Strategic Use of Product Delays to Shape Word-of-Mouth Communication

Alexei Parakhonyak* and Nick Vikander†

March 2023

Abstract

This paper investigates the advantages a seller can gain by strategically creating product scarcity to manipulate consumer word-of-mouth communication. The seller offers a product of uncertain quality and establishes a service speed that determines whether opinion leaders are immediately served or delayed when attempting to purchase the product. These opinion leaders subsequently share their experiences with other consumers, influencing their beliefs about product quality and their purchase decisions. We show that delaying opinion leaders can significantly impact consumer learning by altering both the content and level of word-of-mouth communication. Specifically, the content effect alone can incentivize the seller to delay opinion leaders, except in cases where private information is highly accurate and expected product quality is moderate. In situations where information about purchased products spreads more easily than information about pre-orders, the level effect limits the potential for delay, particularly when expected product quality is high.

Keywords: word of mouth, product delays, scarcity, capacity constraints

JEL codes: D83, L15

*University of Oxford, Department of Economics and Lincoln College. E-mail: aleksei.parakhonyak@economics.ox.ac.uk. Address: Lincoln College, Turl Street, OX1 3DR, Oxford, UK.

†University of Copenhagen, Department of Economics. E-mail: nick.vikander@econ.ku.dk. Address: Øster Farimagsgade 5, Building 26, 1353 K, Copenhagen, Denmark.

We gratefully acknowledge funding from the Independent Research Fund Denmark (grant no. 8019-00075B).

1 Introduction

This paper considers how a seller may profitably shape consumer 'buzz' about its product by strategically creating product scarcity. The idea is that initial scarcity limits the number of opinion leaders who can make early purchases. As such, it can influence both *what* leaders can communicate about the product and *how much* they communicate to other consumers in the market. We explore how creating initial product scarcity can either help or hurt the seller by affecting subsequent sales, via this word-of-mouth channel.

In practice, consumers can communicate in many ways, such as face-to-face conversation, social media posts, online discussion groups, and product reviews. An established literature has shown how word-of-mouth communication can affect many consumer decision, including television viewing (Godes and Mayzlin, 2004), book purchases (Chevalier and Mayzlin, 2006), movie viewing (Liu, 2006; Duan et al., 2008), even adoption of microfinance (Banerjee et al., 2013), and much more.

Word of mouth is a valuable tool that enables consumers to learn about product quality. For example, when considering purchasing a mobile phone, individuals may seek out information from acquaintances who have ordered or received the same model. Additionally, online product reviews can provide insights from other consumers that aid in the decision-making process. By leveraging these sources, individuals can effectively learn from the experiences of others before making a purchase decision.

Given the power of word of mouth, it is no wonder that firms may try to strategically influence consumer communication. This can be through so-called 'buzz marketing', where digital advertising aims to help get people taking and build the buzz around a particular product (Mohr, 2017). This can also be via pricing, as in frequent zero-price sales for smartphone apps to attract new consumers, who will then spread awareness to others (Ajorlou et al., 2018). Firms may also want to shape 'pre-release buzz' by influencing how consumers talk before receiving the product (Xiong and Bharadwaj, 2014). They may even go so far as to directly manipulate consumer communication, by posting fake product reviews on platforms such as TripAdvisor (Mayzlin et al., 2014), Yelp (Luca and Zervas, 2016), and Amazon (He et al., 2022).

Our paper considers a novel way a seller may strategically influence word of mouth to shape consumer beliefs about product quality: by using scarcity to delay the purchase of opinion leaders. Delays and stocks outs are common in practice. Prominent examples include the Nintendo Switch in 2020[1], Tesla in 2021[2] and the Sony PlayStation 5 over a similar period.[3] While delays can certainly arise for many reasons, such as unexpected production problems, our focus is on how delay can allow a seller to shape consumer buzz.

We show that delaying opinion leaders can sometime be profitable, precisely because of its impact on word-of-mouth communication. We also describe how the profitability of delay depends on factors such as ex ante beliefs about quality, the precision of consumers' private information, and the relative number of opinion leaders in the market.

