The Role of Messenger in Advertising Content: Bayesian Persuasion Perspective

Jiwoong Shin and Chi-Ying Wang*

June 24, 2023

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 costless 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

*Jiwoong Shin: Professor of Marketing, School of Management, Yale University, jiwoong.shin@yale.edu, Chi-Ying Wang: Ph.D. candidate, School of Management, Yale University, chi-ying.wang@yale.edu. We 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 their valuable feedback and comments. The authors contributed equally, and their names are listed in alphabetical order.
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 forty 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 (Eisend and Langner, 2010; Joseph, 1982; Liu et al., 2007) documented that physically attractive messengers positively affect consumers’ attitudes, product evaluation, and purchase intention. Also, a stream of literature (Ohanian, 1991; Schouten et al., 2020; Till and Busler, 1998) suggests that the messenger’s perceived relevance to the advertised product is essential in affecting purchase behavior. Thus, using attractive or more relevant messengers seems appropriate to increase advertising effectiveness.

\(^1\)https://hashtagpaid.com/banknotes/ugly-ads-dont-mean-bad-ads-try-these-expert-tips-for-high-intent-ads.
\(^2\)https://ssir.org/articles/entry/finding_the_right_messenger_for_your_message
The industry practices are mixed. Traditionally, many advertising campaigns tend to have endorsements from physically attractive actors or celebrities. However, not all marketers have agreed that the idealized images of such endorsers are the most effective in advertising. In 2004, Unilever developed several Real Beauty campaigns that feature real women, not models.\(^3\) Although the positive results of those campaigns may be due to their provocative nature of novelty (Vézina and Paul, 1997), advertising industries have witnessed the rise of ads featuring real people for their products, notwithstanding the lack of novelty it used to garner in the early days. For example, Stitch Fix has implemented a social media advertising campaign that highlights “real” or “normal” individuals rather than relying on traditional models. However, despite the growing trend towards more realistic advertisements, it is worth noting that nearly 20% of television ads still feature physically attractive celebrities.\(^4\) Furthermore, companies like H&M persist in using attractive models to promote their products on platforms such as Instagram. (See Figure 1).\(^5\)

This raises an important question: which types of messengers, attractive or normal individuals, 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. In this study, our objective is 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

\(^4\)https://www.marketing-schools.org/consumer-psychology/marketing-with-celebrities/#section-1
\(^5\)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/.
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, while 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, since 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.\(^6\) 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 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 proceed to compare the expected profits in order 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 as the product cost rises, the need to raise the price becomes more significant. 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 in order to maintain a greater demand. Hence, inducing consumer attention is deemed optimal when the consumer’s deliberation cost decreases or when the product cost increases.

After establishing the optimal decision to induce consumer attention, we demonstrate that the optimal choice of messenger depends on the signal elasticity (or informativeness) of each messenger 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, due to 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 product 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, due to 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 utilize 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. While 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 the specific manner in which 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. We present the main analyses and results focusing only on the persuasion role of the messenger in Section 4. In Section 5, we demonstrate the robustness of the main insights by analyzing a number of extensions that relax several key model assumptions. Section 6 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 (see Nelson 1974 and 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 (Chaiken and Maheswaran, 1994; Till and Busler, 1998; Zajonc, 1968). For example, Chaiken and Maheswaran (1994) suggest that having the fit between the messengers and the endorsed products tends to make their audience assume they will make valid arguments and, therefore, is predisposed to believe what they say and like what they are selling. Also, many studies on the attractiveness model have established the notion that attractive messengers can have a positive effect on advertising effectiveness (see Pornpitakpan 2004 for the review on this topic). According to this model, the physical attractiveness of a messenger acts as a persuasive cue, resulting in heightened attention, liking, and credibility of the message and their impact on persuasion (Eisend and Langner, 2010; Joseph, 1982; Liu et al., 2007). For example, Till and Busler (1998) have shown that attractive messengers tend to elicit more positive responses from consumers and can enhance the persuasive impact of advertising. Building on these findings, we assume two types of messengers: a high-type (such as highly attractive) messenger and a low-type one. We then analyze how these two types of messengers influence the persuasiveness of advertising communication.

---

7 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 (Woltman 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.
Our study also incorporates the dual-mode of communication perspective for persuasion in the psychology literature (Chaiken and Maheswaran, 1994; Kahneman, 2011; Petty and Cacioppo, 1986). 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). On the other hand, the peripheral route often entails issue-irrelevant cues and requires little to no cognitive effort from the decision-makers.

Drawing on insights from the psychology literature, our study contributes to a deeper understanding of the multifaceted nature of persuasion and provides valuable implications for designing persuasive advertising strategies that effectively influence consumers.

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 (Gardete and Guo, 2021; Guo, 2016; Guo and Wu, 2016; Guo and Zhang, 2012; Kuksov and Villas-Boas, 2010; Li et al., 2019; Lu and Shin, 2018; Wathieu and Bertini, 2007). 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 upon these streams of existing literature. We recognize that consumers can costlessly observe the advertising messenger through the peripheral route, which is influenced by...
attractiveness. On the other hand, evaluating the product-match signal requires more effort and careful consideration through the central route, where consumers incur deliberation costs to process the message content. By integrating these two modes, we investigate how costly deliberation impacts 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 (Bergemann and Morris, 2019; Gentzkow and Kamenica, 2016, 2017; 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. In 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) studies 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, while 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, such as attractiveness. These characteristics influence 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 \( F[0, \bar{v}] \) 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.\(^\text{11}\) 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. While the brand’s reputation may establish the product’s quality, the crucial factor lies in the product’s fit or match. However, since 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.

The 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,

\(^\text{11}\)In practice, consumers perceive the product either as a good match (a good signal) or not a good fit for their needs (a bad signal) after deliberation. Those impressions are signals that consumers receive.
credibility, and expertise, can influence the effectiveness of an ad (Dhar and Stillman, 2019; Till and Busler, 2000). 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. For example, many beauty industry firms tend to feature good-looking celebrities to affect consumers’ attitudes toward the brand by borrowing the celebrities’ personal images or attractiveness (Eisend and Langner, 2010; Joseph, 1982; Till and Busler, 1998).\textsuperscript{12} 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:

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

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

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

This assumption implies that consumers with a higher (true) match value will more likely receive a good signal. Therefore, a positive signal $s_g$ is really “good news” regarding the true match value (Milgrom, 1981). Also, a high-type messenger is more likely to generate a good signal than a low-type messenger for any level of true value $v$. It is important to clarify that the assumption is only focused on high-type messengers being more likely to generate private good signals about the product match values. However, this does not necessarily imply that high-type messengers are inherently better at convincing or persuading consumers to purchase the product. The Bayesian persuasion framework offers valuable insights into understanding the circumstances under which different types of messengers can be more effective in influencing consumer behavior.

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

\textsuperscript{12}In the realm of social media marketing, social influencers can be seen as messengers. They 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 AE for this suggestion.
messenger as

\[ \tilde{\sigma}_H \equiv \Pr(s_i = s_g|m_H) = \int_0^\theta \sigma_H(v) dF, \quad \tilde{\sigma}_L \equiv \Pr(s_i = s_g|m_L) = \int_0^\theta \sigma_L(v) dF. \] (2)

Then, \( \tilde{\sigma}_H > \tilde{\sigma}_L \). We use these two terms \( \tilde{\sigma}_H, \tilde{\sigma}_L \) in our subsequent analysis.

We also note that 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.\(^{13}\) 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.\(^{14}\)

**Dual mode of communication**

Consumers need different cognitive loads to process different types of information in advertising. We take a dual-mode of communication perspective, where the information is conveyed through two different routes (Petty and Cacioppo, 1986). 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 might be the first thing consumers notice even before

\(^{13}\) 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.

\(^{14}\) 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 in order 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.
they start thinking about the product information. Such ostensible information about the messenger can be communicated through mindless processing (Kahneman, 2011; Petty and Cacioppo, 1986). On the other hand, the more costly central route of persuasion requires thoughtful consideration of advertising contents, such as the personal relevance to a product or the product’s match value (Guo, 2016; Guo and Zhang, 2012; Wathieu and Bertini, 2007). 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, they 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, \( c > 0 \).\(^{15}\) 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, she 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

![Figure 2: Game Sequence](image-url)

---

\(^{15}\)We impose the assumption that \( c < \bar{c} \equiv \min_{j \in \{L, H\}} \int v_i \sigma_j(v) - \tilde{\sigma}_j dF \), so the firm can set prices to induce consumers’ attention for any given type of messenger. Otherwise, the deliberation cost is too high, and it might never be optimal for consumers to pay attention to the ad regardless of the price. As a result, the problem becomes trivial.
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.\(^{16}\) If a consumer decides to pay attention, she incurs a deliberation cost \( c \) and receives a noisy private signal \( s_i \in \{s_g, s_b\} \) 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 the consumer actually pays attention to the advertising by incurring a deliberation cost, let alone what signal the consumer ultimately receives if she chooses 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 introduce and summarize all the basic notations in this paper in Table 1.

\(^{16}\)In a model extension in Section 5.3, we consider a case where the price is ex-ante unobservable and can be known only after they incur a search cost. We demonstrate that the key findings and insights remain consistent in this setup.

<table>
<thead>
<tr>
<th>Notations</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>( i )</td>
<td>consumer ( i )</td>
</tr>
<tr>
<td>( j \in {L, H} )</td>
<td>messenger type (low and high)</td>
</tr>
<tr>
<td>( k )</td>
<td>production cost</td>
</tr>
<tr>
<td>( c, \bar{c} )</td>
<td>deliberation cost and the highest deliberation cost</td>
</tr>
<tr>
<td>( v_i, \tilde{v} )</td>
<td>product match value and the highest product match value</td>
</tr>
<tr>
<td>( F, f )</td>
<td>CDF and PDF of product match value distribution</td>
</tr>
<tr>
<td>( s_i \in {s_g, s_b} )</td>
<td>consumer’s private signal about product match value: a good signal ( s_g ) and a bad signal ( s_b )</td>
</tr>
<tr>
<td>( m_j \in {m_L, m_H} )</td>
<td>low type messenger ( m_L ) and high type messenger ( m_H )</td>
</tr>
<tr>
<td>( \sigma_j(v) )</td>
<td>probability that the private signal is good ( (s_g) ) when the type of messenger is ( m_j ) and the product match value is ( v )</td>
</tr>
<tr>
<td>( \bar{\sigma}_j )</td>
<td>average probability that the private signal is good ( (s_g) ) when the type of messenger is ( m_j )</td>
</tr>
<tr>
<td>( \varepsilon_j )</td>
<td>signal elasticities when the type of messenger is ( m_j )</td>
</tr>
<tr>
<td>( \bar{p}_j, p_j )</td>
<td>price premium under consumer attention ( (\bar{p}_j) ) and under no consumer attention ( (p_j) ) when the type of messenger is ( m_j )</td>
</tr>
<tr>
<td>( \lambda, \kappa )</td>
<td>thresholds of deliberation cost under consumer attention ( (c^* \equiv \frac{\sigma_L(v)}{\sigma_H(v)} \bar{p} \int_0^v \sigma_H(v)dF - \sigma_L(v) \int_0^v \sigma_L(v)dF) ), and under no consumer attention ( (c^{**} \equiv \frac{(1-\sigma_H) \int_0^v [1-\sigma_L(v)]dF - (1-\sigma_L) \int_0^v [1-\sigma_H(v)]dF)}{\sigma_H - \sigma_L} )</td>
</tr>
<tr>
<td>( k_n(c), n \in {1, 2, 3, \ldots} )</td>
<td>thresholds of production cost</td>
</tr>
</tbody>
</table>

Table 1: Summary of Notations
4 Analysis

When the firm evaluates the effectiveness of a messenger, it compares the expected profits from different types of messengers, which are affected by (1) the signal structure generated by a messenger, (2) the messenger’s ability to attract consumer attention, and (3) the 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. Then, we study the mechanism of how a messenger can attract consumer attention and potentially raise price premiums. In our main model, we fix the messenger’s attraction ability, which directly affects the consumer’s deliberation cost. Nevertheless, as we show in 4.2, the messenger’s type can still affect the consumers’ attention decision rationally through influencing the consumers’ inference. In our extension, we consider the direct attention-grabbing effect by incorporating the messenger’s differential 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.

