambda} := 2(2q-1)$. In fact, when $\lambda_2 \leq \bar{\lambda}$, no informative equilibrium exists of any form—-not only those in which the good type reports truthfully.

When $\lambda_2 \in (\bar{\lambda}, 1)$, there are exactly two types of informative equilibria in which the good type reports truthfully, one of which is the clickbait equilibrium.

The other equilibrium has the DM never clicking and choosing the action that matches the headline. This equilibrium is sustained by the opportunistic type's headline choice, which is pinned down solely by the truth-telling benefit. However, this class of period-2 equilibria is not robust: its existence hinges on the *direction* of an infinitesimal headline bias. For example, if the opportunistic type instead received an arbitrarily small benefit from *anti*-truth-telling, the no-click equilibrium would fail to exist.[4] By contrast, the existence of the clickbait equilibrium does not depend on the direction of such a bias, provided that the benefit remains sufficiently small. I therefore rule out this non-robust no-click informative equilibrium from consideration.

[4] With a small anti-truth-telling benefit, the opportunistic type chooses $m_2 \neq \sigma_2$. Then upon observing $m_2 = 1$ the DM can no longer infer the content from the headline. By clicking on $m_2 = 1$ the DM observes σ_2 and can perfectly infer the outlet's type (since the good type is truthful while the opportunistic type is anti-truthful), which makes clicking strictly profitable. Consequently, the class of equilibria in which the DM never clicks but acts on the headline disappears.

3.2 Uninformative Equilibrium in Period 2

Suppose the DM believes that both types choose zero effort in period 2, so that signals are babbling. Then he always chooses $a_2 = 0$ since his posterior over the state coincides with the prior no matter which headlines or contents he might observe. This implies that exerting no effort is indeed rational for the good type and the good type chooses headlines truthfully in the uninformative equilibrium.

A click could at most reveal the outlet's type, but such information has no continuation value in this final period, and thereby the DM never clicks on any headline, implying that the opportunistic type chooses truthful headlines as well.

Proposition 3 (**Uninformative equilibrium in period 2**). *For any $\lambda_2 \in [0, 1]$, there uniquely exists an uninformative equilibrium in period 2. In the uninformative equilibrium for each λ_2,*

- *Both types chooses truthful headlines. The DM never clicks any of them and chooses $a_2 = 0$.*
- *Equilibrium payoffs are 0 for the DM, k_G for the good type, and k_O for the opportunistic type.*

Combining Proposition 2 and 3, the period-2 equilibrium for each $\lambda_2 \in [0, 1]$ is specified as follows:

- If $\lambda_2 \leq \overline{\lambda}$, an uninformative equilibrium is played (Proposition 3).
- If $\overline{\lambda} < \lambda_2 < 1$, the clickbait equilibrium is played (Proposition 2 (ii)).
- If $\lambda_2 = 1$, the informative equilibrium where the good type reports the headline truthfully is played (Proposition 2 (iv)).

Given this specification, the value functions of the DM, the good type, and the opportunistic type—denoted $V_{\text{DM}}, V_{\text{G}}$, and V_{O}, respectively—are:

$$V_{\text{DM}}(\lambda_2) = \begin{cases} \frac{1}{2}\left(\frac{3}{4} - q\right) & \text{if } \lambda_2 = 1, \\ \frac{1}{2}\left(\frac{1}{2} + \frac{\lambda_2}{4} - q\right) - \left(1 - \frac{\lambda_2}{2}\right)c & \text{if } \lambda_2 \in (\overline{\lambda}, 1), \\ 0 & \text{if } \lambda_2 \leq \overline{\lambda}, \end{cases}$$

$$V_{\text{G}}(\lambda_2) = \begin{cases} \frac{5}{16} - \frac{q}{2} + k_{\text{G}} & \text{if } \lambda_2 > \overline{\lambda}, \\ k_{\text{G}} & \text{if } \lambda_2 \leq \overline{\lambda}, \end{cases}$$

$$V_{\text{O}}(\lambda_2) = \begin{cases} k_{\text{O}} & \text{if } \lambda_2 = 1, \\ 1 + \frac{k_{\text{O}}}{2} & \text{if } \lambda_2 \in (\overline{\lambda}, 1), \\ k_{\text{O}} & \text{if } \lambda_2 \leq \overline{\lambda}. \end{cases}$$

The opportunistic type needs some reputation for being good to get clicked, but if she is thought to be perfectly good, clicking does not occur.

4 Equilibrium of the Entire Game

From now on, I examine the equilibrium of the entire game, including the first period. Hereafter, the equilibrium refers to the equilibrium of the entire game. An equilibrium is *informative* if at least one type exerts a positive effort in the first period. Otherwise, it is referred to as *uninformative*.[5]

I compare the equilibria of two environments—one with fact-checking, and one without. To conduct a meaningful comparison, I restrict attention to equilibria that employ identical headline strategies in both settings. The second-period headline strategies, already specified at the end of Section 3, are as follows:

- **Uninformative equilibrium in period 2** ($\lambda_2 \leq \overline{\lambda}$): both types of the outlet report the headlines truthfully.
- **Informative equilibrium in period 2** ($1 \geq \lambda_2 > \overline{\lambda}$): the good type reports the headlines truthfully, whereas the opportunistic type always sets $m_2 = 1$.

The first-period headline strategies remain to be specified. For both settings, I restrict attention to the class of informative equilibria of the entire game where the good type reports

[5]Note that even in an uninformative equilibrium, the outlet may exert positive effort in the second period, since the clickbait equilibrium will be played when the game proceeds to the second period with $\overline{\lambda} < \lambda_2 < 1$. However, within the class of equilibria that this paper focuses on, this will not happen.

the headlines truthfully, while the opportunistic type chooses $m_1 = 1$. For the uninformative equilibria, I focus in both settings on the one where both types report the headlines truthfully in period 1. Thus, the first-period headline strategies are fully specified for both the informative and uninformative equilibria, in each of the two environments—with and without fact-checking.

I denote the pair of effort levels in period 1 as μ_G and μ_O, instead of $\mu_G^{t=1}$ and $\mu_O^{t=1}$, since period-2 equilibria have been specified.

4.1 Equilibria of the Entire Game without Fact-Checking

Let $o_1 \in \mathcal{M} \cup (\mathcal{M} \times C)$ denote the outcome in period 1, where $o_1 = m_1 \in \mathcal{M}$ corresponds to the situation in which the DM observes headline m_1 but does not click on it, while $o_1 = (m_1, \sigma_1) \in \mathcal{M} \times C$ is the situation where the DM clicks on m_1 and observes the content σ_1.

When there is no fact-checking, the reputation of the media outlet at the beginning of period 2 after the outcome o_1, $\lambda_2(o_1)$, is

$$\lambda_2(o_1) = \frac{\lambda_1 \Pr(o_1 \mid \mathrm{G})}{\lambda_1 \Pr(o_1 \mid \mathrm{G}) + (1 - \lambda_1) \Pr(o_1 \mid \mathrm{O})},$$

where the probability that o_1 comes from the good outlet, denoted by $\Pr(o_1 \mid \mathrm{G})$, is given by

$$\Pr(o_1 \mid \mathrm{G}) = \begin{cases} \dfrac{1}{2} \Pr(m_1 \mid \mathrm{G}, \sigma_1) & \text{if } o_1 = (m_1, \sigma_1), \\ \dfrac{1}{2} \displaystyle\sum_{\sigma_1'} \Pr(m_1 \mid \mathrm{G}, \sigma_1') & \text{if } o_1 = m_1. \end{cases}$$

Note that $\Pr(m_1 \mid \mathrm{G}, \sigma_1)$ is the headline strategy for the good type with σ_1, and $1/2$ is the probability that the good type receives each signal. One can rewrite the probability that o_1 comes from the opportunistic outlet, $\Pr(o_1 \mid \mathrm{O})$, analogously. Therefore, in the absence of fact-checking, the mapping $o_1 \mapsto \lambda_2(o_1)$ is determined by the headline strategies as well as the DM's strategy, and $\lambda_2(o_1)$ does not depend on the equilibrium effort level.

This observation, together with the fact that the probability distribution over the first-period outcome o_1 is independent of effort, implies the following result: in the absence of fact-checking, the opportunistic type never exerts any effort in period 1. This conclusion holds across all equilibria, no matter how the period-2 equilibrium is specified or which headline strategy is used in period 1.

Proposition 4. *Suppose there is no fact-checking. For any $\lambda_1 \in (0, 1)$, opportunistic type exerts no effort in period 1 in any equilibrium of the entire game.*

Proof. Suppose by contradiction that in the equilibrium of the entire game, the opportunistic type exerts effort in period 1. I argue that deviation to zero effort is profitable for her. In period 1, whether or not being clicked on is independent of the effort, as it is unobservable to the DM. Thus, it suffices to show that her period-2 payoff $V_O(\lambda_2)$ is independent of her effort as well. Since λ_2 is solely determined by the equilibrium outcome o_1, it suffices to show that the probability distribution over the outcomes is independent of her effort, which is true due to the fact that the probability that the opportunistic type receives each signal is independent of her effort. Consequently, the opportunistic type can reduce her cost of effort without affecting payoffs in both periods. □

This behavior of the opportunistic type is in sharp contrast to her behavior under fact-checking, where she may exert positive effort in the first period.

4.1.1 Uninformative Equilibrium in the Absence of Fact-Checking

Consider an uninformative equilibrium in which both types report headlines truthfully. Since neither type exerts effort in period 1, no information about the state ω_1 is generated, and the DM optimally chooses $a_1 = 0$. Moreover, because $m_1 = \sigma_1$ under truth-telling, the headline already reveals the content, so clicking yields no additional information about the outlet's type; the DM therefore does not click. Given these best responses, truth-telling and zero effort are optimal for both types.

Proposition 5 (Uninformative Equilibria in the Absence of Fact-Checking). *Suppose there is no fact-checking. For any $\lambda_1 \in (0, 1)$, there exists an uninformative equilibrium in which both types of the outlet choose the headline truthfully in period 1. In this equilibrium, the DM does not click any headline in period 1 and always chooses $a_1 = 0$.*

Proof. See Appendix A.2. □

4.1.2 Informative Equilibrium in the Absence of Fact-Checking

Now I investigate an informative period-1 equilibrium in which the good type reports truthfully and the opportunistic type always chooses the headline $m_1 = 1$ in period 1.

Theorem 1 (Informative Equilibrium in the Absence of Fact-Checking). *Suppose that fact-checking is absent. Consider the class of informative equilibria in which the good type reports truthfully and the opportunistic type always chooses $m_1 = 1$ in period 1.*

(i) *If $\lambda_1 \le \overline{\lambda}$, there is no equilibrium in this class.*

(ii) *If $\lambda_1 > \overline{\lambda}$, then for sufficiently small $c > 0$, There exists a unique equilibrium in this class. In this equilibrium, the good type's effort is $\mu_G^* = 1/4$; the DM clicks on $m_1 = 1$ and, upon observing σ_1, chooses $a_1 = \sigma_1$; the DM does not click $m_1 = 0$ and chooses $a_1 = 0$.*

Proof. See Appendix A.3. □

Combining Proposition 5 with Theorem 1 yields the following specification of the no-fact-checking equilibrium for each $\lambda_1 \in (0, 1)$:

- $\lambda_1 \leq \overline{\lambda}$: Uninformative equilibrium in which both types choose headline truthfully.
- $\lambda_1 > \overline{\lambda}$: Informative equilibrium in which the good type reports truthfully, the opportunistic type sends $m_1 = 1$, and $(\mu_G^*, \mu_O^*) = (1/4, 0)$.

4.2 Equilibrium of the Entire Game with Fact-Checking

I now turn to the model with fact-checking. Under fact-checking, the true state in period 1, ω_1, is revealed at the end of the period with probability $\beta \in (0, 1]$. Recall that, without fact-checking, the opportunistic type exerts no effort: from her perspective, the probability distribution over the period-1 outcomes $o_1 \in \mathcal{M} \cup (\mathcal{M} \times C)$ is invariant to her effort choice. By contrast, with fact-checking, the reputation carried into period 2, $\lambda_2(o_1, \gamma)$, depends on both the period-1 outcome o_1 and the fact-checking result γ, and the probability distribution of γ is itself dependent on her effort choice. This creates an additional incentive for both types to exert costly effort.

4.2.1 Informative Equilibria in the Presence of Fact-Checking

I construct an informative equilibrium in which the good type reports headlines truthfully, while the opportunistic type always sets $m_1 = 1$.

Lemma 1 specifies the DM's equilibrium strategy.

