

MANAGING BUZZ

Arthur Campbell*, Dina Mayzlin** and Jiwoong Shin****

June 2016

Abstract

We model the incentives of individuals to engage in word of mouth (or buzz) about a product, and how a firm may strategically influence this process through its information release and advertising strategies. Individuals receive utility by improving how others perceive them. A firm restricts access to information, advertising may crowd out word of mouth, and a credible commitment not to engage in advertising is valuable for a firm.

Keywords: Buzz, word of mouth, self-enhancement, signaling, advertising, diffusion.

*Monash University; arthur.campbell@monash.edu, **USC Marshall School of Business; mayzlin@marshall.usc.edu, ****Yale School of Management; jiwoong.shin@yale.edu.

We thank Dirk Bergemann, Alessandro Bonatti, Odilon Camara, Florian Ederer, Johannes Horner, Yuichiro Kamada, Duncan Simester, Birger Wernerfelt, and seminar participants at Brown University, CKGSB, Columbia University, Indiana University, MIT, Monash University, Paris School of Economics, Purdue University, UC Berkeley, University of Cambridge, University of Florida, University of Melbourne, University of Toronto, Wharton, Yale University, the 2012 SICS Conference at Berkeley, and the 2011 Conference on the Economics of Advertising in Moscow.

1 Introduction

Word of mouth is an important driver of consumers' purchase decisions.¹ However, despite its importance, the existing literature typically treats word of mouth as a costless and mechanical process and offers little insight into an individual's incentive to engage in word of mouth. In this article, we introduce a framework to model how consumers' concerns about social status drive their incentives to engage in word of mouth. We focus on a particular reputational motive – word of mouth as “self-enhancement” (Baumeister 1998) or the idea that an individual engages in word of mouth to improve how she is perceived by the listener. In contrast to a setting where consumers mechanically undertake word of mouth, we show that a firm may improve the incentives for individuals to engage in word of mouth by restricting early access to product information.

As a real-life illustration of our finding that firms may benefit from actively restricting early access to product information, consider the US launch of the European music streaming site Spotify in July 2011. At first, Spotify's free US version was available by invitation only. Interestingly, obtaining an invitation was non-trivial, and direct invitations were limited to certain groups: consumers could receive one either through current users or through other channels. For example, the company sent invitations to users who interacted with Spotify on Twitter, and Coca-Cola gave out invitations to users who submitted their email address.² After a few weeks, anyone could download the free version of Spotify through the company's website.³ By November 2011, Spotify was able to attract 4 million users, although it undertook almost no advertising. Media sources speculated that the initial exclusivity surrounding the site contributed to early buzz and high adoption rates.⁴

The Spotify example is by no means unique. Many marketing practitioners recommend and use similar strategies which limit access to information in order to spur word of mouth. For example, Hughes (2005) states, “Sometimes withholding can work better than flooding. Limit supply and everybody's interested. Limit those in the know of a secret, those not in the know want the currency of knowing - they want to be part of the exclusive circle.” David Balter, the founder of the buzz

¹Word of mouth has been shown to affect purchasing behavior in restaurant choices (Luca 2013), book sales (Chevalier and Mayzlin 2006), banking (Keaveney 1995), entertainment (Chintagunta et al. 2010), technological products (Herr et al. 1991), and appliances and clothing (Richins 1983). These studies are also consistent with recent industry research: for example, according to *Word of Mouth Marketing Association* (2011), 54% of purchase decisions are influenced by word of mouth. Also, “Word of mouth is the primary factor behind 20 to 50 percent of all purchase decisions” in *McKinsey Quarterly* (Bughin et al. 2010), and “word of mouth remains the biggest influence in people's electronics (43.7%) and apparel (33.6%) purchases,” *National Retail Federation* (2009).

²http://news.cnet.com/8301-13845_3-20081418-58/get-a-quick-and-easy-invitation-to-spotify/

³One possible reason for a firm's initial limited release could be a beta version of the product in a test market for the purposes of collecting feedback from users about the product's functionality before its wide release. This explanation is less applicable to the Spotify case, given its presence and operational volume in Europe by the time it launched in the US market in 2011 – Spotify had already become the most popular service of its kind in the world; it had 1.6 million paid subscribers and more than 10 million registered users in total (<http://www.nytimes.com/2011/07/14/technology/spotify-music-streaming-service-comes-to-us.html>).

⁴See, for example, “Spotify's ascension can be largely attributed to word of mouth” (<http://www.theverge.com/2013/3/25/4145146/spotify-kicks-off-ad-blitz-as-rumors-hint-of-video-service>).

marketing firm BzzAgent, considers exclusivity to be one of the necessary ingredients for a successful word of mouth campaign. “Exclusivity is the velvet rope of social media: everyone wants to be special enough to be on the right side of it.”⁵ Also, Sernovitz (2011) observes that “Many people are more likely to talk about a product if there is some kind of insider access or privileged status” and provides a number of examples where firms use exclusivity strategy to increase word of mouth about their products. For instance, retailers sometimes offer private shopping hours for their select customers the night before new products are available to the public, and software companies send pre-release versions of new software to active message board users. Our model explains why purposely limiting the number of people who are initially exposed to product information may in fact maximize overall product adoption.

We develop a model where consumers meet one another at a Poisson rate over time. Consumers are heterogeneous in that they are either high or low type. The most straightforward interpretation of high type here is being knowledgeable about a particular product area or having category expertise: for example, having good taste in wine, being technologically savvy, knowing the best restaurants and bars, or having good taste in music. The key element of the model is that the utility an individual receives during a social interaction is an increasing function of her peer’s belief that she is the high type. Prior to meeting others, individuals choose whether or not to undertake a search to acquire information about the firm’s product at a certain individual-specific cost. Then, during each social interaction, individuals decide whether or not to engage in costly word of mouth. We focus on a signaling equilibrium where word of mouth about a product serves as a credible signal of high type. The central focus of our analysis is how the firm can manage the extent of the information diffusion in this context.

We broadly consider two types of strategies by the firm. First, we consider information release strategies where the firm imposes differential costs for information acquisition on different types of consumers. When the costs of acquiring information for the low type are high enough, there exists a signaling equilibrium where individuals acquire information and then pass it on through word of mouth to people they meet. As information diffuses in the population, and low social types acquire information through word of mouth, the signaling value of the information becomes diluted, and diffusion stops when the signaling value is equal to the cost of engaging in word of mouth. Hence, the firm can influence the extent of the diffusion by manipulating the asymmetry in the cost of acquiring information across different groups of consumers. The optimal information release strategy of the firm is to maximally increase the costs of the low type and minimize the costs of the high type. This enables us to explain why exclusivity strategies by a firm that restrict (in particular ways) who has access to information about the product may in fact increase the total amount of information that is shared in the population.

⁵<http://blog.hubspot.com/blog/tabid/6307/bid/6183/4-Social-Media-Methods-for-Generating-Word-of-Mouth.aspx>

Second, we introduce advertising⁶ by the firm and consider how the amount of advertising affects the incentives for consumers to undertake a costly search to acquire information and engage in word of mouth. We find that advertising by the firm crowds out the incentives of individuals to acquire information and engage in word of mouth. Hence, a commitment by the firm not to engage in advertising can increase the diffusion of information. We show that a natural way for a firm to commit to advertise less is to release the information a sufficient amount of time prior to the product's release. This decreases the present value of the benefit of advertising for the firm, thereby allowing a sufficient amount of word of mouth to occur. Our results accord well with observations that word of mouth tends to occur prior to mass advertising being undertaken by a firm. Thus, we characterize consumer search and word of mouth behaviors, and their interactions with a firm's advertising activities. We also present a general result that any source of information (such as increased advertising, media attention or public relations that distribute information in an untargeted way) will reduce the amount of information acquired by consumers and the extent of information diffusion.

The rest of the article is organized as follows. In the next section, we discuss the related literature. Section 3 presents a model of buzz based on the self-enhancement motive and analyzes how firm interacts with this motive to maximize the diffusion of information through its information releasing strategy. In Section 4, we examine the effect of advertising on word of mouth generation. In Section 5, we discuss several alternative approaches, and we conclude in Section 6.

2 Literature Review

Our work is most closely related to the literature in social network theory that studies a firm's optimal strategy in the presence of learning or adoption externality through some forms of local interactions by consumers, including word of mouth. Typically, these articles analyze how characteristics of the social network interact with a firm's pricing (Galeotti 2010, Candogan, Bimpikis and Ozdaglar 2010, and Ifrach, Maglaras and Scarsini 2011), or advertising strategy (J. Campbell 2013, Galeotti and Goyal 2007, Bimpikis, Ozdaglar and Yildiz 2015 and Chatterjee and Dutta 2010), or both (A. Campbell 2013).

In particular, Galeotti (2010) considers the costly search for pricing information by consumers in a model where two firms engage in Bertrand competition. Similarly to our work, it is costly for consumers to acquire information directly. In contrast to the current article - where we assume that it is costly for the sender to pass on information but not for the receiver to receive it - Galeotti models word of mouth as costly for the receiver but not for the sender. The difference in assumptions from

⁶There are several different perspectives on advertising, such as the informative, persuasive and complementary views of advertising (see Bagwell 2007 for a comprehensive survey of the literature). In this article, we only consider the informative view of advertising; thus, the purpose of advertising is purely to provide information about the existence of the product to consumers.

our work is due to the difference in the earlier article’s focus. The current article is focused on the motivations of individuals with information to engage in word of mouth (the senders of information). On the other hand, Galeotti is concerned with the equilibrium of consumer search (conducted by the receivers) either directly or through friends and firm pricing.

Many of the other articles in the existing literature have detailed models of the social network, but they treat word of mouth generation as a mechanical process, whereby a consumer passes on information upon acquiring it, and the word of mouth stops after a certain number of steps (or with some probability after each step). Galeotti and Goyal (2007) and J. Campbell (2010) assume that firms initially advertise to consumers, and then word of mouth travels a distance of one in the social network.⁷ Chatterjee and Dutta (2010) assume that the firm can pay individuals to engage in word of mouth. Ifrach et al. (2011), Candogan et al. (2010), and A. Campbell (2013) consider settings where consumers pass on information if they are prepared to purchase the product. This line of analysis has been successful at relating characteristics of the social environment (such as frequency of connections/interactions, distribution of friendships, and clustering of friendships between members of the population) to a firm’s strategy. Given the assumption of word of mouth as a mechanical process in these models, any firm strategy which increases the propensity of any consumer to hold information will facilitate a greater amount of information diffusion. In contrast, our work addresses the issue of why individuals engage in word of mouth about a firm’s product by explicitly modeling an individual’s social motivation. In addition, there are a number of articles which examine information diffusion through endogenous (privately motivated) communication between individuals in social networks (Galeotti and Mattozzi 2011, Calvó-Armengol et al. 2012, Stein 2008, Niehaus 2011). However, none of these articles study the effect of exclusivity on the information diffusion. Our social signaling mechanism for word of mouth leads to novel insights into how a firm may *increase* the diffusion of information through *restrictions* on information acquisition and advertising: optimally the firm imposes restrictions in ways that increase the signaling value of the information.

Our article deals with the incentives for individuals to engage in word of mouth. Prior research in psychology and marketing has proposed several distinct psychological motives that can drive word of mouth communication. For example, some studies have found that word of mouth can be driven by altruism (Henning-Thurau et al. 2004, Sundaram, Mitra, and Webster 1998), reciprocity (Dichter 1966, Dellarocas et al. 2004) or the desire to signal expertise to others (Wojnicki and Godes 2011). Although any of these motives can independently drive word of mouth, in this article, we focus on the latter desire to signal to others in a social setting. The starting point of our model is that consumers derive benefits during social interactions from making themselves “look good.” This assumption is motivated by the psychological theory of “self-enhancement” or the tendency to “affirm the self” (Baumeister 1998, Fiske 2001, Sedikides 1993) and includes the tendency to draw attention to one’s skills and

⁷Galeotti and Goyal (2007) also extend their untargeted advertising results to a generalized maximum distance.

talents (Baumeister 1998, Wojnicki and Godes 2011). There is a growing empirical literature in marketing that focuses on the self-presentation motive as a driver of word of mouth (Berger and Milkman 2012, Berger and Schwartz 2011, Hennig-Thurau, et. al. 2004, Sundaram et al. 1998, Wojnicki and Godes 2011, and De Angelis et al. 2012).⁸

Finally, our model studies how social interactions between consumers can be influenced by the firm’s strategy. Pesendorfer (1995) analyzes the interaction between a firm’s design innovation and pricing strategy and consumers’ social matching behavior. In his model, owning a fashionable product serves as a matching mechanism: it allows high-type consumers to interact with other high types who also own the product. Pesendorfer’s focus is on product cycles and pricing whereas our focus is on how a firm’s information release and advertising strategy interacts with consumers’ social concerns. Yoganarasimhan (2012) also models a fashion firm’s desire to withhold the identity of its “hottest” product in order to enable consumers to signal to each other that they are “in the know” in social interactions. Our work is similar to Yoganarasimhan’s in that both articles model the firm’s incentive to restrict information in their communication strategy, in order to facilitate the social interaction. However, Yoganarasimhan analyzes the firm’s pricing strategy to extract consumer surplus in a static setting, whereas we focus on the effect of initial exclusive release on the extent of information diffusion in a dynamic setting.

3 A Model of Buzz Management

Model Set-up

A monopolist sells a product to a mass 1 of consumers. A consumer i may be one of two types: high or low $\theta = h, l$ where $\Pr[\theta_i = h] = \alpha < \frac{1}{2}$ (high types are relatively scarce). Consumers are privately informed of their own type. In social situations, it is valuable for either type of person to be perceived as the high type by others - a reputational motive that has been referred to as “self-enhancement” in the word of mouth literature. One can think of the high-type consumers as being broadly knowledgeable about the product category (wine enthusiast, technology-savvy, knows all the fashionable/trendy restaurants and bars). Importantly, it is valuable to be perceived as a high type, regardless of the consumer’s true type. Thus, our model fits particularly well in many entertainment, technology and fashion product categories where being perceived as knowledgeable about these areas is desirable.

We assume that consumers can purchase only if they are informed about the product, and thus the firm’s objective is to maximize the fraction of the population that is informed about the product. We denote the fraction of the population that has received the information m about the product at time

⁸The concept of self-enhancement is closely related to the concept of reputation-building. We use the term “self-enhancement” to be consistent with the existing word of mouth literature.

t by $S(t)$. Initially, we assume no advertising before introducing it in Section 4. Without advertising, consumers can obtain the information in two ways. First, they may undertake a costly search to learn about the product themselves. Second, they may costlessly hear about the product from another person.

When a consumer i decides to undertake a costly search to learn about the product, she incurs a search cost for obtaining information c_i , which is i.i.d. uniformly on $[0, \bar{c}]$.⁹ One can think of this as the minimum amount of time and effort an individual must expend to understand the information. The firm, in addition to this cost, may impose further costs on either or both types through its information release strategy. This is modeled as an additional cost $\nu_h, \nu_l \geq 0$, which is type-specific. Hence, the total cost that an individual i of type $\theta = \{h, l\}$ bears to obtain information m is $c_i + v_\theta$. Here, v_θ is the type-specific part of the information acquisition cost that is under the firm's control, and c_i is the person-specific part of the cost that is outside of the firm's control.

