

# BUZZ MANAGEMENT

Arthur Campbell, Dina Mayzlin and Jiwoong Shin\*

*Yale University*

September 2012

## **Abstract**

We study how a firm uses structured information release, word of mouth marketing and advertising to maximize the diffusion of information about its product. Individuals are either a high or a low social type. During social interactions it is valuable for any individual to increase another person's posterior that they are a high social type. We develop a model to study how a firm interacts with this motive to maximize the diffusion of information. We find that a firm will increase the cost for low types to acquire information and restrict the ability of low types to pass on information. A commitment by the firm not to advertise is beneficial which may take the form of releasing information a sufficient amount of time before the product is released.

**Keywords:** Word of mouth, advertising, buzz, diffusion.

VERY PRELIMINARY DRAFT: DO NOT CITE WITHOUT PERMISSION

---

\*Correspondence: 135 Prospect St., P.O. Box 208200, New Haven, CT 06520, e-mail: arthur.campbell@yale.edu, dina.mayzlin@yale.edu, jiwoong.shin@yale.edu. We thank seminar participants at the SICS Conference, Workshop on the Economics of Advertising and Marketing, Marketing Science Conference, Yale Economics, Yale SOM, for helpful feedback.

# 1 Introduction

Word of mouth is widely recognized as an important source of information for consumer's purchase decisions both in the academic literature<sup>1</sup> as well as in industry research.<sup>2</sup> Although a number of papers have studied the effect of word of mouth on sales, the issue of what motivates consumers to spread information has received less attention, and no work to date has examined the implications of these motives on firm's communications and advertising strategy. Most of existing literature treats 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). 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. Naturally, when treated as a mechanical process, it leads one to predict that any strategy on behalf of a firm which increases the propensity of consumers to hold the information will facilitate a greater amount of word of mouth and diffusion of information. In contrast, we explicitly model the motivation of consumers to engage in word of mouth and hence treat the word of mouth generation process as an outcome of strategic consumer behavior. This has very different implications on the firm's optimal information release and advertising strategy: we find that a firm will increase the costs to consumers to acquire information, restrict who can pass on information and will benefit from a commitment not to undertake advertising.

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), or desire to signal expertise to others (Wojnicki and Godes 2011). Although any of these motives can independently drive word of mouth, in this article, we focus on a particular psychological motive. The starting point of our model is that consumers derive benefits during social interactions from making themselves look good. It is motivated by the psychological theory of "self-enhancement" or the tendency to "affirm the self" (Baumeister 1998). Self-enhancement is recognized as one of the basic social needs (Fiske 20010, Sedikides 1993) and it refers to the tendency to seek experiences that improve or bolster the self-concept, for example by drawing 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).

Building on the notion that people often seek to enhance the self, we develop a model where consumers are either a high or low social type and mix (meet one another) at a Poisson rate over time. Consumers receive utility during these social interactions. The key assumption is that the

---

<sup>1</sup>Several studies have found that word of mouth communications influence nearly 70% of all buying decisions (Balter 2008) and are considered the primary driving force of all industries (Dye 2000). For example, word of mouth has been shown to affect purchasing behavior in restaurant choice (Luca 2012), book sales (Chevalier and Mayzlin), banking (Keaveney 1995), entertainment (Chintagunta, Gopinath, and Venkatarman 2010), technological products (Herr, Kardes, and Kim 1991), and appliances and clothing (Richins 1983).

<sup>2</sup>"54% of purchase decisions are influenced by word of mouth" Word of Mouth Marketing Association (2011), "word of mouth is the primary factor behind 20 to 50 percent of all purchase decisions" Bughin et. al. (2010), "There are 3.3 billion brand impressions created each and every day in America via word of mouth" and "The leading categories for word of mouth are food/dining and media/entertainment, with a majority of all Americans talking about these on a given day" Keller and Libai (2009); "word of mouth remains the biggest influence in people's electronics (43.7%) and apparel (33.6%) purchases" National Retail Federation (2009).

utility an individual receives during a social interaction is an increasing function of her peer's belief that she is the high social type. Prior to mixing, individuals choose whether or not to acquire information about the firm's product at a certain cost. Then, during the mixing, individuals decide whether or not to engage in costly word of mouth during each social interaction. In this context, knowing the information and passing it on via word of mouth may serve as a credible signal that one is a high social type and thus, is valuable during the subsequent social interaction. Hence, this may compensate an individual for both the cost of initially acquiring the information and for engaging in word of mouth. Individuals upon hearing the information, during a social interaction, may themselves engage in word of mouth during subsequent interactions. The central focus of our analysis is on a signaling equilibrium and how a firm may affect the signaling motive of consumers through its information release and advertising strategies. This allows us to characterize the optimal strategies for a firm to maximize the extent of information diffusion.

We find that a firm may optimally increase (decrease) the cost for low (high) social types to acquire information about the product. When a firm creates a cost asymmetry between the social types a signaling equilibrium exists whereby high social types are more likely to initially acquire information, and their ability to pass it on, through word of mouth, can serve as signal of their high social status. As information diffuses in the population, and low social types acquire information through word of mouth, the signaling value of the information becomes diluted. Eventually, the diffusion stops when the signaling value is smaller than the cost to engage in word of mouth. When a firm cares about the speed at which the diffusion occurs, a basic trade-off is between increasing the initial speed of diffusion and maximizing the total spread of information about the product. We find that untargeted advertising by the firm serves to crowd out the incentives of individuals to acquire information and engage in word of mouth communication. A commitment by the firm not to engage in advertising increases the diffusion of information. We also show that one natural source of commitment power is for the firm to release information about the product a sufficient amount of time prior to a product's release, thereby allowing word of mouth to occur.

The theoretical predictions of our model are consistent with recommendations of marketing practitioners. For instance, 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." There are also some examples of marketing campaigns where initial restriction on consumer access to product information resulted in a large amount of "buzz" about the product. For example, Ty used exclusivity in its initial launch of bean-stuffed toys.<sup>3</sup> In July of 2011, the European music site Spotify launched in the US. At first, its free version was available by invitation only.<sup>4</sup> Obtaining the invitation was non-trivial: consumers could receive either through current users or through other channels. For example, Coca Cola,<sup>5</sup> gave out invitations to users who submitted their email address, and interacting with Spotify over Twitter could also result in an invitation. Eventually, anyone could download the free version of Spotify through the company's Web site. By November of 2011, 4 million users adopted Spotify.<sup>6</sup> 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.<sup>7</sup> On October 2011, Google

---

<sup>3</sup>See Dye (2000)

<sup>4</sup><http://www.nytimes.com/2011/07/14/technology/spotify-music-streaming-service-comes-to-us.html>

<sup>5</sup>[http://news.cnet.com/8301-13845\\_3-20081418-58/get-a-quick-and-easy-invitation-to-spotify/](http://news.cnet.com/8301-13845_3-20081418-58/get-a-quick-and-easy-invitation-to-spotify/)

<sup>6</sup><http://articles.latimes.com/2011/nov/10/business/la-fi-ct-facetime-spotify-20111110>

<sup>7</sup>An example of this is Ken Hess who wrote on the tech news site ZDNet on June 30, 2011, "Dear Google, I want

announced Google+'s user base to be at more than 40 million users.<sup>8</sup> Both of these sites initially made information and access to their products scarce and benefited from large amounts of word of mouth and buzz about these sites. Finally, our result on advertising displacing the incentive to talk explains why companies who try to generate buzz typically do so several months prior to a mass advertising campaign.<sup>9</sup>

## 2 Literature Review

Self-enhancement is acknowledged as one of the five key human social motivations along with belonging, understanding, controlling and trusting (Basumeister 1998, Sedikides 1993, Fiske 2001, Sirgy 1982). In psychology and marketing literature, several papers show that word of mouth is influenced by motives related to self-presentation. Berger and Milkman (2011) find that positive content is more likely to be shared, as is content that evokes high-arousal emotions. The authors speculate 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 motivation behind WOM. Also, Sundaram et. al.(1998) infer that self-enhancement accounts for 20% of positive word of mouth, 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.

Also, our study is related to the models 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. The most related studies are Galeotti and Goyal (2007), J. Campbell (2009), Chatterjee and Dutta (2010), which consider the advertising strategy of the firm, Ifrach, Maglaras and Scarsini (2011) and Candogan, Bimpikis and Ozdaglar (2010), which focus on pricing, and A. Campbell (2012) which considers both. Within this literature our paper is the first to explicitly model a social motivation for individuals to engage in word of mouth. Our social signaling mechanism for word of mouth is distinct from the previous literature. Galeotti and Goyal (2007) and J. Campbell (2009) assume that firms initially advertise to consumers and then word of mouth travels a distance of one in the social network, and Galeotti and Goyal (2007) also extend their untargeted advertising results to a generalized maximum distance. 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 (2012) consider settings where consumers pass on information if they are prepared to purchase the product. In contrast to this literature, our social signaling mechanism leads to novel insights into how a firm may optimally manage 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.

---

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. Come on, I really want one." <http://www.zdnet.com/blog/bt1/dear-google-where-the-hell-is-my-google-invitation/51640>

<sup>8</sup><http://www.pcmag.com/article2/0,2817,2398114,00.asp>

<sup>9</sup>In the case of Mountain Dew Code Red, early on the company engaged in actions such as reaching out to influentials and organizing parties at college campuses, and did not engage in mass advertising until a few months later (Kotler (2002)).

Our model studies how social interactions between consumers may be influenced by a firm’s strategy. A paper in this vein is Pesendorfer (1995) where the focus is on the interaction between consumers’ social matching behavior and the pricing strategy and design innovations of a firm. In that model owning a particular product serves as a signal for a consumer to others that they are a high social type. This is valuable for high types to identify one another during a matching process in order to ensure they are matched with another high type. The premise that there are socially desirable types and consumers may benefit from interacting with high types, and thus the ability to signal their types in a social context does connect both the models. However, our focus is very different as we consider how a firm’s information release and advertising strategy interacts with social concerns.

Similarly to our work, two recent papers have analytically modeled a signaling motive on word of mouth (Wojnicki and Godes 2011) or the firm’s communication strategy (Yoganarasimhan 2012).<sup>10</sup> Wojnicki and Godes (2011) present a model where the desire for impression management motivates experts to suppress conversations about their negative experiences. A major difference between our work and Wojnicki and Godes (2011) is the nature of the signal: while in Wojnicki and Godes (2011) the valence of experts’ experiences signals their ability to choose, in our model the signal regards an individual’s social type. Our work also relates to Yoganarasimhan (2012), which 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 a static setting. Besides the differences in context - while Yoganarasimhan (2012) studies fashion goods, our model concerns itself with all products - one notable difference in the models is our focus on the effect of initial exclusive release on the extent of diffusion in a dynamic setting.

### 3 Model

#### 3.1 A Model of Buzz

A monopolist is selling a product to a mass 1 of agents. An agent  $i$  may be one of two types: high and low  $\theta = h, l$  where  $\Pr[\theta_i = h] = \alpha < \frac{1}{2}$  where high types are relatively scarce. We think of these as high- and low-status consumers. The firm’s objective is to maximize the fraction of the population which receives information  $m$  about its product.<sup>11</sup> We shall denote the fraction of the population that has received the information by  $S$ . Initially we assume no advertising before introducing it in Section 5. Without advertising, consumers can find out about the product 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 who themselves will incur a cost to pass the information to them. Note once an agent has found out about the product through either channel, they are able to pass on information about it themselves. In the case of Spotify and

<sup>10</sup>Kuksov (2007) also studies the incentives of consumers to reveal or conceal information about themselves to others through brand choices in the consumer matching context.

