Do Online Social Networks Increase Welfare? *

Manuel Mueller-Frank[†] Mallesh M. Pai[‡]

August 9, 2014

Abstract

Agents who are active on an online social network select among competing products of unknown quality by costly sequential search. We study the impact of advertising by firms, and the incentives of the underlying social network. We consider display advertising, which is standard firm to consumer advertising, and social advertising, in which agents who purchased that firm's product are highlighted to their friends. We show that in equilibrium, the heterogeneous firms spend the same amount on advertising. Social advertising may be more lucrative than standard display advertising. However, both forms of advertising have no effect on consumer welfare. A social network motivated by advertising revenues may limit the amount of information agents see about actions by other agents, since this will increase advertising revenue. This reduces consumer welfare relative to the first best, since early movers' purchases are informative about relative quality.

Keywords: social networks, advertising, search

*We would like to thank Ken Burdett, Kalyan Chatterjee, Rahul Deb, Hanming Fang, Sanjeev Goyal, Hanna Halaburda, Maher Said, Jack Stecher and participants of the First Pennsylvania Theory Conference, the Searle Center Conference on Internet Search and Advertising, and the ACM/Sigecom Advertising Auctions Workshop for helpful comments and suggestions.

[†]IESE Business School. Email: mmuellerfrank@iese.edu. Website

[‡]Department of Economics, University of Pennsylvania Email: mallesh@econ.upenn.edu. Website

1 Introduction

People are influenced by their friends and acquaintances. Influential literatures suggest that social networks affect important personal outcomes such as health (e.g. Christakis and Fowler (2007)), economic outcomes such as income and employment (e.g. Montgomery (1991) or Calvo-Armengol and Jackson (2004)), and may distort market outcomes.[1] In a majority of this literature, the social network is modeled as an inert conduit for information flow and communication between agents, rather than as a strategic agent. In this paper, motivated by commercial online social networks such as Facebook and Twitter, we study outcomes taking into account the motivations of the underlying social network.

Online social network platforms have developed a spectacular user base in recent years and provide a rich layer of social interaction for its users. For example, Facebook is reported to have over 1 billion active users, while Twitter is reported to have over 250 million active users.[2] These users generate millions of pieces of content per day—posts, photos, discussions etc.[3] Due to the vast amount of information being generated, online platforms use algorithms to select and filter what is displayed to users. For example Facebook's Newsfeed, displayed when a user visits the site, summarizes the recent activity taken by the user's online "friends," shows advertisements. What is displayed, and the order in which it is displayed, is determined by an opaque algorithm designed by Facebook to optimize *its* objectives.

Online network platforms are monetized mainly through advertising, which may affect the information that is displayed. Biasing the organic information displayed in favor of (say) advertisers might have significant welfare effects—for example, herds on inferior products might form. In this context, and due to the tremendous scale of online networks, it is paramount to understand the welfare effects of having a financially motivated firm controlling the dispersion of information. We take a first step towards this.

We address two main questions: First, what are the welfare effects of advertising in a social context? Second, how might a advertising revenue maximizing social network influence social information dispersion relative to a benevolent social network? At a high level, we find that in equilibrium, advertising is a transfer from firms to the platform, with no effect on overall economic welfare. However, we show that a revenue motivated platform may limit the amount of information users see organically, since this increases the revenue from advertising. This distortion reduces consumer welfare.

[1] For example due to inefficient herding: see Banerjee (1992) and Bikhchandani, Hirshleifer, and Welch (1992) for the underlying theory, and Salganik, Dodds, and Watts (2006) for experimental evidence.

[2] As reported by each in their 1st Quarter 2014 financial results, see e.g. `http://goo.gl/4kytSK`, `http://goo.gl/loqoPw`

[3] See e.g. `http://goo.gl/8yJrzo`.

We conduct our analysis in a stylized model. Two firms produce goods which are substitutes and compete for consumers. The firms' products have different qualities which are common knowledge among them, but not known to the consumers. Consumers decide which product to buy through costly sequential search among the products. Consumers are of two kinds: early movers and late movers. Both are active on a social network platform. Early movers make their purchases after costly search, and announce their purchase decision on the network platform. Each late mover observes the decision of some early mover/ movers. This observation influences his beliefs about the qualities of the two products, and therefore his search and purchase decision.[4] The probability with which a late mover observes the decision of an early mover is referred to as the baseline virality of the network.

There are two types of advertising that the platform offers to firms. The first, display advertising, is the conventional firm to consumer communication. This is the standard form of advertising across the internet, where the firm displays a banner containing a logo, message or image on a webpage the user is viewing. It is also referred to as banner advertising. We assume that the advertisement itself is uninformative, but seeing it may effect consumer beliefs or actions in equilibrium.

The second kind is social advertising, which influences the information late movers see about the early movers' actions. This form of advertising is unique to social networks—examples include Facebook's "Sponsored Stories" and Google's "social ads." Here, a firm pays the social network to make posts by consumers with relevant content more visible to the online "friends" of these consumers.

Our approach to display advertising is taken from the literature (an early application is by Friedman (1958)) and is essentially a Tullock contest. Each consumer observes exactly one ad. Both firms simultaneously choose how much to spend on advertising. Any consumer observes the ad of a given firm independently with a probability which equal to the proportion of the expenditure of that firm to the total advertising expenditures.

Social advertising by a firm distorts the information seen by late movers about the purchase decisions of the early movers. In the absence of social advertising, recall that late movers observe the purchase decision of a randomly chosen early mover. We assume that social advertising by a firm increases the probability that the late mover observes an early mover who purchased that firm's product.

A Motivating Example To motivate our model consider a simple example of duopolists who each make a consumer product (such as cellphones or cars). Every consumer is in the market for a single unit of this product. The duopolists each know the quality of both

[4] This is a variant of the social learning setting considered in Mueller-Frank and Pai (2013).

products. Early movers have the ex-ante beliefs about the qualities of these products. Late movers may observe the choices made by his predecessors—for example, he sees what brands of cars people drive or what kind of phones they carry etc, and update their beliefs based on this.

Given these beliefs a consumer then may choose to acquire more information, e.g. take the cars on a test drive, read reviews of specific phones online, and so on. Sampling an alternative (test driving a car, reading a review of a phone) is costly due to the time and effort involved, and reveals information about the quality of that alternative. After sampling the first alternative, the agent decides whether to sample the second (given his opportunity cost of sampling versus the expected benefit given his beliefs). If he chooses not to sample further he is concluding that the first alternative is "good enough" and purchases that. If he has sampled both he picks the higher quality product. An early mover's choice of product to buy is therefore noisily informative about the relative quality of these products (since he may have sampled both). A late mover who observes an early mover's purchase will update his beliefs, and this will affect his choice of which product to sample etc.

Display advertising in this setting is a standard advertisement. In our setting, in equilibrium, it influences consumers who are otherwise indifferent between the two products to sample that product first. This makes it more likely that this consumer purchases that product, since the consumer might have high search costs and therefore may not sample further. Social advertising by a firm biases the information seen by late movers by making it more likely that the late movers see an early mover who purchased that product. Since purchase decisions are informative, this makes it more likely that the consumer chooses that product. We study the influence of each.

Discussion of Results We first consider display advertising and compare the case of our social network platform (with positive virality) to an "unsocial" platform benchmark where late movers receive no information from early movers. In either setting, we find that in equilibrium, both firms spend the same amount on advertising, and the display ads do not introduce any informational distortion. This implies that display ads cannot be used to signal quality and are a mere transfer from firms to the platform. A platform uniquely interested in maximizing display advertising revenue would set its baseline virality to zero. In the countervailing direction, the consumer welfare of the social platform is strictly greater than the "unsocial" benchmark—this is because the purchase decisions of early movers is an informative signal, which helps late movers make better informed search and purchase decisions.

Next we introduce social advertising. We show that social advertising is welfare neutral.

That is, welfare is the same as in the case where the baseline virality is set equal to one and each consumer receives organic information. Additionally, we show that advertising revenue is maximized if the baseline virality is set equal to zero and consumers observe only sponsored social information. Nevertheless, no informational distortions occur. Further, a social network platform that offers a combination of display and social ads to firms may generate strictly higher advertising revenue than an unsocial platform. Hence, the introduction of social advertising might make a social platform more profitable than an unsocial platform. In regards to welfare, while social network platforms are strictly better than unsocial ones and social advertising introduces no welfare distortions, they do not achieve the first best solution. We show that, as the number of early consumers that a given late consumer observes grows large, the social welfare of late movers converges to the first best, while the total advertising revenue of the social platform decreases.

1.1 Related Literature

The broader literature on social networks is too large to comprehensively cite here, we refer the interested reader to Jackson (2010) for an overview. We restrict ourselves to more closely connected papers.

