Tell Your Friends!
Word of Mouth and Percolation in Social Networks

Arthur Campbell*

November 9, 2008

JOB MARKET PAPER

Abstract

This paper studies the optimal strategies of a monopolist selling a good to consumers who engage in word of mouth communication. In the model consumers may spread news about the monopolist's good to uninformed consumers through a social network. The monopolist may use the price it charges to influence both the proportion of the population that is willing to purchase the good and the pattern of communication that takes place within the social network. I find a number of results: (i) demand is more elastic in the presence of word of mouth and this induces a downward bias in estimates of consumers' valuation for the good which ignore word of mouth; (ii) the monopolist reduces the price to induce additional word of mouth for regular goods however for goods whose valuation is greater for well connected individuals the price may, in fact, be greater; (iii) the optimal pattern of diffusion involves introductory prices which fluctuate up and down; and (iv) exclusive (high priced) products will optimally target advertising towards individuals with many friends whereas common (low priced) products will target individuals with fewer friends.

*Massachusetts Institute of Technology, Department of Economics, E52, 50 Memorial Drive, Cambridge, MA, 02142. email: ards@mit.edu. I am particularly indebted to Glenn Ellison. I would also like to thank Muhamet Yildiz, Florian Ederer, Johannes Spinnewijn, Moshe Cohen, Jesse Edgerton, Daniel Gottlieb, Nicolas Arregui, Alex Wolinsky, Alp Simsek and participants at the MIT Theory and Industrial Organization lunches.

1 Introduction

A widely recognized phenomenon is the diffusion of products and innovation within populations. A key conduit for this diffusion is often word of mouth (WOM hereafter) between members of the population. A large number of studies have found that WOM is an important source of information for consumers' purchase decisions.[1] The significant influence of WOM on purchasing decisions raises a number of questions pertaining to an environment where consumers share their experience of a firm's good or service with each other. How does WOM affect demand for a product? What strategies does a firm employ in the presence of WOM? Do these strategies differ across different product categories? How are traditional advertising strategies affected by WOM? This paper characterizes product demand, and the firm's optimal pricing and advertising behavior by considering how a firm may strategically affect the probability consumers engage in WOM through the price and the subsequent pattern of communication which takes place. It combines a model of a monopolist and a percolation model of WOM in a social network. The percolation process describes the pattern of WOM that takes place in the social network as a function of a firm's pricing and advertising strategies, and consumer's valuations.

The paper studies a monopolist selling a good to an initially uninformed population with heterogeneous valuations for the good. Consumers are connected within the population by a social network which is modelled as a random graph with an arbitrary degree distribution. Consumers may communicate with their friends in the social network. The content of communication is to inform the receiver that the good exists. In order to purchase the good, consumers must first find out about the good and second be prepared to purchase it at the price charged by the monopolist. The analysis assumes that an infinitessimal fraction of the population become informed and the remainder of the population may only find out about the good via WOM diffusing through the social network. Later, I also consider the case where consumers may become informed from costly advertising undertaken by the monopolist.

WOM is modeled as a percolation process on the social network. Representing the social network by a random graph with arbitrary degree distributions makes the analysis of the percolation process particularly tractable and maintains a great

[1]See Bass (1969), Sheth (1979), Arndt (1967), Day (1971) and Richins (1983), Mobius et al (2006), Godes and Mayzlin (2003,2004), and Reichheld (2003).

deal of freedom in the distribution of friendships across the population. The percolation process assumes individuals are prepared to engage in WOM with a certain probability, which is a function of the individual's valuation for the good and the price charged by the monopolist. This probability is modelled by a step function, whereby the consumer only engages in WOM if she is prepared to purchase the product. When the price is zero everyone is willing to engage in WOM and the potential pathways for communication correspond to the social network, however as the price increases fewer people are prepared to engage in WOM and there are fewer pathways for communication. The connectedness of the network over which communication may take place becomes more and more disconnected at higher prices, mitigating the effect of WOM. The analysis proceeds in two steps: firstly the formulation of WOM as a percolation process allows one to map the primitive of the model, the social network, and firm's strategy, price, to the network describing the potential communication pathways between individuals; second the demand and profit of the monopolist may then be derived as a function of this communication network, price and advertising strategies.

The model is able to provide insights into how the social network affects the nature and shape of demand for a good, what pricing strategies a monopolist may use in a static and dynamic setting, the individuals to target advertising towards and how an owner of the rights to advertise to individuals within a social network can benefit from utilizing information about a consumer's relative connectivity. Demand has two regions, one at high prices, where very few people hear about the good, and another at lower prices, where there is a significant fraction of the population communicating about the good through WOM. At these lower prices demand is more elastic than demand when the population is fully informed. Estimates of consumers' valuation for the good are biased downwards and estimated counterfactual responses to price increases are overstated when WOM is ignored. Regular goods are priced below what a monopolist would charge absent WOM; however for goods whose valuation is greater for people with many friends, the price can in fact be greater. Introductory prices may include periods of sales as prices fluctuate up and down to facilitate a more effective spread of the good through the population. Increasing advertising costs may benefit consumers. Exclusive (high priced) products will optimally target advertising towards individuals with many friends, whereas common (low priced) products will target individuals with fewer friends.

1.1 Related literature

This paper is related to a recent economics literature which considers the optimal strategy for an outside party trying to maximize an objective which is a function of agents' actions in a social network (see for instance Goyal and Galleotti (2007), Ballester et al. (2006), Banerji and Dutta (2006) and Galleotti and Mattozzi (2008)). The most related of these is Goyal and Galleotti (2007) which considers the optimal advertising decisions of a monopolist in the presence of local information sharing and local adoption externalities. In contrast to these papers the present paper is the first to use a model of percolation to capture the pattern of communication and endogenize the probability that individuals engage in WOM as a function of their valuation for the good and monopolist's strategies. The model addresses new questions concerning the optimal strategies the monopolist employs when it can affect the diffusion rate of information, and gains fresh insights into the shape and nature of demand and the effects of diffusion of information via WOM on the pricing and advertising behavior of the monopolist.

There are also a number of papers which consider diffusion of an action or adoption decision of agents interacting in social networks. In these papers an agent's payoff is a function of the actions of agents connected to them in the social network. Some of these papers, like this paper, find that there is some critical threshold which determines whether an action or behavior will successfully propagate through a population (for instance Ellison (1993), Morris (2000), Jackson and Yariv (2007), Lopez-Pintado (2007)). In these papers the probability an agent is prepared to propagate/pass on the action/information is a function of the decisions of other agents, in this paper the focus is different, it is on the strategic decision making of an outside party, the monopolist, when it can influence this probability (also known as the percolation probability) and hence the rate of diffusion.

Within the broad literature that considers percolation processes, some other papers, as this paper does, consider the spread of phenomenon on social networks which are modelled by random graphs with arbitrary degree distributions, for epidemic diseases (Newman (2002), Sander et al. (2002)) and fads/innovations (Watts (2002)). In contrast to these papers, the innovation of this paper is to endogenize the percolation probability itself by making it a function of the strategy (price) chosen by the monopolist. In doing so I am able to relate the strategy of the monopolist to the characteristics of the network and diffusion process.

2 Model

There is a monopolist selling a good to a population of consumers $N = \{1, .., n\}$ who have heterogeneous preferences for this good and are initially unaware that it exists. A fraction $\varepsilon \approx 0$ of these people will find out about the good exogenously, everyone else must find out about it either through WOM from one of their friends or from informative advertising undertaken by the monopolist. Consumers have a uniform valuation for the good $\theta_i \sim U[0, 1]$ and they derive utility $\theta_i - P$ if they purchase the good and 0 otherwise. The individuals who desire the product will be those for whom $\theta_i \geq P$. Hence the demand for the good if the population is fully informed is $1 - P$.

The population is connected by a social network described by a graph (N, Ξ) with n nodes and a set of edges $\Xi \subseteq \{(i, j) | i \neq j \in N\}$ where an element $(i, j) \in \Xi$ indicates there is a friendship between individuals i and j. The social network considered here is an undirected network so if $(i, j) \in \Xi$ then $(j, i) \in \Xi$. Each person may engage in WOM with their friends. I assume that the probability an individual i passes on information about the good to her friends is a function $\nu(\theta, P)$ of the individual's valuation for the good and the price charged by the monopolist. Specifically:

$$\begin{aligned} \nu(\theta, P) &= 1 \text{ if } \theta_i - P \geq 0 \\ &= 0 \text{ if } \theta_i - P < 0 \end{aligned} \tag{1}$$

The key characteristic of this function is that it is increasing in the individual's willingness to purchase the good $\theta - P$. It is the relationship between the probability and the price, that allows the monopolist to affect the rate, and distance which WOM about the good spreads within the social network.

All consumers are initially unaware of the good, so the fraction of the population that eventually buy it is in part determined by how many people find out about it. The timing of the model is as follows:

1. Each person in the population becomes informed with independent probability $\varepsilon \approx 0$

2. Monopolist chooses a fraction of the population ω to inform directly through a costly advertising technology

3. Informed consumers purchase the product if $\theta_i \geq P$ and tell all their friends about the product through WOM according to $\nu(\theta, P)$

4. Step 3 is repeated for newly informed consumers until there are no more consumers being informed

An important part of the analysis will be to describe the social network and how this network affects the number of people who become informed about the product. The study of random graphs goes back to the influential work of Erdös and Renyí (1959, 1960, 1961). One of the key insights of Erdös and Renyí is to consider the properties of a "typical" graph in a probability space consisting of graphs of a particular type. I assume the social network is described by random graphs with an arbitrary defined degree distribution $\{p_k\}$ (as per Newman, Strogatz and Watts (2001)) where p_k represents the fraction of individuals in the population with k friends and $\sum p_k = 1$. There are several different algorithms for constructing random graphs of this type, one is the "configuration model". Consider the following formation process for the configuration model. For a given N consider forming a sequence of n numbers which are i.i.d. draws from p_k. This is known as the "degree sequence" where the ith number k_i is the number of friends of individual i. One can think of individual i as having k_i stubs of friendships to be. Stubs are then chosen at random and connected together until there are no stubs left.[2] It has been shown that this produces every possible graph with the given degree sequence with equal probability (Molloy and Reed (1995)). The configuration model is the ensemble of graphs $\Omega_{N,\{p_k\}}$ produced via this procedure and the properties derived in the analysis are for the average over this ensemble of graphs in the limit as $n \to \infty$.

I now define a number of characteristics of networks.

Definition 1 *A path exists between two individuals i and j if there exists a sequence of individuals where i is the first member of the sequence and j is the last member of the sequence such that for the $(t+1)$th member of the sequence $(t, t+1) \in \Xi$*

Using this definition of a path I define a component.

[2] This process assumes there is an even number of stubs to begin with and does not rule out two stubs from the same individual connecting to one another or multiple links existing between two individuals. Under some regularity conditions on $\{p_k\}$ the instances of own or multiple links become small in a variety of senses as the size of the network $n \to \infty$. For an excellent discussion of these issues see Jackson (2008).

Definition 2 *A component $C(i)$ of individual i is the set $\{j | \exists$ path from i to $j\}$*

The size of a component $\#C$ is the number of individuals in it. In undirected networks components are connected subsets of the population, who may all reach one another by following friendships in the network, such that $j \in C(i) \Leftrightarrow i \in C(j)$. The set of components in a network represents a partition of the set N. Denote this partition of N induced by Ξ as $\Pi(N, \Xi)$. An important part of the analysis will be the distribution of component sizes in the partition $\Pi(N, \Xi)$. Define the size of the largest component $\bar{s}$ in a graph (N, Ξ) by

$$\bar{s} = \max_{C \in \Pi(N,\Xi)} \#C$$

Definition 3 *A giant component is said to exist in a random graph with degree distribution $\{p_k\}$ if $\lim_{n\to\infty} E_{\Omega_{N,\{p_k\}}}[\bar{s}] = O\left(n^{2/3}\right)$.*

In subsequent sections the question of the existence and size of a giant component in a network will be central to the analysis. In the next section I explain how to represent a social network with a probability generating function and some of the characteristics of generating functions which will become useful for deriving the distribution of component sizes.

3 Representing social networks with random graphs with arbitrary degree distributions

A social network with an arbitrary degree distribution given by $\{p_k\}$ can be described by a probability generating function. The probability generating function $G_0(x)$ for the social network is written as:

$$G_0(x) = \sum_{k=0}^{\infty} p_k x^k$$

This is a polynomial in the generating function argument x where the coefficient on the kth power is the probability p_k that a randomly chosen individual has k friends. Generating functions have a number of useful properties that can allow one to calculate a variety of local and global properties of the social network. A good exposition of these and the formalism for calculating various properties can also be found in Newman, Strogatz and Watts (2001). I will briefly reproduce some of them here for clarity.

Derivatives The probability p_k is given by the kth derivative of G_0 according to:

$$p_k = \frac{1}{k!}\frac{d^k G_0}{dx^k}\bigg|_{x=0}$$

Moments Moments of the probability distribution can be calculated from the derivative of the generating function. The mth moment equals:

$$\textstyle\sum_k k^m p_k = \left[\left(x\frac{d}{dx}\right)^m G_0\left(x\right)\right]_{x=1}$$

Where the average degree, which I denote by z_1, is given by $z_1 = G_0^{'}\left(1\right) = \sum_k p_k k$ and the terminology $\left(x\frac{d}{dx}\right)^m$ means repeating m times the operation: differentiate with respect to x and then multiply by x.

Powers The distribution of the sum of m independent draws from the probability distribution $\{p_k\}$ is generated by the mth power of the generating function $G_0\left(x\right)$. For example, if I choose two individuals at random from the population and sum together the number of friends each person has then the distribution of this sum is generated by the function $\left[G_0\left(x\right)\right]^2$. To see this, consider the expansion of $\left[G_0\left(x\right)\right]^2$:

$$\begin{aligned}
\left[G_0\left(x\right)\right]^2 &= \left[\sum_k p_k x^k\right]^2 \\
&= \sum_{j,k} p_j p_k x^{j+k} \\
&= p_0 p_0 x^0 + \left(p_0 p_1 + p_1 p_0\right) x^1 \\
&\quad + \left(p_0 p_2 + p_1 p_1 + p_2 p_0\right) x^2 \\
&\quad + \left(p_0 p_3 + p_1 p_2 + p_2 p_1 + p_3 p_0\right) x^3 ...
\end{aligned}$$

In this expression the coefficient of the power of x^l is the sum of all products $p_k p_j$ such that $k + j = l$ and is thus the probability that the sum of the degrees of the two individuals will be l. This property can be extended to any power m of the generating function.

The distribution of the number of friends of a person found by following a randomly chosen friendship will be important in the analysis to come. This is not the same as the distribution of the number of friends of a person chosen at random because people with many friends are more likely to be found, when selected in this way, since they have more friendships. A person with k friends is k times more likely to be found than a person with 1 friend. Therefore the probability of finding a person with k friends is proportional to kp_k. After the correct normalization the generating function for this distribution is:

$$\frac{\sum_k kp_k x^k}{\sum_k kp_k} = x\frac{G_0'(x)}{z_1}$$

Now consider choosing a person randomly and looking at each of her friends. Then for each friend, the distribution of the number of friendships these people have, which do not lead back to the originally chosen person (this is $k-1$ if the friend has k friends themselves since one must lead back to original individual chosen), is generated by the function $G_1(x)$:

$$G_1(x) = \frac{G_0'(x)}{z_1}$$

The assumption that friendships between individuals are independent of one another means that as the network becomes large, $(n \to \infty)$, then the probability that any of the neighbors also know one another goes as n^{-1} and can be ignored in the limit of large n. Making use of the powers property of generating functions the probability distribution of second neighbors of the individual is given by:

$$\sum_k p_k [G_1(x)]^k = G_0(G_1(x))$$

In the analysis I utilize results regarding the robustness of random graphs with arbitrary degree distributions. Part of the analysis will consider the resultant network of individuals when the individuals who do engage in WOM are removed according to $\nu(\theta, P)$ (those people with $\theta_i - P < 0$)[3]. This is equivalent to a percolation problem on a random graph. A depiction of this process is given in Figure 1. Consider the original social network shown in Figure 1 where the individuals who desire the product $(\theta_i \geq P)$ are represented as the black nodes. The process

[3]This process of percolation is a variant of the Reed-Frost model in the epidemiology literature

of percolation takes one from this network to the network of WOM on the right where only the friendships between individuals who are willing to engage in WOM is shown.

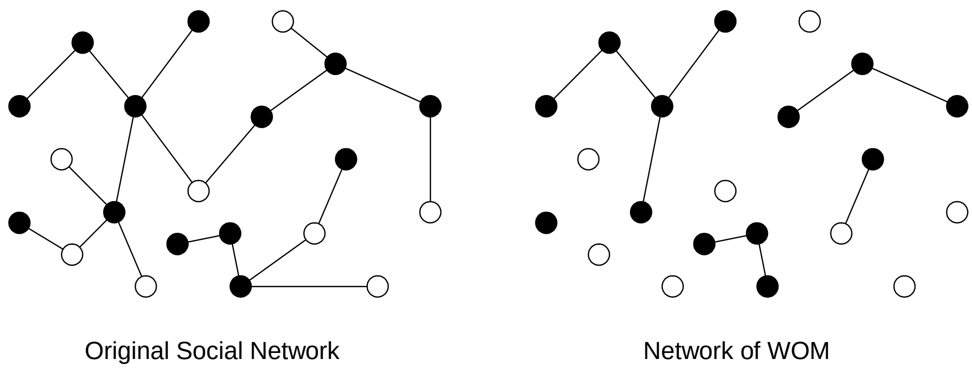

Figure 1: Percolation process

In general a giant component does not always exist in either social network. If it exists in the network of WOM then it will in the original social network but not vice versa. One of the most important and well studied topics in the random graph literature is characterizing the conditions under which the giant component does or does not exist when a parameter(s) describing the network is varied. Typically the cases, also known as phases, where the giant component does and does not exist, are separated by a critical threshold. In the model developed here this quantity will be the price which affects the probability individuals will engage in WOM. When the giant component does exist in the network of WOM every individual who knows someone in the giant component will become informed about the good. This will occur as a result of the tiny fraction ε of the population who find out about the good independently. By the law of large numbers in the limit as $n \to \infty$ at least one of these individuals will be in the giant component, and thus WOM will spread out from this person to all the people in the giant component.