We assume that imposing a type-specific cost is costless (or involves a very small cost) for the firm. For example, the firm can explicitly increase consumer information acquisition costs through the use of technical jargon (which the high type more easily understands), or equivalently can decrease the cost of the high type relative to the low type by releasing information on blogs, at events, or in venues that are frequented by high types but not low types. What is important for the model is that the firm may differentially affect the costs of each type and that these costs are common knowledge. Using the Spotify example, the firm created an asymmetry in acquisition costs by posting the invitation to register on its Twitter feed. Hence, all consumers could potentially obtain the invitation, but the cost of acquiring it was lower for the tech-savvy consumers who were already familiar with Twitter. Another prominent example of this occurs when technology companies such as Apple or Samsung make product announcements at industry events, which are broadcast through live feeds. Again in this instance, tech-savvy individuals have lower costs to find, monitor, and even understand these sources. Thus, being able to engage in word of mouth about the information serves as a signal of an individual's tech-savvy.

Once the consumer has acquired the information about the product m through either her own costly search or word of mouth from the other person, she too is able to pass on the hard information to others at a cost k during the interactions with others, where we assume $\alpha < k < 1 - \alpha$. One can think of this as the cost of time and effort spent transmitting the information to others. For example, consider an early user of Spotify who tells about the service to her friend. She has to explain what makes the music service distinct from others, which may take both time and effort. Note that we do not specify whether the conversation must take place in a face-to-face interaction or electronically,

⁹In some situations such as fashion or technology product categories, it would be natural to assume that these costs would be higher for the low type than the high type. However, we do not do this because this assumption is not a necessary condition for the finding that the firm has a strict incentive to treat each type asymmetrically through its information release strategy.

but we do require that the communication takes some effort. That is, our model does not extend to a situation where a person effortlessly communicates by liking a Facebook page for example.

To summarize, there are two distinct types of costs in our model: the cost of communication (k) and the cost of information acquisition ($c_i + v_\theta$). It is important to note that both costs are essential to our model and results. On the one hand, the cost of communication determines the stopping rule of word of mouth diffusion. This cost is driven by the individual's costs of time and effort and is outside the control of the firm. On the other hand, the cost of information acquisition determines the proportion of consumers of each type who are informed about the product. In contrast, this cost is partially under the firm's control. Relaxing either one would render the main problem trivial: (1) costless information acquisition would result in the entire population being informed even in the absence of word of mouth, and (2) costless communication would lead to word of mouth being diffused to the entire population if a single individual acquires the information even in the absence of any asymmetry between types.

Timing

At time $t = -1$, each individual chooses whether to obtain information m about a firm's product. We assume that this information is hard and verifiable, and a consumer is not able to fabricate information. Note that in our model, the content of information acquired by the consumer does not affect its signaling value. It is the mere fact that the consumer is able to acquire the information that serves as a signal. As we discussed above, an individual i of type $\theta = \{h, l\}$ incurs total costs of $c_i + v_\theta$ to obtain information about the product, where c_i is the individual's search costs and v_θ is a type-specific cost imposed by the firm.

From time $t > 0$, individuals meet others at rate λ in continuous time. During each meeting, an individual who has acquired the information previously may pass on the hard information m at a cost k , or pass on no information \emptyset at zero cost. We assume that communication is done simultaneously during the meeting so that each individual has the ability to do so without seeing the other individual's information first.

Utility During Social Interactions

A central element of our model is that individuals derive a benefit from word of mouth due to reputational concerns. In other words, an individual engages in word of mouth in order to make herself look good. We capture this idea through the social utility U_{ij} that an individual i receives from an interaction with another individual j , where the utility is an increasing function of the beliefs the other consumer has about the focal consumer's type. In particular, consumer i receives instantaneous utility $U_i(b_j(\theta_i = h|m, t))$ if consumer i passes a message m at time t , where $b_j(\theta_i = h|m, t)$ is j 's belief that consumer i is a high type upon receiving the information m . And similarly, i receives

instantaneous utility $U_i(b_j(\theta_i = h|\varnothing, t))$ if she does not pass on information, where $b_j(\theta_i = h|\varnothing, t)$ is the belief if no signal (denoted by \varnothing) is sent.¹⁰ Given our notion of high and low types, we assume $\frac{dU_i}{db_j} > 0$. Also, note that the signaling benefit at a time t is the difference between sending a signal and not sending a signal at that time t . We assume that utility is linear in beliefs; thus,

$$\begin{aligned}\Delta U_i(t) &= U_i(b_j(\theta_i = h|m, t)) - U_i(b_j(\theta_i = h|\varnothing, t)) \\ &= b_j(\theta_i = h|m, t) - b_j(\theta_i = h|\varnothing, t)\end{aligned}\tag{1}$$

We also assume that an individual passes on information only when the benefit of doing so is strictly positive: $\Delta U_i(t) > k$. In other words, we assume that if an individual is indifferent between passing or not passing on the information, ($\Delta U_i(t) = k$), she chooses not to pass it on. This assumption precludes some unreasonable equilibria where only a fraction of individuals with information choose to pass it on due to indifference, and these fractions differ across types such that $\Delta U_i(t) = k$ is maintained over time. These types of equilibria are unreasonable, as they arbitrarily introduce an asymmetry between the types by manipulating indifference in a very specific manner.¹¹

Finally, we assume that the firm extracts the entire consumer surplus from the sale of the product to the consumer. Thus, from the consumer's perspective there is no value in obtaining product information for the purposes of making a purchase decision; the only value from information is as a potential signaling device to others. This allows us to focus on the incentives of individuals to acquire information in order to pass it on to others.

Analysis

We focus on a signaling equilibrium where word of mouth serves as a credible signal of high type. We focus on equilibria where individuals engage in word of mouth while the signaling benefit is strictly greater than the costs of passing on information.

The fraction of each type who acquires information at $t = -1$ determines the total amount of information diffusion. We denote the fraction of each type who becomes informed at $t = -1$ by φ_h, φ_l . These will be endogenously determined in equilibrium, but for now we take the vector $\varphi = (\varphi_h, \varphi_l)$ as given. For the moment, we also assume that $\varphi_h > \varphi_l$, which will be confirmed in equilibrium subsequently.

The fraction of the population which is informed about the product, $S(t)$, evolves over time as individuals mix at rate λ and pass on information. The initial condition at $t = -1$ is $S_0 =$

¹⁰More precisely, the belief b_j is also affected by the initial fraction of each type who acquires information at $t = -1$. We introduce this in the next subsection, but we drop this notation here for simplicity.

¹¹Furthermore, the extent of diffusion found in this way is not robust to incorporating individual specific costs of passing on information, which are drawn from $[k, k + \epsilon]$ where ϵ is arbitrarily small. In our equilibria, all individuals with information act in the same way: they either all pass it on or do not.

$\varphi_h \alpha + \varphi_l (1 - \alpha)$, and the rate of change of the informed population is given by $\frac{dS}{dt} = \lambda S(t) (1 - S(t))$. This results in the following path for $S(t)$:

$$S(t) = \frac{1}{1 + a e^{-\lambda t}}, \text{ where } a = \frac{1 - S_0}{S_0} \quad (2)$$

which continues to grow until the benefit of passing on the information is less than or equal to the cost of doing so.

Beliefs change over time as the message diffuses through the population. At $t = 0$, consumers' beliefs are

$$\begin{aligned} b_j(\theta_i = h|m, 0, \varphi) &= \frac{\varphi_h \alpha}{\varphi_h \alpha + \varphi_l (1 - \alpha)} \\ b_j(\theta_i = h|\emptyset, 0, \varphi) &= \frac{(1 - \varphi_h) \alpha}{(1 - \varphi_h) \alpha + (1 - \varphi_l) (1 - \alpha)} \end{aligned} \quad (3)$$

Note that no word of mouth (no signal) at any time t implies that the sender did not receive the information at $t = -1$. The h and l type are equally likely to hear (or not to hear) about the product through others' word of mouth for any $t \geq 0$. Hence, the belief on the consumer's type, conditional on no signal, is constant over time:

$$b_j(\theta_i = h|\emptyset, t, \varphi) = b_j(\theta_i = h|\emptyset, 0, \varphi). \quad (4)$$

On the other hand, a sender who shares word of mouth at time t either obtained the information on her own at $t = -1$ or heard it from a friend. Hence, the belief on the sender's type upon receiving word of mouth at time t is a mixture of the beliefs at time 0 ($b_j(\theta_i = h|m, 0, \varphi)$ and $b_j(\theta_i = h|\emptyset, 0, \varphi)$), weighted by the relative fraction of the individuals who acquired information on their own, $\frac{S_0}{S(t)}$, and individuals who heard it from a friend, $1 - \frac{S_0}{S(t)}$:

$$b_j(\theta_i = h|m, t, \varphi) = \frac{S_0}{S(t)} b_j(\theta_i = h|m, 0, \varphi) + \left[1 - \frac{S_0}{S(t)}\right] b_j(\theta_i = h|\emptyset, 0, \varphi). \quad (5)$$

The instantaneous signaling value at t is:

$$\begin{aligned} \Delta U(t) &= U(b_j(\theta_i = h|m, t, \varphi)) - U(b_j(\theta_i = h|\emptyset, t, \varphi)) \\ &= \frac{S_0}{S(t)} [b_j(\theta_i = h|m, 0, \varphi) - b_j(\theta_i = h|\emptyset, 0, \varphi)] \end{aligned} \quad (6)$$

While word of mouth is taking place, $S(t)$ is strictly increasing over time, and thus the instantaneous benefit of signaling is strictly decreasing over time. That is, as information diffuses through the population, and as more low types receive information through word of mouth, the signaling value

of passing on information decreases. Under our assumption that $\alpha < k < 1 - \alpha$, there exists a time t^* when the instantaneous benefit of word of mouth is exactly equal to the cost of transmission $\Delta U(t^*) = k$, at which point word of mouth stops.¹² The extent of diffusion is

$$S^*(t^*) = \frac{S_0}{k} [b_j(\theta_i = h|m, 0) - b_j(\theta_i = h|\emptyset, 0)] \quad (7)$$

[Insert Figure 1 Here]

Figure 1 illustrates the effect of consumer beliefs on information diffusion. The top figure shows the evolution of consumer beliefs over time, and the bottom figure shows the growth of the informed population $S(t)$. Note that $S(t)$ stops growing when $\Delta U(t^*) = k$, which defines the extent of the diffusion $S^* = S(t^*)$.

We first examine how the extent of information diffusion depends on the fraction of high and low types who acquire information.

Proposition 1. *The total information diffusion is increasing (decreasing) in the fraction of high (low) types who acquire information: $\frac{\partial S^*(\varphi)}{\partial \varphi_h} \geq 0$, $\frac{\partial S^*(\varphi)}{\partial \varphi_l} \leq 0$, where the former inequality is strict if $\varphi_l < 1$, and the latter is strict if $\varphi_h < 1$.*

From Equation (7), we can see that the total diffusion is the product of the proportion of the population who acquires the information ($\frac{S_0}{S(t)}$) and the instantaneous signaling benefit ($b_j(\theta_i = h|m, 0) - b_j(\theta_i = h|\emptyset, 0)$) at $t = 0$. The latter term can also be interpreted as the informativeness of word of mouth as a signal of h type: the difference in the posterior belief following a message versus no message. Higher φ_h increases both the initial spread of information and the informativeness of word of mouth as a signal. Hence, the total diffusion of information is increasing in φ_h . Similarly, lower φ_l increases the informativeness of word of mouth as a signal of h type. However, lower φ_l also decreases the initial spread of information. In our model, the former effect dominates the latter: the extent of the diffusion of the firm's message is decreasing in φ_l . That is, increasing the initial information asymmetry between the two types (increasing the proportion of high types who are informed and decreasing the proportion of low types who are informed) benefits the firm in the long run by maximizing the overall diffusion of information. An immediate result of the proposition is that the diffusion of information is maximized in the event that all the high types acquire information and none of the low types do, $\varphi_h = 1$ and $\varphi_l = 0$: maximal asymmetry in information. Proposition 1 presents the results for exogenously given levels of information acquisition φ_h and φ_l . In the next section, we consider the incentives of each consumer type to acquire information at $t = -1$, and the optimal way for the firm to influence these incentives through the additional costs it imposes.

¹²We restrict our analysis to strict equilibria. For $t \geq t^*(\varphi)$, word of mouth does not occur. As long as the out-of-equilibrium belief is such that $b_j(\theta_i = h|m, t) < b_j(\theta_i = h|m, t^*(\varphi))$ for all $t \geq t^*(\varphi)$, consumers prefer not to spread word of mouth upon reaching $t^*(\varphi)$.

Consumers' Incentives to Acquire Information

We solve for a perfect Bayesian equilibrium of the model by solving for the consumers' decision to acquire product information at $t = -1$. The decision to acquire the information at $t = -1$ depends on the total signaling benefit of word of mouth during the diffusion process. If this benefit is greater than an individual's cost $c_i + \nu_\theta$, then the consumer will acquire information. Denoting the time at which the diffusion process ends by t^* , a consumer that acquires information at $t = -1$ derives the following expected signaling benefit (aggregated over time)::

$$W^{acq} = \lambda \left(\int_0^{t^*} (\Delta U(t) - k) dt \right) \quad (8)$$

Note that even if the consumer does not acquire information at $t = -1$, she may still at some point learn about the product through word of mouth diffusion and then go on to derive a signaling benefit from passing this information on to others. The fraction of the population that did not acquire product information at $t = -1$ is $1 - S_0$. The fraction of the population that is informed at time t through word of mouth is $S(t) - S_0$. Hence, the probability that a consumer who did not acquire information at $t = -1$ would acquire it through word of mouth by time t is $\frac{S(t) - S_0}{1 - S_0}$. Hence, a consumer that does not acquire information at $t = -1$ derives the following expected signaling benefit (aggregated over time):

$$W^{no\ acq} = \lambda \left(\int_0^{t^*} \frac{S(t) - S_0}{1 - S_0} (\Delta U(t) - k) dt \right) \quad (9)$$

Combining the two equations above, the marginal benefit from acquiring information is

$$V \equiv W^{acq} - W^{no\ acq} = \lambda \left(\int_0^{t^*} \left(\frac{1 - S(t)}{1 - S_0} \right) (\Delta U(t) - k) dt \right) \quad (10)$$

where the first term $\frac{1 - S(t)}{1 - S_0}$ is the probability of remaining uninformed at time t for an individual who is uninformed at time 0,¹³ and the second term, $\Delta U(t) - k$, is the signaling benefit at each moment in time. We can further simplify the expression to obtain the following:¹⁴

$$V = \left(\frac{k}{1 - S_0} \right) \left(\frac{S^* - S_0}{S_0} + \ln \frac{S_0}{S^*} \right) \quad (11)$$

Here the total signaling benefit is expressed as a function of only the initial diffusion state (S_0) and the total extent of information diffusion (S^*), both of which are functions of φ_h and φ_l . Thus, $V = V(\varphi)$.

¹³ $1 - S_0$ is the proportion of consumers who are uninformed at $t = 0$; $S(t) - S_0$ is the proportion of consumers who become informed through word of mouth by time t . Therefore, the probability that a consumer who is initially uninformed remains uninformed by time t is $\frac{(1 - S_0) - (S(t) - S_0)}{1 - S_0} = \frac{1 - S(t)}{1 - S_0}$.

