Managing Buzz

Arthur Campbell[†], Dina Mayzlin[‡] and Jiwoong Shin[†*]

October 2013

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. In the model individuals are privately motivated by a desire to signal their type to others. Individuals are either a high or a low type, and during social interactions it is valuable for any individual to increase another person's posterior belief that she is a high type. We find that a firm will restrict access to information by low types at the information release stage. We also find that advertising may crowd out the incentives for consumers to engage in word of mouth, and that a firm can benefit from a credible commitment not to engage in advertising. Finally, we find that the ability by the firm to target advertising to well-connected consumers may be detrimental to the signaling value of word of mouth. Our model provides new insights into the tradeoff a firm may face between spreading information quickly versus maximizing the total spread of information about the product.

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

[*]We thank Dirk Bergemann, Alessandro Bonatti, Odilon Camara, Florian Ederer, Johannes Horner, Duncan Simester, Birger Wernerfelt, and seminar participants at Brown University, CKGSB, Columbia University, Indiana University, MIT, Monash University, Paris School of Economics, Purdue University, University of Cambridge, University of Florida, University of Melbourne, University of Toronto, Yale University, SICS Conference at Berkeley, and the Conference on the Economics of Advertising in Moscow. [†]135 Prospect St., P.O. Box 208200, New Haven, CT 06520, e-mail: arthur.campbell@yale.edu, jiwoong.shin@yale.edu,[‡]USC Marshall School of Business, Los Angeles, CA 90089, mayzlin@marshall.usc.edu.

1 Introduction

In July 2011, the European music streaming site Spotify launched in the US market. At first, its free US version was available by invitation only. Interestingly, obtaining the 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.[1] After a few weeks, anyone could download the free version of Spotify through the company's website.[2] By November 2011, Spotify was able to attract 4 million users, while undertaking almost no advertising.[3] In the same month Google launched its new social network site, Google+, using the same invitation-only method. Users could join either by receiving an invitation from Google or by receiving one from a friend. For a brief period in the summer of 2011, receiving an invitation from Google+ became a status symbol. Here is what Ken Hess wrote on the tech news site ZDNet on June 30, 2011, "Dear Google, I want a Google+ account. I'm an avid Googler and have always been an early adopter of all things Google. Please give me an account before you give them to anyone else on my list so I can gain some real street cred with my fellow ZDNetters."[4] On October 2011, Google announced Google+'s user base to be at more than 40 million users.[5] Similarly to the Spotify example, media sources speculated that the initial exclusivity surrounding the site contributed to early buzz and high adoption rates.[6]

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 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."[7] 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 exclusiv-

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

[2] One 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 such service in the world and 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).

[3] http://www.theverge.com/2013/3/25/4145146/spotify-kicks-off-ad-blitz-as-rumors-hint-of-video-service.

[4] http://www.zdnet.com/blog/btl/dear-google-where-the-hell-is-my-google-invitation/51640

[5] http://www.pcmag.com/article2/0,2817,2398114,00.asp

[6] See, 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).

[7] http://blog.hubspot.com/blog/tabid/6307/bid/6183/4-Social-Media-Methods-for-Generating-Word-of-Mouth.aspx

ity 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 prerelease versions of software to active message board users.

One interesting aspect of these strategies is that they seem to contradict the intuition that wider exposure to product information will lead to more word of mouth and a larger fraction of the population eventually holding the information. In these examples, it is profitable for a firm to increase the number of people that know information about their product, yet these firms adopt strategies to purposefully limit the number of people who are initially exposed to product information. In this paper we explore why a firm that is seeking to maximize the number of people who possess information about its product may undertake strategies which actively restrict early access to the information.

The firms' emphasis on word of mouth management is driven by the recognition that word of mouth significantly impacts consumer purchase decisions. Word 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, Kardes, and Kim 1991), and appliances and clothing (Richins 1983).[8] However, theoretical models of word of mouth largely ignore the incentives to engage in word of mouth by treating it as a costless, mechanical process. Consequently, the existing literature offers little insight into how exclusivity may help diffuse information via word of mouth. A central issue for understanding a firm's exclusivity strategy in this context is the motivation of consumers to acquire information and then engage in word of mouth about it. In this paper, we address this issue by modeling the motivation of individuals to engage in word of mouth and treat the word of mouth generation process as an outcome of strategic consumer behavior. We focus on a particular motive – word of mouth as "self-enhancement" (Baumeister 1998) or the idea that an individual engages in word of mouth to make herself appear knowledgeable or trendy in a social setting. By explicitly modeling this motivation of consumers to engage in word of mouth, we show how exclusivity, through restricting who has access to information, affects the decision to spread word of mouth and how these strategies increase the total diffusion of information. A main lesson from our analysis is that a firm's strategy to generate word of mouth is as much about who does not receive information as it is about who does.

We develop a model where consumers meet one another at a Poisson rate over time. 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 (utility from "self enhancement"). The most straightforward interpretation of high type is being knowledgeable about a particular product area or

[8] 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).

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. Prior to meeting others, individuals choose whether or not 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 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 signaling game.

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. We also highlight a basic trade-off between increasing the initial speed of diffusion and maximizing the total spread of information about the product: while confining the initial spread of information to the high types maximizes the incentive to talk on the part of each exposed consumer, the diffusion process takes longer since the number of exposed individuals is small early on.

Second, we introduce advertising by the firm and consider how the ability of the firm to undertake advertising affects the incentives for consumers 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 not to advertise is to release the information a sufficient amount of time prior to product release. This increases the present value of costs of advertising for the firm, thereby allowing word of mouth to occur.

Finally, we apply our model to a setting where the firm may target advertising at individuals who are more connected (meet people more frequently) than others. This type of targeting strategy appears to be particularly attractive for firms because a significant increase in demand might be possible with limited budget. We find that advertising which is targeted in this way tends to have a particularly strong negative impact on the signaling value of information. This reduces the extent

of total information diffusion even conditional on an amount of information acquired by individuals. Unless an individual's meeting rate is correlated with her own social type, simply reaching out to those high mixing individuals can displace the incentive to spread word of mouth.

The rest of the paper is organized as follows. In the next section, we discuss the related literature, and Section 3 presents a model of buzz based on the self-enhancement motive and analyzes how a 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 and Section 5 extends the model by allowing the heterogeneous mixing patterns among consumers. We conclude in Section 6.

2 Literature Review

Our paper is most closely related to a number of papers in social network theory that study 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 papers are interested in 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, and Chatterjee and Dutta 2010), or both (A. Campbell 2013).

Galeotti (2010) is the most related paper. The author considers costly search for pricing information by consumers in a model where two firms engage in Bertrand competition. Similarly to the current paper it is costly for consumers to acquire information directly. Differently, word of mouth is costly for the receiver of information but it is not costly for the sender of information. In contrast we assume it is costly for the sender to pass on information but not for the receiver to receive the information. The reason for this difference in assumptions is that the focus of the two papers are different. The current paper is focused on the motivations of individuals with information to engage in word of mouth (the senders of information). On the other hand Galeotti (2010) is concerned with the equilibrium of consumer search (the receivers of information) either directly or through friends and firm pricing.

Many of the other papers in the existing literature have detailed models of the social network but 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.[9] Chatterjee and Dutta (2010) assume the firm can pay individuals to engage in word of mouth. Ifrach, Maglaras and Scarsini (2011), Candogan, Bimpikis and Ozdaglar (2010), and A. Campbell (2013)

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

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 or pass on information will facilitate a greater amount of information diffusion. In contrast, our paper addresses the issue of why individuals engage in word of mouth about a firm's product by explicitly incorporating an individual's social motivation. 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, advertising and the ability to engage in word of mouth. The firm will optimally impose restrictions in ways which increase the signaling value of the information.

In a more broad context, our work is also related to a number of papers which examine information diffusion through endogenous (privately motivated) communication between individuals in social networks. Galeotti and Mattozzi (2011) study competition between two political parties when voters acquire information via both advertising and word of mouth. The authors find that richer communication networks lead parties to disclose less political information, voters to be less informed and parties to adopt more extreme policy platforms. Calvó-Armengol, de Martí and Prat (2012) consider endogenous communication in organizations and Stein (2008) considers how far through a population ideas spread. Also, Niehaus (2011) analyzes what type of information may be learned in an environment where local information sharing is efficient but is not necessarily globally efficient. However, none of these papers study the effect of exclusivity on the information diffusion.

Our paper 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, Webster 1998), reciprocity (Dichter 1966, Dellarocas, Fan, and Wood 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 paper, we focus on the latter desire to signal to others about oneself 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 talents (Baumeister 1998, Wojnicki and Godes 2011).

A number of papers provide empirical evidence that word of mouth is influenced by motives related to self-presentation (Berger and Milkman 2011, Berger and Schwartz 2011, Hennig-Thurau, et. al. 2004, Sundaram et al. 1998, Wojnicki and Godes 2011). Berger and Milkman (2011) find that positive content is more likely to be shared, as is content that evokes high-arousal emotions. The

authors conjecture that the sharing of positive content may be due to impression-management. Berger and Schwartz (2011) find that, in the short run, conversations are influenced by how interesting the product is: consumers do not want to appear to be dull. In a survey conducted by Hennig-Thurau et. al. (2004) respondents indicate self-enhancement as one of the primary motivations behind word of mouth. Also in a survey, Sundaram et. al. (1998) find that 20% of positive word of mouth is undertaken "to show connoisseurship, to project themselves as experts, to enhance status, and to seek appreciation," and Wojnicki and Godes (2011) show in a series of experiments that experts are less likely to talk about their negative experiences in an attempt to enhance their self-image since a negative outcome reflects badly on their ability to make choices.

Finally, our model studies how social interactions between consumers can be influenced by a 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 particular product serves as a wealth signal for a consumer to others that they are a high type. This is valuable for high types to identify one another during a matching process. Although both the current paper and Pesendorfer (1995) consider social interactions between consumers, our focus is very different as we consider how a firm's information release and advertising strategy interacts with these social concerns (in particular, the effect of exclusivity on information diffusion) whereas the focus of Pesendorfer (1995) is on product cycles and pricing. 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 paper is similar to Yoganarasimhan (2012) in that both model the firm's incentive to restrict information in communication strategy to facilitate the social interaction. However, Yoganarasimhan (2012) analyze the firm's pricing strategy to extract consumer surplus in a static setting while we focus on the effect of initial exclusive release on the extent of information diffusion in a dynamic setting.

3 A Model of Buzz

3.1 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. The high type consumers are more knowledgeable about the product category, and this is viewed positively by all the consumers. In social situations, it is valuable for either type of person to be perceived as the high type by others (self-enhancement motives). One can think of the high-type consumers are 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.

We assume that the firm's profit is linear in the fraction of consumers who obtain information about its product.[10] The reduced form of the firm's objective is to maximize the fraction of the population which receives a piece of information m about its product. We denote the fraction of the population that has received the information at time 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 costly search to learn about the product themselves. Second, they may costlessly hear about the product from another person. Once a consumer has found out about the product through either channel, she too is able to pass on the information to others. Our focus in this paper is to study how far the information eventually diffuses through the population when passing it on is costly to consumers.

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: a consumer is not able to fabricate information. There is a fixed lower bound on the costs for obtaining information for each consumer, c_i, which is i.i.d. uniformly on $[0, \bar{c}]$. One can think of this as the minimum amount of time and effort an individual must expend to understand the information.[11] 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 $\upsilon_h, \upsilon_l \geq 0$ which is type-specific (that is, the total cost that an individual i of type $\theta = \{h, l\}$ bears to obtain information about the product is $c_i + \upsilon_\theta$).

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, through 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

[10] Although we leave the firm's objective in this reduced form in our main text, one could also model this reduced form. The firm's product is produced at marginal cost c and is ex ante equally likely to be one of n types. Consumers are also one of n types; each type values the corresponding product at $\vartheta > c$ and values the other types at 0. There is an equal mass of each consumer type in the population. Suppose that $c > \frac{\vartheta}{n}$. Absent receiving information, no consumer will purchase the product at any price $p > \frac{\vartheta}{n}$ and the firm will not sell at a price $p < c$. However, any consumer that receives information about the type of the product will purchase it if it matches their preferred type at $p \leq \vartheta$. The firm makes a profit of $\frac{\vartheta - c}{n}$ per consumer who receives information, thus the firm's profit is linear in the fraction of the population who receive information about the product. Note also that the firm captures all the surplus from the purchase and so the information is not valuable to a consumer in its own right. We study a setting where the firm captures the full surplus from any sale so as to isolate the incentives of individuals upon acquiring the information to pass it on to others.

