The Role of Messenger in Advertising Content: Bayesian Persuasion Perspective

Jiwoong Shin,1,* Chi-Ying Wang1

1School of Management, Yale University, New Haven, Connecticut 06511
*Corresponding author

Contact: jiwoong.shin@yale.edu; https://orcid.org/0000-0001-8175-854X (JS); chi-ying.wang@yale.edu (C-YW)

Received: November 15, 2022
Revised: June 23, 2023; September 22, 2023
Accepted: November 4, 2023
Published Online in Articles in Advance: January 16, 2024

https://doi.org/10.1287/mksc.2022.0405

Abstract. We propose a model of advertising content that focuses on messenger selection, where advertising can generate product-match signals for consumers. We consider advertising as a problem of Bayesian persuasion with costly information processing, where the type of communication messenger is costly to observe and determines the information structure consumers will face, thereby affecting their attention decisions. Messengers are classified as high type or low type based on their likelihood of generating positive signals about product match. Our findings highlight that the optimal choice of messengers depends on their signal elasticities and the firm’s decision on whether to induce consumer attention. In particular, we find that when it is crucial to raise prices and high-type messengers overshadow the product match value by providing generally positive signals, a low-type messenger can effectively capture consumer attention and persuade them to pay a higher premium. This holds true even if high-type messengers can better grab consumers’ attention by providing additional entertainment value or when some consumers are naive in belief updating.

Keywords: advertising content • peripheral cues • messenger • dual-mode of communication • deliberation cost • Bayesian persuasion

1. Introduction

“We’ve said all along that the messenger can be as or more important than the actual message itself.”

—Lisa Sherman, the chief executive, Ad Council

Over the past 40 years, economists and marketers have extensively studied the role of advertising in educating consumers about products and persuading them to choose them over competitors. The success of an advertising campaign crucially depends not only on whether the message in its content can clearly communicate the product benefits to the consumers but also on how the message is delivered, taking into account factors such as style and messengers in the ad content. The style or design of an advertisement is often the initial attention-grabbing factor, thus making these choices influential in the campaign’s effectiveness.1 Additionally, the messenger, who delivers the advertising message, holds significant importance. It is well known that it is not just what is said (message) that matters, but who says it (messenger) matters too.2

Effective advertising campaigns carefully consider both the message and its execution through various content choices to clearly communicate product benefits and influence consumer behavior. This paper primarily focuses on exploring the choice of messengers as a central application of content selection. The messenger of communication plays a vital role in attracting the receiver’s attention and convincing the message, and it can influence the effectiveness of communication. Thus, how to deploy an effective messenger strategy in different contexts and whom to use to deliver the message is of great interest to both academics and practitioners.

The academic literature finds some evidence of the positive effects of using a specific type of messenger, focusing on what messenger traits account for the effectiveness of advertising campaigns. Many studies on the attractiveness model (Joseph 1982, Liu et al. 2007, Eisend and Langner 2010) documented that physically attractive messengers positively affect consumers’ attitudes, product evaluation, and purchase intention. Also, a stream of literature (Till and Busler 1998, Ohanian 1991, Schouten et al. 2020) suggests that the messenger’s perceived relevance to the advertised product is essential in affecting purchase behavior. Hence, using attractive messengers or those with substantial expertise in a relevant subject area appears to be a viable strategy for enhancing advertising effectiveness.

The industry practices are mixed. Traditionally, many advertising campaigns lean toward endorsements from experts and physically attractive actors or celebrities. For instance, Sensodyne often spotlights endorsements from
dentists or industry experts in its commercials. Similarly, beauty and cosmetics companies frequently enlist attractive models and celebrities to elevate the appeal and desirability of their products. Consider L’Oréal, which has featured celebrities such as Jennifer Aniston and Beyoncé in their advertising campaigns for beauty products. However, there is no unanimous consensus among marketers regarding the efficacy of idealized or expert endorsers in advertising. In 2004, Unilever pioneered several Real Beauty campaigns, featuring everyday women rather than professional models. Abercrombie & Fitch also embarked on a “Face Your Fierce” campaign in 2020, showcasing a diverse range of real individuals.

Although the success of these campaigns may, in part, stem from their provocative novelty (Vézina and Paul 1997), the advertising industry has witnessed the rise of ads featuring real people for their products, despite the diminished novelty compared with earlier years. For instance, Stitch Fix has adopted a social media advertising strategy that spotlights “real” or “ordinary” individuals instead of relying on conventional models. Nevertheless, although the trend toward more realistic advertisements continues to grow, it is noteworthy that approximately 20% of television ads still feature physically attractive celebrities (Dhotre and Bhola 2010). Furthermore, companies like H&M persist in using attractive models to promote their products on platforms such as Instagram (Figure 1).4

This raises an important question: which types of messengers are more effective in an advertising campaign, and under what conditions? To ensure the success of an advertising campaign, it is crucial to have a deeper understanding of how messengers contribute to persuasion. Our study endeavors to explore the effectiveness of advertising messengers by investigating the process of advertising persuasion. We aim to determine the extent to which messengers can influence consumer attention and purchase decisions. Specifically, we analyze the impact of different messengers, such as attractive models versus everyday individuals, and identify the critical variables associated with their effectiveness using a Bayesian persuasion framework.

We adopt a dual-mode communication perspective in analyzing advertising communication. In this framework, a messenger acts as a cue that is communicated through a less resource-intensive peripheral route, whereas the central issue-relevant information requires more thoughtful deliberation. Consumers can costlessly observe the advertising messenger. However, they only receive a binary signal about their match value (either a good or bad signal) only if they incur a deliberation cost to pay attention to the ad. We first analyze the consumer inference process upon receiving a private signal from an ad featuring different types of messengers. Although the type of messenger does not change the message conveyed by the ad, it determines the informativeness of the signal. Consumers update their beliefs in a Bayesian manner upon receiving a signal from the advertisement. For instance, an attractive messenger may make a product look more appealing yet fail to convince the audience of its product-match values because consumers may be confused about whether the positive appeal is due to the messenger or reflects the product’s genuine underlying value.

To comprehend the intricate relationship between the advertising messenger and persuasion, we approach advertising as a problem of Bayesian persuasion. In this framework, the type of messenger chosen by the firm determines the information structure the consumer will encounter. Specifically, because the product-match values are ex ante unknown to all parties and the firm has no additional private information in our model, the firm can commit to a particular signal structure by selecting the type of messenger, making advertising a Bayesian persuasion device.5 We then delve into investigating the mechanisms through which messengers attract consumer attention and potentially persuade them to pay a higher price for the product. We begin by examining the consumer’s attention decision, which is based on the price at which the benefits of paying attention outweigh the costs of deliberation. We identify the specific

Figure 1. (Color online) Examples of Stitch Fix and H&M Ads on Instagram

Notes. (a) Stitch Fix. (b) H&M.
conditions under which different types of messengers can more effectively induce consumer attention. In particular, we find that low-type messengers can be more effective in attracting consumer attention only when the price of the product is sufficiently high. In such cases, the good signals transmitted by a low-type messenger become more informative due to their lower likelihood of producing such signals. Consequently, despite the lower probability of generating positive signals, the higher expected utility of purchasing the product leads consumers to place greater value on the messages delivered by low-type messengers. Furthermore, we demonstrate the circumstances under which various types of messengers can command higher price premiums both when consumer attention is present and when it is not.

Next, we compare the expected profits to determine the optimal decision for the firm regarding whether to induce consumer attention or not. When making this decision, the firm must take into account both the necessity and the ability to increase the price, effectively balancing sales volume and price premium. Our findings reveal that the need to raise the price becomes more significant as the production cost rises. Consequently, the firm tends to sacrifice demand in order to increase its price. Conversely, when the consumer’s deliberation cost increases, the firm’s ability to charge a higher price without inducing consumer attention also increases. This entices the firm to forgo the profit margin to maintain greater demand. Hence, inducing consumer attention is deemed optimal when the consumer’s deliberation cost decreases or when the production cost increases.

After establishing the optimal decision to induce consumer attention, we demonstrate that the optimal choice of messenger depends on each messenger’s signal elasticity (or informativeness) and the firm’s decision regarding inducing consumer attention. When a high-type messenger’s signals are highly informative, an advertisement featuring a high-type messenger is more effective in attracting consumer attention and increasing the price. However, because of the high informativeness of the high-type messenger’s signals, the firm must set a very low price to discourage consumer attention. Consequently, if the firm chooses not to induce consumer attention, which is the case when the production cost is low or the deliberation cost is high, it selects a low-type messenger. On the other hand, when an advertisement featuring a high-type messenger tends to generate more good signals, a high-type messenger’s signals become less informative. In this case, the high-type messenger effectively overshadows the product and limits the firm’s ability to update consumers’ beliefs about the product match despite good signals that consumers are more likely to receive. In contrast, because of its rarity, the low-type messenger’s good signals have a substantial impact on updating consumers’ beliefs. Therefore, in the case where the low-type messenger’s signals are sufficiently informative, and the production cost is sufficiently high, the firm finds it optimal to use the low-type messenger to attract consumer attention and, more importantly, persuade them to pay a higher premium. This general insight can still carry over even if a high-type messenger can better grab consumers’ attention by providing additional entertainment value or when some consumers are naive in updating their beliefs. Although the optimal range for using a low-type messenger significantly narrows when we consider these effects, a low-type messenger can still serve as a more effective medium for advertisers, particularly when the objective of raising the price holds substantial importance, but the high-type messenger’s informativeness is diminished, thereby overshadowing the product’s characteristics.

The significance of ad content is widely recognized, but how it influences consumer persuasion remains elusive. In this study, we propose one possible mechanism to demonstrate how ad content, especially the messenger, matters in persuading consumers. Thus, this paper’s main contribution lies in providing a new framework to consider advertising—we posit that ad content can be thought of as information structure and leverage Bayesian persuasion as a tool to model advertising content and the role of the messenger.

The paper is organized as follows. The next section reviews the related literature. Section 3 describes a model that characterizes the role of communication messengers in advertising. Then, we first highlight the key intuitions and mechanism of the paper with an illustrative example in Section 4 and present the main analyses and results focusing only on the persuasion role of the messenger in Section 5. In Section 6, we demonstrate the robustness of the main insights by analyzing a number of extensions that relax several key model assumptions. Section 7 concludes.

2. Literature Review

This paper relates to several streams of research: advertising content, persuasion, and information design. First, our research contributes to the burgeoning area of advertising content. The content of advertising provides direct information, such as the existence of the product or its price (Butters 1977, Iyer et al. 2005, Shin 2005). The information can also be indirect, where the mere fact that the firm advertises signals an experience good’s quality (Nelson 1974, Milgrom and Roberts 1986). The latter is known as the “money-burning” theory of advertising. The central argument of the money-burning theory suggests that the level of spending, not the advertising content, signals the quality of the product. That is, content is irrelevant to communicating information about product quality. However, this view has been challenged by a number of recent papers, which explore the role of advertising content in conveying information in a
rational equilibrium framework (Anderson and Renault 2006, Mayzlin and Shin 2011). Anderson and Renault (2006) study the optimal amount of information in advertising content, allowing firms to provide both price and match information. Mayzlin and Shin (2011) show how providing product attributes information in ad content, along with price, can signal product quality when the bandwidth of advertising messages is limited, and consumers can conduct their own search. Our paper contributes to this literature by analyzing the role of advertising content, especially the messenger, under the Bayesian persuasion framework.

For the role of advertising messenger, a large body of research finds several key characteristics of messengers that impact the effectiveness of persuasion, such as attractiveness, relevance, credibility, and familiarity (Zajonc 1968, Chaiken and Maheswaran 1994, Till and Busler 1998). In the attractiveness literature, Baker and Churchill (1977) demonstrate that consumers’ affective attitudes (the appeal and ad impression) are consistently higher when ads feature more attractive messengers. Nonetheless, they also note an ambiguity in the effects on behavioral attitudes (willingness to try/buy the product). Indeed, other studies in the literature yield mixed results regarding the outcomes of messenger attractiveness in advertising, ranging from positive (Kahle and Homer 1985, Till and Busler 1998, Liu et al. 2007) to mixed results, including insignificant or negative effects (Baker and Churchill 1977, Joseph 1982, Kamins 1990). Building on these findings, our model assumes that featuring attractive celebrities or models in ads generates more positive ad impressions (higher affective attitude or “signals”), whereas the behavioral outcomes depend on the messenger’s signal structure. Also, this assumption can be applied beyond the context of messenger attractiveness as long as some types of messengers (e.g., experts) can always generate more positive signals than others. Our framework allows us to rationalize the diverse findings from previous studies by analyzing how messengers influence the persuasiveness of advertising communication through consumer inference.

Our study also incorporates the dual mode of communication perspective for persuasion in the psychology literature (Petty and Cacioppo 1986, Chaiken and Maheswaran 1994, Kahneuman 2011). This perspective suggests that the decision-makers process information through two different routes (the central and the peripheral route) with different amounts of cognitive loads for processing information. The central route involves issue-relevant information that is more cognitively demanding to process and can occur only when one is willing and able to devote substantial mental resources to the message (Petty and Cacioppo 1986). Conversely, the peripheral route often entails issue-irrelevant cues and requires little to no cognitive effort from the decision-makers. The implications of consumers’ costly information processing or deliberation costs on firms’ decision making have been extensively studied in various papers. These studies examine how factors such as pricing, quality signaling, product line decisions, and optimal information disclosure are affected by consumers’ information processing costs (Wathieu and Bertini 2007, Kuksov and Villas-Boas 2010, Guo and Zhang 2012, Guo 2016, Guo and Wu 2016, Lu and Shin 2018, Li et al. 2019). In the context of advertising, researchers have explored the optimal design of media advertising formats, taking into account consumers’ incentives to pay attention to advertisements in the presence of costly attention or opportunity costs (Dukes et al. 2021, Lin 2022). These studies focus on understanding how advertising can effectively capture consumers’ attention despite the potential costs associated with information processing. Our study builds on these streams of existing literature. We recognize that consumers can costlessly observe the advertising messenger through the peripheral route. However, assessing the product-match signal requires more effort and careful consideration through the central route, incurring deliberation costs to process the message content. By integrating these two modes, we explore how costly deliberation influences consumers’ responses to advertising messages and how firms can strategically use peripheral cues (such as messengers) and product-match signals to persuade consumers.