The methodology for describing the network of WOM that I use in this paper was developed in Callaway et al. (2000). For expositional purposes I reproduce part of their analysis here to derive the probability generating function for this second network in terms of the first network and the probability that a person engages in WOM. To this end let q_k be the probability that an individual with k friends is willing to engage in WOM. Note this allows for some correlation between θ_i and the number of friends of individual i. The product $p_k q_k$ is the probability that a randomly chosen individual has k friends and is willing to engage in WOM. The probability generating function $F_0(x)$ for this distribution of people is given by:

$$F_0(x) = \sum_{k=0}^{\infty} p_k q_k x^k \qquad (2)$$

where $\frac{\partial^k}{\partial x} F_0(0) = p_k q_k$ and when $\nu(\theta, P)$ is given by equation 1 $F_0(1) = 1 - P$ which is the fraction of the population with valuations $\theta_i \geq P$. If we again consider following a randomly chosen friendship from the social network, the individual we reach has degree distribution proportional to kp_k rather than just p_k. Hence the probability generating function that an individual has k friends and also desires the good, when she is chosen by randomly following a friendship is:

$$F_1(x) = \frac{\sum_k k p_k q_k x^{k-1}}{\sum_k k p_k} = \frac{F_0'(x)}{z} \qquad (3)$$

3.1 Distribution of component sizes

Now let $H_1(x)$ be the generating function for the probability that one end of a randomly chosen friendship from the original social network in Figure 1 leads to a component of a given size in the network of WOM. Denote the probability that this component is of size s by h_s'. The component may in fact be empty if the individual at the end of the friendship has $\theta_i < P$ which occurs with probability $1 - F_1(1)$ or the individual may purchase the good and have k friends (distributed according to $F_1(x)$) any of whom may also purchase the good. Note that the giant component, if there is one, is *excluded* from $H_1(x)$. When component sizes are finite the chances of a finite component containing a closed loop goes as n^{-1} which becomes negligible as n becomes large. This means that the distribution of components can be represented as in Figure 2 where each component is represented as a tree like structure consisting of the single individual reached at the end of the randomly chosen friendship plus any number, including zero, of other tree-like structures.

Therefore $H_1(x)$ must satisfy the following self consistency condition:

$$\begin{aligned} H_1(x) &= 1 - F_1(1) + xq_0p_0 + xq_1p_1[H_1(x)] + xq_2p_2[H_1(x)]^2 + ... \qquad (4) \\ H_1(x) &= 1 - F_1(1) + xF_1(H_1(x)) \end{aligned}$$

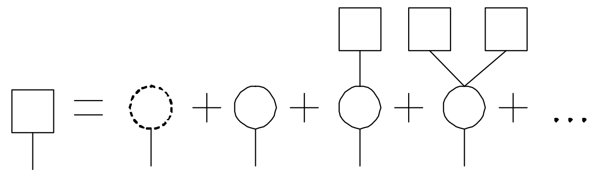

Figure 2: Schematic representation of the sum rule for components found by following a randomly chosen friendship

If an individual is chosen randomly then there is one such component at the end of each friendship of that person. Therefore the generating function for the size of the connected components in the network of WOM that a randomly chosen individual belongs is similarly generated by a function $H_0(x)$ which satisfies:

$$H_0(x) = 1 - F_0(1) + xF_0(H_1(x)) = \sum_{s=0}^{\infty} h_s x^s \tag{5}$$

where h_s is probability a randomly chosen individual from the population belongs to a component of size s. These four relationships, equations 2, 3, 4 and 5, determine the distribution of the sizes of the connected groups of individuals who communicate to one another about the good. The size of the giant component, if it exists, is given by the number of people not in components that are of finite size, $1 - H_0(1)$.

The analysis proceeds in the next section by making assumptions about the correlation between the probability a person has k friends p_k and the individual's valuation θ, and how the probability q_k that individuals pass on information about the good is related to the individual's valuation and the price charged by the monopolist. Denoting the joint distribution of θ and k in the population by $\Phi(\theta, k)$ and the conditional distribution of θ given k as $\phi(\theta|k)$ then p_k and $q_k(P)$ can be calculated as

$$\begin{aligned} p_k &= \int_\theta \Phi(\theta, k)\, d\theta \\ q_k(P) &= \int \nu(\theta, P)\, \phi(\theta|k)\, d\theta \end{aligned}$$

and substituted into the above relationships to derive $F_0(x, P)$, $F_1(x, P)$, $H_0(x, P)$, $H_1(x, P)$ in terms of the price P to describe the network of WOM. One then relates this network to the subsequent demand for the good and hence the profits of the monopolist. In principle the methodology described here may accommodate a variety of functional forms for ν relating valuations, price and even advertising effort to an individuals' probability of talking about the good. However for much of the analysis I assume θ and k are uncorrelated such that

$$q_k(P) = 1 - P$$

Also the methodology can easily incorporate a probability less than 1 that an individual passes on news about the good down any individual friendship. Denoting this probability by q_b the above expressions would be unchanged except that $F_1(x) = q_b \times \frac{\sum_k k p_k q_k x^{k-1}}{\sum_k k p_k}$. For the remainder of the paper I will focus on the case where $q_b = 1$.

4 Demand

In this section I bring the insights of percolation processes on random graphs to bear on the characterization of demand for a good in the presence of WOM. One of the central insights from the random graph literature is that there is a critical threshold which determines whether a giant component does or does not exist. I find that absent any advertising by the monopolist, demand for the good, as measured by the fraction of the population who purchase the good, exhibits two distinct regions: one where demand is zero and the giant component does not exist, and another where demand is non-zero and the giant component exists. These two regions are separated by a critical price P^{crit} below which the giant component exists and above which it does not. Provided a giant component exists in the social network itself, then as prices rise demand shrinks from a positive fraction of the population continuously to a 0 fraction of the population at the critical price where it will in general have a strictly negatively slope. In comparison to the fully informed demand curve the demand curve under WOM is more elastic. Ignoring the effect of WOM introduces a downward bias in welfare calculations and an upward bias of consumers' response to price increases after the population has become informed.

4.1 Critical Price

When there is no advertising the demand for the good is derived from the fraction $\varepsilon \approx 0$ of the population who independently find out about the good. The probability that a component of finite size s in the network of WOM becomes informed via WOM is

$$\Pr\left(C\left(\cdot\right)\text{ is informed}|\#C\left(\cdot\right)=s\right)=1-\left(1-\varepsilon\right)^{s}$$

The probability $u\left(P\right)$ that the person at the end of a randomly chosen friendship from the population does *not* become informed via WOM:

$$\begin{aligned} u\left(P\right) &= \lim_{\varepsilon\to 0}\lim_{n\to\infty}\Pr\left(C\left(i\right)\text{ is not informed}|\left(i,j\right)\text{ chosen randomly from }\Xi\right) \\ &= \lim_{\varepsilon\to 0}H_1\left(1-\varepsilon,P\right)=H_1\left(1,P\right) \end{aligned}$$

This probability is closely related to the size of the giant component. If there is no giant component $u\left(P\right)=H_1\left(1,P\right)=1$. As alluded to earlier, for the monopolist to sell to a non-zero fraction of the population there needs to be a giant component. If this is not the case all of the εn individuals will belong to components whose size is finite so total demand will be approximately a fraction ε of the population and therefore negligible as $\varepsilon\to 0$. If the giant component is of size $O\left(n\right)$, $u\left(P\right)<1$, then almost surely at least one of the εn individuals will belong to the giant component and thus the fraction of the population which becomes informed about the good is the fraction of people who know someone in the giant component. This reasoning implies that demand exhibits two distinct regions one when the giant component exists and the other when it does not depending on the price of the good. The following theorem defines these two regions in terms of a critical price below which the giant component exists and above which it does not.

Theorem 1 *Suppose an individual's valuation is independent of the number of friends, $q_k=q=1-P$ for all k. Then, there exists a critical price P^{crit} such that*

$$\lim_{\varepsilon\to 0}\lim_{n\to\infty}E_{\Omega_{N,\{p_k\}}}\left[\bar{s}\right]=O\left(n\right)\;\; \textit{if } P<P^{crit}$$

$$\lim_{\varepsilon\to 0}\lim_{n\to\infty}E_{\Omega_{N,\{p_k\}}}\left[\bar{s}\right]=O\left(n^{2/3}\right)\;\; \textit{if } P=P^{crit}$$

and

$$\lim_{\varepsilon\to 0}\lim_{n\to\infty}E_{\Omega_{N,\{p_k\}}}\left[\bar{s}\right]=O\left(\log n\right)\;\; \textit{if } P>P^{crit}$$

Moreover the critical price satisfies $1-P^{crit}=\frac{E[k]}{E[k^2]-E[k]}$.

Proof. The result follows immediately from results on percolation thresholds in the statistical physics literature cited in the appendix. ■

The intuition behind the result is best illustrated by considering the number of people who subsequently buy the product after an individual with $\theta_i \geq P$ hears about the good via WOM. If a person has k friends then the expected number of first neighbors who are informed by this person and then purchase the good themselves will be $(1-P)(k-1)$, where it is $k-1$ because the individual hears about the good from one of her friends. Now taking the expectation over the expected number of friends of a person found by following a randomly chosen friendship is $(1-P)\frac{\sum_k p_k k(k-1)}{E[k]} = (1-P)\frac{E[k^2]-E[k]}{E[k]}$. When this quantity is greater than 1 the component will initially grow exponentially, while for values less than 1 the component will decay and die out. This is known as the reproduction rate. The critical price is the price at which this reproduction rate equals 1. Subsequently when $P < P^{crit}$ a giant component exists and when $P > P^{crit}$ it does not.

4.2 Level of demand

The first step of the analysis is to determine the fraction of the population who become informed about the product. This is the fraction of people who know someone in the giant component. The probability that the person at the end of a randomly chosen friendship does *not* belong to the giant component $u(P)$ is the smallest non-negative solution to the self consistency condition equation 4:

$$u = 1 - F_1(1, P) + F_1(u, P)$$

where F_1 is now written as a function of the price P when the percolation probability is written in terms of price, $q_k(P) = \int \nu(\theta, P)\,\phi(\theta|k)\,d\theta$. The probability a person with k friends becomes informed is hence $1 - (u(P))^k$ and the total fraction of the population which is informed is:

$$\sum_k p_k \left(1 - (u(P))^k\right)$$

The second step of the analysis is to determine how many of these people purchase the product $S(P)$ this is given by:

$$S(P) = \sum_k p_k \left(1 - (u(P))^k\right) \int_P^1 \phi(\theta|k)\,d\theta$$

which given the functional form chosen for $\nu(\theta, P)$ is the size of the giant component[4]

$$S(P) = F_0(1, P) - F_0(u(P), P)$$

where again price is now an argument of F_0. Suppose there is no correlation between θ and k then $\int_P^1 \phi(\theta|k)\, d\theta = 1 - P$ for all k and $S(P)$ can be written in terms of the price P as:

$$S(P) = (1 - P)\sum_k p_k \left(1 - (u(P))^k\right) \tag{6}$$

It should be obvious from this expression that the difference between demand as generated here and the standard fully informed demand comes through the $u(P)$ term in equation 6. The distribution of valuations within the fraction of the population who find out about the product is $U[0, 1]$ because the probability a person finds out about the product is independent of her own valuation as she hears about it from a neighbor. Demand is the product of the probability that an individual finds out about the good $\sum_k p_k \left(1 - (u(P))^k\right)$ and the probability a person is prepared to purchase the good $(1 - P)$ given the price and distribution of valuations amongst the informed individuals. When the monopolist chooses a price it influences both the fraction of the population who find out about the product $\sum_k p_k \left(1 - (u(P))^k\right)$ and the proportion of these people $(1 - P)$ who are prepared to purchase it.

The following example considers a Homogeneous and a Hub social network to illustrate some of the characteristics of demand with WOM. I will then formalize these for a more general class of networks later in this section. The mean degree is 3 for both networks, in the first homogeneous social network (triangles) every individual has exactly 3 friends so the generating function is $G_0 = x^3$ and in the second Hub network (asterisks) 98% of the population have 2 friends and 2% have 52 friends so $G_0(x) = 0.98x^2 + 0.02x^{52}$. The inverse demand curves are shown below along with the fully informed inverse demand $P = 1 - Q$.Decreasing the price from $P = 1$, a giant component appears first in the Hub network, where there is greater variance in distribution of friendships, at a price $P \approx 0.94$. Demand grows relatively slowly because it is unlikely that the individuals with $\theta_i \geq P$ and 2 friends become informed when the giant component is very small. As the price falls further the giant component grows faster and the inverse demand curve appears convex in this region.

[4] Callaway et al. (2000) derive this expression using the generating function approach. Molloy and Reed (1995) derive an equivalent expression for the size of the giant component using a different methodology.

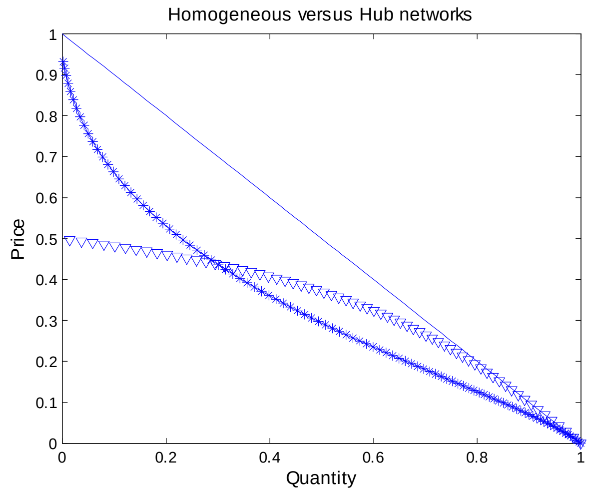

Figure 3: Homogeneous versus Hub networks

When the price reaches $P = 0.5$ a giant component appears in the Homogenous network. Initially the giant component in the Homogeneous network grows very quickly compared to the Hub network because everyone has the same number of friends. In fact the giant component in the homogeneous network becomes larger than in the Hub network at $P \approx 0.45$ and at a price of 0.3 it contains approximately 40% more individuals. This difference is driven by the relative likelihood of a person with 3 friends versus 2 friends becoming informed in this range of prices. Eventually the giant component in the Homogenous network consists of almost all individuals for whom $\theta_i > P$ so $S \approx 1 - P$ and it can only grow at the rate at which new people are willing to purchase the product for a given price change. Since both networks are fully connected eventually the giant component in the Hub network also approaches $1 - P$ and for both networks $S = 1$ at $P = 0$.

The following theorem characterizes demand as price varies absent any direct advertising by the monopolist.

Theorem 2 *Suppose θ and k are uncorrelated then demand for the good $S(P)$ is*

1. *Continuous*

2. $$\begin{array}{l} S(P) = 0 \text{ for } P \geq P^{crit} \\ S(P) > 0 \text{ for } P < P^{crit} \end{array}$$

3. $\frac{dS}{dP} = 0$ *for* $P \geq P^{crit}$
$\frac{dS}{dP} < 0$ *for* $P < P^{crit}$

4. $\lim_{P \to P^{crit-}} \frac{dS}{dP} = -\left(1 - P^{crit}\right) \frac{G_0'''(1)}{\left(G_0''(1)\right)^2} < 0$

5. $\left|\frac{P}{S}\frac{dS}{dP}\right| > \left|\frac{P}{1-P}\right|$ *for* $P < P^{crit}$

Proof. See appendix. ■

This theorem establishes that demand is continuous in price and that at the critical price the slope of demand makes a discontinuous change from zero in the region $P > P^{crit}$ to a strictly negative amount at P^{crit}. This change in the growth rate distinguishes the two regions of demand. This change in the behavior of demand does not come as a result of the fully informed demand having a negative slope at $P = 1$. Indeed provided that a fraction $1 - P^{crit}$ of the population have valuations θ greater than P^{crit} and valuations are locally distributed uniform with density 1 around P^{crit} the above theorem will continue to be true. This means that the fully informed demand may in fact asymptote to 0 as P increases such that for the inverse demand curve $\lim_{Q \to 0} \frac{dP}{dQ} = 0$ and the theorem will be unchanged. The elasticity of demand when there is WOM $\frac{P}{S}\frac{dS}{dP}$ is:

$$\frac{P}{S}\frac{dS}{dP} = \frac{-P}{1-P}\left[1 + \frac{(1-P)}{1 - \sum\limits_k p_k u^k} \frac{du}{dP} \sum_k p_k k u^{k-1}\right]$$

which is the fully informed elasticity adjusted by a factor $1 + \frac{(1-P)}{1-\sum\limits_k p_k u^k} \frac{du}{dP} \sum\limits_k p_k k u^{k-1}$ where the second term comes from the increase in connectivity of the network from lowering the price. This is new customers, with $k \geq 2$, forming a bridge to the giant component to connect previously disjoint components of individuals.

4.3 Biases in estimates which ignore WOM

One can imagine using cross-sectional data to non-parametrically identify the relationship for $S(P)$. In this section I find that failing to recognize the effects of WOM in generating this demand may lead to a several of biases. The first is a downward bias in welfare calculations of consumer surplus of the form $\int_{\tilde{P}}^{\infty} S(P)\, dP$.

Corollary 1 *Suppose the price of the good is* $\tilde{P}$ *then an estimate of consumer surplus* $\widehat{CS}\left(\tilde{P}\right) = \int_{\tilde{P}}^{\infty} S(P)\, dP$ *is biased downwards.*

Proof. See Appendix. ■

An estimate of the valuations of consumers who purchase the product based on the demand curve $S(P)$ understate the valuations of the *purchasing consumers*. It is obvious that the population being uninformed leads to fewer consumers purchasing the product than if they were fully informed, however this corollary implies that even amongst the consumers who do find out about the product and purchase it, an estimate based on $S(P)$ of *their* valuations will be biased downwards. The reason is that the marginal consumers at a price of P are a combination of individuals who know about the product and have valuations $\theta \approx P$, and consumers in previously disjoint components with valuations $\theta \sim U[P, 1]$ who become informed via one of the consumers with $\theta \approx P$. Failing to recognize that demand changes through this second channel induces a downward bias in estimates of consumer valuations because it attributes a valuation of $\theta \approx P$ to a group of consumers with valuations $\theta \sim U[P, 1]$. Thus welfare calculations such as evaluating the introduction of a new good will understate the consumer surplus.