<sup>11</sup>Although we leave the firm’s objective in this reduced form in our main text, one can think of micro-model of 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 firm’s profit is linear in the fraction of the population who receive information about the product.

Google+, a consumer could only learn about the product by using it. However, this generally may not be the case. We first consider the general setting where anyone can pass on information and then consider in Section 4, how a firm can increase the spread of information by restricting who can pass along information as it may be able to do if adopting the product is required in order to learn the information or credibly pass on the information (in the sense that it does not perfectly reveal the individual as someone who heard the information from someone else).

### Timing

At time  $t = -1$  each type 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 agent,  $c_h, c_l$ , 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. The firm, in addition to this cost, may impose further costs on either or both types through its information release strategy. This is modelled as an additional cost  $v_h, v_l \geq 0$  which differentially affects each social type. We assume this is costless (or is a very small cost) for the firm. One can think of these activities as explicitly increasing costs, through the use of technical jargon and language which the high social type more easily understands, or equivalently as decreasing costs of the high social type relative to the low social type, through releasing information on blogs, at events, or in venues that are frequented by high social types but not low social types. What is important for the model is that the firm may differentially affect the costs of each social type. Again, using the Spotify example, the firm could allow anyone who registered on its website to obtain an invitation or it could only send an invitation to those who interacted with it on Twitter.

From time  $t = 0$  onwards individuals mix at rate  $\lambda$ . During each meeting an individual may pass on the hard information  $m$  at a cost  $k$ , where we assume  $\phi < k < 1 - \phi$ , or pass on no information  $\emptyset$  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 individuals information first. This assumption makes the analysis more tractable.

### Social Utility

The motivation for our analysis 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 agent has about the agent's type. In particular, agent  $i$  receives utility

$$U_i(b_j(\theta_i = h|m, t)) \tag{1}$$

if agent  $i$  passes a message  $m$  at time  $t$ , where  $b_j(\theta_i = h|m, t)$  is the other agent  $j$ 's belief that agent  $i$  is a high type upon receiving the information  $m$ . And similarly,

$$U_i(b_j(\theta_i = h|\emptyset, t)) \tag{2}$$

if the agent does not pass information, where  $b_j(\theta_i = h|\emptyset, t)$  is the belief if no signal (denoted by  $\emptyset$ ) 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

$$\Delta U_i(t) = U_i(b_j(\theta_i = h|m, t)) - U_i(b_j(\theta_i = h|\emptyset, t)) \quad (3)$$

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

$$\begin{aligned} U(b_j(\theta_i = h|m, t)) - U(b_j(\theta_i = h|\emptyset, 0)) \\ = \bar{u} [b_j(\theta_i = h|m, t) - b_j(\theta_i = h|\emptyset, t)] \end{aligned} \quad (4)$$

where we normalize  $\bar{u} = 1$ .<sup>12</sup>

### Growth of the informed population

We now map the incentive to talk to the size of the informed population. Denote the fraction of types who become informed at  $t = -1$  by  $\varphi_h, \varphi_l$  where these are going to be endogenously determined in equilibrium. The fraction of the population which are informed  $S(t)$  evolves over time as agents mix at rate  $\lambda$  and pass on information. The initial condition at  $t = -1$  is  $S_0 = \varphi_h\alpha + \varphi_l(1 - \alpha)$  and the rate of change of the informed population is given by:

$$\frac{dS}{dt} = \lambda S(t)(1 - S(t)) \quad (5)$$

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

$$S(t) = \frac{1}{1 + ae^{-\lambda t}}, \text{ where } a = \frac{1 - S_0}{S_0} \quad (6)$$

which continues to grow until the beneficial impact of passing the firm's message is less than the cost of doing so. Hence  $S(t)$  stops growing when  $\Delta U(t^*) = k$  which defines the extent of the diffusion  $S^* = S(t^*)$ . In other words, the consumer stops talking when the signaling benefit of word of mouth equals the cost of passing on the information. We assume the firm's profits are positively related to the extent of the diffusion and it will endeavor to maximize the extent of this diffusion.

### Beliefs

At  $t = 0$  beliefs are

$$b_j(\theta_i = h|m, 0) = \frac{\varphi_h\phi}{\varphi_h\phi + \varphi_l(1 - \phi)} \quad (7)$$

and

$$b_j(\theta_i = h|\emptyset, 0) = \frac{(1 - \varphi_h)\phi}{(1 - \varphi_h)\phi + (1 - \varphi_l)(1 - \phi)}. \quad (8)$$

---

<sup>12</sup>An example of this reduced form model is to assume the utility to an individual  $i$  from an interaction with another individual  $j$  is given by  $U_{ij} = 1[\theta_j = h]e_i + e_j - \frac{e_i^2}{2}$  where  $e_i$  and  $e_j$  are investments made by each individual at personal cost  $\frac{e_i^2}{2}$ . Expected utility maximization results in an individual making an investment equal to the posterior about the other agent's type  $e_j = \Pr_j[\theta_i = h]$ . A individual  $i$  may increase the posterior of individual  $j$  through the message she sends  $\Pr_j[\theta_i = h|m] > \Pr_j[\theta_i = h|\emptyset]$  and thus increases the investment  $e_j$  the agent makes.

Beliefs change over time as the message diffuses through the population. The belief when a person receives a signal at a time  $t$  is given by

$$\begin{aligned} b_j(\theta_i = h|m, t) &= \frac{S(t) - S_0}{S(t)} [b_j(\theta_i = h|\emptyset, 0)] + \frac{S_0}{S(t)} [b_j(\theta_i = h|m, 0)] \\ &= b_j(\theta_i = h|\emptyset, 0) + \frac{S_0}{S(t)} [b_j(\theta_i = h|m, 0) - b_j(\theta_i = h|\emptyset, 0)]. \end{aligned} \quad (9)$$

The beliefs upon not receiving a signal do not change over time and hence are given by

$$b_j(\theta_i = h|\emptyset, t) = b_j(\theta_i = h|\emptyset, 0) = \frac{(1 - \varphi_h) \phi}{(1 - \varphi_h) \phi + (1 - \varphi_l)(1 - \phi)}. \quad (10)$$

### Extent of diffusion

The diffusion of the signal thus stops when the marginal value of signaling equals the marginal cost of passing on the information:

$$\begin{aligned} U(b_j(\theta_i = h|m, t)) - U(b_j(\theta_i = h|\emptyset, t)) \\ = \frac{S_0}{S(t)} [b_j(\theta_i = h|m, 0) - b_j(\theta_i = h|\emptyset, 0)] = k \end{aligned} \quad (11)$$

Hence, the diffusion stops at

$$S^* = \frac{S_0}{k} [b_j(\theta_i = h|m, 0) - b_j(\theta_i = h|\emptyset, 0)] \quad (12)$$

We can now replace  $b_j(\theta_i = h|m, 0)$  and  $b_j(\theta_i = h|\emptyset, 0)$ ,

$$\begin{aligned} S^*(\varphi_h, \varphi_l) &= \frac{1}{k} (\varphi_h \phi + \varphi_l (1 - \phi)) \left[ \frac{\varphi_h \phi}{\varphi_h \phi + \varphi_l (1 - \phi)} - \frac{(1 - \varphi_h) \phi}{(1 - \varphi_h) \phi + (1 - \varphi_l)(1 - \phi)} \right] \\ &= \frac{\phi}{k} \left[ \varphi_h - (1 - \varphi_h) \frac{\varphi_h \phi + \varphi_l (1 - \phi)}{(1 - \varphi_h) \phi + (1 - \varphi_l)(1 - \phi)} \right] \end{aligned} \quad (13)$$

The derivative of  $S^*$  with respect to  $\varphi_h, \varphi_l$  are:

$$\begin{aligned} \frac{dS^*}{d\varphi_h} &= \frac{1}{k} \left[ \frac{(1 - \varphi_l)(1 - \phi)}{((1 - \varphi_h) \phi + (1 - \varphi_l)(1 - \phi))^2} \right] \geq 0 \text{ if } \varphi_l < 1 \text{ then } > 0 \\ \frac{dS^*}{d\varphi_l} &= -\frac{\phi}{k} \frac{(1 - \varphi_h)(1 - \phi)}{((1 - \varphi_h) \phi + (1 - \varphi_l)(1 - \phi))^2} \leq 0 \text{ if } \varphi_h < 1 \text{ then } < 0 \end{aligned} \quad (14)$$

Without considering the ex ante incentives for consumers to acquire information at  $t = -1$ , the diffusion of the firm's message is increasing in the number of high type informed consumers ( $\varphi_h$ ) and decreasing in the number of low type informed customers ( $\varphi_l$ ). It is, therefore, easy to see that the diffusion of information is maximized at  $\varphi_h = 1$  and  $\varphi_l = 0$  in the absence of ex ante incentives for consumers to acquire information. However this may not be a feasible solution if the ex ante information acquisition constraints bind for some high types.



### Ex Ante Incentives

Consumers decide whether to acquire information at  $t = -1$  (for example, to scour the Internet for the Spotify invitation). The decision to acquire the information at  $t = -1$  depends on the total signaling benefit the agent will acquire during the diffusion process. If this benefit is above an individual's cost  $c$  then the agent will acquire information. Denoting the time at which the diffusion process ends by  $t^*$ . The signaling benefit for an agent is then

$$V = \lambda \left( \int_0^{t^*} \left( \frac{1 - S(t)}{1 - S_0} \right) \left( \frac{S_0}{S(t)} [b_j(\theta_i = h|m, 0) - b_j(\theta_i = h|\emptyset, 0)] - k \right) dt \right) \quad (15)$$

where  $\frac{1-S(t)}{1-S_0}$  is the probability of remaining uninformed at time  $t$  for an individual uninformed at time 0. It may be further simplified by making a change of variable using

$$\begin{aligned} \frac{dS}{dt} &= \lambda S(t) (1 - S(t)) \\ dt &= \frac{dS}{\lambda S(t) (1 - S(t))} \end{aligned}$$

making this substitution for  $dt$  and substituting  $\Delta b = b_j(\theta_i|m, 0) - b_j(\theta_i|\emptyset, 0)$ , where  $\Delta b = \frac{S^*}{S_0}k$  (from Equation 12), we get:

$$\begin{aligned} V &= \int_{S_0}^{S^*} \left( \frac{1}{1 - S_0} \right) \frac{1}{S(t)} \left( \frac{S_0}{S(t)} [b_j(\theta_i = h|m, 0) - b_j(\theta_i = h|\emptyset, 0)] - k \right) dS \quad (16) \\ &= \frac{1}{1 - S_0} \int_{S_0}^{S^*} \left[ \frac{\Delta b \cdot S_0}{S^2(t)} - \frac{k}{S(t)} \right] dS = \left( \frac{k}{1 - S_0} \right) \left( S^* \left[ -\frac{1}{S} \right]_{S_0}^{S^*} - [\ln S]_{S_0}^{S^*} \right) \\ &= \left( \frac{k}{1 - S_0} \right) \left( \frac{S^* - S_0}{S_0} + \ln \frac{S_0}{S^*} \right) \end{aligned}$$

The total signaling benefit is a function of the initial diffusion state ( $S_0$ ) and the total extent of information diffusion ( $S^*$ ).