Finally, the problem we study is closely related to the Bayesian persuasion and information design literature (Gentzkow and Kamenica 2016, Gentzkow and Kamenica 2017, Bergemann and Morris 2019, Guo 2022, Iyer and Zhong 2022), which seeks to identify the optimal information environment to affect the receiver’s decisions through influencing the posterior beliefs about the state of the world. The seminal paper by Kamenica and Gentzkow (2011) considers a model with symmetric information where a sender can only affect a receiver’s action by choosing and committing to a particular information structure. In recent papers, Iyer and Zhong (2022) study a firm’s optimal information notification design to maximize consumer engagement using an information design framework. Guo (2022) also applies the Bayesian persuasion framework in the collaborative customization setting, focusing on the effect of the choice of information structure on customers’ engagement decisions in the customization process. Additionally, several papers have explored scenarios involving different information provision or processing costs within the Bayesian persuasion framework. Nguyen and Tan (2021) investigate cases where sending messages incurs costs for the sender, whereas Lipnowski et al. (2020) focus on situations where it is costly for the receiver to process information. In our study, we view advertising as a Bayesian persuasion device, where the firm can strategically select a specific information structure by choosing a messenger.
with specific characteristics. The messenger influences the information structures that generate signals about the match values. Unlike many existing papers in this literature that solely optimize the information structure, we examine a setting where the firm simultaneously determines both the information structure and the product’s price to maximize its expected profit.

3. Model

3.1. Strategic Players and Information Environment

We consider a market with a monopolistic firm that sells a single product with a constant production cost $k$ to a unit mass of consumers. Consumers are unaware of the product and can buy it only if they receive an ad. Each consumer has unit demand, and consumer $i$ can obtain a utility of $v_i$ by consuming the product, where $v_i$ is the individual-specific product match value drawn from a distribution $F[0, \sigma]$ with density $f(\cdot)$. Both the probability distribution of the product match $F$ and the production cost $k$ are common knowledge, and we assume that the production cost $k$ is less than an average consumer’s consumption utility such that $k < E(v)$.

Upon deliberation, advertising yields a binary product-match signal $s_i \in \{s_g, s_b\}$, where $s_g$ denotes good news, and $s_b$ denotes bad news about the match. In this setting, the signal relates to the horizontal match information, which exhibits a symmetric nature. This implies that neither the firm nor the consumers possess additional information regarding the specific match value between the product and each individual consumer. Take fashion brands like Zara or H&M advertising a new product, for instance. Although the brand’s reputation may establish the product’s quality, the crucial factor lies in the product’s fit or match. However, because the match value is consumer specific, neither party possesses any extra information regarding the product’s match value for an individual consumer. Thus, we can consider the role of the advertisement as a Bayesian persuasion device (Kamenica and Gentzkow 2011). That is, consumers update their beliefs in a Bayesian fashion upon seeing the signal from the ad instead of updating beliefs based on the firm’s equilibrium advertising strategies. Next, we focus on the characteristics of messengers, which can influence the realization of signals and consumer inference.

3.1.1. Role of Communication Messengers. Who delivers the message is important in attracting attention and convincing the receiver about the product match. Specifically, some characteristics of the messenger, such as attractiveness, credibility, and expertise, can influence the effectiveness of an ad (Till and Busler 2000, Dhar and Stillman 2019). We postulate that a messenger’s type (either a high-type $m_H$ or a low-type $m_L$) can probabilistically affect the realization of the signal that consumers obtain from the ad (Baker and Churchill 1977). Formally, given the true product match value $v$ and the messenger $m$, consumers obtain private signals upon paying attention to the ad, with the following probability:

$$\sigma_j(v) = \Pr(s_i = s_j | v, m_j), \quad \text{where } j \in \{L, H\}. \quad (1)$$

The firm knows that the consumer can obtain the signal $s_i$ with the previous probabilities but does not observe what signal a consumer ultimately receives. However, the signal structure consumers encounter is affected by the messenger’s type as follows.

**Assumption.** The probability of receiving a good signal $\sigma_L(v)$ is increasing in the true match value $v$ and the messenger’s type $m_L$: $\frac{\partial \sigma_L}{\partial v} \geq 0$, and $\sigma_H(v) > \sigma_L(v)$ for all $v$.

In broad terms, there are two key characteristics defining a signal structure: the intercept, closely associated with the signal’s overall positivity, and the slope, which governs its sensitivity or responsiveness to the true match value. The assumption implies that individuals with higher true match values are more inclined to receive positive signals (indicated by positive slopes). Consequently, a positive signal, denoted as $s_g$, essentially conveys “good news” regarding the true match value (Milgrom 1981). Furthermore, a high-type messenger possesses a greater propensity to generate positive signals compared with a low-type messenger across all levels of the true value $v$ (illustrated by the high-type messenger’s signal structure having a larger intercept). It is important to emphasize that this assumption specifically centers on high-type messengers being more adept at generating private positive signals related to product match values. However, it does not necessarily imply that high-type messengers excel at persuading or convincing consumers to make a purchase. For a more comprehensive understanding of when and how various messenger types can be more effective in influencing consumer behavior, the Bayesian persuasion framework provides valuable insights.

Let us define the average probability of receiving a good signal from a particular type of messenger as

$$\bar{\sigma}_H \equiv \Pr(s_i = s_g | m_H) = \int_0^\sigma \sigma_H(v) dF,$$

$$\bar{\sigma}_L \equiv \Pr(s_i = s_g | m_L) = \int_0^\sigma \sigma_L(v) dF. \quad (2)$$

Then, $\bar{\sigma}_H > \bar{\sigma}_L$. We use these two terms $\bar{\sigma}_H, \bar{\sigma}_L$ in our subsequent analysis.

A messenger’s type may determine how easily the messenger can command attention and motivate people to think about the information in the advertising message. In particular, high-type messengers can lower the deliberation cost consumers incur to pay attention to the ad. For example, the firm can feature a celebrity to provide consumers with additional entertaining value to
offset the cost. Formally, the deliberation cost can be a function of a messenger’s type, $c = c(m)$, where $0 \leq c(m_H) < c(m_L)$. However, we first focus on the relationship between a messenger’s type and the signal structure in the main model by fixing the deliberation cost constant for both types of messengers ($c(m_H) = c(m_L) = c$) to convey the insights cleanly. We later relax this assumption by incorporating the effect of a messenger’s type on the deliberation cost to capture the messenger’s attention-grabbing role. Moreover, the deliberation cost also plays a significant role in influencing the effectiveness of different types of messengers in commanding prices and the optimal decision making by firms to attract consumer attention. We will delve into a more in-depth analysis of these aspects in the subsequent sections of our study.

### 3.1.2. Dual Mode of Communication

Consumers exhibit varying cognitive loads when processing different types of information in advertising. In this context, we adopt a dual mode of communication perspective from the model proposed by Petty and Cacioppo (1986). This perspective posits that information is conveyed through two distinct routes. First, there is a less costly peripheral route delivering issue irrelevant cues such as the expertise or attractiveness of the messenger (Petty and Cacioppo 1986). Because messenger’s attractiveness or expertise might be the first thing consumers notice even before they start thinking about the product information. Such ostensible information about the messenger can be communicated through mindless processing (Petty and Cacioppo 1986, Kahneman 2011). In contrast, the second route, known as the central route of persuasion, demands more cognitive resources. It involves a thoughtful evaluation of advertising content, including its personal relevance to a product and the product’s match value (Wathieu and Bertini 2007, Guo and Zhang 2012, Guo 2016). Therefore, detailed information about the actual product would only be examined when consumers pay close attention to the ad.

Here, we assume that consumers can costlessly observe the advertising messenger. In contrast, consumers can receive an issue-relevant signal through the central route of communication only if they decide to pay attention to the ad by incurring costly deliberation costs, $c > 0$. For example, when an ad appears on TV, consumers can mindlessly see the ad delivered to their eyes and recognize the appearance of a messenger in it. Then, consumers can decide whether to pay attention by incurring deliberation costs to process the central information in the ad, thereby receiving a private signal about the product match.

### 3.2. Sequence of the Game

The firm’s action space consists of its choice of the product price, $p_j$, and the type of advertising messenger, $m_j \in \{m_L, m_H\}$. Consumers’ strategies are deciding 1) whether to pay attention to the advertising message after observing the messenger and 2) whether to buy a product.

The timeline of the game is summarized in Figure 2. In stage 0, Nature draws each consumer $i$’s product match value $v_i$. In stage 1, the firm $j$ chooses the type of messenger $m_j \in \{m_L, m_H\}$ and a price $p_j$ to maximize its expected profit. In stage 2, consumers decide whether to pay attention to the ad upon seeing the messenger $m_j$ and the price $p_j$ chosen by the firm. If a consumer decides to pay attention, she incurs a deliberation cost $c$ and receives a noisy private signal $s_i \in \{s_j, s_k\}$ about the product match. The firm knows that the consumer will receive a private good signal about the product match with probabilities $\sigma_j(v)$ but does not observe whether consumers engage with the advertising by incurring deliberation costs, let alone the specific signals received should they choose to do so. Finally, consumers decide whether to purchase the product based on the price and their updated beliefs about the product match value in stage 3. We summarize all the basic notations in this paper in the appendix (Table A.1).

### 4. Numerical Examples: An Illustration

Let us first consider a simple numerical example to illustrate our model and its underlying mechanism. The firm sells a product at a price $p$ at cost $k$ to consumers whose true match values follow a binary distribution $Pr(v_i = 1) = Pr(v_i = 0) = 1/2$. The firm can feature either a celebrity (high-type, $m_H$) or a normal person (low-type, $m_L$) in an ad.

Now, let us consider two scenarios where each messenger $j$ has varying probabilities, denoted as $\sigma_j$, of generating a favorable signal $s_j$ as shown in Table 1. In both cases, the high-type messenger is more likely to generate a good signal than a low type in either state ($\sigma_H(v_i) > \sigma_L(v_i)$).
\(\sigma_L(v_i)\) for \(v_i = 0, 1\), and both messengers are more likely to generate a good signal when the true match is good (\(\sigma_L(v_i = 1) > \sigma_L(v_i = 0)\) for \(j \in \{L, H\}\)). However, in case 1, the high-type messenger’s signal demonstrates greater sensitivity to changes in the true match state compared with the low type (\(\sigma_H(v_i = 1) - \sigma_H(v_i = 0) = 0.4 > \sigma_L(v_i = 1) - \sigma_L(v_i = 0) = 0.1\)), whereas case 2 represents a scenario where the low type’s signal is more sensitive (\(\sigma_H(v_i = 1) - \sigma_H(v_i = 0) = 0.1 < \sigma_L(v_i = 1) - \sigma_L(v_i = 0) = 0.4\)).

Additionally, consumers incur a deliberation cost of \(c = 0.02\) to receive the signal. Consumers face a choice regarding their attention to the ad when they receive an ad. If consumers opt not to pay attention, their expected utilities are either \(E(v_i) - p = 0.5 - p\) or zero, selecting the higher of the two depending on the price. In contrast, paying attention to the ad commits them to buy the product upon a good signal and abstain from making a purchase upon a bad signal. The expected probability of receiving a favorable signal is denoted as \(\tilde{\sigma}_j\) and can be calculated as follows: \(\tilde{\sigma}_j = 0.5 \cdot \sigma_j(v_i = 1) + 0.5 \cdot \sigma_j(v_i = 0)\), as both states have equal probabilities of 1/2. Furthermore, when consumers pay attention, they update their belief about the true match value upon receiving a favorable signal using Bayes’ rule:

\[
E(v_i | s_k, m_i) = \frac{0.5 \cdot \sigma_L(v_i = 1)}{0.5 \cdot \sigma_L(v_i = 1) + 0.5 \cdot \sigma_L(v_i = 0)}
\]

In Case 1, we have \(\tilde{\sigma}_H = 0.4, \tilde{\sigma}_L = 0.15,\) and \(E(v_i | m_{H}, s_k) = \frac{0.5 \cdot 0.6}{0.5 \cdot 0.6 + 0.5 \cdot 0.2} = 3/4, E(v_i | m_{L}, s_k) = \frac{0.5 \cdot 0.6}{0.5 \cdot 0.6 + 0.5 \cdot 0.2} = 2/3\). Then, we can evaluate consumers’ expected utility when they choose to pay attention to an ad. When the ad features a celebrity, their expected utility is given by

\[
EU = \tilde{\sigma}_H \cdot (E(v_i | m_{H}, s_k) - p) - c = 0.4 \cdot \left(\frac{3}{4} - p\right) - 0.02 = 0.28 - 0.4 \cdot p.
\]

Conversely, when the ad features a normal person, the expected utility becomes

\[
EU = 0.15 \cdot \left(\frac{2}{3} - p\right) - 0.02 = 0.08 - 0.15 \cdot p.
\]

Comparing these expected utilities with those without paying attention, we find that consumers pay attention to the ad featuring a celebrity if and only if \(0.28 - 0.4p \geq \max(0.5 - p, 0) \iff 0.37 < p < 0.7\). In this case, the firm can adopt two pricing strategies to maximize its profit. It can set the price at 0.37, encouraging all consumers to purchase the product without paying attention to the ad, resulting in a profit of \(\Pi_{H0}^L = 0.37 - k\), where the superscript no denotes no consumer attention. Alternatively, the firm can set the price at 0.7, ensuring that all consumers pay attention to the ad and purchase the products only if their private signals are good. This strategy yields an expected profit of \(\Pi_{H0}^H = 0.4(0.7 - k) = 0.28 - 0.4k\), where the superscript aft denotes with consumer attention.