More specifically, we present a two-period model, where a seller first faces a cohort of opinion leaders and then a cohort of followers who can potentially talk to the leaders. All parties are initially unsure whether product quality is high or low and each consumer receives a boundedly informative private signal. The seller chooses a service speed which influences how quickly leaders can be served, after which leaders arrive and simultaneously choose whether to place orders. All demand from leaders up to the service speed is immediately served while all other leaders who place orders are delayed. Each of these leaders can then tell some followers about their experience; whether they were delayed or served, and in the latter case, their utility from buying. Followers then simultaneously choose whether to buy. Those that buy receive the product, as do leaders who were delayed.[4]

We first consider a baseline case, where leaders talk to the same number of followers regardless of whether they are immediately served or delayed. Delaying leaders therefore does not affect the overall level of word-of-mouth communication but only its informational content. Rather than telling followers hard information about product quality, delayed leaders can only convey soft information about their intention to buy and the fact that they could not get the product immediately. It is via this *content effect* that a lower service speed reduces learning.

We show that delaying leaders can be strictly profitable. That is, for certain parameter

[1]See https://www.forbes.com/sites/davidthier/2020/04/19/the-real-reason-nintendo-switch-is-out-of-stock-everywhere/?sh=4e08af9d5694, accessed on March 13 2023.

[2]See https://www.cnbc.com/2021/08/18/months-long-delivery-delays-confound-would-be-tesla-owners.html, accessed on March 13 2023.

[3]See https://www.npr.org/2023/01/05/1147157065/sony-playstation-5-shortage-over, accessed on March 13 2023.

[4]Assuming delayed leaders also eventually receive the product allows us to isolate the informational impact of delay on profits. Reducing service speed then does not directly cost sales but only affects word of mouth. As such, in the absence of word of mouth or quality uncertainty, the seller would be indifferent between immediately serving or delaying leaders.

values, there exists a pure strategy equilibrium where all leaders are delayed. Perhaps surprisingly, the seller will hide information via delay whenever product quality is likely high. Meeting a delayed leader is good news because demand is higher for a high-quality product. When prior beliefs about quality are high, this good news may be enough to convince a follower to buy regardless of their private signal. The seller may also use delay when product quality is likely low, but only if signals are not too accurate. Delaying leaders then only ends up paying off when quality turn out to be low, and demand for a low-quality product is decreasing in signal precision.

In contrast, no pure strategy equilibrium exists in which leaders are immediately served. If followers expect a high service speed, then meeting a delayed leader is particularly convincing, since followers reason that demand must exceed the high service speed. This inference allows the seller to mislead followers by unexpectedly reducing service speed to delay all leaders. By the same token, it is never strictly profitable in equilibrium to immediately serve leaders rather than delay them. However, in large markets there exists a mixed strategy equilibrium where the seller is indifferent and serves all leaders almost surely, and where followers who meet delayed leaders randomize between always buying and acting on their private signal.[5]

We then consider a more general case, where immediately-served leaders talk to more followers than leaders who are delayed.[6] Reducing service speed still affects the informational content of word of mouth for followers who receive news. Now, it also affect how many followers receive any news at all. This *level effect* on word of mouth reduces learning, as did the content effect, but with different implications for seller behavior.

We show that the level effect on word of mouth makes it less attractive for the seller to use delay, compared to the baseline. Meeting a delayed leader still constitutes good news, but when few delayed leaders talk, then this good news is not received by many followers. This can potentially make it profitable for the seller to immediately serve leaders to maximize the amount of word off mouth. More specifically, we show that a pure strategy equilibrium sometimes exists where the all leaders are served. The less that

[5]Our baseline results are robust to changes in the amount of word of mouth that leaders engage in, or in the number of leaders relative to followers. These results are also unchanged if we assume that word of mouth not only transmits utility information but also spreads product awareness, i.e. if we assume that only followers who hear from leaders are able to buy.

[6]Arguably, consumers are more likely to share information about the product they actually have, than about the fact that they put a pre-order. Such information sharing, e.g. can take the form of a product review.

delayed leaders talk, the larger the parameter region for which serving all leaders is an equilibrium. Nonetheless, the seller will still delay leaders in equilibrium when expected quality is moderate low and signals are relatively inaccurate.

