Targeted Information Revelation through Word of Mouth Marketing and Advertising

Arthur Campbell, Dina Mayzlin, and Jiwoong Shin

Yale University

June 14, 2011

Abstract

In this paper we explore how the firm can use word of mouth marketing and advertising to optimally target information to different groups of consumers in order to maximize the diffusion of information about its product. We show that the firm may benefit from the commitment to not reveal information to the low-type group throughout the information diffusion process. On the other hand, if such a commitment is not possible, the firm will choose to reveal some information to the low-type consumers. Finally, we explore the question of when a firm may choose to invest resources by communicating product information to the consumers versus having the consumers engage in costly search on their own.

1 Basic model

1.1 No advertising case

Suppose that a monopolist is selling a product to a mass 1 of agents. An agent i may be one of two types high and low $\theta = h, l$. The fraction of high types in the population ϕ. The firm's objective is to maximize the fraction of the population which receives information m^* about its product. We shall denote this as S^*. Assume no advertising for the moment. agents 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.

Timing

At time $t = -1$ each type chooses whether to obtain information m^* about a firm's product, this information is hard. The cost c_h, c_l are distributed uniformly on $[0, \bar{c}]$. However the firm can increase the costs for each type above this level if it wishes. From time $t = 1$ onwards individuals mix at rate λ. During each meeting an individual may pass on the hard information m^* at a cost k or pass on no information $\varnothing$ at zero cost. Assume 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 social utility derived from the meeting is a function of the beliefs the other agent has about the agent's type. In particular agent i receives utility

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

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

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

if the agent does not pass information where $b_j\left(\theta_i = h|\varnothing, t\right)$ is the belief if no signal (denoted by $\varnothing$) is sent. Given our notions of high and low types we assume $U' > 0$. Also note the signaling benefit at a time t is

$$\Delta U\left(t\right) = U_i\left(b_j\left(\theta_i = h|m^*, t\right)\right) - U_i\left(b_j\left(\theta_i = h|\varnothing, t\right)\right) \tag{3}$$

which is the difference between sending a signal and not sending a signal at that time$t$.

Growth of the informed population

Denote the fraction of types who become informed at $t = -1$ by φ_h, φ_l where these are going to be endogenously determined in equilibrium. The fraction of the population which are informed $S\left(t\right)$ evolves over time as agents mix at rate λ and pass on information. The initial condition is $S_0 = \varphi_h\phi + \varphi_l\left(1-\phi\right)$ and the rate of change of the informed population is given by:

$$\frac{dS}{dt} = \lambda S\left(t\right)\left(1 - S\left(t\right)\right) \tag{4}$$

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

$$S\left(t\right) = \frac{1}{1 + \left(\frac{1}{S_0} - 1\right)e^{-\lambda t}} \tag{5}$$

which continues to grow until the beneficial impact of passing the firm's message is less than the cost of doing so. Hence $S\left(t\right)$ stops growing when $\Delta U\left(t^*\right) = k$ which defines the extent of the diffusion $S^* = S\left(t^*\right)$. The firm's objective is to maximize the extent of this diffusion.

Beliefs

At $t = 0$ beliefs are

$$b_j\left(\theta_i = h|m^*, 0\right) = \frac{\varphi_h\phi}{\varphi_h\phi + \varphi_l\left(1-\phi\right)} \tag{6}$$

and

$$b_j\left(\theta_i = h|\varnothing, 0\right) = \frac{\left(1-\varphi_h\right)\phi}{\left(1-\varphi_h\right)\phi + \left(1-\varphi_l\right)\left(1-\phi\right)}. \tag{7}$$

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

$$b_j\left(\theta_i = h|m^*, t\right) = \frac{S\left(t\right) - S_0}{S\left(t\right)}\left[b_j\left(\theta_i = h|\varnothing, 0\right)\right] + \frac{S_0}{S\left(t\right)}\left[b_j\left(\theta_i = h|m^*, 0\right)\right] \tag{8}$$

$$= b_j\left(\theta_i = h|\varnothing, 0\right) + \frac{S_0}{S\left(t\right)}\left[b_j\left(\theta_i = h|m^*, 0\right) - b_j\left(\theta_i = h|\varnothing, 0\right)\right]. \tag{9}$$