¹⁴We do so by making a change of variables for dt , $\frac{dS}{dt} = \lambda S(t)(1 - S(t)) \Leftrightarrow dt = \frac{dS}{\lambda S(t)(1 - S(t))}$

The firm can increase the costs of the high and low types for acquiring information differentially by imposing $\nu_l \geq 0$ and $\nu_h \geq 0$. This can create information asymmetry between the two types, which enables an equilibrium where word of mouth can serve as a signaling device. At $t = -1$, a customer i of type θ chooses to acquire information if $c_i + v_\theta \leq V(\varphi)$. The proportion of consumers of type θ who choose to acquire information is

$$\varphi_\theta = \Lambda(V(\varphi), v_\theta) = \begin{cases} 0 & \text{if } V - v_\theta < 0 \\ \frac{V - v_\theta}{\bar{c}} & \text{if } 0 \leq V - v_\theta \leq \bar{c} \\ 1 & \text{if } V - v_\theta > \bar{c} \end{cases} \quad (12)$$

Note that the relationship between v_θ and φ_θ is not one-to-one at the extremes of φ_θ and we assume that if $\Lambda(V(\varphi), \bar{v}_\theta) = \Lambda(V(\varphi), \tilde{v}_\theta)$ for $\bar{v}_\theta \leq \tilde{v}_\theta$, the firm chooses \bar{v}_θ .

Optimal Information Release Strategy

The firm's optimization problem is

$$\max_{\nu_h, \nu_l} S^*(\varphi) \quad (13)$$

subject to

$$\varphi_h = \Lambda(V(\varphi), \nu_h) \quad (14)$$

$$\varphi_l = \Lambda(V(\varphi), \nu_l) \quad (15)$$

$$\nu_h \geq 0, \nu_l \geq 0 \quad (16)$$

We refer to Equation (14) as the high-type consumer's information acquisition constraint and to Equation (15) as the low-type's information acquisition constraint. Before finding the optimal strategy for the firm, we note that asymmetry in costs is a necessary condition for word of mouth to take place.

Lemma 1. *No word of mouth occurs under symmetric costs ($\nu_h = \nu_l$).*

The underlying driver for the word of mouth diffusion is the signaling benefit which arises from the initial asymmetry in information between the high and low types ($\varphi_h > \varphi_l$) in Equation (7). Under symmetric costs (which leads to $\varphi_h = \varphi_l$ as c_i for both types is from the same *i.i.d.* uniform $[0, \bar{c}]$), there is no signaling benefit to word of mouth. Because in our model the only purpose of acquiring information is as a signaling device, this implies that $S^* = 0$. Therefore, no word of mouth arises.

The following proposition characterizes the optimal level of firm-imposed asymmetry in costs across types.

Proposition 2. *When $\bar{c} \geq \frac{1-k+k \ln k}{1-\alpha}$, the optimal strategy for the firm is to impose a sufficiently large cost on the low type and zero cost on the high type (large ν_l and $\nu_h = 0$) such that $\varphi_l = 0$ and $0 < \varphi_h < 1$.*

The proposition demonstrates that when there exists a sufficient amount of heterogeneity in consumers' information acquisition costs ($\bar{c} \geq \frac{1-k+k \ln k}{1-\alpha}$), the optimal strategy is to restrict information to the low-type consumers by setting the costs of information acquisition sufficiently high (large ν_l) and minimizing the costs for the high-type consumers ($\nu_h = 0$). The result demonstrates that the firm benefits from maximal asymmetry in initial information acquisition between the two types. Note that in this case, some of the high types' total acquisition costs ($c_h + \nu_h$) are so high that they choose not to acquire information about the product: $\varphi_h < 1$. However, if $\bar{c} < \frac{1-k+k \ln k}{1-\alpha}$, any cost strategy that implements $\varphi_h = 1$ is optimal. In particular, although the strategy in the proposition (large ν_l and $\nu_h = 0$) is clearly optimal as it implements $\varphi_h = 1$, it is not uniquely optimal.¹⁵

Note that it is not only the costs of a given individual but also the costs of others that provide the incentive to engage in word of mouth and to acquire information about the product. For example, suppose that the firm increases the cost to the low type (ν_l). The direct effect of this is that each low-type consumer now faces a higher cost of information acquisition. However, the indirect effect is that the benefit of acquiring information is increased for the high-type consumer. As we showed in Proposition 1, an increase in information asymmetry between the two types results in an increase of informativeness of word of mouth as a signal, which increases the benefit of acquiring information. Hence, an increase in the cost to the low type results in higher rates of information acquisition by the high-type consumers. In conclusion, an increase in asymmetry in costs (and information) incentivizes consumers to acquire information in a manner that benefits the firm. A firm can manipulate this asymmetry by foregoing opportunities to decrease the costs of the low type or even increasing the costs of this type. The firm can do this by restricting how and what it communicates and where it makes available information about its product. These types of activities are hard to rationalize in the more mechanical models of word of mouth. This highlights the importance of including the motivation of consumers when analyzing a firm's optimal strategies in these environments.

Another way to interpret the asymmetry of costs is as being a characteristic of the product category. For some product categories, socially desirable characteristics may be negatively correlated with the costs of acquiring information. Our results suggest that we will observe more word of mouth in product categories where it is desirable to be perceived as knowledgeable. The observation that categories such as wine, technology, fashion, entertainment and dining satisfy this criteria in many social settings and

¹⁵This is a corner solution case. If $\bar{c} < \frac{1-k+k \ln k}{1-\alpha}$, $S^*(\varphi_h, \varphi_l) = S(1, 0)$ is feasible. From Proposition 1, we know that this is the maximum value of S^* for $0 \leq \varphi_h, \varphi_l \leq 1$. Furthermore, $\frac{\partial S^*}{\partial \varphi_l} = 0$ when $\varphi_h = 1$. Thus, the overall diffusion is not affected by φ_l as long as $\varphi_h = 1$. Therefore, any feasible pair $(1, \varphi_l)$ will also result in the same level of optimal diffusion. .

are also amongst those where there is significant word of mouth - is consistent with our result.

4 Seeding Information

In our main model, there are two ways in which consumers can acquire information about the product: (1) individual consumers can engage in a costly search for product information; (2) consumers can learn about the product through word of mouth. In practice, the firm can also seed product information to the consumer through marketing communications such as advertising or through public relations efforts that result in media coverage. We spend most of this section analyzing a simple model of advertising, which is the most common marketing action for the firm, and its effect on consumers' incentives to search for and spread information via word of mouth. We show that when the marginal cost of advertising is larger, the level of information acquisition is greater, and the overall level of diffusion of information increases. We discuss some available sources of commitment not to advertise. At the end of the section, we establish a result that any information that is distributed in a way that is independent of the type of consumers decreases the extent of information diffusion.

Advertising

In the remainder of this section, we simplify the exposition by assuming that the firm chooses ν_h and ν_l optimally such that $\nu_h = 0$ and ν_l is sufficiently large, which results in no low types and some high types acquiring information initially: $\varphi_h > 0$ and $\varphi_l = 0$.

Timing

At $t = -1$, consumers choose whether to search for information, resulting in $\varphi_h(\nu_h, \nu_l) > 0$ and $\varphi_l(\nu_h, \nu_l) = 0$. At $t = 0$, the firm exposes a fraction γ of the population to the advertising message about the product.¹⁶ From $t > 0$, consumers engage in word of mouth as before. We restrict our analysis to the firm choosing a pure strategy. Advertising is costly to the firm: $C(\gamma) \geq 0$, $C'(\gamma) > 1$, $C''(\gamma) > 0$ for all $\gamma \geq 0$. Our assumption that the marginal cost of advertising at $\gamma = 0$ is greater than 1 implies that advertising absent word of mouth is not worthwhile for the firm. We assume that consumers only observe whether they themselves receive ($a_i = 1$) or do not receive ($a_i = 0$) the ad; that is, the firm's total advertising spending (γ) is not observable to the consumers, who instead infer it in equilibrium. This extra layer of inference on the consumer side implies an additional source of complexity and multiplicity of equilibria. The word-of-mouth generation process that occurs at $t > 0$

¹⁶We model advertising as a one-time pulse for simplification purposes. Later in this section, we discuss alternative assumptions, such as a continuous advertising technology that gradually informs people over time, and alternative timing, such as the advertising occurring at or prior to $t = -1$.

is the same as in Section 3, except that consumers' inference and optimal stopping strategy is now also conditional on their beliefs about the level of advertising undertaken by the firm.

Characterizing the Word-of-Mouth Signaling Equilibrium in the Presence of Advertising

We solve for a perfect Bayesian Nash equilibrium in which (as in the earlier section) the diffusion of the information stops when the marginal value of signaling equals the marginal cost of passing on the information. But unlike the main model in Section 3, there is another source of information acquisition for consumers, advertising. Here, the level of advertising is unobserved by individuals, and consumers form beliefs about the level of advertising conditional on seeing or not seeing an advertisement. We analyze the firm's advertising decision and the consumer's choice of a time to stop spreading information as a simultaneous move game, and focus on the signaling equilibrium that results in the largest diffusion of information subject to the refinements described in Appendix A.8. The equilibrium is described by the firm's and consumers' strategies $\{\varphi_h^*, \gamma^*, \tau_i^*\}$ and their beliefs $\{\tilde{\gamma}_i(a_i), b_j^*(\theta_i = h|m, t), b_j^*(\theta_i = h|\emptyset, t)\}$. Here, φ_h^* is the fraction of high-type consumers who choose to search for information at $t = -1$, γ^* is the optimal level of advertising undertaken by the firm at $t = 0$, and τ_i^* is the equilibrium stopping time of word of mouth. The set of beliefs consists of (1) $\tilde{\gamma}_i(a_i)$, the consumers' belief on the amount of advertising undertaken by the firm, which is conditional on the consumer's personal exposure to advertising, a_i ; and (2) b_j^* , j 's belief on i 's type following a social interaction with i at time t during which i either does or does not pass on information.

First, we solve the game beginning with the consumer-to-consumer word of mouth signaling game that occurs after advertising exposure at $t = 0$.

Consumers' Word-of-Mouth Decision

When the firm engages in advertising, the inference a consumer makes from an individual passing on information depends on both the level of information acquisition $\varphi_h(\nu_h, \nu_l)$ and consumers' beliefs about the level of advertising undertaken by the firm $\tilde{\gamma}$. In particular, $\tilde{\gamma}$ affects how the signaling value of information $\Delta U(t)$ evolves over time, and hence how long consumers will continue to spread the information until $\Delta U(t) = k$. We denote the consumers' inference on the fraction of the population with information about the product due to search and advertising at $t = 0$ by \tilde{S}_0 , and the consumers' inferred level of total information diffusion by \tilde{S}_c^* . The length of time consumers engage in word of mouth $\tau(\varphi_h, \tilde{\gamma})$ is then given by

$$\tau(\varphi_h, \tilde{\gamma}) = \frac{1}{\lambda} \ln \frac{\tilde{S}_c^*}{1 - \tilde{S}_c^*} \frac{1 - \tilde{S}_0}{\tilde{S}_0} \quad (17)$$

The term \tilde{S}_0 is increasing in the conjectured level of advertising:

$$\tilde{S}_0(\varphi_h, \tilde{\gamma}) = \varphi_h \alpha + \tilde{\gamma}((1 - \varphi_h)\alpha + 1 - \alpha) \quad (18)$$

The effect of advertising is to increase the initial fraction of the population with information at $t = 0$ from $S_0 = \varphi_h \alpha$ in the basic model with no advertising (where $\varphi_l = 0$) to $\tilde{S}_0 = \varphi_h \alpha + \tilde{\gamma}((1 - \varphi_h)\alpha + 1 - \alpha)$ in the model with advertising. Note that the consumer-to-consumer signaling game through word of mouth remains the same as in the basic model, with the only difference being in the initial level of exposure prior to word of mouth diffusion. Suppose that consumers' beliefs about the level of advertising the firm undertakes $\tilde{\gamma}$ are not so large that no word of mouth takes place ($\tilde{\gamma} < \bar{\gamma} \equiv \frac{\varphi_h \alpha}{k(1 - \varphi_h \alpha)} \left(1 - k - \frac{(1 - \varphi_h)\alpha}{1 - \varphi_h \alpha}\right)$). The following lemma summarizes the characteristics of the consumers' conjectured level of total diffusion $\tilde{S}_c^*(\varphi_h, \tilde{\gamma})$.

Lemma 2. *For a given amount of initial information acquisition $\varphi_h = \varphi_h(\nu_h, \nu_l)$,*

1. *The consumers' conjectured level of information diffusion (\tilde{S}_c^*) is independent of the consumers' beliefs about the level of advertising ($\tilde{\gamma}$): $\tilde{S}_c^*(\varphi_h, \tilde{\gamma}) = \tilde{S}_c^*(\varphi_h, 0)$ for all $0 \leq \tilde{\gamma} \leq \bar{\gamma}$,*
2. *The duration of word of mouth is decreasing in the conjectured level of advertising: $\frac{d\tau}{d\tilde{\gamma}} < 0$.*

For a given level of initial information acquisition by consumers $\varphi_h(\nu_h, \nu_l)$ at $t = 0$, the consumers' conjectured level of total diffusion is independent of the consumers' beliefs about the level of advertising $\tilde{\gamma}$, provided that $\tilde{\gamma}$ is not so large that $\tilde{\gamma} < \bar{\gamma}$. The incentives to engage in word of mouth at any point in time are governed by the signaling value of the information. The signaling value is determined by the initial asymmetry in information acquisition between high and low type consumers, and the number of consumers of each type who have acquired the information subsequently through word of mouth or advertising. More importantly, for a given level of initial information acquisition through consumer search, the change in the signaling value of information is the same whether an additional individual is informed via either word of mouth or advertising. The mechanics of diffusion are the same as in both information channels individuals are informed randomly - individuals receive information irrespective of their type. Because both advertising and word of mouth reduce asymmetry between high and low types in the same way, advertising has the same diluting effect on the signaling value of information as word of mouth. Hence, advertising is a pure substitute for word of mouth, which stops at the point where $\Delta U(t) = k$. That is, conditional on a given $\varphi_h(\nu_h, \nu_l)$, the consumer's beliefs do not affect the conjectured extent of information diffusion, \tilde{S}_c^* .

Although the conjectured level of advertising does not affect the extent of information diffusion, it does affect the duration of word of mouth. The effect comes through the initial level of information $\tilde{S}_0(\varphi_h, \tilde{\gamma})$, which is increasing in $\tilde{\gamma}$. Thus, the time $\tau(\varphi_h, \tilde{\gamma})$ for information to spread from this level

to the cutoff is decreasing in \tilde{S}_0 and the conjectured level of advertising $\tilde{\gamma}$. That is, as more consumers are initially informed about the product through advertising, the duration of word of mouth decreases.

[Insert Figure 2 Here]

Figure 2 illustrates the difference in information diffusion with and without advertising conditional on a given level of information acquisition. It shows that advertising reduces the asymmetry in information spread between high and low types at $t = 0$ by randomly distributing information to more consumers. Because the asymmetry is reduced in exactly the same way as would otherwise have taken place through word of mouth, beliefs are lower at each moment in time but are the same conditional on the fraction of the population that has the information. Hence, conditional on a given level of information acquisition, advertising does not change the extent of total information diffusion (S^*), but the diffusion occurs for a shorter length of time ($t^{**} < t^*$).