[11] We think our model fits particularly well many entertainment, technology and fashion product categories where being perceived as knowledgeable about these areas is desirable. In these cases, 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 to find that the firm has a strict incentive to treat each type asymmetrically through its information release strategy.

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 is lower for the tech-savvy consumers who are already familiar with Twitter. Another prominent example of this occurs when technology companies such as Apple or Samsung make product announcements at 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.

From time $t = 0$ onwards 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, where we assume $\alpha < k < 1 - \alpha$, or pass on no information $\varnothing$ at zero cost. We assume that this 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 "self-enhancement." We capture this idea through a 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 $U_i\left(b_j\left(\theta_i = h|m,t\right)\right)$ if consumer i passes a message m at time t, where $b_j\left(\theta_i = h|m,t\right)$ is the other consumer $j's$ belief that consumer i is a high type upon receiving the information m. And similarly, $U_i\left(b_j\left(\theta_i = h|\varnothing,t\right)\right)$ if the consumer does not pass information, where $b_j\left(\theta_i = h|\varnothing,t\right)$ is the belief if no signal (denoted by $\varnothing$) is sent. Given our notions of high and low types, we assume $\frac{dU_i}{db_j} > 0$. Also note the signaling benefit at a time t is

$$\Delta U_i\left(t\right) = U_i\left(b_j\left(\theta_i = h|m,t\right)\right) - U_i\left(b_j\left(\theta_i = h|\varnothing,t\right)\right) \tag{1}$$

which 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,

$$\begin{aligned} &U\left(b_j\left(\theta_i = h|m,t\right)\right) - U\left(b_j\left(\theta_i = h|\varnothing,0\right)\right) \\ &= \overline{u}\left[b_j\left(\theta_i = h|m,t\right) - b_j\left(\theta_i = h|\varnothing,t\right)\right] \end{aligned} \tag{2}$$

where we normalize $\bar{u} = 1$. We assume that an individual passes on information only when $\Delta U_i\left(t\right) > k$.[12]

[12]This procludes equilibria where only a fraction of individuals with the information choose to pass it on due to indifference and these fractions happen to differ in such a way across types that $\Delta U_i\left(t\right) = k$ is maintained over time.

Finally note that we assume that the firm extracts the entire consumer surplus from the sale of the product to a consumer, see footnote 10. Thus, there is no value from obtaining information for the purposes of making a purchase decision. This setup allows us to only focus on the utility that the consumer derives from the information, which is accomplished through signaling.

3.2 Analysis

We focus on a signaling equilibrium where consumers engage in word of mouth in order to signal to each other that they are a 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.

We analyze how 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 are going to be endogenously determined in equilibrium, but for now we take (φ_h, φ_l) (which we denote by $\vec{\varphi}$) as given. For the moment, we also assume that $\varphi_h > \varphi_l$ (which will be confirmed in equilibrium subsequently). The initial condition of the informed population at $t = 0$ is $S_0 = \varphi_h \alpha + \varphi_l (1 - \alpha)$ and the rate of growth of the informed population is given by:

$$\frac{dS}{dt} = \lambda S(t)(1 - S(t)) \tag{3}$$

This results in the following path for $S(t)$:

$$S(t) = \frac{1}{1 + ae^{-\lambda t}}, \text{ where } a = \frac{1 - S_0}{S_0} \tag{4}$$

which continues to grow while word of mouth is taking place. We characterize the extent of the diffusion (when $S(t)$ stops growing) below.

Next, we characterize the evolution of consumers' belief over time while individuals with the information engage in word of mouth.

Beliefs

At $t = 0$, consumers' beliefs are

$$b_j(\theta_i = h | m, 0, \vec{\varphi}) = \frac{\varphi_h \alpha}{\varphi_h \alpha + \varphi_l (1 - \alpha)} \tag{5}$$

We feel these types of equilibria are unreasonable since they arbitrarily introduce an asymmetry between the types by manipulating indifference in a very specific manner. Furthermore the extent of diffusion found in this way is not robust to incorporating individual specific costs of passing on information which are draws from $[k, k + \epsilon]$ where ϵ is arbitrarily small. In the equilibria we analyze all individuals with information act in the same way, they either all pass it on or do not.

and

$$b_j\left(\theta_i = h|\varnothing, 0, \vec{\varphi}\right) = \frac{\left(1-\varphi_h\right)\alpha}{\left(1-\varphi_h\right)\alpha + \left(1-\varphi_l\right)\left(1-\alpha\right)}. \tag{6}$$

Beliefs change over time as the message diffuses through the population. The Bayesian belief on the sender's type when a consumer receives a signal at a time t (receives word of mouth) is given by:

$$\begin{aligned} b_j\left(\theta_i = h|m, t, \vec{\varphi}\right) &= \frac{S\left(t\right) - S_0}{S\left(t\right)}\left[b_j\left(\theta_i = h|\varnothing, 0, \vec{\varphi}\right)\right] + \frac{S_0}{S\left(t\right)}\left[b_j\left(\theta_i = h|m, 0, \vec{\varphi}\right)\right] \\ &= b_j\left(\theta_i = h|\varnothing, 0, \vec{\varphi}\right) + \frac{S_0}{S\left(t\right)}\left[b_j\left(\theta_i = h|m, 0, \vec{\varphi}\right) - b_j\left(\theta_i = h|\varnothing, 0, \vec{\varphi}\right)\right]. \end{aligned} \tag{7}$$

The belief on the sender's type upon not receiving a signal is:

$$b_j\left(\theta_i = h|\varnothing, t, \vec{\varphi}\right) = \frac{\left(1-\varphi_h\right)\alpha}{\left(1-\varphi_h\right)\alpha + \left(1-\varphi_l\right)\left(1-\alpha\right)} = b_j\left(\theta_i = h|\varnothing, 0, \vec{\varphi}\right). \tag{8}$$

Note that 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 the same over time.

Extent of diffusion

We focus on equilibria where the diffusion of the information stops at a time t^* when the marginal value of signaling equals the marginal cost of passing on the information.[13] This allows us to describe all consumers' decision about whether to pass on the information, conditional on having acquired it, by the time t^* at which consumers stop passing on information. The instantanous signaling value at t is:

$$\begin{aligned} \Delta U\left(t\right) &= U\left(b_j\left(\theta_i = h|m, t, \vec{\varphi}\right)\right) - U\left(b_j\left(\theta_i = h|\varnothing, t, \vec{\varphi}\right)\right) \\ &= \frac{S_0}{S(t)}\left[b_j\left(\theta_i = h|m, 0, \vec{\varphi}\right) - b_j\left(\theta_i = h|\varnothing, 0, \vec{\varphi}\right)\right] \end{aligned} \tag{9}$$

$S(t)$ is strictly increasing over time and thus the instantaneous signaling benefit of signaling strictly decreases over time. That is, as information diffuses through the population, and 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.[14]

[13] Formally we require that in equilibrium individuals pass on information at all times $r \leq t$ while $\lim_{r \to t^-} \Delta U\left(r\right) > k$ and stop at any time t^* where $\lim_{r \to t^{*-}} \Delta U\left(t^*\right) = k$.

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

Hence, the extent of diffusion is

$$S^*(t^*) = \frac{S_0}{k}\left[b_j\left(\theta_i = h|m, 0\right) - b_j\left(\theta_i = h|\varnothing, 0\right)\right] \tag{10}$$

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:* $\frac{dS^*(\vec{\varphi})}{d\varphi_h} \geq 0$, $\frac{dS^*(\vec{\varphi})}{d\varphi_l} \leq 0$ *where the inequalities are strict if* $\varphi_l < 1$ *and* $\varphi_h < 1$ *respectively.*

From Equation (10) we can see that the total diffusion is the product of the proportion of the population who acquire the information (S_0) and the instantaneous signaling benefit ($b_j\left(\theta_i = h|m, 0\right) - b_j\left(\theta_i = h|\varnothing, 0\right)$) at $t = 0$. The latter term can also be interpreted as the informativeness of word of mouth as a signal of h type; that is the difference in the posterior belief following a message versus no message. Increasing φ_h increases both the inital 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, decreasing φ_l increases the informativeness of word of mouth as a signal of h type. However, decreasing φ_l also decreases the initial spread of information. In our model, the former effect dominates the latter: the diffusion of the firm's message is decreasing in φ_l. Hence, increasing the initial asymmetry between the two types benefits the firm in the long run by maximizing the over-all diffusion of information. An immediate result of the Proposition is that the diffusion of information is maximized at $\varphi_h = 1$ and $\varphi_l = 0$ in the partial equilibrium where we do not consider the incentives for consumers to acquire information at $t = -1$. Next, we show that maximal asymmetry remains the optimal solution in the full equilibrium.

Consumers' Incentives to Acquire Information

We solve for the 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, then the consumer will acquire information. For simplicity, we assume no time discounting for consumers.[15] Denoting the time at which the diffusion process ends by t^*, the

word of mouth upon reaching $t^*(\vec{\varphi})$.

[15] None of the results hinge on this assumption. The instantaneous signaling utility at time t decreases as information diffuses in the population. Adding time discount further reduces this instantaneous utility, but does not change our main results.

total signaling benefit for an agent is then[16]

$$V = \lambda \left(\int_0^{t^*} \left(\frac{1 - S(t)}{1 - S_0} \right) (\Delta U(t) - k)\, dt \right) \tag{11}$$

where the first term $\frac{1-S(t)}{1-S_0}$ is the probability of remaining uninformed at time t for an individual uninformed at time 0 and the second term, $\Delta U(t) - k$, is the signaling benefit at each moment of time. We can further simpify the expression to obtain the following:[17]

$$V = \left(\frac{k}{1 - S_0} \right) \left(\frac{S^* - S_0}{S_0} + \ln \frac{S_0}{S^*} \right) \tag{12}$$

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, φ_l. Thus, $V = V(\overrightarrow{\varphi})$. As we stated previously, the firm can increase the costs of the high and low types for acquiring information by $v_l \geq 0$ and $v_h \geq 0$ to create asymmetry between two types, enabling a signaling equilibrium where word of mouth can serve as a signaling device. At $t = -1$, the customer i of type θ chooses to acquire information if $c_i + v_\theta \leq V(\overrightarrow{\varphi})$. Since c_i is assumed to be i.i.d. uniformly on $[0, \bar{c}]$, the proportion of consumers of type θ who choose to acquire information is

$$\varphi_\theta = \Lambda(V(\overrightarrow{\varphi}), v_\theta) = \begin{cases} 0 & if\ V - v_\theta < 0 \\ \frac{V - v_\theta}{\bar{c}} & if\ 0 \leq V - v_\theta \leq \bar{c} \\ 1 & if\ V - v_\theta > \bar{c} \end{cases} \tag{13}$$

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

3.3 Optimal Information Release Strategy

The firm's optimization problem is

$$\max_{v_h, v_l} S^*(\vec{\varphi})$$

[16] More precisely, the total signaling benefit is the difference between the expected benefit with and without information at $t = -1$. The expected benefit with information at $t = -1$ is $W^{acq}(\vec{\varphi}) = \lambda \left(\int_0^{t^*} (\Delta U(t) - k)\, dt \right)$. Even if the consumer does not acquire the information initially, she may still acquire it through others' word of mouth and therefore, the expected benefit without information at $t = -1$ is $W^{no\,acq}(\vec{\varphi}) = \int_0^{t^*} \lambda \frac{S(t)(1-S(t))}{1-S_0} \int_t^{t^*} \lambda(\Delta U(\tau) - k) d\tau dt = \frac{\lambda}{1-S_0} \int_0^{t^*} (\Delta U(t) - k)[S(t) - S_0] dt$. Therefore, the total signaling benefit for an agent is $V = W^{no\,acq}(\vec{\varphi}) - W^{no\,acq}(\vec{\varphi}) = \frac{\lambda}{1-S_0} \int_0^{t^*} (\Delta U(t) - k)((1 - S_0) - [S(t) - S_0]) dt = \int_0^{t^*} \frac{\lambda(1-S(t))}{1-S_0} (\Delta U(t) - k) dt$.