There has been a recent interest in understanding how social networks may affect commercial activity. For instance, a strand considers settings where a monopolist seller sells a good to agents on a network, and agents' purchases have (positive or negative) externalities on their neighbors. In this class of settings, these papers study pricing by the seller and the distortions this network introduces—see e.g. Candogan, Bimpikis, and Ozdaglar (2012), Bloch and Quérou (2013) or Feldman, Kempe, Lucier, and Paes Leme (2013) for recent papers in the area, and Cabral, Salant, and Woroch (1999) for a classic reference. Fainmesser and Galeotti (2013) explicitly consider the value of the underlying network to be in selling information to the monopolist so that it can price discriminate. Kircher and Postlewaite (2008) observe that firms may offer higher quality products to "influential" agents in the network so that they may influence their connections. Chatterjee and Dutta (2014) study the adoption of a new product in a network when there are both "innovators" who immediately adopt the product, and rational agents who adopt only when expected gains exceed costs. They characterize the structure of networks in which good new products are adopted.

The increasing amount of commerce conducted on the internet has led to some seminal investigations of the business models of firms on the internet. This literature broadly studies questions raised by the ability to use novel mechanisms on the internet (real time auctions), or gather specific information about individual consumers. Most notably Edelman, Ostrovsky, and Schwarz (2005) and Varian (2007) study the advertising auction used by major search engines and its properties. Athey and Ellison (2011) consider the impact of consumer search among an ordered list of ads. Gomes (2014) also models the preferences of the users (whom the ads are being shown to) for organic content, and therefore studies it as a problem of designing a two-sided market. Bergemann and Bonatti (2011) and Bergemann and Bonatti (2014) study targeting, and the sale of consumer specific information on the internet. We add to this literature by considering the ability of firms on the internet (social network) to control communication between individual consumers.

In terms of papers related to our model, the idea of considering search in a social setting was first considered in Mueller-Frank and Pai (2013), here we consider a variant of the more general model there. That paper provides a characterization asymptotic learning of asymptotic learning for the case of endogenous private information. Our basic model of advertising is, as we pointed out, a Tullock contest, and was first seen in Friedman (1958).

2 Model

There are two competing firms, 1 and 2, each of which produces a product of quality $q_i \in Q = [0, 1]$, $i = 1, 2$. The product qualities q_i are independently drawn at time $t = 0$ according to a probability measure with cumulative distribution function F_Q and density f_Q. The set of possible pairs of quality realizations is denoted $\mathbf{Q} = [0, 1]^2$. The firms commonly learn the realized product qualities of both firms.[5]

There are two exogenously given groups of consumers that differ in the timing of their purchase decision. A continuum of early movers E decide among the two products in time period $t = 1$, their mass is normalized to 1. A continuum of mass λ of late movers L decide in time period $t = 2$. A consumer's utility of purchasing firm i's product is equal to its quality q_i. The gross profit Π_i of firm i is equal to the measure of consumers purchasing its product.

Finally, there is a social network platform on which late consumers might observe the purchase decisions of early movers and on which firms might communicate with consumers via advertising.

The platform is assumed to have complete control of the communication taking place. That is, the platform controls whether or not early choices are observed by late movers, which choices are observed, by whom and whether and how firms can communicate to consumers.[6]

[5]This assumption is relatively standard in the literature studying online advertising, and is normally defended on the grounds that the repeated interaction between the firms would publicly reveal any private information. For a standard reference, see Edelman, Ostrovsky, and Schwarz (2005).

[6]The assumption of complete control is made mainly for notational convenience, but there are several examples of online systems that approximate this, notably Facebook's Newsfeed, discussed earlier.

Let $v_B \in [0,1]$ denote the baseline virality of the social network platform, i.e. with probability v_B each late mover independently observes exactly one early mover that is drawn uniformly from the group of early movers. Later we will consider the case where a late mover sees the actions of a number $k > 1$ of early movers.

2.1 Consumer Search

The decision of each consumer is based on costly sequential search among these products. The sequential search model is as in Weitzman (1979). Each consumer has a probability distribution on $\mathbf{Q}$. This might be the prior distribution or a Bayesian update based upon additional information. For example, an early mover who observes no other information will view the products as ex-ante identical draws from F_Q. By contrast, a late mover might observe the purchase decision of some early mover or a firm's advertising, and will update on this information appropriately.

At time t each consumer acting in the given period decides which product to sample first $s_j^1 \in \{1,2\}$. Sampling a product perfectly reveals its quality to the consumer. After observing this quality consumer j decides whether to sample the remaining product, $s_j^2 \in \{1,2\}$, or to discontinue searching, $s_j^2 = n$. For simplicity, the first product is sampled at no cost while sampling the second involves a cost of $c_j \in C = [0,1]$ to consumer j. The search costs c_j are independently drawn according to a probability measure with cumulative distribution function F_C and density f_C.

Consumer j then decides to purchase one of the products he sampled. The purchase decision of consumer j is denoted by $a_j \in A = \{1,2\}$. The net-utility of agent j is therefore the quality of the product he selects less the search cost if he chooses to sample a second time. Total consumer welfare is consumer utility less effort expended on search.

2.2 Advertising

Display Advertising Display advertising is a traditional form of advertising as it consists of firms communicating with consumers. In this form each consumer sees exactly one ad for one of the competing products. An ad contains no direct information in regards to the quality of the product but intuitively might serve to raise the awareness for its product.

Let $m_i^d \in \mathbb{R}$ be firm i's expenditure on display advertising. Both firms simultaneously select their display advertising expenditures in time $t = 0$ after observing the product qualities. The banner advertising revenue of the platform is the sum of the amount spent by each. Given these chosen advertising levels, each consumer independently sees an ad for product

1 with probability

$$\frac{m_1^d}{m_1^d + m_2^d},$$

and otherwise an ad for product 2. If both display advertising expenditures are equal to zero, consumers do not see any display advertising. Let Θ denote the set of possible ads, $\Theta = \{1, 2, x\}$, where x describes the case of no ad, and let $\theta_j^d \in \Theta$ denote the display ad seen by consumer j.

Social Advertising Social advertising of firm i influences the probability with which a late consumer observes an early consumer who purchased product i and as such centers around distorting consumer-to-consumer communication rather than the traditional firm-to-consumer communication.

Let ϕ_i be the measure of early consumers that purchased product i. Absent social advertising, the (independent) probability of a late consumer i observing a purchase of product i is then given by $v_B \phi_i$. At time $t = 0$, both firms simultaneously decide on the amount social advertising. Let $m_i^s \in \mathbb{R}$ be firm i's expenditure on social advertising. For social advertising expenditures m_1^s, m_2^s the probability of a late consumer i observing a purchase of product i by an early mover is then given by

$$v_B \phi_i + (1 - v_B) v_S \frac{\phi_i m_i^s}{\phi_1 m_1^s + \phi_2 m_2^s}.$$

This term can be interpreted as follows. With the probability given by the baseline virality v_B a consumer receives organic social information, and conditional on not receiving organic social information the consumer observes sponsored social information with probability v_S. Again, if both firms spend zero on social advertising the probability of observing a purchase of product i is equal to $v_B \phi_i$. To match the reality of social and search engine advertising we assume that the consumer knows whether the social information he observes is organic or sponsored.[7] The formal nature of our social advertising is inspired by the "Sponsored Stories" on Facebook.[8]

[7]For example, the FTC in the US requires social networks to clearly label and distinguish any advertising or "promoted" items where they have a financial interest: see http://goo.gl/2w4zrP.

[8]See e.g. the two minute video introducing this product at: http://goo.gl/6bZyQ.

2.3 Strategies and Equilibrium

We assume that the structure of the game described above is commonly known among all participants. The strategy of firm i is given by

$$\sigma_i : \mathbf{Q} \to \mathbb{R}_+ \times \mathbb{R}_+.$$

That is, for each possible realization of the product qualities firm i decides how much to spend on display and social advertising.

Next consider an early consumer $e \in E$. His strategy σ_e is a two-tuple consisting of the first sampling decision, the subsequent decision to sample further or not. The following approach to search and the notation is taken from the companion paper Mueller-Frank and Pai (2013). Consumer e's initial sampling strategy is given by[9]

$$\sigma_e^1 : \Theta \times C \to \{1, 2\}.$$

The subsequent sampling strategy is formalized as

$$\sigma_e^2 : \Theta \times C \times Q \to \left\{\neg s_e^1, n\right\}$$

where $\neg s_e^1$ denotes the product not sampled initially. His purchase decision is mechanical, we omit the formal notation: if the consumer only samples one product, he purchases that product, if he samples both, he purchases the product with the higher quality.

Next consider a late consumer $l \in L$. His strategy σ_l is again a two-tuple with the difference that the sampling decisions capture the possibility of having observed the purchase decision of an early agent. Let $H_l = \{1, 2, x\}$ denote the set of possible histories that consumer l can observe (x denotes the case where consumer l observes no purchase decision). Consumer l's initial sampling strategy is given by

$$\sigma_l^1 : H_l \times \Theta \times C \to \{1, 2\}.$$

His subsequent sampling strategy satisfies

$$\sigma_l^2 : H_; \times \Theta \times C \times Q \to \left\{\neg s_l^1, n\right\}$$

where $\neg s_k^1$ denotes the product not sampled initially. The purchase strategy is identical to an early mover's and omitted.