4.1 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 $s_i$, a consumer’s posterior beliefs about the underlying match value distribution follow the next two densities:\footnote{Note that the posterior beliefs are Bayes plausible as defined in Kamenica and Gentzkow (2011). That is, the expected posterior belief equals the prior belief: $\bar{\sigma}_j f(v|m_j, s_b)+(1-\bar{\sigma}_j) f(v|m_j, s_g) = f(v)\sigma_j(v)+(f(v)[1-\sigma_j(v)] = f(v)$.}

$$f(v|m_j, s_g) = \frac{f(v) \cdot \sigma_j(v)}{Pr(s_i = s_g|m_j)} = \frac{f(v) \cdot \sigma_j(v)}{\bar{\sigma}_j}, \quad (3)$$

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

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 $s_g$, 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, \( f(v|m_H,s_g) \) is increasing in \( v \)) if and only if, for all \( v \),

\[
\varepsilon_H(v) \equiv \frac{d\sigma_H(v)/dv}{\sigma_H(v)/v} \geq \frac{d\sigma_L(v)/dv}{\sigma_L(v)/v} \equiv \varepsilon_L(v). \tag{5}
\]

On the other hand, if \( \varepsilon_H(v) < \varepsilon_L(v) \) for all \( v \), the posterior belief from the low-type messenger $m_L$ satisfies the MLRP with respect to the high-type messenger $m_H$: \( f(v|m_H,s_g) \) is decreasing in \( v \).

The elasticities \( \varepsilon_H(v) \) and \( \varepsilon_L(v) \) measure how sensitive a messenger’s good signal is to changes in match values, and their relationship determines the informational environments that consumers face.\(^1\) 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 \( v \): \( \frac{\partial \sigma}{\partial v} \geq 0 \), and (2) \( \sigma_H(v) > \sigma_L(v) \) for all \( v \). Therefore, receiving a good signal \( s_g \) always improves the posterior belief about the product match value \( f(v|m_j,s_g) \) 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

\(^{1}\)Alternatively, the signal structures can be characterized by assuming the signal elasticities, denoted as \( \varepsilon_H(v) \) and \( \varepsilon_L(v) \). Nevertheless, we believe it would be more effective to start from the entire posterior distribution, as presented in Equations (3) and (4), in order to establish a stronger and intuitive connection between the model and tangible, managerially relevant variables. In reality, firms may have a general understanding of which type of messengers are more or less likely to generate positive impressions about the product, which our current approach captures. However, it is challenging to determine how sensitive a messenger’s positive signal is to changes in match values. It is not immediately clear how we can apply such situations suggested by the model to practical scenarios or precisely define the elasticity of the messenger in practice. To address this issue, a more elaborate explanation is required, which necessitates exploring the relationship between the messenger’s underlying signal-generating characteristics, as demonstrated in Figures 3 and 4.
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 upon the interplay between the two elasticities of signal structures.

First, when $\varepsilon_H(v) \geq \varepsilon_L(v)$ (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 $v$. 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 $s_g$ 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. Under this information environment, high-type messengers tend to produce predominantly positive signals irrespective of the match value, while 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 ($s_g$) 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 $v$.

### 4.2 Consumer attention

When a consumer receives an ad, she decides whether to pay attention to the ad by costly deliberating about the content (Guo, 2016; Guo and Zhang, 2012; Wathieu and Bertini, 2007). If the additional information from the advertisement won’t change her 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

$$EU(\text{attention}|m_j) = \tilde{\sigma}_j \cdot [E(v|m_j, s_g) - p] - c$$

(6)

On the other hand, even when she 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

$$EU(\text{no attention}) = \max\{0, E(v) - p\}$$

(7)

We characterize the consumer’s attention decision in the following lemma.$^{19}$

$^{19}$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
Lemma 1. Consumers pay attention to the ad if and only if the price falls within a moderate range such that

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

where \( p_j < \overline{p}_j \) 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, i.e., \( c < (1-\bar{\sigma}_j)[p - E(v|m_j, s_b)] \), which can be simplified to \( p > E(v|m_j, s_b) + \frac{c}{1-\sigma_j} \). On the other hand, 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 \( c \leq \bar{\sigma}_j[E(v|m_j, s_g) - p] \), which can be simplified to \( p \leq E(v|m_j, s_g) - \frac{c}{\bar{\sigma}_j} \).

Combining these cases, consumers pay attention to the ad if and only if the price falls within a moderate range given by the lemma: \( p_j < p \leq \overline{p}_j \). This holds under the assumption \( c < c \equiv \min_{j \in \{L,H\}} \int_0^v v[\bar{\sigma}_j - \sigma_j(v)] dF \). 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 decision. 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 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.
utility and their attention decision. Irrespective of the actual signal realized, whether it is $s_i \in \{s_g, s_b\}$, a high-type messenger can generate higher ex-ante expected utility compared to a low-type messenger if $EU(\text{attention}|m_H) \geq EU(\text{attention}|m_L)$, which is equivalent to

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

(9)

**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 $p \leq \tilde{p} \equiv \frac{\int_{0}^{\tilde{\sigma}_H(v) - \tilde{\sigma}_L(v)} dF}{\tilde{\sigma}_H - \tilde{\sigma}_L}$.

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 $\tilde{\sigma}_j = \Pr(s_i = s_g|m_j)$, 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. Lemma 2 demonstrates that this holds when the price falls below a certain threshold $\tilde{p}$ that satisfies $\int_{0}^{\tilde{p}} (v - \tilde{p}) \cdot [\sigma_H(v) - \sigma_L(v)] dF = 0$.

Thus far, we have established the conditions for consumer attention, which are based on the price range where the benefits of paying attention outweigh the costs (Lemma 1). Furthermore, we have examined the conditions under which different types of messengers can more effectively induce consumer attention (Lemma 2). These factors provide valuable insights into understanding the optimal choice of messenger, but we still need to consider one more factor: the price premium associated with each messenger type. By taking this additional factor into account, we can gain a comprehensive understanding of the optimal messenger choice.

### 4.3 Price premiums

Given messenger type $m_j \in \{m_L, m_H\}$, consumers decide to pay attention when $p_j < p \leq \bar{p}_j$. 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 $\bar{p}_j = E(v|m_j, s_g) - \frac{\tilde{\sigma}_j}{\tilde{\sigma}_j}$. These highest prices are denoted
as \( \bar{p}_H \) and \( \bar{p}_L \), corresponding to different messenger types. We note that the highest price \( p_j = E(v|m_j, s_g) − \frac{\sigma_j}{\sigma_j} \) 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 \( \frac{\sigma_j}{\sigma_j} \). This component captures the trade-off 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 \( (\frac{\sigma_H}{\sigma_L} < \frac{\sigma_L}{\sigma_L}) \). This is because the average positive signal is always higher for the high-type messenger, resulting in a higher effective expected deliberation cost for the low-type messenger. This difference negatively affects the low-type messenger’s ability to set a higher price compared to 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 \( \lambda \equiv \frac{\sigma_L(v)}{\sigma_H(v)} \), a low-type messenger can command a higher price \( (\bar{p}_L > \bar{p}_H) \) if and only if the deliberation cost is sufficiently small such that \( c < c^* \equiv \frac{\bar{\sigma}_H \int_{v_0}^{\bar{v}} v \sigma_L(v) dF - \bar{\sigma}_L \int_{v_0}^{\bar{v}} v \sigma_H(v) dF}{\bar{\sigma}_H - \bar{\sigma}_L} \).
3. When \( \lambda \varepsilon_L(v) > \varepsilon_H(v) \), a low-type messenger can always command a higher price \( (\bar{p}_L > \bar{p}_H) \).

The proposition highlights that different messengers can induce a higher price premium depending on the signal environments under consumer attention. First, when \( \lambda \varepsilon_L(v) \leq \varepsilon_H(v) \) (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 \( \bar{p}_H \geq \bar{p}_L \). 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 \( \bar{p}_L \) to be larger than \( \bar{p}_H \) 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. Lastly,
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 $p_H$ and $p_L$.

**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 $\kappa \equiv \frac{\sigma_L(v)(1-\sigma_H(v))}{\sigma_H(v)(1-\sigma_L(v))}$, 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 $c < c^{**} \equiv \frac{(1-\sigma_H) \int_0^\infty v[1-\sigma_L(v)] dF - (1-\sigma_L) \int_0^\infty v[1-\sigma_H(v)] dF}{\sigma_H - \sigma_L}$.

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

Again, the highest price premium $p_j \equiv E(v|m_j, s_b) + \frac{c}{1-\sigma_j}$ for each messenger type $m_j$ 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 $E(v|m_j, s_b)$. This component reflects how consumers adjust their expectations about the match value based on the messenger’s signal. Since a bad signal tends to make consumers more pessimistic about the match value, this component has a negative impact on the price premium. The second component is the net deliberation cost per unit of bad signal, represented as $\frac{c}{1-\sigma_j}$. This component captures the cost consumers incur in the decision-making process when they receive 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 \( \frac{c}{1-\sigma_H} > \frac{c}{1-\sigma_L} \). This implies that more elastic or informative signals from the messenger tend to exert downward pressure on the price, while 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 to 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 to the low-type messenger under no consumer attention. Lastly, when \( \kappa \varepsilon_L(v) > \varepsilon_H(v) \), 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.

20Note that \( |\varepsilon_b^j(v)| = \frac{\sigma_j}{1-\sigma_j} \varepsilon_j(v) \), where \( \varepsilon_b^j(v) = \frac{d(1-\sigma_j(v))/dv}{(1-\sigma_j(v))/v} \) 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. So, \( \lambda \varepsilon_L(v) \leq \varepsilon_H(v) \) for all \( v \) also implies that high-type messengers’ bad signals are sufficiently sensitive to changes in match values compared to that of low-type messengers.

Figure 6: Price premiums under no consumer attention
4.4 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:

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

(10)

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.\(^{21}\) 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.

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

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

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 price

\(^{21}\)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 5.3.
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 IR 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 $p_L < p \leq p_H$ 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 $D(p,m_j) = \tilde{\sigma}_j = \Pr(s_i = g|m_j)$, 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:

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

subject to $p_L < p \leq p_H$ and $k \leq p$.

Similar to the full coverage case, the optimal price for partial market coverage must satisfy the constraint $p \leq p_j$ 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 trade-off between demand $\tilde{\sigma}_j$ (where $\tilde{\sigma}_L < \tilde{\sigma}_H$) and profit margin $p_j - k$. When $p_H \geq p_L$, it is optimal for the firm to choose a high-type messenger. However, when $p_H < p_L$, 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 $p_L - k$ is sufficiently larger than $p_H - k$. This condition can be expressed as

$$E\Pi_{\text{part}}(m_H) = \tilde{\sigma}_H \cdot (p_H - k) < E\Pi_{\text{part}}(m_L) = \tilde{\sigma}_L \cdot (p_L - k) \Leftrightarrow p < k$$

Since 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 \( \lambda \equiv \frac{\sigma_L(v)}{\sigma_H(v)} \), we have \( \tilde{p} < E(v) \). Otherwise, when \( \varepsilon_H(v) \geq \lambda \varepsilon_L(v) \), \( \tilde{p} \geq E(v) \).

From the above discussion, we can summarize the optimal choice of messenger and price under full and partial coverage cases in the following table.

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

Table 2: Optimal choice of messenger and price \((m^*_j, p^*_j)\) under full and partial coverage

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.**
1. When \( \lambda \varepsilon_L(v) < \varepsilon_H(v) \), inducing partial coverage with \((m_H, \bar{p}_H)\) is optimal if \( k \geq k_1(c) \), otherwise inducing full coverage with \((m_L, \bar{p}_L)\) is optimal.
2. When \( \varepsilon_L(v) \leq \varepsilon_H(v) < \lambda \varepsilon_L(v) \), inducing partial coverage is optimal if \( k \geq k_2(c) \). In this case, \((m_L, \bar{p}_L)\) is optimal if and only if \( k > \tilde{p} \). Otherwise, if \( k < k_2(c) \), inducing full coverage is optimal, where \((m_L, \bar{p}_L)\) is optimal if and only if \( c < c^{**} = \frac{(1-\sigma_H) \int_0^v [1-\sigma_L(v)]dF - (1-\sigma_L) \int_0^v [1-\sigma_H(v)]dF}{\sigma_H-\sigma_L} \).
3. When \( \kappa \varepsilon_L(v) > \varepsilon_H(v) \), inducing partial coverage is optimal if \( k \geq k_3(c) \). In this case, \((m_L, \bar{p}_L)\) is optimal if and only if \( k > \tilde{p} \). Otherwise, if \( k < k_3(c) \), inducing full coverage with \((m_H, \bar{p}_H)\) is optimal.

Moreover, all thresholds are increasing in \( c \): \( \frac{\partial k_2(c)}{\partial c} \geq 0 \), \( \frac{\partial k_3(c)}{\partial c} \geq 0 \), and \( \frac{\partial k_3(c)}{\partial c} \geq 0 \).