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

subject to

$$\begin{aligned}\varphi_h =& \Lambda(V(\overrightarrow{\varphi}), \upsilon_h)\\ \varphi_l =& \Lambda(V(\overrightarrow{\varphi}), \upsilon_l)\\ \upsilon_h \geq& 0\\ \upsilon_l \geq& 0\end{aligned}$$

Before finding the optimal strategy for the firm, we note that asymmetry is a necessary condition for word of mouth to take place.

Lemma 1. *No word of mouth occurs under symmetric costs.*

The underlying driver for the word of mouth diffusion is the signaling benefit which arises from the asymmetry between the high and low types ($\varphi_h > \varphi_l$) in Equation (10). It is obvious that under symmetric costs (which leads to $\varphi_h = \varphi_l$), $S^* = 0$. Therefore, no word of mouth arises.

The following Proposition characterizes the optimal strategy (optimal level of asymmetry) that maximizes information diffusion in the full equilibrium.

Proposition 2. *The optimal strategy for the firm is to set a sufficiently large cost for the low type such that $\varphi_l = 0$, and minimize the costs to the high type, $\upsilon_h = 0$ such that $0 < \varphi_h \leq 1$. Moreover, $\varphi_h < 1$ if $\overline{c} \geq \frac{1-k+k\ln k}{1-\alpha}$ and $\varphi_h = 1$ if $\overline{c} < \frac{1-k+k\ln k}{1-\alpha}$.*

We find that the optimal strategy is to restrict information to the low-type consumers by setting the cost of information to be high while minimizing the costs for the high-type consumers. The result demonstrates that the firm benefits from maximal asymmetry in initial information acquisition between the two types. In the model, 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/learn information about the product in equilibrium. 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, through 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. We also note that our "self-enhancement" mechanism is consistent with many of the product categories where we tend to observe significant word of mouth occurring, such as fashion, entertainment and dining. From casual observation these are also categories where it is often perceived to be desirable to be knowledgeable about products in these categories.

Clearly, waiting for information to spread through word of mouth takes time, and the firm may not be in a position to be patient. For example, a newly-released movie typically stays at a multiplex

cinema for only one to three weeks, and thus the timing of the information diffusion as well as the amount of diffusion can be critical. We revisit the firm's optimization problem with the discount factor $\beta^t = \exp(-rt)$, where r is the discount rate. In this case, the firm's objective function becomes:

$$\max_{\nu_h,\nu_l} \left\{ S(0) + \int_0^{t^*} \frac{dS}{dt} e^{-rt} dt \right\}$$

We find that when the firm is not too patient (r is sufficiently large), it may want to disseminate the information even to the low type customers by choosing a low enough ν_l such that some low types acquire information ($\varphi_l > 0$) to achieve a higher level of information diffusion at an early stage.

Proposition 3. *When the discount rate r is large enough, the firm may choose an intermediate level of costs for the low type such that $\varphi_l > 0$.*

The Proposition demonstrates the tradeoff a firm may face between spreading information quickly versus maximizing the spread of information: by confining the initial acquisition of information to high type consumers only, it maximizes an individual's incentive to engage in word of mouth but the process of information diffusion takes a longer time. On the other hand, by allowing some low types to also gain access to the information the firm may achieve a greater level of initial adoption. However the incentives to engage in word of mouth are reduced and the final extent of the information diffusion is smaller. When the firm is sufficiently impatient, it prefers to initially allow the information to be more widely accessible to consumers than to wait for a wider diffusion by further restricting access to information. This trade-off is illustrated in Figure 1. Allowing some low types access to information ("telling more people") dilutes the signaling value of word of mouth (the upper graph in Figure 1) and results in shorter period of diffusion ($t^* < t^{**}$) and a lower level of over-all diffusion ($S^* < S^*$) compared to the case of "telling few people" (only high types access to the information). However, it does yield a higher level of early diffusion (S_0'), which may be particularly valuable to the firm.

4 Advertising

In Section 3, we find that the firm optimally restricts initial access to information in order to increase word of mouth among consumers. In this Section, we show that attempts by the firm to jump-start the diffusion process through traditional marketing actions such as advertising lowers the signaling value of word of mouth to the consumer. As a consequence, "rushing" diffusion through advertising crowds out the incentives for individuals to acquire information. We add to our basic set-up in Section 3 a simple advertising technology, which is costly to the firm, that exposes consumers early on to information about the product. Here, we simplify the exposition by reverting to the setting with no discounting and by assuming that the firm chooses υ_h and υ_l optimally such that $\upsilon_h = 0$ and υ_l is large so that

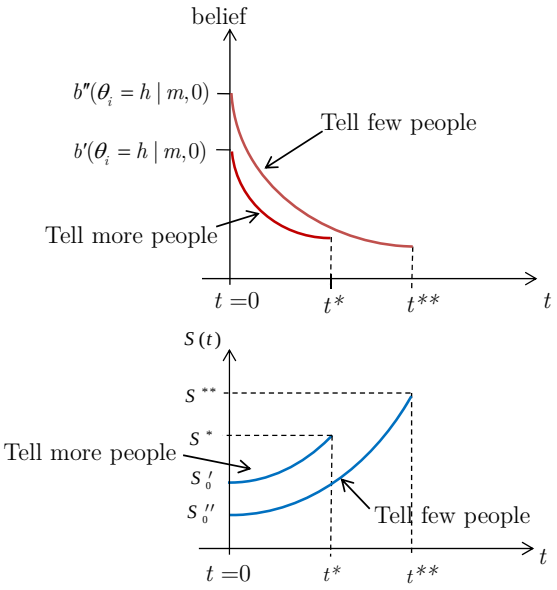


Figure 1: Tradeoff between Spreading Information Quickly versus Maximizing the Total Spread

$\varphi_h \geq 0$ and $\varphi_l = 0$.

Timing

At $t = -1$ consumers choose whether to search for information. At $t = 0$ the firm exposes a fraction β of the consumer population to the advertising message about the product,[18] and from then on consumers engage in word of mouth as before. Advertising is costly to the firm: $C(\beta) \geq 0, C'(\beta) > 1, C''(\beta) > 0$ for all $\beta \geq 0$. Our assumption that the marginal cost of advertising at $\beta = 0$ 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 (β) is not observable to the consumers, who rather infer it in equilibrium. The word of mouth generation process that occurs at $t > 0$ is the same as in Section 3, with the only exception that consumers' inference and optimal stopping strategy is now also conditional on their belief on 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 where, 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. We focus on the signaling equilibrium which results in the largest diffusion of information

[18] We model advertising as a one-time pulse is for simplification purposes. We discuss alternative assumptions such as a continuous advertising technology which gradually informs people over time and alternative timing such as the advertising occuring at or prior to $t = -1$.

subject to the refinements described below. The equilibrium is described by the firm's and consumers' strategies $\{\varphi_h^*, \beta^*, \tau_i^*\}$ and their beliefs $\left\{\tilde{\beta}_i(a_i), b_j^*\left(\theta_i = h|m,t\right), b_j^*\left(\theta_i = h|\varnothing,t\right)\right\}$. Here, φ_h^* is the fraction of high-type consumers who choose to search for information at $t=-1$, β^* 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) $\tilde{\beta}_i\left(a_i\right)$, 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.

In any equilibrium where $0<\beta^*<1$ receiving or not receiving an ad are both consistent with the firm's equilibrium strategy. Thus a consumer's beliefs are the same in both scenarios $\tilde{\beta}_i\left(a_i=0\right) = \tilde{\beta}_i\left(a_i=1\right) = \beta^*$. For example, if the consumer believes that the firm sent out an ad to 10% of the population, the fact that she did or did not receive an ad does not change her prior belief. In contrast, in the case of $\beta^*=1$, not seeing an ad ($a_i=0$) is not on the equilibrium path, as is the case for $\beta^*=0$ and exposure to the ad ($a_i=1$). In these instances we impose a trembling hand refinement on the set of equilibria, which is defined in more detail in Appendix B. We assume that there is a tremble associated with advertising. That is, when the firm chooses a level $\beta \in \{0,1\}$, the actual fraction that receive the advertisement is $\beta\left(1-\epsilon\right)+\epsilon\left(1-\beta\right) = \beta+\epsilon-2\epsilon\beta$. We show that the limit $\epsilon \to 0$ of these "trembling" equilibria corresponds to the signaling equilibrium we find here.

As before, if there are multiple solutions to the consumer's problem, we assume that the firm can implement the solution which results in the greatest level of information diffusion. The beliefs are Bayesian on the equilibrium path for $t \leq \tau^*$ and satisfy $b_j^*\left(\theta_i = h|m,t\right) - b_j^*\left(\theta_i = h|\varnothing,t\right) < k$ for all $t > \tau^*$. Finally, we maintain the assumptions that $\overline{c} \geq \frac{1-k+k\ln k}{1-\alpha}$, 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.

4.1 Word of Mouth Signaling Continuation Game ($t \geq 0$)

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 φ_h and the beliefs of consumers about the level of advertising undertaken by the firm $\tilde{\beta}$. In particular, $\tilde{\beta}$ affects how the signaling value of information $\Delta U\left(t\right)$ evolves over time, and hence how long consumers will continue to spread the

information until $\Delta U(t) = k$. We denote the consumers' inference of 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 $\tau\left(\varphi_h, \tilde{\beta}\right)$ is given by

$$\tau\left(\varphi_h, \tilde{\beta}\right) = \frac{1}{\lambda} \ln \frac{\tilde{S}_c^*}{1 - \tilde{S}_c^*} \frac{1 - \tilde{S}_0}{\tilde{S}_0} \tag{14}$$

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

$$\tilde{S}_0\left(\varphi_h, \tilde{\beta}\right) = \varphi_h \alpha + \tilde{\beta}\left(\left(1 - \varphi_h\right)\alpha + 1 - \alpha\right)$$

The effect of advertising is to increase the initial fraction of population with information at $t = 0$ from $S_0 = \varphi_h \alpha$ in the basic model with no advertising (where $\varphi_l = 0$) to $\tilde{S}_0 = \varphi_h \alpha + \beta\left(\left(1 - \varphi_h\right)\alpha + 1 - \alpha\right)$ 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 the initial level of exposure prior to word of mouth diffusion. The conjectured level of total diffusion $\tilde{S}_c^*(\varphi_h, \tilde{\beta})$ is determined by when consumer beliefs result in the signaling value of passing on information being equal to the cost of passing on information $\Delta U(t) = k$. Interestingly, for a given level of initial information acquisition by consumers φ_h, the consumers' conjectured level of total diffusion is independent of the consumers' conjectured level of advertising $\tilde{\beta}$, provided that $\tilde{\beta}$ is not so large that no word of mouth takes place ($\tilde{\beta} < \bar{\beta} \equiv \frac{\varphi_h \alpha}{k(1-\varphi_h \alpha)}\left(1 - k - \frac{(1-\varphi_h)\alpha}{1-\varphi_h \alpha}\right)$).

Lemma 2. *For a given amount of initial information acquisition through search* (φ_h)*, the consumers' conjectured level of information diffusion* $(\tilde{S}_c^*)$ *is independent of the amount of the conjectured level of advertising* $(\tilde{\beta})$ *for all* $0 \leq \tilde{\beta} \leq \bar{\beta}$ *:* $\tilde{S}_c^*(\varphi_h, \tilde{\beta}) = \tilde{S}_c^*(\varphi_h, 0)$.

The incentives to engage in word of mouth at any point in time is governed by the signaling value of the information. The signaling value in turn 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. Therefore, for a given level of initial information acquisition through consumer search, both advertising and word of mouth affect the signaling value by diffusing the information through the population. The mechanics of diffusion is the same across advertising and word of mouth; in both information channels individuals are informed randomly - individuals receive information irrespective of their type. Since 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, the evolution of the consumer's belief, as a function of the set of people who are informed, is the same when diffusion occurs through either channel. 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 φ_h, 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.

Lemma 3. *For a given amount of initial information acquisition through search* (φ_h)*, the duration of word of mouth is decreasing in the conjectured level of advertising*: $\frac{d\tau}{d\tilde{\beta}} < 0$.

The conjectured level of advertising $\tilde{\beta}$ has no direct effect on the extent of the diffusion $\tilde{S}_c^*$ and thus has no effect on τ through this term. The only effect comes through the initial level of information $\tilde{S}_0(\varphi_h, \tilde{\beta})$ which is increasing in $\tilde{\beta}$. The time $\tau\left(\varphi_h, \tilde{\beta}\right)$ is strictly decreasing in $\tilde{S}_0$ and is thus strictly decreasing in the conjectured level of advertising $\tilde{\beta}$.