[9]For ease of notation we only describe pure strategies here. The paper considers mixed strategies.

We study the Perfect Bayesian equilibria of this game as a function of the parameters controlled by the network—namely, v_B, v_S and k. Our main results correspond to comparative statics of outcomes of interest—advertising revenue, consumer welfare in these parameters.

3 Benchmarks

Some basic results about how rational consumers search in the absence of advertising, and display advertising in a social network will be useful, and are collected here for use in later analysis.

3.1 Search in the Absence of Advertising

First, we consider outcomes in a model where there is no advertising, and v_B is set to 1. The exposition is borrowed from Mueller-Frank and Pai (2013)—we refer the reader to that for further results.

First consider an early consumer e. In the absence of any additional information, the marginal distributions of the qualities of both products are identical. According to the optimal search strategy characterized by Weitzman (1979) either product might be sampled first. Let us assume that he randomizes uniformly over which of the two products to sample first. If he samples product i first, he learns the quality q_i of this product. Next, he must decide whether to sample further or not. He will only sample if it is rational to do so, i.e. if the expected additional gain from searching exceeds his cost of an additional search. Formally, he searches further if:

$$c_e \leq \int_{q_i}^{1} (q - q_i) dF_Q(q).$$

We denote the cutoff cost that just leaves an early consumer indifferent from searching further, given the observed quality of the product he sampled first by $c_e(q_i)$, i.e.:

$$c_e(q_i) = \int_{q_i}^{1} (q - q_i) dF_Q(q). \tag{1}$$

Let p_{σ_e} denote the ex-ante probability of early consumer e to buy the better product given his strategy σ_j and let π_{σ_e} the posterior probability that product i is the better product conditional on consumer e having bought i and his strategy being σ_e.

Some equilibrium properties of the search behavior of early and late consumer will be useful.[10] First, note that both the ex-ante probability p_{σ_e} and the posterior probability π_{σ_e}

[10]These equilibrium properties are not formally proven here but follow from the analysis in Mueller-Frank

are greater than half in any equilibrium: intuitively, the early mover samples both products with positive probability, in which case he purchases the better product.

Next consider the case of a late consumer l who observes that a (randomly selected) early mover e has purchased product i, i.e. $a_e = i$. Based upon this observation, the late consumer then updates his probability distribution on the space of product qualities. Bayesian updating has the following implication: if a late consumer l observes the action of an early consumer, then he samples the observed product first, in any equilibrium. The updated posterior distribution of the quality of the observed product first order stochastically dominates the updated distribution of the other product's quality, again because with a certain probability both products were sampled by the early consumer in which case the observed product is optimal. The claim then follows from the characterization of the optimal sampling strategy in Weitzman (1979), which implies that (first order stochastic) dominant options are sampled first.

Finally, the cost cutoff $c_l(q_i)$ of a late consumer l who observed the choice of an early consumer has the following characteristic. If a late consumer l observes the purchase of product i of an early consumer, then the cost cutoff $c_l(q_i)$ satisfies $c_l(q_i) < c_e(q_i)$.

3.2 Display Advertising in the absence of a social network

As we pointed out earlier, our model of advertising has a long pedigree in the literature—see Friedman (1958) for an early application. Here, however, consumers are strategic: they may update their beliefs about the relative qualities of the products based on the advertisement they see. These updated beliefs will then influence their search decisions. To build intuition we consider the equilibrium when there is a single consumer whom the two firms may advertise to.

So in this simpler setting, first the qualities of the two firms $\mathbf{q} = (q_1, q_2)$ are realized and revealed to the firms. Each firm i chooses its level of advertising $m_i(\mathbf{q})$. The consumer then sees an ad for firm i with probability $m_i(\mathbf{q})/(m_1(\mathbf{q}) + m_2(\mathbf{q}))$, and updates his beliefs about the quality of each product. Based on these updated beliefs, he chooses which product to sample first, and whether or not to sample further.

Intuitively, there are three possible "kinds" of equilibria—ones where seeing an ad for a product is, respectively, "good news," uninformative, and "bad news."

In lieu of analyzing each of these in turn, consider each possible strategy for the consumer, i.e. a $\alpha \in [0, 1]$ which denotes the customer's probability of sampling the product corresponding to the ad he sees, and his cutoff for sampling further given the product he

and Pai (2013).

sampled and the ad he saw, i.e. $c(\cdot|\theta)$. We solve for the optimal advertising level for the firms for every realization of $\mathbf{q} \in \mathbf{Q}$ given the strategy of the consumer. Equilibria of this game correspond to those where the α and cutoffs $c(\cdot)$ are consistent with Bayesian updating given the firms' derived advertising strategies.

Consider the customer strategy $\alpha, c(\cdot)$, and fix a realization of qualities $\mathbf{q}$ such that wlog $q_1 > q_2$. Consider first the incentives of the inferior firm 2 to advertise. Firm 2 only makes a sale if the customer samples its product first *and* then does not sample any further. Therefore, firm 2 chooses its level of advertising to maximize its profits

$$\frac{m_2}{m_1+m_2}\alpha(1-F_C(c(q_2|\theta=2))) + \frac{m_1}{m_1+m_2}(1-\alpha)(1-F_C(c(q_2|\theta=1))) - m_2.$$

In all other cases, the consumer purchases from firm 1, which therefore maximizes

$$1-\frac{m_2}{m_1+m_2}\alpha(1-F_C(c(q_2|\theta=2))) - \frac{m_1}{m_1+m_2}(1-\alpha)(1-F_C(c(q_2|\theta=1))) - m_1.$$

Taking first order conditions, and substituting in that each is maximized at the equilibrium level $m_i(\mathbf{q})$, we have that

$$\frac{m_1(\mathbf{q})}{(m_1(\mathbf{q})+m_2(\mathbf{q}))^2}\left(\alpha(1-F_C(c(q_2|\theta=2))) - (1-\alpha)(1-F_C(c(q_2|\theta=1)))\right) \leq 1,$$
$$\frac{m_2(\mathbf{q})}{(m_1(\mathbf{q})+m_2(\mathbf{q}))^2}\left(\alpha(1-F_C(c(q_2|\theta=2))) - (1-\alpha)(1-F_C(c(q_2|\theta=1)))\right) \leq 1,$$

with the inequality strict at an interior solution.

By observation, therefore, $m_1(\mathbf{q}) = m_2(\mathbf{q})$ for every $\mathbf{q}$, i.e. advertising is uninformative. On seeing the ad for product i, the consumer's posterior belief about the qualities is the same as his prior. Consider the case where he breaks ties by sampling first the product whose advertisement he saw. Further, since the ad is uninformative, which ad he saw is irrelevant to his subsequent search decision, which is therefore given by $c_e(\cdot)$ defined in (1) in the previous section. Substituting into the first order conditions above, we have

$$m_1(\mathbf{q}) = m_2(\mathbf{q}) = \frac{1}{4}(2\alpha-1)(1-F_C(c_e(q_2))),$$

for $\alpha \in (\frac{1}{2}, 1]$.

The intuition is roughly this—both firms are advertising for the same marginal consumers, i.e. ones who upon seeing the ad for product 2 would sample it but not sample further. The consumers' purchase decisions are determined completely by the ad they see (while all other consumers' purchase decisions are uninfluenced by advertising). Therefore both firms spend the same amount, such that the marginal incremental probability of purchase from increased ad spend equals 1.

Finally, there is an equilibrium where seeing a corner equilibrium where, both firms spend 0 on advertising and consumers randomly choose which product to sample first. This is supported by the consumer sampling the advertised product with probability $\alpha \leq \frac{1}{2}$ in the event he sees an ad, making advertising unprofitable.

To summarize:

1. In every equilibrium, we have that for every realization of qualities $\mathbf{q}$, the two firms spend the same amount on advertising, i.e. $m_1(\mathbf{q}) = m_2(\mathbf{q})$.

2. In equilibrium, therefore, the consumer's posterior beliefs about the qualities of the products are the same as his prior.

3. There is a continuum of equilibria based on how the customer breaks ties on which product to sample first. The tie breaking rule that maximizes advertising expenditure is one where the consumer deterministically first samples the product corresponding to the advertisement he saw.

4. Since advertising is uninformative, all equilibria have the same consumer welfare, differing only in the advertising expenditure.

4 Advertising on Online Social Networks

4.1 Display advertising

We begin our analysis focusing on display advertising on a social network. In particular, we are interested in how the baseline virality v_B interacts with the incentives of firms to advertise, the display advertising revenue of the platform and the overall social welfare. To focus on display advertising, we shut down social advertising, i.e. $v_S = 0$.

Our first theorem generalizes the findings of Section 3.2.

Theorem 1. *In every equilibrium both firms spend the same amount on display advertising, i.e.,* $m_1^d(\mathbf{q}) = m_2^d(\mathbf{q})$ *for every* $\mathbf{q} \in \mathbf{Q}$.