---

^22 We illustrate the main results with a numerical example. Suppose the consumers’ product match values are drawn from a Bernoulli distribution with the parameter 0.3. That is, \( v_i = 1 \) with probability 0.3 and \( v_i = 0 \) with probability 0.7, so \( E(v) = 0.3 \). The messengers are associated with the signal structures: \( \sigma_L(0) = 0.2, \sigma_L(1) = 0.3, \sigma_H(0) = 0.6, \) and \( \sigma_H(1) = \sigma \). Therefore, \( \sigma_L = 0.23 \) and \( \sigma_H = 0.42 + 0.3\sigma \). Also, \( \varepsilon_H(v) = \varepsilon_L(v) \iff \sigma = 0.9 \), \( \varepsilon_H(v) = \lambda \varepsilon_L(v) \iff \sigma = 0.7 \), and \( \varepsilon_H(v) = \kappa \varepsilon_L(v) \iff \sigma = 0.65 \). First, when \( \sigma = 0.8 \), \( \lambda \varepsilon_L(v) \leq \varepsilon_H(v) \). So, partial coverage with \((m_j^*, p^*) = (m_H, 0.36 - 1.52c)\) is optimal if \( k \geq k_1(c) \equiv 0.1 + 6.76c \). Otherwise, full coverage with \((m_j^*, p^*) = (m_L, 0.27 + 1.3c)\) is optimal. Next, when \( \sigma = 0.68 \), \( \kappa \varepsilon_L(v) \leq \varepsilon_H(v) < \lambda \varepsilon_L(v) \). Thus, partial coverage...
Figure 7: Optimal choice of market coverage and messenger

Figure 7 visually presents the insights from Proposition 4, with the darker areas in each sub-figure 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 in order to increase profit margin. On the other hand, 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 $\overline{p}_j$ 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 elasticity and the firm’s coverage decision, which, in turn, depends on the consumer’s deliberation cost and production cost. is optimal if $k \geq k_3(c) \equiv \min\{\max\{0.24 + 2.99c, 0.21 + 4.75c\}, \max\{0.18 + 6.11c, 0.14 + 9.73c\}\}$. Otherwise, full coverage is optimal. Finally, when $\sigma = 0.6$, $k_{\epsilon L}(v) > \epsilon_H(v)$. Therefore, partial coverage is optimal if $k \geq k_3(c) \equiv \min\{0.27 + 4.55c, 0.3 + 8.75c\}$. Otherwise, full coverage with $(m^*_j, p^*) = (m_H, 0.3 + 2.5c)$ is optimal.
In cases where a high-type messenger is highly informative \((\varepsilon_H(v) \geq \lambda \varepsilon_L(v))\), as shown in Figure 7-(a)), using a high-type messenger ad is more effective in raising prices. Therefore, it is optimal for the firm to attract consumer attention and increase prices when the production cost \(k\) is relatively high compared to the deliberation cost \(c\) (i.e., \(k \geq k_1(c)\)). However, due to the high informativeness of the high-type messenger’s signals, the firm needs to set the price very low in order to discourage consumer attention. As a result, using a low-type messenger becomes optimal for achieving full coverage without consumer attention.

In cases where a high-type messenger becomes less informative \((\varepsilon_H(v) < \lambda \varepsilon_L(v))\), as shown in both Figure 7-(b) and 7-(c), the effectiveness of a high-type messenger ad in updating consumers’ beliefs about the product match is limited. This is because the high-type messenger primarily provides good signals, which can overshadow the product itself. 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 \((k > \tilde{p})\). 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 \(\kappa \varepsilon_L(v) \leq \varepsilon_H(v) < \lambda \varepsilon_L(v)\) and \(c \leq c^{**}\) (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 to H&M. Consequently, utilizing normal people or non-attractive 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, while H&M uses attractive models to induce full coverage without relying on consumer attention.
5 Extensions

5.1 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. Due to 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 $E(v|m_H, s_g)$ when receiving a good signal or $E(v|m_H, s_b)$ when receiving a bad signal. The demand function associated with a high-type messenger ad can be defined as follows:

$$D(p, m_H) = \begin{cases} 
1, & p \leq E(v|m_H, s_b) \\
\bar{\sigma}_H, & E(v|m_H, s_b) < p \leq E(v|m_H, s_g) \\
0, & p > E(v|m_H, s_g) 
\end{cases}.$$  

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 $E(v|m_H, s_b) \equiv \bar{p}_H^*$ under full coverage and $E(v|m_H, s_g) \equiv \bar{p}_H^*$ under partial coverage. Under full coverage, a low-type messenger ad is optimal if (1) $\varepsilon_H(v) \geq \kappa\varepsilon_L(v)$ or (2) $\varepsilon_H(v) < \kappa\varepsilon_L(v)$ and $c > (1 - \bar{\sigma}_L)(E(v|m_H, s_b) - E(v|m_L, s_b)) \equiv c^{**}$. Also, under partial coverage with consumer attention, a low-type messenger ad is optimal if

$$\bar{\sigma}_L \left( E(v|m_L, s_g) - \frac{c}{\bar{\sigma}_L} - k \right) > \bar{\sigma}_H(E(v|m_L, s_g) - k) \iff k > \bar{p} + \frac{c}{\bar{\sigma}_H - \bar{\sigma}_L},$$  

where $\bar{p} = \frac{\int v^\varepsilon_L(v) - \varepsilon_L(v)]dF}{\bar{\sigma}_H - \bar{\sigma}_L}$ 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.**

1. If \( \lambda \varepsilon_L(v) \leq \varepsilon_H(v) \), inducing partial coverage with \((m_H, \bar{p}_H^*)\) is optimal when \( k \geq k_4(c) \). Otherwise, full coverage with \((m_L, \bar{p}_L^*)\) 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 \((m_L, \bar{p}_L^*)\) is optimal. Furthermore, under partial coverage, choosing a low-type messenger with \((m_L, \bar{p}_L)\) is optimal if and only if \( k > \tilde{p} + \frac{c}{\sigma_H - \sigma_L} \).

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 \( k > \tilde{p} + \frac{c}{\sigma_H - \sigma_L} \). Also, under full 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 large 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, while 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 Figures 8-(b) and 8-(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 ($\varepsilon_H(v) < \lambda \varepsilon_L(v)$).

### 5.2 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, they 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 $E^*(v|m_j, s_g) = \bar{v}$ and $E^*(v|m_j, s_b) = 0$ 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.

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\textsuperscript{23}:

$$D(m_j, p) = \begin{cases} 
\alpha \tilde{\sigma}_j + (1 - \alpha), & p \leq \underline{p}_j \\
\hat{\sigma}_j, & \underline{p}_j < p \leq \bar{p}_j \\
\alpha \hat{\sigma}_j, & \bar{p}_j < p \leq \bar{\tilde{v}}.
\end{cases}$$

(16)

\textsuperscript{23}Note that as $\alpha \to 1$, $D(m_j, p) \to \hat{\sigma}_j$ 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 ($\tilde{\sigma}_H > \tilde{\sigma}_L$) and the lack of consideration of different signal-generating processes by naive consumers.
full coverage case without (rational) consumer attention \((p \leq \bar{p}_j)\), the optimal price given the messenger \(m_j\) is \(\bar{p}_j\). A high-type messenger is optimal if and only if \([\alpha \bar{\sigma}_H + (1 - \alpha)](\bar{p}_L - k) \geq \left[\alpha \bar{\sigma}_L + (1 - \alpha)\right](\bar{p}_j - k) \iff k \leq \frac{\bar{p}_H - \bar{p}_L}{\alpha(\bar{\sigma}_H - \bar{\sigma}_L)} + \tilde{\bar{p}} \equiv \tilde{k}(c, \alpha)\). Second, in the partial coverage case with (rational) consumer attention \((\bar{p}_j < p \leq \bar{p})\), the optimal price given the messenger \(m_j\) is \(\bar{p}_j\). A high-type messenger is optimal if and only if \(\bar{\sigma}_H(\bar{p}_j - k) \geq \bar{\sigma}_L(\bar{p}_j - k) \iff k \leq \bar{p}\), which always holds when \(\varepsilon_H(v) \geq \lambda \varepsilon_L(v)\) for all \(v\). Finally, there is a new case where no (rational) consumers purchase – no coverage case \((\bar{p}_j < p \leq \bar{v})\), 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 \(\bar{v}\).

If \(\alpha\) becomes large, the case reverts to a trivial case. We assume that \(\alpha\) is small enough \((\alpha \leq \min \left\{\frac{1 - \bar{\sigma}_H}{1 - \bar{\sigma}_L}, \frac{\bar{\sigma}_L}{\bar{\sigma}_H}\right\})\) such that the general demand structure in the main model holds. 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 \(\alpha \leq \min \left\{\frac{1 - \bar{\sigma}_H}{1 - \bar{\sigma}_L}, \frac{\bar{\sigma}_L}{\bar{\sigma}_H}\right\}\). If \(\varepsilon_H(v) < \lambda \varepsilon_L(v)\) and \(\max\{k_{10}(c, \alpha), \tilde{\bar{p}}\} < k < k_9(c, \alpha)\), partial coverage with rational consumers’ attention and a low-type messenger \(m_L\) is optimal.

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 \(\varepsilon_H(v) < \lambda \varepsilon_L(v)\) for all \(v\) and \(\max\{k_{10}(c, \alpha), \tilde{\bar{p}}\} < k < k_9(c, \alpha)\) (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 \(p = \bar{v}\) upon receiving good signals, it becomes more profitable for the firm to forego the attention of rational consumers (i.e., no coverage from rational

---

\(^{24}\)Intuitively, no coverage case where the firm targets only naive consumers become 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: \(\frac{\partial \delta}{\partial \alpha} |_{k_0} = \frac{(\bar{\sigma}_H - (\bar{\sigma}_L - k))}{1 - \bar{\sigma}_H} > 0\), constant in the partial coverage case: \(\frac{\partial \delta}{\partial \sigma} |_{k_0} = 0\), and decreasing in the full coverage case: \(\frac{\partial \delta}{\partial \alpha} |_{k_0} = \frac{(\bar{\sigma}_L - k)}{\alpha (\bar{\sigma}_H - (\bar{\sigma}_L - k))} < 0\). Thus, as \(\alpha\) becomes larger, the profit under no coverage case dominates the other cases.

\(^{25}\)Without naive consumers, the consumer demand for full coverage without inducing consumer attention is always greater than partial coverage without 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: \(\alpha \bar{\sigma}_L + (1 - \alpha) \geq \bar{\sigma}_H \iff \alpha \leq \frac{1 - \bar{\sigma}_L}{\bar{\sigma}_H} \) and \(\bar{\sigma}_L \geq \alpha \bar{\sigma}_H \iff \alpha \leq \frac{\bar{\sigma}_H}{\bar{\sigma}_L}\).

\(^{26}\)We characterize the entire parameter space for the optimal choice of coverage and messenger in the Online Appendix.
5.3 Unobservable Pricing

In this section, we introduce a new setting where consumers cannot costlessly observe the price of the product. Instead, they can only observe the type of messenger. Based on this limited information and their expectation of the product price, denoted as $p^e$, 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 in order 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 search ex-ante. To address such a problem, we introduce consumer heterogeneity by assuming that a portion $\beta$ 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, while $(1 - \beta)$ 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: \( p \leq E(v|m_j, s_i) \). 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: \( E(v|m_g, s_i) - p^e - \xi \geq 0 \). On the other hand, 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: \( E(v) - p^e - \xi \geq 0 \). Overall, regular consumers would pay attention to the ad with messenger \( m_j \) if and only if \( EU(\text{attention}|m_j) = \tilde{\sigma}_j[E(v|m_j, s_g) - p^e - \xi] - c \geq EU(\text{no attention}) = \max\{E(v) - p^e - \xi, 0\} \). So, regular consumers would pay attention to the ad and engage in price search if and only if \( p_j < p^e + \xi \leq \bar{p}_j \), not pay attention to the ad and engage in price search if and only if \( p^e + \xi \leq p_j \), and not pay attention to the ad and not engage in price search if and only if \( p^e + \xi > \bar{p}_j \).

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: \( D(m_j, p^e, p) = \beta D_s(m_j, p) + (1 - \beta) D_i(m_j, p^e, p) \), where the demand from shoppers is \( D_s(m_j, p) = \tilde{\sigma}_j\mathbb{I}\{p \leq E(v|m_j, s_g)\} + (1 - \tilde{\sigma}_j)\mathbb{I}\{p \leq E(v|m_j, s_b)\} \) and the demand from regular consumers is

\[
D_i(m_j, p^e, p) = \begin{cases} 
\mathbb{I}\{p \leq E(v)\}, & p^e + \xi \leq p_j \\
\tilde{\sigma}_j\mathbb{I}\{p \leq E(v|m_j, s_g)\} + (1 - \tilde{\sigma}_j)\mathbb{I}\{p \leq E(v|m_j, s_b)\}, & p_j < p^e + \xi \leq \bar{p}_j \\
0, & p^e + \xi > \bar{p}_j
\end{cases}
\]  

(17)

Given the consumer demand, the profit-maximizing price is either \( E(v|m_j, s_b) \), \( E(v) \), or \( E(v|m_j, s_g) \). However, \( p = E(v) \) can never be consistent with the expected price \( p^e \) because \( p = E(v) \) can be optimal only when \( p^e + \xi \leq p_j < E(v) \). Thus, only \( E(v|m_j, s_b) \) and \( E(v|m_j, s_g) \) 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 \( E(v|m_j, s_b) + \xi \leq p_j \) 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 \( p = E(v|m_j, s_b) \) and the partial coverage case with \( p = E(v|m_j, s_g) \).\(^{27}\)

\(^{27}\)We assume that \( \beta \) is large enough so that either \( p = E(v|m_j, s_b) \) or \( p = E(v|m_j, s_g) \) is optimal. Specifically, we
firm’s pricing strategy for regular consumers is influenced by the two extremes: either giving up on cost. This is because no price can credibly attract attention from regular consumers. Therefore, the consumers. The regions where a messenger-price pair is optimal are independent of the deliberation However, there are some notable differences due to the introduction of deliberation costs for regular be an effective medium to attract consumer attention and potentially increase prices for the firm.