**Lemma 1.** *The total signaling benefit is increasing in  $S^*$  and decreasing in  $S_0$ :*

$$\frac{\partial V}{\partial S_0} < 0 \text{ and } \frac{\partial V}{\partial S^*} > 0.$$

The lemma shows that the total benefit of signaling from passing the firm's message is increasing in  $S^*$ . As the information will spread greater portion of population, the more benefit the consumer can enjoy over longer period time. The lemma also finds that the more people are informed about the message at the initial stage, holding constant the total diffusion, then there are fewer opportunities for signaling for each individual and thus the value of obtaining the signal ex ante is decreasing.

### Firm's problem

We assume that the firm can increase the costs of the high and low types for acquiring information. Denote the amount by which the firm increases the cost of each type by  $v_l \geq 0$  and  $v_h \geq 0$ . The cutoff type for each is then given by

$$\begin{aligned}\varphi_h \bar{c} + v_h &\leq V(\varphi_h, \varphi_l) \\ \varphi_l \bar{c} + v_l &\leq V(\varphi_h, \varphi_l)\end{aligned}$$

This is equivalent to the firm choosing  $\varphi_l$  and  $\varphi_h$  subject to the feasibility constraints

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

$$\begin{aligned}0 &\leq \varphi_h \leq 1 \\ 0 &\leq \varphi_l \leq 1\end{aligned}$$

We can now write the firm's optimization problem as maximizing the spread of information through the choice of  $\varphi_l$  and  $\varphi_h$  subject to these feasibility constraints:

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

subject to

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

$$\begin{aligned}0 &\leq \varphi_h \leq 1 \\ 0 &\leq \varphi_l \leq 1\end{aligned}$$

**Proposition 1.** *The optimal strategy for the firm has the following characteristics:*

$$0 < \varphi_h \leq 1 \text{ and } \varphi_l = 0.$$

Moreover,

$$\begin{aligned}\varphi_h &< 1 \text{ if } \bar{c} \geq \frac{1 - k + k \ln k}{1 - \phi} \\ \varphi_h &= 1 \text{ if } \bar{c} < \frac{1 - k + k \ln k}{1 - \phi}\end{aligned}$$

Summarizing the result, we find that even with ex ante incentives for consumers to acquire information, the optimal strategy for the firm that maximizes the diffusion of the firm's message is to restrict information to the socially low type agents by choosing  $\varphi_l = 0$  by setting  $v_l > \bar{c}$  and minimizing the costs for the socially high type agents  $v_h = 0$  such that  $0 < \varphi_h \leq 1$ .

### 3.2 Time Discounting

Next we show that the timing of the information diffusion matters. We revisit the firm’s optimization problem with the discount factor  $\beta^t = \exp(-rt)$ , where  $r$  is the discount rate. Writing out the firm’s optimization problem:

$$\max_{\varphi_h, \varphi_l} \int_0^{t^*} \frac{dS}{dt} e^{-rt} dt$$

subject to

$$\varphi_h \bar{c} \leq V(\varphi_h, \varphi_l)$$

$$\varphi_l \bar{c} \leq V(\varphi_h, \varphi_l)$$

$$0 \leq \varphi_h \leq 1$$

$$0 \leq \varphi_l \leq 1$$

**Proposition 2.** *When discount rate  $r$  is large enough,*

$$\varphi_l > 0.$$

Hence, there is a tradeoff between the amount of information diffusion ( $S^*$ ) and the timing of diffusion – as long as 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  $\varphi_l > 0$  to achieve a higher level of information diffusion at early stage.

#### Discussion

The analysis above demonstrates that the firm faces the following trade-off: by confining the initial spread of information to high-status consumers only, it maximizes the individual’s incentive to spread the information. On the other hand, by spreading the information more widely (allowing some low-types access to it), the firm is guaranteeing itself an initial high level of adoption, but the total number of conversations over time decreases. Of course, if the firm is impatient, it prefers to make the information more widely accessible to consumers, even if they happen to be low-status. The trade-off can be summarized in Figure 1. That is, telling more people initially results in a lower over-all diffusion but a higher level of early diffusion of information, while telling few people results in higher overall level of diffusion in the long run.

## 4 Restricted Diffusion

In this subsection we consider a technology that the firm can employ which restricts who can credibly spread buzz about the product. For instance, Facebook initially only allowed individuals with email addresses from a small number of elite schools to adopt its product and without adopting the product, one cannot learn the information. We find that the firm can increase the diffusion of information by restricting who can spread information about the product. There are two potentially beneficial effects. The first is that if the restriction is biased towards high types in that they are more able to spread the information, then this will improve the inference from being able to spread the information. The second is that provided that the initial adopters aren’t restricted by the

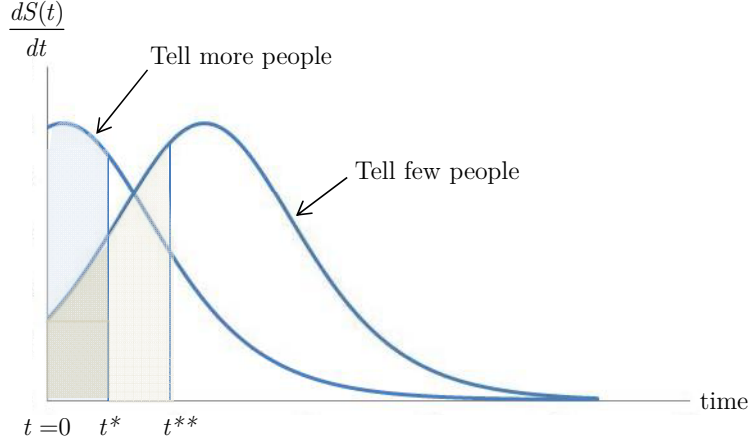


Figure 1: Diffusion of Information Based on Firm's Strategy

policy, then restricting the subsequent imitators during the diffusion of the information improves ex ante incentives to acquire information.

We assume that the firm can exclude some low types from being able to spread information about the product. The subset of the population who can spread information conditional on acquiring it is denoted by  $\Omega$ . We assume that all high types are in  $\Omega$ ,  $\Pr(i \in \Omega | \theta_i = h) = 1$ ; however, the firm excludes some low types,  $\Pr(i \in \Omega | \theta_i = l) < 1$ , and hence  $\phi \leq \Pr(i \in \Omega) = \omega < 1$ . We distinguish the fraction of the population which are informed by  $S^I(t)$  from the fraction which can diffuse information by  $S^R(t) = S^I(t) \cap \Omega$ . The informed population  $S^I(t)$  increases due to the word of mouth behavior of the population  $S^R(t)$ . In this section we take as given the results of the previous section that the firm increases the costs for low social types to acquire information ex ante. Rather we ask whether the firm can do better by restricting the set of individuals who may diffuse information about the product to  $\Omega$  and answer it in the affirmative. Therefore, even if the firm strictly prefers to change its information acquisition structure, it will only strengthen the result of this section. We denote the cutoff type of high value consumer by  $\varphi^\Omega$  and  $\varphi^*$  for the cases when the firm does and does not restrict the set of individuals who can diffuse the information, respectively. Similarly, define for each case the total diffusion by  $S^\Omega(\varphi_h)$  and  $S^*(\varphi_h)$  conditional on a level of ex ante information acquisition and denote beliefs by  $b^\Omega$  and  $b$ .

The equations describing the evolution of these variables over time as agents mix at rate  $\lambda$  are given by:

$$\frac{dS^I}{dt} = \lambda S^R(t) (1 - S^I(t)) \quad (17)$$

$$\frac{dS^R}{dt} = \lambda \omega S^R(t) (\omega - S^R(t)) \quad (18)$$

we proceed by proving three lemmas.

**Lemma 2.** Under initial conditions  $S_0^I = S_0^R = \varphi_h \phi$ . The diffusion path for  $S^R(t)$  and  $S^I(t)$  are

$$\begin{aligned} S^R(t) &= \frac{\omega}{1 + ae^{-\lambda\omega t}} \\ S^I(t) &= \frac{1}{1 + ae^{-\lambda\omega t}} - \frac{(1 - \omega)e^{-\lambda\omega t}}{1 + ae^{-\lambda\omega t}} \end{aligned}$$

where  $a = \frac{\omega - \varphi_h \phi}{\varphi_h \phi}$ .

*Proof.* See the Appendix. □

We now prove that under a restricted diffusion process the signaling value of information is greater.

**Lemma 3.** For any given  $S_0 = \varphi_h$ , the signaling value of information is greater when diffusion is restricted to  $\Omega$ :

$$b^\Omega(\theta_i = h|m, t) - b^\Omega(\theta_i = h|\emptyset, t) > b(\theta_i = h|m, t) - b(\theta_i = h|\emptyset, t)$$

*Proof.* See the Appendix. □

We now prove that the firm increases the value of information ex-ante by restricting the diffusion.

**Lemma 4.** For any given  $\varphi_h$ , the value of acquiring the information ex ante is greater when diffusion is restricted to  $\Omega$ :

$$V^\Omega(\varphi_h) > V(\varphi_h).$$

Thus,

$$S^\Omega(\varphi_h) > S^*(\varphi_h).$$

*Proof.* See the Appendix. □

We can now present our main result of this section. Even if we endogenize the amount of information acquisition at  $t = -1$ ,  $\varphi_h$ , we can show that when the firm restricts the set of people who can diffuse information to  $\Omega$ , the extent of the diffusion is greater.

**Proposition 3.** The diffusion of information is greater when the firm restricts the diffusion of information to individuals in  $\Omega$ ,

$$S^{\Omega*} > S^*.$$

*Proof.* See the Appendix. □

The firm benefits from restricting who can spread information about the product because it strengthens the signaling value of the information increasing both the ex ante incentives to acquire the information and the ex post incentives to spread it. This form of information management is consistent with marketing campaigns by Facebook and Spotify which endeavored to create artificial scarcity for their products by restricting who can initially adopt them and credibly spread information about them as an initial adopter rather than an imitator.

## 5 Advertising

In this section we allow the firm to engage in untargeted advertising. We find that advertising crowds out the incentives for individuals to acquire information and engage in word of mouth. Advertising is a source of information which is not correlated with an individual's type and so weakens the signaling value of the information. We find that a higher cost of advertising for the firm leads to a greater diffusion of information. Higher costs of advertising are a credible commitment by the firm not to advertise widely which crowds out the consumers' incentives to acquire and spread information. We discuss a source of commitment available to a firm, to release information a sufficient amount of time prior to the product release, and find that it is optimal for a firm to release information a sufficient amount of time prior to the product release and undertake any advertising only after word of mouth has spread. This accords well with observed behavior of firms which successfully generate buzz for a product.

### Setup

In this section, as before, consumers choose whether or not to acquire information at  $t = -1$ . We simplify the exposition by assuming that the firm maximizes  $\varphi_h$  subject to the incentive constraint by choosing  $v_h = 0$  and minimizes  $\varphi_l$  by setting it equal to 0 through choosing a large  $v_l$ . At  $t = 0$  the firm has access to a costly advertising technology. The firm chooses the fraction  $\beta$  of the population to advertise to at cost  $C(\beta) \geq 0$ ,  $C'(\beta) > 1$ ,  $C''(\beta) > 0$  for all  $\beta \geq 0$ . The technology is untargeted so is equally likely to inform high or low types. We assume that consumers only observe whether they themselves receive or don't receive the advertisement, in particular consumers do not observe the level of  $\beta$  but must infer it in equilibrium. Consumers then meet at rate  $\lambda$  from  $t = 0$  onwards until the information diffusion stops at the moment when the signaling value is equal to the cost of passing on word of mouth. Consumers do not observe  $\beta$  so form beliefs and make inferences during the mixing based on a conjectured level of advertising  $\tilde{\beta}$  by the firm. Consumers make inferences based on this conjecture and so the length of time  $\tau(\varphi_h, \tilde{\beta})$  that consumers spread the word of mouth is a function of beliefs based on this conjectured level of advertising by the firm  $\tilde{\beta}$  and the actual level of information acquisition  $\varphi_h$  at  $t = -1$ . Similarly the amount of advertising  $\beta(\varphi_h, \tilde{\tau})$  a firm will undertake will also depend on how long they conjecture consumers are prepared to pass on word of mouth  $\tilde{\tau}$  and the actual level of information acquisition  $\varphi_h$  at  $t = -1$ .<sup>13</sup>

We solve for a subgame perfect equilibrium of the game. We work backwards from the period  $t \geq 0$  where consumers choose a time  $\tau(\varphi_h, \tilde{\beta})$  at which they will stop spreading word of mouth.