Hence display ads cannot be used by firms to signal quality, in any equilibrium. Effectively, therefore, advertising is a transfer from firms to the online platform, the amount depending on how consumers respond to ads. For example, consumers might sample first the observed product in which case the advertising revenue of the platform is maximized.

Conversely, consumers might sample first the product independent of the display ad they observe which leads to an advertising revenue of zero.

So how does baseline virality v_B influence advertising revenue? As we alluded to earlier, a first intuition suggests that display advertising might be more valuable in a high virality environment. Getting the early consumer to purchase can cause the good to go 'viral' since late consumers who see this purchase decision sample the observed product first and are less likely to engage in further costly search. Following this intuition, a higher baseline virality would induce a higher incentive to advertise which would lead to a higher advertising revenue. The following theorem shows that this intuition is incorrect.

THEOREM 2. *In every equilibrium expected social welfare is strictly increasing in the baseline virality. If the equilibrium advertising revenue is positive then the revenue is strictly decreasing with the baseline virality.*

To provide intuition for our result, consider a simple social environment with two consumers, i.e. consumer e moves first and consumer l sees e's purchase decision prior to his own search. Further suppose that the early consumer samples the product first for which he has seen an ad. Renaming firms if necessary, let us assume that firm 1 has the superior product, i.e. $q_1 > q_2$.

A consumer who does not observe the purchase decision of another consumer buys product 2 if and only if his search costs were high *and* he sampled 2 first, i.e. saw an ad for product 2. A late consumer l who observes the early consumer buys product 2 with the following probability

$$\mathbb{P}[a_e = 2 \,|q_1 > q_2] \,(1 - F_C(c_l(q_2)))\,.$$

Instead, if consumer l does not observe e he purchases product 2 with the identical probability as consumer e, i.e.

$$\begin{aligned}\mathbb{P}[a_l = 2 \,|q_1 > q_2, v_B = 0] &= \mathbb{P}[a_e = 2 \,|q_1 > q_2] \\ &> \mathbb{P}[a_e = 2 \,|q_1 > q_2] \,(1 - F_C(c_l(q_2)))\,.\end{aligned}$$

An increased baseline virality means that a larger fraction of the late movers will see an (informative) social signal about the product purchased by an early mover. These consumers will be uninfluenced by the display ad directly. Of course they may be influenced indirectly, i.e. the early mover may have been influenced by the display ad, who in turn influences this consumer. But this influence is imperfect—the late mover searches again if his search cost is low enough, which occurs with probability $F_C(c_l(q_2))$. Hence the influence of display advertising decreases as v_B increases, decreasing advertising revenue.

Next let us consider the welfare implications on display advertising. The welfare externality results from limiting, or omitting socially generated information. The intuition for the result can be seen from the discussion of the purchase decision of a late consumer in Section 3.1. There, we argued that the purchase decision of an early consumer is informative, since the early consumer searches among the products with positive probability. This remains true in the presence of advertising since both firms spend the same amount on advertising (Theorem 2). Therefore, the late consumer purchases the superior product with strictly higher probability when observing the decision of an early consumer than he would if he didn't observe this decision. Further, the late consumer spends less on search costs. As a result the net utility of a late customer is strictly higher when he observes an early mover's decision.

To conclude, for a fixed baseline virality display advertising is welfare neutral. However, the profit maximization objectives may induce the platform to reduce the baseline virality.

4.2 Combined Display & Social Advertising

In this section we consider the equilibria of the game when both display and social advertising is possible for firms. The first result echos those in the previous section with just display advertising, in that it shows that both firms spend the same amount on each form of advertising. Formally,

Theorem 3. *For any pair of network parameters (v_B, v_S), and any realized qualities $\mathbf{q} \in \mathbf{Q}$, both firms spend the same amount on each form of advertising, i.e.,*

$$\forall \mathbf{q} \in \mathbf{Q}, m_1^d(\mathbf{q}) = m_2^d(\mathbf{q}) \text{ and } m_1^s(\mathbf{q}) = m_2^s(\mathbf{q}).$$

In the introduction, we had suggested that a major concern was that an advertising revenue motivated social network may "bias" information towards the highest bidder. Our result shows that this concern is moot in our setting in equilibrium. As before this implies that in equilibrium, consumers do not infer anything from any display ad that they see, though this might influence their decision on how to break their indifference on which good to sample first. Further, consumers seeing social ads treat this equivalently to organic social information.

Given this result, the counterpart of Theorem 2 for the case of social welfare is apparent. Both users who see organic social information (with probability v_B), and users who see social ads (with probability $(1 - v_B)v_S$), receive the same information in equilibrium. Therefore we have the following easy corollary of Theorems 2 and 3:

Corollary 1. *In every equilibrium expected social welfare is strictly increasing in the fraction of people who see some social information, i.e. strictly increasing in $v_B + (1 - v_B)v_S$.*

The counterpart of Theorem 2 for revenue is less immediate. Social and display advertisements have revenues, and changing the parameters of the network, v_B and v_S, effects each. The theorem that follows details the comparative statics of these parameters on total advertising revenue.

Theorem 4. *For any pair of viralities (v_B, v_S) consider the equilibrium that maximizes the total advertising revenue. Advertising revenue is decreasing in v_B. For any $v_B < 1$, and any realized $\mathbf{q} = (q_1, q_2) \in \mathbf{Q}$ such that wlog $q_1 > q_2$, ad revenue is increasing in v_S if and only if*

$$\frac{1 + F_C(c_e(q_2))}{2 + F_C(c_e(q_2))} \geq F_C(c_l(q_2|a_e = 2)). \tag{2}$$

The corollary below follows trivially from Theorem 4:

Corollary 2. *If (2) is satisfied, then the social network's revenues are maximized at $(v_B, v_S) = (0, 1)$, otherwise they are maximized at $(v_B, v_S) = (0, 0)$.*

This bears some discussion. Firstly, note that a revenue motivated social network should always set the baseline virality v_B to 0. Organic social information competes with social ads (since the user sees only one or the other), and reduces the effectiveness of display advertisements.

The more interesting trade-off therefore is the choice of v_S—should the social network show "social ads," or simply surpress all social information and only show display ads to the consumers? The analysis shows that this trade-off reduces to (2). To interpret this, note that the left hand side is a (decreasing) function of $F_C(c_e(q_2))$, the fraction of consumers who would search further after sampling the inferior product first. The right hand side is exactly the fraction of late movers who would search further after seeing an early mover who had chosen the inferior product (and therefore sampled it first). The condition says that social ads are more lucrative than display ads if and only if the social information is sufficiently "convincing," i.e. the probability that a late mover samples further from the inferior product, after seeing an early mover purchase it, is sufficiently low.

5 Increasing the density of the social network

In our analysis we have so far assumed that each late mover observes the action of at most one early consumer. However, in observed real world social networks, users have, on average, a large number of friends. We therefore extend our results to consider a setting where late movers now observe k early movers—the preceding analysis therefore considers the special case where $k = 1$.

Formally, the advertising stage and early movers remain as previously. Each late consumer independently observes organic social information with probability v_B. Now the organic social information consists of the purchase decisions of k early consumers who are drawn uniformly, independently from the set of early consumers. In other words, the late mover sees k_1 customers who purchased product 1 and k_2 customers who purchased product 2, where k_1 is a draw from the Bernoulli distribution $B(k, \phi_1)$, where ϕ_1 is the fraction of early movers who purchased firm 1's product.

With probability $(1 - v_B)v_S$ the late consumer observes social advertising. These are the purchase decisions of k early consumers. As in the previous sections, instead of being drawn uniformly from the set of early consumers, each is drawn according to a advertising influenced distribution. In particular, the late mover sees k_1 customers who purchased product 1 and k_2 customers who purchased product 2, where k_1 is a draw from the Bernoulli distribution $B(k, p)$, with

$$p = \frac{\phi_1 m_1^s}{\phi_1 m_1^s + \phi_2 m_2^s},$$

where ϕ_i is the fraction of early movers who purchased firm i's product, and m_i^s is firm $i's$ expenditure on social ads.

Theorem 5. *For any pair of viralities (v_B, v_S) consider the equilibrium that maximizes the total advertising revenue. As the size of the social observation set k grows large, the social advertising expenditures converge to zero. Further, the display advertising revenues converge to those in a setting with only a mass of $1 + \lambda(1 - v_B)(1 - v_S)$ early movers.*

The basic intuition of this result is simple. Just as in the earlier setting, both firms will spend the same amount on each form of advertising. Therefore a consumer seeing social advertising will still be seeing unbiased information about the purchase decision of k early movers. If k is large, the product purchased by the majority of the early movers is the higher quality product with high probability. The late movers can therefore "free-ride" on this information, sampling the higher quality product first (with high probability) and rarely searching further. On the margin, therefore, advertising has no impact on the late movers' choices, and therefore firms do not spend on it. Clearly, however, social information is welfare improving for exactly the same reason. The following Corollary summarizes this.