For the scenario in which consumers pay attention to the ad featuring a normal person, it occurs if and only if \(0.08 - 0.15p \geq \max(0.5 - p, 0) \iff 0.49 < p \leq 0.53\). In this case, the firm can choose to set the price at 0.49, prompting all consumers to purchase the product without paying attention to the ad, resulting in a profit of \(\Pi_{H0}^L = 0.49 - k\). Alternatively, the firm can set the price at 0.53, ensuring that all consumers pay attention to the ad and purchase the products only if their private signals are good, with an expected profit of \(\Pi_{H0}^H = 0.15(0.53 - k) = 0.08 - 0.15k\). It’s noteworthy that featuring a normal person can command a higher price compared with featuring a celebrity when discouraging consumer attention (0.49 versus 0.37) but not when encouraging consumer attention (0.53 versus 0.7).

The firm’s profit hinges on the production cost, denoted as \(k\). For our analysis, we examine two production cost levels: high \((k = 0.4)\) and low \((k = 0.1)\). Using these cost values, we can calculate the expected profits associated with using various messenger types to either induce or discourage consumer attention. The results are summarized in Table 2. Notably, when production costs are high, we find that consumers prefer the ad featuring a celebrity at a price of 0.37, whereas they prefer the ad featuring a normal person at a price of 0.49 when production costs are low.

### Table 1. Two Different Signal Structures

<table>
<thead>
<tr>
<th>Case 1 (H-type signal is more sensitive)</th>
<th>True match value</th>
<th>Low-type’s (\sigma_L)</th>
<th>High-type’s (\sigma_H)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(v_i = 0)</td>
<td>0.1</td>
<td>0.2</td>
<td></td>
</tr>
<tr>
<td>(v_i = 1)</td>
<td>0.2</td>
<td>0.6</td>
<td></td>
</tr>
<tr>
<td>Case 2 (L-type signal is more sensitive)</td>
<td>(v_i = 0)</td>
<td>0.4</td>
<td>0.8</td>
</tr>
<tr>
<td>(v_i = 1)</td>
<td>0.8</td>
<td>0.9</td>
<td></td>
</tr>
</tbody>
</table>

### Table 2. Comparison of Firm’s Expected Profits

<table>
<thead>
<tr>
<th>Production cost</th>
<th>Consumer attention</th>
<th>Low-type</th>
<th>High-type</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>Yes</td>
<td>0.02</td>
<td>0.12</td>
</tr>
<tr>
<td>((k = 0.4))</td>
<td>No</td>
<td>0.09</td>
<td>-0.03</td>
</tr>
<tr>
<td>Low</td>
<td>Yes</td>
<td>0.065</td>
<td>0.24</td>
</tr>
<tr>
<td>((k = 0.1))</td>
<td>No</td>
<td>0.39</td>
<td>0.27</td>
</tr>
<tr>
<td>(b)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>Yes</td>
<td>0.14</td>
<td>0.09</td>
</tr>
<tr>
<td>((k = 0.4))</td>
<td>No</td>
<td>-0.1</td>
<td>0.13</td>
</tr>
<tr>
<td>Low</td>
<td>Yes</td>
<td>0.32</td>
<td>0.35</td>
</tr>
<tr>
<td>((k = 0.1))</td>
<td>No</td>
<td>0.2</td>
<td>0.43</td>
</tr>
</tbody>
</table>

Notes. (a) Firm’s expected profits: Case 1. (b) Firm’s expected profits: Case 2.
opting for a celebrity as the messenger is the optimal choice to attract consumer attention. Conversely, when production costs are low, using a regular person as the messenger is ideal for discouraging consumer attention.

In the subsequent Case 2, we repeat the same procedure outlined in Case 1 to determine the firm’s expected profits, as displayed in Table 2(b). In contrast to Case 1, we observe that in Case 2, a normal person’s signals exhibit a higher sensitivity to true match states \((\sigma^*_T (v) = 1) - \sigma^*_L (v) = 0.4\) compared with a celebrity \((\sigma^*_H (v) = 1) - \sigma^*_H (v) = 0.1\). Remarkably, this reversal in signal sensitivity leads to different optimal messenger choices. Specifically, when production costs are high, it is advantageous for the firm to use a regular person to attract consumer attention, whereas when production costs are low, featuring a celebrity is the optimal strategy to discourage consumer attention.

5. Analysis

In this section, we generalize the analyses from Section 4. As demonstrated in the illustrative examples, when the firm evaluates a messenger’s effectiveness, it compares the expected profits from different messenger types, which are affected by (1) the signal structure generated by a messenger, (2) the messenger’s ability to attract consumer attention, and (3) the potential price premium it can command based on updated beliefs of consumers after deliberating the ad. We start by analyzing consumers’ inference process upon receiving a private signal. Subsequently, we study the mechanism of how a messenger can attract consumer attention and potentially raise price premiums. In our main model, we maintain a fixed level of messenger attraction ability, which directly impacts consumers’ deliberation costs. Nonetheless, as detailed in Section 5.2, the messenger’s type can still rationally influence consumers’ attention decisions by shaping their inferences. In our extension, we introduce a direct attention-grabbing effect by incorporating variations in the messenger’s attraction ability. After analyzing consumers’ inference processes and attention decisions, we turn to the firm’s decisions about pricing and optimal choice of messenger in equilibrium. Our solution concept is subgame perfect equilibrium, and all proofs can be found in the appendix.

5.1. Consumer Inference

Consumers update their beliefs in a Bayesian fashion upon receiving a signal from the ad (conditional on paying attention to it). Formally, upon receiving a private signal \(s_j\) a consumer’s posterior beliefs about the underlying match value distribution follow the next two densities\(^{16}\):

\[
f(v|m_j, s_j) = \frac{f(v) \cdot \sigma_j(v)}{\Pr(s_i = s_j | m_j)} = \frac{f(v) \cdot \sigma_j(v)}{\sigma_j}, \quad (3)
\]

\[
f(v|m_j, s_b) = \frac{f(v) \cdot [1 - \sigma_j(v)]}{\Pr(s_i = s_b | m_j)} = \frac{f(v) \cdot [1 - \sigma_j(v)]}{1 - \sigma_j}. \quad (4)
\]

To compare the informativeness of the signal from different messengers, we first establish the following proposition, which connects consumers’ posterior beliefs to the different signal structures generated by different messengers.

**Proposition 1.** Upon receiving a good signal \(s_j\), the posterior belief generated from the high-type messenger \(m_H\) satisfies the MLRP with respect to the low-type messenger \(m_L\), that is, \(f(v|m_H, s_j)\) is increasing in \(v\) if and only if, for all \(v\),

\[
\epsilon_H(v) = \frac{d \sigma_H(v)}{d \sigma_H(v)/v} = \frac{d \sigma_L(v)}{d \sigma_L(v)/v} = \epsilon_L(v). \quad (5)
\]

Conversely, if \(\epsilon_H(v) < \epsilon_L(v)\) for all \(v\), the posterior belief from the low-type messenger \(m_L\) satisfies the MLRP with respect to the high-type messenger \(m_H\); \(f(v|m_H, s_j)\) is decreasing in \(v\).

The elasticities \(\epsilon_H(v)\) and \(\epsilon_L(v)\) measure how sensitive a messenger’s good signal is to changes in match values, and their relationship determines the informational environments that consumers face.\(^{17}\) Figure 3 demonstrates two possible cases suggested by the lemma. It is important to note that both cases depicted in the figures align with our assumption: (1) the probability of receiving a good signal always increases in \(v\), \(\frac{d \sigma_j}{d v} \geq 0\), and (2) \(\sigma_H(v) > \sigma_L(v)\) for all \(v\). Therefore, receiving a good signal \(s_j\) always improves the posterior belief about the product match value \(f(v|m_H, s_j)\) regardless of the messenger’s type. Moreover, one may intuit that a good signal from a low-type messenger may update consumers’ posterior beliefs about the match value more simply because it is less likely to generate a good signal (i.e., \(\sigma_H(v) > \sigma_L(v)\) for all \(v\)). However, this lay intuition overlooks a more nuanced consumer inference process. In fact, the impact of messengers’ signals on posterior beliefs is contingent on the interplay between the two elasticities of signal structures.

First, when \(\epsilon_H(v) \geq \epsilon_L(v)\) (as shown in Figure 3(a)), the likelihood ratio of generating a good signal between high and low-type messengers actually increases as \(v\) increases. Moreover, the signals from different messengers become more divergent only when the product match reaches a sufficiently large value. Notably, in the case of low-type messenger ads, the likelihood of generating a good signal remains relatively similar across all match values \(v\). This can occur when the intercepts of different messengers’ signal structures are similar, allowing more room for improvement in the high-type messenger’s signal along the matching value. Consequently, although a good signal positively updates the consumer belief about the match value, it is not so informative about the match value of \(v\). In contrast, receiving a good signal \(s_j\) from a high-type messenger signifies unequivocally positive news regarding the potential match value. This is because the probability of generating a good signal, \(\sigma_H(v)\), is highly sensitive to the match value \(v\). A high-type messenger is more likely to produce a good signal only under high match values. To illustrate this
extreme scenario, refer to Figure 4(a), where \( \sigma_L(v) = 0 \) for all \( v \). In such an information environment, signals from low-type messengers provide no informative insights regarding the true match value.

Second, when \( \varepsilon_H(v) < \varepsilon_L(v) \) (as depicted in Figure 3(b)), the pattern is reversed. In this case, the likelihood ratio of generating a good signal between high- and low-type messengers actually decreases as \( v \) increases. Consequently, the signals from different messengers exhibit significant divergence only when the product match value is sufficiently small. This can happen when the intercepts of the messengers’ signal structures are quite different, with the high-type messenger’s intercept being significantly higher, limiting the room for improvement for its signal. Under this information environment, high-type messengers tend to produce predominantly positive signals irrespective of the match value, whereas low-type messengers’ signals are highly sensitive to the true match value \( (\varepsilon_H(v) < \varepsilon_L(v)) \). Therefore, if a consumer receives a good signal from a low-type messenger, it is highly indicative of a high match value. As a result, receiving a good signal \( s_g \) from a low-type messenger becomes more discriminatory and informative than a signal from a high-type messenger, aiding in resolving uncertainties in this information environment. Once again, consider Figure 4(b), where \( \sigma_H(v) = 1 \) for all \( v \). This extreme case demonstrates an information environment where a signal from a high-type messenger is always positive \( (s_h) \) but provides no informative insights about the true match value.

The signal structures observed in practice typically lie somewhere between the two extreme cases depicted in Figure 4. The proposition characterizes these general signal structures. When a specific messenger type’s signals exhibit greater sensitivity to changes in match values, these elastic signals become more discriminating and consequently more valuable for consumers in resolving uncertainties regarding the true match value \( v \).

### 5.2. Consumer Attention

When a consumer receives an ad, she decides whether to pay attention to the ad by costly deliberating about the content (Wathieu and Bertini 2007, Guo and Zhang 2012, Guo 2016). If the additional information from the advertisement will not change the consumer’s decision, there is no value in incurring the deliberation cost. Thus, upon deliberation, a consumer purchases the product only if she receives a good signal \( s_g \) and does not purchase a product if she receives a bad signal \( s_b \). The expected utility of paying attention is

\[
EU(\text{attention}|m_j) = \sigma_j \cdot [E(v|m_j, s_g) - p] - c. \tag{6}
\]

Conversely, even when the consumer decides not to pay attention to the advertisement, she would still purchase a product if \( E(v) - p \geq 0 \). Thus, the expected utility of not paying attention is

\[
EU(\text{no attention}) = \max\{0, E(v) - p\}. \tag{7}
\]

We characterize the consumer’s attention decision in the following lemma.\(^{18}\)

**Lemma 1.** Consumers pay attention to the ad if and only if the price falls within a moderate range such that

\[
p_j \equiv E(v|m_j, s_b) + \frac{c}{1 - \sigma_j} < p \leq \bar{p}_j \equiv E(v|m_j, s_g) - \frac{c}{\bar{\sigma}_j}, \tag{8}
\]

where \( p_j < \bar{p}_j \) always holds.

The decision rule is intuitive: If the price is too low, consumers find it not worth incurring the deliberation...
costs and prefer to buy the product without paying attention to the ad. Conversely, when the price is too high, even if consumers receive a good signal, they still choose not to purchase the product, making it not worth incurring the deliberation cost to pay attention to the ad.

The lemma, therefore, demonstrates how the price can affect the consumer’s attention decision. Consumers compare the cost and benefit of paying attention to the ad after observing the messenger. The benefit of paying attention depends on their default decision without the ad. When \( E(v) > p \), the default action is to purchase the product. In this case, the marginal benefit of the private signal is avoiding unnecessary purchases when receiving a bad signal \( s_b \). Consumers decide to pay attention to the ad if the cost of paying attention is smaller than the marginal benefit of a bad signal, that is, \( c < (1 - \tilde{\sigma})[p - E(v|m_j, s_b)] \), which can be simplified to \( p > E(v|m_j, s_b) + \frac{c}{1 - \tilde{\sigma}} \). Conversely, when \( E(v) \leq p \), the default action is not to purchase the product. In this case, the marginal benefit of the private signal is helping with the decision to purchase when receiving a good signal \( s_g \). Consumers pay attention to the ad if \( c < \tilde{\sigma}[E(v|m_j, s_g) - p] \), which can be simplified to \( p < E(v|m_j, s_g) - \frac{\tilde{\sigma}}{1 - \tilde{\sigma}} \).