Firm's Advertising Decision

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 of 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:

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

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

$$\tilde{S}_f^*(\varphi_h, \beta, \tilde{\tau}) = \frac{1}{1 + ae^{-\lambda\tilde{\tau}}}, \quad \text{where } a = \frac{1 - S_0(\varphi_h, \beta)}{S_0(\varphi_h, \beta)}. \tag{15}$$

For a given level of initial information acquisition by consumers φ_h, a firm is able to influence $\tilde{S}_f^*$ through advertising activity which changes the informed share of the population $S_0(\varphi_h, \beta)$ at $t = 0$.

Lemma 4. When the marginal cost of advertising $C'(0)$ is not too large,[19] there exists a cutoff $\hat{\tau}$ such that for all $0 \leq \tilde{\tau} \leq \hat{\tau}$, the optimal level of advertising is $\beta^* = 0$, and for all $\tilde{\tau} \geq \hat{\tau}$, the firm's optimal choice of advertising is $\beta^* > 0$, where $\frac{\partial \beta^*}{\partial \tilde{\tau}} > 0$.

The actual amount of advertising does not affect the length of time that word of mouth occurs because it is not directly observable to consumers. The length of time is only affected by consumers' own expectations of the level of advertising $\tilde{\beta}$ in equilibrium. Thus, advertising increases the number of individuals engaging in word of mouth for a fixed amount of time $\tilde{\tau}$. Hence, when the conjectured

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

length of time $\tilde{\tau}$ is smaller than a cutoff $\bar{\tau}$, it is not optimal to undertake any costly advertising. The firm will only undertake a strictly positive amount of advertising if its conjecture is above this level. Moreover, in the case where advertising is positive ($\beta^* > 0$), the best response function is increasing in the conjectured amount of time consumers engage in word of mouth $\frac{\partial \beta^*}{\partial \tilde{\tau}} > 0$ and satisfies the first order condition $C'(\beta) = \frac{d\tilde{S}_f^*}{d\beta} = \frac{d\tilde{S}_f^*}{\partial S_0}\frac{\partial S_0}{d\beta}$. 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 $t = 0$. Given the actual level of information acquisition φ_h, and the firm's conjectured length of time that consumers will engage in word of mouth $\tilde{\tau}$, the firm chooses an optimal level of advertising that balances the marginal costs of advertising $C'(\beta)$ against the marginal impact on the firm's conjectured level of diffusion $\frac{d\tilde{S}_f^*}{d\beta}$.

Next, we find the equilibrium of the word of mouth signaling continuation game ($t \geq 0$).

Equilibrium of the Word of Mouth Signaling Game $t \geq 0$

The equilibrium of the continuation game for a given level of information acquisition is a pair of strategies $\{\tau^*(\varphi_h), \beta^*(\varphi_h)\} = \{\tau(\varphi_h, \beta^*(\varphi_h)), \beta(\varphi_h, \tau^*(\varphi_h))\}$ and beliefs $\Big\{\beta^*, b_j^*(\theta_i = h|m, t, \varphi_h, \beta^*), b_j^*(\theta_i = h|\varnothing, t, \varphi$
The firm and consumers choose best responses to each other's actions and beliefs are correct.

Proposition 4. *For a given level of information acquisition φ_h, there is a unique PBNE $\{\beta^*(\varphi_h), \tau^*(\varphi_h)\}$ for the word of mouth signaling game $t \geq 0$.*[20]

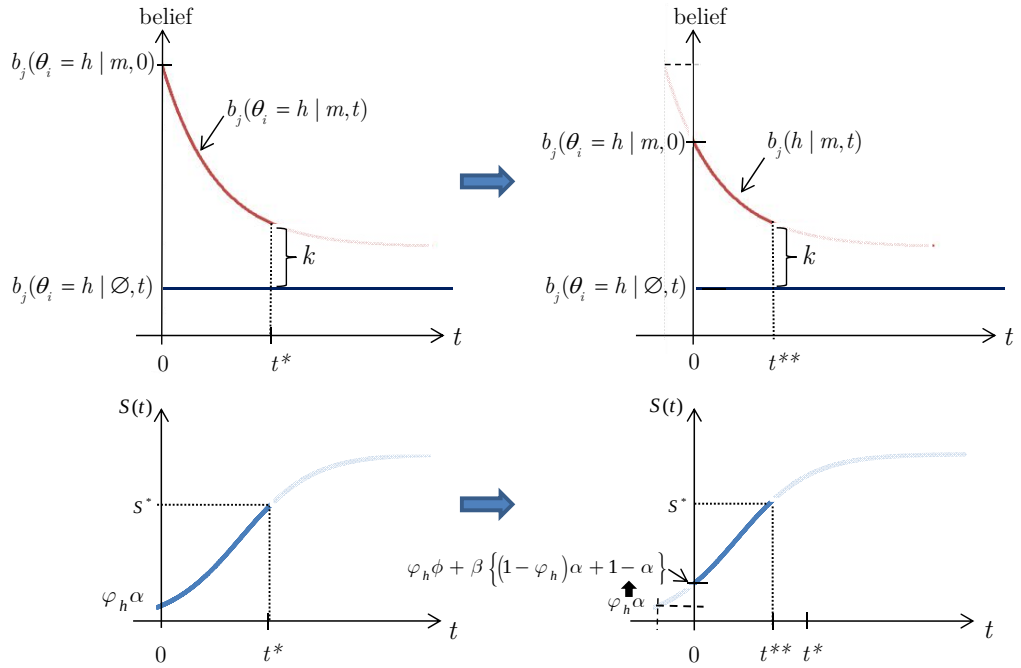
We further characterize the equilibrium of this continuation game for a given φ_h in the following lemma.

Lemma 5. *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 costs of advertising are larger, the level of advertising is lower and the length of diffusion is longer : $\beta_1^*(\varphi_h) \geq \beta_2^*(\varphi_h)$ and $t_1^*(\varphi_h) \leq t_2^*(\varphi_h)$.*

In equilibrium, both the firm and the consumers correctly anticipate the equilibrium extent of diffusion $S^*(\varphi_h)$, thus $\tilde{S}_c^*(\varphi_h, \beta^*) = \tilde{S}_f^*(\varphi_h, \tau^*, \beta^*) = S^*(\varphi_h)$. We also note that for a given level of information acquisition φ_h, the equilibrium level of advertising β^* does not change the equilibrium extent of diffusion $S^*(\varphi_h)$, which would have been reached without advertising. Consumers' expectations of the level of advertising are correct in equilibrium and from Lemma 2, $\tilde{S}_c^*(\varphi_h, \beta^*)$ is independent of β^*. Therefore, the equilibrium extent of diffusion $S^*(\varphi_h)$ is independent of the level of β^*, and only depends on the level of initial information acquisition of consumers φ_h.

Figure 2 illustrates the difference in information diffusion with and without advertising conditional on a given level of information acquisition. The top two panels show the evolution of beliefs over time.

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



(a) Without advertising (b) With advertising

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

Advertising does not change the extent of total diffusion S^*, but it expedites the diffusion process by increasing the initial level of informed individuals at $t = 0$ from $\varphi_h\alpha$ to $\varphi_h\alpha + \tilde{\beta}\left((1-\varphi_h)\alpha + 1 - \alpha\right)$. Thus, the word of mouth reaches the level of S^* at which individuals stop spreading the information (i.e., $b_j(\theta_i = h|m,t) - b_j(\theta_i = h|\varnothing,t) = k$) earlier with advertising than without advertising ($t^{**} < t^*$).

In both cases, we see that the diffusion stops when the difference in beliefs is equal to k . It also shows that advertising clearly dilutes the signaling value of information by randomly distributing information to more consumers, which reduces the asymmetry in information spread between high and low types at $t = 0$. But it does so exactly the same way as would otherwise have taken place through word of mouth and thus, beliefs are lower at each moment in time. Hence, advertising does not change the extent of total information diffusion (S^*), but the diffusion occurs for a shorter length of time. We emphasize that this result holds for a fixed level of information acquisition φ_h. In the full game, where the level of information acquisition is endogenous, advertising will affect the ex ante incentive for consumers to acquire the information, and thus the level of information acquired by consumers at $t = -1$. This, in turn, affects the extent of total information diffusion (S^*).

4.2 Perfect Bayesian Equilibrium of Full Game from $t = -1$

We now consider the ex ante incentives for consumers to acquire information at $t = -1$. 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 5. *Consider two cost functions of advertising C_1 and C_2 where $C_1' < C_2'$ and $C_1'(0)$ is not too large.*[21] *The level of information acquisition is greater and thus, the extent of information diffusion is 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 can lead to an unambiguously greater diffusion of information for the firm. When the cost of advertising is greater, consumers anticipate that the level advertising will be lower and the length of the diffusion will be longer for a given amount of information acquired φ_h (Lemma 5). Although advertising does not affect the level of diffusion for a given level of φ_h (Lemma 2), it reduces the duration that consumers continue to spread the information (Lemma 3). In this way, advertising removes signaling opportunities for individuals since it substitutes for word of mouth and thus reduces the overall benefit to an agent from acquiring information ex-ante. Hence, higher advertising costs results in an equilibrium with lower level of advertising and more information acquisition φ_h^*. [22]

Proposition 5 shows that a commitment not to advertise through higher costs is valuable to the firm. Next, we show that the release date of the product is a source of such commitment for a firm. For instance, film studios release movies on certain holidays during the year, and technology companies often release information and announce the future release date for the product concurrently. The key is that the cost of advertising at the time of information release in dollars calculated at the product release date is increased by e^{rT} where T is the amount of time between the information release and the product release and r is the interest rate. We can couch the determination of T as a mechanism design problem for the monopolist. Provided that the monopolist chooses a large enough T, then this can serve as a credible commitment for the firm not to advertise and maximize the diffusion of information due to word of mouth.

Proposition 6. *There exists a $\bar{T}$ such that for $T > \bar{T}$ the firm undertakes no advertising, and the maximum possible diffusion occurs.*

This Proposition highlights that early information release can serve as a commitment not to undertake advertising during the period when individuals engage in word of mouth. Of course, the firm may also undertake advertising upon the product being released which would not affect the word of mouth, if it occurs after the diffusion has stopped. The result is consistent with the observation

[21] When $C'(0)$ is so large that $C_1'(0) > \bar{C}$, the equilibrium level of advertising $\beta^* = 0$ for both cost functions. We define the cutoff $\bar{C}$ more precisely in the proof in the Appendix.

[22] It is important to note that the decrease in φ_h is not just mechanical outcome obtained from our particular setup of advertising timing. In the current setup, advertising takes place at $t = 0$ only after consumers' information acquisition. Hence, consumers anticipating the possbility of acquiring information through advertising, have less incentives to acquire information ex ante, which is present in the current model. However, on top of this mechanical effect of advertising, there exists more robust effect of decreasing value of acquiring information from removing the signaling opporunity, which we elaborated in the text. We dicuss the robustness of this result to the advertising timing in more detail subsequently.

that delaying advertising may promote word of mouth generation. For example, in an influential 2000 Harvard Business Review article, Dye 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.

Robustness to Advertising Timing

While we feel the assumption regarding the timing of advertising is the most reasonable given our application, there are a number of ways for modeling the timing of advertising in our setting all of which lead to similar qualitative conclusions. For instance, one might consider making the advertising decision occur before consumers acquire information at $t = -2$, simultaneously at the moment of information acquisition $t = -1$, or modeling advertising as a continual process over time during the word of mouth diffusion $t \geq 0$. We argue that in all of these cases, the qualitative nature of advertising is the same; namely that advertising substitutes for word of mouth, thereby removing signaling opportunities and reducing the value of acquiring information.

First, when advertising takes place prior to information acquisition, there is no mechanical effect of reducing the incentives for consumers to acquire information from anticipating the possibility of receiving information through advertising. In this case, even though the overall level of advertising remains unobservable, a consumer may condition the information acquisition decision on whether or not he/she received information via advertising. Hence, the fraction of the population with information at $t = 0$ (S_0) would change. This in turn affects the value of acquiring information $V = \frac{k}{1-S_0}\left[\frac{S^*-S_0}{S_0} - \ln\frac{S^*}{S_0}\right]$. The cutoff type which acquires information satisfies:

$$\varphi_h \bar{c} = \frac{k}{1-S_0}\left[\frac{S^*}{S_0} - 1 - \ln\frac{S^*}{S_0}\right]$$

where S_0 and S^* are the same expressions as earlier. The term $\frac{k}{1-S_0}\left[\frac{S^*}{S_0} - 1 - \ln\frac{S^*}{S_0}\right]$ is decreasing in the conjectured level of advertising (we show this in the proof of proposition 5 in the Appendix). 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 same way as earlier, the extent of diffusion in equilibrium is entirely pinned down by the level of information acquisition. It is thus larger when the costs of advertising are higher.