Firm's Advertising Decision

At $t = 0$, the firm chooses an optimal level of advertising spending, which maximizes the firm's conjectured level of diffusion \tilde{S}_f^* . This conjectured level of diffusion depends on the firm's conjecture on the length of time consumers engage in word of mouth $\tilde{\tau}$, the fraction of high types who acquire information φ_h , and its own level of advertising:

$$\gamma(\varphi_h, \tilde{\tau}) = \arg \max_{\gamma \in [0,1]} \tilde{S}_f^*(\varphi_h, \gamma, \tilde{\tau}) - C(\gamma), \quad (19)$$

The function $\tilde{S}_f^*(\varphi_h, \gamma, \tilde{\tau})$ can be written in terms of the actual level of initial information acquisition $S_0(\varphi_h, \gamma)$ and the conjectured amount of time consumers will spread information $\tilde{\tau}$:

$$\tilde{S}_f^*(\varphi_h, \gamma, \tilde{\tau}) = \frac{1}{1 + ae^{-\lambda\tilde{\tau}}}, \quad \text{where } a = \frac{1 - S_0(\varphi_h, \gamma)}{S_0(\varphi_h, \gamma)}. \quad (20)$$

For a given $\varphi_h(\nu_h, \nu_l)$ at $t = 0$, a firm is able to influence \tilde{S}_f^* through advertising activity, which changes the informed share of the population $S_0(\varphi_h, \gamma)$. It is important to note that the actual amount of advertising γ does not affect the length of time that word of mouth occurs, because it is not directly observable to consumers. Only consumers' own expectations of the level of advertising $\tilde{\gamma}$ in equilibrium affect the length of time. Thus, advertising increases the number of individuals engaging in word of mouth for a fixed amount of time $\tilde{\tau}$.

Given φ_h (the actual level of information acquisition) and $\tilde{\tau}$ (the firm's conjectured length of time that consumers will engage in word of mouth), the firm chooses an optimal level of advertising that balances the marginal costs of advertising $C'(\gamma)$ against the marginal impact on the firm's conjectured

level of diffusion $\frac{d\tilde{S}_f^*}{d\gamma}$ from Equation (19).¹⁷ Moreover, in the case where advertising is positive ($\gamma^* > 0$), the best response function is increasing in the conjectured amount of time consumers engage in word of mouth $\frac{\partial \gamma^*}{\partial \bar{\tau}} > 0$, and satisfies the first order condition $C'(\gamma) = \frac{d\tilde{S}_f^*}{d\gamma} = \frac{d\tilde{S}_f^*}{dS_0} \frac{\partial S_0}{d\gamma}$. Here, the second equality relationship follows from the observation that advertising only affects the level of diffusion through the level of information S_0 at $t = 0$.

Consumers' Information Acquisition at $t = -1$

As before, consumers choose to acquire information if the signaling value of doing so exceeds the cost. In this section, we assume that the firm chooses ν_h and ν_l such that only high types have low enough costs to acquire information. The cutoff type φ_h^* and value of acquiring information V^* satisfy:

$$\varphi_h^* = \begin{cases} 0 & \text{if } V^* < 0 \\ \frac{V^*}{\bar{c}} & \text{if } 0 \leq V^* \leq \bar{c} \\ 1 & \text{if } V^* > \bar{c} \end{cases} \quad (21)$$

Perfect Bayesian Equilibrium

We find that the consumers' anticipation of advertising by the firm affects the total value of signaling and the incentives for individuals to acquire the information. The following proposition gives the main result of the section: the extent of diffusion is greater when the firm has larger marginal costs of advertising.

Proposition 3. *Consider two cost functions of advertising C_1 and C_2 where $C_1' < C_2'$.*

1. *When the marginal cost of advertising is larger, the level of advertising is lower and the length of diffusion is longer: $\gamma_1^*(\varphi_h) \geq \gamma_2^*(\varphi_h)$ and $t_1^*(\varphi_h) \leq t_2^*(\varphi_h)$ for any ν_h and ν_l such that $\varphi_h = \varphi_h(\nu_h, \nu_l) \leq 1$.*
2. *Moreover, the level of initial information acquisition and the extent of information diffusion are also greater when the marginal costs of advertising are larger: $\varphi_{h1}^* \leq \varphi_{h2}^*$ and $S^*(\varphi_{h1}^*) \leq S^*(\varphi_{h2}^*)$.*

The proposition shows that a higher marginal cost of advertising leads to an unambiguously greater diffusion of information for the firm. When the cost of advertising is greater, consumers anticipate that the level of advertising will be lower, and thus the length of the diffusion will be longer for a given amount of information acquired $\varphi_h(\nu_h, \nu_l)$. Although advertising does not affect the level of diffusion for a given level of $\varphi_h(\nu_h, \nu_l)$, it reduces the duration that consumers continue to spread the

¹⁷The exact condition for the firm's optimal advertising strategy is derived in Lemma 4 in Appendix A.

information. Advertising removes signaling opportunities for individuals by serving as a substitute for word of mouth and thus reduces the overall benefit to a consumer from acquiring information ex-ante. Hence, higher advertising costs result in an equilibrium with a lower level of advertising, more initial information acquisition and greater overall diffusion.

As we see from Proposition 3, a firm may benefit from higher advertising costs as these costs serve as a credible commitment to advertise less in equilibrium. For example, smaller and less established firms with smaller advertising budgets (and higher costs of capital) have higher marginal costs of advertising. Our results suggest that word of mouth will more readily occur for products of these firms than for products of more established firms.

Untargeted Exposure to Product Information

We have shown that advertising displaces consumers' incentives to engage in word of mouth and results in smaller overall diffusion. However, consumers may acquire information through a wide variety of channels, which to a greater or lesser extent may be influenced by a firm. In this section, we show that consumer search and information diffusion are decreasing in the amount of untargeted information that the population receives. The key driver behind our earlier result is not the advertising context per se, but the fact that the delivery of information is untargeted.

To generalize our finding to contexts other than advertising, we consider any untargeted exposure to product information. That is, suppose that prior to acquiring information on their own and engaging in word of mouth, a fraction γ of consumers in the population are exposed to the product information in an untargeted manner. Here, the level of γ is exogenously given and we assume it is common knowledge. One can think of this common knowledge as deriving from knowledge of the costs of advertising as in the previous section, or from knowledge of other processes of information dissemination such as media attention or public relations. The next proposition characterizes consumer search $\varphi_h^*(\gamma)$ and information diffusion $S^*(\varphi_h^*(\gamma), \gamma)$ as a function of γ .

Proposition 4. *Suppose that $\varphi_h^*(0) < 1$ and $0 \leq \gamma' < \gamma'' \leq S^*(\varphi_h^*(\gamma''), 0)$. An increase in the level of untargeted exposure of consumers to product information results in a lower level of information acquisition by the consumer, and a lower level of overall information diffusion: $\varphi_h(\gamma') > \varphi_h(\gamma'')$ and $S^*(\varphi_h^*(\gamma'), \gamma') > S^*(\varphi_h^*(\gamma''), \gamma'')$.*

In the proposition above, we show that the untargeted exposure of consumers to product information ultimately crowds out information diffusion via word of mouth more than one-for-one. That is, an increase in γ results in a decrease of both the amount of information acquired φ_h , and a decrease in the extent of information diffusion S^* . Similar to the advertising case, the intuition for the result comes from two observations. First, diffusion of information via word of mouth informs individuals in an untargeted manner, because individuals randomly mix. Thus, for a given amount of information

acquisition φ_h , the change in the signaling value of information is the same whether an additional individual is informed via either word of mouth or exogenous information γ . The extent of diffusion $S^*(\varphi_h, \gamma)$, where the signaling value is equal to the cost of passing on information $\Delta U(t) = k$, is therefore independent of γ for a given level of information acquisition φ_h . Second, although the extent of diffusion is unchanged, exogenous information removes opportunities for individuals who have acquired information to signal their type. Hence, their anticipated signaling value $V(\varphi_h, \gamma)$ of acquiring information ex-ante is lower for greater levels of exogenous information. The extent of diffusion will go down because fewer high types will acquire information, making word of mouth a weaker signal.

For the firm, the greatest level of diffusion via word of mouth can be achieved when there is no exogenous information, i.e., $\gamma = 0$. This general result implies that a firm's traditional marketing actions (such as advertising or public relations) crowd out the incentives for individuals to acquire information and engage in word of mouth when the exposure is not correlated with individual type.

Extensions

In our main analysis, we show that advertising (or more generally, any untargeted information exposure) interferes with word of mouth in such a way that it crowds out word of mouth information diffusion. Here, we discuss the robustness of this result by looking at the extent to which it is affected by the exact assumption on the timing of advertising exposure. We find that the results are qualitatively similar under a variety of assumptions. For instance, one might consider making the advertising decision occur before consumers acquire information at $t = -2$, simultaneously at the moment of information acquisition at $t = -1$, or modeling advertising as a continual process during word of mouth diffusion, $t \geq 0$. In this section, we argue that in all of these cases the qualitative nature of the effect of advertising is the same, namely, that advertising substitutes for word of mouth, thereby removing signaling opportunities and reducing the benefit of acquiring information. In fact, we argue that the timing where consumers are exposed early on in the word of mouth diffusion process may occur as the equilibrium of the game where the firm chooses the timing as well as the amount of advertising.

Of the three potential alternative specifications mentioned above, the only one that fundamentally affects the structure of the game is the case when we assume that advertising takes place prior to information acquisition at $t = -2$, because here the consumer may condition the information acquisition decision on whether or not she received information via advertising. This in turn implies that the fraction of the population with information at $t = 0$ (i.e., S_0) would change, which in turn affects the value of acquiring information $V = \frac{k}{1-\tilde{S}_0} \left[\frac{\tilde{S}^*}{\tilde{S}_0} - 1 - \ln \frac{\tilde{S}^*}{\tilde{S}_0} \right]$. The cutoff type which acquires information satisfies:

$$\varphi_h \bar{c} = \frac{k}{1-\tilde{S}_0} \left[\frac{\tilde{S}^*}{\tilde{S}_0} - 1 - \ln \frac{\tilde{S}^*}{\tilde{S}_0} \right] \quad (22)$$

where \tilde{S}_0 and \tilde{S}^* are the same expressions as earlier. The term $V = \frac{k}{1-\tilde{S}_0} \left[\frac{\tilde{S}^*}{\tilde{S}_0} - 1 - \ln \frac{\tilde{S}^*}{\tilde{S}_0} \right]$ is decreasing in the conjectured level of advertising ($\frac{\partial V}{\partial \tilde{\gamma}} = \frac{\partial V}{\partial \tilde{S}_0} \frac{\partial \tilde{S}_0}{\partial \tilde{\gamma}} + \frac{\partial V}{\partial \tilde{S}^*} \frac{\partial \tilde{S}^*}{\partial \tilde{\gamma}} < 0$ because $\frac{\partial V}{\partial \tilde{S}_0} < 0$, from Lemma 3 in the proof of Proposition 2, $\frac{\partial \tilde{S}_0}{\partial \tilde{\gamma}} > 0$, $\frac{\partial \tilde{S}^*}{\partial \tilde{\gamma}} = 0$). Hence, the qualitative effect of increasing the costs of advertising is the same: larger costs lead consumers to conjecture that the firm has undertaken less advertising, and this increases the value of acquiring the information and leads more high types to acquire it.

In the cases where advertising occurs at the same time as information acquisition at $t = -1$ or at the same time as word of mouth diffusion, the effect is the same. The extent of information acquisition and diffusion in equilibrium are determined by the consumers' belief about the amount of advertising that the firm undertakes. In both cases, this belief is smaller when the costs of advertising are higher. More specifically, if we assume that advertising occurs simultaneously with information acquisition at $t = -1$, the analysis is unchanged. Also, there are a number of ways one might model a firm that advertises gradually over time. However, provided that the model generates the prediction that lower costs of advertising lead to greater advertising at each moment of time during the diffusion of information, that will reduce the value of acquiring information. Hence, we do not expect that the basic results will change in this setting.

Next, we consider the incentives of a firm choosing the timing for a one-time advertising campaign. As we saw in Equation (20), the benefit of advertising to a firm is that it increases the mass of consumers who are passing on information until a fixed time. The length of time is not directly affected by the actual level of advertising, because it is unobservable. Instead, the length of time is determined by consumers' equilibrium conjecture. If a firm also chooses the timing of advertising, it would want to advertise as early as possible to allow the additional mass of individuals informed by advertising to spread information for as long as possible.¹⁸ Hence, in an equilibrium, consumers will infer that a firm will advertise at $t = 0$, and the analysis will be otherwise unchanged. Hence, even if the firm prefers to commit not to advertise early on, it may have trouble doing so credibly.

However, in the presence of time discounting, the firm may be able to credibly commit to delaying advertising by separating the release of information about the product (and the ensuing word of mouth) and the actual product release. For instance, suppose that the firm releases product information far ahead of the announced product release date. If we define T as the amount of time between the information release and the product release, the monopolist can choose a large enough T to credibly commit not to advertise and maximize the diffusion of information due to word of mouth. Of course, the separation in time is an issue only in the presence of time discounting.

Proposition 5. *In the presence of time discounting, there exists a \bar{T} such that for $T > \bar{T}$ the firm*

¹⁸In developing the model along these lines, one would also need to use a similar type of trembling refinement for consumer beliefs following the off-equilibrium timing choices by the firm.

undertakes no advertising, and the maximum possible diffusion occurs.

Due to time discounting (and because the firm does not earn profits until a future date), an early release of information lowers the benefit the firm derives from current period advertising. The result is consistent with the observation that delaying advertising may promote word-of-mouth generation. For example, in an influential Harvard Business Review article, Dye (2000) recommends that “while the media and advertising can help fan the flames of buzz, involving them too early can help undermine buzz. Indeed, the vanguard will often reject a heavily promoted product merely because of overexposure.” Also consistent with this logic, Spotify started its mass advertising in March 2013, only after word of mouth had sufficiently spread.

5 Discussion of Alternative Approaches

In this section, we discuss several alternative specifications of the main model regarding the two different costs involved in communication – the costs of information acquisition and transmission.

First, in our model, the consumer bears a fixed cost of k of information transmission, which plays an important role for determining the stopping rule of word of mouth diffusion – the diffusion of information stops when the difference in beliefs is no longer larger than this positive cost. However, one can consider an alternative modeling specification without imposing a positive cost of communication k , which is not isomorphic to ours, that instead utilizes a threshold for the posterior belief itself. For example, consider a setting where the focal consumer derives utility from whether the other consumer will take a particular action in the future; for instance, invites them to dinner parties, functions, etc. The receiver will only do so if his belief is above an exogenous threshold (given by the relative benefit of having the high or low type guest at the dinner party). In this setting, a consumer will only derive a benefit from signaling if she can increase the receiver’s posterior inference above the threshold. Here the beliefs only matter to the extent to which they drive actions. That is, unlike our model, the focal consumer’s utility is not continuous in the other consumer’s belief. In this scenario, an individual would only choose to engage in word of mouth when the signaling benefit is large enough (i.e., the posterior belief is above a threshold that triggers a change in the receiver’s behavior). Although this is an equally natural mechanism, it does not add any additional insight to our results. Our simpler assumption of fixed cost of communication captures the idea that diffusion of information halts when the signal associated with word of mouth is weak enough.

Next, we clarify the separate roles that the two costs play in our main model by considering another modeling specification where the cost structure is reversed: a specification with a uniform cost of acquisition and a type-dependent cost of communication. In our main model, the heterogeneous costs of acquisition drive differences in information acquisition across types, and the uniform communication cost implies that both types stop generating word of mouth at the same time. Clearly, having the

type-dependent cost of communication complicates equilibrium construction in the word-of-mouth-generation period of the game. That is, given asymmetric incentives to stop, we would have a period where only one type would choose to communicate, which may lead to either a separation (only one type talks in equilibrium) or technical issues in a pooling equilibrium (only one type has an incentive to talk in the “silent” period; we demonstrate that this issue could destroy the equilibrium in more detail using a model in the Appendix B). In contrast, in our main model with identical communication costs, the incentives to generate word of mouth and to stop generating word of mouth are identical for both types which lead to a clean stopping rule and no issues in the “silent” period after word of mouth stops. . A model with differential communication costs may yield further insights; however, it is not necessary for our main results about how product information - how it was acquired, who else also knows the information - influences the signaling motive and the extent of diffusion of information in the population. Our primary research objective is to explain the phenomenon of firms strategically creating the informational asymmetry between different types of potential consumers as exemplified by Spotify intentionally making it difficult for some potential consumers to find out about its service. This is why the focus is on differences in information acquisition costs which are under the firm’s control. In contrast, it appears less realistic to allow the firm to affect communication costs, which are outside of the firm’s control. To summarize, reversing which cost is uniform and which differs across types has some technical issues and qualitatively alters the firm’s action space in a way that we think is less realistic.