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

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

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:

$$U(b_j(\theta_i = h|m^*, t)) - U(b_j(\theta_i = h|\varnothing, 0)) = k \tag{10}$$

for the moment assume U is linear

$$\begin{aligned} & U(b_j(\theta_i = h|m^*, t)) - U(b_j(\theta_i = h|\varnothing, 0)) \\ &= \overline{u}\left[b_j(\theta_i = h|m^*, t) - b_j(\theta_i = h|\varnothing, t)\right] \\ &= \overline{u}\frac{S_0}{S(t)}\left[b_j(\theta_i = h|m^*, 0) - b_j(\theta_i = h|\varnothing, 0)\right] = k \end{aligned} \tag{11}$$

where $\overline{u}$ is just a constant which we can normalize to 1. Hence the diffusion stops at

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

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

Ignoring the ex ante incentive constraints we could take first order conditions wrt φ_h, φ_l.

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

$$\begin{aligned} \frac{dS^*}{d\varphi_l} &= -\frac{\phi}{k}(1-\varphi_h)\left[\frac{(1-\phi)\left[(1-\phi_h)\phi + (1-\varphi_l)(1-\phi) + \varphi_h\phi + \varphi_l(1-\phi)\right]}{((1-\varphi_h)\phi + (1-\varphi_l)(1-\phi))^2}\right] \\ &= -\frac{\phi}{k}\frac{(1-\varphi_h)(1-\phi)}{((1-\varphi_h)\phi + (1-\varphi_l)(1-\phi))^2} \leq 0 \text{ if } \varphi_h < 1 \text{ then } < 0 \end{aligned} \tag{15}$$

It is pretty quick to see this is maximized provided that $\varphi_h = 1$, independent of φ_l. However if the ex ante information acquisition constraint binds for some high types then $\varphi_h < 1$ and we need to worry about the first order condition on φ_l. When $\varphi_h < 1$ we have that $\frac{dS^*}{d\varphi_l} < 0$ hence the optimal policy results in $\varphi_l = 0$ provided that φ_l does not help alleviate the ex ante incentives to acquire information which we still need to check.

Ex Ante Incentives

Of course the previous section ignores agents decision to acquire information at $t = -1$. This may mean that φ_h is bounded strictly below 1. If this is the case we may wish to check whether $\varphi_l = 0$ maximizes S^*. Hence it will be at least weakly optimal to set $\varphi_l = 0$. Hence, we can restrict attention to strategies which specify S_0 where the set of informed individuals is 100% high types or if $S_0 > \phi$ then all high types are informed and the remainder are low types. We can denote the firm's strategy as just a cut off c_h such that all high types with $c \leq c_h$ collect information at the ex ante stage and all $c > c_h$ do not. We will first analyse the case where $c_h \leq 1$ where 0% of low types acquire information.

Going back to the information accumulation stage, the decision to acquire the information at this stage depends on the total signaling benefit the agent will acquire during the diffusion process. If this benefit is above their cost c then the agent acquire information. Denote the time at which the diffusion process ends by t^*. The signaling benefit for an agent is then

$$V = \lambda \left(\int_0^{t^*} \left(\frac{1 - S(t)}{1 - S_0} \right) \left(\frac{S_0}{S(t)} \left[b_j(\theta_i = h | m^*, 0) - b_j(\theta_i = h | \varnothing, 0) \right] - k \right) dt \right) \tag{16}$$