The level effect can also give equilibrium multiplicity, as both a pure strategy equilibrium where the seller delays all leaders, and a pure strategy equilibrium where all leaders are immediately served, can exist for the same parameter values. Our analysis suggest that equilibrium multiplicity is more of an issue in markets with many opinion leaders. In such markets, the behavior of followers who *do not* hear from leaders is particularly sensitive to their beliefs about the service speed, as not hearing about the product when there are many leaders is a strong signal of low demand, and hence low quality. This can affect which service speed the seller finds most profitable.

Our paper's main contribution is to show how a seller can profitably use product scarcity to influence word-of-mouth communication. In so doing, we add to the literature looking at the interaction between consumer word of mouth and seller strategic behavior.[7] By directly modeling consumer communication, our paper differs from the strand of this literature that directly assumes local interaction effects, for which one possible interpretation is word of mouth. See, e.g., Galeotti and Goyal (2009), Galeotti et al. (2020). Our focus on quality uncertainty also differs from work where word-of-mouth communication simply informs consumers about product existence, where different papers explore the interaction between word of mouth and pricing decisions (Campbell (2015), Ajorlou et al. (2018)), information release (Campbell et al., 2017), as well as both pricing and advertising (Campbell, 2013).[8]

Relatively few papers consider how word-of-mouth communication may help consumers learn about product quality. Godes (2017) assume quality is endogenous and focuses on firm incentives to invest in quality. They show that more word of mouth that transmits utility information is associated with higher equilibrium quality. Campbell et al. (2020) assume consumers engage in costly search to learn product quality from earlier buyers. They focus on how the network structure of communication affects the distribution of quality in the market, via high- and low- quality firms' entry and exit. Our

[7]The broader literature on word-of-mouth communications considers how agents can learn from hearing about outcomes from earlier agents in different settings See, e.g., Ellison and Fudenberg (1995), Bala and Goyal (1998), Banerjee and Fudenberg (2004). The focus has often been on whether agents will converge on the same action in the long run and whether this action is efficient.

[8]Galeotti (2010) instead assumes consumers know about product existence but communicate about prices, and shows how this affects equilibrium price dispersion.

modelling approach instead assume that quality is exogenous, and we explore how the seller's strategic choice of scarcity can affect both the informational content and the level of word of mouth.

Our focus on consumer word of mouth and product quality also has some relation to the literature on customer reviews. A major focus there has been on how firms may directly manipulate consumers via fake reviews. In terms of theory, Mayzlin (2006) shows that firm fake reviews can make consumer communication less persuasive, by making it less credible. Relatedly, Smirnov and Starkov (2022) show that firm ability to censor bad reviews can affect the informational content of the bad reviews that do end up appearing in equilibrium.[9] We also look at how a seller may benefit from strategically reducing the informational content of consumer communication, but via product scarcity rather than deception, and without being privately informed about product quality.

By looking at whether a seller prefers to serve consumers sequentially (first leaders, then followers), or effectively to serve all consumers simultaneously (by delaying leaders), our paper also contributes to work on optimal product launch strategies. There the distinction is often made between a sequential 'waterfall' strategy to promote learning, and a simultaneously 'sprinkler' strategy to restrict it. See, e.g. Sgroi (2002); Aoyagi (2010); Liu and Schiraldi (2012); Bhalla (2013); Parakhonyak and Vikander (2019). Whereas these papers consider observational learning, where consumers learn from observing each others choices, we explicitly model word-of-mouth communication, by which consumers can learn each others' utility information. Moreover, serving consumers simultaneously in this literature essentially shuts down any information transmission. This contrasts with our paper, as delayed leaders still talk to followers, which in turn contributes to making delay more attractive for the seller.

Our paper also contributes to the literature on firms' strategic use of product scarcity to influence consumer behavior. Previous work on scarcity has looked at discouraging consumer strategic delay (DeGraba (1995); Nocke and Peitz (2007); Möller and Watanabe (2010)) as well as facilitating price discrimination (Wilson (1988); Bulow and Roberts (1989); Ferguson (1994); Loertscher and Muir (2022)). A few papers in this literature share our focus on how scarcity can affect consumer learning about product quality, but via different mechanisms. Consumers learn either via firm signaling (Stock and Balachander, 2005), observational learning from each others' choice (Parakhonyak and Vikander, 2023) or a

[9] Hauser (2023) look at censorship but without explicitly considering consumer communication. There, censorship reduces the arrival rate of potential bad news, which can affect firm incentives to invest in quality.

combination of the two (Debo et al., 2012). In contrast, we look at how scarcity affects how consumers learn via both the content and the level of word-of-mouth communication.