6 Conclusion

In this article, we study how a firm’s communication strategy (which consists of information release and marketing actions such as advertising) is shaped by the incentives of individuals to acquire information and engage in word of mouth. We develop a model in which a firm’s objective is to maximize the diffusion of information about its product, and consumers are motivated to engage in word of mouth by self-enhancement. A firm maximizes the diffusion of information by structuring its information release strategy such that the act of acquiring and passing on information through word of mouth serves as a signal of a consumer’s type. In our model, the firm chooses to optimally restrict access to information to low social type consumers in order to stimulate word of mouth. Even though these activities seem to restrict the spread of information in the immediate term, they in fact serve to maximize the total diffusion of information. The firm also benefits from a commitment not to undertake advertising, which serves to crowd out word of mouth as a source of information. We highlight that a potential source of this commitment is to coordinate the information release a sufficient amount of time prior to the product release.

A key lesson from our study is that word of mouth is a subtle process for the firm to influence.

In our model, consumers only engage in word of mouth if it can serve as a signal; therefore, if low social types have similar access to the information as high types, little word of mouth will ensue. The existence of asymmetries in information between the two types enables word of mouth to occur. This emphasizes the challenge a firm faces in harnessing the power of word of mouth. Beyond simply getting information into the hands of particular individuals, who may engage in word of mouth, it must do so in such a way that the information may serve as a signal. Thus, a firm's strategy to stimulate word of mouth through information release and advertising is as much about who does not have information as it is about who does.

A Proofs

A.1 Proof of Proposition 1

First, we replace $b_j(\theta_i = h|m, 0)$ and $b_j(\theta_i = h|\emptyset, 0)$ and express the total extent of diffusion in terms of φ_h, φ_l ($\varphi_h > \varphi_l$):

$$S^*(\varphi_h, \varphi_l) = \frac{\alpha}{k} \left[\varphi_h - (1 - \varphi_h) \frac{\varphi_h \alpha + \varphi_l (1 - \alpha)}{1 - \varphi_h \alpha - \varphi_l (1 - \alpha)} \right] \quad (23)$$

The derivatives of S^* with respect to φ_h, φ_l are

$$\begin{aligned} \frac{\partial S^*}{\partial \varphi_h} &= \frac{1}{k} \left[\frac{(1 - \varphi_l)(1 - \alpha)}{((1 - \varphi_h)\alpha + (1 - \varphi_l)(1 - \alpha))^2} \right] \geq 0, \text{ if } \varphi_l < 1 \text{ then } > 0 \\ \frac{\partial S^*}{\partial \varphi_l} &= -\frac{\alpha}{k} \frac{(1 - \varphi_h)(1 - \alpha)}{((1 - \varphi_h)\alpha + (1 - \varphi_l)(1 - \alpha))^2} \leq 0, \text{ if } \varphi_h < 1 \text{ then } < 0 \end{aligned} \quad (24)$$

A.2 Proof of Lemma 1

Under symmetric costs, $\nu_h = \nu_l$, it is immediate that $\varphi_h = \varphi_l$ because c_i for both types is from the same i.i.d. uniform $[0, \bar{c}]$. Then, $b_j(\theta_i = h|m, 0, \varphi) = \frac{\varphi_h \alpha}{\varphi_h \alpha + \varphi_l (1 - \alpha)} = \alpha$, $b_j(\theta_i = h|\emptyset, 0, \varphi) = \frac{(1 - \varphi_h) \alpha}{(1 - \varphi_h) \alpha + (1 - \varphi_l)(1 - \alpha)} = \alpha$. Hence, we have $b_j(\theta_i = h|m, 0, \varphi) - b_j(\theta_i = h|\emptyset, 0, \varphi) = 0$, which implies that $S^* = 0$. Therefore, no word of mouth arises.

A.3 Proof of Proposition 2

The firm's optimization problem is equivalent to maximizing the spread of information through the choice of φ_l and φ_h :

$$\max_{\varphi_h, \varphi_l} S^*(\varphi_h, \varphi_l)$$

subject to the feasibility constraints:

$$\begin{aligned} \varphi_h \bar{c} &\leq V(\varphi_h, \varphi_l) \\ \varphi_l \bar{c} &\leq V(\varphi_h, \varphi_l) \\ 0 &\leq \varphi_h \leq 1 \\ 0 &\leq \varphi_l \leq 1 \end{aligned}$$

First, we consider the case where $\bar{c} < \frac{1-k+k \ln k}{1-\alpha}$. Note that $V(1, 0) = \frac{1-k+k \ln k}{1-\alpha}$. Hence, if $\bar{c} < \frac{1-k+k \ln k}{1-\alpha}$, then $S^*(1, 0)$ is feasible. From Proposition 1, we know that this is the maximum value of S^* for $0 \leq \varphi_h, \varphi_l \leq 1$. Furthermore, $\frac{\partial S^*}{\partial \varphi_l} = 0$ when $\varphi_h = 1$. Therefore, any feasible pair $(1, \varphi_l)$ will also result in the same level of diffusion.

Next, we consider the case where $\bar{c} \geq \frac{1-k+k \ln k}{1-\alpha}$. In this case, $S^*(1, 0)$ is not feasible. Proposition 1 showed that S^* is increasing in φ_h when $\varphi_l < 1$ and decreasing in φ_l if $\varphi_h < 1$. It is immediate that, in the optimum, the firm will choose the maximum feasible φ_h $\varphi_h^* \bar{c} = V(\varphi_h, \varphi_l)$ or $\varphi_h^* = 1$. To conclude that the firm will minimize φ_l , we need to show that increasing φ_l does not relax the feasibility constraint $\frac{\partial V}{\partial \varphi_l} \leq 0$. We proceed with the following lemma.

Lemma 3. *The total signaling benefit is increasing in S^* and decreasing in S_0 : $\frac{\partial V}{\partial S^*} > 0$ and $\frac{\partial V}{\partial S_0} < 0$.*

Proof. The derivatives of V with respect to S^* and S_0 are

$$\begin{aligned} \frac{\partial V}{\partial S^*} &= \left(\frac{k}{1-S_0} \right) \left[\frac{1}{S_0} - \frac{1}{S^*} \right], \\ \frac{\partial V}{\partial S_0} &= \left[k \left(\frac{1}{1-S_0} \right)^2 \left(\frac{S^*}{S_0} - 1 + \ln S_0 - \ln S^* \right) + \left(\frac{k}{1-S_0} \right) \left(-\frac{S^*}{S_0^2} + \frac{1}{S_0} \right) \right] \\ &= \left[\frac{k}{(1-S_0)^2 S_0^2} \left\{ (S^* - S_0)(2S_0 - 1) + S_0^2 \left(\ln \frac{S_0}{S^*} \right) \right\} \right]. \end{aligned}$$

We have that $\left(\frac{k}{1-S_0} \right) \left[\frac{1}{S_0} - \frac{1}{S^*} \right] > 0$. Hence, $\frac{\partial V}{\partial S^*} > 0$.

Next, we show that $\frac{\partial V}{\partial S_0} < 0$. Note that it is immediate that $\frac{\partial V}{\partial S_0} < 0$ when $S_0 < \frac{1}{2}$. In the case that $S_0 \geq \frac{1}{2}$, we need to show that

$$\begin{aligned} (S^* - S_0)(2S_0 - 1) + S_0^2 \left(\ln \frac{S_0}{S^*} \right) < 0 &\Leftrightarrow S_0^2 \left(\ln \frac{S_0}{S^*} \right) < (S^* - S_0)(1 - 2S_0) \\ &\Leftrightarrow \frac{\ln \frac{S_0}{S^*}}{\frac{S^*}{S_0} - 1} < \left(\frac{1}{S_0} - 2 \right). \end{aligned}$$

Now consider the left-hand side, where $x = \frac{S^*}{S_0} > 1$:

$$\frac{\ln \frac{1}{x}}{x-1} = \frac{-\ln x}{x-1} \Leftrightarrow \frac{d\left(\frac{-\ln x}{x-1}\right)}{dx} = \frac{1}{x-1} \left(\frac{\ln x}{x-1} - \frac{1}{x} \right)$$

This expression is greater than 0 for $x > 1$. To see this, note that $\frac{\ln x}{x-1} - \frac{1}{x} > 0 \Leftrightarrow \ln x > 1 - \frac{1}{x}$, which is a known relation for the natural log.

Hence, the left-hand side of the above is increasing in $\frac{S^*}{S_0}$, and an upper bound on the left-hand side is given by $-\frac{\ln \frac{1}{S_0}}{\frac{1}{S_0} - 1}$, and we need only check that

$$-\frac{\ln \frac{1}{S_0}}{\frac{1}{S_0} - 1} < \frac{1}{S_0} - 2 \Leftrightarrow -\ln y - (y-2)(y-1) < 0, \text{ where } y = \frac{1}{S_0}.$$

Now we show that it is a decreasing function of y for $1 \leq y \leq 2$ ($\leftrightarrow \frac{1}{2} \leq S_0 \leq 1$):

$$\begin{aligned} \frac{d(-\ln y - (y-2)(y-1))}{dy} &= -\frac{1}{y} - 2y + 3 = \frac{-2y^2 + 3y - 1}{y} \\ &= \frac{(1-2y)(y-1)}{y} < 0 \text{ for } 1 \leq y \leq 2 \end{aligned}$$

and note that $\lim_{y \rightarrow 1} [-\ln y - (y-2)(y-1)] = 0$. Hence, $-\ln y - (y-2)(y-1) < 0$, which shows that $\frac{\partial V}{\partial S_0} < 0$. \square

Next, consider V as a function of S_0 and S^* , $V(S^*, S_0) = \left(\frac{k}{1-S_0}\right) \left(\frac{S^*-S_0}{S_0} + \ln \frac{S_0}{S^*}\right)$. Taking the derivative $\frac{dV}{d\varphi_l}$,

$$\frac{dV}{d\varphi_l} = \frac{\partial V}{\partial S_0} \frac{dS_0}{d\varphi_l} + \frac{\partial V}{\partial S^*} \frac{dS^*}{d\varphi_l}.$$

From Lemma 3, $\frac{\partial V}{\partial S_0} < 0$, $\frac{\partial V}{\partial S^*} > 0$, and note that $\frac{\partial S_0}{\partial \varphi_l} > 0$, and when $\varphi_h < 1$ $\frac{dS^*}{d\varphi_l} \leq 0$ (from Proposition 1). Hence,

$$\frac{dV}{d\varphi_l} = \underbrace{\frac{\partial V}{\partial S_0} \frac{dS_0}{d\varphi_l}}_{-} + \underbrace{\frac{\partial V}{\partial S^*} \frac{dS^*}{d\varphi_l}}_{+} < 0.$$

We conclude that $\varphi_l = 0$ when $\varphi_h < 1$ because φ_l reduces the objective and tightens the feasibility constraint for φ_h .

Finally, we verify that there exists a $\varphi_h^* > 0$ such that $\varphi_h^* \bar{c} = V(\varphi_h^*, 0)$. When $\varphi_l^* = 0$;

$$\begin{aligned} S_0(\varphi_h^*, 0) &= \varphi_h^* \alpha \\ S^*(\varphi_h^*, 0) &= \frac{\alpha}{k} \left[\frac{\varphi_h^* (1-\alpha)}{1-\varphi_h^* \alpha} \right] \end{aligned}$$

$$\begin{aligned} V(\varphi_h^*, 0) &= \left(\frac{k}{1-S_0}\right) \left(\frac{S^*-S_0}{S_0} + \ln \frac{S_0}{S^*}\right) = \left(\frac{k}{1-\varphi_h^* \alpha}\right) \left(\frac{\frac{1}{k} \left[\frac{\varphi_h^* (1-\alpha)}{1-\varphi_h^* \alpha}\right] - \varphi_h^*}{\varphi_h^*} + \ln \frac{\varphi_h^*}{\frac{1}{k} \left[\frac{\varphi_h^* (1-\alpha)}{1-\varphi_h^* \alpha}\right]}\right) \\ &= \frac{1-\alpha}{(1-\varphi_h^* \alpha)^2} - \left(\frac{k}{1-\varphi_h^* \alpha}\right) \left(1 - \ln \frac{k(1-\varphi_h^* \alpha)}{1-\alpha}\right) \end{aligned} \quad (25)$$

We note $\lim_{\varphi_h \rightarrow 0} V(\varphi_h, 0) = k \left(\frac{1-\alpha}{k} - 1 - \ln \frac{1-\alpha}{k}\right) > 0$, and hence there exists $0 < \varphi_h^* < 1$ such that $\varphi_h^* \bar{c} = V(\varphi_h^*, 0)$. Therefore, the optimal strategy for the firm has the following characteristics: $\varphi_h < 1$ and $\varphi_l = 0$ if $\bar{c} \geq \frac{1-k+k \ln k}{1-\alpha}$; $\varphi_h = 1$ if $\bar{c} < \frac{1-k+k \ln k}{1-\alpha}$.