Lemma 1 (DM's equilibrium strategy). *In this equilibrium, the DM's strategy is*

$$\begin{cases} \textit{if } m_1 = 1: & \textit{click and choose } a_1 = \sigma_1, \\ \textit{if } m_1 = 0: & \textit{not click and choose } a_1 = 0. \end{cases}$$

Proof. See Appendix A.4. □

Opportunistic Type's Effort. Consider the opportunistic type. Let (μ_G^*, μ_O^*) denote the effort levels the DM believes to be chosen in the equilibrium. The opportunistic type chooses μ_O so as to maximize her continuation payoff evaluated at the time effort is chosen.

The continuation payoff from choosing μ_O can be derived as follows. The first-period payoff is straightforward: when she observes $\sigma_1 = 1$ (with probability $1/2$), she reports truthfully, the headline is clicked, and she earns $1 + k_O$. If instead she observes $\sigma_1 = 0$, she sends an untruthful headline that will be clicked, earning 1. Thus, the first-period payoff is $1 + k_O/2$.

She also anticipates the period-2 payoff V_G. Suppose she observes $\sigma_1 = 1$. Fact-checking fails (i.e., $\gamma = \emptyset$) with probability $1 - \beta$, leaving the posterior probability that the outlet is good type as $\lambda_2(m_1 = 1, \sigma_1 = 1, \gamma = \emptyset) = \lambda_1$. Fact-checking confirms the news article (i.e., $\gamma = 1$) with probability

$$\Pr(\gamma = 1 \mid \sigma_1 = 1) = \beta \cdot \Pr(\omega_1 = 1 \mid \sigma_1 = 1) = \beta\left(\frac{1}{2} + \mu_{\mathrm{O}}\right),$$

in which case the DM—having observed the article $(m_1, \sigma_1) = (1, 1)$—shifts his belief to

$$\lambda_2(1,1,1) = \frac{\lambda_1\left(\frac{1}{2} + \mu^*_{\mathrm{G}}\right)}{\lambda_1\left(\frac{1}{2} + \mu^*_{\mathrm{G}}\right) + (1-\lambda_1)\left(\frac{1}{2} + \mu^*_{\mathrm{O}}\right)}.$$

Disconfirmation (i.e., $\gamma = 0$) occurs with probability $\beta(1/2 - \mu_{\mathrm{O}})$, yielding

$$\lambda_2(1,1,0) = \frac{\lambda_1\left(\frac{1}{2} - \mu^*_{\mathrm{G}}\right)}{\lambda_1\left(\frac{1}{2} - \mu^*_{\mathrm{G}}\right) + (1-\lambda_1)\left(\frac{1}{2} - \mu^*_{\mathrm{O}}\right)}.$$

Note that $\lambda_2(1,1,1)$ (resp. $\lambda_2(1,1,0)$) corresponds to the probability that the outlet observing the correct (resp. incorrect) evidence is of good type. Therefore, if the DM believes $\mu^*_{\mathrm{G}} > \mu^*_{\mathrm{O}}$, confirmation (resp. disconfirmation) by fact-checking raises (resp. lowers) the DM's posterior belief that the outlet is of good type.

If instead she receives $\sigma_1 = 0$, she sends an untruthful message that will be clicked, and thus the DM learns that the outlet is opportunistic; hence, $\lambda_2(1, 0, \gamma) = 0$ for all γ.

Her continuation payoff is therefore

$$\frac{1}{2}\Bigg[\overbrace{(1-\beta)V_{\mathrm{O}}\big(\lambda_2(1,1,\emptyset)\big) + \beta\left(\tfrac{1}{2} - \mu_{\mathrm{O}}\right)V_{\mathrm{O}}\big(\lambda_2(1,1,0)\big) + \beta\left(\tfrac{1}{2} + \mu_{\mathrm{O}}\right)V_{\mathrm{O}}\big(\lambda_2(1,1,1)\big)}^{\text{period-2 payoff when } \sigma_1=1}\Bigg]$$
$$+ \frac{1}{2}\underbrace{V_{\mathrm{O}}\big(\lambda_2 = 0\big)}_{\text{period-2 payoff when } \sigma_1=0} - (\mu_{\mathrm{O}})^2 + \underbrace{1 + \tfrac{k_{\mathrm{O}}}{2}}_{\text{period-1 payoff}}.$$

Here λ_2 (and hence $V_{\mathrm{O}}(\lambda_2)$) depends on the effort levels the DM *expects* in equilibrium, $(\mu^*_{\mathrm{G}}, \mu^*_{\mathrm{O}})$, not on μ_{O}. Dropping terms independent of μ_{O}, her problem reduces to

$$\max_{\mu_{\mathrm{O}}}\left[-(\mu_{\mathrm{O}})^2 + \frac{\beta\mu_{\mathrm{O}}}{2}\Big\{V_{\mathrm{O}}\big(\lambda_2(1,1,1)\big) - V_{\mathrm{O}}\big(\lambda_2(1,1,0)\big)\Big\}\right],$$

with first-order condition

$$\mu^*_{\mathrm{O}} = \frac{\beta}{4}\Big[V_{\mathrm{O}}\big(\lambda_2(1,1,1)\big) - V_{\mathrm{O}}\big(\lambda_2(1,1,0)\big)\Big].$$

Hence, the opportunistic type exerts effort only if

$$\overbrace{\frac{\lambda_1\left(\frac{1}{2}+\mu_G^*\right)}{\lambda_1\left(\frac{1}{2}+\mu_G^*\right)+(1-\lambda_1)\left(\frac{1}{2}+\mu_O^*\right)}}^{\lambda_2(1,1,1)} > \overline{\lambda} \geq \overbrace{\frac{\lambda_1\left(\frac{1}{2}-\mu_G^*\right)}{\lambda_1\left(\frac{1}{2}-\mu_G^*\right)+(1-\lambda_1)\left(\frac{1}{2}-\mu_O^*\right)}}^{\lambda_2(1,1,0)}. \tag{2}$$

Using the explicit form of V_O, $\mu_O^* > 0$ is given by

$$\mu_O^* = \frac{\beta\delta_O}{4}, \tag{3}$$

where $\delta_O := V_O(\lambda_2(1,1,1)) - V_O(\lambda_2(1,1,0)) = 1 - k_O/2$.

Equation (2) means the following: if fact-checking confirms the articles ($\gamma = 1$), the posterior reputation exceeds the threshold $\overline{\lambda}$—the DM trusts the outlet enough to click the *suspicious* headline in period 2—and the clickbait equilibrium obtains; if it disconfirms ($\gamma = 0$), the posterior reputation stays at or below $\overline{\lambda}$ and uninformative equilibrium obtains. Therefore, under equation (2), the reputational concern endogenously arises, creating an incentive to exert effort.

If, instead, equation (2) fails—so the period-2 equilibrium is independent of the fact-checking outcome (e.g., clickbait obtains regardless of confirmation/disconfirmation)—the opportunistic outlet has no incentive to exert effort: if reputation would exceed the threshold anyway, shirking is optimal; if it would never reach the threshold, effort is futile.

I say that fact-checking is *diagnostic* about the outlet's motive if equation (2) holds. In that case, the DM's second-period behavior depends on whether the article is confirmed or disconfirmed by fact-checking.

Good Type's Effort. Given that equation (2) is true, the driving forces for the good type to exert effort are now two-fold: the action-related payoff in period 1 and the second-period payoff. Hence, the good type chooses μ_G so as to maximize

$$\max_{\mu_G} \left[-(\mu_G)^2 + \frac{\mu_G}{2}\left\{1 + \beta V_G\big(\lambda_2(1,1,1)\big) - \beta V_G\big(\lambda_2(1,1,0)\big)\right\}\right],$$

yielding

$$\mu_G^* = \frac{1+\beta\delta_G}{4}, \tag{4}$$

where $\delta_G := V_G(\lambda_2(1,1,1)) - V_G(\lambda_2(1,1,0)) = 5/16 - q/2 > 0$.

Consistency. Finally, in any informative equilibrium where both types exert positive effort, equation (2) must hold when (μ_G^*, μ_O^*) are given by equations (4) and (3).

Theorem 2 (**Informative Equilibrium in the Presence of Fact-Checking**). *Suppose fact-checking is present. Consider the class of informative equilibria in which the good type reports headlines truthfully and the opportunistic type always sends $m_1 = 1$ in period 1. Then:*

(i) *If $2\overline{\lambda}/(1+\overline{\lambda}) < \lambda_1$, the unique equilibrium in this class has $(\mu^*_{\mathrm{G}}, \mu^*_{\mathrm{O}}) = (1/4, 0)$.*

(ii) *If $\max\{\underline{\lambda}^*(\beta), \underline{\lambda}^{**}(\beta)\} < \lambda_1 \leq \overline{\lambda}^*(\beta)$, the unique equilibrium in this class has*

$$(\mu^*_{\mathrm{G}}, \mu^*_{\mathrm{O}}) = \left(\frac{1+\beta\delta_{\mathrm{G}}}{4}, \frac{\beta\delta_{\mathrm{O}}}{4}\right),$$

where $\underline{\lambda}^*(\beta)$, $\overline{\lambda}^*(\beta)$, $\underline{\lambda}^{**}(\beta)$ are given by equations (A.9), (A.10), *and* (A.12) *in Appendix A.5, respectively.*

(iii) *Otherwise, no informative equilibrium in this class exists.*

Proof. See Appendix A.5. □

4.2.2 Uninformative Equilibrium in the Presence of Fact-Checking

Focus on uninformative equilibria in which both types report headlines truthfully. Since signals are babbling, no information about the period-1 state is conveyed and the DM always chooses $a_1 = 0$. Moreover, since truthful headlines reveal the content without clicking, the DM does not click any headline.

Proposition 6 (**Uninformative equilibrium in the presence of fact-checking**). *Suppose fact-checking is present. Consider the class of uninformative equilibria in which both types report headlines truthfully. For any $\lambda_1 \in (0,1)$, there exists a unique equilibrium in this class in which the DM never clicks any headline and always chooses $a_1 = 0$ in period 1.*

Combining Theorem 2 and Proposition 6 yields the following equilibrium selection in the presence of fact-checking: the equilibrium will be

$$\begin{cases} \text{Informative one with } \mu^*_{\mathrm{G}} = \dfrac{1}{4}, \mu^*_{\mathrm{O}} = 0 & \text{if } \dfrac{2\overline{\lambda}}{1+\overline{\lambda}} < \lambda_1, \\ \text{Informative one with } \mu^*_{\mathrm{G}} = \dfrac{1+\beta\delta_{\mathrm{G}}}{4},\ \mu^*_{\mathrm{O}} = \dfrac{\beta\delta_{\mathrm{O}}}{4} & \text{if} \max\left\{\underline{\lambda}^*(\beta),\ \underline{\lambda}^{**}(\beta)\right\} < \lambda_1 \leq \overline{\lambda}^*(\beta), \\ \text{Uninformative one} & \text{otherwise.} \end{cases}$$

Figure 1 depicts the parameter regions for each equilibrium. The gray-shaded region marks parameter values under which an informative equilibrium in which both types exert effort arises; the dotted region marks those under which only the good type exerts effort; the remaining region corresponds to the uninformative equilibrium. Note that $\underline{\lambda}^{**}(\beta) \to \overline{\lambda}$ as $\beta \to 0$, $\underline{\lambda}^*(\beta), \overline{\lambda}^*(\beta) \to \overline{\lambda}$ as $\beta \to 1$, and $\overline{\lambda}^*(\beta) \to 2\overline{\lambda}/(1+\lambda)$ as $\beta \to 0$ (see Appendix A.5 for more detail).

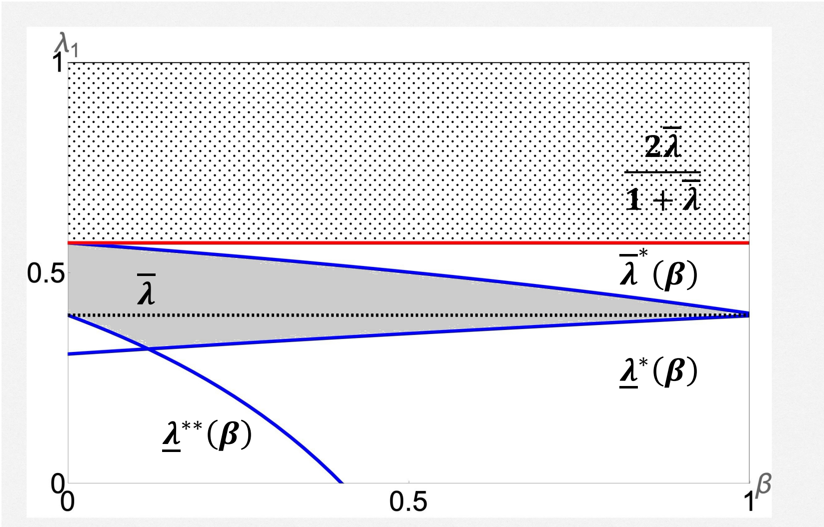


Figure 1: Equilibria with Fact-checking. $q = 0.6, k_O = 0.01$

4.2.3 Comparative Statics with respect to Probability of Fact-Checking being Successful

This subsection studies how the success probability of fact-checking, β, affects the equilibrium under fact-checking. Theorem 3 formalizes the trade-off between the diagnostic and the discipline effects.