Second, when advertising occurs simultaneously with information acquisition at $t = -1$ the analysis is unchanged. The only change that could occur in this case is if an individual consumer strategically chose to acquire or not acquire information in order to affect the advertising decision of the firm.

However, individuals are infintessimally small in the model and there is no such incentive.

Finally, there are a number of ways one might model a firm advertising gradually over time. However, provided that the model generates the prediction that lower costs of advertising leads to greater advertising in the word of mouth continuation game, this will reduce the value of acquiring information. Hence a lower fraction of high types will acquire information in equilibrium and the extent of the diffusion will be lower.

5 Targeting Well-Connected Individuals

A large and growing literature has emphasized the role of a small number of key individuals (referred to variously as "social hubs," "network connectors," "opinion leaders," "influentials," or "mavens") on information diffusion. This idea goes back to Katz and Lazarsfeld (1955) and Lazarsfeld, Berelson, and Gaudet (1968). More recently the asymmetry of influence has been studied across a variety of social settings (Weimann 1994 and Gladwell 2000) and particularly in the context of marketing products (e.g. Coulter, Feick and Price 2002, and Van den Bulte and Joshi 2007). The idea is very attractive for marketing practitioners because it suggests that a significant increase in demand is possible, with limited marketing resources, by targeting a small number of key individuals. Here we point out that, paradoxically, the possibility of targeting well-connected consumers may actually negatively affect consumer beliefs and ultimately decrease the overall amount of diffusion.

To model the idea that that some individuals have more friends or are more social than others, we allow some members of the population to mix at a higher rate. In particular, we assume that an individual mixes at either a high or a low rate, $\lambda_i = \lambda_{high}, \lambda_{low}$ ($\lambda_{high} > \lambda_{low}$), where $\Pr[\lambda_i = \lambda_{high}] = \mu$, and this mixing type λ_i is independent of an individual's social type θ_i. This assumption allows us to separately consider the effect of targeted advertising at well-connected individuals from targeting at socially desirable types, i.e. $\Pr[\lambda_i = \lambda_{high}, \theta_i = h] = \mu\alpha$. We fix the fraction of high types who acquire information at $t = -1$ at φ_h. Finally, we denote the levels of targeted advertising to the high mixing population as β^{High} and the level of untargeted advertising as β^{Untarg}. Thus, the extent of diffusion under no advertising, untargeted advertising, and targeted advertising at high mixing populations can be expressed as $S^*(\varphi_h, 0)$, $S^*\left(\varphi_h, \beta^{Untarg}\right)$, $S^*\left(\varphi_h, \beta^{High}\right)$.

The following Proposition examines the effect of different consumer beliefs about the firm's advertising strategy on the extent of diffusion. That is, we assume that consumer beliefs are consistent with firm advertising strategy in the three different cases and examine how the beliefs impact the extent of diffusion. To facilitate the comparison, we assume that in both targeted and untargeted advertising cases, the firm advertises to the the same number of individuals in total ($\tilde{\beta}^{High} = \frac{\tilde{\beta}^{Untarg}}{\mu}$), and the amount of advertising is such that word of mouth occurs in each case ($\beta \leq \bar{\bar{\beta}} = \frac{\varphi_h\alpha}{1-\varphi_h\alpha}\left[\frac{1-\alpha}{k(1-\varphi_h\alpha)} - 1\right][\lambda_{high}\mu + \lambda_{low}(1-\mu)]$).

Proposition 7. *Suppose the fraction of the population informed through advertising (untargeted or targeted) is the same* $\tilde{\beta}^{High} = \frac{\tilde{\beta}^{Untarg}}{\mu} \leq \bar{\bar{\beta}}$. *For any given* φ_h, *targeted advertising at the high mixing individuals results in the smallest diffusion of information. Furthermore the following relationship holds:*

$$S^*\left(\varphi_h, \tilde{\beta}^{High}\right) < S^*\left(\varphi_h, 0\right) < S^*\left(\varphi_h, \tilde{\beta}^{Untarg}\right).$$

We find that the amount of diffusion is largest under consumer beliefs that the firm engages in untargeted advertising, second largest under no advertising, and smallest under advertising targeted at high mixing types. In particular, note that the result that untargeted advertising generates greater diffusion than no advertising ($S^*\left(\varphi_h, 0\right) < S^*\left(\varphi_h, \tilde{\beta}^{Untarg}\right)$ can be contrasted to the earlier Lemma 2, where in the homogeneous mixing case beliefs by the consumer that advertising takes place did not affect the extent of diffusion for a given φ_h: $\tilde{S}_c^*(\varphi_h, \tilde{\beta}) = \tilde{S}_c^*\left(\varphi_h, 0\right)$.

The intuition for the result in Proposition 7 is the following. The presence of well-connected individuals decreases the signaling value of word of mouth. This is due to the fact that well-connected consumers are more likely to hear about the information through word of mouth since they mix at a higher rate, and, once they gain access to the information, diffuse it further in the population through more frequent social interactions. This of course decreases the signaling value of word of mouth by decreasing asymmetry across the high and low social types. Note that further targeting the well-connected consumers exacerbates this problem. In fact, we find that beliefs by consumers that advertising is targeted at high mixing types offsets word of mouth more than one for one. In contrast, untargeted advertising increases the signaling value of word of mouth. This is due to the fact that the probability that an untargeted ad reaches a well-connected consumer (as opposed to other consumers) is less than the probability that a word of mouth interaction informs a well-connected consumer since, as we argued earlier, word of mouth favors the well-connected types. Since untargeted advertising makes it less likely that information lands in the hands of well-connected consumers relative to word of mouth alone, consumer beliefs that untargeted advertising is taking place offsets word of mouth less than one for one.

One important implication of Proposition 7 is that the possibility of targeting well-connected individuals dilutes the signaling value of word of mouth. Even in the case where the consumer does not observe the firm's targeting strategy directly, she can infer that the firm would choose to target the well-connected types. (For example, suppose that targeting is costless. On the margin the firm prefers to target an additional well-connected consumer (versus other consumers) since the well-connected will engage in more social interactions, resulting in greater diffusion). Hence in equilibrium the consumer will believe that that the firm targets its advertising at the individuals who mix at a high rate. This of course decreases the signaling value of word of mouth and results in less diffusion.

We saw in Section 3 that a greater fraction of high social type individuals obtaining the information prior to diffusion increases the extent of diffusion (Proposition 2). On the other hand, in this Section

we see that consumer beliefs that advertising is targeted at well-connected consumers is detrimental to the diffusion of information. This contrast highlights the economic relevance of recognizing the motivation of individuals to engage in word of mouth, and how different communication strategies impact these incentives. Targeted information release or advertising strategies are effective in so much as they credibly generate (or increase) the asymmetry in information between high and low social types and are detrimental if they decrease this asymmetry.

6 Conclusion

In this paper we study a motive for why individuals engage in word of mouth and how a firm may interact with this motive through its information release and advertising strategies. We develop a model where 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 (Baumeister et al. 1989). A firm maximizes the diffusion of information, by structuring its information release strategy so that the act of passing on information, through word of mouth communication, can serve 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, these 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. Finally, firms are often urged to reach out to "opinion leaders" or "influentials" since they are more likely to talk to others about the product. We revisit this conventional wisdom by allowing heterogeneity of consumer mixing – some individuals meet more people than others. We find that when the mixing rate is uncorrelated with the social type, targeting individuals who mix at a high rate is in fact detrimental to the diffusion of information.

The key lesson from our study is that word of mouth is a subtle process for the firm to influence. Since consumers only engage in word of mouth if it can serve as a signal, if low social types have been given or have similar access to the information as high types, little word of mouth will ensue. What really enables the spread of word of mouth is the existence of asymmetries in information between the two types. 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 $b_j(\theta_i = h|\varnothing, 0)$ and express the total extent of diffusion in terms of φ_h, φ_l $(\varphi_h > \varphi_l)$:

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

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

$$\begin{aligned}
\frac{dS^*}{d\varphi_h} &= \frac{1}{k}\left[\frac{(1-\varphi_l)(1-\alpha)}{((1-\varphi_h)\alpha + (1-\varphi_l)(1-\alpha))^2}\right] \geq 0, \text{ if } \varphi_l < 1 \text{ then } > 0 \\
\frac{dS^*}{d\varphi_l} &= -\frac{\alpha}{k}\frac{(1-\varphi_h)(1-\alpha)}{((1-\varphi_h)\alpha + (1-\varphi_l)(1-\alpha))^2} \leq 0, \text{ if } \varphi_h < 1 \text{ then } < 0
\end{aligned} \tag{16}$$

A.2 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:

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

We proceed by proving the following lemma.

Lemma 6. *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:

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

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 $2S_0-1<0$, which is true for $S_0<\frac{1}{2}$.

If $S_0\geq\frac{1}{2}$, we need 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 $x=\frac{S^*}{S_0}>1$.

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

which is greater than 0 for $x>1$ if $\frac{\ln x}{x-1}-\frac{1}{x}>0\Leftrightarrow\ln x>1-\frac{1}{x}$ which is 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

$$-\frac{\ln\frac{1}{S_0}}{\frac{1}{S_0}-1}<\frac{1}{S_0}-2\Leftrightarrow-\ln y-(y-2)(y-1)<0,\text{ where } y=\tfrac{1}{s_0}.$$

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

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

and note that $\lim_{y\to 1}\left[-\ln y-(y-2)(y-1)\right]=0$.

Hence, $-\ln y-(y-2)(y-1)<0$, which shows that $\frac{\partial V}{\partial S_0}<0$. □

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

Hence,

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

We already know that S^* is maximized when $\varphi_h = 1$, independent of φ_l. However, this may not be a feasible solution if the ex ante information acquisition constraints bind for some high types. When $\varphi_h < 1$, we have that $\frac{dS^*}{d\varphi_l} < 0$. Hence, the optimal policy will result in $\varphi_l = 0$ if φ_l does not increase the ex ante incentives for high type consumers to acquire information; $\frac{dV}{d\varphi_l} < 0$, and $V(\varphi_h, 0) \geq 0$ for $\forall \varphi_h \geq 0$.

Finally, $0 \leq S_0(\varphi_h^*, 0) \leq S^*$ and $\frac{dV}{dS_0} < 0$.

$$\begin{aligned}\lim_{S_0 \to 0} V &= \lim_{S_0 \to 0} \left(\frac{k}{1 - S_0}\right)\left(\frac{S^*}{S_0} - 1 + \ln S_0 - \ln S^*\right) \\ &= k\left(\frac{S^*}{S_0} - 1 - \ln \frac{S^*}{S_0}\right) > 0. \\ \lim_{S_0 \to S^*} V &= 0\end{aligned}$$

Hence, $V\left(\varphi_h^*, 0\right) \geq 0$ for $\forall \varphi_h \geq 0$. This proves that $\varphi_l^* = 0$.

Next, when $\varphi_l^* = 0$;

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

$$\begin{aligned}V(\varphi_h^*, 0) = \left(\frac{k}{1 - S_0}\right)\left(\frac{S^* - S_0}{S_0} + \ln \frac{S_0}{S^*}\right) &= \left(\frac{k}{1 - \varphi_h^* \alpha}\right)\left(\frac{\frac{1}{k}\left[\frac{\varphi_h^*(1-\alpha)}{1-\varphi_h^*\alpha}\right] - \varphi_h^*}{\varphi_h^*} + \ln \frac{\varphi_h^*}{\frac{1}{k}\left[\frac{\varphi_h^*(1-\alpha)}{1-\varphi_h^*\alpha}\right]}\right) \\ = \frac{1-\alpha}{\left(1 - \varphi_h^* \alpha\right)^2} - \left(\frac{k}{1 - \varphi_h^* \alpha}\right)\left(1 - \ln \frac{k\left(1 - \varphi_h^* \alpha\right)}{1 - \alpha}\right) &\qquad (17)\end{aligned}$$

In particular, when $\varphi_h^* = 1$, $V(1, 0) = \frac{1 - k + k \ln k}{1 - \alpha}$. We now verify that an equilibrium exists where $\varphi_h > 0$. First, when $1 \cdot \bar{c} \geq \frac{1 - k + k \ln k}{1 - \alpha}$, we note $\lim_{\varphi_h \to 0} V\left(\varphi_h, 0\right) = k\left(\frac{1-\alpha}{k} - 1 - \ln \frac{1-\alpha}{k}\right) > 0$, and hence there exists $0 < \varphi_h^* < 1$ such that $\varphi_h^* \bar{c} = V\left(\varphi_h^*, 0\right)$. Second, when $1 \cdot \bar{c} < \frac{1 - k + k \ln k}{1 - \alpha}$, the cutoff type is $\varphi_h = 1$ and in this case the optimum only requires that $\varphi_h^* = 1$.