A.4 Proof of Lemma 2

First, let $\bar{\gamma} = \frac{\varphi_h \alpha}{k(1-\varphi_h \alpha)} \left(1 - k - \frac{(1-\varphi_h)\alpha}{1-\varphi_h \alpha} \right)$. The relationship for the extent of diffusion is determined by

$$\tilde{S}_c^*(\varphi_h, \tilde{\gamma}) = \frac{S_0}{k} [b_j(\theta_i = h|m, 0, \varphi_h, \tilde{\gamma}) - b_j(\theta_i = h|\emptyset, 0, \varphi_h, \tilde{\gamma})],$$

where $b_j(\theta_i = h|m, 0, \varphi_h, \tilde{\gamma}) = \frac{\varphi_h \alpha + \tilde{\gamma}(1-\varphi_h)\alpha}{\varphi_h \alpha + \tilde{\gamma}(1-\varphi_h \alpha)}$ and $b_j(\theta_i = h|\emptyset, 0, \varphi_h, \tilde{\gamma}) = \frac{\alpha(1-\varphi_h)}{1-\varphi_h \alpha}$. When $0 \leq \tilde{\gamma} \leq \bar{\gamma}$, then $b_j(\theta_i = h|m, 0, \varphi_h, \tilde{\gamma}) - b_j(\theta_i = h|\emptyset, 0, \varphi_h, \tilde{\gamma}) \geq k$, and there are incentives to engage in word of mouth at $t = 0$. We find that

$$\begin{aligned} \tilde{S}_c^*(\varphi_h, \tilde{\gamma}) &= \frac{\varphi_h \alpha + \tilde{\gamma}(1-\varphi_h \alpha)}{k} \left[\frac{\varphi_h \alpha + \tilde{\gamma}(1-\varphi_h)\alpha}{\varphi_h \alpha + \tilde{\gamma}(1-\varphi_h \alpha)} - \frac{\alpha(1-\varphi_h)}{1-\varphi_h \alpha} \right] \\ &= \frac{1}{k} \left[\varphi_h \alpha + \tilde{\gamma}(1-\varphi_h)\alpha - \frac{\alpha^2 \varphi_h (1-\varphi_h)}{1-\varphi_h \alpha} - \tilde{\gamma}(1-\varphi_h)\alpha \right] \\ &= \frac{1}{k} \left[\varphi_h \alpha - \frac{\alpha^2 \varphi_h (1-\varphi_h)}{1-\varphi_h \alpha} \right] = \frac{\varphi_h \alpha}{k} \left[1 - \frac{\alpha(1-\varphi_h)}{1-\varphi_h \alpha} \right] \\ &= \tilde{S}_c^*(\varphi_h, 0) \end{aligned}$$

Second, when $0 \leq \tilde{\gamma} \leq \bar{\gamma}$ the best response time $\tau(\varphi_h, \tilde{\gamma})$ is given by

$$\tau(\varphi_h, \tilde{\gamma}) = \frac{1}{\lambda} \ln \frac{\tilde{S}_c^*(\varphi_h^*, \tilde{\gamma})}{1 - \tilde{S}_c^*(\varphi_h^*, \tilde{\gamma})} \frac{1 - S_0(\varphi_h, \tilde{\gamma})}{S_0(\varphi_h, \tilde{\gamma})}$$

where $\tilde{S}_0(\varphi_h, \tilde{\gamma}) = \varphi_h \alpha + \tilde{\gamma}((1-\varphi_h)\alpha + 1 - \alpha)$. And we know from above that $\frac{d\tilde{S}_c^*}{d\tilde{\gamma}} = 0$. It is now straightforward to find that the derivative is $\frac{d\tau^*}{d\tilde{\gamma}} = \frac{d\tau^*}{d\tilde{S}_0} \frac{d\tilde{S}_0}{d\tilde{\gamma}}$:

$$\frac{d\tau^*}{d\tilde{\gamma}} = -\frac{1-\varphi_h \alpha}{\lambda} \left[\frac{1}{1-\tilde{S}_0} + \frac{1}{\tilde{S}_0} \right] < 0$$

A.5 Proof of Proposition 3

It is useful to define the equilibrium level of information acquisition without advertising $\varphi_h^{**} = \varphi_h(\nu_h, \nu_l)$ for a given ν_h and ν_l such that only high type consumers acquire information $\varphi_h > 0$ and $\varphi_l = 0$. (from Equation 25). It satisfies the following condition:

$$\varphi_h^{**} \bar{c} = \frac{1-\alpha}{(1-\varphi_h^{**} \alpha)^2} - \left(\frac{k}{1-\varphi_h^{**} \alpha} \right) \left(1 - \ln \frac{k(1-\varphi_h^{**} \alpha)}{1-\alpha} \right).$$

We begin with the following lemmas:

Lemma 4. *When the marginal cost of advertising $C'(0)$ is not too large,¹⁹ there exists a cutoff $\hat{\tau}$ such*

¹⁹If $C'(0)$ is too large, the firm never advertises such that $\gamma = 0$ for all $\tilde{\tau} \geq 0$.

that for all $0 \leq \tilde{\tau} \leq \hat{\tau}$, the optimal level of advertising is $\gamma^* = 0$, and for all $\tilde{\tau} \geq \hat{\tau}$, the firm's optimal choice of advertising is $\gamma^* > 0$.

Proof. The firm's choice of advertising is determined by:

$$\gamma(\varphi_h, \tilde{\tau}) = \arg \max_{\gamma \in [0,1]} \tilde{S}_f^*(\varphi_h, \gamma, \tilde{\tau}) - C(\gamma)$$

A firm is able to influence \tilde{S}_f^* through the initial informed share of the population $S_0(\varphi_h, \gamma)$ but cannot directly influence $\tilde{\tau}$, which is only affected by consumers' expectations of the level of advertising $\tilde{\gamma}$ in equilibrium. Hence, the marginal effect of advertising on \tilde{S}_f^* is $\frac{d\tilde{S}_f^*}{d\gamma} = \frac{d\tilde{S}_f^*}{\partial S_0} \frac{\partial S_0}{d\gamma} = \frac{e^{-\lambda\tilde{\tau}}(1-\varphi_h\alpha)}{[S_0+(1-S_0)e^{-\lambda\tilde{\tau}}]^2}$. Also note that $\frac{d^2\tilde{S}_f^*}{d\gamma^2} = -2\frac{e^{-\lambda\tilde{\tau}}(1-e^{-\lambda\tilde{\tau}})(1-\varphi_h\alpha)^2}{[S_0+(1-S_0)e^{-\lambda\tilde{\tau}}]^3} < 0$, and by assumption $C''(\gamma) > 0$, so that the objective is strictly concave, and a first-order condition can be used for interior solutions for $\gamma \in (0, 1)$. Thus, if there is a γ that satisfies equation

$$C'(\gamma) = \frac{d\tilde{S}_f^*}{d\gamma} = \frac{d\tilde{S}_f^*}{\partial S_0} \frac{\partial S_0}{d\gamma} = \frac{e^{-\lambda\tilde{\tau}}(1-\varphi_h\alpha)}{[S_0+(1-S_0)e^{-\lambda\tilde{\tau}}]^2}, \quad (26)$$

then this is a best response for a given $\tilde{\tau}$. Also note that $C'(1) > 1 > e^{-\lambda\tilde{\tau}}(1-\varphi_h\alpha)$ such that $\gamma = 1$ is never a best response. We now show that there exists a threshold amount of time $\bar{\tau}$ such that we can find an interior solution $\gamma > 0$ that satisfies $C'(\gamma) = \frac{e^{-\lambda\tilde{\tau}}(1-\varphi_h\alpha)}{[S_0+(1-S_0)e^{-\lambda\tilde{\tau}}]^2}$ for $\tau \geq \bar{\tau}$ and the corner solution $C'(0) \geq \frac{e^{-\lambda\tilde{\tau}}(1-\varphi_h\alpha)}{[\varphi_h\alpha+(1-\varphi_h\alpha)e^{-\lambda\tilde{\tau}}]^2}$ for $\tau \leq \bar{\tau}$. We define $\bar{\tau}$ as the time that satisfies $C'(0) = \frac{e^{-\lambda\tilde{\tau}}(1-\varphi_h\alpha)}{[\varphi_h\alpha+(1-\varphi_h\alpha)e^{-\lambda\tilde{\tau}}]^2}$. At $\tilde{\tau} = 0$, the (RHS) is $1 - \varphi_h\alpha \leq 1$, and furthermore, the (RHS) is concave and maximized at $\tilde{\tau} = \frac{1}{\lambda} \ln \frac{1-\varphi_h\alpha}{\varphi_h\alpha}$ at a value of $\frac{1}{2\varphi_h\alpha}$. Therefore, if $C'(0) < \frac{1}{2\varphi_h\alpha}$, then there exists a cutoff $\bar{\tau} \leq \frac{1}{\lambda} \ln \frac{1-\varphi_h\alpha}{\varphi_h\alpha}$ such that for all $\tilde{\tau} \in [0, \bar{\tau})$, $C'(0) > \frac{e^{-\lambda\tilde{\tau}}(1-\varphi_h\alpha)}{[\varphi_h\alpha+(1-\varphi_h\alpha)e^{-\lambda\tilde{\tau}}]^2}$, which implies $\gamma = 0$. Furthermore, for all $\tilde{\tau} \in [\bar{\tau}, \frac{1}{\lambda} \ln \frac{1-\varphi_h\alpha}{\varphi_h\alpha}]$, $C'(0) \leq \frac{e^{-\lambda\tilde{\tau}}(1-\varphi_h\alpha)}{[\varphi_h\alpha+(1-\varphi_h\alpha)e^{-\lambda\tilde{\tau}}]^2}$ and hence, $\exists \gamma > 0$ which satisfies $C'(\gamma) = \frac{e^{-\lambda\tilde{\tau}}(1-\varphi_h\alpha)}{[S_0+(1-S_0)e^{-\lambda\tilde{\tau}}]^2}$. Otherwise, (i.e., $C'(0) \geq \frac{1}{2\varphi_h\alpha}$), $C'(0) \geq \frac{e^{-\lambda\tilde{\tau}}(1-\varphi_h\alpha)}{[\varphi_h\alpha+(1-\varphi_h\alpha)e^{-\lambda\tilde{\tau}}]^2}$ and $\gamma = 0$ for all $\tilde{\tau}$. \square

We now show that there is a unique pair of pure strategies $\{\gamma^*(\varphi_h), \tau^*(\varphi_h)\}$ that satisfy the incentive constraint of the firm and the consumers' word of mouth constraint for the continuation game from time $t \geq 0$.

Lemma 5. *For a given level of information acquisition $\varphi_h > 0$, there is a unique pair of pure strategies $\{\gamma^*(\varphi_h), \tau^*(\varphi_h)\}$ that satisfy the incentive constraint of the firm and the consumers' word of mouth constraint for the continuation game from time $t \geq 0$.²⁰*

²⁰On the equilibrium path, beliefs are pinned down by Bayesian beliefs; however, any beliefs such that $b_j^*(\theta_i = h|\emptyset, t, \varphi_h, \gamma) = \alpha$ and $b_j^*(\theta_i = h|m, t, \varphi_h, \gamma) - b_j^*(\theta_i = h|\emptyset, t, \varphi_h, \gamma) \leq k$ for $t > \tau^*$ are possible.

Proof. By definition, γ is bounded above by 1, and an upper bound on τ for a given φ_h is $\tau(\varphi_h, 0) = \frac{1}{\lambda} \ln \frac{1-\varphi_h\alpha}{\frac{k}{1-\alpha}-\varphi_h\alpha}$. The best response functions $\{\tau(\varphi_h, \tilde{\gamma}), \gamma(\varphi_h, \tilde{\tau})\}$ are continuous and map a compact set to a compact set, therefore there exists a fixed point and hence an equilibrium in pure strategies for any φ_h in the continuation game from $t \geq 0$.

We note from Lemma 2 that the length of time consumers engage in word of mouth is decreasing in the conjectured level of advertising by the firm $\frac{d\tau}{d\tilde{\gamma}} < 0$ for $0 \leq \tilde{\gamma} \leq \frac{\varphi_h\alpha}{k(1-\varphi_h\alpha)} \left(1 - k - \frac{(1-\varphi_h)\alpha}{1-\varphi_h\alpha}\right)$ and constant at zero $\tau = 0$ for $\tilde{\gamma} \geq \frac{\varphi_h\alpha}{k(1-\varphi_h\alpha)} \left(1 - k - \frac{(1-\varphi_h)\alpha}{1-\varphi_h\alpha}\right)$. Note that $\gamma(\varphi_h, 0) = 0$ so that at the point of intersection the diffusion time $\tau^* > 0$ and $\frac{d\tau^*}{d\tilde{\gamma}} < 0$.

If $C'(0) \geq \frac{1}{2\varphi_h\alpha}$, then using Lemma 4 $\gamma = 0$ for all $\tilde{\tau}$. Hence, the unique equilibrium is $\{\tau^*, \gamma^*\} = \left\{ \frac{1}{\lambda} \ln \frac{1-\varphi_h\alpha}{\frac{k}{1-\alpha}-\varphi_h\alpha}, 0 \right\}$.

We now show that for $C'(0) < \frac{1}{2\varphi_h\alpha}$ the slope $\frac{d\gamma}{d\tilde{\tau}} \geq 0$ at any point of intersection where $\tau^* > 0$ and $\gamma^* > 0$.

Implicitly differentiating Equation (26), we find $\frac{d\gamma}{d\tilde{\tau}}$ for $\gamma \in (0, 1)$:

$$\begin{aligned} \frac{d\gamma}{d\tilde{\tau}} &= \frac{\lambda(1-\varphi_h\alpha)e^{\lambda\tilde{\tau}}}{[1-(1-e^{\lambda\tilde{\tau}})S_0]^2} - 2S_0\lambda e^{\lambda\tilde{\tau}} \frac{(1-\varphi_h\alpha)e^{\lambda\tilde{\tau}}}{[1-(1-e^{\lambda\tilde{\tau}})S_0]^3} = \lambda \frac{1 - S_0 - e^{\lambda\tilde{\tau}}S_0}{\frac{[1-(1-e^{\lambda\tilde{\tau}})S_0]^3}{e^{\lambda\tilde{\tau}}(1-\varphi_h\alpha)} C''(\gamma) - 2(1-e^{\lambda\tilde{\tau}})(1-\varphi_h\alpha)} \\ &= \lambda \frac{(1-S_0) \left(1 - \frac{S^*}{1-S^*}\right)}{\frac{\left[\frac{1-S_0}{1-\tilde{S}_f^*}\right]^3}{(1-\varphi_h\alpha) \frac{\tilde{S}_f^*(1-S_0)}{S_0(1-S^*)}} C''(\gamma) + 2 \frac{(1-\varphi_h\alpha)(\tilde{S}_f^*-S_0)}{S_0(1-S^*)}} = \frac{\lambda(1-\varphi_h\alpha) \left(1 - 2\tilde{S}_f^*\right)}{\frac{S_0}{\tilde{S}_f^*} \left(\frac{1-S_0}{1-\tilde{S}_f^*}\right) C''(\gamma) + 2 \left(\frac{\tilde{S}_f^*}{S_0} - 1\right) \frac{(1-\varphi_h\alpha)^2}{1-S_0}} \end{aligned}$$

Hence, a sufficient condition for $\frac{d\gamma}{d\tilde{\tau}} > 0$ is $\tilde{S}_f^* < \frac{1}{2}$. At an equilibrium point,

$$\begin{aligned} \tilde{S}_f^*(\varphi_h, \tau^*, \gamma^*) &= \tilde{S}_c^*(\varphi_h, \gamma^*) = \tilde{S}_c^*(\varphi_h, 0) = S^*(\varphi_h) = \frac{\alpha}{k} \left[\frac{\varphi_h(1-\alpha)}{1-\varphi_h\alpha} \right] < \frac{1}{2} \\ \Leftrightarrow k &> \frac{2\varphi_h\alpha(1-\alpha)}{1-\varphi_h\alpha} = G(\varphi_h|\alpha) \end{aligned}$$

We note that $\frac{\partial G(\varphi_h)}{\partial \varphi_h} = \frac{2\alpha(1-\alpha)}{(1-\varphi_h\alpha)^2} > 0$ for all $\alpha \in (0, 1)$ and thus, $S^* < \frac{1}{2}$ is true when $k > 2\alpha = G(1)$. Moreover, the equilibrium is continuous in φ_h because $\tau(\varphi_h, \tilde{\gamma})$ and $\gamma(\varphi_h, \tilde{\tau})$ are also continuous in φ_h for $\varphi_h > 0$. \square

The following lemma proves the first part of Proposition 3 by providing the comparative statics of the cost of advertising on the continuation game from $t \geq 0$.