Theorem 3. *Suppose fact-checking is available and*

$$\lambda_1 \in \left(\overline{\lambda},\ \frac{2\overline{\lambda}}{1+\overline{\lambda}}\right).$$

Then there exists $\overline{\beta}$ such that the following holds:

(i) *The equilibrium is*

$$\begin{cases} \textit{Informative with } \mu_G^* = \dfrac{1+\beta\delta_G}{4},\ \mu_O^* = \dfrac{\beta\delta_O}{4}, & \textit{if } \beta \le \overline{\beta}, \\ \textit{Uninformative}, & \textit{if } \beta > \overline{\beta}. \end{cases}$$

(ii) *The DM's welfare is increasing in β on $(0, \overline{\beta})$.*

(iii) *For any β', β'' with $\beta' < \overline{\beta} < \beta''$, the DM's welfare under β' is higher than under β''.*

Proof. See Appendix A.6 □

As argued in Section 4.2.1, the opportunistic type exerts effort only if equation (2) holds. Equation (2) means fact-checking is diagnostic in the following sense: if it confirms the article, the DM's posterior that the outlet is good exceeds the threshold, so the period-2 outcome is the clickbait equilibrium—the DM clicks the suspicious headline and follows the outlet's recommendation; if it disconfirms, the posterior falls at or below the threshold, so the period-2 equilibrium is uninformative—no headline is clicked and no recommendation is followed.

Because this *diagnostic* effect creates reputational incentives for both types, even the opportunistic outlet voluntarily exerts effort and raises the likelihood of obtaining the correct evidence, which is captured by $\beta\delta_G/4$ and $\beta\delta_O/4$. These terms are increasing in β—the *discipline* effect—since the marginal return to effort rises with β, the probability that fact-checking reveals whether the article was correct or not. As long as fact-checking remains diagnostic, the discipline effect makes the DM's welfare increase in β, as stated in Theorem 3 (ii).

However, the discipline effect *attenuates* the diagnostic effect. Intuition is as follows: as β increases, both μ_G^* and μ_O^* rise, but the gap $\mu_G^* - \mu_O^*$ shrinks, implying that the likelihoods of obtaining correct/incorrect evidence become more similar across types. Consequently, the posteriors after confirmation and disconfirmation—$\lambda_2(1,1,1)$ and $\lambda_2(1,1,0)$—become closer, so fact-checking conveys less information about type. For sufficiently large β, the diagnostic effect vanishes. Consequently, the opportunistic type stops exerting effort, the equilibrium reverts to the uninformative one, where even the good type exerts no effort, and the DM is worse off, as stated in Theorem 3 (i) and (iii).

5 Comparison of the Presence and Absence of Fact-Checking

Section 4.2.3 provides comparative statics with respect to β, holding λ_1 fixed and assuming the presence of fact-checking. This section conducts comparative statics across institutional settings, fixing both λ_1 and β; namely, it compares the environments with and without fact-checking.

The equilibrium specifications for the two environments are summarized below.

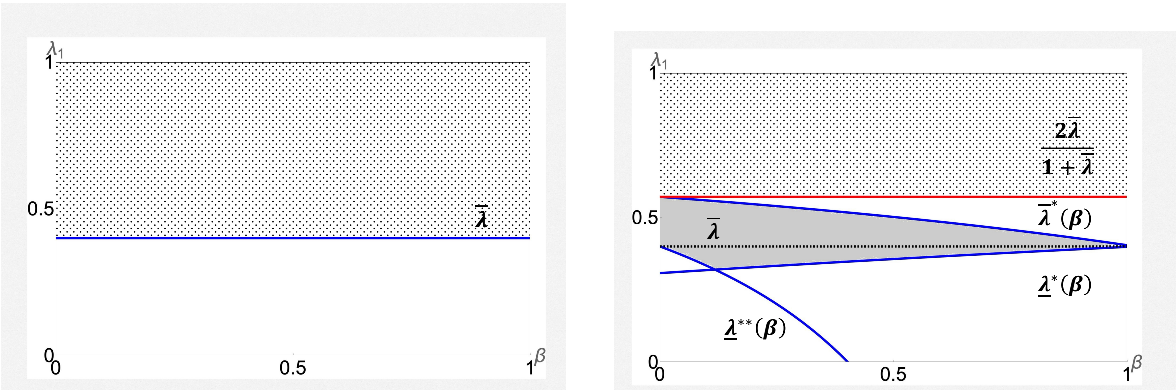
Without fact-checking:

$$\begin{cases} \text{Informative with } \mu_G^* = \frac{1}{4}, \mu_O^* = 0 & \text{if } \overline{\lambda} < \lambda_1, \\ \text{Uninformative} & \text{otherwise.} \end{cases}$$

With fact-checking:

$$\begin{cases} \text{Informative with } \mu_G^* = \frac{1}{4}, \mu_O^* = 0 & \text{if } \frac{2\overline{\lambda}}{1+\overline{\lambda}} < \lambda_1, \\ \text{Informative with } \mu_G^* = \frac{1+\beta\delta_G}{4},\ \mu_O^* = \frac{\beta\delta_O}{4} & \text{if} \max\{\underline{\lambda}^*(\beta), \underline{\lambda}^{**}(\beta)\} < \lambda_1 \leq \overline{\lambda}^*(\beta), \\ \text{Uninformative} & \text{otherwise.} \end{cases}$$

The left panel of Figure 2 plots the threshold $\overline{\lambda}$ in the model without fact-checking in the (β, λ_1) plane. Note that $\overline{\lambda}$ is independent of β. The right panel plots the thresholds $\underline{\lambda}^{**}(\beta), \underline{\lambda}^{*}(\beta), \overline{\lambda}^{*}(\beta), 2\overline{\lambda}/(1+\overline{\lambda})$ which jointly delineate the equilibrium regions when fact-checking is present. In both panels, the dotted area represents the parameter region where the informative equilibrium in which only the good type exerts effort is played. The gray-shaded area in the right panel represents those where the informative equilibrium in which both types exert effort is obtained.



(a) Without Fact-checking

(b) With Fact-checking

Figure 2: Region of parameters for each equilibrium. $q = 0.6, k_{\mathrm{O}} = 0.01$

Combining these panels, I obtain

(a) Beneficial: Suppose $\max\{\underline{\lambda}^{*}(\beta), \underline{\lambda}^{**}(\beta)\} < \lambda_1 \leq \overline{\lambda}^{*}(\beta)$ (dotted area of the top left panel in Figure 3), so λ_1 is intermediate and β is not too high. The equilibria for each setting are as follows.

Without fact-checking Informative equilibrium with $\mu_{\mathrm{G}}^{*} = 1/4, \mu_{\mathrm{O}}^{*} = 0$ or Uninformative equilibrium.

With Fact-checking Informative equilibrium with $\mu_{\mathrm{G}}^{*} = (1 + \beta\delta_{\mathrm{G}})/4, \mu_{\mathrm{O}}^{*} = \beta\delta_{\mathrm{O}}/4$.

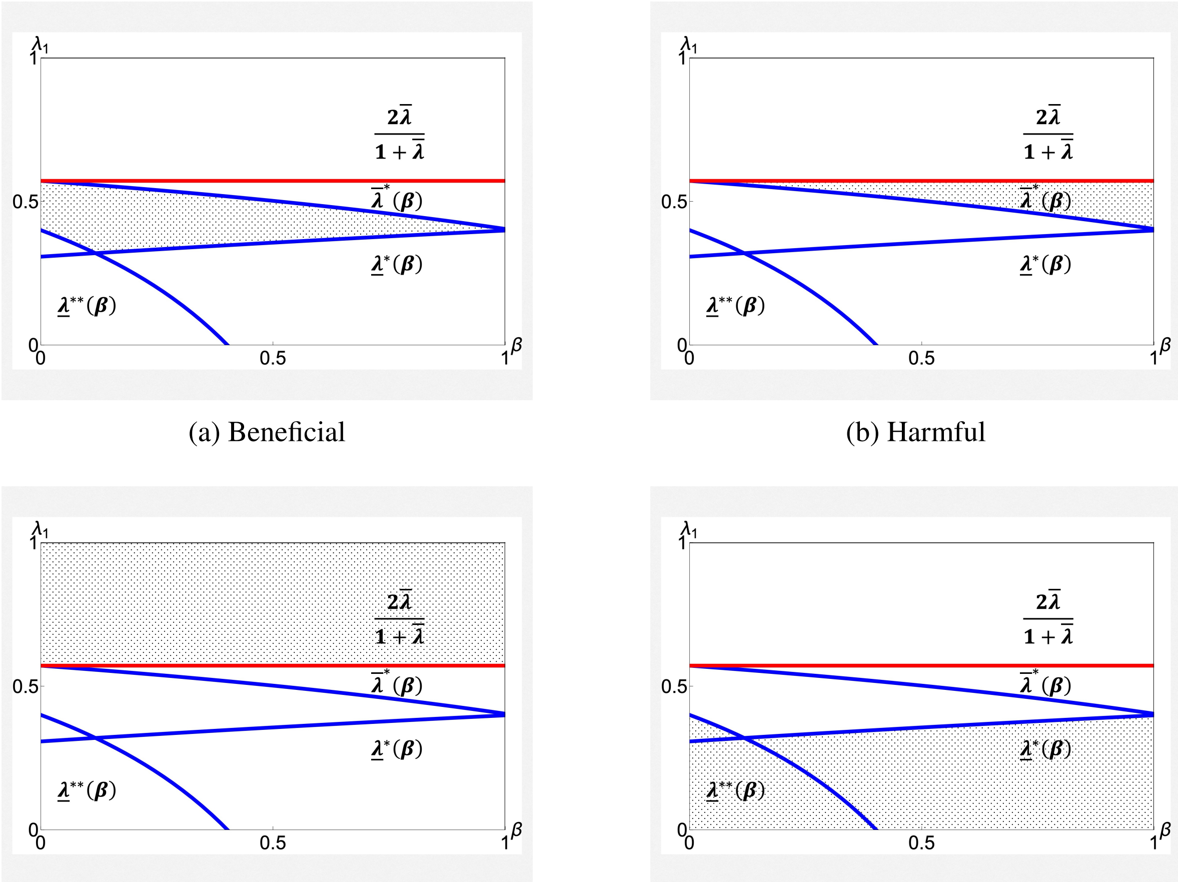
Comparison Introducing fact-checking *increases* the effort levels of both types.

(b) Harmful: Suppose $\overline{\lambda}^{*}(\beta) \leq \lambda_1 < 2\overline{\lambda}/(1+\overline{\lambda})$ (dotted area of the top right panel in Figure 3), so λ_1 is intermediate but β is high enough. The equilibria for each setting are as follows.

Without fact-checking Informative equilibrium with $\mu_{\mathrm{G}}^{*} = 1/4, \mu_{\mathrm{O}}^{*} = 0$.

With Fact-checking Uninformative equilibrium.

Comparison Introducing fact-checking *decreases* the effort level of good type.



(a) Beneficial

(b) Harmful

(c) No effect under informative equilibrium

(d) No effect uninformative equilibrium

Figure 3: Comparison between with and without Fact-checking

(c) No Effect under Informative Equilibrium: Suppose $\lambda_1 \geq 2\bar{\lambda}/\left(1+\bar{\lambda}\right)$ (dotted area of the bottom left panel in Figure 3), so λ_1 is high enough. The equilibria for each setting are as follows.

Without fact-checking Informative equilibrium with $\mu_G^* = 1/4, \mu_O^* = 0$.

With Fact-checking Informative equilibrium with $\mu_G^* = 1/4, \mu_O^* = 0$.

Comparison Introducing fact-checking does not affect the effort level of both types.

(d) No Effect under Uninformative Equilibrium: Otherwise, the equilibria for each setting are as follows. Dotted area of bottom right panel in Figure 3 represents this condition. Namely, λ_1 is low enough.

Without fact-checking Uninformative equilibrium.

With Fact-checking Uninformative equilibrium.

Comparison Introducing fact-checking does not affect the effort level of both types.