**Proposition 7.**

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

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

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

In the full coverage case, the firm’s profit given a messenger \( m_j \) is \( E(v|m_j, s_b) - k \), so the high-type messenger is optimal if and only if \( E(v|m_H, s_b) \geq E(v|m_L, s_b) \), which holds when \( \varepsilon_H(v) < \kappa \varepsilon_L(v) \) for all \( v \). In the partial coverage case, the firm’s profit given a messenger \( m_j \) is \( \beta \tilde{\sigma}_j [E(v|m_j, s_g) - k] \), so the high-type messenger is optimal if and only if \( \beta \tilde{\sigma}_H(E(v|m_H, s_g) - k) \geq \beta \tilde{\sigma}_L(E(v|m_L, s_g) - k) \iff k \leq \tilde{p} \), which is always the case when \( \varepsilon_H(v) \geq \lambda \varepsilon_L(v) \). Table 3 summarizes the results of the firm’s optimal messenger and price under each regime.

**Table 3: Optimal choice of messenger and price \((m^*_j, p^*)\) when pricing is unobservable**

<table>
<thead>
<tr>
<th>Coverage</th>
<th>( \varepsilon_H(v) \geq \lambda \varepsilon_L(v) )</th>
<th>( \kappa \varepsilon_L(v) \leq \varepsilon_H(v) &lt; \lambda \varepsilon_L(v) )</th>
<th>( \varepsilon_H(v) &lt; \kappa \varepsilon_L(v) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full</td>
<td>((m_L, E(v</td>
<td>m_L, s_b)))</td>
<td>((m_H, E(v</td>
</tr>
<tr>
<td>Partial</td>
<td>((m_H, E(v</td>
<td>m_H, s_g)))</td>
<td>(\begin{cases} (m_L, E(v</td>
</tr>
</tbody>
</table>

\[
\text{Table 3: Optimal choice of messenger and price \((m^*_j, p^*)\) when pricing is unobservable}
\]

\[\begin{array}{cccc}
\varepsilon_H(v) & \lambda \varepsilon_L(v) & \kappa \varepsilon_L(v) & \varepsilon_H(v) < \kappa \varepsilon_L(v) \\
\hline
\text{Full} & (m_L, E(v|m_L, s_b)) & (m_H, E(v|m_H, s_b)) & \\
\text{Partial} & (m_H, E(v|m_H, s_g)) & \begin{cases} (m_L, E(v|m_L, s_g)), k > \tilde{p} \\
(m_H, E(v|m_H, s_g)), \text{otherwise} \end{cases} \\
\end{array}\]

\[\text{Table 3: Optimal choice of messenger and price \((m^*_j, p^*)\) when pricing is unobservable}
\]

In the full coverage case, the firm’s profit given a messenger \( m_j \) is \( E(v|m_j, s_b) - k \), so the high-type messenger is optimal if and only if \( E(v|m_H, s_b) \geq E(v|m_L, s_b) \), which holds when \( \varepsilon_H(v) < \kappa \varepsilon_L(v) \) for all \( v \). In the partial coverage case, the firm’s profit given a messenger \( m_j \) is \( \beta \tilde{\sigma}_j [E(v|m_j, s_g) - k] \), so the high-type messenger is optimal if and only if \( \beta \tilde{\sigma}_H(E(v|m_H, s_g) - k) \geq \beta \tilde{\sigma}_L(E(v|m_L, s_g) - k) \iff k \leq \tilde{p} \), which is always the case when \( \varepsilon_H(v) \geq \lambda \varepsilon_L(v) \). Table 3 summarizes the results of the firm’s optimal messenger and price under each regime.

**Proposition 7.**

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

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

3. When \( \varepsilon_H(v) \geq \lambda \varepsilon_L(v) \), it is optimal to induce partial coverage if \( k \geq k_{13} \). Otherwise, it is optimal to induce full coverage with \((m_H, E(v|m_H, s_b))\). Moreover, under partial coverage, the firm chooses a low-type messenger with \((m_L, E(v|m_L, s_g))\) 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 deliberation 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 \( E(v|m_j, s_g) \) or setting a low price \( E(v|m_j, s_b) \) to make all consumers purchase without paying attention to the ad.

Assume that \( \beta \geq \tilde{\beta} \equiv \max \{ \frac{E - \xi - k}{E(v|m_j, s_g) - k}, \min \{ \frac{E - \xi - k}{E(v|m_j, s_b) - k} \} \} \). When \( \beta < \tilde{\beta} \), there can be regions of parameters where a pure-strategy equilibrium does not exist.
5.4 Other extensions

We also carry out other extensions to assess the robustness of our main insight. We overview those extensions and the details 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, while 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 $v_i$, consumers can obtain additional social utilities $\eta > 0$ if the product is endorsed by a high-type messenger. The results are qualitatively the same as the main results. 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 to 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.

### 6 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, while 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. However, 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 a few limitations. Firstly, the results are characterized under specific relationships between signal structures, which may make it challenging to extend the findings to more general cases. However, when considering a binary distribution of match values, we can cover all possible scenarios and show that the results still hold. Secondly, while the paper provides general principles for selecting advertising messengers, other factors can also influence a firm’s choice, such as unconventional messengers for capturing attention or appealing to consumer values.

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.
Appendix

Proof of Proposition 1. Note that \( \frac{f(v|m_H,s_p)}{f(v|m_L,s_p)} = \frac{f(v)\sigma_H(v)}{f(v)\sigma_L(v)} \cdot \frac{\frac{\sigma_L}{\sigma_H}}{\bar{\sigma}_j} \). Because \( \frac{\sigma_L}{\sigma_H} \) does not depend on \( v \), \( \frac{f(v|m_H,s_p)}{f(v|m_L,s_p)} \) is increasing in \( v \) if and only if \( \frac{df}{dv} \left( \frac{f(v)\sigma_H(v)}{f(v)\sigma_L(v)} \right) \geq 0 \Leftrightarrow \frac{df}{dv} \left( \frac{\sigma_H(v)}{\sigma_L(v)} \right) \geq 0 \Leftrightarrow \frac{d\sigma_H(v)/dv}{\sigma_H(v)/v} \geq \frac{d\sigma_L(v)/dv}{\sigma_L(v)/v} \) for all \( v \). Conversely, \( \frac{f(v|m_H,s_p)}{f(v|m_L,s_p)} \) is decreasing in \( v \) if and only if \( \frac{d\sigma_H(v)/dv}{\sigma_H(v)/v} < \frac{d\sigma_L(v)/dv}{\sigma_L(v)/v} \) for all \( v \).

Proof of Lemma 1. First, when \( p \geq E(v) \), a consumer pays attention to the ad if and only if \( EU(\text{attention}|m_j) = \bar{\sigma}_j \cdot [E(v|m_j,s_g) - p] - c \geq EU(\text{no attention}) = \text{max}\{0, E(v) - p\} = 0 \Leftrightarrow p \leq \bar{\sigma}_jE(v|m_j,s_g) - c \). Next, when \( p < E(v) \), consumers pay attention to the ad if and only if \( \bar{\sigma}_j \cdot [E(v|m_j,s_g) - p] > c > \text{max}\{0, E(v) - p\} = E(v) - p \Leftrightarrow p > \frac{E(v) - \bar{\sigma}_jE(v|m_j,s_g) + c}{1 - \bar{\sigma}_j} \). Because \( E(v) = \bar{\sigma}_jE(v|m_j,s_g) + (1 - \bar{\sigma}_j)E(v|m_j,s_b) \), this is equivalent to \( p > E(v|m_j,s_b) + \frac{c}{1 - \bar{\sigma}_j} \).

Proof of Lemma 2. Inequality (9) can be rewritten as \( \int_0^\infty (v - p)\left[\sigma_H(v) - \sigma_L(v)\right]dF \geq 0 \Leftrightarrow p \int_0^\infty \left[\sigma_H(v) - \sigma_L(v)\right]dF \leq \int_0^\infty v\left[\sigma_H(v) - \sigma_L(v)\right]dF \Leftrightarrow p \leq \frac{\int_0^\infty v[\sigma_H(v) - \sigma_L(v)]dF}{1 - \frac{\sigma_L}{\sigma_H}} \equiv \bar{p} \).

Proof of Proposition 2. Note that \( \bar{p}_L > \bar{p}_H \Leftrightarrow c < c^* = \frac{E(v|m_L,s_p) - E(v|m_H,s_p)}{\frac{\sigma_L}{\sigma_H}} \). First, when \( \varepsilon_H(v) \geq \varepsilon_L(v) \) for all \( v \), \( f(v|m_H,s_g) \) satisfies the MLRP with respect to \( f(v|m_L,s_g) \) from Proposition 1, which implies that \( E(v|m_L,s_g) - E(v|m_H,s_g) \leq 0 \). Thus, \( c^* \leq 0 \), and \( \bar{p}_H \geq \bar{p}_L \). Second, when \( \lambda \varepsilon_L(v) \leq \varepsilon_H(v) < \varepsilon_L(v) \) for all \( v \), \( f(v|m_L,s_g) \) satisfies the MLRP with respect to \( f(v|m_H,s_g) \) from Proposition 1, which implies that \( E(v|m_L,s_g) - E(v|m_H,s_g) > 0 \) and \( c^* > 0 \). Moreover, we have \( c^* \leq \bar{c} = \varepsilon_L(v) \int_0^\infty v[\sigma_H(v) - \sigma_j]dF \) (that is, \( c^* \) is within our range of consideration of \( c \)) if and only if \( \frac{\int_0^\infty v\sigma_L(v)dF - \int_0^\infty v\sigma_H(v)dF}{1 - \frac{\sigma_L}{\sigma_H}} \leq \min_{v \in [L,H]} \int_0^\infty v[\sigma_H(v) - \sigma_{\bar{\sigma}_j}]dF \Leftrightarrow \int_0^\infty v[\sigma_H(v) - \sigma_L(v)]dF \geq \int_0^\infty (\bar{\sigma}_j - \sigma_{\bar{\sigma}_j})dF \Leftrightarrow \bar{p} = \frac{\int_0^\infty v[\sigma_H(v) - \sigma_L(v)]dF}{1 - \frac{\sigma_L}{\sigma_H}} \geq \int_0^\infty v(\bar{\sigma}_j - \sigma_L)dF = E(v) \) which holds when \( \varepsilon_H(v) \geq \lambda \varepsilon_L(v) \) from Lemma 3 below. So, \( c^* \in (0, \bar{c}] \), and \( \bar{p}_L > \bar{p}_H \) if and only if \( c < c^* \). Finally, when \( \varepsilon_H(v) < \lambda \varepsilon_L(v) \), we have \( c^* > \bar{c} \) if and only if \( \int_0^\infty v[\sigma_H(v) - \sigma_L(v)]dF \leq \int_0^\infty v(\bar{\sigma}_j - \sigma_L)dF \Leftrightarrow \bar{p} < E(v) \) which holds from Lemma 3 below. Therefore, \( \bar{p}_L > \bar{p}_H \).