Corollary 3. *For any realized qualities $\mathbf{q} \in \mathbf{Q}$, wlog such that $q_1 > q_2$, as the size of the* *social observation set k grows large, the expected net utility of any late mover who observes social information converges to the highest possible value of q_1.*

This theorem and corollary therefore summarize the central tension between a revenue-motivated social network and social welfare. A dense network is welfare improving—the free-riding late movers do not need to spend effort on search, and make better choices. However, the fact that they have so much information of early movers' choices leaves them uninfluenced by advertising. As a result, the social network may wish to limit the amount of information about early movers that late movers see.

6 Discussion and Conclusions

In this paper, we took a first step toward understanding the distortions that may arise when a social network is modeled as having its own commercial interests, rather than an inert conduit. We considered a simple model where agents may conduct costly sequential search to choose between competing products of unknown quality. Information on the social network is thus economically valuable: the choices made by predecessors is informative about the qualities of the products, potentially saving an agent search costs and preventing them from purchasing inferior products.

We considered two forms of advertising the social network may allow. The first, display advertising, is potentially valued by a firm because it may help a product go "viral," i.e. late moving agents may purchase the product purely based on observing that their friends have, rather than search on their own. However, we show that this intuition is not quite correct—advertising spend on display advertising actually decreases relative to a benchmark in which there is no social network.

The second form, social advertising, is motivated by advertising products recently offered by major online social networks (such as Facebook and Google), and allows a firm to highlight activity taken by a user (e.g. buying a product by that firm) to the users friends. We show that this may be a more effective type of advertising, since the fact that the user took the action is informative to other users.

Neither form of advertising directly impacts consumer welfare in our model. Advertising is solely a transfer from firms to the social network, with no resulting distortion. However, a social network focused on advertising revenues may want to limit the amount of information its users see about each others' activities. Users who see the choices of a lot of predecessors will perfectly discern which of the products is better, and therefore there will achieve first best welfare. However, these consumers also cannot be influenced by advertising, and therefore advertising revenues drop to zero. As a result, a social network may do better by limiting such information, so as to better monetize from advertising.

Similar concerns have also been present in search engines,[11] which has led to vigilant antitrust oversight. Such worries are more muted in the social networking space. This may partly be because large online social networks which have advertising as their core business model have emerged only recently. It may also be that the incentives we suggest are more subtle and less focal than those of search engines. However, recent worries voiced by several businesses who advertise on Facebook (the current largest social network) suggest that at the very least,[12] the results of this paper warrant further empirical investigation.

References

Athey, S., and G. Ellison (2011): "Position Auctions with Consumer Search," *The Quarterly Journal of Economics*, p. qjr028.

Banerjee, A. V. (1992): "A simple model of herd behavior," *The Quarterly Journal of Economics*, 107(3), 797–817.

Bergemann, D., and A. Bonatti (2011): "Targeting in advertising markets: implications for offline versus online media," *The RAND Journal of Economics*, 42(3), 417–443.

——— (2014): "Selling Cookies," Discussion paper, Yale University.

Bikhchandani, S., D. Hirshleifer, and I. Welch (1992): "A theory of fads, fashion, custom, and cultural change as informational cascades," *Journal of political Economy*, pp. 992–1026.

Bloch, F., and N. Quérou (2013): "Pricing in social networks," *Games and Economic Behavior*.

Brin, S., and L. Page (1998): "The anatomy of a large-scale hypertextual Web search engine," *Computer networks and ISDN systems*, 30(1), 107–117.

[11] Indeed the original paper describing the organization of Google, Brin and Page (1998), states " clear that a search engine which was taking money for showing cellular phone ads would have difficulty justifying the page that our system returned to its paying advertisers. For this type of reason ... we expect that advertising funded search engines will be inherently biased towards the advertisers and away from the needs of the consumers" (Appendix A).

[12] See e.g. the following article by entrepreneur Mark Cuban, `http://blogmaverick.com/2012/11/19/what-i-really-think-about-facebook/` or this article by food delivery service Eat 24 `http://blog.eat24hours.com/breakup-letter-to-facebook-from-eat24/`.

Cabral, L., D. J. Salant, and G. A. Woroch (1999): "Monopoly pricing with network externalities," *International Journal of Industrial Organization*, 17(2), 199–214.

Calvo-Armengol, A., and M. O. Jackson (2004): "The effects of social networks on employment and inequality," *The American Economic Review*, 94(3), 426–454.

Candogan, O., K. Bimpikis, and A. Ozdaglar (2012): "Optimal pricing in networks with externalities," *Operations Research*, 60(4), 883–905.

Chatterjee, K., and B. Dutta (2014): "Credibility and strategic learning in networks," Discussion paper, Penn State University.

Christakis, N. A., and J. H. Fowler (2007): "The spread of obesity in a large social network over 32 years," *New England journal of medicine*, 357(4), 370–379.

Edelman, B., M. Ostrovsky, and M. Schwarz (2005): "Internet advertising and the generalized second price auction: Selling billions of dollars worth of keywords," *American Economic Review*, 97.

Fainmesser, I. P., and A. Galeotti (2013): "The Value of Network Information," *Available at SSRN 2366077*.

Feldman, M., D. Kempe, B. Lucier, and R. Paes Leme (2013): "Pricing public goods for private sale," in *Proceedings of the fourteenth ACM conference on Electronic commerce*, pp. 417–434. ACM.

Friedman, L. (1958): "Game-theory models in the allocation of advertising expenditures," *Operations Research*, 6(5), 699–709.

Gomes, R. (2014): "Optimal auction design in two-sided markets," *The RAND Journal of Economics*, 45(2), 248–272.

Jackson, M. O. (2010): *Social and economic networks*. Princeton University Press.

Kircher, P., and A. Postlewaite (2008): "Strategic firms and endogenous consumer emulation," *The Quarterly Journal of Economics*, 123(2), 621–661.

Montgomery, J. D. (1991): "Social networks and labor-market outcomes: Toward an economic analysis," *The American economic review*, 81(5), 1408–1418.

Mueller-Frank, M., and M. M. Pai (2013): "Social Learning with Costly Search," Discussion paper, University of Pennsylvania.

Salganik, M. J., P. S. Dodds, and D. J. Watts (2006): "Experimental study of inequality and unpredictability in an artificial cultural market," *science*, 311(5762), 854–856.

Varian, H. R. (2007): "Position auctions," *International Journal of Industrial Organization*, 25(6), 1163–1178.

Weitzman, M. L. (1979): "Optimal search for the best alternative," *Econometrica: Journal of the Econometric Society*, pp. 641–654.

A Appendix

A.1 Proof of Theorem 1

Note that consumers only see an advertisement, not the amount spent on advertising by the firms. To solve for equilibria of this game, we therefore consider each possible strategy of consumers given the ad they see, and solve for the optimal advertising level for firms each $\mathbf{q}$ given these strategies. The equilibria then consist of all configurations where the consumer strategies are consistent with Bayesian updating given the derived advertising strategies of firms.

We will need some notation. First, denote by $\alpha_e \in [0,1]$ the fraction of early movers who sample first the product corresponding to the advertisement that they see. Similarly, denote by α_l^m the fraction of late movers who sample first the product corresponding to the display advertisement they see when they see an early mover's action as well, and the action and display advertisement are of the same product. Finally, denote by α_l^n the fraction of late movers who sample first the product corresponding to the display advertisement they see when they see an early mover's action as well, and the two do not match.

Fix a realized $\mathbf{q}$, wlog suppose that $q_1 > q_2$. Let us consider the incentives of firm 2—note that the only consumers who buy firm 2's product are those who sample it first *and* do not search any further. Therefore, an early mover buys firm 2's product with probability

$$\begin{aligned} P_e = \quad & \frac{m_2}{m_1+m_2}\alpha_e(1-F_C(c_e(q_2|\theta=2))) \\ & + \frac{m_1}{m_1+m_2}(1-\alpha_e)(1-F_C(c_e(q_2|\theta=1))). \end{aligned} \tag{3}$$

The first term on the right hand side corresponds to the probability that a consumer sees an ad for firm 2, samples that firm's product first, and has a search cost such that he does not want to search further. The second corresponds to the probability that a consumer sees an ad for firm 1 but samples 2's product first, and has a search cost such that he does not want to search further.

Note that given the baseline virality, there is an effective mass of $1 + \lambda(1 - v_B)$ of early movers.

Next, note that a late mover who observes an early mover's action has the following probability of purchasing firm 2's product

$$P_l = P_e(\frac{m_2}{m_1 + m_2}\tau_1 + \frac{m_1}{m_1 + m_2}\tau_2) + (1 - P_e)(\frac{m_2}{m_1 + m_2}\tau_3 + \frac{m_1}{m_1 + m_2}\tau_4), \tag{4}$$

where[13]

$$\begin{aligned}
\tau_1 &= \alpha_l^m(1 - F_C(c_l(q_2|\theta = 2, a_e = 2))), \\
\tau_2 &= (1 - \alpha_l^n)(1 - F_C(c_l(q_2|\theta = 1, a_e = 2))), \\
\tau_3 &= \alpha_l^n(1 - F_C(c_l(q_2|\theta = 2, a_e = 1))), \\
\tau_4 &= (1 - \alpha_l^m)(1 - F_C(c_l(q_2|\theta = 1, a_e = 1))).
\end{aligned}$$

Therefore the total profits of firm 2 are

$$\Pi_2 = P_e(1 + \lambda(1 - v_B)) + P_l\lambda v_B - m_2,$$

while the total profits of firm 1 are

$$\Pi_1 = 1 + \lambda - P_e(1 + \lambda(1 - v_B)) + P_l\lambda v_B - m_1.$$

Let us ignore the boundary condition that $m_i \geq 0$, and take first order conditions. The equilibrium advertising strategies therefore must satisfy, for each $i = 1, 2$.

$$\frac{\partial \Pi_i}{\partial m_i} = 0. \tag{5}$$

[13]The additional notation of the τ_i's is introduced purely for readability of the expressions.