Theorem 4 compares the DM's welfare with and without fact-checking in each case.

Theorem 4 (**Comparison Between with and without Fact-Checking**).

(1) *Suppose* $\max\{\underline{\lambda}^*(\beta), \underline{\lambda}^{**}(\beta)\} < \lambda_1 \le \overline{\lambda}^*(\beta)$*, so that prior reputation* λ_1 *is intermediate and the probability of the fact-checking being successful* β *is not too high. Fact-checking increases the effort levels of both types. Consequently, fact-checking increases the DM's equilibrium welfare.*

(2) *Suppose* $\overline{\lambda}^*(\beta) \le \lambda_1 < 2\overline{\lambda}/(1+\overline{\lambda})$*, so that* λ_1 *is intermediate but* β *is high enough. Fact-checking reduces the effort level of the good type. Consequently, fact-checking reduces the DM's welfare.*

(3) *Suppose parameters* (β, λ_1) *satisfies neither of the conditions in* **(1)** *nor* **(2)***, so that* λ_1 *is sufficiently high or sufficiently low. Fact-checking affects effort levels of neither types. Consequently, fact-checking does not affect DM's welfare.*

Proof. See Appendix A.7 □

Theorem 4 establishes that the presence of fact-checking affects the media outlet or the DM's welfare only if the prior reputation of the media outlet λ_1 is intermediate.

The intuition is as follows: Because the second-period payoffs for both types of the outlet are step functions of the posterior reputation λ_2 with threshold $\overline{\lambda}$, reputational concerns arise only if confirmation by fact-checking pushes λ_2 above the threshold while disconfirmation leaves it below. For this to happen, the outlet's prior reputation must not be extreme. If the prior reputation is already high enough (or low enough), fact-checking induces only small belief revision, whereas posterior reputation is most responsive at intermediate priors. Thus, fact-checking affects nothing if the prior reputation is sufficiently high or low.

By contrast, when prior reputation is intermediate, fact-checking matters, but its effect is mixed. Suppose the prior reputation λ_1 is intermediate and the probability of fact-checking being successful β is not too high—for instance, (β, λ_1) lies in the dotted area in the top left panel in Figure 3. In this case, the effort levels of both types are higher in the presence than in the absence of fact-checking, thereby improving the DM's welfare. Notably, even the opportunistic type who cares only about attracting clicks exerts costly effort.

If instead λ_1 is intermediate and β is sufficiently high—for instance (β, λ_1) lies in the dotted area in the top right panel in Figure 3—then, although the informative equilibrium would arise without fact-checking, the presence of fact-checking shifts the game to the uninformative equilibrium, reducing the DM's welfare.

6 Concluding Remarks

This paper studies how fact-checking affects a media outlet's information acquisition in a two-period strategic communication game with a decision maker. Fact-checking reveals the state of the world in the first period with positive probability.

Without fact-checking, the opportunistic media outlet never exerts effort in either period, while the good type exerts costly effort when the prior reputation towards the media industry is not too low. When the prior is either sufficiently high or sufficiently low, fact-checking has no effect on information acquisition or on the DM's welfare. For intermediate prior, fact-checking plays a decisive role. Its effect, however, is mixed: it improves information acquisition and the decision maker's welfare when the probability of successful revelation is not too high, but reduces them when the probability is sufficiently high.

This paper focused on clickbait. The effect of fact-checking on information acquisition by outlets with other objectives, such as ideology, is left for future research.

A Omitted Proofs

A.1 Proof of Proposition 2

I characterise second-period informative equilibria ; accordingly, statements *(i)-(v)* are established jointly. In any informative equilibrium with $\mu_{\mathrm{G}}^* > 0$ and $\mu_{\mathrm{O}}^* = 0$, the DM must choose $a_2 = 0$ upon observing $\sigma_2 = 0$ and—conditional on clicking a headline—must choose $a_2 = 1$ when the revealed content is $\sigma_2 = 1$. Moreover, the DM's equilibrium strategy must satisfy that, for each realization of the good type's signal $\sigma_2 \in \{0, 1\}$, at least one headline induces the DM to choose action $a_2 = \sigma_2$.

This requirement yields $3 \times 3 = 9$ possible configurations. Specifically, for each σ_2, one of the following holds: (1) both headlines implement $a_2 = \sigma_2$; (2) only $m_2 = 1$ implements $a_2 = \sigma_2$; or (3) only $m_2 = 0$ implements $a_2 = \sigma_2$. The table below summarizes these cases.

Good $\sigma_2 = 0$ / Good $\sigma_2 = 1$	$\forall m_2 \Rightarrow a_2 = 0$	$\begin{pmatrix} m_2 = 1 \Rightarrow a_2 = 1 \\ m_2 = 0 \Rightarrow a_2 = 0 \end{pmatrix}$	$\begin{pmatrix} m_2 = 1 \Rightarrow a_2 = 0 \\ m_2 = 0 \Rightarrow a_2 = 1 \end{pmatrix}$
$\forall m_2 \Rightarrow a_2 = 1$	Case 1	Case 3	Case 5
$\begin{pmatrix} m_2 = 1 \Rightarrow a_2 = 1 \\ m_2 = 0 \Rightarrow a_2 = 0 \end{pmatrix}$	Case 2	Case 6	Case 9
$\begin{pmatrix} m_2 = 1 \Rightarrow a_2 = 0 \\ m_2 = 0 \Rightarrow a_2 = 1 \end{pmatrix}$	Case 4	Case 8	Case 7

This matrix pins down the DM's equilibrium strategy. For instance, in Case 2 (the clickbait equilibrium), the DM must click on $m_2 = 1$ and, upon observing σ_2, choose action $a_2 = \sigma_2$. Otherwise a contradiction arises: if, for instance, the DM did not click on $m_2 = 1$ and nevertheless chose $a_2 = 1$, then from the perspective of the good type with $\sigma_2 = 0$—who may publish $m_2 = 1$—the DM would be required to choose $a_2 = 0$ after $m_2 = 1$, which is incompatible with the former behavior.

The below specifies the DM's equilibrium strategy for each of the nine cases.

- Case 1: DM's strategy

$$\begin{cases} \text{if } m_2 = 1: & \text{click, and choose } a_2 = \sigma_2 \text{ on observing } \sigma_2. \\ \text{if } m_2 = 0: & \text{click, and choose } a_2 = \sigma_2 \text{ on observing } \sigma_2. \end{cases}$$

- Case 2: DM's strategy

$$\begin{cases} \text{if } m_2 = 1: & \text{click, and choose } a_2 = \sigma_2 \text{ on observing } \sigma_2. \\ \text{if } m_2 = 0: & \text{not click, and choose } a_2 = 0. \end{cases}$$

- Case 3: DM's strategy

$$\begin{cases} \text{if } m_2 = 1: & \text{not click, and choose } a_2 = 1. \\ \text{if } m_2 = 0: & \text{click, and choose } a_2 = \sigma_2 \text{ on observing } \sigma_2. \end{cases}$$

- Case 4: DM's strategy

$$\begin{cases} \text{if } m_2 = 1: & \text{not click, and choose } a_2 = 0. \\ \text{if } m_2 = 0: & \text{click, and choose } a_2 = \sigma_2 \text{ on observing } \sigma_2. \end{cases}$$

- Case 5: DM's strategy

$$\begin{cases} \text{if } m_2 = 1: & \text{click, and choose } a_2 = \sigma_2 \text{ on observing } \sigma_2. \\ \text{if } m_2 = 0: & \text{not click, and choose } a_2 = 1. \end{cases}$$

- Case 6: DM's strategy

$$\begin{cases} \text{if } m_2 = 1: & \text{not click, and choose } a_2 = 1. \\ \text{if } m_2 = 0: & \text{not click, and choose } a_2 = 0. \end{cases}$$

- Case 7: DM's strategy

$$\begin{cases} \text{if } m_2 = 1: & \text{not click, and choose } a_2 = 0. \\ \text{if } m_2 = 0: & \text{not click, and choose } a_2 = 1. \end{cases}$$

Cases 8 and 9 can be ruled out, as no strategy for the DM is consistent with equilibrium behavior in these cases.[6] I now argue that only Cases 2 and 6 admit informative second-period equilibria in which the good type reports truthfully. To establish this, I first examine the remaining cases other than Cases 2 and 6.

Case 1: In this case, the DM chooses $a_2 = 1$ whenever $\sigma_2 = 1$ is revealed, regardless of its headline. To sustain such behavior in equilibrium, the good type with signal $\sigma_2 = 1$ must randomize between the two headlines. If she does not mix, then one of the outcomes $(m_2, \sigma_2) = (1, 1)$ or $(m_2, \sigma_2) = (0, 1)$ is never sent by the good type with $\sigma_2 = 1$, so choosing $a_2 = 1$ is not optimal there (since the content $\sigma_2 = 1$ must be babbling), contradiction. However, truth-telling benefit prevents the good type with $\sigma_2 = 1$ from randomizing. Hence, Case 1 cannot be sustained in equilibrium.

Case 3: In this case, the DM clicks only on headline $m_2 = 0$. Consequently, the opportunistic outlet must always choose $m_2 = 0$. The good type must choose headline truthfully; the case when she receives $\sigma_2 = 0$ is straightforward: She prefers $a_2 = 0$, which truth-telling ($m_2 = 0$) implements. When she observes $\sigma_2 = 1$, she also reports truthfully, since by construction, either headline ultimately induces $a_2 = 1$. Given these headline strategies, any realization of $(m_2, \sigma_2) = (0, 1)$ must come from the opportunistic outlet, because the good type never sends $m_2 = 0$ when $\sigma_2 = 1$. Hence the DM treats $\sigma_2 = 1$ following $m_2 = 0$ as babbling and does not optimally choose $a_2 = 1$ there—contradicting the assumed DM behavior in this case. Therefore, no informative equilibrium is consistent with Case 3.

Case 4: This case features an off-the-equilibrium-path headline. Since the DM clicks only on $m_2 = 0$, the opportunistic outlet must always choose this headline regardless of her signal. For the good type with $\sigma_2 = 0$, the truth-telling benefit pins down $m_2 = 0$. Consider now the good type with $\sigma_2 = 1$. To induce the DM to choose $a_2 = 1$ upon observing content $\sigma_2 = 1$ following headline $m_2 = 0$, the good type must put positive probability on $m_2 = 0$ when $\sigma_2 = 1$; otherwise, the DM would not update in favor of $\sigma_2 = 1$ in this contingency. In equilibrium, however, the good type does not randomize over headlines; rather, she deterministically chooses $m_2 = 0$. The rationale is as follows. By sending the untruthful headline $m_2 = 0$ (when $\sigma_2 = 1$), she obtains payoff, $1/2 + \mu_G^* - q$, whereas truthful reporting yields k_G. In any informative equilibrium consistent with this case, it must be that $1/2 + \mu_G^* - q > k_G$; otherwise, if $1/2 + \mu_G^* - q \leq k_G$, the good type's action-related payoff becomes 0. She would then find it profitable to deviate by setting her effort level to zero, contradicting the informativeness. Thus, in any equilibrium consistent with Case 4, both types always choose $m_2 = 0$, rendering $m_2 = 1$ off the equilibrium path. Since we are restricting

[6] For example, in Case 8 the DM is required to click on $m_2 = 1$ and, upon observing $\sigma_2 = 0$, choose $a_2 = 1$; yet taking $a_2 = 1$ after $\sigma_2 = 0$ is not optimal for the DM, so such behavior cannot be part of any equilibrium. A symmetric argument rules out Case 9.

attention to equilibria in which the good type reports truthfully in both periods, Case 4 is excluded from consideration.

Case 5: This case is symmetric to Case 4. It likewise features an off-the-equilibrium-path headline, namely $m_2 = 0$. The arguments are analogous and omitted.

Case 7: An equilibrium consistent with Case 7 exists only if the cost of clicking is sufficiently high. I regard this equilibrium as implausible and exclude it from consideration.

In this case, the DM never clicks, so the opportunistic outlet has no incentive to deviate from truthful reporting. By contrast, the good type always misreports, i.e., chooses a headline $m_2 \neq \sigma_2$. The logic is as follows.