Proof of Proposition 3. Note that \( \bar{p}_L > \bar{p}_H \Leftrightarrow c < c^{**} = \frac{E(v|m_L,s_p) - E(v|m_H,s_p)}{\frac{1 - \sigma_H}{1 - \sigma_L}} \). First, when \( \varepsilon_H(v) \geq \lambda \varepsilon_L(v) \), we have \( c^{**} \geq \bar{c} \) (that is, \( c^{**} \) is outside the range of consideration) if and only if \( \int_0^\infty (1 - \sigma_L)v[1 - \sigma_{\bar{\sigma}_j}]dF - (1 - \bar{\sigma}_j) \int_0^\infty v[1 - \sigma_H(v)]dF \geq \min_{v \in [L,H]} \int_0^\infty v[\sigma_H(v) - \sigma_{\bar{\sigma}_j}]dF \leq \int_0^\infty v(1 - \sigma_L)dF \Leftrightarrow \int_0^\infty v[\sigma_H(v) - \sigma_L(v)]dF \geq \int_0^\infty v(1 - \sigma_L)dF \Leftrightarrow \bar{p} \geq E(v) \), which holds according to Lemma 3. Thus, \( \bar{p}_L > \bar{p}_H \).
Second, when $\kappa \varepsilon_L(v) \leq \varepsilon_H(v) < \lambda \varepsilon_L(v)$ for all $v$, we have $\varepsilon_H(v) \geq \kappa \varepsilon_L(v) \iff \frac{d\sigma_H(v)}{dv} [1 - \sigma_L(v)] \geq \frac{d\sigma_L(v)}{dv} [1 - \sigma_H(v)] \iff \frac{d\sigma_H(v)}{dv} / \sigma_H(v) \geq \frac{d\sigma_L(v)}{dv} / \sigma_L(v) \iff \frac{d\sigma_H(v) - \sigma_L(v)}{\sigma_H - \sigma_L} \geq 0$. Also, note that $\tilde{\sigma} = \int_0^v \frac{\sigma_H(v) - \sigma_L(v)}{\sigma_H - \sigma_L} dF \geq E(v) \Leftrightarrow \int_0^\delta \frac{\sigma_H(v)}{v} - \sigma_L(v) dF \geq \int_0^\delta E(v)[\sigma_H(v) - \sigma_L(v)] dF \Leftrightarrow \int_0^\delta [v - E(v)](\sigma_H(v) - \sigma_L(v)) dF \geq 0$. Thus, when $\varepsilon_H(v) \geq \lambda \varepsilon_L(v)$ for all $v$, we have $\int_0^\delta [v - E(v)](\tilde{\sigma} H(v) - \sigma_L(v)) dF \geq \int_0^\delta [v - E(v)](\sigma_H(E(v)) - \sigma_L(E(v))) dF = [\sigma_H(E(v)) - \sigma_L(E(v))] \int_0^\delta [v - E(v)] dF = 0$, where the first inequality holds because the positive part of the integral decreases while the negative part of the integral increases. Thus, $\tilde{\sigma} \geq E(v)$. Conversely, when $\varepsilon_H(v) < \lambda \varepsilon_L(v)$ for all $v$, $\frac{d\sigma_H(v) - \sigma_L(v)}{dv} < 0$, which implies that $\int_0^\delta [v - E(v)](\tilde{\sigma} H(v) - \sigma_L(v)) dF < 0$. Thus, $\tilde{\sigma} < E(v)$.

**Proof of Lemma 3.** When $\varepsilon_H(v) \geq \lambda \varepsilon_L(v)$ for all $v$, we have that $\varepsilon_H(v) \sigma_H(v) \geq \varepsilon_L(v) \sigma_L(v) \iff \frac{d\sigma_H(v)}{dv} / \sigma_H(v) \geq \frac{d\sigma_L(v)}{dv} / \sigma_L(v) \iff \frac{d(\sigma_H(v) - \sigma_L(v))}{dv} \geq 0$. Also, note that $\tilde{\sigma} = \int_0^v \frac{\sigma_H(v) - \sigma_L(v)}{\sigma_H - \sigma_L} dF \geq E(v) \Leftrightarrow \int_0^\delta \frac{\sigma_H(v)}{v} - \sigma_L(v) dF \geq \int_0^\delta E(v)[\sigma_H(v) - \sigma_L(v)] dF \Leftrightarrow \int_0^\delta [v - E(v)](\sigma_H(v) - \sigma_L(v)) dF \geq 0$. Thus, when $\varepsilon_H(v) \geq \lambda \varepsilon_L(v)$ for all $v$, we have $\int_0^\delta [v - E(v)](\sigma_H(v) - \sigma_L(v)) dF \geq \int_0^\delta [v - E(v)](\sigma_H(E(v)) - \sigma_L(E(v))) dF = [\sigma_H(E(v)) - \sigma_L(E(v))] \int_0^\delta [v - E(v)] dF = 0$, where the first inequality holds because the positive part of the integral decreases while the negative part of the integral increases. Thus, $\tilde{\sigma} \geq E(v)$. Conversely, when $\varepsilon_H(v) < \lambda \varepsilon_L(v)$ for all $v$, $\frac{d(\sigma_H(v) - \sigma_L(v))}{dv} < 0$, which implies that $\int_0^\delta [v - E(v)](\sigma_H(v) - \sigma_L(v)) dF < 0$. Thus, $\tilde{\sigma} < E(v)$.

**Proof of Proposition 4.** We first discuss the optimal coverage choice. Given the optimal coverage choice, the optimal choice of the messenger is given by Table 2. First, when $\varepsilon_H(v) \geq \lambda \varepsilon_L(v)$ for all $v$, we know that only $(m_H, p_H)$ can be optimal under partial coverage and only $(m_L, p_L)$ can be optimal under full coverage. Thus, it is optimal to induce partial coverage with $(m_H, p_H)$ if the expected profit that it generates ($\tilde{\sigma}_H[p_H - k]$), is larger than the highest expected profit under full coverage ($p_L - k$): $\tilde{\sigma}_H[p_H - k] = \tilde{\sigma}_H[E(v|m_H, s_g) - k] - c \geq p_L - k = E(v|m_L, s_b) + \frac{\tilde{\sigma}_L}{1 - \tilde{\sigma}_L} - k \iff k \geq \frac{1}{1 - \tilde{\sigma}_H} [2 - \frac{\tilde{\sigma}_H}{1 - \tilde{\sigma}_L} c + E(v|m_L, s_b) - \tilde{\sigma}_H E(v|m_H, s_g)] \equiv k_1(c)$. Otherwise (i.e., $k < k_1(c)$), it is optimal to induce full coverage with $(m_L, p_L)$. Second, when $\kappa \varepsilon_L(v) \leq \varepsilon_H(v) < \lambda \varepsilon_L(v)$ for all $v$, either $m_L$ or $m_H$ can be optimal under any given coverage. Therefore, it is optimal to induce partial coverage if the highest expected profit under partial coverage, $\max_{j \in \{L, H\}} \tilde{\sigma}_j[p_j - k]$, is larger than the highest expected profit under full coverage, $\max_{j' \in \{L, H\}} p_{j'} - k$: $k \geq \min_{j \in \{L, H\}} \left\{ \max_{j' \in \{L, H\}} \frac{1}{1 - \tilde{\sigma}_j} [2 - \frac{\tilde{\sigma}_j}{1 - \tilde{\sigma}_{j'}} c + E(v|m_j, s_b) - \tilde{\sigma}_j E(v|m_j, s_g)] \right\} \equiv k_2(c)$. Otherwise (i.e., $k < k_2(c)$), it is optimal to induce full coverage. Third, when $\varepsilon_H(v) < \kappa \varepsilon_L(v)$ for all $v$, we know from Table 2 that only $(m_H, p_H)$ can be optimal under full coverage. Also, under partial coverage, either $m_L$ or $m_H$ can be optimal. Therefore, it is optimal to induce partial coverage if the highest ex-
Proof of Proposition 5. First, when \( \varepsilon_H(v) \geq \lambda \varepsilon_L(v) \) for all \( v \), we know from the discussion that only \((m_H, \overline{v}_H)\) can be optimal under partial coverage and only \((m_L, \overline{p}_L)\) can be optimal under full coverage. Thus, it is optimal to induce partial coverage with \((m_H, \overline{v}_H)\) if the expected profit that it generates \((\bar{\sigma}_H[\overline{v}_H - k])\), is larger than the highest expected profit under full coverage \((\bar{p}_L - k)\): 

\[
\bar{\sigma}_H[\overline{v}_H - k] = \bar{\sigma}_H[E(v|m_H, s_b) - k] \geq \bar{p}_L - k = E(v|m_L, s_b) + \frac{c}{1-\sigma_L} - k \Leftrightarrow k \geq \frac{1}{1-\sigma_L} \left[ \frac{2-\bar{\sigma}_L}{1-\sigma_L} c + E(v|m_L, s_b) - \bar{\sigma}_LE(v|m_L, s_g) \right] \equiv k_4(c). 
\]

Otherwise (i.e., \( k < k_4(c) \)), it is optimal to induce full coverage with \((m_H, \overline{p}_H)\). Second, when \( \kappa \varepsilon_L(v) \leq \varepsilon_H(v) < \lambda \varepsilon_L(v) \) for all \( v \), we again know from the discussion that only \((m_L, \overline{p}_L)\) can be optimal under full coverage. And, under partial coverage, either \( m_L \) or \( m_H \) can be optimal. Thus, it is optimal to induce partial coverage if the highest expected profit under partial coverage, \( \max\{\bar{\sigma}_L[\overline{p}_L - k], \bar{\sigma}_H[\overline{v}_H - k] \} \), is larger than the highest expected profit under full coverage, \( \bar{p}_L - k \), which holds if and only if \( k \geq \min \left\{ \frac{1}{1-\sigma_L} \left[ \frac{2-\bar{\sigma}_L}{1-\sigma_L} c + E(v|m_L, s_b) - \bar{\sigma}_LE(v|m_L, s_g) \right], \frac{1}{1-\sigma_L} \left[ c + E(v|m_H, s_b) - \bar{\sigma}_LE(v|m_H, s_g) \right] \right\} \equiv k_5(c). \) Otherwise (i.e., \( k < k_5(c) \)), it is optimal to induce full coverage with \((m_L, \overline{p}_L)\). Finally, when \( \varepsilon_H(v) < \kappa \varepsilon_L(v) \) for all \( v \), we know from the discussion that either \( m_L \) or \( m_H \) can be optimal under any given coverage. Thus, it is optimal to induce partial coverage if the highest expected profit under partial coverage, \( \max\{\bar{\sigma}_L[\overline{p}_L - k], \bar{\sigma}_H[\overline{v}_H - k] \} \), is larger than the highest expected profit under full coverage, \( \bar{p}_L - k \), which holds if and only if \( k \geq \min \left\{ \max \left\{ \frac{1}{1-\sigma_L} \left[ \frac{2-\bar{\sigma}_L}{1-\sigma_L} c + E(v|m_L, s_b) - \bar{\sigma}_LE(v|m_L, s_g) \right], \frac{1}{1-\sigma_L} \left[ c + E(v|m_H, s_b) - \bar{\sigma}_LE(v|m_H, s_g) \right] \right\} \right\} \equiv k_6(c). \) Otherwise (i.e., \( k < k_6(c) \)), it is optimal to induce full coverage.

Proof of Proposition 6. We know from the discussion before the proposition that the partial coverage with a low type messenger \((m_L, \overline{p}_L)\) can never be optimal when \( \varepsilon_H(v) \geq \lambda \varepsilon_L(v) \) for all \( v \). Thus, for the low-type messenger to be an optimal choice under partial coverage, it must be the case that \( \varepsilon_H(v) < \lambda \varepsilon_L(v) \) for all \( v \). We also know from the discussion that under no coverage, only \((m_H, \overline{v})\) can be optimal. And, under partial and full coverage, either \( m_L \) or \( m_H \) can be optimal. Therefore, when \( \varepsilon_H(v) < \lambda \varepsilon_L(v) \) for all \( v \), partial coverage is optimal if the highest expected profit under partial
coverage \((\max_{j \in \{L, H\}} \tilde{\sigma}_j(p_j - k))\) is larger than both the highest expected profit under no coverage \((\alpha \tilde{\sigma}_H(\bar{p} - k))\) and the highest expected profit under full coverage \((\max_{j \in \{L, H\}} [\alpha \tilde{\sigma}_j + (1 - \alpha)](p_j - k))\). This holds if and only if \(\max_{j \in \{L, H\}} \tilde{\sigma}_j(p_j - k) > \max \{\alpha \tilde{\sigma}_H(\bar{p} - k), \max_{j' \in \{L, H\}} [\alpha \tilde{\sigma}_{j'}' + (1 - \alpha)](p_{j'} - k)\} \Leftrightarrow \min_{j \in \{L, H\}} \left\{ \max_{j' \in \{L, H\}} \frac{\tilde{\sigma}_{j'}(v) - \tilde{\sigma}_j(v)}{\alpha(\tilde{\sigma}_{j'} - \tilde{\sigma}_j) + (1 - \alpha)} \right\} \equiv k_{10}(c, \alpha) < k < \max_{j \in \{L, H\}} \frac{\tilde{\sigma}_j(E(v|m_j, s_b) - \tilde{\sigma}_j H \bar{p} - c)}{\tilde{\sigma}_j - \alpha \tilde{\sigma}_H} \equiv k_9(c, \alpha).\) In particular, when partial coverage is optimal, we know that \((m_L, \bar{p}_L)\) is optimal if \(k > \bar{p}\). All together, partial coverage with \((m_L, \bar{p}_L)\) is optimal if \(\varepsilon_H(v) < \lambda \varepsilon_L(v)\) for all \(v\) and \(\max\{k_{10}(c, \alpha), \bar{p}\} < k < k_9(c, \alpha).\)