A second bias may occur when considering how consumers will respond to an increase in price once WOM has diffused.

Corollary 2 *Suppose the price of the good is $\tilde{P}$ then an estimate $\Delta\hat{S}$*

$$\Delta\hat{S} = S\left(\tilde{P}\right) - S\left(\tilde{P} + \Delta P\right)$$

of the consumer response to a price increase ΔP overstates the actual response ΔS

$$\Delta\hat{S} < \Delta S$$

Proof. See Appendix ■

The distribution of valuations are distributed $U[0, 1]$ across those people who are informed about the product. An increase in the price by ΔP will change demand by $\frac{\Delta P}{1-\tilde{P}}\%$ however an estimate based on $S(P)$ overstates the elasticity with respect to price of the consumer's preferences for the product and will predict a greater response. A monopolist choosing to increase its price or a policy maker introducing a tax will estimate a larger change in demand than what would actual take place.

5 Static Pricing

In this section I study the optimal static pricing decision of the monopolist. For regular goods, where valuations and number of friends are uncorrelated, I show that the monopolist will set a lower price when there is WOM compared to when consumers are fully informed. However, for goods where there is significant positive correlation between valuation for the good and an individual's number of friends then the monopolist may in fact price above the fully informed level. When the monopolist can price discriminate between consumers based on numbers of friends, then better connected individuals are charged lower prices.

5.1 Regular goods

The first result in this section is that the monopolist will set a lower price when there is WOM compared to when consumers are fully informed.

Theorem 3 *Suppose valuations and number of friends are uncorrelated and marginal costs $c < 1$, then a monopolist facing demand given by $S(P)$ charges a lower price P^*_{WOM} than a monopolist facing a fully informed population P^*_{FI}, where demand is given by $Q(P) = 1 - P$.*

Proof. See appendix. ■

This theorem comes as an immediate consequence of demand being more elastic under WOM in Theorem 2. The WOM monopolist has an additional incentive to stimulate demand through the word of mouth channel and will lower prices below the price that would be charged by the monopolist facing a fully informed population. The effect can be so large that consumers may in fact be better off being uninformed than fully informed.

Corollary 3 *Consumer surplus may be greater when consumers are uninformed and the monopolist charges P^*_{WOM} than if consumers are fully informed and the monopolist charges P^*_{FI}.*

Proof. See appendix ■

This proposition illustrates that consumers may in fact be better off when they are uninformed because the monopolist lowers the price below P^*_{FI} to stimulate word of mouth in the population. The gains to consumers may in fact be quite significant, for instance in a social network where everyone has 3 friends the consumer surplus is 65% larger in the WOM setting than the fully informed setting. Of course a social networks which does not have a significant fraction of the population in the giant component at any price it will no longer be true.

5.2 Correlation Between Valuations and Number of Friends

For goods where there is significant correlation between the connectivity of individuals and their valuation for the product, in contrast to Theorem 3, it can be the case that the monopolist will charge a price higher than it would if everyone is informed. When there is significant positive correlation, the network of WOM is much better connected at higher prices than a network with no correlation. The following proposition illustrates a case where significant positive correlation leads to prices above the fully informed monopoly price $\frac{1+c}{2}$.

Theorem 4 *If $P^{crit} > \underline{\theta}$ and all consumers with $\theta \in [c, \underline{\theta}]$ have $k = 1$ then the monopoly price will be greater than the fully informed monopoly price $\frac{1+c}{2}$.*

Proof. See Appendix ■

The intuition for this result is that when the mix of marginal consumers has a large fraction of individuals with low connectivity then demand will be relatively inelastic. In this theorem the mix contains only individuals with 1 friend. These consumers can not provide a bridge to connect components which are disjoint from the giant component for $c \leq P \leq \underline{\theta}$, thus demand is relatively inelastic compared to the fully informed demand over the range of prices $P \in [c, \underline{\theta}]$ and the monopolist will not price at or below the fully informed monopoly price $\frac{1+c}{2}$.

The types of goods which would naturally have some correlation between valuation and the number of friends are fashion and status products, where the value is, at least in part, increasing in the consumer's ability to display them to others. The example given here suggests that these types of goods will receive a higher mark up than other types of goods all else equal.

5.3 Price discrimination

When the monopolist can discriminatingly price to consumers with different numbers of friends, the optimal set of prices will be decreasing in the number of friends each person has. For the monopolist there is a greater incentive to decrease the price offered to individuals with more friends since these individuals are the most effective at informing others. When the monopolist decreases the price to one of the groups it can increase the number of people informed of all groups through WOM.

Monopolist's maximization problem when it can discriminate between consumers with different numbers of friends is:

$$\pi(\{P_k\}) = \max_{\{P_k\}\in[0,1]^n} \sum p_k q_k \left(1 - u(\{P_k\})^k\right)(P_k - c) \tag{7}$$

where the value of u is now a function of the set of prices $\{P_k\}$

Theorem 5 *If valuations and number of friends are uncorrelated and* $\exists \{P_k\}$ *such that* $\pi(\{P_k\}) > 0$ *then the optimal set of prices* $P_1 = \frac{1+c}{2}$ *and* $\exists \underline{k} : \{P_k\}$ *is decreasing for* $2 \leq k \leq \underline{k}$ *and* $P_k = 0$ *for* $k \geq \underline{k}$.

Proof. See appendix ■

The proof considers the complementarity of demand from the different groups of consumers. In fact the problem is equivalent to a multiproduct monopolist's problem where $p_k q_k \left(1 - u^k\right)$ in equation 7 is the demand for good k and the demands for each good are complementary through the value of $u(\{P_k\})$. When marginally adjusting a price P_k the monopolist faces the usual pricing incentives over the informed population $(1 + c - 2P_k)$ plus the impact of changing the price on the size of the informed population through $u(\{P_k\})$. The relative trade-off between these two effects is proportional to $\frac{k\left(1-u^{k-1}\right)}{1-u^k}$ which is increasing in the number of friends k. Hence P_k is decreasing in the number of friends. In fact it can be profitable to give the good away for free to individuals with sufficiently many friends because of the size of their influence on demand from individuals with fewer friends.

This is a very intuitive result that offering discounts to the individuals who are best able to spread news about the good increases the profits of the monopolist. As discussed earlier the individuals with a large number of friends are very influential because these individuals are both more likely to hear about the good and able to inform more people. There have been a number of authors who have emphasized the importance of market mavens for spreading information about products (for instance Feick and Price (1987) and Gladwell (2000)). Interpreting market mavens as people who are able to influence many people within the social network then this theorem underlines the importance of providing a discount to these types of consumers because of the significant complementarity between their choice to buy the product and the total number of people who hear about the product.

6 Introductory pricing

In this section I find that introductory pricing involves periods of sales. The monopolist may fluctuate the price of the good up and down to optimally diffuse news of the good in the population. The trade off facing the monopolist is to sacrifice immediate profits to facilitate greater WOM today and a larger population of informed consumers in the future. The natural intuition in this situation, is that the dynamic sequence of prices will be increasing because as more and more people become informed there is less incentive for the monopolist to keep the price below the monopoly level. I show that this not necessarily the case for prices soon after the good is introduced, during the early stages of diffusion of WOM.

I will assume the good is non-durable to avoid the added complexity of strategic purchasing decisions by consumers. In this section I also assume for tractability that the marginal cost is 0 and that valuations and number of friends are uncorrelated. In each period consumers who know about the good will purchase it if $\theta_i \geq P_t$. In the first period, $t = 0$, a small number of people M_0, hear about the good and decide whether to purchase it at the price P_0. Those that purchase the good tell their friends, who are then added to the total population of informed consumers in the next period denoted M_1. In this way M_t grows over time. The current period payoff can be written as $M_t P_t (1 - P_t)$ where $1 - P_t$ represents the distribution of valuations ($\theta \sim U[0, 1]$) across this population. The distribution of valuations within M_t does not change because becoming informed via WOM from a friend is independent of an individual's own valuation. Hence it is a random draw from the distribution of valuations within the population. The change from one period to the next $M_{t+1} - M_t$ comes through the number of people who purchase the good for the first time during period t and then tell their friends about it. The number of people who know about the good, but have never purchased it, are the conduit for this change. I will denote this population of people by R_t and the distribution of valuations in it by $F_t(\theta)$. Unlike the distribution of valuations across M_t, $F_t(\theta)$ may change as M_t grows. When a person in R_t purchases the good that person will not be in R_{t+1}, since they have now purchased the good, but all their friends, who are now informed from WOM, will be in R_{t+1}, since they are informed about the good but are yet to have purchased it. If a person is in R_t but does not purchase the good during period t, $\theta < P_t$, then that person will also be in R_{t+1}. Thus after a sequence of prices $P > 0$ a stock of people with low valuations can build up in R_t. Depending on the sequence of prices the distribution of valuations within R_t changes.

The number of friends a person tells when they purchase the good for the first time is both a function of the time at which they found out about the good (since people with many friends find out about the good earlier than those with few) and also the size of the population when they do purchase, when a large fraction of the total population knows about the product there is a probability that more than one of their friends have already found out about the product from someone else in the past or even in the current period. The transition M_t to M_{t+1} is a stochastic process and depends on the distribution of both valuations and number of friends of individuals within R_t. Characterizing how this distribution and R_t evolve over time is a complicated problem. To illustrate why a monopolist may fluctuate the price up and down over time I will consider a simplified problem to avoid a number of the complexities that occur in the more general setting.

I will focus on a branching problem which assumes that the market is a mass of people M_t which can grow without bound such that it never consists of a significant fraction of the population. This is of course unrealistic over long time horizons since, if the market continues to grow, at some point it will be bound by the size of the population. Notwithstanding this, it does allow a much more tractable characterization of the problem, which I argue is a reasonable approximation of behavior close to when the product is first introduced and characterizes the incentives the monopoly faces for introductory prices. This setting allows one to characterize how the change of valuations within R_t can lead the monopolist to fluctuate the price up and down over time.

6.1 Infinite horizon branching problem

At the start of period 0 a unit mass $M_0 = 1$ of individuals find out about the good. During each period the monopolist chooses prices $\{P_0, P_1, P_2...\}$ and in each period the mass of informed individuals M_t chooses whether or not to purchase the good. The monopolist faces a trade off between making profits over the existing population of informed individuals and lowering the price to sell to a greater number of individuals in R_t, who are informed but yet to purchase the good, in order to increase the rate at which the mass of informed individuals grows. The expected number of individuals who become informed when a member of R_t purchases the good for the first time is the reproduction rate $G_1'(1) = \frac{z_2}{z_1}$. The growth rate conditional on price is deterministic because I assume M_0 is a unit mass of consumers.

In this problem there are three state variables and one control variable. The state variables are the number of people informed of the good M_t, the number of people who are both informed about the good but are yet to purchase it R_t and the distribution of valuations within these people F_t. The control variable is the price in each period P_t.

6.1.1 Reducing the number of state variables

In this section I reduce the number of state variables from three to two by considering the ratio of individuals who are informed but have never purchased to those that are informed, this is $\frac{R_t}{M_t}$. I assume that the set of individuals in M_0 are found in such a way that $\frac{R_0}{M_0} = \frac{z_2 - z_1}{z_2}$. Consider how $\frac{R_0}{M_0}$ changes when someone in R_0 purchases the good, the change of the state variables M_0 and R_0 are $\Delta M_0 = \frac{z_2}{z_1}$ and $\Delta R_0 = \frac{z_2}{z_1} - 1$. The reproduction rate $\frac{z_2}{z_1}$ is the expected number of additional people who become informed ΔM_t when a person in R_t purchases the good, and $\frac{z_2}{z_1} - 1$ is the number of additional people in R_t when this happens ΔR_t (the -1 comes from the purchasing individual no longer being in R_t after purchasing). Therefore the new ratio is

$$\begin{aligned}
\frac{R_t + \Delta R_t}{M_t + \Delta M_t} &= \frac{\frac{z_2 - z_1}{z_2} M_0 + \frac{z_2}{z_1} - 1}{M_0 + \frac{z_2}{z_1}} \\
&= \frac{z_2 - z_1}{z_2}
\end{aligned}$$

thus as more and more individuals purchase, the ratio $\frac{R_t}{M_t}$ remains constant. Using this relationship I can eliminate one state variable which I choose to be R_t.

6.1.2 Characterizing the transition functions

In this section I characterize the transition functions for both M_t and F_t which I denote Γ_M and Γ_F respectively. The population of informed individuals next period M_{t+1} is the population last period M_t plus the number of people who hear about the good through WOM from the consumers in R_t. This relationship is:

$$M_{t+1} = M_t + R_t \left(1 - F_t\left(P_t\right)\right) \frac{z_2}{z_1}$$

Using the relationship $\frac{R_t}{M_t} = \frac{z_2 - z_1}{z_2}$and substituting this into the transition function for M_t :

$$\begin{aligned} M_{t+1} &= \left(\left(1 - F_t\left(P_t\right)\right) \frac{z_2}{z_1} + F_t\left(P_t\right) \right) M_t \\ &= \Gamma_M\left(M, F, P\right) \end{aligned}$$

The distribution of valuations across the set of people yet to purchase R_t will depend on the distribution the previous period and the price in the previous period. The cumulative distribution function this period F_t (with associated pdf f_t) will be a weighted combination of the distribution last period f_{t-1} truncated at P which is the set of people in R_{t-1} who didn't buy last period $\left(F_{t-1}\left(P_{t-1}\right) R_{t-1}\right)$ and a uniform distribution over the newly informed people $\left(1 - F_{t-1}\left(P_{t-1}\right)\right) \left(\frac{z_2}{z_1} - 1\right) R_{t-1}$. The relative weights for each are

$$\frac{1}{1 + \left(1 - F_{t-1}\left(P_{t-1}\right)\right) \left(\frac{z_2 - z_1}{z_1}\right)}$$

for f_{t-1} and

$$\frac{\frac{z_2}{z_1}\left(1 - F_{t-1}\left(P_{t-1}\right)\right)}{1 + \left(\frac{z_2}{z_1} - 1\right)\left(1 - F_{t-1}\left(P_{t-1}\right)\right)}$$

on the uniform. Thus the transition function for F_t is

$$\begin{aligned} F_{t+1}\left(\theta\right) &= \frac{\min\left[F_t\left(\theta\right), F_t\left(P_t\right)\right] + \frac{z_2}{z_1}\left(1 - F_t\left(P_t\right)\right)\theta}{1 + \left(\frac{z_2}{z_1} - 1\right)\left(1 - F_t\left(P_t\right)\right)} \\ &= \Gamma_F\left(F, P\right) \end{aligned}$$

Define $\mathcal{F}$ as the set of continuous cdfs on $[0, 1]$ which satisfy $\frac{F(x) - F(x-\delta)}{\delta} \leq \frac{z_2}{z_2 - z_1}$.

Lemma 1 *If $F \in \mathcal{F}$ then $\Gamma_F\left(F, P\right) \in \mathcal{F}$.*

Proof. See appendix ■

This lemma bounds the density of valuations in R_t above and is used to establish the continuity of the mapping Γ_F.

Lemma 2 $\Gamma_M : [1, \infty) \times \mathcal{F} \times [0, 1] \to [1, \infty)$ *and* $\Gamma_F : \mathcal{F} \times [0, 1] \to \mathcal{F}$ *are continuous mappings*

Proof. See appendix ■

The transition functions are single valued mappings and their continuity helps ensure the problem is well behaved. The following lemma derives the limiting distribution of F_t for a constant price P^*

Lemma 3 *If* $P_t = P^* < P^{crit}$ *for all* t *and* $F_t \in \mathcal{F}$ *then the limiting distribution* $f^*(\theta) = \lim_{t\to\infty} f_t(\theta)$ *will be*

$$
\begin{aligned}
f^*(\theta) &= \frac{z_2}{z_2 - z_1} \text{ if } \theta < P^* \\
&= \frac{z_2 - \frac{z_1}{1-P^*}}{z_2 - z_1} \text{ if } \theta \geq P^*
\end{aligned}
$$

Proof. See appendix ■

Given a distribution f_t and price P_t then $\frac{df_t(\theta)}{dt} > 0$ if $f_t^*(\theta) > f_t(\theta)$ and $\frac{df_t(\theta)}{dt} < 0$ if $f_t^*(\theta) < f_t(\theta)$ and $\frac{df_t(\theta)}{dt} = 0$ if $f_t(\theta) = f^*(\theta)$ for all θ. The key characteristic of this problem is that there is a discontinuity in the incentives between marginally increasing vs marginally decreasing the price above and below P^*. When the monopolist charges a price greater than zero there is a stock of people who know about the good but are yet to purchase it. This stock is the difference between the density $f^*(\theta)$ at $\theta < P$ compared to $\theta \geq P$. I show in the following section that it is this characteristic which leads the monopolist to fluctuate the price up and down.

6.2 Introductory pricing problem

The monopolist's problem is the following:

$$
\begin{aligned}
J(M_0, F_0) &= \max_{\{P_t\}} \sum_{t=0}^{\infty} \beta^{t-1} P_t (1 - P_t) M_t \\
& st \\
M_{t+1} &= \left((1 - F_t(P_t)) \frac{z_2}{z_1} + F_t(P_t) \right) M_t \\
F_{t+1}(\theta) &= \frac{\min[F_t(\theta), F_t(P_t)] + \frac{z_2}{z_1}(1 - F_t(P_t))\theta}{1 + \left(\frac{z_2}{z_1} - 1\right)(1 - F_t(P_t))} \\
M_0 &= 1 \\
F_0 &= \theta
\end{aligned}
$$

I make the following assumption about the network structure and discount factor

$\beta < \frac{z_1}{z_2} < \frac{1}{2}$ so that the problem is well posed.