This is a function of beliefs based on the conjectured level of advertising by the firm  $\tilde{\beta}$  and the actual level of information acquisition  $\varphi_h$  at  $t = -1$ . The firm chooses the level of advertising at  $t = 0$  however the level is unobserved by consumers. Consumers' beliefs evolve based on a conjectured level of advertising undertaken by the firm  $\tilde{\beta}$ . This belief is not influenced directly by the actual level of advertising undertaken by the firm for  $\beta \in (0, 1)$  in a subgame perfect equilibrium. The natural way to analyze it is thus a simultaneous move game between the firm and consumers. However if  $\tilde{\beta} = 0, 1$  we need to specify off-equilibrium beliefs for consumers whereby a consumer may observe or not observe advertising when their conjectured beliefs are that no advertising or 100% advertising

<sup>13</sup>The assumption that  $\varphi_h$  is observable to other consumers is without consequence. An individual consumer has infinitesimal effect on  $\varphi_h$  and under an assumption that consumers must form beliefs about  $\varphi_h$  an identical equilibrium would arise.

takes place. In Appendix B we consider a “tremble” in the firm’s advertising strategy, whereby it chooses an advertising level  $\beta$ , then with probability  $\epsilon$ , individuals in the fraction of the population advertised to, do not receive the advertisement, and with the same probability  $\epsilon$ , the fraction not advertised to, do receive the advertisement. Thus, when the firm chooses a level  $\beta \in [0, 1]$ , the actual fraction that receive the advertisement is  $\beta(1 - \epsilon) + \epsilon(1 - \beta) = \beta + \epsilon - 2\epsilon\beta$ . We show that the limit of the “trembling” equilibria for  $\epsilon \rightarrow 0$  correspond to the signaling equilibria, we find in our subsequent analysis, where consumers do not change their beliefs in the off-equilibrium event of receiving an advertisement when  $\beta^* = 0$ . We proceed by assuming that when a consumer with  $\tilde{\beta} = 0$  sees an advertisement their beliefs are unchanged, with the understanding that the equilibria we find are the limits of these “tremble” equilibria. For the purposes of analyzing the subgame from  $t = 0$ , we may thus analyze a simultaneous move game where the firm and consumers choose  $\beta(\varphi_h, \tilde{\tau})$  and  $\tau(\varphi_h, \tilde{\beta})$  respectively.

The equilibrium of the subgame for a given level of information acquisition is a pair of strategies  $\{\tau^*(\varphi_h), \beta^*(\varphi_h)\} = \{\tau(\varphi_h, \beta^*(\varphi_h)), \beta(\varphi_h, \tau^*(\varphi_h))\}$  where the firm and consumers choose best responses to each others actions in the usually Nash equilibrium sense. We define a sub-game perfect equilibrium by a triple  $\{\varphi_h^*, \tau^*(\varphi_h^*, \beta^*), \beta^*(\varphi_h^*, \tau^*)\}$  where the equilibrium level of information acquisition is a cutoff  $\varphi_h^*$  satisfies:

$$\begin{aligned} \varphi_h^* \bar{c} = V(\varphi_h^*, \tau^*(\varphi_h^*, \beta^*), \beta^*(\varphi_h^*, \tau^*)) & \quad \text{if} \quad V(1, \tau^*(1, \beta^*), \beta^*(1, \tau^*)) < \bar{c} \\ \varphi_h^* = 1 & \quad \text{if} \quad V(1, \tau^*(1, \beta^*), \beta^*(1, \tau^*)) \geq \bar{c}. \end{aligned} \quad (19)$$

Where  $V(\varphi_h^*, \tau^*(\varphi_h^*, \beta^*), \beta^*(\varphi_h^*, \tau^*))$  is the value for a consumer from acquiring information at  $t = -1$ . As before, if there are multiple solutions to the above condition, we assume that the firm can implement the solution which results in the greatest level of information diffusion.

We continue by now characterizing the best response functions of the consumer and firm.

### Subgame $t \geq 0$ : Consumers’ Best Response

Consumers inference of the level of advertising by the firm affects how their beliefs evolve over time about the signaling value of information and hence how long consumers will continue to spread the information. The best response for consumers is to continue spreading information, conditional upon having the information, until the signaling value of information is no longer greater than the cost spread the information  $k$ . This time  $\tau(\varphi_h, \tilde{\beta})$  is given by

$$\tau(\varphi_h, \tilde{\beta}) = \frac{1}{\lambda} \ln \frac{S_c^*(\varphi_h, 0, \tilde{\beta})}{1 - S_c^*(\varphi_h, 0, \tilde{\beta})} \frac{1 - S_c^0(\varphi_h, \tilde{\beta})}{S_c^0(\varphi_h, \tilde{\beta})} \quad (20)$$

for  $0 \leq \tilde{\beta} \leq \frac{\varphi_h \phi}{k(1 - \varphi_h \phi)} \left(1 - k - \frac{(1 - \varphi_h) \phi}{1 - \varphi_h \phi}\right)$ , otherwise  $\tau(\varphi_h, \tilde{\beta}) = 0$ . Where the conjectured initial level of diffusion is:

$$S_c^0(\varphi_h, \tilde{\beta}) = \varphi_h \phi + \tilde{\beta} ((1 - \varphi_h) \phi + 1 - \phi)$$

and the conjectured total diffusion  $S_c^*(\varphi_h, \tilde{\beta})$  is:

$$S_c^*(\varphi_h, \tilde{\beta}) = \frac{\varphi_h \phi + \tilde{\beta} (1 - \varphi_h \phi)}{k} \left[ \frac{\varphi_h \phi + \tilde{\beta} (1 - \varphi_h) \phi}{\varphi_h \phi + \tilde{\beta} (1 - \varphi_h \phi)} - \frac{\phi (1 - \varphi_h)}{1 - \varphi_h \phi} \right] \quad (21)$$

which, as before, is determined by the level of diffusion where the signaling value of information is equal to the cost of spreading information.

We now show that for a given level of information acquisition by consumers  $\varphi_h$ , the conjectured level of diffusion  $S_c^*(\varphi_h, \tilde{\beta})$  at which individuals cease to spread information is independent of the conjectured level of advertising  $\tilde{\beta}$  in the following lemma.

**Lemma 5.** For any given  $\varphi_h$ ,  $S_c^*(\varphi_h, \tilde{\beta}) = S_c^*(\varphi_h, 0)$  for all  $0 \leq \tilde{\beta} \leq \frac{1}{k} \left[ \varphi_h \phi - \frac{\phi^2 \varphi_h (1 - \varphi_h)}{1 - \varphi_h \phi} \right] - \frac{\varphi_h \phi}{1 - \varphi_h \phi}$ .

The information diffusion is governed by the signaling value, which is a function of the initial composition of high and low types consumers who acquired the information at  $t = -1$ . Since the advertising is untargeted, the increased population through the advertising does not change the initial composition of customer types and thus does not affect the consumer inference about the type. In this sense, the role of advertising has the same affect on the signaling value of information as the word of mouth diffusion as both are untargeted. Hence, the customer's belief updating is governed by the same mechanism and the advertising is purely substituting the word of mouth process and it does not affect the extent of diffusion ultimately.

Although the advertising does not affect the conjectured extent of information diffusion  $S_c^*$  it does affect the conjectured length of time until that level is reached and so does affect the consumer's best response.

**Lemma 6.** The conjectured level of advertising reduces the time at which consumers stop engaging in word of mouth:  $\frac{d\tau}{d\tilde{\beta}} < 0$ .

The conjectured level of advertising  $\tilde{\beta}$  has no direct effect on  $S_c^*$  and thus has no effect on  $\tau$  through this term. The only effect comes through the initial level of information  $S_c^0(\varphi_h, \tilde{\beta})$ . The timer  $\tau(\varphi_h, \tilde{\beta})$  is strictly decreasing in  $S_0$  and is also strictly decreasing in the conjectured level of advertising  $\tilde{\beta}$ .

#### Subgame $t \geq 0$ : Firm's Best Response Function

A firm's incentive to undertake advertising balances the marginal costs of advertising against the marginal impact on the firm's conjectured level of diffusion  $S_f^*$ . This conjectured level of diffusion depends on the firm's conjecture of the length of time consumers engage in word of mouth  $\tilde{\tau}$  and its own level of advertising:

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

The function  $S_f^*(\varphi_h, \beta, \tilde{\tau})$  may 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  $\tilde{\tau}$

$$S_f^*(\varphi_h, \beta, \tilde{\tau}) = \frac{1}{1 + ae^{-\lambda \tilde{\tau}}}, \quad \text{where } a = \frac{1 - S^0(\varphi_h, \beta)}{S^0(\varphi_h, \beta)}. \quad (22)$$

A firm is able to influence  $S_f^*$  through the initial informed share of the population  $S^0(\varphi_h, \beta)$ , but cannot directly influence  $\tilde{\tau}$ , which is only affected by consumers' expectations of the level advertising  $\tilde{\beta}$  in equilibrium. The level of advertising affects  $S^0$  through the following relationship:

$$S^0(\varphi_h, \beta) = \varphi_h \phi + \beta((1 - \varphi_h) \phi + 1 - \phi). \quad (23)$$

Note we have dropped the argument for  $\varphi_l$  from earlier as we set it to 0. Now the marginal effect of advertising on  $S^0$  is

$$\frac{dS^0}{d\beta} = (1 - \varphi_h) \phi + 1 - \phi = 1 - \varphi_h \phi$$



Now taking the derivative of equation (22) with respect to  $S^0$  while holding  $\tilde{\tau}$  constant

$$\frac{dS_f^*}{dS^0} = \frac{e^{-\lambda\tilde{\tau}}}{[S^0 + (1 - S^0)e^{-\lambda\tilde{\tau}}]^2}.$$

The marginal effect of advertising on  $S_f^*$  is, therefore:

$$\frac{dS_f^*}{d\beta} = \frac{dS_f^*}{dS^0} \frac{dS^0}{d\beta} = \frac{e^{-\lambda\tilde{\tau}} (1 - \varphi_h\phi)}{[S^0 + (1 - S^0)e^{-\lambda\tilde{\tau}}]^2}.$$

**Lemma 7.** If  $C'(0) < \frac{1}{2\varphi_h\phi}$  then  $\exists \hat{\tau} < \frac{1}{\lambda} \ln \frac{1 - \varphi_h\phi}{\varphi_h\phi}$  such that the firm's optimal choice of advertising  $\beta(\varphi_h, \tilde{\tau})$  for a level of information acquisition by consumers  $\varphi_h$  and conjectured spreading time  $\tilde{\tau}$  satisfies

$$C'(\beta) = \frac{e^{-\lambda\tilde{\tau}} (1 - \varphi_h\phi)}{[S^0 + (1 - S^0)e^{-\lambda\tilde{\tau}}]^2} \text{ for } \hat{\tau} \leq \tilde{\tau} \leq \frac{1}{\lambda} \ln \frac{1 - \varphi_h\phi}{\varphi_h\phi} \quad (24)$$

and  $\beta = 0$  for  $0 \leq \tilde{\tau} < \hat{\tau}$ . Otherwise  $\beta = 0$  for  $0 \leq \tilde{\tau} \leq \frac{1}{\lambda} \ln \frac{1 - \varphi_h\phi}{\varphi_h\phi}$ .