First, consider the good type with $\sigma_2 = 1$. To induce the DM to choose $a_2 = 1$ after *seeing only the headline* $m_2 = 0$ (recall: no clicking), the good with $\sigma_2 = 1$ must assign positive probability to $m_2 = 0$; otherwise, the DM would not take action $a_2 = 1$ on the basis of $m_2 = 0$. This implies that the good type's expected payoff, conditional on receiving $\sigma_2 = 1$, is $1/2 + \mu_G^* - q$. Now consider the good type with $\sigma_2 = 0$. To implement her preferred action $a_2 = 0$, she has to misreport—send $m_2 = 1$—since, by construction of this case, the DM chooses $a_2 = 1$ after $m_2 = 0$. Truthful reporting ($m_2 = 0$) would instead induce $a_2 = 1$, which she dislikes. Thus, she secures payoff 0 by misreporting, and $1/2 - \mu_G^* - q + k_G$ by reporting truthfully. In equilibrium, it must be that $1/2 - \mu_G^* - q + k_G$ is negative, so she strictly prefers to misreport. If instead this is non-negative, then her expected payoff conditional on $\sigma_2 = 0$ would be $1/2 - \mu_G^* - q + k_G$, while conditional on $\sigma_2 = 1$ it is $1/2 + \mu_G^* - q$. Evaluated at the effort stage, the action-related component would no longer depend on μ_G^*, making zero effort a profitable deviation—contradicting informativeness. Hence the good type with $\sigma_2 = 0$ must send $m_2 = 1$ with certainty.

Next, consider the good type with $\sigma_2 = 1$, who places positive probability on $m_2 = 0$. She does not randomize: in equilibrium she chooses $m_2 = 0$ with certainty. If she mixed across headlines, her action-related payoff evaluated at the effort stage would be zero, so any positive effort would be wasteful, again contradicting informativeness. Therefore, the good type always sends an untruthful headline $m_2 \neq \sigma_2$, while the opportunistic type reports truthfully. Under these headline strategies, equation (1) yields $\mu_G^* = 1/4$. Strategies are then fully pinned down.

When $\lambda_2 = 1$, this constitutes an informative equilibrium for any $c > 0$. I nevertheless exclude this equilibrium for two reasons: the good type always lies while the DM fully understands the signal, undermining the interpretation of headlines; (ii) when $\lambda_2 = 1$, this equilibrium is Pareto-dominated by Case 6, in which the good type reports truthfully.

Now suppose $\lambda_2 < 1$, so the DM is uncertain about the outlet's type at the beginning of period 2. In this equilibrium, the DM does not click on $m_2 = 0$ but chooses $a_2 = 1$. For this to be happen, two incentive conditions must hold:

(i) Choosing $a_2 = 1$ upon observing $m_2 = 0$ is optimal.

(ii) No clicking on $m_2 = 0$ is optimal.

Condition (i) requires that the the DM's posterior belief on $\omega_2 = 1$ must satisfy $1/2 + \lambda_2/4 - q \geq 0$, which is equivalent to $\lambda_2 \geq \overline{\lambda}$, where $\overline{\lambda} := 2(2q - 1)$. For condition (ii), compare the DM's expected payoff from clicking versus not clicking. If he were to click on $m_2 = 0$, he would see $\sigma_2 = 1$ with probability λ_2 and choose $a_2 = 1$, earning $1/2 + 1/4 - q$; with probability $1 - \lambda_2$, he would see $\sigma_2 = 0$ and chooses $a_2 = 0$, earning 0. The expected instrumental payoff (excluding the cost of clicking) is therefore $\lambda_2(1/2+1/4-q)$. By contrast, without clicking and choosing $a_2 = 1$, the expected payoff is $1/2 + \lambda_2/4 - q$.

Since $\lambda_2(1/2 + 1/4 - q) > 1/2 + \lambda_2/4 - q$ for all $\lambda_2 < 1$, the DM would strictly prefer to click in the absence of costs. Hence, condition (ii) requires a sufficiently high click cost: $c \geq (q - 1/2)(1 - \lambda_2) > 0$.

To summarize, an informative equilibrium consistent with Case 7 exists only if $\lambda_2 \geq \overline{\lambda}$ and the click cost is sufficiently high; moreover, in this equilibrium the good type always misreports. For these reasons, I exclude Case 7 from further consideration.

The remaining cases to consider are Case 2 and Case 6.

Case 2 Clickbait Equilibrium: As established in Section 3.1, the equilibrium headline strategies are: the good type reports truthfully, while the opportunistic type always sends $m_2 = 1$. Thus, the strategy profile is fully specified.

However, when $\lambda_2 = 1$ (the DM is certain the outlet is good), this equilibrium cannot be sustained: because the good typeis truthful, the headline reveals the content, so clicking strictly suboptimal.

Suppose instead $\lambda_2 < 1$, so the DM uncertain about the outlet's type at the start of period 2. I derive the conditions under which the clickbait equilibrium can be sustained.

First, for the DM to choose $a_2 = 1$ after clicking on $m_2 = 1$ and observing $\sigma_2 = 1$, his posterior that $\omega_2 = 1$ must be at least q, namely $1/2+\lambda_2/4 \geq q$, which is equivalent to $\lambda_2 \geq \overline{\lambda}$.

Next, consider the DM's incentive to click on headline $m_2 = 1$ by comparing expected payoffs with and without clicking.

If he clicks on $m_2 = 1$, he observes $\sigma_2 = 1$ with probability $1/(2 - \lambda_2)$, in which case he chooses $a_2 = 1$ and earns $1/2 + \lambda_2/4 - q$; if he observes $\sigma_2 = 0$, he chooses $a_2 = 0$ and earns 0. Hence, the expected instrumental payoff from clicking is:

$$\frac{1}{2 - \lambda_2}\left(\frac{1}{2} + \frac{\lambda_2}{4} - q\right). \tag{A.1}$$

If instead the DM does not click on $m_2 = 1$, his posterior belief on $\omega_2 = 1$ is:

$$\frac{1}{2} + \Pr(G \mid m_2 = 1) \cdot \frac{1}{4} = \frac{1}{2} + \frac{\lambda_2}{2 - \lambda_2}\frac{1}{4}.$$

Hence, he chooses $a_2 = 1$ if and only if $\lambda_2/(2-\lambda_2) \geq \overline{\lambda}$, which is equivalent to $\lambda_2 \leq 2\overline{\lambda}/(1+\overline{\lambda})$. Accordingly, his expected payoff from not clicking is:

$$\begin{cases} 0 & \text{if } \lambda_2 \in \left[\overline{\lambda}, \dfrac{2\overline{\lambda}}{1+\overline{\lambda}}\right], \\ \frac{1}{2} + \dfrac{\lambda_2}{2-\lambda_2} \cdot \frac{1}{4} - q & \text{if } \lambda_2 \in \left(\dfrac{2\overline{\lambda}}{1+\overline{\lambda}}, 1\right). \end{cases} \tag{A.2}$$

We now compare the instrumental payoff from clicking (A.1) with the payoff from not clicking (A.2) to determine when clicking is optimal.

- For $\overline{\lambda} < \lambda_2 \leq 2\overline{\lambda}/(1+\overline{\lambda})$, clicking is optimal if:

$$\frac{1}{2-\lambda_2}\left(\frac{1}{2} + \frac{\lambda_2}{4} - q\right) \geq c, \tag{A.3}$$

- For $2\overline{\lambda}/(1+\overline{\lambda}) \leq \lambda_2 < 1$, clicking is optimal if:

$$\frac{1}{2-\lambda_2}\left(\frac{1}{2} + \frac{\lambda_2}{4} - q\right) - \left(\frac{1}{2} + \frac{\lambda_2}{2-\lambda_2}\frac{1}{4} - q\right) \geq c \Leftrightarrow \frac{(2q-1)(1-\lambda_2)}{2(2-\lambda_2)} \geq c. \tag{A.4}$$

Hence, as long as $\overline{\lambda} < \lambda_2 < 1$, one can choose a sufficiently small $\overline{c} > 0$ such that both inequalities (A.3) and (A.4) hold whenever $c \leq \overline{c}$. This establishes the existence of the clickbait equilibrium and completes the proof of statement *(ii-a)*.

Case 6: In this equilibrium, both types report headlines truthfully. The opportunistic type reports truthfully because no headline is ever clicked. For the good type with $\sigma_2 = 0$, truthful reporting directly induces the desired action $a_2 = 0$.

Consider the good type with $\sigma_2 = 1$. She must put positive probability on reporting $m_2 = 1$; otherwise, the DM would not choose action $a_2 = 1$ upon seeing that headline. Moreover, she does not mix between headlines: if she randomized, her action-related payoff would be zero, making zero effort a profitable deviation and contradicting informativeness. Hence she chooses $m_2 = 1$ with certainty when $\sigma_2 = 1$.

Given that both types report truthfully and the headline perfectly reveals the content, not clicking any headline is optimal. Finally, for the DM to choose $a_2 = 1$ after observing $m_2 = 1$, it must be that: $1 \geq \lambda_2 \geq \overline{\lambda}$. This equilibrium exists even in the fully separating case $\lambda_2 = 1$. In fact, when $\lambda_2 = 1$, it is the only informative period-2 equilibrium in which the good type reports truthfully.

This completes the proof of statements *(i)* through *(v)*. □

A.2 Proof of Proposition 5

Given that both outlet types choose headlines truthfully in period 1, the DM's posterior belief that the outlet is good, conditional on any headline, coincides with the prior λ_1. Namely, $\lambda_2(o_1) = \lambda_1$ for all o_1. Hence, the two-period game collapses to the single-period game analyzed for period 2. This completes the proof.

A.3 Proof of Theorem 1

I construct an informative equilibrium in which the good type reports headlines truthfully while the opportunistic type always sends $m_1 = 1$. To that end, I first derive the DM's equilibrium strategy.

Lemma 2. *DM's equilibrium strategy is:*

$$\begin{cases} m_1 = 1 & \textit{click and, upon observing } \sigma_1, \textit{choose } a_1 = \sigma_1, \\ m_1 = 0 & \textit{do not click and choose } a_1 = 0. \end{cases}$$

Proof. Suppose the DM believes that the good type is truthful and the opportunistic type chooses $m_1 = 1$ in period 1. Upon observing $m_1 = 0$, the DM infers that the outlet is the good type with $\sigma_1 = 0$; hence $\lambda_2(m_1 = 0) = 1$. He therefore does not click on $m_1 = 0$ and chooses $a_1 = 0$.

Now consider $m_1 = 1$. I show the DM clicks. Suppose by contradiction that he does not click on $m_1 = 1$. Then, in equilibrium he must choose $a_1 = 1$ after $m_1 = 1$; otherwise, he would always choose $a_1 = 0$ in period 1, driving the good type's action-related payoff in that period to zero and inducing a deviation zero effort, contradicting informativeness.

Conditional on not clicking, choosing $a_1 = 1$ after $m_1 = 1$ requires

$$\Pr(\omega_1 = 1 \mid m_1 = 1) \geq q \quad \Leftrightarrow \quad \frac{1}{2} + \Pr(\mathrm{G} \mid m_1 = 1)\, \mu_G^* \geq q \quad \Leftrightarrow \quad \frac{1}{2} + \frac{\lambda_1}{2 - \lambda_1}\, \mu_G^* \geq q.$$

Then, since $\Pr(\mathrm{G} \mid m_1 = 1, \sigma_1 = 1) = \lambda_1 > \lambda_1/(2 - \lambda_1)$,

$$\Pr(\omega_1 = 1 \mid m_1 = 1, \sigma_1 = 1) = \frac{1}{2} + \lambda_1 \cdot \mu_{\mathrm{G}}^* > q,$$

so if he were to click on $m_1 = 1$ and observe $\sigma_1 = 1$, the DM would choose $a_1 = 1$.

Clicking therefore strictly improves the DM's expected period-1 payoff: with clicking he conditions his action on realized content, whereas without clicking he must commit to a single action. Moreover, clicking does not decrease the period-2 payoff. Hence, for sufficiently small $c > 0$, clicking yields a strictly higher expected payoff than not clicking, contradicting the hypothesis of no clicking after $m_1 = 1$. Therefore, in equilibrium the DM clicks on $m_1 = 1$.

Moreover, conditional on clicking and observing σ_1, the DM chooses $a_1 = \sigma_1$. The case for $\sigma_1 = 0$ is immediate. If instead the DM choose were to choose $a_1 = 0$ after

$(m_1 = 1, \sigma_1 = 1)$—implying he would always choose $a_1 = 0$ in period 1—the good type's action-related payoff would be zero, making a deviation to $\mu_{\mathrm{G}} = 0$ profitable—contradicting informativeness. □

Next, I derive the good type's equilibrium effort μ^*_{G}. Since her effort does not affect the second-period payoff, the same reasoning as in Section 3.1 yields $\mu^*_{\mathrm{G}} = 1/4$. The equilibrium strategies are thus fully specified.

For the DM's strategy to be an equilibrium, two conditions must hold:

(1) upon clicking on $m_1 = 1$ and observing $\sigma_1 = 1$, choosing $a_1 = 1$ is optimal;

(2) clicking on $m_1 = 1$ delivers a higher expected payoff than not clicking.