The problem is an optimal control problem where the state is an element of $(M, F) \in [1, \infty) \times \mathcal{F}$ and the control is the price $P \in [0, 1]$. Writing it recursively:

$$V(M, F) = \max_{P \in [0,1]} P(1-P) M + \beta V(M', F')$$

subject to

$$\begin{aligned} M' &= \Gamma_M(M, F, P) \\ F' &= \Gamma_F(F, P) \end{aligned}$$

Theorem 6 *The monopolist's problem has a unique solution, the value function is continuous and homogeneous of degree 1 in M and the policy function $P(F)$ is upper hemicontinuous and only a function of the state F.*

Proof. See appendix ∎

A brief outline of the argument is as follows. The proof proceeds by defining a contraction mapping T :

$$(TV)(M, F) = \max_{\substack{P \in [0,1] \\ M' = \Gamma_M(M,F,P) \\ F' = \Gamma_F(F,P)}} P(1-P) M + \beta V(M', F')$$

and looking for a solution in the space of continuous functions $V : [1, \infty) \times \mathcal{F} \to \mathbb{R}$ which are bounded in the norm

$$\|V\| = \max_{\substack{F \in \mathcal{F} \\ M=1}} V(M, F)$$

Letting $H(M, F)$ be the space of these functions. Then the maximization is for a continuous function over a compact set $P \in [0, 1]$ so the maximum exists. Then from the Theorem of the Maximum (Berge 1963) the maximum is continuous and from the homogeneity of the problem with respect to M the contraction T maps $H(M, F) \to H(M, F)$. Using the contraction mapping theorem the contraction has a unique fixed point which satisfies the recursive relationship. The properties of the policy function then follow immediately from the theorem of the maximum and homogeneity of the value function with respect to its first argument.

The value function is linear in M and the policy function is only a function of the distribution of valuations in the set of people who are informed but yet to purchase the good. I am able to further characterize the dynamic set of prices in the following theorem which highlights the incentives of the monopolist to move the price up and down.

Theorem 7 *$\nexists T$ such that for all $t > T$ the optimal price sequence $\{P_t^*\}$ is weakly increasing or decreasing.*

Proof. See appendix ■

The argument is a proof by contradiction. I first show that the optimal prices $P_t^* \in \left[0, \frac{1}{2}\right]$ and that if $\{P_t^*\}$ is weakly increasing or decreasing then the sequence will converge to a price $P^* \in \left[0, \frac{1}{2}\right]$. In this case F_t will converge to the distribution F^* given in lemma 3. This distribution is kinked at the price P^* where the density is discontinuous. The contradiction comes from considering deviations $P_t + \delta$ and $P_t - \delta$. The growth rate $\frac{M_{t+1}}{M_t}$ is $\left((1 - F_t(P_t))\frac{z_2}{z_1} + F_t(P_t)\right)$ thus the marginal change in growth rate is proportional to $\lim_{P \to P_t^+} f^*(P_t)$ for $P_t + \delta$ and $\lim_{P \to P_t^-} f^*(P_t)$ for $P_t - \delta$. The kink in F^* means that $\lim_{P \to P_t^+} f^*(P_t) < \lim_{P \to P_t^-} f^*(P_t)$. The contradiction then comes from showing that for a small enough δ one of the two deviations is profitable.

This theorem shows that the monopolist has the incentive to not hold the price constant but rather fluctuate it up and down. The intuition for the proof illustrates why this is the case. If the price remains constant, or close to constant, for a period of time there is a stock of individuals with valuations slightly below the price who know about the good but are yet to purchase it. At some points in time it becomes worthwhile for the monopolist to drop the price to allow these consumers to purchase the product and subsequently inform their friends. This provides an intuitive explanation of sales whereby the benefit of the sale is reaped in future periods from the increased WOM it induces. This theory of sales is a rather natural one, the sale generates greater future demand through the additional WOM from people who wouldn't normally purchase the good.

7 Advertising

In this section I study the advertising decision of the monopolist by allowing it to engage in informative advertising. Advertising allows the monopolist to spread news of the good to individuals in components outside the giant component. I find that in the presence of WOM, marginal returns to advertising exhibit a peak at the critical price, a monopolist selling an exclusive (high price) good will target advertising at individuals with many friends whereas a monopolist selling a common (low price) good will target advertising at individuals with relatively fewer friends, and an owner of the rights to advertise to people within the social network will optimally allocate advertising for exclusive products to well connected individuals and advertising for common products to less well connected individuals.

Throughout this section I will talk about the returns to advertising not in terms of profit or revenue but rather in terms of how many consumers a specified level of advertising attracts to the product. The effects of direct advertising can be thought of as striking entire components of individuals within the network of WOM represented by F_0 and H_0. Whenever anyone within a component of individuals finds out about the product, the entire component becomes informed via WOM as members of the component pass on news about it. The marginal returns from increasing the level of advertising are the number of additional consumers found by advertising to another individual chosen at random from the population of people not already advertised to. This can be thought of as a traditional advertisement where ω (fraction of the population) represents the level of exposure it gets in the population. For a given level of advertising, the marginal returns from advertising can be written as a function of the distribution of component sizes, where $h_s(P)$ is the probability an individual chosen at random belongs to a component of size s, for a given price P. When the level of advertising is ω the probability that the next person advertised to belongs to a component of size s, which has not already been found via advertising (none of the other members of the component have been advertised to) is $h_s(P) \times (1-\omega)^{s-1}$ where $(1-\omega)^{s-1}$ is the probability that no one else in the component has been advertised to as well. The marginal return is therefore:

$$\sum_s s h_s(P)\,(1-\omega)^{s-1} = H_0'\,(1-\omega, P)$$

and the aggregate return is:

$$\begin{aligned}
& \int_0^w H_0'\,(1-\omega, P)\,d\omega \\
= \quad & H_0\,(1, P) - H_0\,(1-\omega, P) \\
= \quad & 1 - S\,(P) - H_0\,(1-\omega, P)
\end{aligned}$$

Assuming a constant cost per unit of advertising α and marginal cost of production c, the monopolist's profit is defined by

$$\pi(P,\omega) = (P-c)(1-H_0(1-\omega,P)) - \alpha\omega$$

Theorem 8 *For all* $(\omega,P) \in [0,1]^2 \setminus (0,P^{crit})$, $\pi(P,\omega)$ *is continuous and differentiable with respect to both price and advertising, and* $\lim\limits_{(\omega,P)\to(0,P^{crit})} \pi(\omega,P) = 0.$

Proof. See appendix ■

Corollary 4 *If* $\pi(\omega',P') > 0$ *for some* (ω',P') *then* $\exists\varepsilon > 0$ *such that for all* $(\omega,P) \in B_\varepsilon(0,P^{crit})$ *where* B_ε *is an open ball* $\pi(\omega,P) < \pi(\omega',P')$.

Proof. See appendix ■

This theorem and corollary mean that if we find $(\omega^*,P^*) \in [0,1]^2 \setminus B_\varepsilon(0,P^{crit})$ which maximizes $\pi(\omega,P)$ then this is the optimal strategy for the monopolist. The set $[0,1]^2 \setminus B_\varepsilon(0,P^{crit})$ is compact and $\pi(\omega,P)$ is continuous so the optimal strategy exists, and we can apply the theorem of the maximum to the problem hence $\pi(\alpha)$ is continuous and $(\omega(\alpha),P(\alpha))$ is upper hemicontinuous. Necessary conditions for the optimal price and level of advertising are

$$(1-H_0(1-\omega^*,P^*)) - (P^*-c)\left.\frac{\partial H_0}{\partial P}\right|_{(1-w^*,P^*)} \leq 0$$

and

$$(P^*-c)\left.H_0'(x,P)\right|_{(1-w^*,P^*)} - \alpha \leq 0$$

An implication of Corollary 3 is that increasing the advertising cost α, can in fact increase consumer surplus as the monopolist reduces advertising and relies more on the price to stimulate WOM amongst consumers. If advertising is free $\alpha = 0$ then the monopolist chooses $\omega = 1$ and a price $P = \frac{1+c}{2}$, as $\alpha \to \infty$ the monopolist will choose $\omega = 0$ and $P = P^*_{WOM}$ for α high enough. Corollary 3 shows that consumers can in fact be better off in the latter case. Indeed Figure 4 illustrates how consumer surplus can increase or decrease when the advertising cost increases for a social network $G_0 = x^3$ and marginal cost $c = 0.42$. In Figure 4 (a) and (b) increasing the advertising cost corresponds to moving along the curves shown starting from the upper right. As the advertising cost increases the monopolist cuts back on the level of advertising and compensates by decreasing the price to stimulate word of mouth.

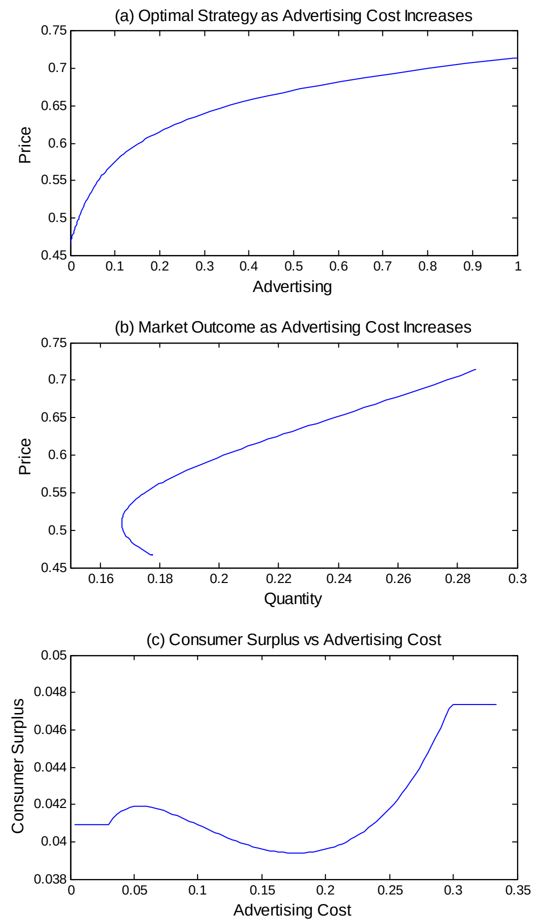

Figure 4: Effects of increasing advertising costs

When advertising costs exceed 0.18 the monopolist starts to dramatically decrease the price which increases the equilibrium quantity and consumer surplus despite the lower levels of advertising taking place.

There is a wide range of potential equilibrium price and advertising pairs depending on the marginal costs of production and advertising. In the following sections I focus on characterizing the marginal returns to advertising.

7.1 Marginal returns to advertising

In this section I find that the marginal returns to advertising exhibit a peak as $P \rightarrow P^{crit}$ $\omega \rightarrow 0$, and are decreasing and convex with respect to advertising. The marginal return to the first unit of advertising is the average size of components containing uninformed individuals. I find that the average size of these components (marginal returns of the first unit of advertising) exhibit a very distinctive feature around the critical price. In particular the average component size asymptotes to infinity as the price advertising strategy pair approaches the critical price with zero advertising. This implies that for low levels of advertising there are regions where marginal returns are sharply increasing and decreasing at prices close to the critical price. I provide examples of how the marginal returns vary across a number of networks.

The following theorem characterizes the marginal returns to advertising close to the critical price and zero advertising.

Theorem 9 *If* $0 < P^{crit} < 1$, *then* $\lim_{(\omega,P)\rightarrow(0,P^{crit})} H_0'(1,P) = \infty$.

Proof. See appendix ■

This theorem implies that around the critical price the marginal returns to the first units of direct advertising increase and decrease very sharply. The sharp increase is caused by the phase transition where the giant component appears. The distribution of component sizes contains more and more very large components as the price approaches the critical price. I conclude that for low levels of advertising marginal returns to advertising will exhibit a peak at a price close to the critical price. I contrast this result to an identical model of advertising without WOM. In such a model if the monopolist advertises to $w\%$ of the population $(1-P)\,w\%$ of the people will end up buying the product if the price is P. The demand in this model is linear in the level of advertising and in contrast to the WOM case the

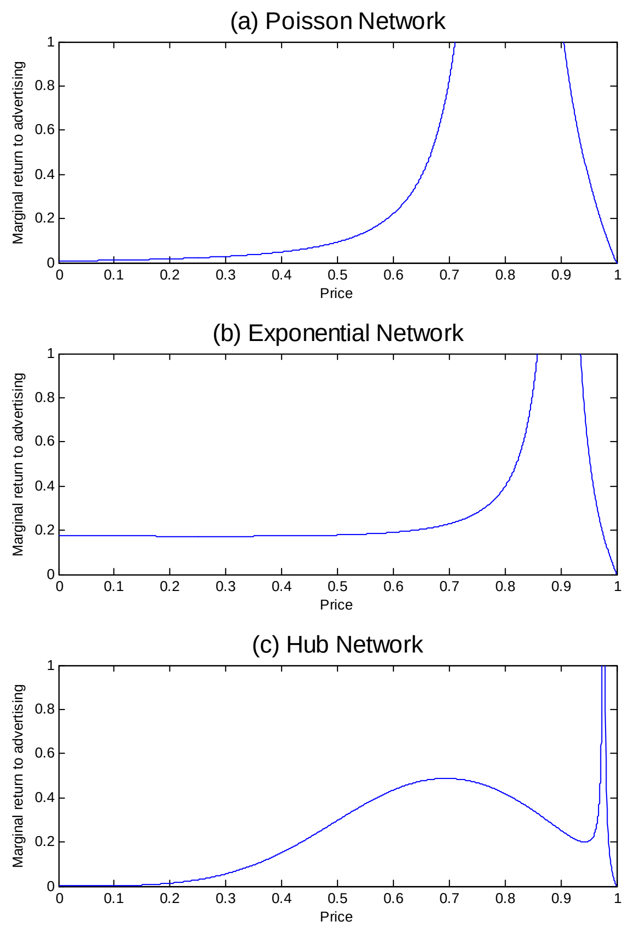

Figure 5: Marginal returns to advertising for different social networks

marginal returns are constant. Figure 5 illustrates how the marginal returns to the first unit of advertising vary across three networks, each with a mean number of friendships per individual of 5, Poisson, Exponential and Hub (2% have 103 friends and 98% have 3 friends). Each has the distinctive spike at the critical price, however for prices below the critical price the networks are very different. In the Poisson network the marginal returns are strictly increasing, in the Exponential network they are approximately constant until $P = 0.6$ before they start to increase, and in the Hub network the marginal returns are non-monotonic.

The following theorem characterizes marginal returns as the level of advertising changes.

Theorem 10 *Advertising exhibits decreasing and convex marginal returns.*

Proof. See appendix ■

As advertising increases the largest components are relatively more likely to be struck first by the advertising because of their size. Thus as the level of advertising increases the marginal returns fall away sharply at first and then flatten out at higher levels of advertising as the mix of unadvertised components contains a greater fraction of small sized components. This can be seen for the Poisson network in Figure 6 where the distinctive spike in marginal returns is evident close to the critical price and zero advertising but as advertising increases the marginal returns fall away sharply and are much flatter at higher levels of advertising.

7.2 Targeted marketing

If the monopolist can target its advertising at individuals with a certain number of friends then how should it do so? In this section I find that when prices are high and the giant component doesn't exist or is small then individuals with many friends should be targeted, if on the other hand the giant component is large then it is more effective to target advertising at those people with few friends who are least likely to be in the giant component.

Consider the question of which individual the monopolist should advertise to first? I fix the price at a level P and assume the monopolist can observe the number of friends of an individual. The return from advertising to an individual with k friends is the size of the component the individual belongs to, $(1-P)\left(1+k\frac{H_1'(1,P)}{(u(P))}\right)$, multiplied by the probability the individual is not in the giant component, $(u(P))^k$.

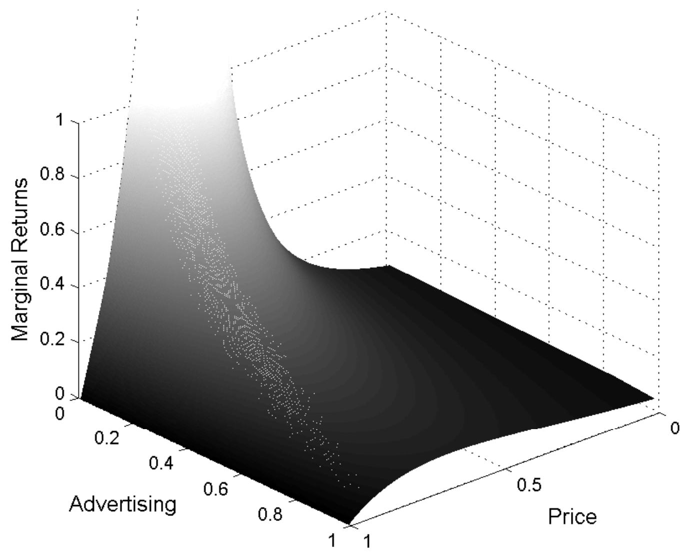

Figure 6: Marginal returns vs Advertising and Price

The optimal target individual is the individual which maximizes the return for a given value of P, this person is:

$$k^* = \arg\max_k (1-P)(u(P))^k \left(1 + k\frac{H_1'(1,P)}{(u(P))}\right)$$

where k is constrained to be an integer.

Theorem 11 *Assuming $p_k > 0$ for all k the highest return type of individual k^* is:*

$$k^* \in \{\lfloor k^{**} \rfloor, \lceil k^{**} \rceil\} \text{ for } P < P^{crit}$$

where

$$k^{**} = \max\left\{0, -\left(\frac{1}{\ln u(P)} + \frac{u(P)}{H_1'(1,P)}\right)\right\}$$

Proof. See appendix ■

This theorem allows one to characterize the optimal target individual for a monopolist charging P. Note that the floor and ceiling functions $(\lfloor\cdot\rfloor, \lceil\cdot\rceil)$ are necessary because k is an integer. The following corollary illustrates how the optimal target, ignoring integer constraints, k^{**} changes as price changes.

Corollary 5 *The optimal target k^{**} is continuous in P for $P < P^{crit}$, $\lim_{P \to P^{crit}} k^{**} = \infty$, $k^{**} \leq \left\lceil \frac{-1}{\ln u(P)} \right\rceil$ for $P < P^{crit}$.*

Proof. See appendix ■

The optimal target individual depends on the price. When there is no giant component at high prices $P > P^{crit}$ $(u(P) = 1)$ then the individuals with the most friends should be targeted. However when the giant component exists $P < P^{crit}$ $(u(P) < 1)$ then individuals with fewer friends should be targeted. The intuition is that as a greater proportion of the population become informed those people with many friends are very likely find out about the good via WOM.