The rest of the paper is organised as follows. Section 2 introduces the model. Section 3 contains all main results of the paper, with the baseline case, corresponding to an equal amount of word of mouth for served and delayed leaders discussed in section 3.1, and the general case discussed in 3.2. Section 4 considers an extension where followers only become aware of the product if they hear about it through word of mouth. Section 5 concludes.

2 Setting

Suppose there is a product of unknown quality and two possible states of the world, $\Omega = \{G, B\}$. In state G, quality is good and each consumer who buys obtains $u_G = 1$. In state B, quality is bad and each consumer who buys obtains $u_B = 0$. A consumer who does not buy gets reservation utility $r \in (0, 1)$. The actual state is known neither to the seller nor to consumers. Prior beliefs of all players are that $P(G) \equiv \beta$ and $P(B) = 1 - \beta$, where $\beta \in (0, 1)$.

There are two cohorts of potential buyers who make their decisions in different stages. We refer to the first cohort as opinion leaders, or simply leaders, and the second cohort as followers. The number of leaders is N whereas the number of followers is nN, with $n, N \in \mathbb{N}$.

Each potential buyer has unit demand and learns about quality based on a noisy private signal, $s \in \{g, b\}$, where $P(g|G) = P(b|B) \equiv \alpha \in (1/2, 1)$. By $\alpha < 1$, signals are *boundedly informative* about the state. Additionally, followers can potentially learn by hearing from leaders, as described more precisely below. We focus on situations where $P(G|s = g) > r > P(G|s = b)$. This means that in the absence of further information, it is optimal for a potential buyer to follow their own private signal.

The timing of the game is as follows. At $t = -1$, the the seller sets a service speed $K \in \mathbb{Z}^+$, with $K \leq N$. This speed will influence how quickly leaders can receive the product. Then, at $t = 0$, nature chooses product quality, $\omega \in \{G, B\}$. The rest of the game consists of three stages.

In stage $t = 1$, the leaders enter the market, each receives their private signal, and they simultaneously decide whether to order the product. If the total quantity ordered is less than K, then all leaders who ordered receive the product immediately. If instead total orders exceed K, then K randomly chosen leaders receive the product immediately, whereas the remaining $N - K$ leaders do not. We will say that the former group of leaders are served immediately and that the latter group are delayed.

In stage $t = 2$, the followers enter the market, each receives their private signal, and each may also learn from a leader via word-of-mouth communication. That is, each follower connects with at most one leader who ordered the product and hears about that leader's experience in stage 1: whether they were delayed or served immediately, and in the latter case, their utility from receiving the product. Each leader who is served, and each leader who is delayed, connects with m_1 and m_0 followers, respectively, where $m_1 + m_0 \leq n$, so that $N(m_1 + m_0)$ followers in total learn via word of mouth. The remaining $N(n - m_1 - m_0)$ followers remain unconnected and do not hear from anybody. The parameters m_1 and m_0 capture the effectiveness of word-of-mouth. We assume $m_1 \geq m_0$ to capture the idea that leaders who immediately receive the product, and discover its quality, plausibly should not communicate less than leaders who have yet to receive the product because they are delayed.[10]

Finally, in stage $t = 3$, followers simultaneously decide whether to buy. All followers who want to buy the product receive it, as do all leaders who were previously delayed.

We normalize the per consumer profit of the seller to 1. The seller sets service speed K so as to maximize expected profits, given the subsequent behavior of consumers. Each consumer makes the purchase decision that is optimal, given their beliefs about product quality. These beliefs about quality follow from Bayes' rule and the other players' equilibrium strategies, whenever possible.