Next, we find that there is a unique Nash equilibrium of the subgame which is continuous in  $\varphi_h$ .

**Lemma 8.** Suppose  $k > 2\phi$ . The equilibrium of the subgame  $\{\tau^*(\varphi_h), \beta^*(\varphi_h)\}$  is unique and continuous in  $\varphi_h$  for  $\varphi_h \in (0, 1]$ .

We further characterize the equilibrium of this subgame in the following lemma. We find that the firm has a higher marginal cost of advertising, the equilibrium of the subgame has a longer diffusion time and lower advertising. Before doing so it is useful to define a cutoff type  $\varphi_h^{**}$  which satisfies the following condition:

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

**Lemma 9.** Suppose  $k > 2\phi$ . Consider two cost functions of advertising  $C_1$  and  $C_2$  where  $C_1' < C_2'$ . Then  $\beta_1^*(\varphi_h) \geq \beta_2^*(\varphi_h)$  and  $t_1^*(\varphi_h) \leq t_2^*(\varphi_h)$  where the inequalities are strict if  $C_1'(0) < \frac{1 - \phi}{k} \left[ 1 - \frac{\varphi_h\phi(1 - \phi)}{k} \right]$ .

In the Nash equilibrium both the firm and the consumers correctly anticipate the equilibrium extent of diffusion  $S^*(\varphi_h)$ , thus  $S_c^*(\varphi_h, \beta^*) = S_f^*(\varphi_h, \tau^*, \beta^*) = S^*(\varphi_h)$ . We also note that for a given equilibrium level of information acquisition,  $\varphi_h$ , the equilibrium level of advertising,  $\beta^*$ , does not affect the equilibrium extent of diffusion  $S^*(\varphi_h)$  in the subgame from  $t \geq 0$ . Consumers' expectations of the level of advertising are correct in equilibrium and from Lemma 5  $S_c^*(\varphi_h, \beta^*)$  is independent of  $\beta^*$  so the equilibrium extent of diffusion in the subgame  $S^*(\varphi_h)$  is independent of the level of  $\beta^*$  and only depends on the level of initial information acquisition of consumers.

Figure 2 illustrates the role of untargeted advertising. In the subgame advertising partially crowds out the diffusion of the information that would otherwise take place via word of mouth but does not change the level of diffusion at which individuals stop spreading the information. However, in the full game, advertising will affect the ex ante incentive for consumers to acquire the information, and thus the level of information acquired by consumers at  $t = -1$ .

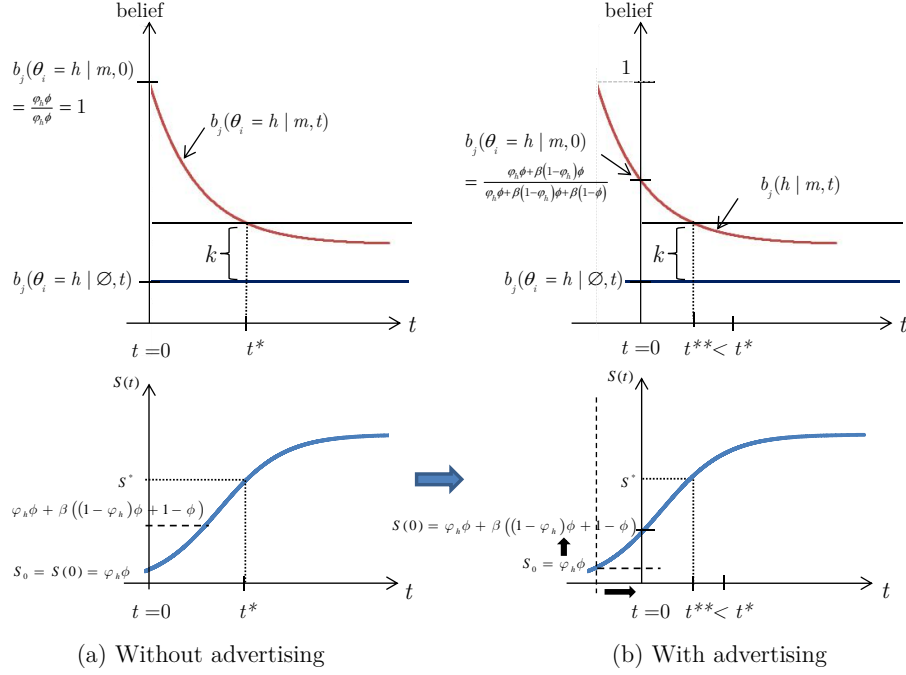


Figure 2: The Role of Untargeted Advertising for a given  $\varphi_h$

## Subgame Perfect Equilibrium

We now consider the ex ante incentives for consumers to acquire information. The value of information acquisition for consumers is

$$V(\varphi_h, \tau^*(\varphi_h), \beta^*(\varphi_h)) = (1 - \beta^*(\varphi_h)) \frac{k}{1 - S^0(\varphi_h, \beta^*(\varphi_h))} \left[ \frac{S^*(\varphi_h)}{S^0(\varphi_h, \beta^*(\varphi_h))} - 1 - \ln \frac{S_c^*(\varphi_h)}{S^0(\varphi_h, \beta^*(\varphi_h))} \right]. \quad (25)$$

We now show that the value of information acquisition is continuous in the amount of information consumers acquire.

**Lemma 10.** *The value of information acquisition  $V(\varphi_h, \tau^*(\varphi_h), \beta^*(\varphi_h))$  is continuous in  $\varphi_h$  for  $\varphi_h \in (0, 1]$ .*

We now establish a sufficient condition under which there exists a sub-game perfect equilibrium of the model.

**Proposition 4.** *Suppose  $k > 2\phi$  then there exists a sub-game perfect equilibrium.*

The conjectured level of advertising by the firm affects the value of information and thus, the ex ante incentive for consumers to acquire the information at  $t = -1$ . The following proposition gives the main result of the section. The extent of diffusion in the best signaling equilibrium is weakly greater when the firm has larger marginal costs of advertising and provides a condition where the inequality is strict.

**Proposition 5.** *Suppose  $k > 2\phi$  and  $\bar{c} \geq \frac{1-k+k \ln k}{1-\phi}$ . Consider two cost functions of advertising  $C_1$  and  $C_2$  where  $1 \leq C'_1 < C'_2$  then,  $\varphi_{h1}^* \leq \varphi_{h2}^*$ ,  $\beta_1^* \geq \beta_2^*$ ,  $t_1^* \leq t_2^*$  and*

$$S^*(\varphi_{h1}^*) \leq S^*(\varphi_{h2}^*)$$

where the inequalities are strict if  $C'_1(0) < \frac{S^*(\varphi_h^{**})(1-S^*(\varphi_h^{**}))}{S^0(\varphi_h^{**},0)(1-S^0(\varphi_h^{**},0))} (1 - \varphi_h^{**}\phi)$ .

The result shows that a higher marginal cost of advertising may lead to an unambiguously greater diffusion of information for the firm. The commitment not to advertise is therefore valuable to the firm. Advertising reduces the signaling benefit for an agent  $V$  and thus, reduces the ex ante incentive for consumers to acquire the information. Hence, advertising acts to crowd out the incentives for consumers to search for information ex ante. The equilibrium level of information acquisition  $\varphi_h^*$ , which is determined in equation (19) decreases since the RHS also decreases as the level of advertising increases ( $\frac{dV}{d\beta} < 0$ ).

The next proposition shows that a source of commitment for a firm is the release date of the product. For instance movie theatres 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 it can maximize the diffusion of information due to word of mouth.

**Proposition 6.** *Suppose  $T > \frac{1}{r} \ln \frac{S^*(\varphi_h^{**})(1-S^*(\varphi_h^{**}))}{S^0(\varphi_h^{**},0)(1-S^0(\varphi_h^{**},0))} \frac{((1-\varphi_h^{**})\phi+1-\phi)}{C'(0)}$  and  $T > \tau^*(\varphi_h^{**})$  then the equilibrium is  $\{\varphi_h^{**}, \tau^*(\varphi_h^{**}), 0\}$ .*

The proposition highlights that early information release can serve as a commitment not to undertake advertising during the period of time that individuals are undertaking word of mouth. Of course the firm could 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 that many instance of buzz/word of mouth occur prior to any advertising campaign by a firm.

## 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 in a population of consumers 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 social type. In our model a firm constrains low-social types' access to information and their ability to pass on information to other consumers, in order to maintain the signaling value of passing on information. 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 may also benefit 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.

## 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. Glibert, 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] 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.
- [8] 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.
- [9] 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.
- [10] Dye, Renee (2000), "The Buzz on Buzz," *Harvard Business Review*, Nov-Dec, 138-146.
- [11] 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.
- [12] Hughes, Mark (2005), "*Buzzmarketing: Get People to Talk About Your Stuff*," Penguin Group, New York, NY.
- [13] Kotler, Philip (2002), "*A Framework for Marketing Management*," Prentice Hall, US.
- [14] Kuksov, Dmitri (2007), "Brand Value in Social Interaction," *Management Science*, 53 (10), 1634-1644.
- [15] Luca, Michael (2011), "Reviews, Reputation, and Revenue: The Case of Yelp.com," *Harvard Business School Working Paper*, No 12-016.
- [16] 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.

- [17] Pesendorfer, W. (1995), "Design Innovation and Fashion Cycles," *American Economic Review*, 85 (4), 771-792.
- [18] Sundaram, D.S., Kaushik Mitra and Cynthia Webster (1998), "Word-of-Mouth Communications: A Motivational Analysis," *Advances in Consumer Research*, 25, 527-531.
- [19] Wojnicki, Andrea and David Godes (2011), "Signaling Success: Strategically-Positive Word of Mouth," University of Maryland, *Working Paper*.
- [20] Word of Mouth Marketing Association (2011) infographic: [http://womma.org/word/wp-content/uploads/2011/10/word\\_of\\_mouth\\_marketing\\_impact\\_and\\_influence\\_womma.png](http://womma.org/word/wp-content/uploads/2011/10/word_of_mouth_marketing_impact_and_influence_womma.png)
- [21] Young, H. Peyton (2009), "Innovation Diffusion in Heterogeneous Populations: Congation, Social Influence, and Social Learning," *American Economic Review*, 99 (5), 1899-1924.

## A Proofs

### A.1 Proof of Lemma1

*Proof.* The first order conditions of  $V$  with respect to  $S^*$  and  $S_0$  are:

$$\begin{aligned}\frac{\partial V}{\partial S^*} &= \left( \frac{k}{1-S_0} \right) \left[ \frac{1}{S_0} - \frac{1}{S^*} \right], \\ \frac{\partial V}{\partial S_0} &= \left[ \begin{aligned} &k \left( \frac{1}{1-S_0} \right)^2 \left( S^* \frac{1}{S_0} - 1 + \ln S_0 - \ln S^* \right) \\ &+ \left( \frac{k}{1-S_0} \right) \left( -\frac{1}{S_0^2} \frac{BS_0}{k} + \frac{1}{S_0} \right) \end{aligned} \right] \\ &= \left[ \frac{k}{(1-S_0)^2 S_0^2} \left\{ (S^* - S_0)(2S_0 - 1) + S_0^2 \left( \ln \frac{S_0}{S^*} \right) \right\} \right].\end{aligned}$$

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  $S_0 < \frac{1}{2}$ . exist

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  $x = \frac{S^*}{S_0} > 1$ .