A firm selling an exclusive product, which is sold at a high price such that only a small fraction of the population is prepared to purchase it, should target its marketing at individuals who can pass on information about the product to as many people as possible. On the other hand if the firm is selling a common product, which a larger fraction of the population is prepared to purchase, then the optimal targets for advertising are individuals with few friends. In this case these are the people most likely to be on the "fringe" of the network, or in other words the least well connected parts of the network. This means that they are unlikely to hear about the good via WOM and in expectation will provide the highest return to direct advertising. As the level of advertising increases then the targeted consumer for the next unit of advertising should be a person with fewer friends than the previous consumer. In other words the targeted individual moves towards the fringes of the network. Again the reason for targeting individuals with fewer and fewer friends is that these are the people least likely to be already informed when a greater proportion of the population are already informed.

7.3 Application to social networking sites: Matching consumers to advertising

In this section I assume there is an owner of the rights to advertise to people within the social network. An example of this entity is an online social networking site such as Myspace or Facebook which can sell the rights to advertise to the individuals on their websites. Much of the value from the rights to advertise on these websites is the additional information that the websites have about each consumer. I find that when the website utilizes the information it has about the number of friends each person has, it will optimally allocate advertising for high priced products to individuals with many friends and low priced products to consumers with relatively few.

I assume that there are m different monopolists each selling m different goods. The problem for the owner of the advertising rights is to allocate advertising rights across the m monopolists to maximize the profits from advertising. Each Consumer's valuation for the goods is represented by a vector $\theta_i = \left(\theta_i^1....\theta_i^m\right)$ where each θ_i^j is an independent draw from $U[0,1]$. Thus the demand for each good is independent of the other $m-1$ goods. All monopolists have different prices P^j. Initially a vanishingly small fraction $\varepsilon \to 0$ of the population find out about each good exogenously and WOM is sufficient for the people in the giant component for each good to become informed. Note there are now different giant components for each good. The opportunity to advertise is scarce and only a small fraction δ of individuals may be advertised to.

To simplify the analysis I will assume that the owner of the advertising rights maximizes its own profits by allocating the rights across the monopolists to maximize the total returns to advertising aggregated across all goods, valuing each good equally. The problem is therefore to allocate the δ consumers to the monopolists where they are expected to be in the largest component outside the giant component. I assume that δ is small relative to the giant component thus if multiple individuals are allocated to the same monopolist the probability that they are in the same component outside the giant component is ≈ 0.

7.3.1 Benefit from knowing an individual's number of friends

The return to advertising to a person chosen at random from the population for a monopolist charging P^j is:

$$H_0'\left(1,P^j\right) = \sum_k p_k\left(1-P^j\right)\left(u\left(P^j\right)\right)^k\left[1+k\frac{H_1'\left(1,P^j\right)}{u\left(P^j\right)}\right]$$

where $u\left(P^j\right)$ is the probability that a randomly chosen link does not lead to an individual in the giant component when the price is P^j. This is the weighted sum over all people with k friends where the return from each person is the product of the probability they will purchase the good $\left(1-P^j\right)$, the probability they are not in the giant component $\left(u\left(P^j\right)\right)^k$, and the expected size of the component they belong to $\left[1+k\frac{H_1'\left(1,P^j\right)}{u\left(P^j\right)}\right]$.

When advertising cannot be targeted consumers are optimally shown an advertisement for product j^{**} defined as:

$$j^{**}=\arg\max_j H_0'\left(1,P^j\right)$$

which is the good with the largest component size outside the giant component. If there is a broad range of prices covered by the m products then for the networks presented in Figure 5, (a) and (b) suggest that for the Poisson and Exponential networks j^{**} is going to be the product with the price closest to the critical price. Unless there is a product with a price almost precisely at the critical price for a very heterogeneous networks such as the Hub network in Figure 5 (c) j^{**} will be a product with a price significantly less than the critical price with a price close to 0.7.

When the advertising is targeted at an individual with k friends, the optimal market j^{*k} for the advertising is the market where

$$j^{*k}=\arg\max_j \left(1-P^j\right)\left(u\left(P^j\right)\right)^k\left(1+k\frac{H_1'\left(1,P^j\right)}{\left(u\left(P^j\right)\right)}\right)$$

This is very similar to the targeted marketing case in the previous section, except instead of choosing an individual with k friends for a given P, we are choosing a product with price P^j for a person with k friends. The optimal target follows the intuition of Corollary 5, the optimal product is a common (lower priced) product for individuals with few friends whereas the optimal product for an individual with many friends is a relatively more exclusive (higher priced) product.

The difference between the return to advertising for product j^{*k} and j^{**} is the allocative benefit from knowing the connectivity information of an individual. For a person with k friends this benefit is

$$\begin{aligned}&\left(1-P^{j^{*k}}\right)\left(u\left(P^{j^{*k}}\right)\right)^k\left(1+k\frac{H_1'\left(1,P^{j^{*k}}\right)}{\left(u\left(P^{j^{*k}}\right)\right)}\right)\\-&\left(1-P^{j^{**}}\right)\left(u\left(P^{j^{**}}\right)\right)^k\left(1+k\frac{H_1'\left(1,P^{j^{**}}\right)}{\left(u\left(P^{j^{**}}\right)\right)}\right)\end{aligned}$$

The benefit from knowing a person's connectivity is that it provides an indication of the probability that person is in the giant component and would find out about the good otherwise. For networks such as the Poisson and Exponential networks j^{**} is likely to be a product close to the critical price where $u\left(P^{j^{**}}\right)$ is close to 1.

Theorem 12 *If $P^{j^{**}} > P^{crit}$ then $\exists \widehat{k}$ such that for $k \geq \widehat{k}$ $j^{**} = j^{*k}$.*

Proof. See Appendix ■

When $P^{j^{**}}$ is above the critical price then this benefit is zero for all individuals with a connectivity above a threshold $\widehat{k}$, and all the benefit comes from allocating the rights to advertise to individuals with few friends, to products with a price below the critical price. This is more likely to be the case in the Poisson and Exponential networks compared to the Hub Network shown earlier.

8 Conclusion

Word of mouth is one of the most influential sources of information for consumers when making purchasing decisions. This paper considers informative WOM and how a monopolist can affect the pattern of WOM when the probability an individual engages in WOM is related to her willingness to purchase the product. A key innovation of the paper is to allow the monopolist to strategically determine the probability an individual is willing to engage in WOM. A model of percolation on a random graph with an arbitrary degree distribution is used in the paper and enables me to relate the pricing strategy of the monopolist to the pattern of communication which takes place in the social network. It allows me to study a number of new questions concerning the effect of WOM on demand, pricing and advertising when a firm can affect the pattern of communication which takes place for its own benefit. The setting is very tractable and I am able to introduce correlation between valuations and friendships, price discrimination, regular and targeted advertising and in an application I extend the model to consider how the owner of the rights to advertise on a social network can optimally allocate advertising for specific individuals to different products.

I find a range of interesting results: (i) demand has two distinct regions separated by a critical price related to the first and second moments of the distribution of friendships in the social network; (ii) estimates of consumers valuations are biased downwards and estimates of consumer responses to counterfactual policy/strategy changes are biased upwards if WOM is ignored; (iii) prices are below the fully informed monopoly level for goods where there is no correlation between an individual's valuation of the good and their number of friends, however the opposite may be true if there is significant positive correlation; (iv) introductory prices may have intermittent periods of sales to optimally diffuse news of the good through the population; (v) increasing advertising costs may benefit consumers; (vi) marginal returns to advertising are peaked close to the critical price; and (vii) targeted advertising should be directed towards individuals with many friends for "exclusive" high priced products and towards people with relatively fewer friends for "common" low priced products.

The tractability of the model suggests a number of avenues for future research. One is to incorporate communication structures which include good and bad quality information about the good. This would lead to different inference problems for agents in components of different sizes. In the case of negative WOM the greater connectivity can be a double-edged sword, on one hand it facilitates a greater diffusion of any negative information but on the other, may also permit better statistical inference by aggregating information in larger components. An aspect of the percolation process explored in sections 5.3 and 7.2 is the targeting of strategies at individuals depending on their degree. The resilience of a network to the targeted removal of individuals has been studied in the context of immunization and computer networks and may offer further insights in economic applications. More broadly percolation processes can provide a great deal of information about the structure and pattern of communication that takes place through the distribution of component sizes and how this changes in response to endogenously chosen variables. It is an open question of the applications in economics where this information is important however this paper highlights its application to the pricing and advertising strategies of a monopolist facing a population which engages in WOM about its good.

References

[1] Ballester, C., Calvo-Armengol, A, and Zenou, Y., "Who's Who in Networks. Wanted: The Key Player," *Econometrica*, 2006, 74, 4, pp1403-1418.

[2] Banerji, A. and Dutta, B., "Local Network Externalities and Market Segmentation," 2005, Working paper.

[3] Berge, C., *Topological Spaces*, Oliver and Boyd Ltd., 1963.

[4] Callaway, D., Newman, M., Strogatz, S. and Watts, D., "Network Robustness and Fragility: Percolation on Random Graphs," *Physical Review Letters*, 2000, Vol. 85 No. 25, pp5468-5471.

[5] Ellison, G., "Learning, Local interaction, and Coordination," *Econometrica*, 1993, Vol. 61, pp1047-1071.

[6] Erdös, P., and Renyí, A, "On Random Graphs," *Publicationes Mathematicae*, 1959, Vol. 6, pp290-297.

[7] Erdös, P., and Renyí, A, "On the Evolution of Random Graphs," *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, 1960, Vol. 5, pp17-61.

[8] Erdös, P., and Renyí, A, "On the Strength of Connectedness of a Random Graph," *Acta Mathematica Scienta Hungary*, 1961, Vol. 12, pp261-267.

[9] Feick, L. and Price, L., "The Market Maven: A Diffuser of Marketplace Information," *Journal of Marketing*, 1987, Vol. 51(1), pp83-97.

[10] Galleatti, A. and Mattozzi, A., "Personal Influence: Social Context and Political Competition," 2008, Working Paper.

[11] Gladwell, M., *The tipping point: How little things can make a big difference,* Little, Brown and Company (2000).

[12] Godes, D., and Mayzlin, D., "Using Online Conversations to Study Word of Mouth Communication," *Marketing Science*, Vol. 23(4), pp545-560.

[13] Goyal, S., and Galleotti, A., "A Theory of Strategic Diffusion," Working paper, 2007.

[14] Jackson, M., *Social and Economics Networks,* Princeton University Press, 2008.

[15] Jackson, M., and Yariv, L., "Diffusion of Behavior and Equilibrium Properties in Network Games," *American Economic Review (Papers and Proceedings)*, 2007, 97, 2, pp92-98.

[16] Mobius, M., Niehaus, P., and Rosenblat, T., "Social Learning and Consumer Demand," Working Paper, 2006.

[17] Moore, C. and Newman, M., "Epidemics and Percolation in Small-World Networks," *Physical Review E*, 2000, Vol.61 No.5, pp5678-5682.

[18] Morris, S., "Contagion," *Review of Economic Studies*, 2000, 67, pp57-78.

[19] Newman, M., "The Spread of Epidemic Diseases on Networks," *Physical Reveiw E*, 2002, Volume 66(1), art. no. 016128.

[20] Newman, M., "Component Sizes in Networks with Arbitrary Degree Distributions," *Physical Review E,* 2007, Vol. 76, art. no. 045101.

[21] Newman, M., Strogatz, S., and Watts, D., "Random Graphs with Arbitrary Degree Distributions and their Applications," *Physical Review E*, 2001, Volume 64, art. no. 026118.

[22] Newman, M., and Watts, D., "Scaling and Percolation in the Small-World Network Model," *Physical Review E*, 1999, Vol. 60 No. 6, pp7332-7342.

[23] Park, J., and Newman, M., "Solution for the Properties of a Clustered Network," *Physical Review E*, 2005, Vol. 72 Art No. 026136.

[24] Reichheld, F., "The One Number You Need to Grow," *Harvard Business Review*, 2003, Vol. 81 No. 12, pp47-54.

[25] Sander, L., Warren, C., Sokolov, I., Simon, C., and Koopman, J., "Percolation on Heterogeneous Networks as a Model for Epidemics," *Mathematical Biosciences,* 2002, Volume 180, Issue 1-2, pp293-305.

[26] Watts, D., "A Simple Model of Cascades on Random Networks," *Proceedings of the National Academy of Sciences of the United States of America*, 2002, 99, 9, pp5766-5771.

[27] Watts, D., Perretti, J., and Fumin, M., "Viral Marketing for the Real World," 2007, Working paper.

A Appendix

A.1 Proof of Theorem 1

Suppose an individual's valuation is independent of the number of friends, $q_k = q = 1 - P$ *for all* k. *Then, there exists a critical price* P^{crit} *such that*

$$\lim_{\varepsilon\to 0}\lim_{n\to\infty} E_{\Omega_{N,\{p_k\}}}[\bar{s}] = O(n) \text{ if } P < P^{crit}$$

$$\lim_{\varepsilon\to 0}\lim_{n\to\infty} E_{\Omega_{N,\{p_k\}}}[\bar{s}] = O\left(n^{2/3}\right) \text{ if } P = P^{crit}$$

and

$$\lim_{\varepsilon\to 0}\lim_{n\to\infty} E_{\Omega_{N,\{p_k\}}}[\bar{s}] = O(\log n) \text{ if } P > P^{crit}$$

Moreover the critical price satisfies $1 - P^{crit} = \frac{E[k]}{E[k^2]-E[k]}$

Proof. Molloy and Reed (1995) show that the critical percolation threshold is $q_c = \frac{\sum_k p_k k}{\sum_k p_k k(k-1)}$. The result follows immediately by substituting $1 - P^{crit} = q_c$:

$$1 - P^{crit} = \frac{E[k]}{E[k^2] - E[k]}$$

■

A.2 Proof of Theorem 2

Suppose θ *and* k *are uncorrelated then demand for the good* $S(P)$ *is*

1. *Continuous*
2. $S(P) = 0$ for $P \geq P^{crit}$
$S(P) > 0$ for $P < P^{crit}$
3. $\frac{dS}{dP} = 0$ for $P \geq P^{crit}$
$\frac{dS}{dP} < 0$ for $P < P^{crit}$
4. $\lim_{P\to P^{crit}-} \frac{dS}{dP} = -\left(1 - P^{crit}\right)\frac{G_0'''(1)}{\left(G_0''(1)\right)^2} < 0$
5. $\left|\frac{P}{S}\frac{dS}{dP}\right| > \left|\frac{P}{1-P}\right|$ for $P < P^{crit}$

Proof. Demand is given by

$$S = F_0(1) - F_0(u)$$

where u is the smallest non-negative solution to the self consistency condition:

$$u = 1 - F_1(1) + F_1(u) \tag{8}$$

$\blacksquare$

The following lemma illustrates some properties of u with respect to the price which I will subsequently use to prove the above theorem.

Lemma 4 *Suppose $u(P)$ is given by equation 8 then*

1. $u(P) = 1$ and $\frac{du}{dP} = 0$ for $P^{crit} \leq P \leq 1$
2. $u < 1$ and $\frac{du}{dP} > 0$ for $0 \leq P < P^{crit}$
3. $u(P)$ is continuous in P

Proof. $u(P)$ is the smallest non-negative solution to:

$$u = P + (1-P)\frac{\sum_k k p_k u^{k-1}}{z}$$

Now consider the function $f(u) = P + (1-P)\frac{\sum_k k p_k u^{k-1}}{z}$ first note $f(1) = 1$ and so $u = 1$ always satisfies the above relationship, second $f(u)$ is a polynomial in u with positive coefficients so it is continuous, increasing and convex in the region $0 \leq u \leq 1$ and thus combined with $f(0) = P$ there is at most one other solution $0 \leq u < 1$.

When $f'(1) \leq 1$ there is no solution for $0 \leq u \leq 1$ and $u = 1$ is the only solution. When $f'(1) > 1$ there is a solution for $0 \leq u < 1$. The condition $f'(1) \leq 1$ is equivalent to $P > P^{crit}$:

$$f'(1) = (1-P)\frac{\sum_k k(k-1)p_k}{z} \leq 1$$

$$\begin{aligned} 1 - \frac{z}{\sum_k k(k-1)p_k} &\leq P \\ P^{crit} &\leq P \end{aligned}$$

Therefore $u = 1$ for $P \geq P^{crit}$ and $0 \leq u < 1$ for $P < P^{crit}$. $u = 1$ for $P \geq P^{crit}$ immediately implies $\frac{du}{dP} = 0$ for $1 \geq P \geq P^{crit}$.
To show that $\frac{du}{dP} > 0$ I look at the derivative for $\frac{du}{dP}$:

$$\frac{du}{dP} = \frac{(1-G_1(u))^2}{1 - G_1(u) - (1-u)G_1'(u)}$$

The numerator is positive for $u < 1$ and the denominator $1 - G_1(u) - (1-u)G_1'(u)$ is continuous and equal to 1 at $u = 0$, equal to 0 at $u = 1$ and is decreasing in u for $0 \leq u \leq 1$ provided $G_1''(1) > 0$ which is a necessary condition for $P^{crit} > 0$. Therefore in the range $P \in [0, P^{crit})$ $u(P)$ is continuous and $\frac{du}{dP} > 0$. $\blacksquare$

Returning to the theorem. Using this lemma I conclude that for $P \geq P^{crit}$ $S(P) = 0$ and for $P \in [0, P^{crit})$ $S(P) = (1-P)\left(1 - \sum_k p_k u^k\right)$ is a continuous function since u is continuous in P. I now prove the continuity of $S(P)$ by showing that as the price approaches the critical price from below $S \to 0$:

If there exists a critical price, $0 < P^{crit} < 1$ *then as price approaches the critical price from below* $\lim_{P \to P^{crit}-} S = 0$

Proof. I rewrite the relationship between P and u

$$P(u) = \frac{u - G_1(u)}{1 - G_1(u)}$$

such that $P(u)$ is a continuous, monotonically increasing (one to one) function $[0,1) \to [-1,1]$. I will now show that $\lim_{u \to 1^-} P(u) = P^{crit}$.