Combining these cases, consumers pay attention to the ad if and only if the price falls within a moderate range given by the lemma: \( p_j < p \leq \bar{p}_j \). This holds under the assumption \( c < \tilde{\sigma} \equiv \min_{m \in \{m_l, m_H\}} \int_{\tilde{\sigma}(\tilde{\sigma} - \sigma(v))dF \} \). Therefore, only within a moderate price range can consumers expect to encounter a useful signal (either \( s_g \) or \( s_b \)) that effectively alters their default action. A signal is deemed useful if it can convince consumers to change their actions.

We next study how the messenger’s type \( m_j \in \{m_l, m_H\} \) can influence consumers’ expected utility and their attention decision. Irrespective of the actual signal realized, whether it is \( s \in \{s_g, s_b\} \), a high-type messenger can generate higher ex ante expected utility compared with a low-type messenger if \( EU(attention|m_H) \geq EU(attention|m_l) \), which is equivalent to

\[
\tilde{\sigma}_H \cdot [E(v|m_H, s_g) - p] \geq \tilde{\sigma}_L \cdot [E(v|m_L, s_g) - p].
\]

**Lemma 2.** A high-type messenger ad generates a higher expected utility of paying attention than a low-type messenger if the price becomes sufficiently low such that \( p \leq \bar{p} \equiv \int_{\tilde{\sigma}_H \cdot \sigma(v) - \tilde{\sigma}_L \cdot \sigma(v)} dF \).

The disparity in the marginal benefit of paying attention between a high-type and a low-type messenger ad stems from two factors: the average probability of receiving good signals, denoted as \( \tilde{\sigma} = Pr(s = s_g|m_j) \), and the net expected utility when receiving good signals, expressed as \( E(v|m_j, s_g) - p \). A high-type messenger possesses a higher likelihood of generating good signals. Thus, if the difference \( E(v|m_H, s_g) - p \) is not significantly smaller than \( E(v|m_L, s_g) - p \), the effect of \( \tilde{\sigma}_H \) prevails, leading to a higher expected utility of paying attention associated with a high-type messenger.

**5.3. Price Premiums**

Given messenger type \( m_j \in \{m_l, m_H\} \), consumers decide to pay attention when \( p_j \leq p \leq \bar{p}_j \). We next compare the highest prices using different messengers with and without consumer attention.

First, we consider the case with consumer attention. If the firm chooses to induce consumer attention, it can set the highest price at \( \bar{p}_j = E(v|m_j, s_g) - \frac{\tilde{\sigma}}{1 - \tilde{\sigma}} \). These highest prices are denoted as \( \bar{p}_j \) and \( \bar{p}_j \), corresponding to different messenger types. The highest price \( \bar{p}_j = E(v|m_j, s_g) - \frac{\tilde{\sigma}}{1 - \tilde{\sigma}} \) consists of two components.

The first component is the consumers’ updated beliefs conditional on receiving a good signal, denoted as \( E(v|m_j, s_g) \). This component is positively correlated with the informativeness of a messenger’s good signal, represented by \( \varepsilon_j \). In other words, a more informative good signal will lead to higher updated beliefs and expectations of value, which can justify a higher price. The second component is the expected net deliberation cost per unit of a good signal, expressed as \( \frac{\tilde{\sigma}}{1 - \tilde{\sigma}} \). This component captures the tradeoff between the cost of deliberation and the perceived value of the good signal. Importantly, the net deliberation cost is always higher for the low-type messenger \( (\frac{\tilde{\sigma}}{1 - \tilde{\sigma}}) \). This is because the average positive signal is always higher for the high-type messenger, resulting in a higher expected effective deliberation cost for the low-type messenger. This difference negatively affects the low-type messenger’s ability to set a higher price compared with the high-type messenger.

Therefore, the combination of consumers’ updated beliefs and the net deliberation cost plays a crucial role in determining the highest price that can be set by each messenger type.

**Proposition 2.** Under consumer attention:

1. When \( \varepsilon_L(v) \leq \varepsilon_H(v) \), a high-type messenger can always command a higher price \( (\bar{p}_H \geq \bar{p}_L) \).
2. When \( \lambda \varepsilon_L(v) \leq \varepsilon_H(v) < \varepsilon_L(v) \), where \( \lambda \equiv \frac{\delta_H(v)}{\delta_l(v)} \), a low-type messenger can command a higher price \((\overline{p}_L > \overline{p}_H)\) if and only if the deliberation cost is sufficiently small such that \( c < c^* \equiv \frac{\int_0^{\infty} \varepsilon_H(v) dF - \int_0^{\infty} \varepsilon_L(v) dF}{\delta_{l-H}} \).

3. When \( \lambda \varepsilon_L(v) > \varepsilon_H(v) \), a low-type messenger can always command a higher price \((\overline{p}_L > \overline{p}_H)\).

The proposition highlights that different messengers can induce a higher price premium depending on the signal environments under consumer attention. First, when \( \varepsilon_L(v) \leq \varepsilon_H(v) \) (Figure 5(a)), the good signals of a high-type messenger are always more informative than those of a low-type messenger. Therefore, it is always the case that \( \overline{p}_L \geq \overline{p}_H \). Second, when \( \lambda \varepsilon_L(v) \leq \varepsilon_H(v) < \varepsilon_L(v) \) (Figure 5(b)), the good signals of a low-type messenger are moderately informative. In this scenario, it is possible for \( \overline{p}_L \) to be larger than \( \overline{p}_H \) when the deliberation cost is low such that the positive effect of the informativeness of the low-type messenger’s signals outweighs the negative effect of the deliberation cost, allowing the low-type messenger to achieve a higher price. Last, when \( \lambda \varepsilon_L(v) > \varepsilon_H(v) \) (Figure 5(c)), the good signals of a low-type messenger are significantly more informative. Consequently, \( \overline{p}_L > \overline{p}_H \) is always true. In summary, the proposition and Figure 5 demonstrate how different messengers can induce a higher price premium under consumer attention based on the relative informativeness of their signals and the deliberation cost.

Next, we compare the highest price premiums that the firm can achieve with different types of messengers under no consumer attention. These price premiums are denoted as \( \overline{p}_L \) and \( \overline{p}_H \).

**Proposition 3.** Under no consumer attention:

1. When \( \lambda \varepsilon_L(v) \leq \varepsilon_H(v) \), a low-type messenger can always command a higher price \((p_L > p_H)\).

2. When \( \kappa \varepsilon_L(v) \leq \varepsilon_H(v) < \lambda \varepsilon_L(v) \) where \( \kappa \equiv \frac{\delta_H(v)}{\delta_l(v)} \), a low-type messenger can command a higher price \((p_L > p_H)\) if and only if the deliberation cost is sufficiently small such that \( c < c^* \equiv \frac{(1-\delta_H) \int_0^{\infty} \varepsilon_H(v) dF - (1-\delta_l) \int_0^{\infty} \varepsilon_L(v) dF}{\delta_{l-H}} \).

3. When \( \kappa \varepsilon_L(v) > \varepsilon_H(v) \), a high-type messenger can always command a higher price \((p_L > p_H)\).

Again, the highest price premium \( p_L \equiv E(v|m_l,s_l) + \frac{\kappa}{1-\kappa} \) for each messenger type \( m_l \) under no consumer attention can be broken down into two components. Similar to the case with consumer attention, the first component is consumers’ updated beliefs conditional on receiving a bad signal, denoted as \( E(v|m_l,s_l) \). This component reflects how consumers adjust their expectations about the match value based on the messenger’s bad signal. The second component is the net deliberation cost per unit of bad signal, represented as \( \frac{\kappa}{1-\kappa} \). This component captures the cost consumers incur in the decision-making process per unit of a bad signal. A higher deliberation cost has a positive effect on the price premium in the absence of consumer attention. This is because a higher cost of deliberation decreases the perceived value of incurring deliberation costs to avoid unnecessary purchases. This effect is particularly pronounced for the high-type messenger as \( \frac{\kappa}{1-\kappa} > \frac{1}{1-\kappa} \). This implies that more elastic or informative signals from the messenger tend to exert downward pressure on the price, whereas a higher deliberation cost has a more significant positive impact on the high-type messenger’s ability to maintain the price under the no consumer attention scenario.

When \( \lambda \varepsilon_L(v) \leq \varepsilon_H(v) \), the bad signals from a high-type messenger become sufficiently informative to influence consumer purchase decisions. As a result, the price premium under no consumer attention is lower for the high-type messenger compared with the low-type messenger: \( p_L > p_H \). In the case of \( \varepsilon_H(v) < \lambda \varepsilon_L(v) \), the bad signals from a high-type messenger become less informative. However, this actually enhances its ability to maintain a higher price under no consumer attention. Specifically, when \( \kappa \varepsilon_L(v) \leq \varepsilon_H(v) < \lambda \varepsilon_L(v) \), the bad signals from the
high-type messenger are moderately uninformative. In this scenario, the deliberation cost plays a crucial role. If the deliberation cost is sufficiently high, the high-type messenger can have a higher price premium compared with the low-type messenger under no consumer attention. Last, when $\kappa \varepsilon(v) > \varepsilon_H(v)$, the bad signals from a high-type messenger are particularly uninformative. In this case, it is always true that $p_H \geq p_L$ regardless of the deliberation cost. Figure 6 depicts these findings.

5.4. Optimal Choice of Messenger

We now turn to the firm’s problem. In the firm’s problem, the objective is to maximize its expected profit, which is given by the expression:

$$\mathbb{E} \Pi = D(p, m_j) \cdot (p - k),$$

where $D(p, m_j)$ represents the demand for the product, and $p - k$ is the profit margin per unit.

The firm’s decision variables are the choice of messenger type $m_j$ and the price $p$. These decisions influence consumer attention and purchasing decisions, which in turn determine the demand for the product. The firm aims to find the optimal combination of messenger type and price that maximizes its expected profit. In analyzing the firm’s profit, we take into account the impact of both the profit margin and the demand function, which are influenced by the choice of messenger type $m_j$ and the price $p$. We assume that consumers can observe the price before making their attention decisions, allowing them to factor it into their decision-making process. Propositions 2 and 3 state which types of messengers can generate a higher profit margin. However, the demand function $D(p, m_j)$ is also a function of the price $p$ and the type of messenger $m_j$. Thus, we need to consider the overall effects of the messenger on the firm’s profit through both channels (profit margin and demand) together. Given the consumer attention decision in Lemma 1, the firm considers two demand regimes: full market coverage and partial market coverage.

Figure 6. Price Premiums Under No Consumer Attention

First, the firm has the option to set a price that ensures all consumers purchase the product without incurring deliberation costs, which is the case of full market coverage without consumer attention. In this case, the demand is $D(p, m_j) = 1$ for any messenger type $m_j$. Thus, under full coverage, the firm’s optimization problem is

$$\max_{p, m_j \in \{m_L, m_H\}} \mathbb{E} \Pi_{\text{full}} = 1 \cdot (p - k),$$

subject to $k \leq p \leq p_j$. (11)

To achieve full market coverage, the firm sets the price below the threshold $p_j$, which is the maximum price at which consumers decide to purchase the product without paying attention, given a messenger type $m_j$. This ensures that all consumers make a purchase decision without incurring deliberation costs (IC constraint). The price must also be set higher than the production cost $k$ to generate a positive profit margin (IR constraint). We assume that $k \leq \min\{p_L, p_H\}$ to satisfy the firm’s IC constraint. Under full coverage, the firm’s optimal price must make the IC constraint bind, meaning that $p = p_j$ for some $j \in \{L, H\}$. Solving the optimization problem reveals that under full coverage without consumer attention, a low-type messenger is optimal if and only if $p_L > p_H$. The conditions for this inequality to hold are provided in Proposition 2.

Next, in the case of partial market coverage with consumer attention, the firm can set the price in the range $p_j < p \leq \bar{p}_j$ to attract consumers’ attention and encourage them to purchase the product only if they receive a good signal. In this scenario, the demand function becomes $D(p, m_j) = \bar{d}_j \cdot \Pr(s_j = g|m_j)$, representing the probability that consumers receive a good signal given a messenger type $m_j$. The firm’s optimization problem under partial market coverage with consumer attention is

$$\max_{p, m_j \in \{m_L, m_H\}} \mathbb{E} \Pi_{\text{part}} = \bar{d}_j \cdot (p - k),$$

subject to $p_j < p \leq \bar{p}_j$ and $k \leq p$. (12)

Similar to the full coverage case, the optimal price for partial market coverage must satisfy the constraint $p \leq \bar{p}_j$.
for some $j \in \{L, H\}$ (IC constraint), and it must be higher than the production cost $k$ (IR constraint). Under partial coverage, the firm faces a tradeoff between demand $\tilde{\sigma}_j$ (where $\tilde{\sigma}_L < \tilde{\sigma}_H$) and profit margin $\tilde{p}_j - k$. When $\tilde{p}_j \geq \bar{p}_j$, it is optimal for the firm to choose a high-type messenger. However, when $\tilde{p}_j < \bar{p}_j$, a low-type messenger ad is not necessarily optimal despite its higher profit margin, as it may result in lower demand. Therefore, the low-type messenger is only optimal when the profit margin $\tilde{p}_L - k$ is sufficiently larger than $\tilde{p}_H - k$. This condition can be expressed as

$$
\mathbb{E} \Pi_{\text{opt}}(m_H) = \tilde{\sigma}_H \cdot (\tilde{p}_H - k) < \mathbb{E} \Pi_{\text{opt}}(m_L) = \tilde{\sigma}_L \cdot (\tilde{p}_L - k) \iff \tilde{p} < k.
$$

(13)

Because the production cost is assumed to be less than an average consumer’s consumption utility $(k < E(v))$, the condition $\tilde{p} < k$ is possible only if $\tilde{p} < E(v)$. The following lemma provides a sufficient condition for this inequality to hold.