**Proof of Proposition 7.** First, when \(\varepsilon_H(v) \geq \lambda \varepsilon_L(v)\) for all \(v\), we know from Table 3 that only \((m_L, E(v|m_L, s_b))\) can be optimal under full coverage and only \((m_H, E(v|m_H, s_g))\) can be optimal under partial coverage. Therefore, partial coverage with \((m_L, E(v|m_L, s_b))\) is optimal if the expected profit that it generates, \((\beta \tilde{\sigma}_L(E(v|m_H, s_g) - k))\) is larger than the highest expected profit under full coverage \((E(v|m_L, s_b) - k)): \(k \geq \frac{E(v|m_L, s_b) - \beta \tilde{\sigma}_L E(v|m_H, s_g)}{1 - \beta \tilde{\sigma}_L} \equiv k_{11}.\) Otherwise (i.e., \(k < k_{11}\)), full coverage with \((m_L, E(v|m_L, s_b))\) is optimal. Next, when \(\kappa \varepsilon_L(v) \leq \varepsilon_H(v) < \lambda \varepsilon_L(v)\) for all \(v\), we know that only \((m_L, E(v|m_L, s_b))\) can be optimal under full coverage. And, under partial coverage, either \(m_L\) or \(m_H\) can be optimal. Therefore, partial coverage is optimal if the highest expected profit under partial coverage, \((\max_{j \in \{L, H\}} \beta \tilde{\sigma}_j(E(v|m_j, s_g) - k))\) is larger than the highest expected profit under full coverage \((E(v|m_L, s_b) - k)): \(k \geq \min_{j \in \{L, H\}} \frac{E(v|m_L, s_b) - \beta \tilde{\sigma}_j E(v|m_j, s_g)}{1 - \beta \tilde{\sigma}_j} \equiv k_{12}.\) Otherwise (i.e., \(k < k_{12}\)), full coverage with \((m_L, E(v|m_L, s_b))\) is optimal. Finally, when \(\varepsilon_H(v) < \kappa \varepsilon_L(v)\) for all \(v\), we know that only \((m_H, E(v|m_H, s_b))\) can be optimal under full coverage. And, under partial coverage, either \(m_L\) or \(m_H\) can be optimal. Therefore, partial coverage is optimal if the highest expected profit under partial coverage, \((\max_{j \in \{L, H\}} \beta \tilde{\sigma}_j(E(v|m_j, s_g) - k))\) is larger than the highest expected profit under full coverage, \((E(v|m_H, s_b) - k)): \(k \geq \min_{j \in \{L, H\}} \frac{E(v|m_H, s_b) - \beta \tilde{\sigma}_j E(v|m_j, s_g)}{1 - \beta \tilde{\sigma}_j} \equiv k_{13}.\) Otherwise (i.e., \(k < k_{13}\)), full coverage with \((m_H, E(v|m_H, s_b))\) is optimal.

**Funding and Competing Interests**

All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript. The authors have no funding to report.
References


Online Appendix for “The Role of Messenger in Advertising Content: Bayesian Persuasion Perspective”

A Optimal Signal Structure/Messenger Design

Usually, firms have to choose from the available messengers in the market, so they have limited ability to design the signal structure that consumers will face when they pay attention to the ad. In this section, we investigate the optimal messenger design when the firm has total freedom over designing the signal structure, as well as the commitment power to such design.

To simplify the analysis, we assume consumers’ match value distribution follows a Bernoulli distribution: \( v_i \in \{0, \bar{v}\} \). Formally, the firm chooses \( \sigma(v) \equiv \Pr(s_i = s_g|v) \), which determines \( E(v|s_g) \) and \( E(v|s_b) \). Because \( E(v|s_g) \) and \( E(v|s_b) \) are sufficient to pin down the consumers’ attention and purchase decisions, we can equivalently solve the problem where the firm finds the optimal posterior pair \( (E(v|s_g), E(v|s_b)) \) such that \( E(v|s_b) \leq E(v) \leq E(v|s_g) \). The following proposition characterizes the optimal messenger design.

**Proposition 1.** It is optimal to have full information (i.e., \( E(v|s_g) = \bar{v} \) and \( E(v|s_b) = 0 \)) and induce partial coverage if and only if the production cost is high relative to the deliberation cost such that \( k \geq \frac{\bar{v}}{\bar{v} - E(v)} c \). Otherwise, it is optimal to have no information (i.e., \( E(v|s_g) = E(v|s_b) = E(v) \)) and induce full coverage.

**Proof.** First, under partial coverage, the firm’s highest profit given the posterior pair \( (E(v|s_g), E(v|s_b)) \) is \( \tilde{\sigma}(E(v|s_g) - k) - c = \frac{E(v) - E(v|s_b)}{E(v|s_g) - E(v|s_b)}(E(v|s_g) - k) - c \).

Because \( \frac{\partial E(v|s_g) - E(v|s_b)}{\partial E(v|s_b)}(E(v|s_g) - k) - c = -\frac{E(v|s_g) - E(v)}{E(v|s_g) - E(v|s_b))^2}(E(v|s_g) - k) \leq 0 \), the optimal signal structure under partial coverage must entail \( E(v|s_b) = 0 \), and the firm’s highest profit given the posterior pair \( (E(v|s_g), 0) \) under partial coverage can be rewritten as \( E(v) - \frac{E(v)}{E(v|s_g)} k - c \).

Again, because \( \frac{\partial E(v) - E(v|s_g) k - c}{\partial E(v|s_g)} = \frac{E(v)}{E^2(v|s_g)} k > 0 \), the optimal design under partial coverage entails \( E(v|s_g) = \bar{v} \) and \( \tilde{\sigma} = \frac{E(v)}{\bar{v}} \). Therefore, under partial coverage, full information (i.e., \( E(v|s_g) = \bar{v} \) and \( E(v|s_b) = 0 \)) is optimal.

Next, because the highest price under no consumer attention, \( E(v|s_g) + \frac{c}{1 - \tilde{\sigma}} \), must be lower than \( E(v) \), the firm’s highest profit given the posterior pair \( (E(v|s_g), E(v|s_b)) \) is

\[
\min \left\{ E(v|s_b) + \frac{c}{1 - \tilde{\sigma}} - k, E(v) - k \right\}
\]

under full coverage. Obviously, the posterior pair \( (E(v|s_g), E(v|s_b)) = (E(v), E(v)) \) dominates other posterior pairs under full coverage because the highest profit under full coverage, \( E(v) - k \), can be achieved by \( (E(v), E(v)) \). Thus, under full coverage, no information (i.e., \( E(v|s_g) = E(v|s_b) = E(v) \)) is optimal.

Finally, partial coverage with full information is optimal if the highest expected profit under partial coverage, \( E(v) - \frac{E(v)}{\bar{v}} k - c \), is larger than the expected profit under full coverage, \( E(v) - k \), which holds if and only if \( k \geq \frac{\bar{v}}{\bar{v} - E(v)} c \). Otherwise (i.e., \( k < \frac{\bar{v}}{\bar{v} - E(v)} c \)), full coverage with no information is optimal.

---

\(^1\)Because the posteriors must satisfy the relationship \( \tilde{\sigma}E(v|s_g) + (1 - \tilde{\sigma})E(v|s_b) = E(v) \), it must be that \( \tilde{\sigma} \equiv \Pr(s_i = s_g) \).

1
messenger has the ability to command a higher price when there is full coverage, while a more informative messenger is more effective in capturing consumer attention and can command a higher price under partial coverage. In the case of complete signal structure design freedom, the optimal choice is to use the least informative messenger (no information) under full coverage and the most informative messenger (full information) under partial coverage. This highlights the importance of tailoring the signal structure to achieve the desired outcomes in different coverage scenarios.

B Optimal Deliberation Cost

In our main analysis, we treated the deliberation cost as an exogenously given parameter. However, in this extension, we examine the case where the firm can endogenously determine the deliberation cost. This occurs when the firm has control over factors such as the ad format, content readability, or other elements that can influence the deliberation cost. We aim to identify the optimal deliberation cost and coverage choice in this scenario. The following proposition provides insights into the optimal deliberation cost and coverage choice when the firm has the ability to determine the deliberation cost.

Proposition 2. It is optimal to set the deliberation cost to 0 \((c = 0)\) and induce partial coverage if \(k \geq \min_{j \in \{L,H\}} E(v|m_j, s_b)\). Otherwise, if \(k < \min_{j \in \{L,H\}} E(v|m_j, s_b)\), then it is optimal to set a sufficiently large deliberation cost such that \(c \geq \min_{j \in \{L,H\}} (1 - \tilde{\sigma}_j)(E(v) - E(v|m_j, s_b))\) and induce full coverage.

Proof. First, under partial coverage, the firm’s highest profit given the messenger \(m_j\) and the deliberation cost \(c\) is \(\tilde{\sigma}_j(E(v|m_j, s_g) - k) - c\). So, given the messenger \(m_j\), the firm can maximize its profit under partial coverage by setting \(c = 0\), which yields the profit of \(\tilde{\sigma}_j(E(v|m_j, s_g) - k)\).

Next, under full coverage, because \(\frac{E_j}{E_j^*} \leq E(v)\), the firm’s highest profit given the messenger \(m_j\) and the deliberation cost \(c\) is \(\min \left\{ E(v|m_j, s_b) + \frac{c}{1 - \tilde{\sigma}_j} - k, E(v) - k \right\}\). Thus, the highest possible profit under full coverage, which equals to \(E(v) - k\), can be achieved by letting \(E(v|m_j, s_b) + \frac{c}{1 - \tilde{\sigma}_j} \geq E(v) \iff c \geq (1 - \tilde{\sigma}_j)(E(v) - E(v|m_j, s_b))\) for any type of messenger \(m_j\).

Finally, it is optimal to induce partial coverage with zero deliberation cost if the highest expected profit under partial coverage, \(\max_{j \in \{L,H\}} \tilde{\sigma}_j(E(v|m_j, s_g) - k)\), is larger than the highest expected profit under full coverage, \(E(v) - k\), which holds if and only if \(k \geq \min_{j \in \{L,H\}} E(v|m_j, s_b)\). Otherwise (i.e., \(k < \min_{j \in \{L,H\}} E(v|m_j, s_b)\)), it is optimal to induce full coverage.

In the context of coverage choice, the main intuition remains the same as in the main model. When considering full coverage, setting a higher deliberation cost enables the firm to command a higher price. This is because a higher deliberation cost leads to increased price premiums for the messengers when consumers do not pay attention. Specifically, if the deliberation cost is sufficiently large, the firm can charge a price that matches consumers’ prior expected value of the product. On the other hand, under partial coverage, setting the deliberation cost to zero allows the firm to provide the most information and charge a higher price. This is because a zero deliberation cost negatively impacts the messenger’s price premiums when consumers do pay attention.

Overall, the relationship between the deliberation cost and coverage choice remains consistent with the main model, with the optimal strategies depending on the desired price premiums and consumer attention.
C Social Utility and Fandom

In this section, we assume that other than the consumption utility $v_i$, consumers can obtain additional social utilities $\eta > 0$ if the product is endorsed by a high-type messenger. Therefore, when facing a high-type messenger ad, a consumer’s utility of paying attention to the ad is $\tilde{\sigma}_H(E(v|m_H, s_g) - p + \eta) - c$, and the utility of not paying attention is $\max\{0, E(v) - p + \eta\}$. By comparing the two utilities, consumers pay attention to the ad if and only if $p_{H} + \eta < p \leq \tilde{p}_{H} + \eta$. Therefore, given the messenger is high-type, the optimal price under full coverage is $p_{H} + \eta$, and the optimal price under partial coverage is $\tilde{p}_{H} + \eta$.

Under full coverage, a high-type messenger ad is optimal if and only if $p_{H} + \eta - k \geq \tilde{p}_{H} + \eta \iff c \geq c^{**} - \frac{\eta}{\sigma_{H} - \sigma_{L}}$, which always holds when $\varepsilon_H(v) < \kappa \varepsilon_L(v)$ for all $v$. Under partial coverage, a high-type messenger ad is optimal if and only if $\tilde{\sigma}_H(p_{H} + \eta - k) \geq \tilde{\sigma}_L(p_{L} - k) \iff k < \tilde{p} + \frac{\tilde{\sigma}_H}{\sigma_{H} - \sigma_{L}} \eta$, which always holds when $\varepsilon_H(v) \geq \lambda \varepsilon_L(v)$ for all $v$. Table 1 summarizes the results of the firm’s optimal messenger and price under each regime.