Condition (1) requires

$$\Pr(\omega_1 = 1 \mid m_1 = 1,\ \sigma_1 = 1) \geq q \quad \Leftrightarrow \quad \frac{1}{2} + \frac{\lambda_1}{4} \geq q \quad \Leftrightarrow \quad \lambda_1 \geq \overline{\lambda}.$$

Hence assume $\lambda_1 \geq \overline{\lambda}$.

For condition (2), strict inequality $\lambda_1 > \overline{\lambda}$ is needed: if $\lambda_1 = \overline{\lambda}$, the DM's optimal action is always $a_1 = 0$ regardless of the realized content σ_1, so his first-period instrumental payoff from clicking equals that from not clicking. The second-period payoff is unchanged. Thus, any $c > 0$ deters clicking, contradicting the proposed strategy. If instead $\lambda_1 > \overline{\lambda}$, clicking strictly raises the period-1 expected payoff, while leaving period-2 payoff unaffected. Therefore, for sufficiently small $c > 0$, clicking yields a strictly higher expected payoff than not clicking, satisfying condition (2).

Finally, it remains to verify that the outlet's strategy is optimal given the DM's strategy in Lemma 2.

Opportunistic type: Upon observing $\sigma_1 = 1$, the period-1 payoff is $1 + k_{\mathrm{O}}$, so deviating is never profitable. Upon observing $\sigma_1 = 0$, the period-2 payoff is k_{O} regardless of the headline, since the DM's posterior that the outlet is good is zero in either case: $\lambda_2(m_1 = 0) = \lambda_2(m_1 = 1, \sigma_1 = 0,) = 0$. Hence, she must choose the headline that will be clicked in period 1.

Good type: Consider the most profitable deviation. A deviation may alter both effort and the headline strategy. In any optimal deviation, $m_1 = 0$ is chosen when $\sigma_1 = 0$, yielding $V_{\mathrm{G}}(\lambda_2 = 1)$ in period 2. Let $p := \Pr(m_1 = 1 \mid \sigma_1 = 1)$ denote the headline strategy when $\sigma_1 = 1$. The most profitable deviation is a pair of an effort level $\tilde{\mu}_G$ and $\tilde{p}$ that solves

$$\max_{\mu_{\mathrm{G}}, p} \Bigg[-(\mu_{\mathrm{G}})^2 + \frac{1}{2}p\left(\frac{1}{2} + \mu_{\mathrm{G}} - q + k_{\mathrm{G}} + V_{\mathrm{G}}(\lambda_2 = \lambda_1)\right) + \frac{1}{2}(1-p)V_{\mathrm{G}}(\lambda_2 = 1) + \frac{1}{2}\left(k_{\mathrm{G}} + V_{\mathrm{G}}(\lambda_2 = 1)\right)\Bigg],$$

The objective function is constructed as follows. With probability $1/2$ the good type observes $\sigma_1 = 0$, yielding $0 + k_G$ in period 1, and $V_G(\lambda_2 = 1)$ in period 2. With probability $1/2$ she observes $\sigma_1 = 1$; sending $m_1 = 1$ then yields $1/2 + \mu_G - q + k_G + V_G(\lambda_2 = \lambda_1)$, while $m_1 = 0$ yields $V_G(\lambda_2 = 1)$.

Observe that mixing over headlines is not optimal. If an optimal deviation involves mixing, indifference would require

$$\frac{1}{2} + \tilde{\mu}_G - q + k_G + V_G(\lambda_2 = \lambda_1) = V_G(\lambda_2 = 1),$$

in which case a small reduction in μ_G would strictly increase the objective, contradicting optimality. Hence $\tilde{p}$ must be either 0 or 1, and the only candidates are $(\tilde{p}, \tilde{\mu}_G) = (0, 0)$ or $(1, 1/4)$, the latter coinciding with the equilibrium strategy. It therefore suffices to compare the payoff from $(\tilde{p}, \tilde{\mu}_G) = (0, 0)$ with the equilibrium payoff. The optimal-deviation payoff equals $k_G/2 + V_G(\lambda_2 = 1)$, whereas the equilibrium payoff equals

$$\left(\frac{5}{16} - \frac{q}{2} + k_G\right) + \frac{1}{2}V_G(\lambda_2 = 1) + \frac{1}{2}V_G(\lambda_2 = \lambda_1) = 2V_G(\lambda_2 = 1),$$

which is strictly higher than the former payoff. Therefore, no profitable deviation exists. This completes the proof of Theorem 1. □

A.4 Proof of Lemma 1

Suppose that the good type chooses headlines truthfully, while the opportunistic type always chooses the headline $m_1 = 1$. When the DM observes $m_1 = 0$, he perfectly infers that the content is $\sigma_1 = 0$ as well as that the outlet is of good type, and hence it is optimal for him to not click on it and choose $a_1 = 0$.

Now consider the case when the DM observes $m_1 = 1$. Suppose by contradiction that the DM does not click on $m_1 = 1$ in the equilibrium. There are two cases to consider:

Case 1: DM chooses $a_1 = 0$ after $m_1 = 1$. In this equilibrium, the good type must exert positive effort. To see this, note that the necessary condition for the opportunistic type to exert effort is

$$\lambda_2(m_1 = 1, \gamma = 1) > \overline{\lambda} \geq \lambda_2(m_1 = 1, \gamma = 0),$$

which can be rewritten as

$$\frac{\lambda_1\left(\frac{1}{2} + \mu_G^*\right)}{\lambda_1\left(\frac{1}{2} + \mu_G^*\right) + 1 - \lambda_1} > \overline{\lambda} \geq \frac{\lambda_1\left(\frac{1}{2} - \mu_G^*\right)}{\lambda_1\left(\frac{1}{2} - \mu_G^*\right) + 1 - \lambda_1}.$$

But this cannot be true unless μ_G^* is positive. Thus, the good type must exert effort in this case. The good type's only incentive to exert positive effort comes from the second-period payoff, since the DM always chooses $a_1 = 0$ in period 1. Thefore, it must be that

$$\lambda_2(m_1 = 1, \gamma = 1) > \overline{\lambda} \geq \lambda_2(m_1 = 1, \gamma = 0),$$

meaning that confirmation of $m_1 = 1$ by fact-checking, i.e., $(m_1 = 1, \gamma_1 = 1)$, pushes the second period to the clickbait equilibrium, while disconfirmation, i.e., $(m_1 = 1, \gamma = 0)$, leads to the uninformative equilibrium. However, if this is true, then the good type with $\sigma_1 = 1$ has profitable deviaiton: by misreporting (i.e., choosing $m_1 = 0$), she can guarantee that the clickbait equilibrium in period 2 while leaving her first-period action-related payoff unchanged (the DM still chooses $a_1 = 0$). Hence, for sufficiently small benefit of truth-telling k_G, is sufficiently small, she would strictly prefer to lie, contradiction.

Case 2: DM chooses $a_1 = 1$ after $m_1 = 1$. Let μ_G^* and μ_O^* denote the equilibrium effort level. Since the DM chooses $a_1 = 1$ after observing $m_1 = 1$, his posterior on $\omega_1 = 1$ must be no less than q:

$$\frac{1}{2} + \frac{\lambda_1}{2 - \lambda_1} \cdot \mu_G^* \geq q.$$

This implies that he would also choose $a_1 = 1$ if he would click on $m_1 = 1$ and observe $\sigma_1 = 1$ since his posterior conditional on $(m_1 = 1, \sigma_1 = 1)$, which is $1/2 + \lambda_1 \mu_G^* + (1 - \lambda_1)\mu_O^*$, is strictly greater than $1/2 + (\lambda_1/(2 - \lambda_1))\mu_G^*$. Therefore, clicking on $m_1 = 1$ yields strictly higher first-period action-related payoff. Moreover, the second-period payoff does not decrease by clicking. The reason is as follows: When he does not click on $m_1 = 1$, his second-period payoff is given by

$$\sum_{\gamma} \Pr(\gamma \mid m_1 = 1) \cdot V_{DM}(\lambda_2(m_1 = 1, \gamma)).$$

By contrast, his second-period payoff from clicking is

$$\begin{aligned} &\sum_{\sigma_1} \Pr(\sigma_1 \mid m_1 = 1) \sum_{\gamma} \Pr(\gamma \mid m_1 = 1, \sigma_1)\, V_{DM}(\lambda_2(m_1 = 1, \sigma_1, \gamma)) \\ = &\sum_{\gamma} \Pr(\gamma \mid m_1 = 1) \sum_{\sigma_1} \Pr(\sigma_1 \mid m_1 = 1, \gamma)\, V_{DM}(\lambda_2(m_1 = 1, \sigma_1, \gamma)) \,. \end{aligned}$$

Thus, it suffices to show that for each γ,

$$V_{DM}(\lambda_2(m_1 = 1, \gamma)) \leq \sum_{\sigma_1} \Pr(\sigma_1 \mid m_1 = 1, \gamma) V_{DM}(\lambda_2(m_1 = 1, \sigma_1, \gamma)). \tag{A.5}$$

Consider first $\gamma = 1$. Suppose $\lambda_2(m_1 = 1, \gamma = 1) > \overline{\lambda}$, so the second-period after

$(m_1 = 1, \gamma = 1)$ is clickbait equilibrium. Then,

$$V_{\mathrm{DM}}(\lambda_2(m_1 = 1, \gamma = 1)) = \frac{1}{2}\left(\frac{1}{2} + \frac{\lambda_2(m_1 = 1, \gamma = 1)}{4} - q\right) - \left(1 - \frac{\lambda_2(m_1 = 1, \gamma = 1)}{2}\right) c.$$

Since the posterior belief on the outlet being the good type is the convex combination, we have

$$\begin{aligned}
&\lambda_2(m_1 = 1, \gamma = 1) = \sum_{\sigma_1} \Pr(\sigma_1 \mid m_1 = 1, \gamma = 1)\lambda_2(m_1 = 1, \sigma_1, \gamma = 1),\\
&\lambda_2(m_1 = 1, \sigma_1 = 1, \gamma = 1) > \lambda_2(m_1 = 1, \gamma = 1),\\
&\lambda_2(m_1 = 1, \sigma_1 = 0, \gamma = 1) = 0.
\end{aligned}$$

Thus, by denoting

$$f(x) = \frac{1}{2}\left(\frac{1}{2} + \frac{x}{4} - q\right) - \left(1 - \frac{x}{2}\right) c,$$

we have

$$\begin{aligned}
V_{\mathrm{DM}}(\lambda_2(m_1 = 1, \gamma = 1)) &= \Pr(\sigma_1 = 1 \mid m_1 = 1, \gamma = 1)\, f(\lambda_2(m_1 = 1, \sigma_1 = 1, \gamma = 1))\\
&\quad + \Pr(\sigma_1 = 0 \mid m_1 = 1, \gamma = 1)\, f(\lambda_2(m_1 = 1, \sigma_1 = 0, \gamma = 1))\\
&= \Pr(\sigma_1 = 1 \mid m_1 = 1, \gamma = 1)\, V_{\mathrm{DM}}(\lambda_2(m_1 = 1, \sigma_1 = 1, \gamma = 1))\\
&\quad + \Pr(\sigma_1 = 0 \mid m_1 = 1, \gamma = 1)\, f(0)\\
&< \Pr(\sigma_1 = 1 \mid m_1 = 1, \gamma = 1)\, V_{\mathrm{DM}}(\lambda_2(m_1 = 1, \sigma_1 = 1, \gamma = 1))\\
&\quad + \Pr(\sigma_1 = 0 \mid m_1 = 1, \gamma = 1) \cdot 0\\
&= \Pr(\sigma_1 = 1 \mid m_1 = 1, \gamma = 1)\, V_{\mathrm{DM}}(\lambda_2(m_1 = 1, \sigma_1 = 1, \gamma = 1))\\
&\quad + \Pr(\sigma_1 = 0 \mid m_1 = 1, \gamma = 1)\, V_{\mathrm{DM}}(\lambda_2(m_1 = 1, \sigma_1 = 0, \gamma = 1))\\
&= \sum_{\sigma_1} \Pr(\sigma_1 \mid m_1 = 1, \gamma = 1)\, V_{\mathrm{DM}}(\lambda_2(m_1 = 1, \sigma_1, \gamma = 1))\,,
\end{aligned}$$

which is the desired result. Now suppose $\lambda_2(m_1 = 1, \gamma = 1) \leq \overline{\lambda}$, so the second-period equilibrium after $(m_1 = 1, \gamma = 1)$ is uninformative, implying $V_{\mathrm{DM}}(\lambda_2(m_1 = 1, \gamma = 1)) = 0$, but in this case, equation (A.5) is obvious. The case for $\gamma = 0$ and $\gamma = \emptyset$ can be analogously proven. Therefore, the DM would be better off by clicking on $m_1 = 1$ as long as the cost of clicking is sufficiently small, contradiction.