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

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

making this substitution for dt and substituting $B = b_j(\theta_i = h|m^*, 0) - b_j(\theta_i = h|\varnothing, 0)$ we get:

$$\begin{aligned} V &= \int_{S_0}^{S^*} \left(\frac{1}{1 - S_0} \right) \frac{1}{S(t)} \left(\frac{S_0}{S(t)} \left[b_j(\theta_i = h|m^*, 0) - b_j(\theta_i = h|\varnothing, 0) \right] - k \right) dS \qquad (17) \\ &= \frac{1}{1 - S_0} \int_{S_0}^{S^*} \left[\frac{B \cdot S_0}{S^2(t)} - \frac{k}{S(t)} \right] dS \quad \left(\text{where } \frac{B \cdot S_0}{S^*} = k\right) \\ &= \left(\frac{k}{1 - S_0} \right) \left(S^* \left[-\frac{1}{S} \right]_{S_0}^{S^*} - [\ln S]_{S_0}^{S^*} \right) \\ &= \left(\frac{k}{1 - S_0} \right) \left(\frac{S^* - S_0}{S_0} + \ln \frac{S_0}{S^*} \right) \end{aligned}$$

Writing out the firm's optimization problem:

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

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

$$0 < \varphi_h \leq 1$$
$$\varphi_l = 0$$

Proof. First, we already showed that $\frac{dS^*}{d\varphi_h} \geq 0$, $\frac{dS^*}{d\varphi_l} \leq 0$ (Equations 14). Now consider V as a function of S_0 and S^* we will show that $\frac{dV}{d\varphi_l} < 0$.

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

now taking the derivative $\frac{dV}{d\varphi_x}$ where $\frac{\partial S_0}{\partial \varphi_x} > 0$ for $x = l, h$ and from earlier $\frac{\partial S^*}{\partial \varphi_l} \leq 0$, $\frac{\partial S^*}{\partial \varphi_h} \geq 0$

$$\begin{aligned}
\frac{dV}{d\varphi_l} &= \frac{dV}{dS_0}\frac{dS_0}{d\varphi_l} + \frac{\partial V}{\partial S^*}\frac{\partial S^*}{\partial \varphi_l} \\
&= \left[\begin{array}{c} k\left(\frac{1}{1-S_0}\right)^2\left(S^*\frac{1}{S_0} - 1 + \ln S_0 - \ln S^*\right) \\ +\left(\frac{k}{1-S_0}\right)\left(-\frac{1}{S_0^2}\frac{BS_0}{k} + \frac{1}{S_0}\right) \end{array}\right]\frac{dS_0}{d\varphi_l} \\
&\quad + \left(\frac{k}{1-S_0}\right)\left[\frac{1}{S_0} - \frac{1}{S^*}\right]\frac{\partial S^*}{\partial \varphi_l} \\
&= \underbrace{\left[\frac{k}{(1-S_0)^2 S_0^2}\left\{(S^* - S_0)(2S_0 - 1) + S_0^2\left(\ln\frac{S_0}{S^*}\right)\right\}\right]}_{?}\underbrace{\frac{dS_0}{d\varphi_l}}_{+} \\
&\quad + \underbrace{\left(\frac{k}{1-S_0}\right)\left[\frac{1}{S_0} - \frac{1}{S^*}\right]}_{+}\underbrace{\frac{\partial S^*}{\partial \varphi_l}}_{-}
\end{aligned}$$

We have immediately that $\left(\frac{k}{1-S_0}\right)\left[\frac{1}{S_0} - \frac{1}{S^*}\right] > 0$. Hence we will show that $\frac{dV}{dS_0} < 0$ to prove that it is $\frac{dV}{d\varphi_l} < 0$.

Note that $\frac{dV}{dS_0}$ is negative provided that

$$2S_0 - 1 < 0$$

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

If $S_0 \geq \frac{1}{2}$, we need that

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

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

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

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

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

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

$$
\begin{aligned}
-\frac{\ln \frac{1}{S_0}}{\frac{1}{S_0} - 1} &< \frac{1}{S_0} - 2 \\
-\ln y - (y-2)(y-1) &< 0.
\end{aligned}
$$

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

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

and note that

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

Hence, $-\ln y - (y-2)(y-1) < 0$, which shows that $\frac{dV}{dS_0} < 0$. This completes the proof that $\frac{\partial V}{\partial \varphi_l} < 0$.

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

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

Hence, $V(\varphi_h^*, 0) \ge 0$. This proves that $\varphi_l^* = 0$. $\square$

The result here is that the strategy which maximizes the diffusion of the firm's message is to restrict information to the socially low type agents by choosing $\varphi_l = 0$ and minimizing the costs for the socially high type agents such that $0 < \varphi_h \leq 1$. Note that this result will also continue to hold with a mass of agents who have 0 search costs provided there are not so many of them that there are more than S^*.