**Lemma 3.** When $\varepsilon_H(v) < \lambda \varepsilon_L(v)$, where $\lambda \equiv \frac{\alpha_v}{\sigma_{vH}}$, we have $\tilde{p} < E(v)$. Otherwise, when $\varepsilon_H(v) \geq \lambda \varepsilon_L(v)$, $\tilde{p} \geq E(v)$.

From the previous discussion, we can summarize the optimal choice of messenger and price under full and partial coverage cases in Table 3.

By comparing the expected profits under partial and full coverage, we can determine the firm’s optimal coverage choice with the optimal messenger and price. This represents the main result of this research, as stated in the following proposition.

**Proposition 4.** The firm’s optimal coverage decisions are:

1. When $\lambda \varepsilon_L(v) \leq \varepsilon_H(v)$, inducing partial coverage with $(m_H, \tilde{p}_H)$ is optimal if $k \geq k_4(c)$, otherwise inducing full coverage with $(m_L, \tilde{p}_L)$ is optimal.

2. When $\lambda \varepsilon_L(v) > \varepsilon_H(v)$, inducing partial coverage is optimal if $k \geq k_3(c)$. In this case, $(m_L, \tilde{p}_L)$ is optimal if and only if $k > \tilde{p}$. Otherwise, if $k < k_3(c)$, inducing full coverage is optimal, where $(m_L, \tilde{p}_L)$ is optimal if and only if $c < c^*$

$$
(1 - \tilde{\sigma}_L) \int_0^{\varepsilon_H} \varepsilon [1 - \varepsilon(\varepsilon)] d\varepsilon - (1 - \tilde{\sigma}_H) \int_0^{\varepsilon_L} \varepsilon [1 - \varepsilon(\varepsilon)] d\varepsilon
$$

3. When $\lambda \varepsilon_L(v) > \varepsilon_H(v)$, inducing partial coverage is optimal if $k \geq k_3(c)$. In this case, $(m_L, \tilde{p}_L)$ is optimal if and only if $k > \tilde{p}$. Otherwise, if $k < k_3(c)$, inducing full coverage with $(m_H, \tilde{p}_H)$ is optimal.

Moreover, all thresholds are increasing in $c$: $\frac{\partial k_3(c)}{\partial c} \geq 0$, $\frac{\partial k_4(c)}{\partial c} \geq 0$, and $\frac{\partial k_5(c)}{\partial c} \geq 0$.

**Table 3.** Optimal Choice of Messenger and Price ($m_i^*, p^*$) Under Full and Partial Coverage

<table>
<thead>
<tr>
<th>Coverage</th>
<th>$\lambda \varepsilon_L(v) \leq \varepsilon_H(v)$</th>
<th>$\lambda \varepsilon_L(v) \leq \varepsilon_H(v) &lt; \lambda \varepsilon_L(v)$</th>
<th>$\lambda \varepsilon_L(v) \geq \varepsilon_H(v)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full (without consumer attention)</td>
<td>$(m_L, \tilde{p}_L)$</td>
<td>$(m_H, \tilde{p}_H)$, $c &lt; c^*$</td>
<td>$(m_H, \tilde{p}_H)$, otherwise</td>
</tr>
<tr>
<td>Partial (with consumer attention)</td>
<td>$(m_L, \tilde{p}_L)$</td>
<td>$(m_H, \tilde{p}_H)$, $k &gt; \tilde{p}$</td>
<td>$(m_H, \tilde{p}_H)$, otherwise</td>
</tr>
</tbody>
</table>

Figure 7 visually presents the insights from Proposition 4, with the darker areas in each subfigure indicating the firm’s choice of partial coverage with consumer attention. When deciding whether to induce consumer attention, the firm considers both the demand and the price premium to optimize its profit. As the production cost $k$ increases, the firm faces a greater need to raise prices. Consequently, the firm tends to sacrifice demand to increase profit margin. Conversely, as the consumer’s deliberation cost $c$ increases, the firm becomes more capable of charging higher prices without inducing consumer attention. This tempts the firm to prioritize retaining greater demand over maximizing profit margin. Thus, full coverage without consumer attention becomes more likely as the consumer’s deliberation cost $c$ increases or the product cost $k$ decreases.

Additionally, the thresholds for inducing consumer attention, represented by $k_n(c)$ for $n \in \{1, 2, 3\}$, are all increasing functions of the deliberation cost $c$: $\frac{\partial k_n(c)}{\partial c} \geq 0$. When the deliberation cost is low, the firm can sufficiently increase prices by inducing consumer attention (recall that $\tilde{p}_j$ decreases with $c$). Therefore, even if the need to raise prices is not significant (small $k$), the firm still prefers to induce consumer attention. Conversely, as the deliberation cost increases, the firm has less incentive to raise prices through consumer attention. As a result, the firm opts for full coverage without consumer attention unless there is a strong need to increase prices (large $k$).

Moreover, the optimal choice of messengers is influenced by their elasticities and the firm’s coverage decision, which, in turn, depends on the consumer’s deliberation cost and production cost. In cases like healthcare, where experts like doctors endorse drugs matching consumers’ needs, trust in the product can significantly increase if the advertised product aligns with those needs. Consequently, when a high-type messenger is highly informative ($\varepsilon_H(v) \geq \lambda \varepsilon_L(v)$, as depicted in Figure 7(a)), using a high-type messenger in advertisements proves more effective in raising prices. Therefore, the firm finds it optimal to capture consumer attention and raise prices when the production cost $k$ significantly outweighs the deliberation cost $c$ (i.e., $k \geq k_4(c)$). However, given the high informativeness of signals from the high-type messenger, the firm must maintain a very low price if it opts to dissuade consumer attention effectively. Consequently, using a low-type messenger emerges as the optimal
strategy to achieve full coverage without attracting undue consumer attention.

In cases where a high-type messenger becomes less informative \( \epsilon_H(v) < \lambda \epsilon_L(v) \), as shown in Figure 7, (b) and (c)), the effectiveness of a high-type messenger ad in updating consumers’ beliefs about the product match is limited. This phenomenon emerges when the high-type messenger’s prominence overshadows product-specific information. A classic illustration of this is the widely acknowledged “halo effect” associated with celebrity endorsements, where the messenger’s presence alone enhances the product’s appeal, irrespective of the actual match value. In contrast, the rare occurrence of a good signal from a low-type messenger can have a significant impact on updating consumers’ beliefs. Therefore, for partial coverage (when \( k \geq k_2(c) \) or \( k \geq k_3(c) \)), a low-type messenger can be optimal for inducing consumer attention and charging a high price to cover the production cost under the following conditions: (1) the low-type messenger’s signals are sufficiently elastic or informative, and (2) the production cost is sufficiently high \( (k > \bar{p}) \). However, under full coverage, the firm should choose the type of messenger that can command a higher price without relying on consumer attention. For example, when \( \kappa \epsilon_L(v) \leq \epsilon_H(v) < \lambda \epsilon_L(v) \) and \( c \geq c^* \) (Figure 7(b)), a high-type messenger is optimal for full coverage without consumer attention. This scenario aligns with the case of Stitch Fix in our motivating example. Both H&M and Stitch Fix operate in the fashion industry, where attractive models or influencers are frequently used in social media posts. Stitch Fix, being a company that provides personalized styling services, may incur higher production costs compared with H&M. Consequently, using normal people or nonattractive models as messengers can be more effective in attracting consumer attention and raising price premiums for Stitch Fix. Thus, the choice of messengers for companies like H&M and Stitch Fix is in line with our models, taking into account production costs and the desired coverage strategy (full or partial). Stitch Fix opts for normal people to attract consumer attention and raise price premiums, whereas H&M uses attractive models to induce full coverage without relying on consumer attention.

6. Extensions

6.1. Attention-Grabbing Role

We assumed the deliberation cost is constant and fixed for both types of messengers \( c(m_H) = c(m_L) = c \). We now relax this assumption to capture the attention-grabbing role of the advertising messenger. In practice, a high-type messenger can directly lower the consumer’s deliberation cost by providing extra entertaining value that offsets the deliberation cost. To capture this idea, we assume that a high-type messenger ad can reduce the deliberation cost to zero, making all consumers pay attention to the ad: \( c(m_H) = 0 < c(m_L) = c \).

The analysis is similar to the main model, with the only difference being the demand generated by a high-type messenger ad. Because of the zero deliberation cost of a high-type messenger ad \( (c(m_H) = 0) \), consumers always receive signals from it. Consequently, the consumer’s posterior belief is either \( E(v|m_{HI}, s_g) \) when receiving a good signal or \( E(v|m_{HI}, s_b) \) when receiving a bad signal. The demand function associated with a high-type messenger ad can be defined as follows:

\[
D(p, m_H) = \begin{cases} 
1, & p \leq E(v|m_{HI}, s_b) \\
\bar{g}_H, & E(v|m_{HI}, s_b) < p \leq E(v|m_{HI}, s_g) \\
0, & p > E(v|m_{HI}, s_g). 
\end{cases} \quad (14)
\]

Again, the firm has to consider both the demand and the profit margin when maximizing the profit. Similar to the analyses in the main model, we first determine the optimal price and messenger under the full and partial coverage cases. According to Equation (14), the optimal price for a high-type messenger ad is \( E(v|m_{HI}, s_b) = \bar{p}_H \) under full coverage and \( E(v|m_{HI}, s_g) = \bar{p}_H \) under partial coverage. Under full coverage, a low-type messenger ad
is optimal if (1) \( \varepsilon_H(v) \geq \lambda \varepsilon_L(v) \) or (2) \( \varepsilon_H(v) < \lambda \varepsilon_L(v) \) and 
\[ c > (1 - \delta_L)(E(v|m_{Hj},s_j) - E(v|m_L,s_j)) \equiv c^* \]. Also, under partial coverage with consumer attention, a low-type messenger ad is optimal if
\[
\bar{\sigma}_L \left( E(v|m_{Lj},s_j) - \frac{c}{\bar{\sigma}_L} - k \right) > \bar{\sigma}_H \left( E(v|m_{Lj},s_j) - k \right) \]
\[ \iff k > \bar{p} + \frac{c}{\bar{\sigma}_H - \bar{\sigma}_L} \]
(15)
where \( \bar{p} = \frac{\int \kappa_{\varepsilon_L(v)} - \varepsilon_H(v) \, dF}{\bar{\sigma}_H - \bar{\sigma}_L} \) as before. When \( \varepsilon_H(v) \geq \lambda \varepsilon_L(v) \), \( \bar{p} \geq E(v) \) and a high-type messenger ad is always optimal under partial coverage with consumer attention.

We compare the full and the partial cases to determine the firm’s optimal coverage with optimal messenger and price in the following proposition.

**Proposition 5.** The firm’s optimal coverage decision with optimal messenger and price are:

1. **If** \( \lambda \varepsilon_L(v) \leq \varepsilon_H(v) \), inducing partial coverage with \( (m_{Hj},\bar{p}_{Lj}) \) is optimal when \( k \geq k_4(c) \). Otherwise, full coverage with \( (m_L,p_L) \) is optimal.

2. **If** \( \kappa \varepsilon_L(v) \leq \varepsilon_H(v) < \lambda \varepsilon_L(v) \), inducing partial coverage is optimal when \( k \geq k_5(c) \). Otherwise, full coverage with \( (m_L,p_L) \) is optimal. Furthermore, under partial coverage, choosing a low-type messenger with \( (m_L,\bar{p}_L) \) is optimal if and only if \( k > \bar{p} + \frac{c}{\bar{\sigma}_H - \bar{\sigma}_L} \).

3. **If** \( \kappa \varepsilon_L(v) > \varepsilon_H(v) \), inducing partial coverage is optimal when \( k \geq k_6(c) \). Otherwise, full coverage is optimal. Moreover, under partial coverage, choosing a low-type messenger with \( (m_L,\bar{p}_L) \) is optimal if and only if \( c < c^* \).

Figure 8 illustrates the results from Proposition 5. Similar to the main analysis, when the production cost is sufficiently close relative to the deliberation cost, inducing partial coverage with a higher price remains the optimal strategy. The attention-grabbing effect of a high-type messenger reduces the range where a low-type messenger is optimal under partial coverage with consumer attention. As the deliberation cost increases, the low-type messenger needs to lower the price to attract attention, whereas the high-type messenger is not subject to the same constraint. Consequently, the parameter region where the low-type messenger is optimal shrinks and eventually disappears with higher deliberation costs. Even with high production costs, a high-type messenger becomes more effective for advertising communication, as evidenced by Figure 8, (b) and (c). Under full coverage, a low-type messenger dominates in most cases, except when the low informativeness of a high-type messenger (small \( \varepsilon_H(v) \)) and a low deliberation cost allow for a higher price from the high-type messenger.

Nevertheless, even with the attention-grabbing effect of a high-type messenger, a low-type messenger can still be a more effective medium for advertisers. That is, the general insight that a low-type messenger can still draw consumer attention and, more importantly, persuade them to pay a higher premium carries over if the importance of boosting the price is high, but a high-type messenger is less informative (\( \varepsilon_H(v) < \lambda \varepsilon_L(v) \)).