$P(1) = \frac{0}{0}$ so applying L'Hopital's rule

$$\begin{aligned}
\lim_{u \to 1^-} P(u) &= \lim_{u \to 1^-} P'(u) \\
&= \frac{1 - G_1'(1)}{G_1'(1)} \\
&= 1 - \frac{E[k]}{E[k^2] - E[k]} \\
&= P^{crit}
\end{aligned}$$

Now $P(u)$ is a one to one function and $0 < P^{crit} < 1$ this implies that $\lim_{P \to P^{crit}-} u = 1$ and hence $\lim_{P \to P^{crit}-} S = 0$ ∎

This completes the argument for the continuity of S. So far we have shown $P \in \left[P^{crit}, 1\right]$ $S(P) = 0$, for $P \in [0, P^{crit})$ $S(P)$ is continuous, and finally $\lim_{P \to P^{crit}-} S = 0$. The next part of the theorem is:

If there exists a critical price, $0 < P^{crit} < 1$ *then as price approaches the critical price from below* $\lim_{P \to P^{crit}-} \frac{dS}{dP} = -\left(1 - P^{crit}\right) \frac{G_0'''(1)}{\left(G_0''(1)\right)^2}$

Proof.

$$S = (1 - P)\left[1 - \sum_k p_k u^k\right]$$

$$\lim_{u \to 1^-} \frac{dS}{dP} = \lim_{u \to 1^-} -\left[1 - \sum_k p_k u^k\right] - \frac{du}{dP}(1 - P)\left(-\sum_k k p_k u^{k-1}\right)$$

$$\left.\frac{dS}{dP}\right|_{u=1} = z\left(1 - P^{Crit}\right) \lim_{u \to 1^-} \left.\frac{du}{dP}\right|$$

$$\lim_{u \to 1^-} \left.\frac{du}{dP}\right| = \lim_{u \to 1^-} \left.\frac{1 - G_1(u) - (1 - u) G_1'(u)}{\left(1 - G_1(u)\right)^2}\right| = \frac{0}{0}$$

using L'Hopitals rule,

$$
\begin{aligned}
& \lim_{u \to 1^-} \frac{1 - G_1(u) - (1-u) G_1'(u)}{(1 - G_1(u))^2} \Bigg| \\
= & \lim_{u \to 1^-} \frac{(1-u) G_1''(u)}{2G_1'(u)(1 - G_1(u))} \Bigg| = \frac{0}{0}
\end{aligned}
$$

and again

$$
\begin{aligned}
= & \lim_{u \to 1^-} \frac{G_1'''(u)(1-u) - G_1''(u)}{2G_1''(u)(1 - G_1(u)) - 2(G_1'(u))^2} \\
= & \frac{G_1''(1)}{2G_1'(1)^2}
\end{aligned}
$$

Furthermore provided that $G_1''(1)$ is non zero (which also implies $G_0'(1)$ is non zero) then the demand curve will exhibit a non-zero slope ($\frac{dS}{dP} < 0$) as the price approaches the critical price from below. $G_1''(1) > 0$ also implies that there are some people with 3 or more friends, which is also necessary for $P^{crit} > 0$ so that for any network where $P^{crit} > 0$ then demand will exhibit a kink at P^{crit} separating the two regions of demand. ∎

At $P < P^{crit}$ $\frac{dS}{dP} < 0$

Proof. Consider the expression for $\frac{P}{S}\frac{dS}{dP}$:

$$
\frac{P}{S}\frac{dS}{dP} = \frac{-P}{1-P}\left[1 + \frac{(1-P)}{1 - \sum\limits_k p_k u^k}\frac{du}{dP}\sum_k p_k k u^{k-1}\right]
$$

the result follows immediately from $u < 1$ and $\frac{du}{dP} > 0$ for $P < P^{crit}$. ∎

The final element of the proof is

For $P < P^{crit}$ $\left|\frac{P}{S}\frac{dS}{dP}\right| > \left|\frac{P}{1-P}\right|$

Proof. From above

$$
\frac{P}{S}\frac{dS}{dP} = \frac{-P}{1-P}\left[1 + \frac{(1-P)}{1 - \sum\limits_k p_k u^k}\frac{du}{dP}\sum_k p_k k u^{k-1}\right]
$$

where the second term inside the brackets is strictly positive from lemma 4 and the result follows immediately. ∎

A.3 Proof of Corollary 1

Suppose the price of the good is $\tilde{P}$ *then an estimate of consumer surplus* $CS\left(\tilde{P}\right) = \int_{\tilde{P}}^{\infty} S(P)\,dP$ *is biased downwards*

Proof. I show that the estimate of the distribution of valuations implied by $S(P)$ is first order stochastically dominated by the actual distribution of valuations of the consumers purchasing the product. Denote the actual distribution of valuations for the consumers who purchase the good by $G(\theta)$ and the estimate by $\tilde{G}(\theta)$. Preferences are distributed uniformly across informed consumers when θ and k are uncorrelated thus the actual distribution of valuations is

$$\begin{aligned} G(\theta) &= \frac{\theta - \tilde{P}}{1 - \tilde{P}} \text{ for } \tilde{P} \leq \theta \leq 1 \\ &= 0 \text{ for } \theta < \tilde{P} \end{aligned}$$

The estimate $\tilde{G}(\theta)$ from $S(P)$ is

$$\begin{aligned} \tilde{G}(\theta) &= 1 - \frac{S(\theta)}{S\left(\tilde{P}\right)} \text{ for } \tilde{P} \leq \theta \leq 1 \\ &= 0 \text{ for } \theta < \tilde{P} \end{aligned}$$

For any $\theta \in \left[\tilde{P}, 1\right]$

$$\tilde{G}(\theta) - G(\theta) = 1 - \frac{S(\theta)}{S\left(\tilde{P}\right)} - \frac{\theta - \tilde{P}}{1 - \tilde{P}}$$

substituting in for $S\left(\tilde{P}\right), S(\theta)$ and rearranging

$$\frac{1-\theta}{S\left(\tilde{P}\right)} \sum p_k \left(u(\theta)^k - u\left(\tilde{P}\right)^k\right) > 0 \text{ for } \theta > \tilde{P}$$

because $\theta \geq \tilde{P}$ and $u(\cdot)$ is an increasing function. First Order Stochastic Dominance implies that estimates of consumer welfare using the distribution of valuations implied by $S(P)$ are going to be too small. ■

A.4 Proof of Corollary 2

Suppose the price of the good is $\tilde{P}$ then an estimate of the consumer response $\Delta\hat{S} = S\left(\tilde{P}\right) - S\left(\tilde{P} + \Delta P\right)$ to an increase in the price by ΔP overstates the actual response ΔS

$$\Delta\hat{S} < \Delta S$$

Proof. Denote the actual distribution of valuations for the consumers who purchase the good at $\tilde{P}$ by $G(\theta)$ and the estimate by $\tilde{G}(\theta)$. Then $\Delta S = -G\left(\tilde{P} + \Delta P\right)$ $\Delta\hat{S} = -\tilde{G}\left(\tilde{P} + \Delta P\right)$. The result follows immediately from Corollary 1 where $\tilde{G}(\theta) - G(\theta) \geq 0$ for any $\theta \geq \tilde{P}$. ■

A.5 Proof of Theorem 3

Suppose valuations and number of friends are uncorrelated and marginal costs $c < 1$, then a monopolist facing demand given by $S(P)$ charges a lower price P^*_{WOM} than a monopolist facing a fully informed population P^*_{FI}, where demand is given by $Q(P) = 1 - P$.

Proof. Define the fully informed monopoly price as P^*_{FI} and the WOM monopoly price as P^*_{WOM}. A monopolist facing a fully informed population has a strictly concave profit maximization problem and charges the unique monopoly price $P^*_{FI} = \frac{1+c}{2}$ provided $c < 1$. If $c \geq 1$ then there is clearly no price where the monopolist can make positive profits. It is also true that

$$\frac{P-c}{P} \geq \frac{1}{\varepsilon_{FI}} \text{ for any } P \geq P^*_{FI}$$

it was shown in Theorem 2 that $|\varepsilon_{WOM}| > |\varepsilon_{FI}|$ which implies that:

$$\frac{P-c}{P} > \frac{1}{\varepsilon_{WOM}} \text{ for any } P \geq P^*_{FI}$$

when demand is positive in the range of prices $P^{crit} > P \geq P^*_{FI}$. The WOM monopolists profit function $(P-c)\,S(P)$ is continuous and differentiable for $P < P^{crit}$. Therefore the first order conditions for the monopolist are necessary and hence $\frac{P-c}{P} > \frac{1}{\varepsilon_{WOM}}$ for all $P \geq P^*_{FI}$ implies $P^{Mon} \ngeq P^*_{FI}$. ■

A.6 Proof of Corollary 3

Consumer surplus may be greater when consumers are uninformed and the monopolist charges P^*_{WOM} than if consumers are fully informed and the monopolist charges P^*_{FI}

Proof. Consider the social networks where everyone has 3 friends $G_0(x) = x^3$. If the marginal cost of the monopolist is 0 then the profit maximizing price when the population is fully informed is 0.5 and consumer surplus is 0.125. On the other hand if the population is uninformed the WOM monopoly price $P^*_{WOM} = 0.3215$ and consumer surplus is 0.2057. If marginal cost is higher, > 0.5, then there is no price above the monopolist's marginal cost where the giant component exists thus consumer surplus is 0. ■

A.7 Proof of Theorem 4

If all consumers with $\theta \in [c, \underline{\theta}]$ have $k = 1$ where $\underline{\theta} > \frac{1+c}{2}$ then provided the giant component exists at $P = \underline{\theta}$ the monopoly price will be greater than the fully informed monopoly price $\frac{1+c}{2}$

Proof. I first show that demand will be linear in the region $P \in [c, \underline{\theta}]$. Consider

$$\begin{aligned} S &= 1 - H_0(1, P) \\ \frac{dS}{dP} &= -\frac{dH_0(1,P)}{dP} = -\frac{d\left(1 - \sum p_k q_k \left(1 - u^k\right)\right)}{dP} \\ &= -(1-u) + \frac{du}{dP} \sum k p_k q_k u^{k-1} \end{aligned}$$

In the range of prices $P \in [c, \underline{\theta}]$, $\frac{dq_k}{dP} = 0$ for $k \neq 1$ and $\frac{dq_1}{dP} = -\frac{1}{p_1}$ for $k = 1$ because all consumers $\theta \in [c, \underline{\theta}]$ have $k = 1$. Now consider the self consistency relationship for $u(P)$:

$$u = 1 - \frac{1}{z_1} \sum_{k=2}^{\infty} k p_k q_k \left(1 - u^{k-1}\right)$$

This is independent of q_1, thus for $P \in [c, \underline{\theta}]$ $u(P)$ is constant, $\frac{dS}{dP} = -(1-u)$ and S is linear. Denote $\underline{u} = u(P)$ for $P \in [c, \underline{\theta}]$.

Consider the first order condition of the monopolist in the range $P \in [c, \underline{\theta}]$

$$\frac{d\pi}{dP} = S - (P - c)(1 - \underline{u})$$

this is decreasing in P and positive if $\frac{S(P)}{1-\underline{u}} > P - c$. Therefore the optimal price cannot be less than or equal to $\frac{1+c}{2}$ if $\frac{S\left(\frac{1+c}{2}\right)}{1-\underline{u}} > \frac{1-c}{2}$ which is equivalent to

$$\underline{\theta} + \frac{S(\underline{\theta})}{1 - \underline{u}} > 1$$

provided $P^{crit} > \theta$ and hence $\underline{u} < 1$, this can be rewritten

$$\begin{aligned} \sum p_k q_k \left(1 - \underline{u}^k\right) - (1 - \underline{\theta})(1 - \underline{u}) &> 0 \\ \sum p_k q_k \left(\underline{u} - \underline{u}^k\right) &> 0 \end{aligned}$$

which is true for $\underline{u} < 1$ hence the monopoly price is greater than $\frac{1+c}{2}$. ■

A.8 Proof of Theorem 5

If valuations and number of friends are uncorrelated then the optimal set of prices $P_0 = P_1 = \frac{1+c}{2}$ and $\exists \underline{k} : \{P_k\}$ is decreasing for $2 \leq k \leq \underline{k}$ and $P_k = 0$ for $k \geq \underline{k}$

Proof. Monopolist's maximization

$$\pi = \max_{\{P_k\}} \sum p_k (1 - P_k)\left(1 - u^k\right)(P_k - c)$$

Assuming $P_k > 0$ for all k. First order condition for price P_k :

$$p_k(1-P_k)\left(1-u^k\right)-p_k\left(1-u^k\right)(P_k-c)-\frac{\partial u}{\partial P_k}\sum_j p_j(1-P_j)(P_j-c)\,ju^{j-1}=0 \text{ for } P_k\in(0,1)$$

Probability that a randomly chosen link is outside the giant component:

$$D(u)=u-1+\frac{\sum kp_k(1-P_k)\left(1-u^{k-1}\right)}{z}=0$$

Implicit function theorem

$$\begin{aligned}\frac{\partial D}{\partial u} &= u-\frac{\sum k(k-1)p_k(1-P_k)u^{k-2}}{z}\\ \frac{\partial D}{\partial P_k} &= -\frac{kp_k\left(1-u^{k-1}\right)}{z}\end{aligned}$$

$$\frac{du}{dP_k}=\frac{kp_k\left(1-u^{k-1}\right)}{zu-\sum k(k-1)p_k(1-P_k)u^{k-2}}$$

where $\frac{du}{dP_1}=0$ so $P_1=\frac{1+c}{2}$. Now defining $\alpha=\frac{\sum_j p_j(1-P_j)(P_k-c)ju^{j-1}}{zu-\sum k(k-1)p_k(1-P_k)u^{k-2}}$ which is the same for all k and going back to the first order condition for P_k

$$\begin{aligned}p_kq_k\left(1-u^k\right)-p_k\left(1-u^k\right)(P_k-c)-\alpha kp_k\left(1-u^{k-1}\right) &= 0 \text{ for } P_k\in(0,1)\\ 1-2P_k+c-\alpha k\frac{\left(1-u^{k-1}\right)}{1-u^k} &= 0 \text{ for } P_k\in(0,1)\end{aligned}$$

$$\frac{d\left(k\frac{\left(1-u^{k-1}\right)}{1-u^k}\right)}{dk}>0$$

and thus P_k is decreasing in k. If $1+c-\alpha k\frac{\left(1-u^{k-1}\right)}{1-u^k}<0$ then $P_k=0$ so defining $\underline{k}=\inf\left\{k|1+c-\alpha k\frac{\left(1-u^{k-1}\right)}{1-u^k}<0\right\}$ then for all $k\geq\underline{k}$ $P_k=0$. ■

A.9 Proof of Lemma 1

If $F\in\mathcal{F}$ *then* $\Gamma_F(F,P)\in\mathcal{F}$

Proof. Let $F_t=F$ and $F_{t+1}=\Gamma_F(F,P)$. Consider the value of f_{t+1} for $\theta<P_t$ as a

function of P_t and f_t, this may be written as:

$$\begin{aligned} f_{t+1}(\theta) &= \frac{f_t(\theta) + \frac{z_2}{z_1}(1 - F_t(P_t))}{1 + \left(\frac{z_2}{z_1} - 1\right)(1 - F_t(P_t))} \\ &= \frac{z_2}{z_2 - z_1} \frac{\frac{f_t(\theta)}{\frac{z_2}{z_2 - z_1}} + \left(\frac{z_2}{z_1} - 1\right)(1 - F_t(P_t))}{1 + \left(\frac{z_2}{z_1} - 1\right)(1 - F_t(P_t))} \end{aligned}$$

Hence if $f_t(\theta) < \frac{z_2}{z_2 - z_1}$ then the second term is < 1 and $f_{t+1}(\theta) < \frac{z_2}{z_2 - z_1}$. For $\theta \geq P_t$ $f_{t+1}(\theta) < 1 < \frac{z_2}{z_2 - z_1}$. There are no mass points in F_t so the cdf $\Gamma_F(F_t, P)$ is also continuous. Thus $\Gamma_F(F, P) \in \mathcal{F}$. ■

A.10 Proof of Lemma 2

$\Gamma_M : \mathcal{F} \times [0,1] \times [1, \infty) \rightarrow [1, \infty)$ *and* $\Gamma_F : \mathcal{F} \times [0,1] \rightarrow \mathcal{F}$ *are continuous mappings*

Proof. Use the sup norm on the space of continuos cdfs on $[0, 1]$. Γ_M and Γ_F are single valued mappings so I will proceed with an ε η proof of continuity. That is for a give $\varepsilon > 0$ there exists $\eta > 0$ such that if $|(F_0, P_0), (F, P)| < \eta$ then $|\Gamma_F(F_0, P_0), \Gamma_F(F, P)| < \varepsilon$ in the case of Γ_F and similarly in the case of Γ_M.

First I prove the continuity of Γ_F. For any $F_0 \in \mathcal{F}$ and $P_0 \in [0, 1]$

$$F'(\theta) = \Gamma_F(F_0, P_0) = \frac{\min[F_0(\theta), F_0(P_0)] + \frac{z_2}{z_1}(1 - F_0(P_0))\theta}{1 + \left(\frac{z_2}{z_1} - 1\right)(1 - F_0(P_0))}$$

For any ε choose $\eta = \frac{1}{2}\sqrt{\frac{\varepsilon}{(\alpha+1)^2 \frac{z_2}{z_1}\left(\frac{z_2}{z_1}+1\right)}}$ where $\alpha = \frac{z_2}{z_2 - z_1}$.