<table>
<thead>
<tr>
<th>Coverage</th>
<th>$\varepsilon_H(v) \geq \kappa \varepsilon_L(v)$</th>
<th>$\kappa \varepsilon_L(v) \leq \varepsilon_H(v) &lt; \kappa \varepsilon_L(v)$</th>
<th>$\varepsilon_H(v) &lt; \kappa \varepsilon_L(v)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full</td>
<td>$(m_L, p_L)$, $c &lt; c^{**} - \frac{\eta}{\sigma_{H} - \sigma_{L}}$, otherwise</td>
<td>$(m_H, p_H + \eta)$</td>
<td>$(m_H, p_H + \eta)$</td>
</tr>
<tr>
<td>Partial</td>
<td>$(m_H, \tilde{p}_H + \eta)$</td>
<td>$(m_L, p_L)$, otherwise</td>
<td>$(m_H, \tilde{p}_H + \eta)$, otherwise</td>
</tr>
</tbody>
</table>

Table 1: Optimal choice of messenger and price $(m^*, p^*)$ under full and partial coverage when the product endorsed by a high-type messenger provides additional utilities

Proposition 3. 1. When $\varepsilon_H(v) \geq \kappa \varepsilon_L(v)$, it is optimal to induce partial coverage with $(m_H, \tilde{p}_H + \eta)$ if $k \geq k_{14}(c)$. Otherwise, full coverage is optimal. Moreover, under full coverage, the firm chooses a low-type messenger with $(m_L, p_L)$ if and only if $c < c^{**} - \frac{\eta}{\sigma_{H} - \sigma_{L}}$.

2. When $\kappa \varepsilon_L(v) \leq \varepsilon_H(v) < \kappa \varepsilon_L(v)$, it is optimal to induce partial coverage if $k > k_{15}(c)$. Otherwise, full coverage is optimal. Moreover, under partial coverage, the firm chooses a low-type messenger with $(m_L, p_L)$ if and only if $k > \tilde{p} + \tilde{\sigma}_H \eta$. And, under full coverage, the firm chooses a low-type messenger with $(m_L, p_L)$ if and only if $c < c^{**} - \frac{\eta}{\sigma_{H} - \sigma_{L}}$.

3. When $\varepsilon_H(v) < \kappa \varepsilon_L(v)$, it is optimal to induce partial coverage with $(m_H, \tilde{p}_H + \eta)$. Moreover, under partial coverage, the firm chooses a low-type messenger with $(m_L, p_L)$ if and only if $k > \tilde{p} + \frac{\tilde{\sigma}_H}{\sigma_{H} - \sigma_{L}} \eta$.

Proof. We first discuss the optimal coverage choice under each case. Then, given the optimal coverage choice, the optimal choice of the messenger is given in Table 1.

First, when $\varepsilon_H(v) \geq \kappa \varepsilon_L(v)$ for all $v$, we know from Table 1 that only $(m_H, \tilde{p}_H + \eta)$ can be optimal under partial coverage. And, under full coverage, either $m_L$ and $m_H$ can be optimal. Therefore, partial coverage with $(m_H, \tilde{p}_H + \eta)$ is optimal if the expected profit it generates, $\tilde{\sigma}_H(p_{H} + \eta - k)$, is larger than the highest expected profit under full coverage, $\max\{p_{H} + \eta - k, p_{L} - k\}$, which holds if and only if $k \geq \max\left\{\frac{2 - \tilde{\sigma}_H}{1 - \tilde{\sigma}_H} \left(\frac{c + E(v|m_H, s_g) - \tilde{\sigma}_H E(v|m_H, s_g) + (1 - \tilde{\sigma}_H) \eta}{1 - \tilde{\sigma}_H}\right), \frac{c + E(v|m_L, s_g) - \tilde{\sigma}_L E(v|m_L, s_g) - \tilde{\sigma}_H \eta}{1 - \tilde{\sigma}_H}\right\}$.
where \( \phi \) cost of featuring a messenger can also affect firms' advertising decisions. We assume that different

So far, we have only discussed the profit generated by different messenger ads. In practice, the

and when they do not. Consequently, the regions where a high-type messenger is deemed optimal

Finally, when \( \varepsilon_H(v) < \kappa \varepsilon_L(v) \) for all \( v \), we know from Table 1 that only \((m_H,p_H + \eta)\)

This analysis underscores the significance of considering social utility in messenger selection,

as a greater emphasis on fandom can lead to more favorable outcomes for high-type messengers and
an expanded range of scenarios where they are the optimal choice.

\[ k_{14}(c). \] Otherwise (i.e., \( k < k_{14}(c) \)), full coverage is optimal.

Next, when \( \kappa \varepsilon_L(v) \leq \varepsilon_H(v) < \lambda \varepsilon_L(v) \) for all \( v \), we again know from Table 1 that either \( m_L \)
or \( m_H \) can be optimal under any given coverage. Thus, partial coverage is optimal if the highest
expected profit under partial coverage, \( \max\{\tilde{\sigma}_H(p_H^L + \eta - k), \tilde{\sigma}_L(p_L - k)\} \), is larger than the highest
expected profit under full coverage, \( \max\{p_H^L + \eta - k, p_L - k\} \), which holds if and only if \( k \geq \)

\[ \min \left\{ \frac{2 - \frac{\eta}{1 - \sigma_H} c + E(v|m_H,s_h) - \tilde{\sigma}_H E(v|m_H,s_h) + (1 - \tilde{\sigma}_H) \eta}{1 - \sigma_H}, \frac{2 - \frac{\eta}{1 - \sigma_L} c + E(v|m_L,s_l) - \tilde{\sigma}_L E(v|m_L,s_l) - \tilde{\sigma}_H E(v|m_H,s_h) - \tilde{\sigma}_L E(v|m_L,s_l) + \eta}{1 - \sigma_L} \right\} \equiv k_{15}(c). \] Otherwise (i.e.,
\( k < k_{15}(c) \)), full coverage is optimal.

Finally, when \( \varepsilon_H(v) < \kappa \varepsilon_L(v) \) for all \( v \), we know from Table 1 that only \((m_H,p_H + \eta)\)
can be optimal under full coverage. And, under partial coverage, either \( m_L \) or \( m_H \) can be optimal.

Therefore, partial coverage is optimal if the highest expected profit under partial coverage,
\( \max\{\tilde{\sigma}_H(p_H^L + \eta - k), \tilde{\sigma}_L(p_L - k)\} \), is larger than the highest expected profit under full coverage,
\( p_H^L + \eta - k \), which holds if and only if \( k \geq \)

\[ \min \left\{ \frac{2 - \frac{\eta}{1 - \sigma_H} c + E(v|m_H,s_h) - \tilde{\sigma}_H E(v|m_H,s_h) + (1 - \tilde{\sigma}_H) \eta}{1 - \sigma_H} + \frac{2 - \frac{\eta}{1 - \sigma_L} c + E(v|m_L,s_l) - \tilde{\sigma}_L E(v|m_L,s_l) + \eta}{1 - \sigma_L} \right\} \equiv k_{16}(c). \] Otherwise,
full coverage with \((m_H,p_H + \eta)\) is optimal.

Figure 1 depicts Proposition 3, which explores the impact of social utility (fandom) on the
optimal messenger choice. Notably, when the social utility parameter \( \eta \) is set to zero, the results
align with the main findings. However, as \( \eta \) increases, the advantage of using a high-type messenger
becomes more pronounced. Specifically, as \( \eta \) grows, the high-type messenger gains the ability to
increase the price by an additional amount of \( \eta \), both when consumers pay attention to the ad
and when they do not. Consequently, the regions where a high-type messenger is deemed optimal
expand, indicating that the benefits of employing a high-type messenger are amplified with higher
levels of social utility.

This analysis underscores the significance of considering social utility in messenger selection,
as a greater emphasis on fandom can lead to more favorable outcomes for high-type messengers and
an expanded range of scenarios where they are the optimal choice.

\[ k_{14}(c). \] Otherwise (i.e., \( k < k_{14}(c) \)), full coverage is optimal.

D Differential Fixed Costs

So far, we have only discussed the profit generated by different messenger ads. In practice, the
cost of featuring a messenger can also affect firms' advertising decisions. We assume that different
messenger ads require different amounts of fixed costs \( \phi(m_j) \). In particular, we discuss the case
where \( \phi \equiv \phi(m_H) \geq \phi(m_L) = 0 \). For example, featuring a celebrity or an expert would generally
be more expensive than featuring normal people.

Under full coverage, a high-type messenger is optimal if and only if \( p_H - k - \phi \geq p_L - k \iff c \geq c^* + \frac{\phi}{1 - \sigma_H - \sigma_L} \), which would never hold when \( \varepsilon_H(v) \geq \lambda \varepsilon_L(v) \) for all \( v \). On the other hand, under
partial coverage, a high-type messenger is optimal if and only if \( \tilde{\sigma}_H(p_H - k) - \phi \geq \tilde{\sigma}_L(p_L - k) - c \iff \)
k \leq \( \tilde{p} - \frac{\phi}{\tilde{\sigma}_H - \tilde{\sigma}_L} \). Table 2 summarizes the results of the firm's optimal messenger and price under
each regime.
Partial coverage is optimal if and only if \( k > \frac{\phi}{1 - \sigma_H - \sigma_L} \). Otherwise (i.e., \( k < \frac{1}{\sigma_H - \sigma_L} \)), full coverage with \((m_L, \bar{p}_L)\) is optimal.

2. When \( \varepsilon_H(v) < \lambda \varepsilon_L(v) \), it is optimal to induce partial coverage if \( k \geq k_{18}(c) \). Otherwise, full coverage is optimal. Moreover, under partial coverage, \((m_L, \bar{p}_L)\) is optimal if and only if \( k > \frac{\phi}{1 - \sigma_H - \sigma_L} \).

Also, under full coverage, \((m_L, \bar{p}_L)\) is optimal if and only if \( c < c^{**} + \frac{\phi}{1 - \sigma_H - \sigma_L} \).

Proof. We first discuss the optimal coverage choice under each case. Then, given the optimal coverage choice, the optimal choice of the messenger is given in Table 2.

First, when \( \varepsilon_H(v) \geq \lambda \varepsilon_L(v) \) for all \( v \), we know from Table 2 that only \((m_L, \bar{p}_L)\) can be optimal under full coverage. And, under partial coverage, either \( m_L \) or \( m_H \) can be optimal. Therefore, partial coverage is optimal if the highest expected profit under partial coverage, \( \max\{\tilde{\sigma}_L(E(v|m_L, s_g) - k) - c - \phi, \tilde{\sigma}_L(E(v|m_L, s_g) - k) - c\} \), is larger than the highest expected profits under full coverage, \( E(v|m_L, s_g) - k \), which holds if and only if \( k \geq \min \left\{ \frac{c + \phi + E(v|m_L, s_g) - \tilde{\sigma}_H E(v|m_H, s_g)}{1 - \sigma_H}, \frac{E(v|m_L, s_g) - \tilde{\sigma}_L E(v|m_L, s_g)}{1 - \sigma_L} \right\} = k_{17}(c) \). Otherwise (i.e., \( k < k_{17}(c) \)), full coverage with \((m_L, \bar{p}_L)\) is optimal.

Next, when \( \varepsilon_H(v) < \lambda \varepsilon_L(v) \) for all \( v \), we again know from Table 2 that either \( m_L \) or \( m_H \) can be optimal under any given coverage. Thus, partial coverage is optimal if the highest expected profit under partial coverage, \( \max\{\tilde{\sigma}_L(E(v|m_L, s_g) - k) - c, \tilde{\sigma}_H(E(v|m_H, s_g) - k) - c - \phi\} \), is larger than
Figure 2: Optimal choice of the messenger when different messengers have different fixed costs

the highest expected profit under full coverage, \( \max \{ E(v|m_L, s_b) - k, E(v|m_H, s_b) - k - \phi \} \), which holds if and only if \( k \geq \min \{ \max \left\{ \frac{c+E(v|m_L, s_b)-\bar{\sigma}_L E(v|m_L, s_b)}{1-\bar{\sigma}_L}, \frac{c-\phi+E(v|m_H, s_b)-\bar{\sigma}_L E(v|m_L, s_b)}{1-\phi} \right\} \}, \max \left\{ \frac{c+\phi+E(v|m_L, s_b)-\bar{\sigma}_H E(v|m_H, s_b)}{1-\bar{\sigma}_H}, \frac{c+E(v|m_H, s_b)-\bar{\sigma}_H E(v|m_H, s_b)}{1-\phi} \right\} \} \equiv k_{18}(c) \). Otherwise (i.e., \( k < k_{18}(c) \)), full coverage is optimal.

Figure 2 illustrates Proposition 4, which examines the impact of varying fixed costs on the optimal messenger choice. It is important to note that when the fixed cost parameter \( \phi \) is set to zero, the results align with the main findings. However, as the fixed cost parameter \( \phi \) increases, the desirability of employing a high-type messenger diminishes, leading to smaller regions where a high-type messenger is considered optimal for maximizing firm profits. This implies that higher fixed costs reduce the attractiveness of using high-type messengers, resulting in a narrower range of scenarios where a high-type messenger is preferred. Conversely, lower fixed costs broaden the regions where a high-type messenger is optimal.