### 6.2. Naive Consumers

In our main model, we assume all consumers are rational. However, we recognize that, in reality, there may be a segment of consumers who do not make rational inferences or attention decisions. Instead, these consumers simply pay attention to the ad and take the private signals at face value, regardless of the messenger generating those signals. In this section, we consider such a realistic scenario where a fraction \( \alpha \) of consumers are naive. These naive consumers’ posteriors upon observing the signals are \( E(v|m_{1j},s_j) = \bar{s} \) and \( E(v|m_{1j},s_j) = 0 \) for any \( j \in \{L,H\} \), which means that they believe that the signals are completely informative about the product match. Therefore, a high-type messenger has the advantage of generating more good signals and being able to convert more naive consumers.

**Figure 8.** Optimal Choice of Messenger When Attention-Grabbing Effect Is Present
Given that naive consumers always pay attention to the ad and purchase the product if and only if they receive good signals, the demand function can be specified as follows:\(^2^:\)

\[
D(m_i, p) = \begin{cases} 
\alpha \tilde{\sigma}_{ij} + (1 - \alpha), & p \leq \tilde{p}_j \\
\tilde{\sigma}_{ij}, & \tilde{p}_j < p \leq \overline{p}_j \\
\alpha \tilde{\sigma}_{ij}, & \overline{p}_j < p \leq \overline{p}.
\end{cases}
\]  

(16)

There are three cases in terms of the coverage of the “rational” consumers. First, in the full coverage case without (rational) consumer attention \((p \leq \tilde{p}_j)\), the optimal price given the messenger \(m_i = \tilde{p}_j\). A high-type messenger is optimal if and only if \([\alpha \tilde{\sigma}_{ij} + (1 - \alpha)][\tilde{p}_j - k] \geq [\alpha \tilde{\sigma}_L + (1 - \alpha)][\tilde{p}_L - k] \equiv k \leq \frac{\tilde{p}_L - \tilde{p}_j}{\alpha \tilde{\sigma}_L - \alpha \tilde{\sigma}_{ij}} + \tilde{p} \equiv \tilde{k}(c, \alpha)\). Second, in the partial coverage case with (rational) consumer attention \((\tilde{p}_j < p \leq \overline{p}_j)\), the optimal price given the messenger \(m_i = \overline{p}_j\). A high-type messenger is optimal if and only if \(\tilde{\sigma}_{ij}(\overline{p}_j - k) \leq \tilde{\sigma}_L(\overline{p}_j - k) \equiv k \leq \overline{p}_j\), which always holds when \(\epsilon_{ij}(\overline{p}_j) \geq \lambda \epsilon_L(v)\) for all \(v\). Finally, there is a new case where no (rational) consumers purchase—no coverage case \((\overline{p}_j < p \leq \overline{p})\), where only naive consumers purchase when they receive private good signals. In this case, a high-type messenger is always optimal, and the optimal price is \(\overline{p}\).

If \(\alpha\) becomes large, the case reverts to a trivial case.\(^2^3\) We assume that \(\alpha\) is small enough \(\left(\alpha \leq \min\left\{\tilde{\sigma}_{ij}, \tilde{\sigma}_L, \frac{1}{\tilde{\sigma}_{ij}}, \tilde{\sigma}_L\right\}\right)\) such that the general demand structure in the main model holds.\(^2^4\) We highlight the conditions under which it is optimal for the firm to attract rational consumer attention with a low-type messenger in the following proposition.

**Proposition 6.** Suppose \(\alpha \leq \min\left\{\tilde{\sigma}_{ij}, \tilde{\sigma}_L, \frac{1}{\tilde{\sigma}_{ij}}, \tilde{\sigma}_L\right\}\). If \(\epsilon_{ij}(\overline{v}) < \lambda \epsilon_L(v)\) and \(\max\{k_{ij}(c, \alpha), \tilde{p}\} < k < k_0(c, \alpha)\), partial coverage with rational consumers’ attention and a low-type messenger \(m_i\) is optimal.\(^2^5\)

Figure 9 provides a comprehensive illustration of the optimal choices for the firm in terms of messenger selection and pricing. As Proposition 6 shows, a low-type messenger can optimally attract rational consumers’ attention when \(\epsilon_{ij}(\overline{v}) < \lambda \epsilon_L(v)\) for all \(v\) and \(\max\{k_{ij}(c, \alpha), \tilde{p}\} < k < k_0(c, \alpha)\) (Figure 9(b)). However, as the proportion of naive consumers \(\alpha\) increases, such region would gradually diminish and can eventually disappear. Intuitively, when there are many naive consumers who are willing to pay the highest price \((p = \overline{p})\) upon receiving good signals, it becomes more profitable for the firm to forego the attention of rational consumers (i.e., no coverage from rational consumers) by using a high-type messenger instead.

6.3. Unobservable Pricing

In this section, we introduce a new setting where consumers cannot costlessly observe the price of the product. Instead, consumers can only observe the type of messenger. Based on this limited information and their expectation of the product price, denoted as \(p'\), consumers make their initial decision on whether to engage in deliberation of the advertising content. This deliberation process incurs a cost of deliberation, denoted as \(c\). After deliberation, consumers update their beliefs about the product matches and then decide whether to search for the actual price of the product to make a purchase decision. This price search process incurs a cost of price search, denoted as \(\xi\). However, this setting introduces a classic hold-up problem. When consumers receive a good signal from the messenger, the firm has the incentive to charge a price higher than the expected price to extract all consumer surplus, as consumers’ deliberation costs and price search costs are considered sunk. Anticipating this opportunistic behavior of the firm, consumers become unwilling to pay attention to the ad and engage in price

Figure 9. Optimal Choice of the Messenger When Some Consumers Are Naive
search ex ante. To address such a problem, we introduce consumer heterogeneity by assuming that a portion β of consumers are shoppers who have no deliberation cost and price search cost so that they always pay attention to the ad and observe the price, whereas (1 − β) of “regular” consumers have a deliberation cost c > 0 and a price search cost $\xi > 0$.

Consumers who do not have deliberation costs or price search costs are able to observe a private signal and the price freely. As a result, they will make a purchase if the actual price is lower than their updated beliefs about the product match: $p \leq E(v|m_j, s_k)$. For regular consumers, their decision to engage in price search depends on whether they pay attention to the ad or not. If they pay attention to the ad, they will engage in price search only if they observe a good signal and the difference between their updated belief about the product match and the expected price is greater than or equal to zero: $E(v|m_j, s_k) - p' \geq \xi > 0$. Conversely, if regular consumers do not pay attention to the ad, they will engage in price search only if their expected benefit with a prior belief is greater or equal to the expected price: $E(v) - p' - \xi \geq 0$. Overall, regular consumers would pay attention to the ad with messenger $m_j$ if and only if $\text{EU}(\text{attention}|m_j) = \tilde{\sigma}(E(v|m_j, s_k) - p' - \xi) - c \geq \text{EU}(\text{no attention}) = \max\{E(v) - p' - \xi, 0\}$. Therefore, regular consumers would pay attention to the ad and engage in price search if and only if $p_j < p' + \xi \leq \overline{p}_j$ and not pay attention to the ad and engage in price search if and only if $p' + \xi > \overline{p}_j$, and not pay attention to the ad and not engage in price search if and only if $p' + \xi > \overline{p}_j$.

The total consumer demand is a combination of the demand from shoppers and the demand from regular consumers, taking into account the different decision rules of these two types of consumers: $D(m_j, p', p) = \beta D_1(m_j, p) + (1 - \beta)D_2(m_j, p', p)$, where the demand from shoppers is $D_1(m_j, p) = \tilde{\sigma}(p \leq E(v|m_j, s_k)) - (1 - \tilde{\sigma})\mathbb{I}(p \leq E(v|m_j, s_k))$ and the demand from regular consumers is

$$D_2(m_j, p', p) = \begin{cases} \mathbb{I}(p \leq E(v)), & \text{if } p' + \xi \leq p_j \\ \tilde{\sigma}(p \leq E(v|m_j, s_k)), & \text{if } p' < p' + \xi \leq \overline{p}_j \\ (1 - \tilde{\sigma})\mathbb{I}(p \leq E(v|m_j, s_k)), & \text{if } p' + \xi > \overline{p}_j. \end{cases}$$

(17)

Given the consumer demand, the profit-maximizing price is either $E(v|m_j, s_k)$, $E(v)$, or $E(v|m_j, s_k)$. However, $p = E(v)$ can never be consistent with the expected price $p'$ because $p = E(v)$ can be optimal only when $p' + \xi \leq p_j < E(v)$. Thus, only $E(v|m_j, s_k)$ and $E(v|m_j, s_k)$ can possibly be consistent with the expected price in the equilibrium. To simplify the analysis, we also assume that the price search cost is small such that $E(v|m_j, s_k) + \xi < p_j$ for all $j$. There are two cases in terms of the coverage of the regular consumers: the full coverage case without regular consumer attention case with $p = E(v|m_j, s_k)$ and the partial coverage case with $p = E(v|m_j, s_k)$.20

In the full coverage case, the firm’s profit given a messenger $m_j$ is $E(v|m_j, s_k) - k$, so the high-type messenger is optimal if and only if $E(v|m_j, s_k) \geq E(v|m_j, s_k)$, which holds when $\varepsilon_H(v) < \xi E(1)$ for all $v$. In the partial coverage case, the firm’s profit given a messenger $m_j$ is $\tilde{\beta}j(E(v|m_j, s_k) - k)$, so the high-type messenger is optimal if and only if $\tilde{\beta}j(E(v|m_j, s_k) - k) \geq \tilde{\beta}j(E(v|m_j, s_k) - k) \iff \tilde{p} \geq \tilde{p}_j$, which is always the case when $\varepsilon_H(v) \geq \xi E(1)$. Table 4 summarizes the results of the firm’s optimal messenger and price under each regime.

**Proposition 7.** The firm’s optimal messenger and price decisions are:

1. When $\lambda E(1) \leq \varepsilon_H(v)$, it is optimal to induce partial coverage with $(m_H, E(v|m_H, s_k))$ if $k \geq k_{11}$. Otherwise, full coverage with $(m_L, E(v|m_L, s_k))$ is optimal.

2. When $\kappa E(1) \leq \varepsilon_H(v) < \lambda E(1)$, it is optimal to induce partial coverage if $k \geq k_{12}$. Otherwise, full coverage with $(m_L, E(v|m_L, s_k))$ is optimal. Moreover, under partial coverage, the firm chooses a low-type messenger with $(m_L, E(v|m_L, s_k))$ if and only if $k > \tilde{p}$.

3. When $\kappa E(1) > \varepsilon_H(v)$, it is optimal to induce partial coverage if $k \geq k_{13}$. Otherwise, it is optimal to induce full coverage with $(m_H, E(v|m_H, s_k))$. Moreover, under partial coverage, the firm chooses a low-type messenger with $(m_L, E(v|m_L, s_k))$ if and only if $k > \tilde{p}$.

The results are similar to the main results, indicating that a low-type messenger can still be an effective medium to attract consumer attention and potentially increase prices for the firm. However, there are some notable differences due to the introduction of price search costs for regular consumers. The regions where a messenger-price pair is optimal are independent of the deliberation cost. This is because no price can credibly attract attention from regular consumers. Therefore, the firm’s pricing strategy for regular consumers is influenced by the two extremes: either giving up on them by setting a high price $E(v|m_j, s_k)$ or setting a low price

<table>
<thead>
<tr>
<th>Coverage</th>
<th>$\varepsilon_H(v) \geq \lambda E(1)$</th>
<th>$\kappa E(1) \leq \varepsilon_H(v) &lt; \lambda E(1)$</th>
<th>$\varepsilon_H(v) &lt; \kappa E(1)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full</td>
<td>$(m_L, E(v</td>
<td>m_L, s_k))$</td>
<td>$(m_L, E(v</td>
</tr>
<tr>
<td>Partial</td>
<td>$(m_H, E(v</td>
<td>m_H, s_k))$</td>
<td>$(m_H, E(v</td>
</tr>
</tbody>
</table>

Table 4. Optimal Choice of Messenger and Price ($m'_j, p'$) When Pricing Is Unobservable
6.4. Other Extensions

We also carry out other extensions to assess the robustness of our main insight. We overview those extensions, and the details are relegated to the online appendix. Overall, we find that the qualitative insights from the main model carry over to these extension models.

1. **Optimal signal structure**: In many real-world scenarios, firms have some control over the messengers they use, but they may have limited ability to shape the specific signal structure. The type of messengers available in the market limits firms’ ability to design the signal structure that consumers encounter when they pay attention to advertisements. In this extension, we explore the implications of giving firms complete freedom in designing the signal structure. The intuition for coverage choice is similar to the main model. A less informative messenger commands a higher price under full coverage, whereas a more informative messenger attracts consumer attention and commands a higher price under partial coverage. With complete design freedom, the optimal choice is the least informative messenger for full coverage and the most informative messenger for partial coverage.

2. **Optimal deliberation cost**: In our previous analysis, we assumed that the deliberation cost was determined externally. However, in practice, firms have the ability to manipulate the deliberation cost through various means, such as adjusting the ad format or content readability. In this extension, we investigate the optimal deliberation cost when it is endogenously determined by the firm. Despite this strategic manipulation, we find that the main intuition for the firm’s coverage choice remains unchanged.

3. **Social utility and fandom**: Sometimes, high-type messengers such as celebrities can directly affect consumers’ preference for a product because consumers might project their feelings about celebrities toward the product. Thus, we assume that other than the consumption utility \( v_p \), consumers can obtain additional social utilities \( \eta > 0 \) if the product is endorsed by a high-type messenger. The qualitative results are the same. However, as the magnitude of the social utility \( \eta \) increases, high-type messengers gain an advantage in raising prices. Consequently, the regions where a high-type messenger is optimal expand, indicating their increased effectiveness in influencing consumer preferences and commanding higher prices.