Therefore, the optimal strategy for the firm has the following characteristics: $0 < \varphi_h \leq 1$ and $\varphi_l = 0$. Moreover, $\varphi_h < 1$ if $\bar{c} \geq \frac{1 - k + k \ln k}{1 - \alpha}$ and $\varphi_h = 1$ if $\bar{c} < \frac{1 - k + k \ln k}{1 - \alpha}$.

A.3 Proof of Proposition 3

Let

$$R\left(\varphi_l, \varphi_h\right) = \int_0^{t^*(\varphi_h,\varphi_l)} \left(S_0 + \frac{dS}{dt} e^{-rt}\right) dt$$

We have immediately that

$$\lim_{r\to 0} R\left(\varphi_l, \varphi_h\right) = S^*\left(\varphi_l, \varphi_h\right)$$
$$\lim_{r\to \infty} R\left(\varphi_l, \varphi_h\right) = S_0\left(\varphi_l, \varphi_h\right).$$

When $r = 0$; $R\left(0, \varphi_h\right) > R\left(\varphi_l, \varphi_h\right)$ for all $\varphi_l > 0$ since $\frac{dS^*}{d\varphi_l} \leq 0$ (Equations 16). When $r = \infty$; $R = S_0\left(0, \varphi_h\right) < R = S_0\left(\varphi_l, \varphi_h\right)$ for any $0 < \varphi_h \leq 1$.

Furthermore,

$$\frac{dR\left(\varphi_l, \varphi_h\right)}{dr} = -\int_0^{t^*(\varphi_h,\varphi_l)} \frac{dS}{dt} t e^{-rt} dt < 0 \quad \text{for all } 0 \leq \varphi_l, \varphi_h \leq 1.$$

Hence, for any $\varphi_l > 0$, there exists $r^*(\varphi_l)$ such that for all $r > r^*(\varphi_l)$, $R\left(0, \varphi_h\right) < R\left(\varphi_l, \varphi_h\right)$.

A.4 Proof of Lemma 2

The relationship for the extent of diffusion is then determined by

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

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

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

A.5 Proof of Lemma 3

When $0 \leq \tilde{\beta} \leq \frac{\varphi_h\alpha}{k(1-\varphi_h\alpha)}\left(1 - k - \frac{(1-\varphi_h)\alpha}{1-\varphi_h\alpha}\right)$ the best response time $\tau\left(\varphi_h, \tilde{\beta}\right)$ is given by:

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

where $\tilde{S}_0\left(\varphi_h,\tilde{\beta}\right)=\varphi_h\alpha+\tilde{\beta}\left(\left(1-\varphi_h\right)\alpha+1-\alpha\right)$.

And we know from Proposition 2 that $\frac{d\tilde{S}_c^*}{d\tilde{\beta}}=0$. It is now straightforward to find the derivative is $\frac{d\tau^*}{d\tilde{\beta}}=\frac{d\tau^*}{d\tilde{S}_0}\frac{d\tilde{S}_0}{d\tilde{\beta}}$:

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

A.6 Proof of Lemma 4

The optimal level of advertising is determined by:

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

A firm is able to influence $\tilde{S}_f^*$ through the initial informed share of the population $S_0\left(\varphi_h,\beta\right)$, but cannot directly influence $\tilde{\tau}$, which is only affected by consumers' expectations of the level of advertising $\tilde{\beta}$ in equilibrium. Hence, the marginal effect of advertising on $\tilde{S}_f^*$ is $\frac{d\tilde{S}_f^*}{d\beta}=\frac{d\tilde{S}_f^*}{\partial S_0}\frac{\partial S_0}{d\beta}=\frac{e^{-\lambda\tilde{\tau}}(1-\varphi_h\alpha)}{\left[S_0+(1-S_0)e^{-\lambda\tilde{\tau}}\right]^2}$.

Also note that $\frac{d^2\tilde{S}_f^*}{d\beta^2}=-2\frac{e^{-\lambda\tilde{\tau}}\left(1-e^{-\lambda\tilde{\tau}}\right)(1-\varphi_h\alpha)^2}{\left[S_0+(1-S_0)e^{-\lambda\tilde{\tau}}\right]^3}<0$, and by assumption $C''\left(\beta\right)>0$, so that the objective is strictly concave and a first order condition can be used for interior solutions for $\beta\in(0,1)$. Thus, if there is a β that satisfies equation

$$C'\left(\beta\right)=\frac{d\tilde{S}_f^*}{d\beta}=\frac{d\tilde{S}_f^*}{\partial S_0}\frac{\partial S_0}{d\beta}=\frac{e^{-\lambda\tilde{\tau}}\left(1-\varphi_h\alpha\right)}{\left[S_0+\left(1-S_0\right)e^{-\lambda\tilde{\tau}}\right]^2},\tag{18}$$

then, this is the best response. Also, note that $C'\left(1\right)>1>e^{-\lambda\tilde{\tau}}\left(1-\varphi_h\alpha\right)$ such that $\beta=1$ is never a best response. We now show that there exists a $\bar{\tau}$ such that $\exists\beta$ that satisfies $C'\left(\beta\right)=\frac{e^{-\lambda\tilde{\tau}}(1-\varphi_h\alpha)}{\left[S_0+(1-S_0)e^{-\lambda\tilde{\tau}}\right]^2}$ for $\tau\geq\bar{\tau}$ and $C'\left(0\right)\geq\frac{e^{-\lambda\tilde{\tau}}(1-\varphi_h\alpha)}{\left[\varphi_h\alpha+(1-\varphi_h\alpha)e^{-\lambda\tilde{\tau}}\right]^2}$ for $\tau\leq\bar{\tau}$.

We define $\bar{\tau}$ as the time which satisfies $C'\left(0\right)=\frac{e^{-\lambda\hat{\tau}}(1-\varphi_h\alpha)}{\left[\varphi_h\alpha+(1-\varphi_h\alpha)e^{-\lambda\hat{\tau}}\right]^2}$. At $\tilde{\tau}=0$, the (RHS) is $1-\varphi_h\alpha\leq1$, 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 $C'\left(0\right)<\frac{1}{2\varphi_h\alpha}$, then there exists a cutoff $\bar{\tau}\leq\frac{1}{\lambda}\ln\frac{1-\varphi_h\alpha}{\varphi_h\alpha}$ such that for all $\tilde{\tau}\in\left[0,\bar{\tau}\right),C'\left(0\right)>\frac{e^{-\lambda\tilde{\tau}}(1-\varphi_h\alpha)}{\left[\varphi_h\alpha+(1-\varphi_h\alpha)e^{-\lambda\tilde{\tau}}\right]^2}$, which implies $\beta=0$. Furthermore, for all $\tilde{\tau}\in\left[\bar{\tau},\frac{1}{\lambda}\ln\frac{1-\varphi_h\alpha}{\varphi_h\alpha}\right]$, $C'\left(0\right)\leq\frac{e^{-\lambda\tilde{\tau}}(1-\varphi_h\alpha)}{\left[\varphi_h\alpha+(1-\varphi_h\alpha)e^{-\lambda\tilde{\tau}}\right]^2}$ and hence, $\exists\beta>0$ which satisfies $C'\left(\beta\right)=\frac{e^{-\lambda\tilde{\tau}}(1-\varphi_h\alpha)}{\left[S_0+(1-S_0)e^{-\lambda\tilde{\tau}}\right]^2}$. Otherwise (i.e., $C'\left(0\right)\geq\frac{1}{2\varphi_h\alpha}$), $C'\left(0\right)\geq\frac{e^{-\lambda\tilde{\tau}}(1-\varphi_h\alpha)}{\left[\varphi_h\alpha+(1-\varphi_h\alpha)e^{-\lambda\tilde{\tau}}\right]^2}$ and $\beta=0$ for all $\tilde{\tau}$.

A.7 Proof of Proposition 4

When the best response functions $\left\{\tau\left(\varphi_h,\tilde{\beta}\right),\beta\left(\varphi_h,\tilde{\tau}\right)\right\}$ are continuous and map a compact set to a compact set, then there exists fixed point and hence an equilibrium of the continuation game from $t \geq 0$. By definition β is bounded above by 1, also an upper bound on τ for a given φ_h is $\tau\left(\varphi_h,0\right)=\frac{1}{\lambda}\ln\frac{1-\varphi_h\alpha}{\frac{k}{1-\alpha}-\varphi_h\alpha}$. Both best response functions are continuous over the appropriate domains, so we have established the existence of an equilibrium. We also know that $\tau\left(\varphi_h,\tilde{\beta}\right)=0$ for $\tilde{\beta}\geq\frac{\varphi_h\alpha}{k(1-\varphi_h\alpha)}\left(1-k-\frac{(1-\varphi_h)\alpha}{1-\varphi_h\alpha}\right)$ as this level of advertising will result in $S_0\geq\tilde{S}_c^*$. We note that from Lemma 3 that $\frac{d\tau^*}{d\tilde{\beta}}<0$ for $0\leq\tilde{\beta}\leq\frac{\varphi_h\alpha}{k(1-\varphi_h\alpha)}\left(1-k-\frac{(1-\varphi_h)\alpha}{1-\varphi_h\alpha}\right)$ and $\tau=0$ for $\tilde{\beta}\geq\frac{\varphi_h\alpha}{k(1-\varphi_h\alpha)}\left(1-k-\frac{(1-\varphi_h)\alpha}{1-\varphi_h\alpha}\right)$, also that $\beta\left(\varphi_h,0\right)=0$. So, the equilibrium diffusion time $\tau^*>0$. Thus, at the equilibrium point, $\frac{d\tau^*}{d\tilde{\beta}}<0$. Using Lemma 4, if $C'\left(0\right)\geq\frac{1}{2\varphi_h\alpha}$, then $\beta=0$ for all $\tilde{\tau}$. Hence, the unique equilibrium is $\{\tau^*,\beta^*\}=\left\{\frac{1}{\lambda}\ln\frac{1-\varphi_h\alpha}{\varphi_h\alpha},0\right\}$. We now prove uniqueness for the case where $C'\left(0\right)<\frac{1}{2\varphi_h\alpha}$ by showing that $\frac{d\beta}{d\tilde{\tau}}\geq0$ at any equilibrium point where $\tau^*>0$ and $\beta^*>0$.

Implicitly differentiating the equation (18) to find $\frac{d\beta}{d\tilde{\tau}}$ for $\beta\in(0,1)$:

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

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

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

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 since $\tau\left(\varphi_h,\tilde{\beta}\right)$ and $\beta\left(\varphi_h,\tilde{\tau}\right)$ are also continuous in φ_h for $\varphi_h>0$.

A.8 Proof of Lemma 5

Let $F(\beta)=\frac{e^{-\lambda\tilde{\tau}}(1-\varphi_h\alpha)}{\left[S_0+(1-S_0)e^{-\lambda\tilde{\tau}}\right]^2}$. $\frac{\partial F(\beta)}{\partial\beta}=2\left(1-e^{\lambda\tilde{\tau}}\right)\frac{(1-\varphi_h\alpha)^2e^{-\lambda\tilde{\tau}}}{\left[S_0+(1-S_0)e^{-\lambda\tilde{\tau}}\right]^3}>0$. Again, the best response advertising level β^* satisfies $C'\left(\beta\right)=\frac{e^{-\lambda\tilde{\tau}}(1-\varphi_h\alpha)}{\left[S_0+(1-S_0)e^{-\lambda\tilde{\tau}}\right]^2}$. Then, $C_1'(\beta_1^*)=F(\beta_1^*)<C_2'(\beta_2^*)=F(\beta_2^*)$.