Lemma 6. *For a given level of information acquisition φ_h , consider two cost functions of advertising C_1 and C_2 where $C_1' < C_2'$. When the marginal cost of advertising is larger, the level of advertising is lower and the length of diffusion is longer: $\gamma_1^*(\varphi_h) \geq \gamma_2^*(\varphi_h)$ and $t_1^*(\varphi_h) \leq t_2^*(\varphi_h)$.*

Proof. Let $F(\gamma) = \frac{e^{-\lambda\tilde{\tau}}(1-\varphi_h\alpha)}{[S_0+(1-S_0)e^{-\lambda\tilde{\tau}}]^2} \cdot \frac{\partial F(\gamma)}{\partial \gamma} = 2(1 - e^{\lambda\tilde{\tau}}) \frac{(1-\varphi_h\alpha)^2 e^{-\lambda\tilde{\tau}}}{[S_0+(1-S_0)e^{-\lambda\tilde{\tau}}]^3} > 0$. Again, the best-response advertising level γ^* satisfies $C'(\gamma) = \frac{e^{-\lambda\tilde{\tau}}(1-\varphi_h\alpha)}{[S_0+(1-S_0)e^{-\lambda\tilde{\tau}}]^2}$. Then, $C'_1(\gamma_1^*) = F(\gamma_1^*) < C'_2(\gamma_2^*) = F(\gamma_2^*)$. Hence, $F(\gamma_1^*) < F(\gamma_2^*) \Leftrightarrow \gamma_1(\varphi_h, \tilde{\tau}) > \gamma_2(\varphi_h, \tilde{\tau})$ for all $\tilde{\tau} > \hat{\tau}_1$. From Lemmas 2 and 5, if $\tau(\varphi_h, 0) = \frac{1}{\lambda} \ln \frac{1-\varphi_h\alpha}{\frac{k}{1-\alpha}-\varphi_h\alpha} > \hat{\tau}$, then $\gamma_1^*(\varphi_h) > \gamma_2^*(\varphi_h)$ and $\tau_1^*(\varphi_h) < \tau_2^*(\varphi_h)$. Otherwise, $\gamma^*(\varphi_h) = \gamma_2^*(\varphi_h) = 0$ and $\tau_1^*(\varphi_h) = \tau_2^*(\varphi_h) = \frac{1}{\lambda} \ln \frac{1-\varphi_h\alpha}{\frac{k}{1-\alpha}-\varphi_h\alpha}$. Finally, $\tau(\varphi_h, 0) = \hat{\tau}$ when $C'_1(0) = \frac{1-\alpha}{k} \left[1 - \frac{\varphi_h\alpha(1-\alpha)}{k}\right]$, and hence the inequalities hold strictly for $C'_1(0) < \frac{1-\alpha}{k} \left[1 - \frac{\varphi_h\alpha(1-\alpha)}{k}\right]$. \square

Finally, the second part of Proposition 3 follows from the next lemma, which provides the comparative statics of the cost of advertising on the value of acquiring information.

Lemma 7. *Consider two cost functions of advertising C_1 and C_2 , where $1 \leq C'_1 < C'_2$ and $C'_1(0) < \bar{C} = \frac{S^*(\varphi_h^{**}, 0)(1-S^*(\varphi_h^{**}, 0))}{S_0(\varphi_h^{**}, 0)(1-S_0(\varphi_h^{**}, 0))} (1 - \varphi_h^{**}\alpha)$. Then*

$$V(\varphi_h, \tau_1^*(\varphi_h), \gamma_1^*(\varphi_h)) < V(\varphi_h, \tau_2^*(\varphi_h), \gamma_2^*(\varphi_h)).$$

Proof. We note that from Lemma 6, $\gamma_1^*(\varphi_h) > \gamma_2^*(\varphi_h)$ and $\tau_1^*(\varphi_h) < \tau_2^*(\varphi_h)$, and from Lemma 2, $S_1^* = S_2^*$. From Equation (11), we can now write $V(\varphi_h, \tau^*(\varphi_h), \gamma^*(\varphi_h))$ as:

$$V(\varphi_h, \tau^*(\varphi_h), \gamma^*(\varphi_h)) = (1 - \gamma^*) \frac{k}{1 - S_0} \left[\frac{S^* - S_0}{S_0} + \ln \frac{S_0}{S^*} \right].$$

Taking the derivative with respect to γ^* holding S_0 and S^* constant,

$$\frac{\partial V}{\partial \gamma^*} = -\frac{k}{1 - S_0} \left[\frac{S^* - S_0}{S_0} - \ln \frac{S^*}{S_0} \right] < 0.$$

And from Lemma 3 in the proof of Proposition 2, $\frac{\partial V}{\partial S_0} < 0$, and the lemma follows immediately by noting that $\frac{dS_0}{d\gamma^*} > 0$ and $\frac{dS_0}{d\tau^*} = 0$. \square

Now, $V(\varphi_h, \tau^*(\varphi_h, \gamma^*), \gamma^*(\varphi_h, \tau^*))$ is continuous in φ_h and $\bar{c} \geq \frac{1-k+k \ln k}{1-\alpha} = V(1, \tau^*(1, 0), 0)$. Hence, it follows that $1 > \varphi_{h2}^* > \varphi_{h1}^*$. Also, $S_1^* < S_2^*$ follows from recalling $\frac{dS^*}{d\varphi_h} > 0$ from Equation (24). Finally, if $C'_1(0) \geq \bar{C} = \frac{S^*(\varphi_h^{**}, 0)(1-S^*(\varphi_h^{**}, 0))}{S_0(\varphi_h^{**}, 0)(1-S_0(\varphi_h^{**}, 0))} (1 - \varphi_h^{**}\alpha)$, the equilibrium for both cost functions is the same, $\{\varphi_h^{**}, \tau^*(\varphi_h^{**}), 0\}$.

A.6 Proof of Proposition 4

Suppose $0 \leq \gamma' < \gamma'' \leq S^*(\varphi_h, 0)$. The value of acquiring information may be written as

$$V(\varphi_h, \gamma) = \frac{k}{1 - S_0(\varphi_h, \gamma)} \left[\frac{S^*(\varphi_h, \gamma) - S_0(\varphi_h, \gamma)}{S_0(\varphi_h, \gamma)} - \ln \frac{S^*(\varphi_h, \gamma)}{S_0(\varphi_h, \gamma)} \right].$$

We know from Lemma 2 that S^* is independent of γ . We also know from Proposition 2 that V is decreasing in S_0 , and hence also decreasing in γ because

$$S_0(\varphi_h, \gamma) = \varphi_h \alpha + \gamma((1 - \varphi_h) \alpha + 1 - \alpha). \quad (27)$$

Therefore, $V(\varphi_h, \gamma') > V(\varphi_h, \gamma'')$.

The cutoff high type that acquires information is the largest solution to

$$\varphi_h^*(\gamma) = \frac{V(\varphi_h^*, \gamma)}{\bar{c}}$$

It is then immediate from the above that for higher values of γ , $V(\varphi_h, \gamma)$ is smaller for all φ_h . Hence, the fixed point satisfying the relationship in Equation (27) is also lower, thus $\varphi_h(\gamma') > \varphi_h(\gamma'')$. We also know that $\frac{dS^*}{d\varphi_h} > 0$ from the proof of Proposition 1, so we can also conclude $S^*(\varphi_h^*(\gamma'), \gamma') = S^*(\varphi_h^*(\gamma'), 0) > S^*(\varphi_h^{**}(\gamma'), 0) = S^*(\varphi_h^*(\gamma''), \gamma'')$.

A.7 Proof of Proposition 5

If $C'(0) \geq \frac{\tilde{S}_f^*(\varphi_h^{**}, \tau^*(\varphi_h^{**}), 0)(1 - \tilde{S}_f^*(\varphi_h^{**}, \tau^*(\varphi_h^{**}), 0))}{S_0(\varphi_h^{**}, 0)(1 - S_0(\varphi_h^{**}, 0))} (1 - \varphi_h^{**} \alpha)$, where $\tau^*(\varphi_h^{**}) = \frac{1}{\lambda} \ln \frac{S^*(\varphi_h^{**})}{1 - S^*(\varphi_h^{**})} \frac{1 - \varphi_h^{**} \alpha}{\varphi_h^{**} \alpha}$, the equilibrium is $\{\varphi_h^{**}, \tau^*(\varphi_h^{**}), 0\}$.

Note that choosing $T > \tau^*(\varphi_h^{**})$ will ensure that the diffusion of word of mouth is completed prior to the product being released so as to be consistent with our assumption about discounting. The proposition follows from noting that the marginal cost of advertising at the time of information release is $e^{rT} C'(\gamma)$, where $r > 0$ is the discount rate. And when $T > \frac{1}{r} \ln \frac{S^*(\varphi_h^{**}, 0)(1 - S^*(\varphi_h^{**}, 0))}{S_0(\varphi_h^{**}, 0)(1 - S_0(\varphi_h^{**}, 0))} \frac{((1 - \varphi_h^{**}) \alpha + 1 - \alpha)}{C'(0)}$, we have $e^{rT} C'(0) > \frac{S^*(\varphi_h^{**}, 0)(1 - S^*(\varphi_h^{**}, 0))}{S_0(\varphi_h^{**}, 0)(1 - S_0(\varphi_h^{**}, 0))} (1 - \varphi_h^{**} \alpha)$ and we have that $\gamma = 0$ and $\varphi_h = \varphi_h^{**}$.

A.8 Trembling Hand Refinement in Advertising Section

The analysis in section 4 assumed that the level of advertising is unobserved by individuals, and analyzes the firm's advertising decision and the consumer's choice of a time to stop spreading information as a simultaneous move game. In any equilibrium in which $0 < \gamma^* < 1$, receiving or not receiving an advertisement are both consistent with the firm's equilibrium strategy. Thus, consumers' beliefs upon seeing or not seeing an advertisement are pinned down by Bayesian updating, and their beliefs are the same in both scenarios. For example, if the consumer believes that the firm sent out an advertisement to 10% of the population, the fact that she did or did not receive an advertisement does not change her prior belief. Therefore, the equilibria we find when analyzing the subgame from $t \geq 0$ onward as a simultaneous move game are perfectly consistent with this analysis. On the other hand, in an equilibrium in which $\gamma^* = 0$, consumers expect a zero probability of receiving an advertisement. Hence, Bayesian updating does not provide any guidance for pinning down consumers' off-equilibrium beliefs

when consumers actually observe an advertisement in these equilibria. In this appendix, we analyze a trembling-hand perfect equilibrium.

Prior to mixing, consumers may or may not receive the advertisement from the firm and may potentially condition the time at which they choose to stop passing on word of mouth on this. Thus, the strategy τ of each consumer is a function of whether the consumer received an advertisement. We denote this occurrence by $a = 0, 1$, where $a = 1$ indicates a consumer who received an advertisement.

The tremble we analyze is on the firm's advertising strategy. We assume that when the firm chooses an advertising level γ , then with probability ϵ , individuals in the fraction of the population advertised do not receive the advertisement, and with the same probability ϵ , the fraction not advertised to do receive the advertisement. Thus, when the firm chooses a level $\gamma \in [0, 1]$, the actual fraction that receives the advertisement is $\gamma(1 - \epsilon) + \epsilon(1 - \gamma) = \gamma + \epsilon - 2\epsilon\gamma$. This ensures that when the firm chooses $\gamma^* = 0, 1$, the actual level of advertising is ϵ and $1 - \epsilon$, respectively, and Bayesian updating can be used at each information set $a = 0, 1$ of consumers. Hence, the updated beliefs of consumers $\tilde{\gamma}(a)$ will be the same for $a = 0, 1$, and we can proceed by denoting it by just $\tilde{\gamma}$. We can now write out the consumers' best response in terms of $\tilde{\gamma}$ and ϵ . Importantly, the conjectured level of "effective" advertising after accounting for the tremble is the same at both information sets $a = 0, 1$, and the best response is independent of a . The best-response function of consumers is now given by

$$\tau^*(a, \varphi_h, \tilde{\gamma} + \epsilon - 2\epsilon\tilde{\gamma}) = \frac{1}{\lambda} \ln \frac{\tilde{S}_c^*(\varphi_h, 0, \tilde{\gamma} + \epsilon - 2\epsilon\tilde{\gamma})}{1 - \tilde{S}_c^*(\varphi_h, 0, \tilde{\gamma} + \epsilon - 2\epsilon\tilde{\gamma})} \frac{1 - S_0(\varphi_h, 0, \tilde{\gamma} + \epsilon - 2\epsilon\tilde{\gamma})}{S_0(\varphi_h, 0, \tilde{\gamma} + \epsilon - 2\epsilon\tilde{\gamma})}. \quad (28)$$

The firm's choice of advertising cannot directly influence consumers' strategy and is very similar to earlier. But we replace γ with the "effective" advertising $\tilde{\gamma} + \epsilon - 2\epsilon\tilde{\gamma}$ on the RHS:

$$C'(\gamma(\varphi_h, \tilde{\tau})) \quad (29)$$

$$= \frac{e^{-\lambda\tilde{\tau}}(1 - \varphi_h\alpha)}{\left[\begin{array}{c} S_0(\varphi_h, 0, \gamma(\varphi_h, \tilde{\tau}) + \epsilon - 2\epsilon\gamma(\varphi_h, \tilde{\tau})) \\ + (1 - S_0(\varphi_h, 0, \gamma(\varphi_h, \tilde{\tau}) + \epsilon - 2\epsilon\gamma(\varphi_h, \tilde{\tau})))e^{-\lambda\tilde{\tau}} \end{array} \right]^2} \text{ if } \gamma^* \in (0, 1)$$

and

$$C'(0) \geq \frac{e^{-\lambda\tilde{\tau}}(1 - \varphi_h\alpha)}{[S_0(\varphi_h, 0, \epsilon) + (1 - S_0(\varphi_h, 0, \epsilon))e^{-\lambda\tilde{\tau}}]^2} \text{ if } \gamma^* = 0. \quad (30)$$

The information acquisition choice of consumers satisfies:

$$\begin{aligned} \varphi_h^* \bar{c} &= V(\varphi_h^*, \tau^*(\varphi_h^*), \gamma^*(\varphi_h^*) + \epsilon - 2\epsilon\gamma^*(\varphi_h^*)) & \text{if } V(1, \tau^*(1), \gamma^*(1)) < \bar{c} \\ \varphi_h^* &= 1 & \text{if } V(1, \tau^*(1), \gamma^*(1) + \epsilon - 2\epsilon\gamma^*(1)) \geq \bar{c}. \end{aligned} \quad (31)$$

We are interested in the limit of the equilibria as $\epsilon \rightarrow 0$. We note the Equations (28), (29), and

(30) are continuous in ϵ at $\epsilon = 0$. Thus, the best-response functions go to the best-response functions in section 4, and the limit of the equilibria is the same. Similarly, Equation (31) is continuous in all its arguments and thus also attains the same limit as in section 4.

A.9 Other Refinements

As before, if there are multiple solutions to the consumer's problem, we assume that the firm can implement the solution that results in the greatest level of information diffusion. The beliefs are Bayesian on the equilibrium path for $t \leq \tau^*$, and satisfy $b_j^*(\theta_i = h|m, t) - b_j^*(\theta_i = h|\emptyset, t) < k$ for all $t > \tau^*$. Finally, we maintain the assumptions that $\bar{c} \geq \frac{1-k+k \ln k}{1-\alpha}$, which guarantees that not all the high social types acquire information in equilibrium, and $k \geq 2\alpha$, which is a sufficient condition for uniqueness of an equilibrium in the continuation game from $t \geq 0$. We require uniqueness in order to undertake comparative static analysis of the firm's costs of advertising on the equilibrium.

B Model with Uniform Cost of Acquisition and Heterogeneous Cost of Communication

We consider a simple model where the acquisition cost is uniform and communication costs are heterogeneous across types. Our goal is to demonstrate the technical issues that can arise as the result of introducing heterogeneous communication costs. In particular we show that in a pooling equilibrium (where both types generate word of mouth for a fixed period and then remain silent from then on), different communication costs imply a different incentive to deviate across types in the "silent period."