This completes the proof of Lemma 1. □

A.5 Proof of Theorem 2

Step 1: Deriving Equilibrium Effort: Lemma 1 has already characterized the DM's strategy in the informative equilibrium where the good type reports truthfully, while the opportunistic type always sets $m_1 = 1$. I now derive equilibrium effort. As argued in Section 4.2.1, equation (2), restated below for clarity, is necessary for the opportunistic type to exert effort:

$$\overbrace{\frac{\lambda_1\left(\frac{1}{2}+\mu_G^*\right)}{\lambda_1\left(\frac{1}{2}+\mu_G^*\right)+(1-\lambda_1)\left(\frac{1}{2}+\mu_O^*\right)}}^{\lambda_2(1,1,1)} > \overline{\lambda} \geq \overbrace{\frac{\lambda_1\left(\frac{1}{2}-\mu_G^*\right)}{\lambda_1\left(\frac{1}{2}-\mu_G^*\right)+(1-\lambda_1)\left(\frac{1}{2}-\mu_O^*\right)}}^{\lambda_2(1,1,0)}. \tag{2}$$

This condition states that confirmation by fact-checking leads to the clickbait equilibrium, whereas disconfirmation leads to uninformative one in period 2. We recall that when equation (2) is true, the opportunistic type's equilibrium effort is

$$\mu_O^* = \frac{\beta\delta_O}{4}, \tag{3}$$

where $\delta_O := (1 - k_O/2)$.

Now consider the good type. Her equilibrium effort level μ_G^* must be a solution to the maximization of her continuation payoff, holding all other strategies fixed. Under equation (2), her continuation payoff from choosing μ_G is derived as follows: when she observes $\sigma_1 = 1$, which occurs with probability $1/2$, she truthfully publishes and earn first-period payoff of $1/2 + \mu_G - q + k_G$. Moreover, she obtains the second-period payoff of

$$\beta\left(\tfrac{1}{2}+\mu_G\right)V_G\big(\lambda_2(1,1,1)\big)+\beta\left(\tfrac{1}{2}-\mu_G\right)V_G\big(\lambda_2(1,1,0)\big)+(1-\beta)V_G\big(\lambda_2(1,1,\emptyset)\big).$$

When instead the good type observes $\sigma_1 = 0$, she obtains k_G in period 1 and $V_G(\lambda_2 = 0)$ in period 2.

Thus, μ_G^* solves the following problem:

$$\max_{\mu_G}\left[\frac{1}{2}\,\mu_G+\frac{1}{2}\,\beta\,\mu_G\Big[V_G(\lambda_2(1,1,1))-V_G(\lambda_2(1,1,0))\Big]-(\mu_G)^2\right],$$

with the first-order condition

$$\mu_G^* = \frac{1+\beta\delta_G}{4}, \tag{4}$$

where $\delta_G := 5/16 - q/2$.

Now suppose that equation (2) fails, so that the period-2 payoff is not responsive to the outcome of fact-checking, that is, $V_O(\lambda_2(1,1,1)) - V_G(\lambda_2(1,1,0)) = 0$. Then second-period payoffs are independent of the first-period effort. It follows that the opportunistic type exerts no effort in period 1. For the good type, when considering optimal effort level, she isolates the second period from the first period, and hence chooses $\mu_G^* = 1/4$.

Step 2: Consistency between λ_2 and Effort level: In the informative equilibrium in which both types exert effort, equation (2) must be true when μ_G^* and μ_O^* are given by equation (4) and 3, respectively.

On the other hand, in the equilibrium where only the good type exerts effort, either

$$\lambda_2(1,1,0) = \frac{\lambda_1\left(\frac{1}{2}-\mu_G^*\right)}{\lambda_1\left(\frac{1}{2}-\mu_G^*\right)+(1-\lambda_1)\left(\frac{1}{2}-\mu_O^*\right)} > \overline{\lambda} \tag{A.6}$$

or

$$\overline{\lambda} \geq \frac{\lambda_1\left(\frac{1}{2}+\mu_G^*\right)}{\lambda_1\left(\frac{1}{2}+\mu_G^*\right)+(1-\lambda_1)\left(\frac{1}{2}+\mu_O^*\right)} = \lambda_2(1,1,1) \tag{A.7}$$

must be true when $\mu_G^* = 1/4$ and $\mu_O^* = 0$.

Consider the former informative equilibrium. Substituting equations (4) and (3) into equation (2) yields

$$\underline{\lambda}^*(\beta) < \lambda_1 \leq \overline{\lambda}^*(\beta), \tag{A.8}$$

where thresholds are given by

$$\underline{\lambda}^*(\beta) := \frac{\overline{\lambda}\left(2+\beta\delta_O\right)}{3+\beta\delta_G-\overline{\lambda}\left(1+\beta\delta_G-\beta\delta_O\right)}, \tag{A.9}$$

$$\overline{\lambda}^*(\beta) := \frac{\overline{\lambda}\left(2-\beta\delta_O\right)}{1-\beta\delta_G+\overline{\lambda}\left(1+\beta\delta_G-\beta\delta_O\right)}. \tag{A.10}$$

Now consider instead the latter informative equilibrium. Setting $\mu_G^* = 1/4$ and $\mu_O^* = 0$ in equation (A.6) gives

$$\frac{4(2q-1)}{4q-1} = \frac{2\overline{\lambda}}{1+\overline{\lambda}} < \lambda_1, \tag{A.11}$$

whereas equation (A.7) becomes

$$\lambda_1 \leq \frac{4(2q-1)}{5-4q} = \frac{2\overline{\lambda}}{3-\overline{\lambda}},$$

but, the latter cannot hold here because $\lambda_1 \geq \overline{\lambda}$ is necessary to induce the DM choose $a_1 = 1$ after $(m_1 = 1, \sigma_1 = 1)$. Thus, equation (A.7) can be discarded. Note that $\underline{\lambda}^*(\beta), \overline{\lambda}^*(\beta) \to \overline{\lambda}$ as $\beta \to 1$, and $\overline{\lambda}^*(\beta) \to 2\overline{\lambda}/(1+\lambda)$ as $\beta \to 0$.

In summary,

- **Equilibrium where both types exert effort:** It must be that

$$\underline{\lambda}^*(\beta) < \lambda_1 \le \overline{\lambda}^*(\beta), \tag{A.8}$$

where $\underline{\lambda}^*(\beta)$ and $\overline{\lambda}^*(\beta)$ are given by equation (A.9) and (A.10), respectively.

- **Equilibrium where only the good type exerts effort:** It must be that

$$\frac{2\overline{\lambda}}{1+\overline{\lambda}} < \lambda_1. \tag{A.11}$$

Step 3: Other Conditions Consider the informative equilibrium where both types exert effort, so suppose equation (A.8) is true. Here, the DM chooses $a_1 = 1$ after $(m_1 = 1, \sigma_1 = 1)$. For this to happen, his posterior belief on $\omega_1 = 1$ must be no less than q:

$$\frac{1}{2} + \lambda_1 \mu_G^* + (1-\lambda_1)\mu_O^* \ge q,$$

which can be rewritten as

$$\begin{aligned}
&\frac{1}{2} + \lambda_1 \cdot \frac{1+\beta\delta_G}{4} + (1-\lambda_1)\cdot\frac{\beta\delta_O}{4} \ge q \\
\Leftrightarrow\ & \lambda_1\,(1+\beta\delta_G) + (1-\lambda_1)\beta\delta_O \ge \overline{\lambda} \\
\Leftrightarrow\ & \lambda_1 \ge \underline{\lambda}^{**}(\beta),
\end{aligned}$$

where

$$\underline{\lambda}^{**}(\beta) := \frac{\overline{\lambda} - \beta\delta_O}{1+\beta\delta_G - \beta\delta_O}. \tag{A.12}$$

Now consider the informative equilibrium where only the good type exerts effort in period 1. Here, for the DM to choose $a_1 = 1$ after $(m_1 = 1, \sigma_1 = 1)$, his posterior belief on $\omega_1 = 1$ must be no less than q; this requires $\lambda_1 \ge \overline{\lambda}$, which is true when equation (A.11) holds. □

A.6 Proof of Theorem 3

The first part (i) immediately follows from the fact that $\overline{\lambda}^*(\beta)$ is strictly decreasing and continuous in β, with $\overline{\lambda}^*(\beta) \to \overline{\lambda}$ as $\beta \to 1$, and $\overline{\lambda}^*(\beta) \to 2\overline{\lambda}/(1+\overline{\lambda})$ as $\beta \to 0$.

I shall prove (ii). In the informative equilibrium, the DM's payoff in period 1 is

$$\frac{1}{2}\left(\frac{1}{2} + \lambda_1 \cdot \frac{1+\beta\delta_G}{4} + (1-\lambda_1)\cdot\frac{\beta\delta_O}{4} - q\right) - \left(1 - \frac{\lambda_1}{2}\right)c. \tag{A.13}$$

Consider the period-2 payoff. The posterior reputation after $m_1 = 0$ becomes $\lambda_2 = 1$ regardless of the outcome of fact-checking γ, yielding $V_G(\lambda_2 = 1) = (1/2)(3/4-q)$. Likewise, the posterior reputation after $(m_1 = 1, \sigma_1 = 0)$ is $\lambda_2 = 0$ regardless of γ, yielding 0. By contrast, the posterior reputation after $(m_1 = 1, \sigma_1 = 1)$ depends on γ, and this satisfies

$$\lambda_2(m_1 = 1, \sigma_1 = 1, \gamma = 1) > \overline{\lambda} \geq \lambda_2(1,1,0) \quad \text{and} \quad \lambda_2(1,1,\gamma = \emptyset) = \lambda_1 > \overline{\lambda}.$$

Therefore, the clickbait equilibrium will be played if the outcome is either $(m_1 = 1, \sigma_1 = 1, \gamma = 1)$ or $(m_1 = 1, \sigma_1 = 1, \gamma = \emptyset)$, while the uninformative equilibrium will be played after $(m_1 = 1, \sigma_1 = 1, \gamma = 0)$. Thus, his period-2 payoff is

$$\begin{aligned} &\Pr(m_1 = 0)\tfrac{1}{2}\left(\tfrac{3}{4} - q\right) \\ &\quad + \Pr(m_1 = 1)\Pr(\sigma_1 = 1 \mid m_1 = 1) \sum_{\gamma \in \{1,\emptyset\}} \Pr(\gamma \mid \sigma_1 = 1, m_1 = 1)\, V_{\mathrm{G}}\big(\lambda_2(1,1,\gamma)\big), \end{aligned}$$

where the second term can be rewritten as

$$\begin{aligned} &\sum_{\gamma \in \{1,\emptyset\}} \Pr(\sigma_1 = 1, m_1 = 1, \gamma)\, V_{\mathrm{G}}\big(\lambda_2(1,1,\gamma)\big) \\ &= \Pr(\gamma = 1)\ \Pr(\sigma_1 = 1, m_1 = 1 \mid \gamma = 1)\, V_{\mathrm{G}}\big(\lambda_2(1,1,1)\big) \\ &\quad + \Pr(\gamma = \emptyset)\ \Pr(\sigma_1 = 1, m_1 = 1 \mid \gamma = \emptyset)\, V_{\mathrm{G}}\big(\lambda_2(1,1,\emptyset)\big) \\ &= \Pr(\gamma = 1)\ \Pr(\sigma_1 = 1, m_1 = 1 \mid \omega_1 = 1)\, V_{\mathrm{G}}\big(\lambda_2(1,1,1)\big) \\ &\quad + \Pr(\gamma = \emptyset)\ \Pr(\sigma_1 = 1, m_1 = 1)\, V_{\mathrm{G}}\big(\lambda_2(1,1,\emptyset)\big) \\ &= \tfrac{\beta}{2}\ \Pr(\sigma_1 = 1, m_1 = 1 \mid \omega_1 = 1)\, V_{\mathrm{G}}\big(\lambda_2(1,1,1)\big) \\ &\quad + \left\{\tfrac{1-\beta}{2}\ \Pr(\sigma_1 = 1, m_1 = 1 \mid \omega_1 = 1) + \tfrac{1-\beta}{2}\ \Pr(\sigma_1 = 1, m_1 = 1 \mid \omega_1 = 0)\right\} V_{\mathrm{G}}\big(\lambda_2(1,1,\emptyset)\big) \\ &= \tfrac{\beta}{2} \cdot \left(\tfrac{1}{2} + \lambda_1 \tfrac{1+\beta\delta_{\mathrm{G}}}{4} + (1-\lambda_1)\tfrac{\beta\delta_{\mathrm{O}}}{4}\right) V_{\mathrm{G}}\big(\lambda_2(1,1,1)\big) \ + \ \tfrac{1-\beta}{2}\, V_{\mathrm{G}}\big(\lambda_2(1,1,\emptyset)\big). \end{aligned}$$

Thus, the second-period DM's payoff is

$$\tfrac{\lambda_1}{2} \cdot \tfrac{1}{2}\left(\tfrac{3}{4} - q\right) \ + \ \tfrac{\beta}{2} \cdot \left(\tfrac{1}{2} + \lambda_1 \tfrac{1+\beta\delta_{\mathrm{G}}}{4} + (1-\lambda_1)\tfrac{\beta\delta_{\mathrm{O}}}{4}\right) V_{\mathrm{G}}\big(\lambda_2(1,1,1)\big) \ + \ \tfrac{1-\beta}{2}\, V_{\mathrm{G}}\big(\lambda_1\big). \tag{A.14}$$

Thus, the DM's equilibrium payoff is the sum of equation (A.13) and equation (A.14). Tedious calculation—summing and then differentiating with respect to β—implies that the DM's payoff in this informative equilibrium is increasing in β.