Note that

$$\begin{aligned}
\frac{\partial P_e}{\partial m_2} =& \frac{m_1}{(m_1+m_2)^2}(\alpha_e(1-F_C(c_e(q_2|\theta=2)))-(1-\alpha_e)(1-F_C(c_e(q_2|\theta=1)))), \\
\equiv& \frac{m_1}{(m_1+m_2)^2}\tau_5, \qquad (6)\\
\frac{\partial P_e}{\partial m_1} =& -\frac{m_2}{(m_1+m_2)^2}(\alpha_e(1-F_C(c_e(q_2|\theta=2)))-(1-\alpha_e)(1-F_C(c_e(q_2|\theta=1)))), \\
\equiv& -\frac{m_2}{(m_1+m_2)^2}\tau_5. \qquad (7)
\end{aligned}$$

Further, note that

$$\begin{aligned}
\frac{\partial P_l}{\partial m_2} =& \frac{\partial P_e}{\partial m_2}\left((\frac{m_2}{m_1+m_2}\tau_1+\frac{m_1}{m_1+m_2}\tau_2)-(\frac{m_2}{m_1+m_2}\tau_3+\frac{m_1}{m_1+m_2}\tau_4)\right) \\
&+\frac{m_1}{(m_1+m_2)^2}\left(P_e(\tau_1-\tau_2)+(1-P_e)(\tau_3-\tau_4)\right) \\
\equiv& \frac{\partial P_e}{\partial m_2}\tau_6+\frac{m_1}{(m_1+m_2)^2}\tau_7 \\
\implies \frac{\partial P_l}{\partial m_2} =& \frac{m_1}{(m_1+m_2)^2}\left(\tau_5\tau_6+\tau_7\right). \qquad (8)
\end{aligned}$$

Similarly,

$$\begin{aligned}
\frac{\partial P_l}{\partial m_1} =& \frac{\partial P_e}{\partial m_1}\left((\frac{m_2}{m_1+m_2}\tau_1+\frac{m_1}{m_1+m_2}\tau_2)-(\frac{m_2}{m_1+m_2}\tau_3+\frac{m_1}{m_1+m_2}\tau_4)\right) \\
&-\frac{m_2}{(m_1+m_2)^2}\left(P_e(\tau_1-\tau_2)+(1-P_e)(\tau_3-\tau_4)\right) \\
\equiv& \frac{\partial P_e}{\partial m_1}\tau_6-\frac{m_2}{(m_1+m_2)^2}\tau_7 \\
\implies \frac{\partial P_l}{\partial m_1} =& -\frac{m_2}{(m_1+m_2)^2}\left(\tau_5\tau_6+\tau_7\right). \qquad (9)
\end{aligned}$$

Finally, substituting (6–9) into (5), we have the first order conditions below. For Firm 2 :

$$\begin{aligned}
&\frac{\partial P_e}{\partial m_2}(1+\lambda(1-v_B))+\lambda v_B\frac{\partial P_l}{\partial m_2}-1=0 \\
\implies &\frac{m_1}{(m_1+m_2)^2}\left((1+\lambda(1-v_B))\tau_5+\lambda v_B\left(\tau_5\tau_6+\tau_7\right)\right)=1 \qquad (10)
\end{aligned}$$

Similarly, for firm 1, we have:

$$\begin{aligned}
&-\frac{\partial P_e}{\partial m_1}(1+\lambda(1-v_B)) - \lambda v_B \frac{\partial P_l}{\partial m_1} - 1 = 0, \\
\implies &\frac{m_2}{(m_1+m_2)^2}\left((1+\lambda(1-v_B))\tau_5 + \lambda v_B\left(\tau_5\tau_6+\tau_7\right)\right) = 1 \qquad (11)
\end{aligned}$$

Finally, by observation, (10) and (11) can only have an interior solution if $m_1 = m_2$. So far we have ignored the boundary conditions. Again by observation of (10) and (11), the boundary condition either binds for both or neither.

Therefore in any equilibrium, we must have that $m_1^d(\mathbf{q}) = m_2^d(\mathbf{q})$ for every $\mathbf{q} \in \mathbf{Q}$.

Let us continue to fully describe the equilibria of this model. Since both firms advertise the same amount at every $\mathbf{Q}$, the consumer does not infer anything from the fact that he sees an advertisement. Therefore we have that:

$$c_e(\cdot|\theta=2) = c_e(\cdot|\theta=1) = c_e(\cdot),$$

where $c_e(\cdot)$ is as described in (1), i.e. the choice on whether to sample further does not depend on the ad the consumer saw. Again, since the display ad contains no information, as established in Section 3.1, a late consumer who observes an early consumer samples first the product chosen by the early consumer. Therefore $\alpha_l^m = 1$ and $\alpha_l^n = 0$. Finally, we have that,

$$c_l(\cdot|a_e=i,\theta=2) = c_l(\cdot|a_e=i,\theta=1) = c_l(\cdot|a_e=i),$$

i.e., again that the choice on whether to sample further does not depend on the ad the consumer saw.

Substituting this into the respective definitions of the τ_i's, we have that:

$$\begin{aligned}
&\tau_1 = \tau_2 = \tau_6 = (1-F_C(c_l(q_2|a_e=2))), \\
&\tau_3 = \tau_4 = \tau_7 = 0, \\
&\tau_5 = (2\alpha_e - 1)(1-F_C(c_e(q_2))).
\end{aligned}$$

Substituting these into (10), we have:

$$\frac{m_1}{(m_1+m_2)^2}\left(1+\lambda(1-v_B)+\lambda v_B(1-F_C(c_l(q_2|a_e=2)))\right)(2\alpha_e-1)(1-F_C(c_e(q_2))) = 1$$

Since we have already concluded that $m_1 = m_2$,

$$(1 + \lambda(1 - v_B) + \lambda v_B(1 - F_C(c_l(q_2|a_e = 2))))\,(2\alpha_e - 1)(1 - F_C(c_e(q_2))) = 4m,$$

Finally for an interior solution to be possible, it must be that $2\alpha_e - 1 > 0$, i.e. $\alpha_e > \frac{1}{2}$.

Therefore, there are a continuum of equilibria. In any equilibrium, we have that $m_1^d(\mathbf{q}) = m_2^d(\mathbf{q})$ for every $\mathbf{q} \in \mathbf{Q}$. There is one equilibrium where $m_1^d(\mathbf{q}) = m_2^d(\mathbf{q}) = 0$, i.e. no firm spends anything on advertising. Consumers do not see any ads and randomly choose which firm to sample first. If an ad was seen, consumers would sample first the product corresponding to that ad with probability $\alpha_e \leq \frac{1}{2}$, making a deviation to advertising unprofitable.

Further, there is a continuum of equilibria indexed by $\alpha_e \in (\frac{1}{2}, 1]$, where at any $\mathbf{q}$, each firm spends

$$m_i^d(\mathbf{q}) = \frac{1}{4}\,(1 + \lambda(1 - v_B) + \lambda v_B(1 - F_C(c_l(q_2|a_e = 2))))\,(2\alpha_e - 1)(1 - F_C(c_e(q_2))) \quad (12)$$

on advertising. □

A.2 Proof of Theorem 2

Revenue Recall from (12), we have that in any equilibrium with positive advertising expenditure, we have

$$m_i^d(\mathbf{q}) = \frac{1}{4}\,(1 + \lambda(1 - v_B) + \lambda v_B(1 - F_C(c_l(q_2|a_e = 2))))\,(2\alpha_e - 1)(1 - F_C(c_e(q_2))),$$

for some $\alpha_e \in (\frac{1}{2}, 1]$. Taking a partial derivative with respect to v_B, we have the desired result since $(1 - F_C(c_l(q_2|a_e = 2))) < 1$.