4. **Differential fixed costs**: In this analysis, we relax the assumption of fixed costs for featuring messengers and consider the case where different messengers incur different fixed costs. Specifically, we explore the scenario where featuring a celebrity as a messenger is generally more expensive compared with featuring ordinary individuals. By incorporating varying fixed costs, we examine the implications of messenger selection when cost differentials are present.

5. **Exogenous pricing case**: In this analysis, we consider a scenario where prices of different products within a specific category are generally similar and known to consumers through their past interactions. For instance, many mobile apps are priced at $0.99, establishing a price level that consumers are familiar with. We assume an exogenous price setting, whereby the price is predetermined and not influenced by the firm. By incorporating this assumption, we investigate the implications of consumer attention decisions based on their existing knowledge of the price within the product category.

7. Conclusion

Understanding the impact of advertising messengers on persuasion is crucial for successful advertising campaigns. This paper presents a framework to analyze the effectiveness of advertising messengers. We adopt a dual-mode communication perspective, where the type of messenger is observable at no cost while paying attention to product-related information incurs a cost for consumers. By committing to a specific type of messenger, firms can influence consumers’ attention decisions. Using a Bayesian persuasion framework, we study the role of messengers in persuading consumers through product match signals.

We identify key variables that influence messenger choice, including the signal structures of different messengers, the deliberation cost, and the production cost. The signal structures determine the sensitivity of each messenger’s signals to consumers’ product matches, whereas the deliberation and production costs determine the firm’s ability and incentive to raise prices. We find that a more informative messenger is optimal when raising prices is crucial, as it effectively attracts consumer attention. Particularly, a low-type messenger can be optimal when a high-type messenger would overshadow product-related information, and the benefits of raising prices outweigh the potential loss in consumer demand due to deliberation. Conversely, a less informative messenger can be beneficial when maximizing demand is the priority, as it discourages consumer attention even at a higher price. These findings have managerial implications for firms in selecting appropriate messengers based on various factors. We also extend our main model to incorporate various important factors, including the attention-grabbing role of messengers, the presence of naive consumers, the unobservability of price information, the freedom to design signal structures, the endogenization of deliberation cost, the inclusion of social utility from fandom, differential fixed costs, and exogenous pricing. Across these different scenarios, we find that the main insight remains consistent and holds true.

The paper acknowledges several limitations. First, although we present one plausible explanation for messenger choices within the Bayesian persuasion framework, it is essential to recognize that various other
factors may influence a firm’s decisions. These factors could encompass brand positioning, where firms opt for different messengers to set themselves apart from competitors, or the specific target audience, leading firms to select messengers who align with the audience’s values or effectively capture their attention. Although these factors hold significant practical relevance, we have omitted them for the scope of this research, leaving them as promising avenues for future exploration. Investigating these effects in future studies promises to yield intriguing insights. Second, the results are characterized under specific relationships between signal structures, which may make it challenging to extend the findings to more general cases. The relationship between $\varepsilon_H(v)$ and $\varepsilon_L(v)$ could vary across different ranges of $v$ (e.g., when $v$ is small, $\varepsilon_H(v) > \varepsilon_L(v)$, but when $v$ is high, $\varepsilon_H(v) < \varepsilon_L(v)$). However, the practical implications of modeling these nonlinear variations remains uncertain, particularly when considering how firms perceive and comprehend such nuanced relationships. Unfortunately, we lack a clear theoretical framework that specifies when and how these complex variations would manifest in practical scenarios. Instead, we adopt an assumption of a monotonic relationship where $\varepsilon_H(v)$ and $\varepsilon_L(v)$ consistently maintain their relative behavior across all values of $v$. This assumption represents a tradeoff for tractability and helps build intuitions about the relative informativeness of different messengers’ signal structures.  

Our paper has primarily focused on the influence of the messenger on persuasion, specifically examining “who is saying it.” However, the Bayesian persuasion framework can be expanded to encompass other aspects of ad content and the choice of medium. This includes considering factors like ad style and design, such as sophisticated computer graphics versus simpler visuals or engaging scripts versus plain text. Additionally, the choice of the medium on which an ad appears can be analyzed using the Bayesian persuasion mechanism. For instance, an ad for a health remedy can be placed on a site dedicated to serious health news (high-type messenger) or on a site for celebrity news (low-type messenger). Furthermore, the framework can be applied to settings beyond advertising, such as salespeople employing different tactics (aggressive or conservative) when selling products, generating signals with varying levels of informativeness. Investigating how these factors impact information structures, consumer attention, purchase decisions, and the extent of commitment power presents valuable opportunities for future research. By exploring these avenues, we can deepen our understanding of advertising dynamics and its broader implications.

Acknowledgments

The authors thank Jason Choi, Liang Guo, Tony Ke, Eddie Ning, Aniko Oery, Woocheol Shin, Jungju Yu, Jidong Zhou, and seminar participants at KAIST, Yale University, and Marketing Science Conference for valuable feedback and comments. The authors contributed equally, and their names are listed in alphabetical order.

Appendix

Proof of Proposition 1. Note that $\frac{f^R(c|m_{L}, s_k)}{f^R(c|m_{L}, s_k)} = \frac{f^R(c|m_{R}, s_k)}{f^R(c|m_{R}, s_k)}$. 

Because $\frac{d}{dv} \frac{f^R(c|m_{L}, s_k)}{f^R(c|m_{L}, s_k)}$ does not depend on $v$, $f^R(c|m_{L}, s_k)$ is increasing in $v$ if and only if $\frac{d}{dv} \frac{f^R(c|m_{L}, s_k)}{f^R(c|m_{L}, s_k)} \geq 0 \Leftrightarrow \frac{d}{dv} \frac{f^R(c|m_{L}, s_k)}{f^R(c|m_{L}, s_k)} \geq \frac{d^2f(c|m_{L}, s_k)}{dv^2} \geq 0$ for all $v$. Conversely, $f^R(c|m_{L}, s_k)$ is decreasing in $v$ if and only if $\frac{d}{dv} \frac{f^R(c|m_{L}, s_k)}{f^R(c|m_{L}, s_k)} \geq \frac{d^2f(c|m_{L}, s_k)}{dv^2} < 0$ for all $v$. □

Proof of Lemma 1. First, when $p \geq E(v)$, a consumer pays attention to the ad if and only if $EU(\text{attention}|m_l) = \frac{1}{\gamma} \cdot [E(v|m_l, s_k) - p] - c \geq EU(\text{no attention}) = \max(0, E(v) - p)$, $0 \Leftrightarrow p \geq \frac{b(E(v|m_l, s_k) - c)}{\gamma} = E(v|m_l, s_k) - \frac{c}{\gamma}$. Next, when $p < E(v)$, consumers pay attention to the ad if and only if $\frac{1}{\gamma} \cdot E(v|m_l, s_k) - p] - c \geq \max(0, E(v) - p) = E(v) - p \Leftrightarrow p > \frac{E(v) - b(E(v|m_l, s_k) + c)}{\gamma}$. Because $E(v) = \frac{1}{\gamma} \cdot E(v|m_l, s_k) + (1 - \frac{1}{\gamma})E(v|m_r, s_k)$, this is equivalent to $p \geq E(v|m_l, s_k) + \frac{c}{\gamma}$. □

Proof of Lemma 2. Inequality (9) can be rewritten as $\int_0^\infty (v - p) |a(v) - a(v)| dv \geq 0 \Leftrightarrow \int_0^\infty a(v) dv \frac{dv}{2} \leq \int_0^\infty |a(v)| dv with respect to $f(v|m_l, s_k)$ from Proposition 1, which implies that $E(v|m_l, s_k) - E(v|m_l, s_k) \leq 0$. Thus, $c', 0 \leq c$, and $\mathcal{P}_H \geq \mathcal{P}_L$. Second, when $\lambda_{E_L(v) \leq E_L(v) < E_L(v)}$, for all $v$, $f(v|m_l, s_k)$ satisfies the MLRP with respect to $f(v|m_l, s_k)$ from Proposition 1, which implies that $E(v|m_l, s_k) - E(v|m_l, s_k) > 0$ and $c' > 0$. Moreover, we have $c' \leq \mathcal{P} = \min_{p \in \mathcal{P}_H} \int_0^\infty |a(v)| dv with respect to $f(v|m_l, s_k)$ from Lemma 3. Therefore, $c' \in (0, \infty)$, and $\mathcal{P}_L > \mathcal{P}_H$ if and only if $c < c'$. Finally, when $\lambda_{E_H(v) < \lambda_{E_H(v)}}$, we have $c' > 0$. Therefore, $\mathcal{P}_L > \mathcal{P}_H$. □

Proof of Proposition 3. Note that $p_L > p_R \Leftrightarrow c' < c'' \Leftrightarrow \frac{f^R(c|m_{L}, s_k)}{f^R(c|m_{L}, s_k)}$. First, when $\lambda_{E_L(v) < \lambda_{E_L(v)}}$, when $\lambda_{E_L(v) \geq \lambda_{E_L(v)}}$, we have $c'' \geq \mathcal{P}$ (that is, $c''$ is outside the range of consideration) if and only if $c' \geq \mathcal{P}$. When $\lambda_{E_L(v) \leq \lambda_{E_L(v)}}$, we have $c'' \geq \mathcal{P}$ (that is, $c''$ is outside the range of consideration) if and only if $c' \geq \mathcal{P}$. Therefore, $p_L \geq p_R$. Therefore, $\mathcal{P}_L > \mathcal{P}_H$. □
optimal under full coverage. Thus, it is optimal to induce partial coverage with \((\alpha, \beta)\) if the expected profit that it generates is larger than the highest expected profit under full coverage \((p_L - k)\). Therefore, when \(\alpha \leq \beta < \epsilon\), the highest expected profit under full coverage \((p_L - k)\) is optimal under partial coverage with \((\alpha, \beta)\). Second, when \(\beta \leq \alpha\), it is optimal to induce full coverage. Also, under partial coverage, either \(m_{\alpha, \beta}\) or \(m_{\beta, \alpha}\) can be optimal. Therefore, it is optimal to induce partial coverage if the highest expected profit under partial coverage generates a larger expected profit under full coverage than \((\alpha, \beta)\). This holds if and only if \(m_{\alpha, \beta} \geq \min_{(\alpha, \beta)} [\epsilon + (1 - \alpha) \sigma(\beta - k)]\). Otherwise, if \(m_{\alpha, \beta} < \min_{(\alpha, \beta)} [\epsilon + (1 - \alpha) \sigma(\beta - k)]\), it is optimal to induce full coverage with \((\alpha, \beta)\). Finally, it is clear from the specifications of \(k_3(c), k_2(c), k_1(c)\) that they are all increasing in \(c\). □

Proof of Proposition 5. First, when \(\epsilon_0(v) \geq \lambda_{\epsilon_2}(v)\) for all \(v\), we know from the discussion that only \((p_L, p_L)\) can be optimal under partial coverage and only \((m_{\alpha, \beta}, m_{\alpha, \beta})\) can be optimal under full coverage. Thus, it is optimal to induce partial coverage with \((\alpha, \beta)\) if the expected profit that it generates is larger than the highest expected profit under full coverage \((p_L - k)\). Therefore, when \(\alpha \leq \beta < \epsilon\), the highest expected profit under full coverage \((p_L - k)\) is optimal under partial coverage with \((\alpha, \beta)\). Second, when \(\beta \leq \alpha\), it is optimal to induce full coverage with \((\alpha, \beta)\). Finally, when \(\epsilon_0(v) < \lambda_{\epsilon_2}(v)\) for all \(v\), we know from the discussion that only \(m_{\alpha, \beta}\) or \(m_{\beta, \alpha}\) can be optimal under partial coverage. Also, under partial coverage, either \(m_{\alpha, \beta}\) or \(m_{\beta, \alpha}\) can be optimal. Therefore, it is optimal to induce partial coverage if the highest expected profit under partial coverage generates a larger expected profit under full coverage than \((\alpha, \beta)\). This holds if and only if \(m_{\alpha, \beta} \geq \min_{(\alpha, \beta)} [\epsilon + (1 - \alpha) \sigma(\beta - k)]\). Otherwise, if \(m_{\alpha, \beta} < \min_{(\alpha, \beta)} [\epsilon + (1 - \alpha) \sigma(\beta - k)]\), it is optimal to induce full coverage with \((\alpha, \beta)\). Finally, it is clear from the specifications of \(k_3(c), k_2(c), k_1(c)\) that they are all increasing in \(c\). □
### Table A.1. Summary of Notations

<table>
<thead>
<tr>
<th>Notations</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>$v, \pi$</td>
<td>Consumer $i$'s product match value and the highest product match value</td>
</tr>
<tr>
<td>$s_i \in {s_j, s_0}$</td>
<td>Consumer $i$'s private signal about product match value: a good signal $s_j$ and a bad signal $s_0$</td>
</tr>
<tr>
<td>$m_i \in {m_{l1}, m_{l2}}$</td>
<td>Low type messenger $m_{l1}$ and high type messenger $m_{l2}$</td>
</tr>
<tr>
<td>$\Pi, \Pi_{tut}, \Pi_{part}$</td>
<td>Firm's profit: profit under full coverage ($\Pi_{full}$) and partial coverage ($\Pi_{part}$)</td>
</tr>
<tr>
<td>$k$</td>
<td>Production cost</td>
</tr>
<tr>
<td>$c, \gamma$</td>
<td>Deliberation cost and the highest deliberation cost</td>
</tr>
<tr>
<td>$F, f$</td>
<td>CDF and PDF of product match value distribution</td>
</tr>
<tr>
<td>$\eta(v)$</td>
<td>Probability that the private signal is good ($s_j$) when the type of messenger is $m_i$ and the product match value is $v$</td>
</tr>
<tr>
<td>$\delta_i$</td>
<td>Average probability that the private signal is good ($s_j$) when the type of messenger is $m_i$</td>
</tr>
<tr>
<td>$\lambda, \kappa$</td>
<td>Signal elasticities when the type of messenger is $m_i$</td>
</tr>
<tr>
<td>$P_i, p$</td>
<td>Price premium under consumer attention ($\gamma$) and under no consumer attention ($p_2$) when the type of messenger is $m_i$</td>
</tr>
<tr>
<td>$\beta$</td>
<td>Threshold of price where $p \equiv \frac{\int_{v=\eta(v)-\eta_0}^{\eta(v)} df}{\eta(v)-\eta_0}$</td>
</tr>
<tr>
<td>$\begin{pmatrix} c^-, c^+ \end{pmatrix}$</td>
<td>Thresholds of deliberation cost under consumer attention $\gamma$ and under no consumer attention $p_2$</td>
</tr>
<tr>
<td>$k_n(c), n \in {1, 2, 3, \ldots}$</td>
<td>Thresholds of production cost</td>
</tr>
</tbody>
</table>