Hence, $F(\beta_1^*) < F(\beta_2^*) \Leftrightarrow \beta_1(\varphi_h, \tilde{\tau}) > \beta_2(\varphi_h, \tilde{\tau})$ for all $\tilde{\tau} > \hat{\tau}_1$. From Lemma 3 and 4, if $\tau(\varphi_h, 0) = \frac{1}{\lambda} \ln \frac{1-\varphi_h\alpha}{\frac{k}{1-\alpha}-\varphi_h\alpha} > \hat{\tau}$, then, $\beta_1^*(\varphi_h) > \beta_2^*(\varphi_h)$ and $\tau_1^*(\varphi_h) < \tau_2^*(\varphi_h)$. Otherwise, $\beta^*(\varphi_h) = \beta_2^*(\varphi_h) = 0$ 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 $C_1'(0) < \frac{1-\alpha}{k}\left[1 - \frac{\varphi_h\alpha(1-\alpha)}{k}\right]$.

A.9 Proof of Proposition 5

It is useful to define the equilibrium level of information acquisition without advertising φ_h^{**} (from Equation 17) , which satisfies the following condition:

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

We begin with the following Lemma:

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

$$V\left(\varphi_h, \tau_1^*(\varphi_h), \beta_1^*(\varphi_h)\right) < V\left(\varphi_h, \tau_2^*(\varphi_h), \beta_2^*(\varphi_h)\right).$$

Proof. We note that from Lemma 5, $\beta_1^*(\varphi_h) > \beta_2^*(\varphi_h)$ and $\tau_1^*(\varphi_h) < \tau_2^*(\varphi_h)$, and from Lemma 2, $S_1^* = S_2^*$. From Equation (12), we can now write $V\left(\varphi_h, \tau^*(\varphi_h), \beta^*(\varphi_h)\right)$ as:

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

Taking the derivative with respect to β^* holding S_0 and S^* constant:

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

And from Lemma 6 in the proof of Proposition 2, $\frac{\partial V}{\partial S_0} < 0$, and the lemma follows immediately by noting that $\frac{dS_0}{d\beta^*} > 0$ and $\frac{dS_0}{d\tau^*} = 0$. □

Now, $V\left(\varphi_h, \tau^*(\varphi_h, \beta^*), \beta^*(\varphi_h, \tau^*)\right)$ is continuous in φ_h and $\bar{c} \geq \frac{1-k+k\ln k}{1-\alpha} = V\left(1, \tau^*(1,0), 0\right)$. Hence, it follows that $1 > \varphi_{h2}^* > \varphi_{h1}^*$. Also, $S_1^* < S_2^*$ follows from recalling that $\frac{dS^*}{d\varphi_h} > 0$ from equation (16). Finally, if $C_1'(0) \geq \bar{C} = \frac{S^*\left(\varphi_h^{**},0\right)\left(1-S^*\left(\varphi_h^{**},0\right)\right)}{S_0\left(\varphi_h^{**},0\right)\left(1-S_0\left(\varphi_h^{**},0\right)\right)}\left(1-\varphi_h^{**}\alpha\right)$, then the equilibrium for both cost functions is the same, $\{\varphi_h^{**}, \tau^*(\varphi_h^{**}), 0\}$.

A.10 Proof of Proposition 6

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

Note that by choosing $T > \tau^*\left(\varphi_h^{**}\right)$, this will ensure that the diffusion of word of mouth is completed prior to the product being released 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 $e^{rT}C'(\beta)$. And when $T > \frac{1}{r}\ln\frac{S^*\left(\varphi_h^{**},0\right)\left(1-S^*\left(\varphi_h^{**},0\right)\right)}{S_0\left(\varphi_h^{**},0\right)\left(1-S_0\left(\varphi_h^{**},0\right)\right)}\frac{\left(\left(1-\varphi_h^{**}\right)\alpha+1-\alpha\right)}{C'(0)}$, we have $e^{rT}C'(0) > \frac{S^*\left(\varphi_h^{**},0\right)\left(1-S^*\left(\varphi_h^{**},0\right)\right)}{S_0\left(\varphi_h^{**},0\right)\left(1-S_0\left(\varphi_h^{**},0\right)\right)}\left(1-\varphi_h^{**}\alpha\right)$ and we have that $\beta = 0$ and $\varphi_h = \varphi_h^{**}$.

A.11 Proof of Proposition 7

We denote the mass of informed individuals who are high and low mixing by $S_h(t) \in [0,\mu]$ and $S_l(t) \in [0,1-\mu]$. The diffusion is governed by rate of change of these populations. This is given by

$$\begin{aligned}
\frac{dS_{high}}{dt} &= \lambda_{high}\mu\left(1-\frac{S_{high}}{\mu}\right)\left(\frac{\lambda_{high}S_{high}}{\lambda_{high}\mu+\lambda_{low}(1-\mu)}+\frac{\lambda_{low}S_{low}}{\lambda_{high}\mu+\lambda_{low}(1-\mu)}\right)\\
&= \frac{\lambda_{high}\mu}{\lambda_{high}\mu+\lambda_{low}(1-\mu)}\left(1-\frac{S_{high}}{\mu}\right)\left(\lambda_{high}S_{high}+\lambda_{low}S_{low}\right),\\
\frac{dS_{low}}{dt} &= \lambda_{low}(1-\mu)\left(1-\frac{S_{low}}{1-\mu}\right)\left(\frac{\lambda_{high}S_{high}}{\lambda_{high}\mu+\lambda_{low}(1-\mu)}+\frac{\lambda_{low}S_{low}}{\lambda_{high}\mu+\lambda_{low}(1-\mu)}\right)\\
&= \frac{\lambda_{low}(1-\mu)}{\lambda_{high}\mu+\lambda_{low}(1-\mu)}\left(1-\frac{S_{low}}{1-\mu}\right)\left(\lambda_{high}S_{high}+\lambda_{low}S_{low}\right).
\end{aligned}$$

Then, the total informed population evolves according to

$$\frac{dS_{high}}{dt}+\frac{dS_{low}}{dt} = \left(\lambda_{high}S_{high}+\lambda_{low}S_{low}\right)\left(1-\frac{\lambda_{high}S_{high}+\lambda_{low}S_{low}}{\lambda_{high}\mu+\lambda_{low}(1-\mu)}\right).$$

The ratio of the rates of change can be found by

$$\frac{dS_{high}}{dS_{low}} = \frac{\lambda_{high}(\mu-S_{high})}{\lambda_{low}(1-\mu-S_{low})}.$$

Solving this, we find

$$\begin{aligned}
\frac{dS_{high}}{\lambda_{high}(\mu-S_{high})} = \frac{dS_{low}}{\lambda_{low}(1-\mu-S_{low})} &\Leftrightarrow -\frac{1}{\lambda_{high}}\ln(\mu-S_{high}) = -\frac{1}{\lambda_{low}}\ln\Psi^{\frac{\lambda_{low}}{\lambda_{high}}}(1-\mu-S_{low})\\
&\Leftrightarrow S_{high} = \mu-\Psi(1-\mu-S_{low})^{\frac{\lambda_{high}}{\lambda_{low}}}
\end{aligned}$$

where Ψ is the constant of integration which is determined by the starting conditions $\Psi = \frac{\mu-S_{high}(0)}{(1-\mu-S_{low}(0))^{\frac{\lambda_{high}}{\lambda_{low}}}}$.

Further, we assume that $\frac{S_{low}(0)}{1-\mu} = \frac{S_{high}(0)}{\mu} = \rho$, then

$$\Psi = \frac{\mu}{(1-\mu)^{\frac{\lambda_{high}}{\lambda_{low}}}} \frac{1}{(1-\rho)^{\frac{\lambda_{high}}{\lambda_{low}}-1}} \quad \text{and} \quad \frac{d\Psi}{d\rho} > 0. \tag{19}$$

Take as given $\varphi_h > 0$ and $\varphi_l = 0$. The belief upon receiving information conditional on a level of S_{high} and S_{low} is

$$\begin{aligned} b_j(\theta_i = h|m,t) &= \frac{\frac{\lambda_{high}S_{high}}{\lambda_{high}\mu+\lambda_{low}(1-\mu)} \frac{\varphi_h\alpha\mu+\left(S_{high}-\varphi_h\alpha\mu\right)\frac{(1-\varphi_h)\alpha}{1-\varphi_h\alpha}}{S_{high}} + \frac{\lambda_{low}S_{low}}{\lambda_{high}\mu+\lambda_{low}(1-\mu)} \frac{\varphi_h\alpha(1-\mu)+\frac{(1-\varphi_h)\alpha}{1-\varphi_h\alpha}(S_{low}-\varphi_h\alpha(1-\mu))}{S_{low}}}{\frac{\lambda_{high}S_{high}}{\lambda_{high}\mu+\lambda_{low}(1-\mu)} + \frac{\lambda_{low}S_{low}}{\lambda_{high}\mu+\lambda_{low}(1-\mu)}} \\ &= \frac{\left(\lambda_{high}S_{high}\left(\frac{\varphi_h\alpha\mu}{S_{high}}\frac{1-\alpha}{1-\varphi_h\alpha} + \frac{(1-\varphi_h)\alpha}{1-\varphi_h\alpha}\right) + \lambda_{low}S_{low}\left(\frac{\varphi_h\alpha(1-\mu)}{S_{low}}\frac{1-\alpha}{1-\varphi_h\alpha} + \frac{(1-\varphi_h)\alpha}{1-\varphi_h\alpha}\right)\right)}{\lambda_{high}S_{high}+\lambda_{low}S_{low}} \\ &= (1-\alpha)\frac{\varphi_h\alpha}{1-\varphi_h\alpha}\frac{\lambda_{high}\mu+\lambda_{low}(1-\mu)}{\lambda_{high}S_{high}+\lambda_{low}S_{low}} + \frac{(1-\varphi_h)\alpha}{1-\varphi_h\alpha} \end{aligned}$$

The belief upon not receiving the information remains constant:

$$b_j(\theta_i = h|\emptyset,t) = \frac{\lambda_{high}(\mu - S_{high})\frac{(1-\varphi_h)\alpha\mu}{(1-\varphi_h\alpha)\mu} + \lambda_{low}(1-\mu-S_{low})\frac{(1-\varphi_h)\alpha(1-\mu)}{(1-\varphi_h\alpha)(1-\mu)}}{\lambda_{high}(\mu - S_{high}) + \lambda_{low}(1-\mu-S_{low})} = \frac{(1-\varphi_h)\alpha}{1-\varphi_h\alpha}$$

The extent of the diffusion is determined by the point where the signaling value is equal to the cost of word of mouth, $b_j(\theta_i = h|m,t) - b_j(\theta_i = h|\emptyset,t) = k$. This is

$$\begin{aligned} &\frac{(1-\alpha)\varphi_h\alpha}{1-\varphi_h\alpha}\frac{\lambda_{high}\mu+\lambda_{low}(1-\mu)}{\lambda_{high}S^*_{high}+\lambda_{low}S^*_{low}} = k \\ \Leftrightarrow\ &\lambda_{high}S^*_{high} + \lambda_{low}S^*_{low} = \frac{(1-\alpha)\varphi_h\alpha}{1-\varphi_h\alpha}\frac{\lambda_{high}\mu+\lambda_{low}(1-\mu)}{k}. \end{aligned}$$

There is a set of possible solutions $\left\{S^*_{high}, S^*_{low}\right\}$ to this equation. The unique solution is determined by the initial conditions $S_{high}(0)$ and $S_{low}(0)$ inclusive of any advertising undertaken by the firm. The firm prefers solutions with higher S^*_{low} since $\lambda_{high}S^*_{high} + \lambda_{low}S^*_{low}$ is constant for all solutions. Hence, larger values of S^*_{low} imply larger total diffusion $S^* = S^*_{high} + S^*_{low}$.

Now consider three different starting points at $t = 0$. The first is the no advertising case $S_{high}(0) = \varphi_h\alpha\mu$ and $S_{low}(0) = \varphi_h\alpha(1-\mu)$. The second is the case of untargeted advertising $S_{high}(0) = \left(\varphi_h\alpha + \tilde{\beta}^{Untarg}(1-\varphi_h\alpha)\right)\mu$ and $S_{low}(0) = \left(\varphi_h\alpha + \tilde{\beta}^{Untarg}(1-\varphi_h\alpha)\right)(1-\mu)$. The third is targeted advertising at high mixing types $S_{high}(0) = \left(\varphi_h\alpha + \tilde{\beta}^{High}(1-\varphi_h\alpha)\right)\mu$ and $S_{low}(0) = \varphi_h\alpha(1-\mu)$. Note in both advertising cases, the same mass of individuals is advertised to in total, where $\tilde{\beta}^{High} = \frac{\tilde{\beta}^{Untarg}}{\mu}$. It is immediate that no advertising results in a larger diffusion than targeted advertising because the initial condition for the low mixing population is the same but the high mixing population

is greater under the targeted advertising case.