Welfare The early stage welfare is unaffected by the baseline virality. Hence we restrict attention to the expected welfare of late movers. Without loss of generality, let 2 be the inferior product. For a given late mover who observes the action of a early mover, his expected utility is:

$$\begin{aligned} U_o =& P_e(1 - F_C(c_l(q_2|a_e = 2)))q_2 + P_e F_C(c_l(q_2|a_e = 2))(q_1 - E[c|c < c_l(q_2|a_e = 2)]) \\ &+ (1 - P_e)(q_1 - E[c|c < c_l(q_1|a_e = 1)]) \end{aligned}$$

where P_e is the probability that the early mover purchases the inferior product. For a late mover who does not observe an early mover's product, the expected utility is:

$$\begin{aligned}U_n =& \frac{1}{2}(1 - F_C(c_e(q_2)))q_2 + \frac{1}{2}F_C(c_l(q_2))(q_1 - E[c|c < c_e(q_2)]) \\ &+ \frac{1}{2}(q_1 - E[c|c < c_e(q_1)])\end{aligned}$$

Note that since $P_e < \frac{1}{2}$, and $c_e(\cdot) > c_l(\cdot)$, we have that $U_o > U_n$. Therefore the net social welfare is increasing in v_B, i.e. the fraction of users who observe an early mover's action.

A.3 Proof of Theorem 3

We now need repeat the arguments in the Proof of Theorem 1 with a few more cases. In interests of brevity, we borrow notation and arguments from there.

In particular, there, we had set $v_S = 0$, so any late mover not seeing any organic social information is akin to an early mover. Here, this will no longer be the case: some late movers may see social ads. So, additionally, define by α_s^m (respectively α_s^n) the probability with which a late mover observing a social ad and display ad for the same (respectively, different) products samples first the product corresponding to the display ad he sees.

Fix as before a realized $\mathbf{q}$ with $q_1 > q_2$. P_e remains as defined in (3), similarly P_l remains as defined in (4). Additionally, we must now consider users who only see a social ad. These purchase firm 2's product with probability

$$\begin{aligned}P_s = & \frac{P_e m_2^s}{P_e m_2^s + (1-P_e)m_1^s}\left(\frac{m_2}{m_1+m_2}\gamma_1 + \frac{m_1}{m_1+m_2}\gamma_2\right) \\ &+ \frac{(1-P_e)m_1^s}{P_e m_2^s + (1-P_e)m_1^s}\left(\frac{m_2}{m_1+m_2}\gamma_3 + \frac{m_1}{m_1+m_2}\gamma_4\right).\end{aligned}$$

where,

$$\begin{aligned}\gamma_1 &= \alpha_s^m(1 - F_C(c_l(q_2|\theta = 2, a_s = 2))), \\ \gamma_2 &= (1-\alpha_s^n)(1 - F_C(c_l(q_2|\theta = 1, a_s = 2))), \\ \gamma_3 &= \alpha_s^n(1 - F_C(c_l(q_2|\theta = 2, a_s = 1))), \\ \gamma_4 &= (1-\alpha_s^m)(1 - F_C(c_l(q_2|\theta = 1, a_s = 1))).\end{aligned}$$

Therefore, the total revenues of firm 2 are

$$R_2 = P_e(1 + \lambda(1-v_B)(1-v_S)) + P_s\lambda(1-v_B)v_S + P_l\lambda v_B,$$

resulting in profits to firm 2 of

$$\Pi_1 = R_2 - m_2 - m_2^s,$$

while the total profits of firm 1 are

$$\Pi_1 = 1 + \lambda - R_2 - m_1 - m_1^s.$$

Taking first order conditions with respect to m_1 and m_2, and collecting terms, we have, for firm 2:

$$\frac{m_1}{(m_1+m_2)^2}\left((1+\lambda(1-v_B)(1-v_S))\tau_5 + \lambda(1-v_B)v_S(\tau_5\tau_8+\tau_9)\lambda v_B\left(\tau_5\tau_6+\tau_7\right)\right) = 1$$

where

$$\tau_8 = \frac{\partial \frac{P_e m_2^s}{P_e m_2^s + (1-P_e)m_1^s}}{\partial P_e}\left(\left(\frac{m_2}{m_1+m_2}\gamma_1 + \frac{m_1}{m_1+m_2}\gamma_2\right) - \left(\frac{m_2}{m_1+m_2}\gamma_3 + \frac{m_1}{m_1+m_2}\gamma_4\right)\right),$$
$$\tau_9 = \frac{P_e m_2^s}{P_e m_2^s + (1-P_e)m_1^s}(\gamma_1-\gamma_2) + \frac{(1-P_e)m_1^s}{P_e m_2^s + (1-P_e)m_1^s}(\gamma_3-\gamma_4).$$

Similarly, for firm 1 we have the first order condition:

$$\frac{m_2}{(m_1+m_2)^2}\left((1+\lambda(1-v_B)(1-v_S))\tau_5 + \lambda(1-v_B)v_S(\tau_5\tau_8+\tau_9)\lambda v_B\left(\tau_5\tau_6+\tau_7\right)\right) = 1.$$

Therefore, once again, $m_1 = m_2$.

We are now left to derive expenditure on social ads. Taking first order conditions for Π_2 with respect to m_2^s we have,

$$\frac{P_e(1-P_e)m_1^s}{(P_e m_2^s + (1-P_e)m_1^s)^2}\left(\left(\frac{m_2}{m_1+m_2}\gamma_1 + \frac{m_1}{m_1+m_2}\gamma_2\right) - \left(\frac{m_2}{m_1+m_2}\gamma_3 + \frac{m_1}{m_1+m_2}\gamma_4\right)\right) = \kappa,$$

while similarly for Π_1 with respect to m_1 we have,

$$\frac{P_e(1-P_e)m_2^s}{(P_e m_2^s + (1-P_e)m_1^s)^2}\left(\left(\frac{m_2}{m_1+m_2}\gamma_1 + \frac{m_1}{m_1+m_2}\gamma_2\right) - \left(\frac{m_2}{m_1+m_2}\gamma_3 + \frac{m_1}{m_1+m_2}\gamma_4\right)\right) = \kappa,$$

where $\kappa = \frac{1}{\lambda(1-v_B)v_S}$. Once again, by observation, it must be that $m_1^s = m_2^s$.

It follows that social ads are sampled from the same distribution as as organic information, so by Bayesian updating, it must be that $\alpha_s^m = 1$, and further that $\alpha_s^n = 0$, and $\gamma_1 = \gamma_2 = \tau_1 (= \tau_2)$.

Finally, to index the equilibria, in terms of display advertising, the set of equilibria is similar to those identified in the proof of Theorem 1. There is the zero-advertising equilibrium as before, but also a continuum of equilibria indexed by $\alpha_e \in (\frac{1}{2}, 1]$ (the probability with which a consumer who does not see any social information or ad samples the product corresponding to the ad he sees first), where at any $\mathbf{q}$ each firm spends

$$m_i^d(\mathbf{q}) = \frac{1}{4}\left(1 + \lambda(1 - v_B)(1 - v_S) + \lambda\left(v_B + (1 - v_B)v_S\right)\left(1 - F_C(c_l(q_2|a_e = 2))\right)\right)$$
$$(2\alpha_e - 1)(1 - F_C(c_e(q_2))), \tag{13}$$
$$m_i^s(\mathbf{q}) = \lambda(1 - v_B)v_S\phi(1 - \phi)(1 - F_C(c_l(q_2|a_e = 2))), \tag{14}$$

on advertising, where $\phi = \frac{1}{2}(1 - F_C(c_e(q_2)))$ is the equilibrium probability that an early mover buys the inferior product. □

A.4 Proof of Theorem 4

Recall that (13, 14) describe the expenditure on display and social ads respectively in any equilibrium, indexed by $\alpha_e \in [\frac{1}{2}, 1]$. Since the display on social ads is independent of α_e, and the display advertising expenditure is increasing in α_e, the total expenditure on ads is maximized at $\alpha_e = 1$. The expenditures at any $\mathbf{q} \in \mathbf{Q}$ are given by

$$\begin{aligned}
\overline{m}_i^d(\mathbf{q}) &= \frac{1}{4}\left(1 + \lambda(1 - v_B)(1 - v_S) + \lambda\left(v_B + (1 - v_B)v_S\right)\left(1 - F_C(c_l(q_2|a_e = 2))\right)\right)\left(1 - F_C(c_e(q_2))\right) \\
&= \frac{1}{2}\phi\left(1 + \lambda(1 - v_B)(1 - v_S) + \lambda\left(v_B + (1 - v_B)v_S\right)\left(1 - F_C(c_l(q_2|a_e = 2))\right)\right) \\
\overline{m}_i^s(\mathbf{q}) &= \lambda(1 - v_B)v_S\phi(1 - \phi)(1 - F_C(c_l(q_2|a_e = 2))),
\end{aligned}$$

where $\phi = \frac{1}{2}(1 - F_C(c_e(q_2)))$ is the equilibrium probability that an early mover buys the inferior product. The total advertising expenditure of each firm therefore equals

$$\frac{1}{2}\phi\left(1 + \lambda(1 - v_B)(1 - v_S) + \lambda(v_B + (1 - v_B)v_S(3 - 2\phi))(1 - F_C(c_l(q_2|a_e = 2)))\right). \tag{15}$$