For Part (iii), consider the uninformative equilibrium. The DM's period-1 payoff here is zero. For period 2, note that his posterior belief that the outlet is good remains the prior, λ_1, namely $\lambda_2(o_1, \gamma) = \lambda_1$ for any first-period outcome $o_1 \in \mathcal{M} \cup \mathcal{M} \times \mathcal{C}$. Thus, his period 2 (and thus total) payoff in the uninformative equilibrium is

$$\tfrac{1}{2}\left(\tfrac{1}{2}+\tfrac{\lambda_1}{4}-q\right)\;-\;\left(1-\tfrac{\lambda_1}{2}\right)c.$$

This is strictly smaller than equation (A.13), the period-1 payoff in the informative equilibrium. Therefore, the DM's welfare under any $\beta'' > \overline{\beta}$ (uninformative) is strictly lower than under any $\beta' < \overline{\beta}$ (informative), proving (iii). □

A.7 Proof of Theorem 4

(1): Suppose $\max\{\underline{\lambda}^*(\beta), \underline{\lambda}^{**}(\beta)\} < \lambda_1 \leq \overline{\lambda}^*(\beta)$. Now I compare the DM's equilibrium welfare in both settings—with and without fact-checking. There are two cases: $\lambda_1 \leq \overline{\lambda}$ and $\overline{\lambda} < \lambda_1$.

Case $\lambda_1 \leq \overline{\lambda}$: Suppose first that $\lambda_1 \leq \overline{\lambda}$. In this case, DM's equilibrium welfare in the absence of fact-checking is 0. To see this, note that in period 1, the DM never clicks any headline and always chooses $a_1 = 0$, and his posterior belief that the media outlet is good type λ_2 remains to be the prior λ_1 after any headline, implying that the second-period equilibrium is also uninformative one with probability 1. By contrast, when there is fact-checking, his equilibrium payoff is positive since by equation (2) implies that clickbait equilibrium will be played in period 2 with positive probability.

Case $\overline{\lambda} < \lambda_1$: Suppose instead that $\overline{\lambda} < \lambda_1$, so that in the absence of fact-checking informative equilibrium is played. Consider first the DM's equilibrium payoff when there is no fact-checking. In period 1, the only case where he chooses $a_1 = 1$ is after $(m_1 = 1, \sigma_1 = 1)$, which occurs with probability $1/2$ and in which case he obtains

$$\frac{1}{2} + \Pr(\mathrm{G} \mid m_1 = 1, \sigma_1 = 1)\mu^*_{\mathrm{G}} + \Pr(\mathrm{O} \mid m_1 = 1, \sigma_1 = 1)\mu^*_{\mathrm{O}} - q = \frac{1}{2} + \frac{\lambda_1}{4} - q.$$

In all other cases, he chooses $a_1 = 0$, earning payoff of 0. In period 2, his posterior belief on the good type is $\lambda_2 = \lambda_1$ if the outcome is $(m_1 = 1, \sigma_1 = 1)$ and $\lambda_2 = 0$ otherwise. Thus, he enjoys the payoff of the clickbait equilibrium with $\lambda_2 = \lambda_1$ after the outcome $(m_1 = 1, \sigma_1 = 1)$ and his second-period payoff is 0 otherwise.

Now consider his equilibrium payoff in the presence of fact-checking. Likewise, the only case he chooses $a_1 = 1$ is after $(m_1 = 1, \sigma_1 = 1)$, which occurs with probability $1/2$, and he chooses $a_1 = 0$ otherwise. But in the presence of fact-checking, in the outcome

$(m_1 = 1, \sigma_1 = 1)$, he earns payoff of

$$
\begin{aligned}
&\frac{1}{2} + \Pr(\mathrm{G} \mid m_1 = 1, \sigma_1 = 1)\mu_{\mathrm{G}}^* + \Pr(\mathrm{O} \mid m_1 = 1, \sigma_1 = 1)\mu_{\mathrm{O}}^* - q \\
=&\frac{1}{2} + \frac{1+\beta\delta_{\mathrm{G}}}{4}\lambda_1 + \frac{\beta\delta_{\mathrm{O}}}{4}(1-\lambda_1) - q,
\end{aligned}
$$

which is strictly greater than $1/2 + \lambda_1/4 - q$.

Moreover, the second period payoff increases as well. To see this, note that in the presence of fact-checking, if he observes $m_1 = 0$ or observes $\sigma_1 = 0$ by clicking $m_1 = 1$, his posterior belief on the good type reduces to $\lambda_2 = 0$ and this does not change by the result of fact-checking. Thus, it suffices to consider the case where he observes $\sigma_1 = 1$ by clicking $m_1 = 1$. In this case, his expected second-period payoff is

$$
\begin{aligned}
&\sum_{\gamma} \Pr(\gamma \mid m_1 = 1, \sigma_1 = 1) V_{\mathrm{G}}(\lambda_2(m_1 = 1, \sigma_1 = 1, \gamma)) \\
=&\Pr(\gamma = 1 \mid m_1 = 1, \sigma_1 = 1) f(\lambda_2(1,1,1)) + \Pr(\gamma = \emptyset \mid m_1 = 1, \sigma_1 = 1) f(\lambda_1) + 0 \\
>&\sum_{\gamma} \Pr(\gamma \mid m_1 = 1, \sigma_1 = 1) f(\lambda_2(1,1,\gamma)) \\
=&V_{\mathrm{G}}(\lambda_2(m_1 = 1, \sigma_1 = 1)),
\end{aligned}
$$

where

$$
f(x) = \frac{1}{2}\left(\frac{1}{2} + \frac{x}{4} - q\right) - \left(1 - \frac{x}{2}\right)c.
$$

Thus the second-period payoff increases as well, proving the (1) of Theorem 4.

(2) Suppose $\overline{\lambda}^*(\beta) \le \lambda_1 < 2\overline{\lambda}/(1+\overline{\lambda})$. Consider first the DM's payoff in the absence of fact-checking. In this case, his first-period payoff is the same as he earn in the clickbait equilibrium. Thus, his first-period payoff is $f(\lambda_1)$. For period 2, he enters the second period with reputation $\lambda_2 = \lambda_1$ after $(m_1 = 1, \sigma_1 = 1)$, which occurs with probability $\Pr(m_1 = 1, \sigma_1 = 1) = 1/2$, while with $\lambda_2 = 0$ otherwise. Thus, his second-period payoff is $f(\lambda_1)/2$.

Consider now the DM's welfare in the presence of fact-checking. In this case, since uninformative equilibrium is played in period 1, his first-period payoff is 0, but he enters the second period with $\lambda_2 = \lambda_1$ with certainty so that he earns the second-period payoff of $f(\lambda_1)$ with certainty. Comparing $f(\lambda_1) + f(\lambda_1)/2$ with $f(\lambda_1)$ yields that DM's payoff is higher in the absence of fact-checking.

(3) There are two cases to consider: when $\lambda_1 \ge 2\overline{\lambda}/(1+\overline{\lambda})$ and when $\lambda_1 < \max\{\underline{\lambda}^*(\beta), \underline{\lambda}^{**}(\beta)\}$.

Suppose first that $\lambda_1 \ge 2\overline{\lambda}/(1+\overline{\lambda})$. In both settings, informative equilibrium with $(\mu_{\mathrm{G}}^*, \mu_{\mathrm{O}}^*) = (1/4, 0)$ is played. Therefore, the first-period payoff is the same in both settings. Thus, consider the second-period payoff. In the absence of fact-checking, the DM enters the second period with $\lambda_2 = \lambda_1$ after $(m_1 = 1, \sigma_1 = 1)$, and with $\lambda_2 = 0$ otherwise. In the presence of fact-checking, λ_2 depends on the outcome $o_1 \in \mathcal{M} \cup \mathcal{M} \times C$ as well as the result of fact-checking γ. However, in equilibrium, λ_2 depends on γ only after $o_1 = (m_1 = 1, \sigma_1 = 1)$, since otherwise, he can perfectly infer that the outlet is the opportunistic type only by seeing either the headline or the content, $\lambda_2 = 0$. However, since $\lambda_2(m_1 = 1, \sigma_1 = 1, \gamma) > \overline{\lambda}$ for all γ, meaning that clickbait equilibrium will be played regardless of the result of fact-checking, his expected second-period payoff evaluated at the timing $(m_1 = 1, \sigma_1 = 1)$ is convex combination:

$$\begin{aligned}\sum_{\gamma} \Pr(\gamma \mid m_1 = 1,\, \sigma_1 = 1)\, V_{\mathrm{G}}(\lambda_2(1,1,\gamma)) &= \sum_{\gamma} \Pr(\gamma \mid m_1 = 1,\, \sigma_1 = 1)\, f(\lambda_2(1,1,\gamma)) \\ &= f(\lambda_2(m_1 = 1,\, \sigma_1 = 1)) \\ &= V_{\mathrm{G}}(\lambda_2(m_1 = 1,\, \sigma_1 = 1))\,.\end{aligned}$$

Thus, DM's second-period payoff is the same in both settings.

Analogous argument can prove the case for $\lambda_1 < \max\{\underline{\lambda}^*(\beta), \underline{\lambda}^{**}(\beta)\}$. □

References

Allcott, Hunt, and Matthew Gentzkow. 2017. "Social Media and Fake News in the 2016 Election." *Journal of Economic Perspectives*, 31(2): 211–36.

Argenziano, Rossella, Sergei Severinov, and Francesco Squintani. 2016. "Strategic Information Acquisition and Transmission." *American Economic Journal: Microeconomics*, 8(3): 119–55.

Balbuzanov, Ivan. 2019. "Lies and consequences." *International Journal of Game Theory*, 48(4): 1203–40.

Bergemann, Dirk, and Stephen Morris. 2019. "Information Design: A Unified Perspective." *Journal of Economic Literature*, 57(1): 44–95.

Florian, Ederer, and Min Weicheng. 2021. "Bayesian Persuasion with Lie Detection." *mimeo*.

Hughes, Heather C., and Israel Waismel-Manor. 2021. "The Macedonian Fake News Industry and the 2016 US Election." *PS: Political Science & Politics*, 54(1): 19–23.

Kamenica, Emir, and Matthew Gentzkow. 2011. "Bayesian Persuasion." *American Economic Review*, 101(6): 2590–2615.

Levkun, Aleksandr. 2022. "COMMUNICATION WITH STRATEGIC FACT-CHECKING." *mimeo*.

Luo, Zhaotian, and Arturas Rozenas. 2025. "Lying in persuasion." *Games and Economic Behavior*, 152: 93–112.

Morris, Stephen. 2001. "Political Correctness." *Journal of Political Economy*, 109(2): 231–265.

Pei, Harry Di. 2015. "Communication with endogenous information acquisition." *Journal of Economic Theory*, 160: 132–149.

Subramanian, Samanth. 2017. "Inside the Macedonian Fake-News Complex." https://www.wired.com/2017/02/veles-macedonia-fake-news/.

Suen, Wing. 2004. "The Self-Perpetuation of Biased Beliefs." *The Economic Journal*, 114(495): 377–396.

Tam, Yin Chi, and Hitoshi Sadakane. 2023. "Cheap Talk and Lie Detection." *mimeo*.

Wu, Jiemai, and John H. Nachbar. 2024. "Headlines as Truncated Experiments." *mimeo*.