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

which is greater than 0 for  $x > 1$  if

$$\begin{aligned}\frac{\ln x}{x-1} - \frac{1}{x} &> 0 \\ \Leftrightarrow \ln x &> 1 - \frac{1}{x}\end{aligned}$$

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

$$-\frac{\ln \frac{1}{S_0}}{\frac{1}{S_0}-1} < \frac{1}{S_0} - 2$$

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

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

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

and note that

$$\lim_{y \rightarrow 1} [-\ln y - (y-2)(y-1)] = 0.$$

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

## A.2 Proof of Proposition 1

*Proof.* First, 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  $\frac{dV}{d\varphi_l}$ ,

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

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

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

We already know that  $S^*$  is maximized when  $\varphi_h = 1$ , independent of  $\varphi_l$ . 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  $\varphi_l$  does not help alleviate 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  $\frac{dV}{dS_0} < 0$ .

$$\begin{aligned}\lim_{S_0 \rightarrow 0} V &= \lim_{S_0 \rightarrow 0} \left( \frac{k}{1-S_0} \right) \left( \frac{BS_0}{k} \frac{1}{S_0} - 1 + \ln S_0 - \ln \frac{BS_0}{k} \right) \\ &= k \left( \frac{B}{k} - 1 - \ln \frac{B}{k} \right) > 0. \\ \lim_{S_0 \rightarrow S^*} V &= 0\end{aligned}$$

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

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

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

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

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

### A.3 Proof of Proposition 2

*Proof.* Let

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

We have immediately that

$$\begin{aligned}\lim_{r \rightarrow 0} R(\varphi_l, \varphi_h) &= S^*(\varphi_l, \varphi_h) \\ \lim_{r \rightarrow \infty} R(\varphi_l, \varphi_h) &= S_0(\varphi_l, \varphi_h).\end{aligned}$$

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

Furthermore,

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

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

#### A.4 Proof of Lemma 2

*Proof.* The derivation of  $S^R(t)$  is the same as earlier for  $S(t)$  taking it as given substitute it into the equation for  $S^I(t)$ :

$$\begin{aligned} \frac{dS^I}{dt} &= \lambda \frac{\omega}{1 + ae^{-\lambda\omega t}} (1 - S^I(t)) \Leftrightarrow \frac{dS^I}{(1 - S^I(t))} = \lambda \frac{\omega}{1 + ae^{-\lambda\omega t}} dt \\ &\Leftrightarrow -\ln(1 - S^I(t)) = \ln(-a - e^{\lambda\omega t}) + \ln(A) \Leftrightarrow \frac{1}{(1 - S^I(t))} = -A(a + e^{\lambda\omega t}) \\ &\Leftrightarrow 1 - S^I(t) = \frac{-1}{A(a + e^{\lambda\omega t})}. \end{aligned}$$

Redefine  $A = \frac{-1}{d}$ ;

$$S^I(t) = 1 - \frac{d}{a + e^{\lambda\omega t}}$$

solving for  $d$  using the initial condition ( $S_0^I = \varphi_h \phi$ ) and  $a = \frac{\omega - \varphi_h \phi}{\varphi_h \phi}$ , we find

$$\varphi_h \phi = 1 - \frac{d}{\frac{\omega - \varphi_h \phi}{\varphi_h \phi} + 1} \Leftrightarrow d = \frac{\omega(1 - \varphi_h \phi)}{\varphi_h \phi}.$$

Substituting this back in

$$\begin{aligned} S^I(t) &= \frac{a - d + e^{\lambda\omega t}}{a + e^{\lambda\omega t}} = \frac{e^{\lambda\omega t} - 1 + \omega}{a + e^{\lambda\omega t}} \\ &= \frac{1 - (1 - \omega)e^{-\lambda\omega t}}{1 + ae^{-\lambda\omega t}} \\ &= \frac{1}{1 + ae^{-\lambda\omega t}} - \frac{(1 - \omega)e^{-\lambda\omega t}}{1 + ae^{-\lambda\omega t}}. \end{aligned}$$

$\square$

#### A.5 Proof of Lemma 3

*Proof.* Note that for any  $t \geq 0$ ,

$$b^\Omega(\theta_i = h|m, t) S^R(t) + b^\Omega(\theta_i = h|\emptyset, t) (1 - S^R(t)) = b(\theta_i = h|m, t) S^I(t) + b(\theta_i = h|\emptyset, t) (1 - S^I(t)) = \phi,$$

where  $S^R(t) \leq S^I(t)$ ,  $b^\Omega(\theta_i = h|m, t) > b^\Omega(\theta_i = h|\emptyset, t)$ , and  $b(\theta_i = h|m, t) > b(\theta_i = h|\emptyset, t)$ .

First, we show that

$$b^\Omega(\theta_i = h|\varnothing, t) < b(\theta_i = h|\varnothing, t).$$

The inference from not receiving a signal is:

$$\begin{aligned} b^\Omega(\theta_i = h|\varnothing, t) &= \frac{1 - S^I(t)}{1 - S^R(t)} \times b(\theta_i = h|\varnothing, t) + \frac{S^I(t) - S^R(t)}{1 - S^R(t)} \times 0 \\ &= \frac{1 - S^I(t)}{1 - S^R(t)} \times b(\theta_i = h|\varnothing, t) < b(\theta_i = h|\varnothing, t). \end{aligned}$$

Next, from the fact that  $b^\Omega(\theta_i = h|m, t) S^R(t) + b^\Omega(\theta_i = h|\varnothing, t) (1 - S^R(t)) = \phi$ :

$$\begin{aligned} b^\Omega(\theta_i = h|m, t) &= \frac{1}{S^R(t)} [\phi - b^\Omega(\theta_i = h|\varnothing, t) (1 - S^R(t))] \\ &= \frac{1}{S^R(t)} [\phi - b(\theta_i = h|\varnothing, t) (1 - S^I(t))] \\ &> \frac{1}{S^I(t)} [\phi - b(\theta_i = h|\varnothing, t) (1 - S^I(t))] = b(\theta_i = h|m, t). \end{aligned}$$

□

## A.6 Proof of Lemma 4

*Proof.* The value of acquiring a signal ex ante when diffusion is unrestricted is

$$V(\varphi_h) = \lambda \left( \int_{S_0}^{S^*(\varphi_h)} \left( \frac{1}{1 - \varphi_h \phi} \right) \frac{1}{S(t)} ([b(\theta_i = h|m, t) - b(\theta_i = h|\varnothing, t)] - k) dS \right).$$

The value of acquiring a signal when diffusion of information is restricted is

$$V^\Omega(\varphi_h) = \lambda \left( \int_{S_0}^{S^\Omega(\varphi_h)} \left( \frac{1}{1 - \varphi_h \phi} \right) \frac{1}{S(t)} ([b^\Omega(\theta_i = h|m, t) - b^\Omega(\theta_i = h|\varnothing, t)] - k) dS \right),$$

where the term inside the integral is positive at all points along the path. Now, from Lemma ?? we have  $b^\Omega(\theta_i = h|m, t) - b^\Omega(\theta_i = h|\varnothing, t, \varphi_h) > b(\theta_i = h|m, t) - b(\theta_i = h|\varnothing, t)$ , and hence,  $S^\Omega(\varphi_h) > S^*(\varphi_h)$  for a given  $\varphi_h$ . Thus,  $V^\Omega(\varphi_h) > V(\varphi_h)$ . □

## A.7 Proof of Proposition 3

*Proof.* The result now follows from noting that

$$\begin{aligned} S^*(\varphi_h^*) &< S^*(\varphi_h^\Omega) \\ S^*(\varphi_h^\Omega) &< S^\Omega(\varphi_h^\Omega) \end{aligned}$$

where the first inequality follows by noting that an immediate consequence of lemma 4 is that the cutoff type is greater when diffusion is restricted  $\varphi_h^\Omega > \varphi_h^*$  and from earlier  $\frac{dS^*}{d\varphi_h} > 0$ . The second inequality follows immediately from lemma 3. □

## A.8 Proof of Lemma 5

*Proof.* Using Equation 21

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

□

## A.9 Proof of Lemma 6

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

$$\tau (\varphi_h, \tilde{\beta}) = \frac{1}{\lambda} \ln \frac{S_c^* (\varphi_h^*, \tilde{\beta})}{1 - S_c^* (\varphi_h^*, \tilde{\beta})} \frac{1 - S^0 (\varphi_h, \tilde{\beta})}{S^0 (\varphi_h, \tilde{\beta})}$$

where

$$S^0 (\varphi_h, \tilde{\beta}) = \varphi_h \phi + \tilde{\beta} ((1 - \varphi_h) \phi + 1 - \phi).$$

and we know from Lemma 5 that  $\frac{dS_c^*}{d\tilde{\beta}} = 0$ . It is now straightforward to find the derivative from

$$\frac{d\tau^*}{d\tilde{\beta}} = \frac{d\tau^*}{dS^0} \frac{dS^0}{d\tilde{\beta}}:$$

$$\frac{d\tau^*}{d\tilde{\beta}} = -\frac{1 - \varphi_h \phi}{\lambda} \left[ \frac{1}{1 - S^0} + \frac{1}{S^0} \right] < 0 \tag{26}$$

□

## A.10 Proof of Lemma 7

*Proof.* The best response function is defined as:

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

Also note that

$$\frac{d^2 S_f^*}{d\beta^2} = -2 \frac{e^{-\lambda \tilde{\tau}} (1 - e^{-\lambda \tilde{\tau}}) (1 - \varphi_h \phi)^2}{[S^0 + (1 - S^0) e^{-\lambda \tilde{\tau}}]^3} < 0$$

and by assumption  $C'' (\beta) > 0$  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  $\beta$  that satisfies  $C' (\beta) =$