Next, we can compare no advertising and untargeted advertising by considering the size of the high mixing population conditional on the starting level of the untargeted advertising case $S_{low}(0) = \left(\varphi_h\alpha + \tilde{\beta}(1-\varphi_h\alpha)\right)(1-\mu)$. To reach this level in the no advertising case, information must diffuse. In the time it take the low mixing population to increase its proportional size by $\tilde{\beta}(1-\varphi_h\alpha)$ the high mixing population will increase by proportionally more because they mix at a higher rate. We can see this from equation (19) for the constant of integration for the joint diffusion $A = \frac{\mu}{(1-\mu)^{\frac{\lambda_{high}}{\lambda_{low}}}} \frac{1}{(1-\rho)^{\frac{\lambda_{high}}{\lambda_{low}}-1}}$, which determines the relationship between the high and low mixing populations. We also note that $\frac{dA}{d\rho} > 0$. Hence, A is larger for the untargeted advertising case. Conditional on reaching a given level for S_{low}, then S_{high} will be small. The untargeted advertising population will have a smaller population of informed high mixing types and the extent of the diffusion will be reached with a greater level of S^*_{low}.

Note that advertising does not exceed the extent of diffusion provided that $\tilde{\beta} \leq \bar{\beta}$ where $\bar{\beta}$ satisfies the following condition

$$\begin{aligned}\lambda_{high}\left(\varphi_h\alpha + \tfrac{\bar{\beta}}{\mu}(1-\varphi_h\alpha)\right)\mu + \lambda_{low}\varphi_h\alpha(1-\mu) &= (1-\alpha)\tfrac{\varphi_h\alpha}{1-\varphi_h\alpha}\tfrac{\lambda_{high}\mu+\lambda_{low}(1-\mu)}{k} \\ \Leftrightarrow\ \bar{\beta} &= \tfrac{\varphi_h\alpha}{1-\varphi_h\alpha}\left[\tfrac{1-\alpha}{k(1-\varphi_h\alpha)} - 1\right]\left[\lambda_{high}\mu + \lambda_{low}(1-\mu)\right]\end{aligned}$$

This condition is found by considering targeted advertising because this results in the smallest diffusion and thus provides the tightest bound on the amount of advertising.

B Advertising section under trembling hand refinement

The analysis in Section 4 assumed that the level of advertising is unobserved by individuals and analyzes the firm's advertising decision and the consumers choice of a time to stop spreading information as a simultaneous move game. In an equilibrium where $\beta^* \in (0,1)$ receiving or not receiving an advertisement are both possible on the equilibrium path and thus beliefs of consumers upon seeing or not seeing an advertisement are pinned down by Bayesian updating. Thus the equilibria we find when analyzing the subgame from $t \geq 0$ onwards as a simultaneous move game are perfectly consistent with this analysis. On the other hand, in an equilibrium where $\beta^* = 0$, consumers expect there is 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 β 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 $\beta \in [0, 1]$, the actual fraction of that receive the advertisement is $\beta(1-\epsilon) + \epsilon(1-\beta) = \beta + \epsilon - 2\epsilon\beta$. This ensures that when the firm chooses $\beta^* = 0, 1$ then 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 $\tilde{\beta}(a)$ will be the same for $a = 0, 1$ and we can proceed by denoting it by just $\tilde{\beta}$. We can now write out the consumers best response in terms of $\tilde{\beta}$ 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:

$$\tau^*\left(a, \varphi_h, \tilde{\beta} + \epsilon - 2\epsilon\tilde{\beta}\right) = \frac{1}{\lambda} \ln \frac{\tilde{S}_c^*(\varphi_h, 0, \tilde{\beta} + \epsilon - 2\epsilon\tilde{\beta})}{1 - \tilde{S}_c^*(\varphi_h, 0, \tilde{\beta} + \epsilon - 2\epsilon\tilde{\beta})} \frac{1 - S_0\left(\varphi_h, 0, \tilde{\beta} + \epsilon - 2\epsilon\tilde{\beta}\right)}{S_0\left(\varphi_h, 0, \tilde{\beta} + \epsilon - 2\epsilon\tilde{\beta}\right)} \tag{20}$$

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

$$\begin{aligned} & C'(\beta(\varphi_h, \tilde{\tau})) \\ &= \frac{e^{-\lambda\tilde{\tau}}(1 - \varphi_h\alpha)}{\left[\begin{array}{c} S_0(\varphi_h, 0, \beta(\varphi_h, \tilde{\tau}) + \epsilon - 2\epsilon\beta(\varphi_h, \tilde{\tau})) \\ + (1 - S_0(\varphi_h, 0, \beta(\varphi_h, \tilde{\tau}) + \epsilon - 2\epsilon\beta(\varphi_h, \tilde{\tau}))) e^{-\lambda\tilde{\tau}} \end{array}\right]^2} \text{ if } \beta^* \in (0, 1) \end{aligned} \tag{21}$$

and

$$C'(0) \geq \frac{e^{-\lambda\tilde{\tau}}(1 - \varphi_h\alpha)}{\left[S_0(\varphi_h, 0, \epsilon) + (1 - S_0(\varphi_h, 0, \epsilon)) e^{-\lambda\tilde{\tau}}\right]^2} \text{ if } \beta^* = 0 \tag{22}$$

The information acquisition choice of consumers satisfies:

$$\begin{aligned} \varphi_h^* \bar{c} &= V(\varphi_h^*, \tau^*(\varphi_h^*), \beta^*(\varphi_h^*) + \epsilon - 2\epsilon\beta^*(\varphi_h^*)) \quad \text{if } V(1, \tau^*(1), \beta^*(1)) < \bar{c} \\ \varphi_h^* &= 1 \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{if } V(1, \tau^*(1), \beta^*(1) + \epsilon - 2\epsilon\beta^*(1)) \geq \bar{c}. \end{aligned} \tag{23}$$

We are interested in the limit of the equilibria as $\epsilon \to 0$. We note equations (20), (21), and (22) 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 (23) is continuous in all its arguments and thus also attains the same limit as in Section 4.

References

[1] Bass, Frank M. (1969), "A New Product Growth Model for Consumer Durables," *Management Science*, 15 (5), 215-227.

[2] Bass, Frank M. (1980), "The Relationship between Diffusion Rates, Experience Curves, and Demand Elasticities for Consumer Durable Technological Innovations," *Journal of Business*, 53 (3), S51-67.

[3] Baumeister, Roy F. (1998), "The Self," in *The Handbook of Social Psychology*, Vol 1, D. T. Giblert, S. T. Fiske, and G. Lindzey, eds. Hoboken, NJ: John Wiley & Sons, 680-740.

[4] Baumeister, Roy F., D. M. Tice, and D. G. Hutton (1989), "Self-presentational motivations and personality differences in self-esteem," *Journal of Personality*, 57, 547-579.

[5] Berger, Jonah and Katherine Milkman (2011), "What Makes Online Content Viral?" *Journal of Marketing Research*, forthcoming.

[6] Berger, Jonah and Eric Schwartz (2011), "What Drives Immediate and Ongoing Word of Mouth?" *Journal of Marketing Research*, Vol. XLVIII (October 2011), 869–880.

[7] Bughin, Jacques, Jonathan Doogan, and Ole Jorgen Vetvik (2010), "A New Way to Measure Word-of-Mouth Marketing," *McKinsey Quarterly*, 2, 113-116.

[8] Calvó-Armengol, Antoni, Joan de Martí, and Andrea Prat (2012), "Endogenous Communication in Complex Organizations," working paper.

[9] Campbell, Arthur (2013), "Word-of-Mouth Communication and Percolation in Social Networks," *American Economic Review,* forthcoming.

[10] Campbell, James D. (2010), "Signaling to a Network of Consumers," *working paper.*

[11] Chatterjee, Kalyan and Bhaskar Dutta (2010), "Word of Mouth Advertising, Credibility and Learning in Networks," *working paper.*

[12] Chen, Yubo and Jinhong Xie (2008), "Online Consumer Review: Word-of-mouth as A New Element of Marketing Communication Mix," *Management Science*, 54 (3), 477-491.

[13] Chintagunta, Pradeep, Gopinath, Shyam, and Sriram Venkataraman (2010), "The Effects of Online User Reviews on Movie Box Office Performance: Accounting for Sequential Rollout and Aggregation Across Local Markets," *Marketing Science*, 30 (6), 398-419.

[14] Chevalier, Judith and Dina Mayzlin (2006), "The Effect of Word of Mouth on Sales: Online Book Reviews," *Journal of Marketing Research*, 43 (3), 345-354.

[15] Dichter, Ernest (1966), "How Word-of-Mouth Advertising Works," *Harvard Business Review*, Nov-Dec, 131-144.

[16] Dye, Renee (2000), "The Buzz on Buzz," *Harvard Business Review*, Nov-Dec, 138-146.

[17] Galeotti, Andrea (2010), "Talking, Searching and Pricing," *International Economic Review*, 50 (4), 159-174.

[18] Galeotti, Andrea and Sanjeev Goyal (2007), "Network Multipliers and the Optimality of Indirect Communication," *working paper.*

[19] Galeotti, Andrea and Sanjeev Goyal (2009), "Influencing the influencers: a theory of strategic diffusion," *RAND Journal of Economics*, 40(3), 509-532.

[20] Galeotti, Andrea, and Mattozzi, Andrea (2011), "Strategic Information Transmission in Networks," *American Economic Journal: Microeconomics,* 3 (1), 307-27.

[21] Hennig-Thurau, Thorsten, K. Gwinner; G. Walsh, and D. Gremler (2004), "Electronic Word of Mouth via Consumer Opinion platforms: What motivates consumers to articulate themselves on the Internet?" *Journal of Interactive Marketing*, 18(1) 38-52.

[22] Hughes, Mark (2005), "*Buzzmarketing: Get People to Talk About Your Stuff*," Penguin Group, New York, NY.

[23] Katz, Elihu and Paul F. Lazarsfeld (1955), *Personal Influence*. Glencoe, IL: Free Press.

[24] Kotler, Philip (2002), "*A Framework for Marketing Management*," Prentice Hall, US.

[25] Luca, Michael (2011), "Reviews, Reputation, and Revenue: The Case of Yelp.com," *Harvard Business School Working Paper*, No 12-016.

[26] Mahajan, Vijay, Eitan Muller, and Frank M. Bass (1990), "New Product Diffusion Models in Marketing: A Review and Directions for Further Research," *Journal of Marketing*, 54 (1), 1-26.

[27] Niehaus, Paul (2011), "Filtered Social Learning," *Journal of Political Economy*, 119 (4), 686-720.

[28] Pesendorfer, Wolfgang (1995), "Design Innovation and Fashion Cycles," *American Economic Review,* 85 (4), 771-792.

[29] Sernovitz, Andy (2011), *Word of Mouth Marketing*, PressBox, Chicago, IL.

[30] Stein, Jeremy (2008), "Conversations Among Competitors," *American Economic Review*, 98 (5), 2150-2162.

[31] Sundaram, D.S., Kaushik Mitra and Cynthia Webster (1998), "Word-of-Mouth Communications: A Motivational Analysis," *Advances in Consumer Research*, 25, 527-531.

[32] Watts, Duncan J. and Peter S. Dodds (2007), "Influentials, networks, and public opinion formation," *Journal of Consumer Research*, 34 (4), 441-458.

[33] Weimann, Gabriel (1994), *The Influentials: People Who Influence People,* Albany, SUNY Press.

[34] Wojnicki, Andrea and David Godes (2011), "Signaling Success: Strategically-Positive Word of Mouth," *working paper.*

[35] Word of Mouth Marketing Association (2011) infographic: http://word of mouthma.org/word/wp-content/uploads/2011/10/word_of_mouth_marketing_impact_and_influence_wor of mouthma.png

[36] Young, H. Peyton (2009), "Innovation Diffusion in Heterogeneous Populations: Contagion, Social Influence, and Social Learning," *American Economic Review,* 99 (5), 1899-1924.