Let's consider a model with $c > 0$ (uniform acquisition cost), $k_h < k_l$ (it is less costly for high-type consumer to communicate word of mouth), and $V_{i,\theta} = \lambda \left(\int_0^{t^*} \left(\frac{1-S(t)}{1-S_0} \right) (\Delta U(t) - k_\theta) dt \right) + v_i$, where v_i is i.i.d. uniformly on $[0, \bar{v}]$. The outline of establishing the equilibrium is the following. The total signaling benefit (V) consists of the instantaneous benefit of signaling that is integrated over time plus a random error. Note that the term $\Delta U(t)$, which reflects the change in the receiver's beliefs about the sender's type, does not vary by type θ . Let's consider a pooling equilibrium where both types communicate until some point in time and stop communicating afterward.

For example, suppose that both types stop transmitting when $\Delta U(t^*) - k_l = 0$. Therefore, in this equilibrium $V_{i,h} = \lambda \left(\int_0^{t^*} \left(\frac{1-S(t)}{1-S_0} \right) (\Delta U(t) - k_h) dt \right) + v_i$ and $V_{i,l} = \lambda \left(\int_0^{t^*} \left(\frac{1-S(t)}{1-S_0} \right) (\Delta U(t) - k_l) dt \right) + v_i$. This clearly implies that the total signaling benefit is higher for h -type on average as it is less costly for her to communicate in each period. Hence, even with uniform acquisition cost, the h -type would acquire more information at the beginning of the game. Hence, a permutation in the cost structure can mechanically yield the same result as the main model – a word of mouth transmission still implies that the sender is more likely to be high type. However, the interpretation is quite different – the high

types invest more in information acquisition not because it is less costly for them to find out about restaurants, but because they bear a lower cost of transmitting the information.

Note that the pooling equilibrium above fails the D1 refinement (Fudenberg and Tirole 1990) in the “silent” period after word of mouth stops. (The Cho-Kreps intuitive criterion (Cho and Kreps 1986) here does not restrict the possible set of beliefs). The intuition for this is that if someone deviates and transmits word of mouth, the high type is more likely to benefit from this than the low type because the high type bears a lower cost of communication.

More formally, consider the time period after which word of mouth is no longer transmitted in equilibrium. Hence in equilibrium the payoff that potential senders of word of mouth receive is zero. Next consider a deviation where word of mouth *is* transmitted. If we can show that type h makes higher profit under deviation than in equilibrium under a strictly bigger set of responses from the receiver than type l does, then D1 requires that the receiver does not believe that the sender could be of type l . (In other words, because there is a wider set of receiver’s beliefs about the sender’s type for which the high type is better off sending a message, then observing an off-equilibrium message results in the receiver believing that the sender is a high type). Note that in our model the receiver’s response is simply the belief on the sender’s type ($b_j^{dev}(\theta_i = h|m, t)$). For h -type, the per-period signaling benefit in deviation (the sender’s profit) is $\Delta U_{i,h}(t) - k_h = b_j^{dev}(\theta_i = h|m, t) - b_j(\theta_i = h|\emptyset, t) - k_h$. This is strictly greater than 0 (the equilibrium payoff) if $b_j^{dev}(\theta_i = h|m, t) > b_j(\theta_i = h|\emptyset, t) + k_h$. For l -type, the per-period signaling benefit in deviation (the sender’s profit) is $\Delta U_{i,l}(t) - k_l = b_j^{dev}(\theta_i = h|m, t) - b_j(\theta_i = h|\emptyset, t) - k_l$. This is strictly greater than 0 (the equilibrium payoff) if $b_j^{dev}(\theta_i = h|m, t) > b_j(\theta_i = h|\emptyset, t) + k_l$. Clearly the set $b_j^{dev}(\theta_i = h|m, t)$ (best responses by the receiver) under which the sender makes higher payoff than in equilibrium is strictly bigger for h than for l because $k_h < k_l$.

Hence, we showed that D1 requires that the belief following word of mouth in the “silent period” is that the type is *not* l , which implies that $b_j^{dev}(\theta_i = h|m, t) = 1$. But, of course, this destroys the equilibrium because both types would prefer to deviate and transmit word of mouth given such a positive belief. Note that it is exactly the difference in communication cost between types that is destroying this equilibrium. Because in our original model both types bear the same cost of communication, D1 has no “bite” in the silent period.

References

- [1] ANGELIS, M., BONEZZI, A., PELUSO, A., RUCKER, D., AND CONSTABILE, M. "On Braggarts and Gossips: A Self-Enhancement Account of Word-of-Mouth Generation and Transmission." *Journal of Marketing Research*, Vol. 49 (2012), pp. 551-563.
- [2] BAGWELL, K. "The Economic Analysis of Advertising," In M. Armstrong and R. Porter, eds., *Handbook of Industrial Organization*. Amsterdam: North-Holland, 2007.
- [3] BAUMEISTER, R. "The Self," In D. T. Gilbert, S. T. Fiske, and G. Lindzey, eds., *The Handbook of Social Psychology*. Hoboken, NJ: John Wiley & Sons, 1998.
- [4] BAUMEISTER, R., TICE, D., AND HUTTON, D. "Self-presentational Motivations and Personality Differences in Self-esteem." *Journal of Personality*, Vol. 57 (1989), pp. 547-579.
- [5] BERGER, J. AND MILKMAN, K. "What Makes Online Content Viral?" *Journal of Marketing Research*, Vol. 49 (2012), pp. 192-205.
- [6] BERGER, J. AND SCHWARTZ E. "What Drives Immediate and Ongoing Word of Mouth?" *Journal of Marketing Research*, Vol. 48 (2011), pp. 869–880.
- [7] BIMPIKIS, K., OZDAGLAR, A., AND YILDIZ, E. "Competing over Networks." Working paper, available at asu.mit.edu/sites/default/files/documents/publications/network-marketingOR.pdf. 2015.
- [8] BUGHIN, J., DOOGAN, J., AND VETVIK, O. "A New Way to Measure Word-of-Mouth Marketing." *McKinsey Quarterly*, Vol. 2 (2010), pp. 113-116.
- [9] Calvó-Armengol, A., Martí, J., and Prat, A. "Endogenous Communication in Complex Organizations." *SAET Conference*, 2012.
- [10] CAMPBELL, A. "Word-of-Mouth Communication and Percolation in Social Networks." *American Economic Review*, Vol. 103 (2013), pp. 2466-2498.
- [11] CAMPBELL, J. "Signaling to a Network of Consumers." Working paper, available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1710877. 2010.
- [12] CANDOGAN, O., BIMPIKIS, K., AND OZDAGLAR, A. "Optimal Pricing in the Presence of Local Network Effects." In *International Workshop on Internet and Network Economics*. Heidelberg: Springer, 2010.

- [13] CHATTERJEE, K. AND DUTTA, B. "Word of Mouth Advertising, Credibility and Learning in Networks." Working Paper, no. 941, Department of Economics, University of Warwick, Coventry, 2010.
- [14] CHEVALIER, J. AND MAYZLIN, D. "The Effect of Word of Mouth on Sales: Online Book Reviews." *Journal of Marketing Research*, Vol. 43 (2006), pp. 345-354.
- [15] CHINTAGUNTA, P., GOPINATH, S., AND VENKATARAMAN, S. "The Effects of Online User Reviews on Movie Box Office Performance: Accounting for Sequential Rollout and Aggregation Across Local Markets." *Marketing Science*, Vol. 30 (2010), pp. 398-419.
- [16] CHO, I. AND KREPS, D. "Signaling Games and Stable Equilibria," *Quarterly Journal of Economics*, Vol. 102 (1987), pp. 179-221.
- [17] DELLAROCAS, C., FAN, M., AND WOOD, C. "Self-Interest, Reciprocity, and Participation in Online Reputation Systems." Working Paper, no. 4500-04, Sloan Management School, MIT, Cambridge, 2004.
- [18] DICHTER, E. "How Word-of-Mouth Advertising Works." *Harvard Business Review*, Vol. 44 (1966), pp. 131-144.
- [19] DYE, R. "The Buzz on Buzz." *Harvard Business Review*, Vol. 78 (1999), pp. 139-146.
- [20] FISKE, S. "Five Core Social Motives, Plus or Minus Five," In S. Spencer, S. Fein, M. Zanna and J. Olsen, eds., *Motivated Social Perception: The Ontario Symposium*. New York: Psychology Press, 2003.
- [21] FUDENBERG, D. AND TIROLE, J. *Game Theory*. Cambridge, MA: MIT Press, 1991.
- [22] GALEOTTI, A. "Talking, Searching and Pricing." *International Economic Review*, Vol. 50 (2010), pp. 159-174.
- [23] GALEOTTI, A. AND GOYAL, S. "Network Multipliers and the Optimality of Indirect Communication," Discussion Paper, no. 642, Department of Economics, University of Essex, Colchester, 2007.
- [24] GALEOTTI, A. AND MATTOZZI, A. "Strategic Information Transmission in Networks." *American Economic Journal: Microeconomics*, Vol. 3 (2011), pp. 307-327.
- [25] HENNIG-THURAU, T., GWINNER, K., WALSH, G., AND GREMLER, D. "Electronic Word of Mouth via Consumer Opinion platforms: What Motivates Consumers to Articulate Themselves on the Internet?" *Journal of Interactive Marketing*, Vol. 18 (2004) pp. 38-52.

- [26] HERR, P., KARDES, F., AND KIM, J. "Effects of Word-of-Mouth and Product-Attribute Information on Persuasion: An Accessibility-Diagnosticity Perspective." *Journal of Consumer Research*, Vol. 17 (1991), pp. 454-462.
- [27] HUGHES, M. "*Buzzmarketing: Get People to Talk About Your Stuff.*" New York, NY: Penguin Group, 2005.
- [28] IFRACH, B., MAGLARAS, C., AND SCARSINI, M. "Monopoly pricing in the Presence of Social Learning." Working Paper, no. 12-01, NET Institute, New York, 2015.
- [29] KEAVENEY, S. "Customer Switching Behavior in Service Industries: An Exploratory Study." *Journal of Marketing*, Vol. 59 (1995), pp. 71-83.
- [30] LUCA, M. "Reviews, Reputation, and Revenue: The Case of Yelp.com." Working Paper, no. 12-016, NOM Unit, Harvard Business School, Cambridge, 2016.
- [31] NIEHAUS, P. "Filtered Social Learning." *Journal of Political Economy*, Vol. 119 (2011), pp. 686-720.
- [32] PESENDORFER, W. "Design Innovation and Fashion Cycles," *American Economic Review*, Vol. 85 (1995), pp. 771-792.
- [33] RICHINS, M. "Negative Word-of-Mouth by Dissatisfied Consumers: A Pilot Study." *Journal of Marketing*, Vol. 47, pp. 68-78.
- [34] SEDIKIDES, C. "Assessment, Enhancement, and Verification Determinants of the Self-Evaluation Process." *Journal of Personality and Social Psychology*, Vol. 65 (1993), pp. 317-338.
- [35] SERNOVITZ, A. *Word of Mouth Marketing*. Chicago, IL: PressBox, 2011.
- [36] STEIN, J. "Conversations Among Competitors." *American Economic Review*, Vol. 98 (2008), pp. 2150-2162.
- [37] SUNDARAM, D., MITRA, K., AND WEBSTER, C. "Word-of-Mouth Communications: A Motivational Analysis." *Advances in Consumer Research*, Vol. 25 (1998), pp. 527-531.
- [38] WOJNICKI, A. AND GODES, D. "Signaling Success: Strategically-Positive Word of Mouth." Working Paper, Rotman School of Management, University of Toronto, Toronto, 2011.
- [39] Word of Mouth Marketing Association (2011) infographic: http://wordofmouthma.org/word/wp-content/uploads/2011/10/word_of_mouth_marketing_impact_and_influence_word_of_mouthma.png

[40] YOGANARASIMHAN, H. “Cloak or Flaunt? The Fashion Dilemma.” *Marketing Science*, Vol. 31 (2012), pp. 74–95.

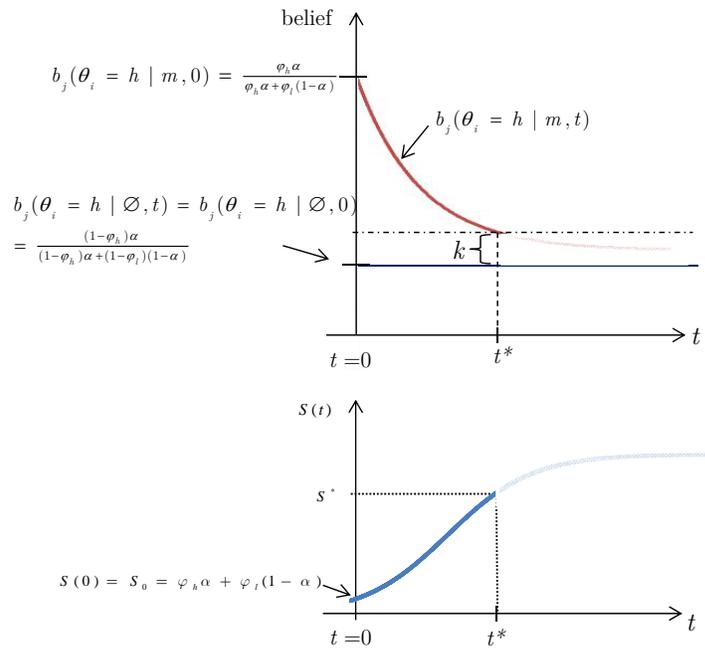


Figure 1: The effect of receiver's beliefs on diffusion of information

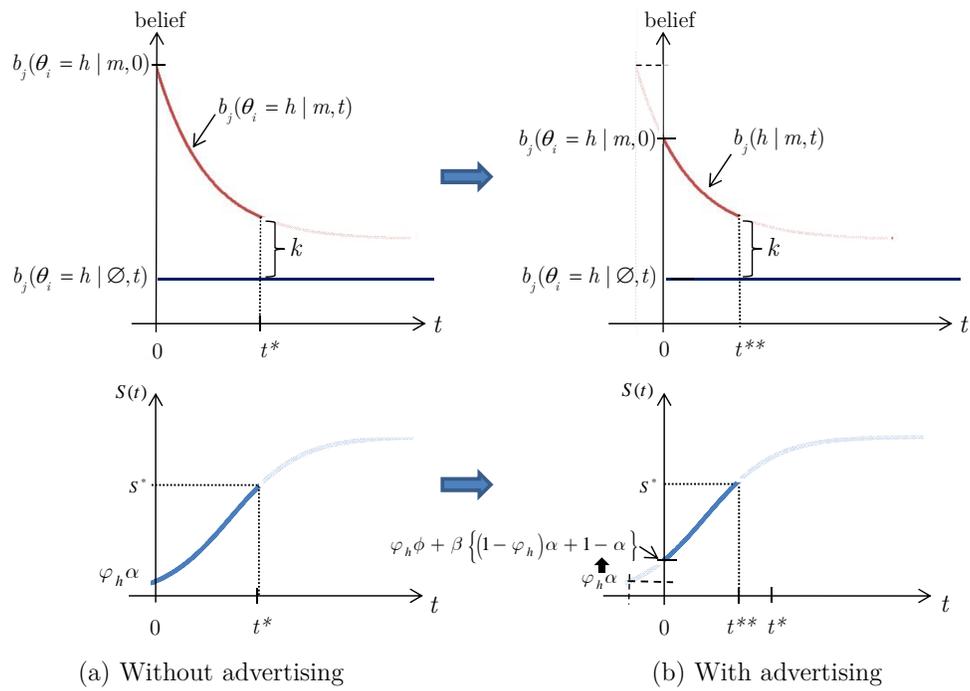


Figure 2: The Effect of Advertising on Word of Mouth and Diffusion, Given φ_h .