$\frac{e^{-\lambda\tilde{\tau}}(1-\varphi_h\phi)}{[S^0+(1-S^0)e^{-\lambda\tilde{\tau}}]^2}$  then this is the best response. Also note that  $C'(1) > 1 > e^{-\lambda\tilde{\tau}}(1-\varphi_h\phi)$  such that  $\beta = 1$  is never a best response. We now show that there exists a  $\hat{\tau}$  such that  $\exists\beta$  that satisfies  $C'(\beta) = \frac{e^{-\lambda\tilde{\tau}}(1-\varphi_h\phi)}{[S^0+(1-S^0)e^{-\lambda\tilde{\tau}}]^2}$  for  $\tau \geq \hat{\tau}$  and  $C'(0) \geq \frac{e^{-\lambda\tilde{\tau}}(1-\varphi_h\phi)}{[\varphi_h\phi+(1-\varphi_h\phi)e^{-\lambda\tilde{\tau}}]^2}$  for  $\tau \leq \hat{\tau}$ . We define  $\hat{\tau}$  as the time which satisfies  $C'(0) = \frac{e^{-\lambda\tilde{\tau}}(1-\varphi_h\phi)}{[\varphi_h\phi+(1-\varphi_h\phi)e^{-\lambda\tilde{\tau}}]^2}$ . The rhs at  $\tilde{\tau} = 0$  is  $1 - \varphi_h\phi \leq 1$  furthermore the rhs is concave and maximized at  $\tilde{\tau} = \frac{1}{\lambda} \ln \frac{1-\varphi_h\phi}{\varphi_h\phi}$  at a value of  $\frac{1}{2\varphi_h\phi}$  thus if  $C'(0) < \frac{1}{2\varphi_h\phi}$  then  $\exists\tilde{\tau} < \frac{1}{\lambda} \ln \frac{1-\varphi_h\phi}{\varphi_h\phi}$  such that  $C'(0) \geq \frac{e^{-\lambda\tilde{\tau}}(1-\varphi_h\phi)}{[\varphi_h\phi+(1-\varphi_h\phi)e^{-\lambda\tilde{\tau}}]^2}$  and  $\beta = 0$  for  $0 \leq \tilde{\tau} < \hat{\tau}$  and  $C'(0) < \frac{e^{-\lambda\tilde{\tau}}(1-\varphi_h\phi)}{[\varphi_h\phi+(1-\varphi_h\phi)e^{-\lambda\tilde{\tau}}]^2}$  for  $\hat{\tau} < \tilde{\tau} \leq \frac{1}{\lambda} \ln \frac{1-\varphi_h\phi}{\varphi_h\phi}$  and hence  $\exists\beta > 0$  which satisfies  $C'(\beta) = \frac{e^{-\lambda\tilde{\tau}}(1-\varphi_h\phi)}{[S^0+(1-S^0)e^{-\lambda\tilde{\tau}}]^2}$ . Otherwise  $C'(0) \geq \frac{e^{-\lambda\tilde{\tau}}(1-\varphi_h\phi)}{[\varphi_h\phi+(1-\varphi_h\phi)e^{-\lambda\tilde{\tau}}]^2}$  and  $\beta = 0$  for  $0 \leq \tilde{\tau} \leq \frac{1}{\lambda} \ln \frac{1-\varphi_h\phi}{\varphi_h\phi}$ .  $\square$

### A.11 Proof of Lemma 8

*Proof.* When the best response functions  $\{\tau(\varphi_h, \beta), \beta(\varphi_h, \tau)\}$  are continuous and map a compact set to a compact set then there exists an equilibrium. By definition  $\beta$  is bounded above by 1, also an upper bound on  $\tau$  for a given  $\varphi_h$  is  $\tau(\varphi_h, 0) = \frac{1}{\lambda} \ln \frac{1-\varphi_h\phi}{\frac{1-k}{1-\phi}-\varphi_h\phi}$ . Both best response functions are continuous over the appropriate domains so we have established the existence of an equilibrium. We also know that  $\tau(\varphi_h, \tilde{\beta}) = 0$  for  $\tilde{\beta} \geq \frac{\varphi_h\phi}{k(1-\varphi_h\phi)} \left(1 - k - \frac{(1-\varphi_h\phi)}{1-\varphi_h\phi}\right)$  as this level of advertising will result in  $S^0 \geq S_c^*$ . We note that from Lemma 6 that  $\frac{d\tau^*}{d\beta} < 0$  for  $0 \leq \tilde{\beta} \leq \frac{\varphi_h\phi}{k(1-\varphi_h\phi)} \left(1 - k - \frac{(1-\varphi_h\phi)}{1-\varphi_h\phi}\right)$  and  $\tau = 0$  for  $\tilde{\beta} \geq \frac{\varphi_h\phi}{k(1-\varphi_h\phi)} \left(1 - k - \frac{(1-\varphi_h\phi)}{1-\varphi_h\phi}\right)$ , also that  $\beta(\varphi_h, 0) = 0$  so the equilibrium  $\tau^* > 0$ . Thus at the equilibrium point  $\frac{d\tau^*}{d\beta} < 0$ . Using Lemma 7 if  $C'(0) \geq \frac{1}{2\varphi_h\phi}$  then  $\beta = 0$  for all  $0 \leq \tilde{\tau} \leq \frac{1}{\lambda} \ln \frac{1-\varphi_h\phi}{\varphi_h\phi}$  hence the unique equilibrium is  $\{\tau^*, \beta^*\} = \left\{\frac{1}{\lambda} \ln \frac{1-\varphi_h\phi}{\varphi_h\phi}, 0\right\}$ . We now prove uniqueness for the case where  $C'(0) < \frac{1}{2\varphi_h\phi}$  by showing that  $\frac{d\beta}{d\tilde{\tau}} \geq 0$  at any equilibrium point where  $\tau^* > 0$  and  $\beta^* > 0$ .

Implicitly differentiating the equation (24) to find  $\frac{d\beta}{d\tilde{\tau}}$  for  $\beta \in (0, 1)$ :

$$\begin{aligned} \frac{d\beta}{d\tilde{\tau}} &= \frac{\frac{\lambda(1-\varphi_h\phi)e^{\lambda\tilde{\tau}}}{[1-(1-e^{\lambda\tilde{\tau}})S^0]^2} - 2S^0\lambda e^{\lambda\tilde{\tau}} \frac{(1-\varphi_h\phi)e^{\lambda\tilde{\tau}}}{[1-(1-e^{\lambda\tilde{\tau}})S^0]^3}}{C''(\beta) - 2(1-e^{\lambda\tilde{\tau}}) \frac{(1-\varphi_h\phi)^2 e^{\lambda\tilde{\tau}}}{[1-(1-e^{\lambda\tilde{\tau}})S^0]^3}} = \frac{\lambda e^{\lambda\tilde{\tau}}(1-\varphi_h\phi) \frac{1-(1+e^{\lambda\tilde{\tau}})S^0}{[1-(1-e^{\lambda\tilde{\tau}})S^0]^3}}{C''(\beta) - 2(1-e^{\lambda\tilde{\tau}}) \frac{(1-\varphi_h\phi)^2 e^{\lambda\tilde{\tau}}}{[1-(1-e^{\lambda\tilde{\tau}})S^0]^3}} \\ &= \lambda \frac{1-S^0 - e^{\lambda\tilde{\tau}}S^0}{\frac{[1-(1-e^{\lambda\tilde{\tau}})S^0]^3}{e^{\lambda\tilde{\tau}}(1-\varphi_h\phi)} C''(\beta) - 2(1-e^{\lambda\tilde{\tau}})(1-\varphi_h\phi)} = \lambda \frac{(1-S^0) \left(1 - \frac{S^*}{1-S^*}\right)}{\frac{\left[\frac{1-S^0}{1-S_f^*}\right]^3}{(1-\varphi_h\phi) \frac{S_f^*(1-S^0)}{S^0(1-S^*)}} C''(\beta) + 2 \frac{(1-\varphi_h\phi)(S_f^*-S^0)}{S^0(1-S^*)}} \\ &= \frac{\lambda(1-\varphi_h\phi) \left(1 - 2S_f^*\right)}{\frac{S_0}{S_f^*} \left(\frac{1-S_0}{1-S_f^*}\right) C''(\beta) + 2 \left(\frac{S_f^*}{S_0} - 1\right) \frac{(1-\varphi_h\phi)^2}{1-S_0}} \end{aligned}$$

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

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

We note that  $\frac{\partial G(\varphi_h)}{\partial \varphi_h} = \frac{2\phi(1-\phi)}{(1-\varphi_h\phi)^2} > 0$  for all  $\phi \in (0, 1)$  and thus,  $S^* < \frac{1}{2}$  is true when  $k > 2\phi = G(1)$ .

Continuity of the solution in  $\varphi_h$  follows from noting that  $\tau(\varphi_h, \tilde{\beta})$  and  $\beta(\varphi_h, \tilde{\tau})$  are continuous in  $\varphi_h$  for  $\varphi_h > 0$ .  $\square$

### A.12 Proof of Lemma 9

*Proof.* Let  $F(\beta) = \frac{e^{-\lambda\tilde{\tau}}(1-\varphi_h\phi)}{[S^0+(1-S^0)e^{-\lambda\tilde{\tau}}]^2} \cdot \frac{\partial F(\beta)}{\partial \beta} = 2(1-e^{\lambda\tilde{\tau}}) \frac{(1-\varphi_h\phi)^2 e^{-\lambda\tilde{\tau}}}{[S^0+(1-S^0)e^{-\lambda\tilde{\tau}}]^3} > 0$ . From equation (24), the best response advertising level  $\beta^*$  satisfies  $C'(\beta) = \frac{e^{-\lambda\tilde{\tau}}(1-\varphi_h\phi)}{[S^0+(1-S^0)e^{-\lambda\tilde{\tau}}]^2}$ . 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  $\tilde{\tau} > \hat{\tau}_1$ . Hence from Lemma 6 and the proof of Lemma 8 if  $\tau(\varphi_h, 0) = \frac{1}{\lambda} \ln \frac{1-\varphi_h\phi}{\frac{k}{1-\phi}-\varphi_h\phi} > \hat{\tau}_1 \Rightarrow \beta^*(\varphi_h) > \beta_2^*(\varphi_h)$  and  $\tau_1^*(\varphi_h) < \tau_2^*(\varphi_h)$ , otherwise  $\beta^*(\varphi_h) = \beta_2^*(\varphi_h) = 0$  and  $\tau_1^*(\varphi_h) = \tau_2^*(\varphi_h) = \frac{1}{\lambda} \ln \frac{1-\varphi_h\phi}{\frac{k}{1-\phi}-\varphi_h\phi}$ . Finally  $\tau(\varphi_h, 0) = \hat{\tau}_1$  when  $C'_1(0) = \frac{1-\phi}{k} \left[ 1 - \frac{\varphi_h\phi(1-\phi)}{k} \right]$  and hence the inequalities hold strictly for  $C'_1(0) < \frac{1-\phi}{k} \left[ 1 - \frac{\varphi_h\phi(1-\phi)}{k} \right]$ .  $\square$

### A.13 Proof of Lemma 10

*Proof.*  $V(\varphi_h, \tau^*(\varphi_h), \beta^*(\varphi_h)) = (1-\beta^*) \frac{k}{1-S^0} \left[ \frac{S^*}{S^0} - 1 - \ln \frac{S^*}{S^0} \right]$  is continuous in  $S^*$  and  $S^0$  for  $0 < S^0 < 1$  and  $S^* \geq S^0$ . From earlier, we have already seen that  $\tau^*(\varphi_h)$  and  $\beta(\varphi_h)$  are continuous in  $\varphi_h$ . Hence both  $S^*$  and  $S^0$  are continuous in  $t$ ,  $\beta$  and  $\varphi_h$  and thus  $V(\varphi_h, \tau^*(\varphi_h), \beta^*(\varphi_h))$  is continuous in  $\varphi_h$  for  $\varphi_h \in (0, 1]$ .  $\square$

We define  $\tau^*(0) = \lim_{\varphi_h \rightarrow 0} \tau^*(\varphi_h)$ ,  $\beta^*(0) = \lim_{\varphi_h \rightarrow 0} \beta^*(\varphi_h)$ , and  $V(0, \tau^*(0), \beta^*(0)) = \lim_{\varphi_h \rightarrow 0} V(\varphi_h, \tau^*(\varphi_h), \beta^*(\varphi_h))$  to ensure that the value of information acquisition is continuous over the closed interval  $\varphi_h \in [0, 1]$ .