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, $\varphi_l > 0$.*

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

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

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 14). 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 exist $r^*(\varphi_l)$ such that for all $r > r^*(\varphi_l)$, $R(\varphi_l = 0, \varphi_h) < R(\varphi_l > 0, \varphi_h)$. (this is a little loose)... $\square$

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 agents by choosing $\varphi_l > 0$.

1.2 Advertising

Consider an advertising technology which is informing people randomly in the population at rate γ. (This is essentially a model of mass advertising). When this is the case we can write the relation for $S(t)$ as follows:

$$\frac{dS}{dt} = (\lambda S(t) + \gamma)(1 - S(t))$$
$$dt = \frac{dS}{(\lambda S(t) + \gamma)(1 - S(t))}$$

We can then go through a similar procedure as earlier to calculate the utility from signaling. Doing that in the case that the firm advertises all the time one can get:

$$\begin{aligned} V &= \lambda \left(\int_0^{t^*} \left(\frac{1 - S(t)}{1 - S_0} \right) \left(\frac{S_0}{S(t)} \left[b_j(\theta_i = h | m^*, 0) - b_j(\theta_i = h | \varnothing, 0) \right] - k \right) dt \right) \\ &= \lambda \left(\int_{S_0}^{S^*} \left(\frac{1}{1 - S_0} \right) \frac{1}{(\lambda S(t) + \gamma)} \left(\frac{S_0}{S(t)} B - k \right) dS \right) \end{aligned} \tag{18}$$

one can calculate the closed form for V as follows

$$\begin{aligned} V &= \frac{\lambda}{1 - S_0} \left(S_0 B \int_{S_0}^{S^*} \frac{1}{S(t)(\lambda S(t) + \gamma)} dS - k \left[\frac{1}{\lambda} \ln(\lambda S(t) + \gamma) \right]_{S_0}^{S^*} \right) \\ &= \frac{\lambda}{1 - S_0} \left(S_0 B \left[\frac{1}{\gamma} \ln \left(\frac{S(t)}{\lambda S(t) + \gamma} \right) \right]_{S_0}^{S^*} - k \left[\frac{1}{\lambda} \ln(\lambda S(t) + \gamma) \right]_{S_0}^{S^*} \right) \\ &= \frac{\lambda}{1 - S_0} \left(k S^* \left[\frac{1}{\gamma} \ln(S(t)) - \ln(\lambda S(t) + \gamma) \right]_{S_0}^{S^*} - k \left[\frac{1}{\lambda} \ln(\lambda S(t) + \gamma) \right]_{S_0}^{S^*} \right) \\ &= \frac{k}{1 - S_0} \left(S^* \frac{\lambda}{\gamma} \ln \left(\frac{S^*}{S_0} \right) - (1 + \lambda S^*) \ln \left(\frac{\lambda S^* + \gamma}{\lambda S_0 + \gamma} \right) \right) \end{aligned} \tag{19}$$

However it is relatively straightforward from the earlier expression to show that $\frac{dV}{d\gamma} < 0$.

Proposition 3. $\frac{dV}{d\gamma} < 0$

Proof. Let

$$V = \lambda \left(\int_{S_0}^{S^*} \left(\frac{1}{1 - S_0} \right) \frac{1}{(\lambda S(t) + \gamma)} \left(\frac{S_0}{S(t)} B - k \right) dS \right)$$

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

$$\frac{dV}{d\gamma} = -\lambda \left(\int_{S_0}^{S^*} \left(\frac{1}{1 - S_0} \right) \frac{1}{(\lambda S(t) + \gamma)^2} \left(\frac{S_0}{S(t)} B - k \right) dS \right) < 0$$

every term to the right of the integral is positive hence the derivative is negative. As was shown earlier one wants to set $\varphi_l = 0$ and φ_h has high as possible. Advertising therefore reduces V and hence φ_h. □

Hence advertising acts to crowd out the incentives for agents to search for information ex ante.