### Proof of Proposition 7
First, when $\epsilon_i(v) \geq \lambda \epsilon_i(v)$ for all $v$, we know from Table 4 that only $(m_{l1}, E[v|m_{l1}, s_j])$ can be optimal under full coverage and only $(m_{l2}, E[v|m_{l2}, s_j])$ can be optimal under partial coverage. Therefore, partial coverage with $(m_{l1}, E[v|m_{l1}, s_j])$ is optimal if the expected profit that it generates, $(\beta \bar{\beta}_i E[v|m_{l1}, s_j] - k)$ is larger than the highest expected profit under full coverage $(E[v|m_{l2}, s_j] - k)$: $k \geq \frac{E[v|m_{l1}, s_j] - \bar{\beta}_i E[v|m_{l1}, s_j]}{1 - p_0} \equiv k_{l1}$. Otherwise (i.e., $k < k_{l1}$), full coverage with $(m_{l2}, E[v|m_{l2}, s_j])$ is optimal. Next, when $\kappa \epsilon_i(v) \leq \epsilon_i(v) < \lambda \epsilon_i(v)$ for all $v$, we know that only $(m_{l1}, E[v|m_{l1}, s_j])$ can be optimal under full coverage. Therefore, partial coverage is optimal if the highest expected profit under partial coverage, $\max_{\epsilon_i(v) \leq \lambda \epsilon_i(v)} \int \beta \bar{\beta}_i E[v|m_{l1}, s_j] - k$ is larger than the highest expected profit under full coverage $(E[v|m_{l2}, s_j] - k)$: $k \geq \min_{\epsilon_i(v) \leq \lambda \epsilon_i(v)} \frac{E[v|m_{l2}, s_j] - \bar{\beta}_i E[v|m_{l2}, s_j]}{1 - p_0} \equiv k_{l2}$. Otherwise (i.e., $k < k_{l2}$), full coverage with $(m_{l1}, E[v|m_{l1}, s_j])$ is optimal. Finally, when $\epsilon_i(v) < \kappa \epsilon_i(v)$ for all $v$, we know that only $(m_{l1}, E[v|m_{l1}, s_j])$ can be optimal under partial coverage. Therefore, partial coverage is optimal if the highest expected profit under partial coverage, $(\max_{\epsilon_i(v) \leq \lambda \epsilon_i(v)} \beta \bar{\beta}_i E[v|m_{l1}, s_j] - k)$ is larger than the highest expected profit under full coverage $(E[v|m_{l2}, s_j] - k)$: $k \geq \min_{\epsilon_i(v) \leq \lambda \epsilon_i(v)} \frac{E[v|m_{l2}, s_j] - \bar{\beta}_i E[v|m_{l2}, s_j]}{1 - p_0} \equiv k_{l3}$. Otherwise (i.e., $k < k_{l3}$), full coverage with $(m_{l1}, E[v|m_{l1}, s_j])$ is optimal. □

### Endnotes

2. See https://ssir.org/articles/entry/finding_the_right_messenger_for_your_message.
4. For more examples, see H&M’s Instagram account: https://www.instagram.com/hm/ and Stitch Fix’s Instagram account: https://www.instagram.com/stitchfix/.
5. Alternatively, one can consider advertising as a signaling device where the choice of messenger signals the product quality because different messengers might require different amounts of costs to employ (for example, celebrities cost more to feature). While this may apply to many situations where firms need to communicate (vertical) quality signals, for many situations, firms also need to communicate horizontal attributes (i.e., matching value), which firms do not necessarily have any superior information than consumers. Furthermore, in our previous examples of Sensodune, L’Oréal, Abercrombie & Fitch, and Unilever, there could be other potential explanations for firms’ messenger choices, such as considerations of authenticity, brand positioning (using different messengers to differentiate from competitors to guide consumer search; Ke et al., 2022), or targeting specific audiences. However, it is important to note that messenger choice consistently impacts the information environment consumers encounter in all these campaigns. Thus, the Bayesian persuasion framework complements, rather than competes with, various roles of advertising by explaining the communication of horizontal matching information.
6. A stream of consumer psychology research has identified different elements of advertising content influence consumer’s attitudes and actions, such as fear (Johnston et al. 2015) and humor (Wolman Elpers et al. 2004). For celebrity endorsement contexts, Dhar and Stillman (2019) suggest that likability and relevance (or fit) are two critical characteristics of messenger that drive persuasion.
7. Recent papers by Ning et al. (2023) and Shin and Yu (2021) also examines how advertising affects consumer inference. Although we investigate the role of the messenger for consumer inference, they focus on how the mere fact that being targeted by advertisements can affect consumer inference about product match.
8. Dewatripont and Tirole (2005) also model the processing of both issue-relevant information and cues. In their paper, the cue is the
degree of preference alignment between a sender and a receiver, and the receiver either successfully assimilates the issue-relevant information or fails to do so. In our model, the cue, which is the type of communication messenger, determines the information structure consumers face, and consumers update their beliefs upon receiving signals drawn from that information structure.

The cognitive process occurring through peripheral and central routes is also referred to as System 1 and System 2 thinking in the literature (Kahneman 2011).

In the realm of social media marketing, social influencers can be seen as messengers who possess a public record of signals, such as product reviews which can vary significantly among different influencers, even within the same product category and price range. This variation in signals stems from the influencers’ types, characterized by their consistent tendencies to either provide overly positive (praising) or overly negative (criticizing) reviews of products. As a result, the types and signal structures of influencers are readily observable. We thank the associate editor for this suggestion.

Other ways to manipulate the deliberation cost include adjusting the ad’s length, changing the content’s readability (format), and so on. We abstract away those features in ad content to focus on the role of the messenger.

On the contrary, in Kamenica and Gentzkow (2011), the sender only chooses a signal structure and does not consider the price. Thus, intuitively, the deliberation cost plays little to no role in their paper because once the receiver’s deliberation cost is sunk, the optimal signal structure should be the same as the one without deliberation cost. However, in our model, the firm simultaneously chooses both the type of messenger and the price to maximize its profit. In such a setting, it becomes crucial to carefully examine the impact of deliberation cost on the price premiums associated with different messengers and the expected profit of the firm.

We impose the assumption that \( e < \sigma = \min_{j \in [L, H]} \left( \int_{v} c[v|\sigma(v) - \hat{\sigma}_j]dF \right) \), so the firm can set prices to induce consumers’ attention for any given type of messenger. Otherwise, if the deliberation cost is too high, it might never be optimal for consumers to pay attention to the ad regardless of the price. As a result, the problem becomes trivial.

In a model extension in Section 6.3, we consider a case where the price is ex ante unobservable and can be known only after consumers incur a search cost. We demonstrate that the key findings and insights remain consistent in this setup.

If consumers make identical purchase decisions upon different signals, they would choose not to incur any deliberation cost. Hence, if consumers ever choose to pay attention, they must make different choices upon receiving different signals.

The posterior beliefs are Bayes plausible as defined in Kamenica and Gentzkow (2011). That is, the expected posterior belief equals the prior belief: \( \sigma_j(f|m_j, s_j) + (1 - \sigma_j)f(m_j, s_j) = f(m_j | \sigma_j + f(v)(1 - \sigma_j)f(v)) \).

We begin with the entire posterior distribution, as outlined in Equations (3) and (4), to establish a stronger and more intuitive connection between the model and managerially relevant interpretation. As previously mentioned, signal structures can alternatively be characterized by considering intercepts (related to the general positivity of the signal) and slopes (reflecting signal elasticities and informativeness). In practice, firms naturally recognize that a high-type (more attractive) messenger typically has a higher probability of generating favorable signals, similar to a higher intercept (as exemplified by Baker and Churchill (1977)). Nonetheless, what remains less clear is the specific sensitivity of the slope concerning a product’s match value—essentially, how much the probability changes as the match value increases. This level of sensitivity (or slope) is not immediately evident to firms and necessitates a more comprehensive explanation to delve into the relationship between the messenger’s underlying signal-generating characteristics, as illustrated in Figures 3 and 4.

Mayzlin and Shin (2011) also characterize the consumers’ search decisions and highlight the relationship between the uncertainty and price range that can encourage consumer search. In their model, a consumer can engage in a costly search to acquire an additional signal about the product’s quality. The quality is the firm’s private information, and the additional information comes from an external source, such as word of mouth. Thus, advertising serves as a signaling device. In contrast, we focus on the role of advertising messenger in the same ad, and the firm has no additional private information. Therefore, the advertising serves as a Bayesian persuasion (not signaling) device, where the firm commits to a certain signal structure by choosing the type of messenger.

Note that \( E[\sigma_j|v] = \frac{\partial \psi_d}{\partial \sigma_j} \), where \( \psi_d = \frac{\partial}{\partial \sigma}(\beta \sigma - \gamma \sigma - \beta \sigma) \) is the elasticity of bad signal from messenger \( m_j \). Therefore, the more elastic a type of messenger’s good signals are, the more elastic their bad signals are.

In contrast, we impose a signal-eliciting device, where the firm commits to a certain signal structure by choosing the type of messenger.

In certain situations, the firm may possess the flexibility to design a messenger’s elasticities. Nevertheless, our current model specifically centers on scenarios in which the elasticities of different messengers are predetermined. In such instances, the firm’s central decision variable revolves around messenger selection. This indeed presents a limitation in our study. To address this limitation, we introduce a streamlined model in Online Appendix A, where the firm can indeed design the optimal signal structure.

We first assume that consumers can costlessly observe the price before they make attention decisions, and later we analyze an alternative scenario where the price is unobservable prior to deliberation in Section 6.3.

As \( \alpha \rightarrow 1 \), \( D(m_j|p) \rightarrow \hat{\sigma}_j \) for all prices \( p \leq \bar{v} \), making \( \bar{v} \) the optimal price. Then, a high-type messenger is optimal due to its higher likelihood to generate good signals (\( \hat{\sigma}_H > \hat{\sigma}_L \)) and the lack of consideration of different signal-generating processes by naive consumers.

Intuitively, no coverage case where the firm targets only naive consumers becomes an optimal choice, making the high-type messenger a more promising option. It is easy to see that the profit is increasing in \( \alpha \) under the no coverage case: \( \frac{\partial p|m_j|v|\sigma_j}{\partial \alpha} = \hat{\sigma}_j(\bar{v} - k) > 0 \), constant in the partial coverage case: \( \frac{\partial p|m_j|v|\sigma_j}{\partial \alpha} = 0 \), and decreasing in the full coverage case: \( \frac{\partial p|m_j|v|\sigma_j}{\partial \alpha} = -(1 - \hat{\sigma}_j)p(k - k) < 0 \). Thus, as \( \alpha \) becomes larger, the profit under no coverage case dominates the other cases.

Without naive consumers, the consumer demand for full coverage without inducing consumer attention is always greater than partial coverage with consumer attention. However, introducing naive consumers can disrupt such hierarchy. Thus, the condition ensures the demand hierarchy among different coverage cases stays constant. The partial coverage demand is smaller than the full coverage, and the demand is smaller in the no coverage than in the partial coverage: \( \hat{\sigma}_j + (1 - \alpha) \geq \hat{\sigma}_H \Rightarrow \hat{\sigma}_j \leq \frac{1}{1 - \alpha} \) and \( \hat{\sigma}_j > \bar{\alpha} \hat{\sigma}_H \Rightarrow \hat{\sigma}_j \leq \frac{1}{1 - \alpha} \).

We characterize the entire parameter space for the optimal choice of coverage and messenger in the online appendix.

We assume that \( \beta \) is large enough so that either \( p = E(v|m_j, s_j) \) or \( p = E(v|m_j, s_j) \) is optimal. Specifically, we assume that \( \beta \geq \beta \equiv \max_{j \in [L, H]} \left( \beta \right. \left. \min \left\{ \frac{\sigma - \gamma}{\partial \sigma_j v(m_j | s_j)}, \min \left\{ \frac{\partial \sigma_j v(m_j | s_j)}{\partial \sigma_j v(m_j | s_j)} \right. \right\} \right) \).

When \( \beta < \beta \), there can be regions of parameters where a pure-strategy equilibrium does not exist.
Our monotonic relationship assumption implies that, regardless of the true matching value, one type’s slope is consistently steeper than the other’s. It can serve as a pragmatic approximation aligning with the reality we aim to capture. For example, it notably can represent situations like the halo effect, where a positive impression of the messenger influences an individual’s overall perception of the product irrespective of the precise match value. Moreover, when considering a binary distribution of match values, we can cover all possible scenarios and show that the results still hold.

References
Lin S (2022) The medium of advertising. Working paper, Hong Kong University of Science and Technology, Hong Kong.