### A.14 Proof of Proposition 4

*Proof.*  $V(\varphi_h, \tau^*(\varphi_h), \beta^*(\varphi_h))$  is continuous in  $\varphi_h$  and bounded below by 0. Hence there is always a value  $\varphi_h$  which satisfies equation 25. We now prove that there is always a solution  $\varphi_h > 0$ . Consider first what happens to the best response function of consumers we consider the extreme points of the best response function for  $\varphi_h \rightarrow 0$ : the best response to zero advertising  $\lim_{\varphi_h \rightarrow 0} \tau(\varphi_h, 0) = \frac{1-\phi}{k}$  and the conjectured level of advertising  $\tilde{\beta}^u(\varphi_h)$  where  $\tau(\varphi_h, \tilde{\beta}^u) = 0$ ,  $\lim_{\varphi_h \rightarrow 0} \tilde{\beta}^u(\varphi_h) = \frac{S^*-S^0}{1-S^0} = \frac{\varphi_h\phi}{k(1-\varphi_h\phi)} \left[ 1 - k - \frac{(1-\varphi_h)\phi}{1-\varphi_h\phi} \right] = 0$  but is strictly positive for any  $\varphi_h > 0$ . Now consider the cut off time  $\hat{t}$  where the firms best response  $\beta = 0$  for all  $0 \leq t \leq \hat{t}$  and  $\beta > 0$  for all  $\hat{t}(\varphi_h) < t \leq \frac{1}{\lambda} \ln \frac{1-\varphi_h\phi}{\frac{k}{1-\phi}-\varphi_h\phi}$ .

The cut off  $\hat{t}$  : satisfies  $C'(0) = \frac{e^{-\lambda\hat{t}}(1-\varphi_h\phi)}{[\varphi_h\phi+(1-\varphi_h\phi)e^{-\lambda\hat{t}}]^2}$  noting that  $\lim_{\varphi_h \rightarrow 0} \frac{e^{-\lambda\hat{t}}(1-\varphi_h\phi)}{[\varphi_h\phi+(1-\varphi_h\phi)e^{-\lambda\hat{t}}]^2} = e^{\lambda\hat{t}}$  thus  $\lim_{\varphi_h \rightarrow 0} \hat{t}(\varphi_h) = \frac{1}{\lambda} \ln C'(0) > 0$ . We recall that  $\frac{d\tau}{d\beta} < 0$  and that  $\{\tau^*(\varphi_h), \beta^*(\varphi_h)\}$  is unique and hence conclude that  $\lim_{\varphi_h \rightarrow 0} \tau^*(\varphi_h) \geq \frac{1}{\lambda} \ln C'(0)$  and  $\lim_{\varphi_h \rightarrow 0} \beta^*(\varphi_h) = 0$ . We also know  $\lim_{\varphi_h \rightarrow 0} \tau^*(\varphi_h) = \lim_{\varphi_h \rightarrow 0} \frac{1}{\lambda} \ln \frac{S^*(1-S^0)}{S^0(1-S^*)} = \lim_{\varphi_h \rightarrow 0} \frac{1}{\lambda} \ln \frac{S^*}{S^0}$  hence  $\lim_{\varphi_h \rightarrow 0} \frac{S^*}{S^0} \geq C'(0) > 1$ . Using this it is straightforward to note the limit for the value of information acquisition  $\lim_{\varphi_h \rightarrow 0} V(\varphi_h, \tau^*(\varphi_h), \beta^*(\varphi_h)) = k \left[ \lim_{\varphi_h \rightarrow 0} \frac{S^*}{S^0} - 1 - \ln \left( \lim_{\varphi_h \rightarrow 0} \frac{S^*}{S^0} \right) \right] > 0$ . Thus there is a strictly positive amount of information acquired in the best signaling subgame perfect equilibrium.  $\square$

### A.15 Proof of Proposition 5

*Proof.* We begin with the following Lemma:  $\square$

**Lemma 11.** *Suppose  $k > 2\phi$ . Consider two cost functions of advertising  $C_1$  and  $C_2$  where  $1 \leq C'_1 < C'_2$  and  $C'_1(0) < \frac{S^*(\varphi_h^{**}, 0)(1-S^*(\varphi_h^{**}, 0))}{S_0(\varphi_h^{**}, 0)(1-S_0(\varphi_h^{**}, 0))} (1 - \varphi_h^{**}\phi)$ . Then,*

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

*Proof.* First, we note that from Lemma 9 that  $\beta_1^*(\varphi_h) > \beta_2^*(\varphi_h)$  and  $\tau_1^*(\varphi_h) < \tau_2^*(\varphi_h)$ , and from ...  $S_1^* = S_2^*$ . Now we can write  $V(\varphi_h, \tau^*(\varphi_h), \beta^*(\varphi_h))$  as:

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

Taking the derivative with respect to  $\beta^*$  holding  $S_0$  and  $S^*$  constant:

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

and from Proposition

$$\frac{\partial}{\partial S_0} \left( \frac{k}{1 - S_0} \left[ \frac{S^*}{S_0} - 1 - \ln \frac{S^*}{S_0} \right] \right) < 0,$$

hence

$$\frac{\partial V}{\partial S_0} < 0,$$

and the lemma follows immediately by noting that  $\frac{dS_0}{d\beta^*} > 0$  and  $\frac{dS_0}{d\tau^*} = 0$ .

Now  $V(\varphi_h, \tau^*(\varphi_h, \beta^*), \beta^*(\varphi_h, \tau^*))$  is continuous in  $\varphi_h$  also  $\bar{c} \geq \frac{1-k+k \ln k}{1-\phi} = V(1, \tau^*(1, 0), 0)$ . Hence it follows that  $1 > \varphi_{h2}^* > \varphi_{h1}^*$ . Finally  $S_1^* < S_2^*$  follows from recalling that  $\frac{dS^*}{d\varphi_h} > 0$  from equation 14. Finally if  $C'_1(0) \geq \frac{S^*(\varphi_h^{**}, 0)(1-S^*(\varphi_h^{**}, 0))}{S_0(\varphi_h^{**}, 0)(1-S_0(\varphi_h^{**}, 0))} (1 - \varphi_h^{**}\phi)$  then the equilibrium for both cost functions is the same,  $\{\varphi_h^{**}, \tau^*(\varphi_h^{**}), 0\}$ .  $\square$

## A.16 Proof of Proposition 6

*Proof.* If  $C'(0) \geq \frac{S_f^*(\varphi_h^{**}, \tau^*(\varphi_h^{**}), 0)(1 - S_f^*(\varphi_h^{**}, \tau^*(\varphi_h^{**}), 0))}{S_0^*(\varphi_h^{**}, 0)(1 - S_0^*(\varphi_h^{**}, 0))} (1 - \varphi_h^{**} \phi)$ , where  $\tau^*(\varphi_h^{**}) = \frac{1}{\lambda} \ln \frac{S^*(\varphi_h^{**})}{1 - S^*(\varphi_h^{**})} \frac{1 - \varphi_h^{**} \phi}{\varphi_h^{**} \phi}$ , then the subgame perfect equilibrium is  $\{\varphi_h^{**}, \tau^*(\varphi_h^{**}), 0\}$ .

Note that by choosing  $T > \tau^*(\varphi_h^{**})$  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  $T > \frac{1}{r} \ln \frac{S^*(\varphi_h^{**}, 0)(1 - S^*(\varphi_h^{**}, 0))}{S_0^*(\varphi_h^{**}, 0)(1 - S_0^*(\varphi_h^{**}, 0))} \frac{((1 - \varphi_h^{**})\phi + 1 - \phi)}{C'(0)}$ , we have  $e^{rT} C'(0) > \frac{S^*(\varphi_h^{**}, 0)(1 - S^*(\varphi_h^{**}, 0))}{S_0^*(\varphi_h^{**}, 0)(1 - S_0^*(\varphi_h^{**}, 0))} (1 - \varphi_h^{**} \phi)$  and from the proposition we have that  $\beta = 0$  and  $\varphi_h = \varphi_h^{**}$ .  $\square$

## B Advertising section under trembling hand refinement

The analysis in Section 5 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 by the firm choosing an advertising level and consumers choosing a time to engage in word of mouth until, are perfectly consistent with this analysis. On the other hand in an equilibrium where  $\beta^* = 0$  then consumers expect there is a zero probability of receiving an advertisement and so Bayesian updating does not provide any guidance for pinning down consumers off-equilibrium beliefs in these equilibria. In this appendix we analyze

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  $\tau$  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.

In this section we analyze a trembling hand equilibrium. The tremble we analyze is on the firm's advertising strategy. We assume that when the firm chooses an advertising level  $\beta$  then with probability  $\epsilon$  individuals in the fraction of the population advertised do not receive the advertisement and with the same probability  $\epsilon$  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  $\epsilon$  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  $\epsilon$ . Importantly the conjectured level of "effective" advertising after accounting for the tremble is the same at both information sets  $a = 0, 1$  and the best response is independent of  $a$ .

The best response function of consumers is now given by:

$$\tau^*(a, \varphi_h, \tilde{\beta} + \epsilon - 2\epsilon\tilde{\beta}) = \frac{1}{\lambda} \ln \frac{S_c^*(\varphi_h, 0, \tilde{\beta} + \epsilon - 2\epsilon\tilde{\beta})}{1 - S_c^*(\varphi_h, 0, \tilde{\beta} + \epsilon - 2\epsilon\tilde{\beta})} \frac{1 - S_0(\varphi_h, 0, \tilde{\beta} + \epsilon - 2\epsilon\tilde{\beta})}{S_0(\varphi_h, 0, \tilde{\beta} + \epsilon - 2\epsilon\tilde{\beta})} \quad (27)$$



The firm's choice of advertising can not directly influence consumers strategy and is very similar to earlier but where we replace  $\beta$  with the "effective" advertising on the rhs.

$$C'(\beta(\varphi_h, \tilde{\tau})) \tag{28}$$

$$= \frac{e^{-\lambda\tilde{\tau}}(1 - \varphi_h\phi)}{[S_0(\varphi_h, 0, \beta(\varphi_h, \tilde{\tau}) + \epsilon - 2\epsilon\beta(\varphi_h, \tilde{\tau})) + (1 - S_0(\varphi_h, 0, \beta(\varphi_h, \tilde{\tau}) + \epsilon - 2\epsilon\beta(\varphi_h, \tilde{\tau})))e^{-\lambda\tilde{\tau}}]^2} \text{ if } \beta^* \in (0, 1)$$

and

$$C'(0) \geq \frac{e^{-\lambda\tilde{\tau}}(1 - \varphi_h\phi)}{[S^0(\varphi_h, 0, \epsilon) + (1 - S^0(\varphi_h, 0, \epsilon))e^{-\lambda\tilde{\tau}}]^2} \text{ if } \beta^* = 0 \tag{29}$$

The information acquisition choice of consumers satisfies:

$$\begin{aligned} \varphi_h^* \bar{c} &= V(\varphi_h^*, \tau^*(\varphi_h^*), \beta^*(\varphi_h^*) + \epsilon - 2\epsilon\beta^*(\varphi_h^*)) \text{ if } V(1, \tau^*(1), \beta^*(1)) < \bar{c} \\ \varphi_h^* &= 1 \quad \text{if } V(1, \tau^*(1), \beta^*(1) + \epsilon - 2\epsilon\beta^*(1)) \geq \bar{c}. \end{aligned} \tag{30}$$

We are interested in the limit of the equilibria as  $\epsilon \rightarrow 0$ . We note Equations 27, 28, and 29 are continuous in  $\epsilon$  at  $\epsilon = 0$ . Thus the best response functions go to the best response functions in Section 5 and the limit of the equilibria is the same. Similarly equation 30 is continuous in all its arguments and thus also attains the same limit as in Section 5 and so the limiting subgame perfect equilibrium is the same as the one found in Section 5.