This analysis sheds light on the sensitivity of messenger selection to varying fixed costs, highlighting the trade-off between the effectiveness of high-type messengers and the associated costs.

### E Exogenous pricing

In many situations, the prices of different products are similar within a product category. For example, most mobile apps are priced at $0.99, and consumers know the price level through past interactions with the product category. So, they can make attention decisions based on this knowledge of the price of the product category. We explore this case by assuming that the price is exogenously given. This assumption allows us to set aside the price signaling issue and focus on the role of the messenger cleanly as a main medium of information.

When the price is exogenously given, the profit margin \( p - k \) is fixed, so the firm’s goal is to maximize the demand for the product \( D(p, m_j) \). When \( p > E(v) \), consumers’ default action is not to purchase the product, so attracting consumer attention is beneficial for capturing part of the demand from those who receive private good signals. However, when \( p \leq E(v) \), consumers’ default action is to purchase the product, and attracting consumer attention can be harmful to the firm.
because some consumers would opt not to purchase the product upon deliberation, thereby lowering the demand. We characterize the conditions under which a low-type messenger ad can be optimal in the proposition below.

**Proposition 5.**
1. When \( E(v) < p \), it is optimal for the firm to attract consumer attention. In this case, a low-type messenger ad generates a higher profit if \( p_H < p \leq \bar{p}_L \), which is possible only if (1) \( \varepsilon_H(v) < \lambda \varepsilon_L(v) \) for all \( v \) or (2) \( \lambda \varepsilon_L(v) \leq \varepsilon_H(v) < \varepsilon_L(v) \) for all \( v \) and \( c < c^* \).

2. When \( E(v) \geq p \), it is optimal for the firm to dissuade consumer attention. In this case, a low-type messenger ad generates a higher profit if \( \varepsilon_H(v) \geq \lambda \varepsilon_L(v) \) for all \( v \) or (2) \( \kappa \varepsilon_L(v) \leq \varepsilon_H(v) < \lambda \varepsilon_L(v) \) for all \( v \) and \( c < c^{**} \).

*Proof.* First, when \( E(v) < p \), attracting consumer attention is optimal because consumers's default actions are not to purchase the product otherwise. Thus, a low-type messenger can be optimal when \( p_H < p \leq \bar{p}_L \), in which case the price is too high for consumers to pay attention to a high-type messenger, but still low enough for them to pay attention to a low-type one. From Proposition 2 in the main paper, this is possible only when (1) \( \varepsilon_H(v) < \lambda \varepsilon_L(v) \) for all \( v \) or (2) \( \lambda \varepsilon_L(v) \leq \varepsilon_H(v) < \varepsilon_L(v) \) for all \( v \) and \( c < c^* \).

Next, when \( E(v) \geq p \), dissuading consumer attention is optimal because consumers are willing to pay a higher price \( E(v) \) without paying attention. Thus, a low-type messenger can be optimal when \( p_H \leq p < \bar{p}_L \), in which case the price is high enough for consumers to pay attention to the high-type messenger but not high enough for them to pay attention to a low-type one. From Proposition 3 in the main paper, this is possible only when (1) \( \varepsilon_H(v) \geq \lambda \varepsilon_L(v) \) for all \( v \) or (2) \( \kappa \varepsilon_L(v) \leq \varepsilon_H(v) < \lambda \varepsilon_L(v) \) for all \( v \) and \( c < c^{**} \).


Figure 3 illustrates the proposition 5. The dark grey regions are where a high-type messenger ad generates a higher profit, and the light grey regions are where a low-type messenger ad generates a higher profit. In the white regions, because consumers either never pay attention or always pay attention regardless of the type of messenger, the firm is indifferent between a high-type and a low-type messenger ad.

To get a clearer intuition, we discuss the two extreme cases (1) \( \varepsilon_H(v) \geq \varepsilon_L(v) \) for all \( v \) and (2) \( \varepsilon_H(v) < \kappa \varepsilon_L(v) \) for all \( v \). First, when a high-type messenger’s signals are sufficiently informative such that \( \varepsilon_H(v) \geq \varepsilon_L(v) \) for all \( v \), a high-type messenger is more able to attract consumer attention, making them useful when the price is high such that consumers’ default action is not to purchase the product. Therefore, when the deliberation cost \( c \) is not too large, it is optimal to use a high-type messenger ad to attract consumer attention when the price is high and dissuade consumer attention with a low-type messenger ad when the price is low.

On the other hand, when a high-type messenger overshadows the product information such that \( \varepsilon_H(v) < \kappa \varepsilon_L(v) \) for all \( v \), a low-type messenger’s signals can update consumers’ beliefs more significantly, making them more able to attract consumer attention. Thus, when the deliberation cost \( c \) is not too large, it is optimal to use a low-type messenger ad to attract consumer attention when the price is high and dissuade consumer attention with a high-type messenger ad when the price is low.
Figure 3: Optimal choice of messenger when pricing is exogenous

F Full equilibrium characterization when some consumers are naive

In Proposition 6 of the main paper, we only highlighted the conditions under which it is optimal for the firm to attract rational consumer attention with a low-type messenger. We now give a full characterization of the optimal choice of coverage and messengers for the firm when some consumers are naive (Section 5.2)

Proposition 6. 1. When \( \lambda \epsilon_L(v) \leq \epsilon_H(v) \), inducing no coverage with \((m_H, \bar{v})\) is optimal if \( k \geq k_7(c, \alpha) \), inducing full coverage is optimal if \( k \leq k_8(c, \alpha) \), and inducing partial coverage with \((m_H, \bar{p}_H)\) is optimal if \( k_8(c, \alpha) < k < k_7(c, \alpha) \). Moreover, when full coverage is optimal, \((m_L, \bar{p}_L)\) is optimal if \( k > \tilde{k}(c, \alpha) \), and \((m_H, \bar{p}_H)\) is optimal if \( k \leq \tilde{k}(c, \alpha) \).

2. When \( \epsilon_H(v) < \lambda \epsilon_L(v) \), inducing no coverage with \((m_H, \bar{v})\) is optimal if \( k \geq \tilde{k}_9(c, \alpha) \), inducing full coverage is optimal if \( k \leq \tilde{k}_{10}(c, \alpha) \), and inducing partial coverage is optimal if \( \tilde{k}_{10}(c, \alpha) < k < \tilde{k}_9(c, \alpha) \). When partial coverage is optimal, \((m_L, \bar{p}_L)\) is optimal if \( k > \tilde{p} \), and \((m_H, \bar{p}_H)\) is optimal if \( k \leq \tilde{p} \). Moreover, when full coverage is optimal, \((m_L, \bar{p}_L)\) is optimal if \( k > \tilde{k}(c, \alpha) \), and \((m_H, \bar{p}_H)\) is optimal if \( k \leq \tilde{k}(c, \alpha) \).

Proof. We first discuss the optimal coverage choice under each case. Then, given the optimal coverage choice, the optimal choice of the messenger is characterized by the discussion before Proposition 6 in the main paper.

First, when \( \epsilon_H(v) \geq \lambda \epsilon_L(v) \) for all \( v \), we know from the discussion that only \((m_H, \bar{v})\) can be optimal under the no coverage case and only \((m_H, \bar{p}_H)\) can be optimal under the partial coverage case. And, under full coverage case, either \( m_L \) or \( m_H \) can be optimal. Therefore,
no coverage with \((m_H, \tilde{v})\) is optimal if the expected profit it generates, \(\alpha \tilde{\sigma}_H(\tilde{v} - k)\), is larger than both the highest expected profit under partial coverage, \(\tilde{\sigma}_H(\tilde{p}_H - k)\), and the highest expected profit under full coverage, \(\max_{j \in \{L, H\}} [\alpha \tilde{\sigma}_j + (1 - \alpha)](\tilde{p}_j - k)\). This holds if and only if \(k \geq \max \left\{ \tilde{\sigma}_H E(v|m_H, s_h) - \alpha \tilde{\sigma}_H \tilde{v} - c, \max_{j \in \{L, H\}} \frac{[\alpha \tilde{\sigma}_j + (1 - \alpha)]E(v|m_j, s_h) - \alpha \tilde{\sigma}_H \tilde{v} + \frac{\alpha \tilde{\sigma}_j + (1 - \alpha)}{1 - \tilde{\sigma}_j} c}{\alpha \tilde{\sigma}_j + (1 - \alpha) - \alpha \tilde{\sigma}_H} \right\} \equiv k_7(c, \alpha).\) Similarly, full coverage is optimal if the highest expected profit under full coverage is larger than both the highest expected profit under no coverage and the highest expected profit under partial coverage, which holds if and only if \(\max_{j \in \{L, H\}} [\alpha \tilde{\sigma}_j + (1 - \alpha)](\tilde{p}_j - k) \geq \max \{\tilde{\sigma}_H(\tilde{p}_H - k), \alpha \tilde{\sigma}_H(\tilde{v} - k)\} \Leftrightarrow k \leq \max_{j \in \{L, H\}} \left\{ \min \left\{ \frac{[\alpha \tilde{\sigma}_j + (1 - \alpha)]E(v|m_j, s_h) - \alpha \tilde{\sigma}_H \tilde{v} + \frac{\alpha \tilde{\sigma}_j + (1 - \alpha)}{1 - \tilde{\sigma}_j} c}{\alpha \tilde{\sigma}_j + (1 - \alpha) - \alpha \tilde{\sigma}_H}, \frac{[\alpha \tilde{\sigma}_j + (1 - \alpha)]E(v|m_j, s_h) - \alpha \tilde{\sigma}_H \tilde{v} - c}{\alpha \tilde{\sigma}_j + (1 - \alpha) - \alpha \tilde{\sigma}_H} \right\} \right\} \equiv k_8(c, \alpha).\) Otherwise (i.e., \(k_8(c, \alpha) < k < k_7(c, \alpha)\), partial coverage with \((m_H, \bar{p}_H)\) is optimal.

Next, when \(\varepsilon_H(v) < \lambda \varepsilon_L(v)\) for all \(v\), again we know from the discussion that only \((m_H, \bar{v})\) is optimal under no coverage. And, under partial and full coverage, either \(m_L\) or \(m_H\) can be optimal. So, no coverage with \((m_H, \bar{v})\) is optimal if the expected profit it generates, \(\alpha \tilde{\sigma}_H(\bar{v} - k)\), is larger than both the highest expected profit under partial coverage, \(\max_{j \in \{L, H\}} \tilde{\sigma}_j(\bar{p}_j - k)\), and the highest expected profit under full coverage, \(\max_{j \in \{L, H\}} [\alpha \tilde{\sigma}_j + (1 - \alpha)](\bar{p}_j - k)\). This holds if and only if \(k \geq \max \left\{ \max_{j \in \{L, H\}} \tilde{\sigma}_j E(v|m_j, s_h) - \alpha \tilde{\sigma}_H \tilde{v} - c, \max_{j \in \{L, H\}} \frac{[\alpha \tilde{\sigma}_j + (1 - \alpha)]E(v|m_j, s_h) - \alpha \tilde{\sigma}_H \tilde{v} + \frac{\alpha \tilde{\sigma}_j + (1 - \alpha)}{1 - \tilde{\sigma}_j} c}{\alpha \tilde{\sigma}_j + (1 - \alpha) - \alpha \tilde{\sigma}_H} \right\} \equiv \tilde{k}_9(c, \alpha).\) And, full coverage is optimal if the highest expected profit under full coverage is larger than both the highest expected profit under no coverage and the highest expected profit under partial coverage, which holds if and only if \(\max_{j \in \{L, H\}} [\alpha \tilde{\sigma}_j + (1 - \alpha)](\bar{p}_j - k) \geq \max \{\max_{j' \in \{L, H\}} \tilde{\sigma}_{j'}(\bar{p}_{j'} - k), \alpha \tilde{\sigma}_H(\bar{v} - k)\} \Leftrightarrow k \leq \max_{j \in \{L, H\}} \left\{ \min \left\{ \frac{[\alpha \tilde{\sigma}_j + (1 - \alpha)]E(v|m_j, s_h) - \alpha \tilde{\sigma}_H \tilde{v} + \frac{\alpha \tilde{\sigma}_j + (1 - \alpha)}{1 - \tilde{\sigma}_j} c}{\alpha \tilde{\sigma}_j + (1 - \alpha) - \alpha \tilde{\sigma}_H}, \frac{[\alpha \tilde{\sigma}_j + (1 - \alpha)]E(v|m_j, s_h) - \alpha \tilde{\sigma}_H \tilde{v} - c}{\alpha \tilde{\sigma}_j + (1 - \alpha) - \alpha \tilde{\sigma}_H} \right\} \right\} \equiv \tilde{k}_{10}(c, \alpha).\) Otherwise (i.e., \(\tilde{k}_{10}(c, \alpha) < k < \tilde{k}_9(c, \alpha)\), partial coverage with \((m_H, \bar{p}_H)\) is optimal.