Taking a partial derivative w.r.t. v_S, (15) is increasing in v_S if and only if

$$\begin{aligned}
&\frac{1}{2}\phi\lambda(1-v_B)\left((3-2\phi))(1-F_C(c_l(q_2|a_e=2)))-1\right)\geq 0\\
\iff\ &(3-2\phi))(1-F_C(c_l(q_2|a_e=2)))\geq 1\\
\iff\ &(2+F_C(c_e(q_2)))(1-F_C(c_l(q_2|a_e=2)))\geq 1\\
\iff\ &(1+F_C(c_e(q_2)))\geq F_C(c_l(q_2|a_e=2))(2+F_C(c_e(q_2)))\\
\iff\ &\frac{1+F_C(c_e(q_2))}{2+F_C(c_e(q_2))}\geq F_C(c_l(q_2|a_e=2)).
\end{aligned}$$

Similarly, (15) is decreasing in v_B if and only if

$$\begin{aligned}
&\frac{1}{2}\phi\lambda\Big(-(1-v_S)+(1-v_S(3-2\phi))(1-F_C(c_l(q_2|a_e=2)))\Big)\leq 0\\
\iff\ &-(1-v_S)+(1-v_S(3-2\phi))(1-F_C(c_l(q_2|a_e=2)))\leq 0\\
\iff\ &\frac{1-v_S(3-2\phi)}{1-v_S}(1-F_C(c_l(q_2|a_e=2)))\leq 1
\end{aligned}$$

Note that since $\phi\in[0,1]$, $3-2\phi\geq 1$. Therefore we have that $\frac{1-v_S(3-2\phi)}{1-v_S}\leq 1$. Further, by observation $(1-F_C(c_l(q_2|a_e=2)))\in[0,1]$. Therefore this inequality is always satisifed. and (15) is always descreasing in v_B.

A.5 Proof of Theorem 5

First, we will show that in this setting as well, both firms will spend the same amount on each type of advertising.

So once again, let α_e be the probability an early mover samples the product corresponding to the display ad he saw first. Similarly, let α_l^j be the probability that a late mover who sees organic social information samples the product corresponding to the display ad he saw first when he also sees exactly $j\leq k$ early movers took the same action. Finally, let α_s^j be the probability that a late mover who sees social advertising samples the the product corresponding to the display ad he saw first when he also sees exactly $j\leq k$ early movers took the same. The search cutoffs are similarly denoted $c_e(\cdot|\theta)$,$c_l(\cdot|\theta,j)$, and $c_s(\cdot|\theta,j)$.

Once again fix realized qualities $\mathbf{q}\in\mathbf{Q}$, wlog such that $q_1>q_2$. Fix the firms' display advertising levels m_i^d and social advertising levels m_i^s.

First note that the probability an early mover purchases the inferior product is given by:

$$\phi_2=\frac{m_2^d}{m_1^d+m_2^d}\alpha_e(1-F_C(c_e(q_2|\theta=2)))+\frac{m_1^d}{m_1^d+m_2^d}(1-\alpha_e)(1-F_C(c_e(q_2|\theta=1))).$$

Recall that the probability of a randomly chosen social ad being for the inferior product 2 given the early movers' purchase decisions and the given social advertising expenditures is:

$$p = \frac{\phi_2 m_2^s}{\phi_1 m_1^s + \phi_2 m_2^s}.$$

Finally, note that the revenues of firm 2 can be written as:

$$\begin{aligned} R_2 = {} & (1 + \lambda(1 - v_B)(1 - v_S))\phi_2 \\ & + \lambda(1 - v_B)v_S \sum_{j=0}^{k} B(k, j, p) \left(\alpha_s^j (1 - F_C(c_s(q_2|\theta = 2, j))) + (1 - \alpha_s^{k-j})(1 - F_C(c_s(q_2|\theta = 1, k - j)))\right) \\ & + \lambda v_B \sum_{j=0}^{k} B(k, j, \phi_2) \left(\alpha_l^j (1 - F_C(c_l(q_2|\theta = 2, j))) + (1 - \alpha_l^{k-j})(1 - F_C(c_l(q_2|\theta = 1, k - j)))\right). \end{aligned}$$

The profits of each firm therefore are:

$$\begin{aligned} \Pi_2 &= R_2 - m_2^d - m_2^s, \\ \Pi_1 &= 1 + \lambda - R_2 - m_1^d - m_1^s. \end{aligned}$$

Taking first order conditions and collecting terms, we have that $m_1^d = m_2^d$ and $m_1^s = m_2^s$ as before.

Recall that the first order conditions with respect to m_2^s are:

$$\begin{aligned} & \frac{\partial p}{\partial m_2^s} \sum_{j=0}^{k} \frac{\partial B(k, j, p)}{\partial p} \left(\alpha_s^j (1 - F_C(c_s(q_2|\theta = 2, j))) + (1 - \alpha_s^{k-j})(1 - F_C(c_s(q_2|\theta = 1, k - j)))\right) \\ & = \frac{1}{\lambda(1 - v_B)v_S} \end{aligned}$$

Recalling that $m_1^x = m_2^x$ we further have that $p = \phi$. Plugging this in, we have:

$$\frac{\partial p}{\partial m_2^s} = \frac{\phi_1 \phi_2 m_1^s}{(\phi_1 m_1^s + \phi_2 m_2^s)^2} = \frac{\phi_1 \phi_2}{m_2^s},$$

which implies

$$\begin{aligned} & \sum_{j=0}^{k} \frac{\partial B(k, j, \phi_2)}{\partial \phi_2} \left(\alpha_s^j (1 - F_C(c_s(q_2|\theta = 2, j))) + (1 - \alpha_s^{k-j})(1 - F_C(c_s(q_2|\theta = 1, k - j)))\right) \\ & = \frac{m_2^s}{\lambda(1 - v_B)v_S \phi_1 \phi_2} \end{aligned}$$

Next, since social ads are unbiased, note that for any $j > \frac{k}{2}$, $\alpha_s^j = \alpha_l^j = 1$, i.e. the late consumers sample first the product chosen by the majority of the early movers they saw. So our FOC can be written as:

$$\sum_{\lceil \frac{k}{2} \rceil}^{k} \frac{\partial B(k,j,\phi_2)}{\partial \phi_1}(1 - F_C(c_s(q_2|\theta = 2, j))) = \frac{m_2^s}{\lambda(1-v_B)v_S\phi_1\phi_2}.$$

Since $1 - F_C \leq 1$, we have

$$\sum_{\lceil \frac{k}{2} \rceil}^{k} \frac{\partial B(k,j,\phi_2)}{\partial \phi_2} \geq \frac{m_2^s}{\lambda(1-v_B)v_S\phi_1\phi_2}.$$

Differentiating and collecting terms, the left hand side equals

$$\begin{aligned} &k\phi_2^k + k(k-1)\sum_{\lceil \frac{k}{2} \rceil}^{k-1} B(k-2, j-1, \phi_2)\left(\frac{1-\phi_2}{k-j} + \frac{\phi_2}{j}\right) \\ \leq &k\phi_2^k + k(k-1)\sum_{\lceil \frac{k}{2} \rceil}^{k-1} B(k-2, j-1, \phi_2) \end{aligned}$$

By an application of Hoeffding's Inequality, we know that

$$\begin{aligned} \sum_{\lceil \frac{k}{2} \rceil}^{k-1} B(k-2, j-1, \phi_2) &\leq \exp\left(-2\frac{(k(1-\phi_2) - \frac{k}{2})^2}{k}\right) \\ &= \exp\left(-2\left(\frac{1}{2} - \phi_2\right)k\right) \end{aligned}$$

Therefore, collecting, we have that

$$\frac{m_2^s}{\lambda(1-v_B)v_S\phi_1\phi_2} \leq k\phi_2^k + k(k-1)\exp\left(-2\left(\frac{1}{2} - \phi_2\right)k\right),$$

implying that $m_2^s \to 0$ as $k \to \infty$.

Similarly the FOC with respect to m_2^d, which requires the following expression to equal

1.

$$\frac{\partial \phi_2}{\partial m_2^d}\Bigg((1+\lambda(1-v_B)(1-v_S))\phi_2$$
$$+\lambda(1-v_B)v_S\sum_{j=0}^{k}\frac{\partial B(k,j,\phi_2)}{\partial \phi_2}\left(\alpha_s^j(1-F_C(c_s(q_2|\theta=2,j)))+(1-\alpha_s^{k-j})(1-F_C(c_s(q_2|\theta=1,k-j)))\right)$$
$$+\lambda v_B\sum_{j=0}^{k}\frac{\partial B(k,j,\phi_2)}{\partial \phi_2}\left(\alpha_l^j(1-F_C(c_l(q_2|\theta=2,j)))+(1-\alpha_l^{k-j})(1-F_C(c_l(q_2|\theta=1,k-j)))\right)\Bigg).$$

Repeating the arguments above, we can show that the latter two term terms tend to 0 as k grows large. Therefore, the FOC reducces to

$$\frac{\partial \phi_2}{\partial m_2^d}(1+\lambda(1-v_B)(1-v_S))\phi_2 = 1,$$

which is identical to a setting with $(1+\lambda(1-v_B)(1-v_S))$ mass of early movers. $\square$