For any (F, P) where $\|(F_0, P_0), (F\ P)\| < \eta$ we have $|F(\theta) - F_0(\theta)| < \eta$ and $|P - P_0| < \eta$. Hence

$$\begin{aligned} \left|\frac{z_2}{z_1}(1 - F_0(P_0))\theta - \frac{z_2}{z_1}(1 - F(P))\theta\right| &= \frac{z_2}{z_1}\theta\,|F_0(P_0) - F(P)| \\ &< \frac{z_2}{z_1}\theta\left(|F_0(P_0) - F_0(P)| + |F_0(P) - F(P)|\right) \\ &< \eta\frac{z_2}{z_1}\theta(\alpha + 1) \end{aligned}$$

,

$$\begin{aligned} \left|\frac{1}{1 + \left(\frac{z_2}{z_1} - 1\right)(1 - F_0(P_0))} - \frac{1}{1 + \left(\frac{z_2}{z_1} - 1\right)(1 - F(P))}\right| &< \frac{\frac{z_2}{z_1}\theta\,|F_0(P_0) - F(P)|}{1 + \left(\frac{z_2}{z_1} - 1\right)(1 - \min[F_0(P_0), F(P)])} \\ &< \eta\frac{z_2}{z_1}\theta(\alpha + 1) \end{aligned}$$

$$\left|\min\left[F_0(\theta), F_0(P_0)\right] - \min\left[F(\theta), F(P)\right]\right|$$

wlog say $P_0 \geq P$ now if $\theta < P$ then

$$\begin{aligned}\left|\min\left[F_0(\theta), F_0(P_0)\right] - \min\left[F(\theta), F(P)\right]\right| &= \left|F_0(\theta) - F(\theta)\right| \\ &< \eta\end{aligned}$$

if $\theta > P_0$

$$\begin{aligned}\left|\min\left[F_0(\theta), F_0(P_0)\right] - \min\left[F(\theta), F(P)\right]\right| &= \left|F_0(P_0) - F(P)\right| \\ &< \eta(\alpha+1)\end{aligned}$$

if $P \leq \theta \leq P_0$

$$\begin{aligned}\left|\min\left[F_0(\theta), F_0(P_0)\right] - \min\left[F(\theta), F(P)\right]\right| &= \left|F_0(\theta) - F(P)\right| \\ &< \left|F_0(P_0) - F(P)\right| \\ &< \eta(\alpha+1)\end{aligned}$$

hence

$$\left|\min\left[F_0(\theta), F_0(P_0)\right] - \min\left[F(\theta), F(P)\right]\right| < \eta(\alpha+1)$$

Now

$$\begin{aligned}\left|\Gamma_F(F_0, P_0) - \Gamma_F(F, P)\right| &< \eta\frac{z_2}{z_1}\theta(\alpha+1)\left(\eta(\alpha+1) + \eta\frac{z_2}{z_1}\theta(\alpha+1)\right) \\ &< \eta^2(\alpha+1)^2\frac{z_2}{z_1}\left(\frac{z_2}{z_1}+1\right)\end{aligned}$$

And therefore

$$\left|\Gamma_F(F_0, P_0) - \Gamma_F(F, P)\right| < \frac{\varepsilon}{2}$$

and $\Gamma_F(F, P)$ is a continuous mapping.

For $M' = \Gamma_M(M_0, F_0, P_0) = \left((1 - F_0(P_0))\frac{z_2}{z_1} + F_0(P_0)\right)M_0$. For any ε choose $\eta = \frac{\varepsilon/2}{M_0\left(\frac{z_2}{z_1}+1\right)(\alpha+1)+\left(\frac{z_2}{z_1}+1\right)}$. Any (M, F, P) for where:

$$\left\|(M_0, F_0, P_0), (M, F, P)\right\| < \eta$$

$$\begin{aligned}&\Rightarrow \left|M_0 - M\right| < \eta \\ &\Rightarrow \left|F(\theta) - F_0(\theta)\right| < \eta \\ &\Rightarrow \left|P - P_0\right| < \eta\end{aligned}$$

and from earlier

$$|F_0(P_0) - F(P)| < \eta(\alpha+1)$$

Now

$$\begin{aligned}
|\Gamma_M(M_0,F_0,P_0) - \Gamma_M(M,F,P)| &< |\Gamma_M(M_0,F_0,P_0) - \Gamma_M(M_0,F,P)| \\
&\quad + |\Gamma_M(M_0,F,P) - \Gamma_M(M,F,P)| \\
&< M_0\left(\frac{z_2}{z_1}+1\right)\eta(\alpha+1) + \left(\frac{z_2}{z_1}+1\right)\eta \\
&< \frac{\varepsilon}{2}
\end{aligned}$$

■

A.11 Proof of Lemma 3

If $P_t = P^* < P^{crit}$ for all t and $F_t \in \mathcal{F}$ then the limiting distribution $f_t^*(\theta) = \lim_{t\to\infty} f_t(\theta)$ will be

$$\begin{aligned}
f_t^*(\theta) &= \frac{z_2}{z_2 - z_1} \text{ if } \theta < P \\
&= \frac{z_2 - \frac{z_1}{1-P}}{z_2 - z_1} \text{ if } \theta \geq P
\end{aligned}$$

Proof. When P_t remains constant each period $1 - F_t(P)$ fraction of people purchase and inform $\frac{z_2}{z_1}$ others. For $\theta < P$ we have the following expression for $F_t(\theta)$:

$$F_t(\theta) - F_{t-1}(\theta) = \frac{\frac{z_2}{z_1}(1 - F_{t-1}(P_{t-1}))\theta - F_{t-1}(\theta)\left(\frac{z_2}{z_1}-1\right)(1 - F_{t-1}(P_{t-1}))}{1 + \left(\frac{z_2}{z_1}-1\right)(1 - F_{t-1}(P_{t-1}))}$$

$F_t(\theta) < \frac{z_2}{z_2 - z_1}\theta$ and $F_t(\theta) > F_{t-1}(\theta)$ when $\frac{z_2}{z_2-z_1}\theta > F_{t-1}(\theta)$ and $F_t(\theta) - F_{t-1}(\theta) \to 0$ as $F_t(\theta) \to \frac{z_2}{z_2-z_1}\theta$. Thus $\lim_{t\to\infty} F_t(\theta) = \frac{z_2}{z_2-z_1}\theta$ for $\theta < P$

For $\theta \geq P$ the only people with $\theta \geq P$ are those that have been newly informed from the period before, so the distribution is uniform for $\theta \geq P$ hence $F_t(\theta)$ can be written as $1 - \alpha_t(1-\theta)$. Substituting this into the transition function Γ_F :

$$\begin{aligned}
1 - \alpha_t(1-\theta) &= 1 - \frac{\frac{z_2}{z_1}(\alpha_{t-1}(1-P))(1-\theta)}{1 + \left(\frac{z_2}{z_1}-1\right)(\alpha_{t-1}(1-P))} \quad \text{for } \theta \geq P \\
\alpha_t &= \frac{\frac{z_2}{z_1}(\alpha_{t-1}(1-P))}{1 + \left(\frac{z_2}{z_1}-1\right)(\alpha_{t-1}(1-P))}
\end{aligned}$$

Hence $\frac{z_2 - \frac{z_1}{1-P}}{z_2 - z_1} < a_t < \alpha_{t-1}$ for any $\alpha_{t-1} > \frac{z_2 - \frac{z_1}{1-P}}{z_2 - z_1}$. Thus $\lim_{t\to\infty} f_t(\theta) = \frac{z_2 - \frac{z_1}{1-P}}{z_2 - z_1}$. ■

A.12 Proof of Theorem 6

The monopolist's problem has a unique solution, the value function is homogeneous of degree 1 in M and the policy function $P(F)$ is u.h.c and only a function of the state F.

Proof. The proof involves defining a contraction mapping on the recursive problem and using this to show that there is a unique solution to it. The continuity of the value function and u.h.c of the policy function come from the theorem of the maximum.

I first prove the homogeneity of the problem ■

Lemma 5 *$J(\cdot, F)$ is homogeneous of degree one in its first argument*

Proof. Note that the state variable M does not appear in the transition equation Γ_F thus for a given sequence of prices the states F_t will be unaffected by changing M_0 to λM_0. Also note that $\frac{M_{t+1}}{M_t} = \left(\left(1 - F(P)\frac{z_2}{z_1} + F(P)\right)\right)$ is also unchanged. The objective function can therefore be rewritten

$$J(M_0, F_0) = M_0 \times \max_{\{P_t\}} \sum_{t=0}^{\infty} \beta^{t-1} P_t (1 - P_t) \left(\Pi_{i=0}^{t} \frac{M_i}{M_{i-1}}\right)$$

Thus $J(\lambda M_0, F_0) = \lambda J(M_0, F_0)$ ■

Now define the set of continuous cdfs on $[0, 1]$ which satisfy

$$\frac{F(x) - F(x - \delta)}{\delta} \leq \alpha$$

for some finite $\alpha > 0$ by $\mathcal{F}$. From Lemma 1any cdf $\Gamma_F(F, P)$ satisfies this property provided F does. Also note the space $\mathcal{F}$ with the sup norm is complete.

Let $H(M, F)$ be the space of functions $V : [1, \infty) \times \mathcal{F} \to R$ which are continuous, homogeneous of degree one with respect to their first argument and bounded in the norm $\max_{F \in \mathcal{F}} \frac{V(M,F)}{M}$. Define an operator T on $H(M, F)$ by

$$(TV)(M, F) = \max_{\substack{P \in [0,1] \\ M' = \Gamma_M(M,F,P) \\ F' = \Gamma_F(F,P)}} P(1 - P) M + \beta V(M', F')$$

where $F \in \mathcal{F}$ and $M \in [1, \infty)$. Note that the objective and transition functions are continuous and the maximization is over a compact set so the maximum is achieved and by the theorem of the maximum (Berge 1963) TV is also continuous. Also note that M' is a linear function of M so TV will be homogenous of degree 1 in M. Thus TV maps $H(M, F) \to H(M, F)$.

Define the function $(V + a)(M, F) = V(M, F) + aM$

Lemma 6 *Let $(M, F) \subseteq [1, \infty) \times \mathcal{F}$ and let $H(M, F)$ be as above, with the associated norm. Let $T : H(M, F) \to H(M, F)$ satisfy* (monotonicity) $V, W \in H$ and $V \leq W$ implies $TV \leq TW$

(discounting) there exists $\gamma \in (0,1)$ such that for all $V \in H$ and all $a \geq 0$, $T(V+a) \leq TV + \gamma a$

Then T is a contraction with modulus γ

Proof. By homogeneity of degree 1,

$$V(M,F) = MV(1,F) \text{ for all } V \in H$$

Choose any $V, W \in H(M,F)$. Then

$$\begin{aligned} V(M,F) &= W(M,F) - [V(M,F) - W(M,F)] \\ &= W(M,F) - M[V(1,F) - W(1,F)] \\ &\leq W(M,F) - M\|V - W\| \end{aligned}$$

Hence monotonicity and discounting imply

$$TV \leq TW + \gamma \|V - W\|$$

Reversing the roles of V and W and combining the two results we get

$$\|TV - TW\| \leq \gamma \|V - W\|$$

■

I can now prove the following:

The operator T as defined above has a unique fixed point $V \in H(M,F)$ in addition

$$\|T^n V_0 - V\| \leq (\alpha\beta)^n \|V_0 - V\|, \quad n = 0, 1, 2, ..., \text{ all } V_0 \in H(M,F)$$

and the associated policy correspondence $G : (M,F) \to P$ is compact valued and u.h.c. Moreover, G is homogeneous of degree one in its first argument

$$P \in G(M,F) \text{ implies } P \in G(\lambda M, F), \text{ all } \lambda > 0$$

Proof. $H(M,F)$ is a complete normed vector space and $T : H(M,F) \to H(M,F)$. Clearly T satisfies the monotonicity property of 6. Choose $V(M,F) \in H(M,F)$ and $a > 0$. Then

$$\begin{aligned}
T(V+a)(M,F) &= \sup_{\substack{P\in[0,1] \\ M'=\Gamma_M(M) \\ F'=\Gamma_F(F)}} P(1-P)M + \beta(V+a)(M',F') \\
&= \sup_{\substack{P\in[0,1] \\ M'=\Gamma_M(M) \\ F'=\Gamma_F(F)}} P(1-P)M + \beta V(M',F') + \beta a M' \\
&\leq \sup_{\substack{P\in[0,1] \\ M'=\Gamma_M(M) \\ F'=\Gamma_F(F)}} P(1-P)M + \beta V(M',F') + \beta a \frac{z_2}{z_1} M \\
&= (TV)(M,F) + \beta \frac{z_2}{z_1} a M
\end{aligned}$$

where the third line uses $M' \leq \frac{z_2}{z_1} M$. Since the V was chosen arbitrarily, if follows that $T(V+a) \leq TV + \beta \frac{z_2}{z_1} a$. Hence given the assumption that $\beta \frac{z_2}{z_1} < 1$ T satisfies the discounting condition in 6 and is a contraction of modulus $\beta \frac{z_2}{z_1}$. It then follows from the Contraction Mapping Theorem that T has a unique fixed point in $H(M,F)$ and that

$$\|T^n V_0 - V\| \leq (\alpha\beta)^n \|V_0 - V\|, \quad n = 0,1,2,..., \text{ all } V_0 \in H(M,F)$$

holds.

That the policy function G is compact valued and u.h.c. follows from the Theorem of the Maximum (Berge 1963). Finally if $P \in G(M,F)$ then $P \in G(\lambda M, F)$ otherwise $\lambda V(M,F) < V(\lambda M, F)$ which by the homogeneity of degree 1 must hold with equality. ■

A.13 Proof of Theorem 7

$\nexists T$ such that for all $t > T$ the optimal price sequence $\{P_t^*\}$ is weakly increasing or decreasing

Proof. It is useful to have the following two lemmas before proceeding ■

Lemma 7 *If F_0 FOSD F_0' then $V(M_0, F_0) \geq V(M_0, F_0')$*

Proof. For any sequence of prices $\{P_t\}$, it suffices to show that $M_t \geq M_t'$. Γ_F preserves FOSD so if F_0 FOSD F_0' then for a set of prices $\{P_t\}$ F_t FOSD F_t'. The growth rate each period $\left((1 - F_t(P_t)) \frac{z_2}{z_1} + F_t(P_t)\right) \geq \left((1 - F_t'(P_t)) \frac{z_2}{z_1} + F_t'(P_t)\right)$ hence $M_t \geq M_t'$. ■

Lemma 8 *The optimal price each period $P_t^* \in [0, \frac{1}{2}]$*

Proof. Consider a price sequence $\{P'\}$ where $P_t' > \frac{1}{2}$. A price sequence $\left\{P_0'...P_{t-1}', \frac{1}{2}, P_{t+1}'...\right\}$ will result in higher profits. In period t the one period profits are strictly greater because $P_t = \frac{1}{2}$ is the one period monopoly price and for all periods $M_{t+i} > M_{t+i}'$ $i = 1,2,....$ ■

Now returning to the proof of the theorem. The proof is by contradiction. Suppose there is a weakly increasing or decreasing price sequence $\{P_t^*\}$. Every P_t^* will be an element of a compact set $\left[0, \frac{1}{2}\right]$ and any sequence which is weakly increasing or decreasing will converge to an element of this set. Call this price $P^* = \lim_{t\to\infty} P_t^*$.

The value function is linear in M_t so I will write it as the product of M_t and a function of $F_t : V(M_t, F_t) = M_t V(F_t)$.

I first rule out that $P^* = 0$. A constant $P_t = 0$ is not optimal because any deviation to a price above 0 gives a positive payoff. Now take a decreasing price sequence for which $\lim_{t\to\infty} P_t = 0$ then an upper bound for $V(F_t^*)$ is $\frac{P_t(1-P_t)}{1-\beta\frac{z_2}{z_1}}$. Therefore $\lim_{t\to\infty} V(F_t^*) = 0$ because $\lim_{t\to\infty} P_t = 0$. However $V(F)$ is also bound from below by $\frac{1}{4}$ from charging the one period monopoly price $P = \frac{1}{2}$. $\lim_{t\to\infty} V(F_t^*) = 0$ is therefore a contradiction and $P^* = 0$ is never the case.

$\Gamma_F(F_t, P_t)$ is continuous in P_t which implies that

$$\begin{aligned}
\lim_{t\to\infty} f_t^*(\theta) &= f^*(\theta) = \frac{z_2}{z_2 - z_1} \text{ for } \theta < P^* \\
\lim_{t\to\infty} f_t^*(\theta) &= f^*(\theta) = \frac{z_2 - \frac{z_1}{1-P^*}}{z_2 - z_1} \text{ for } \theta \geq P^*
\end{aligned}$$

Define the discounted sum of profits from a sequence $P_t = P^*$ and $f_t = f^*$ for all t as:

$$\begin{aligned}
\Pi(P^*) &= \sum_{t=0} \beta^t P^* (1 - P^*) M_t \\
& st \\
M_0 &= 1 \\
M_{t+1} &= \left[\left(1 - \frac{z_2}{z_2 - z_1} P^*\right) \frac{z_2}{z_1} + \frac{z_2}{z_2 - z_1} P^*\right] M_t
\end{aligned}$$

From the optimality of $\{P_t^*\}$ and continuity of V in F $\lim_{t\to\infty} V(F_t^*) = V(F^*) = \Pi(P^*)$. Therefore for any $\varepsilon > 0$ t can be chosen high enough such that $\Pi(P^*) + \varepsilon > V(F_t^*) > \Pi(P^*) - \varepsilon$.

Now consider the following one period deviation from $\{P_t^*\}$, in period t charge $P_t^* - \delta$. This strategy cannot be better than $\{P_t^*\}$ so:

$$(P_t^* - \delta)(1 - (P_t^* - \delta)) + \beta\left((1 - F_t^*(P_t^* - \delta))\frac{z_2}{z_1} + F_t^*(P_t^* - \delta)\right) V(F_{\delta,t+1}) - V(F_t^*) \leq 0$$

where $F_{\delta,t+1} = \Gamma_F(F_t^*, P_t^* - \delta)$.

$F_t^* \to F^*$ $P_t^* \to P^*$ so for any $\sigma > 0$ t may be chosen large enough such that $F_t^*(P_t^*) - F_t^*(P_t^* - \delta) \geq \frac{z_2}{z_2 - z_1} - \sigma$. Since $P_t^* - \delta < P_t^*$, F_δ FOSD F_{t+1}^* and $V(F_{\delta,t+1}) \geq V\left(F_{t+1}^*\right)$.

Combining these two facts the following is true:

$$-\delta\left(1-2P_t\right)-\delta^2-\beta\left(\delta\left(\frac{z_2}{z_2-z_1}-\sigma\right)\left(\frac{z_2}{z_1}-1\right)\right)V\left(F_{t+1}^*\right)\geq 0$$

rearranging

$$-\delta\left(1-2P_t\right)-\delta^2-\beta\left(\delta\left(\frac{z_2}{z_1}-\sigma\left(\frac{z_2}{z_1}-1\right)\right)\right)V\left(F_{t+1}^*\right)\geq 0$$

where the first two terms are the change in this periods profits and the second term is a lower bound on the change in future profits from selling to more people today. Finally the continuity of V implies that for any $\varepsilon>0$, t may be chosen large enough such that $\Pi\left(P^*\right)-\varepsilon\leq V\left(F_{t+1}^*\right)$

$$\Longrightarrow \beta\leq\frac{1-2P_t}{\left(\frac{z_2}{z_1}-\sigma\left(\frac{z_2}{z_1}-1\right)\right)\left(\Pi\left(P^*\right)-\varepsilon\right)}-\frac{\delta}{\left(\frac{z_2}{z_1}-\sigma\left(\frac{z_2}{z_1}-1\right)\right)\left(\Pi\left(P^*\right)-\varepsilon\right)}$$

for any $\omega>0$ $\exists\delta,\varepsilon,\sigma>0$ such that

$$\begin{aligned}&\frac{1-2P_t}{\left(\frac{z_2}{z_1}-\sigma\left(\frac{z_2}{z_1}-1\right)\right)\left(\Pi\left(P^*\right)-\varepsilon\right)}-\frac{\delta}{\left(\frac{z_2}{z_1}-\sigma\left(\frac{z_2}{z_1}-1\right)\right)\left(\Pi\left(P^*\right)-\varepsilon\right)}\\ \leq\ &\frac{1-2P_t}{\frac{z_2}{z_1}\Pi\left(P^*\right)}-\omega\end{aligned}$$

so

$$\beta\leq\frac{1-2P_t}{\frac{z_2}{z_1}\Pi\left(P^*\right)}-\omega \tag{9}$$

for any $\omega>0$. Note for $\beta>0$ this also rules out $P^*=\frac{1}{2}$.

Now consider a different deviation during period t to $P_t^*+\delta$:

$$\left(P_t+\delta\right)\left(1-\left(P_t+\delta\right)\right)+\beta\left(\left(1-F^*\left(P_t+\delta\right)\right)\frac{z_2}{z_1}+F^*\left(P_t+\delta\right)\right)V\left(F_{\delta,t+1}\right)-V\left(F_t^*\right)$$

Note

$$F^*=\Gamma_F\left(\Gamma_F\left(F^*,P_t^*+\delta\right),P^*\right)$$

now $V\left(F_{\delta,t+1}\right)>\Pi\left(P^*\right)-\varepsilon$ since a feasible strategy is to charge P^* in every period after $P_t^*+\delta$. Since Γ_F,V are continuous for any $\sigma,\varepsilon>0$ t can be chosen high enough such that

$\left|F^* - \Gamma_F\left(\Gamma_F\left(F_t^*, P_t^* + \delta\right), P^*\right)\right| < \sigma$ and hence that $V\left(F_{\delta,t+1}\right) > \Pi\left(P^*\right) - \varepsilon$

$$\begin{aligned} & \delta\left(1-2P_t\right) + \delta^2 + \beta\delta\left(\frac{z_2 - \frac{z_1}{1-P^*}}{z_2 - z_1} - \sigma\right)\left(\frac{z_2}{z_1} - 1\right)\left(\Pi\left(P^*\right) - \varepsilon\right) \\ \geq \quad & 0 \\ \implies \quad & \beta \geq \frac{1-2P_t}{\left(\frac{z_2 - \frac{z_1}{1-P^*}}{z_2 - z_1} - \sigma\right)\left(\frac{z_2}{z_1} - 1\right)\left(\Pi\left(P^*\right) - \varepsilon\right)} - \frac{\delta}{\left(\frac{z_2 - \frac{z_1}{1-P^*}}{z_2 - z_1} - \sigma\right)\left(\frac{z_2}{z_1} - 1\right)\left(\Pi\left(P^*\right) - \varepsilon\right)} \end{aligned}$$

for any $\omega > 0$ $\exists \delta, \varepsilon, \sigma > 0$ such that

$$\begin{aligned} & \frac{1-2P_t}{\left(\frac{z_2 - \frac{z_1}{1-P^*}}{z_2 - z_1} - \sigma\right)\left(\frac{z_2}{z_1} - 1\right)\left(\Pi\left(P^*\right) - \varepsilon\right)} - \frac{\delta}{\left(\frac{z_2 - \frac{z_1}{1-P^*}}{z_2 - z_1} - \sigma\right)\left(\frac{z_2}{z_1} - 1\right)\left(\Pi\left(P^*\right) - \varepsilon\right)} \\ \geq \quad & \frac{1-2P_t}{\left(\frac{z_2 - \frac{z_1}{1-P^*}}{z_1}\right)\Pi\left(P^*\right)} - \omega \end{aligned}$$

so

$$\beta \geq \frac{1-2P_t}{\frac{z_2 - \frac{z_1}{1-P^*}}{z_1}\Pi\left(P^*\right)} - \omega \tag{10}$$

For

$$\omega < \frac{\left(1-2P_t\right)\left(z_1\right)^2}{\Pi\left(P^*\right)\left(1-P^*\right)z_2\left(z_2 - z_1\right)}$$

both conditions given by equations 9 and 10 cannot be met so either $P_t + \delta$ or $P_t - \delta$ is a profitable deviation which is a contradiction that there exists a T such that $\{P_t^*\}$ is weakly increasing or decreasing for all $t \geq T$.

A.14 Proof of Theorem 8

For all (ω, P) *except* $\left(0, P^{crit}\right)$ *the profit function is continuous and differentiable with respect to both price and advertising, and* $\lim_{(\omega,P)\to\left(0,P^{crit}\right)} \pi\left(\omega, P\right) = 0$

Proof. Provided H_0 is differentiable with respect to P, ω then so is π.

$$H_0\left(1-\omega, P\right) = P + \left(1-\omega\right)\left(1-P\right)\sum p_k^k\left(u^*\right)^k$$

where u^* is the smallest non-negative solution to

$$u = P + \left(1-\omega\right)\left(1-P\right)\sum \frac{kp_k^{k-1}\left(u^*\right)^{k-1}}{z}$$

H_0 is differentiable if u^* is differentiable in ω and P. The right hand side of the equation for u is continuous, increasing and convex in u. I will show the differentiability of u^* for the 3 cases $\omega > 0$; $P > P^{crit}$;and $P < P^{crit}$.

Using the implicit function theorem we have

$$\begin{aligned}\frac{du^*}{dP} &= \frac{1-G_1(u^*)}{1-(1-\omega)(1-P)G_1'(u^*)}\\ \frac{du^*}{d\omega} &= \frac{-(1-P)G_1(u^*)}{1-(1-\omega)(1-P)G_1'(u^*)}\end{aligned}$$

$\frac{du^*}{dP}$ and $\frac{du^*}{d\omega}$ exist provided $1-(1-\omega)(1-P)G_1'(u^*)>0$

When $\omega>0$ the right-hand side of $u=P+(1-\omega)(1-P)\sum\frac{kp_k u^{k-1}}{z}$ is strictly less than 1 if $u=1$ and strictly greater than 0 when $u=0$. Therefore the solution is strictly less than 1, and at the solution $1-(1-\omega)(1-P)G_1'(u)>0$.

When $P>P^{crit}$ by the definition of P^{crit}

$$\begin{aligned}P^{crit} &= 1-\frac{1}{G_1'(1)}\\ &\Rightarrow (1-P)G_1'(u)<1 \text{ for } u\le 1 \text{ and } P>P^{crit}\end{aligned}$$

so again $1-(1-\omega)(1-P)G_1'(u)>0$ and $\frac{du}{dP}$ and $\frac{du}{d\omega}$ exist.

When $P<P^{crit}$ consider $P+(1-\omega)(1-P)\sum\frac{kp_k u^{k-1}}{z}$. This is strictly convex in u for $0\le P<P^{crit}$ and equal to 1 at $u=1$. At any solution $u^*<1$ $(1-\omega)(1-P)G_1'(u^*)<1$ otherwise $P+(1-\omega)(1-P)G_1(1)\neq 1$. Hence $1-(1-\omega)(1-P)G_1'(u)>0$ and $\frac{du}{dP}$ and $\frac{du}{d\omega}$ exist.

Finally consider

$$\begin{aligned}\lim_{(\omega,P)\to(0,P^{crit})}\pi(\omega,P) &= \lim_{(\omega,P)\to(0,P^{crit})}(P-c)(1-H_0(1-\omega,P))-\alpha\omega\\ &= \left(P^{crit}-c\right)\left(1-\lim_{(\omega,P)\to(0,P^{crit})}H_0(1-\omega,P)\right)\end{aligned}$$

It was shown in Theorem 2 that for the case $\omega=0$ $\lim_{P\to P^{crit}}1-H_0(1,P)=0$ so the theorem holds for this case. Now considering the case $\omega>0$ take any sequence $(\omega,P)\to\left(0,P^{crit}\right)$ where $\omega>0$ the expression $P+(1-\omega)(1-P)\sum\frac{kp_k^{k-1}(u)^{k-1}}{z}<1$ at $u=1$ and $\lim_{(\omega,P)\to(0,P^{crit})}P+(1-\omega)(1-P)\sum\frac{kp_k^{k-1}(u)^{k-1}}{z}=1$ furthermore $\lim_{(\omega,P)\to(0,P^{crit})}u^*=1$. Hence $\lim_{(\omega,P)\to(0,P^{crit})}H_0(1-\omega,P)=1$ and $\lim_{(\omega,P)\to(0,P^{crit})}\pi(\omega,P)=0$. ■

A.15 Proof of Corollary 4

If $\pi(\omega,P)>0$ *for some* (ω',P') *then* $\exists\varepsilon>0$ *such that for all* $(\omega,P)\in B_\varepsilon\left(0,P^{crit}\right)$ *where* B_ε *is an open ball* $\pi(\omega,P)<\pi(\omega',P')$

Proof. From theorem 8 $\pi(\omega,P)$ is continuous in (ω,P) for $(\omega,P)\neq\left(0,P^{crit}\right)$ and $\lim_{(\omega,P)\to(0,P^{crit})}\pi(\omega,P)=0$. The result follows immediately for ε small enough $(\omega,P)\in B_\varepsilon\left(0,P^{crit}\right)\Rightarrow\pi(\omega,P)<\pi(\omega',P')$. ■

A.16 Proof of Theorem 9

If $0 < P^{crit} < 1$, for any sequence of strategies $\lim_{(\omega,P)\to(0,P^{crit})} H_0'(1,P) = \infty$
Proof. $H_0'(1,P)$ is given by:

$$H_0'\left(\omega, P^{crit}\right) = (1-P)\left[G_0(u) + \frac{z_1(1-P)\left[G_1'(u)\right]^2}{1-(1-\omega)(1-P)G_1'(u)}\right]$$

where u is the smallest non-negative solution to

$$u = P + (1-P)(1-\omega)G_1(u)$$

Theorem 8 proves that $H_0'(\omega,P)$ is defined everywhere except $\left(0, P^{crit}\right)$. Now consider any sequence $\{(\omega,P)\} \to \left(0, P^{crit}\right)$ then

$$\begin{aligned} & \lim_{\{(\omega,P)\}\to(0,P^{crit})} (1-P)\left[G_0(u) + \frac{z_1(1-P)\left[G_1'(u)\right]^2}{1-(1-\omega)(1-P)G_1'(u)}\right] \\ = & \left(1-P^{crit}\right)\left[G_0(1) + \frac{z_1\left(1-P^{crit}\right)\left[G_1'(1)\right]^2}{1-\left(1-P^{crit}\right)G_1'(1)}\right] \end{aligned}$$

where $\left(1-P^{crit}\right), G_0(1)$ and $z_1\left(1-P^{crit}\right)\left[G_1'(1)\right]^2$ are finite and from the definition of P^{crit} $1-\left(1-P^{crit}\right)G_1'(1) = 0$. Hence $\lim_{(\omega,P)\to(0,P^{crit})} H_0'(1,P) = \infty$. ■

A.17 Proof of Theorem 10

Advertising exhibits decreasing and convex marginal returns
Proof. Returns to advertising are given by $H_0'(1-w,P)$. The rate of change of the returns with respect to advertising level w is given by $\frac{dH_0'(1-w,P)}{dw} = -H_0''(1-w,P)$ where $-H_0''(1-w,P) < 0$ and $H_0'''(1-w,P) > 0$ because $H_0''(1-w,P)$ is a polynomial in $(1-w)$with positive coefficients. ■

A.18 Proof of Theorem 11

Assuming $p_k > 0$ for all k the highest return type of individual k^* is found as the solution to:

$$k^* \in \{\lfloor k^{**} \rfloor, \lceil k^{**} \rceil\} \text{ for } P < P^{crit}$$

where

$$k^{**} = \max\left\{0, -\left(\frac{1}{\ln u(P)} + \frac{u(P)}{H_1'(1,P)}\right)\right\}$$

Proof. The probability generating function of component sizes an individual with k friends belongs to, conditional on not being in the giant component, is given by $\left(\frac{H_1(x,P)}{u}\right)^k$. The expected component size is $1+k\frac{H_1'(1,P)}{u}\left(\frac{H_1(1,P)}{u}\right)^{k-1}=1+k\frac{H_1'(1,P)}{u}$. Also the probability a person with k friends is not in the giant component is u^k. Therefore

$$k^* = \arg\max_{k\in\{0,1\ldots\}}\left(1+k\frac{H_1'(1,P)}{u}\right)u^k$$

note that for $0<u<1$ $b>0$ the function $f(k)=(1+kb)\,u^k$ is continuous in k; has a maximum at $k^{**}=\max_{k\geq 0}\left\{0,-\left(\frac{1}{\ln u(P)}+\frac{1}{b}\right)\right\}$ and $f'(k)>0$ for $k<k^{**}$ and $f'(k)<0$ for $k>k^{**}$. Hence k^* is either the greatest integer below $\lfloor k^{**}\rfloor$ or the smallest integer above k^{**}, $\lceil k^{**}\rceil$. Thus

$$k^*\in\{\lfloor k^{**}\rfloor,\lceil k^{**}\rceil\}\ \text{ for } P<P^{crit}$$

■

A.19 Proof of Corollary 5

The optimal target k^{**} is continuous in $u(P)$ for $u(P)<1$, $\lim_{P\to P^{crit}}k^{**}=\infty$, $k^{**}\leq\frac{-1}{\ln(u(P))}$ for $P<P^{crit}$

Proof. We have

$$\begin{aligned}
k^{**} &= -\left(\frac{1}{\ln u(P)}+\frac{u(P)}{H_1'(1,P)}\right)\\
&= -\left(\frac{1}{\ln u(P)}+\left(\frac{(1-P)\,G_1(u)}{u(P)\,(1-(1-P)\,G_1'(u))}\right)^{-1}\right)\\
&= -\left(\frac{1}{\ln u(P)}+\left(\frac{(1-P)\sum kp_ku^{k-2}}{(z_1-(1-P)\sum k\,(k-1)\,p_ku^{k-2})}\right)^{-1}\right)
\end{aligned}$$

where $\frac{1}{H_1'(1,P^{crit})}$ is finite so immediately $\lim_{P\to P^{crit}}\frac{-1}{\ln u(P)}=\infty\Rightarrow\lim_{P\to P^{crit}}k^*=\infty$. Also by definition $z_1-(1-P)\sum k\,(k-1)\,p_ku^{k-2}>0$ for $P<P^{crit}$ $u>0$ and $(1-P)\sum kp_ku^{k-2}>0$ for $u(P)>0$ so k^* is continuous in u and hence P for $P<P^{crit}$. Finally $\frac{(1-P)\sum kp_ku^{k-2}}{(z_1-(1-P)\sum k(k-1)p_ku^{k-2})}>0$ so $-\frac{1}{\ln u(P)}$ is an upper bound on k^{**}. ■

A.20 Proof of Theorem 12

If $P^{j^{**}}>P^{crit}$ then $\exists\widehat{k}$ such that for $k\geq\widehat{k}$ $j^{**}=j^{*k}$

Proof. If $P^{j^{**}}>P^{crit}$ then $u\left(P^{j**}\right)=1$ and $H_0'\left(1,P^{j^{**}}\right)=\sum p_k\left(1-P^{j^{**}}\right)\left(1+kH_1'\left(1,P^{j^{**}}\right)\right)$.

For any $P^j > P^{j*}$

$$\left(1 - P^{j^{**}}\right)\left(1 + kH_1'\left(1, P^{j^{**}}\right)\right) > \left(1 - P^j\right)\left(1 + kH_1'\left(1, P^j\right)\right) \text{ for all } k$$

because $H_1^{'}(1, P)$ is decreasing in P. For any $P^j < P^{crit}$ $u(P^{j^{*k}}) < 1$ as k increases the returns to advertising to a person with k friends has the following properties

$$\lim_{k \to \infty} \left(u\left(P^j\right)\right)^k \left(1 - P^j\right)\left(1 + k\frac{H_1'\left(1, P^j\right)}{u\left(P^j\right)}\right) = 0$$

and

$$\frac{\partial\left(\left(u\left(P^j\right)\right)^k \left(1 - P^j\right)\left(1 + k\frac{H_1'\left(1, P^j\right)}{u\left(P^j\right)}\right)\right)}{\partial k} < 0 \text{ for all } k > -\left(\left(\frac{H_1'\left(1, P^j\right)}{u\left(P^j\right)}\right)^{-1} - (\ln u)^{-1}\right)$$

which implies that $\exists \widehat{k}$ such that

$$\left(1 - P^{j^{**}}\right)\left(1 + kH_1'\left(1, P^{j^{**}}\right)\right) > \left(u\left(P^j\right)\right)^k \left(1 - P^j\right)\left(1 + k\frac{H_1'\left(1, P^j\right)}{u\left(P^j\right)}\right) \text{ for all } k \geq \widehat{k}$$

■