We will need to characterise beliefs when the seller is expected to immediately serve all leaders in stage 1, but a follower nonetheless meets a delayed leader in stage 2. Our approach is to allow for seller 'trembles', where the implemented service speed, with small probability, can differ slightly from the seller's profit-maximizing choice of K, in order to pin down these out-of-equilibrium beliefs. Specifically, if the seller chooses $K \geq 1$, then we assume that the implemented service speed is $K - 1$ with probability ϵ and K with probability $1 - \epsilon$, where $\epsilon > 0$ can be arbitrarily small.

Our analysis will focus on the limiting case of large markets, $N \to \infty$. We therefore

[10]We could also allow for each leader who did not order to connect with say $m' > 0$ followers and communicate their own private signal. Doing so will not qualitatively change our results, as long as such leaders do not communicate more broadly than leaders who ordered: $m' \leq m_0 \leq m_1$.

normalize both service speed and profits by the total market size, where the seller sets service speed $k \equiv K/N$ to maximizes expected profits per consumer. This service speed $k \in [0, 1]$ should be thought of as the limiting case of the profit-maximizing normalized service speed K/N in a finite market, as N become large.[11]

There is no discounting between stages. Moreover, unless otherwise stated, we assume that followers observe neither the seller's choice of service speed K nor the total number of leaders who are immediately observed.

3 Analysis

Since leaders follow their private signals, the probabilities that j leaders order the product in the good and bad states are, respectively,

$$Q_G(j) = \binom{N}{j}\alpha^j(1-\alpha)^{N-j}, \quad Q_B(j) = \binom{N}{j}(1-\alpha)^j\alpha^{N-j}. \tag{1}$$

Followers who meet a served leader will learn the state by hearing the leader's utility, and therefore only buy if the state is good. The decision of other followers may depend on what they learn from meeting a delayed leader or remaining unconnected, as well as their own private signal.

As a first step to write out the seller profit function, we will assume a follower buys with probability γ_ω if they meet a delayed leader, and buys with probability δ_ω if they remain unconnected. These probabilities may depend on the state $\omega \in \{G, B\}$, since they take into account both possible realizations of the follower's private signal $s \in \{g, b\}$, and followers with different signals may take different decisions. We later derive what values γ_ω and δ_ω should take in equilibrium, given follower beliefs.

Denote the number of leaders served in state ω as $S_\omega(K)$, and the number of leaders delayed as $D_\omega(K)$, where

$$S_\omega(K) = \sum_{j=0}^{N} \min\{j, K\}Q_\omega(j), \quad D_\omega(K) = \sum_{j=0}^{N} \max\{j-K, 0\}Q_\omega(j).$$

Profits from setting service speed $K \geq 1$, normalized by market size N, are then

$$\pi(K) = (1-\varepsilon)\tilde{\pi}(K) + \varepsilon\tilde{\pi}(K-1)$$

[11]Note that the assumption of 'trembles' has no direct effect on profits per consumer in a large market.

where

$$\tilde{\pi}(K) = \frac{\beta}{N}\left[\sum_{j=0}^{N} jQ_G(j) + m_1 S_G(K) + m_0\gamma_G D_G(K) + \delta_G\left(nN - m_1 S_G(K) - m_0 D_G(K)\right)\right]$$
$$+ \frac{1-\beta}{N}\left[\sum_{j=0}^{N} jQ_B(j) + 0 + m_0\gamma_B D_B(K) + \delta_B\left(nN - m_1 S_B(K) - m_0 D_B(K)\right)\right], \qquad (2)$$

and we define profits from setting service speed $K = 0$ as $\pi(0) = \tilde{\pi}(0)$. In each square bracket in expression (2), the first term refers to the number of leaders who order the product and therefore receive it either in stage 1 (if served immediately) or in stage 2 (if delayed); the second term refers to followers who meet immediately-served leaders and therefore learn the state; the third term refers to receivers who meet delayed leaders; and the fourth term refers to followers who remain unconnected.

We can now state the following result.

Lemma 1. *For any market size N and any $\epsilon > 0$, the seller maximizes expected profits by setting service speed $K = 0$ or $K = N$, or by randomizing between these two values.*

Lemma 1 implies in particular that any equilibrium will involve the seller either delaying all leaders, immediately serving all leaders, or possibly randomizing between the two. Intuitively, a small increase in service speed can help the seller when the state is good and hurt when the state is bad, as more immediately-served leaders may then reveal the state to followers. However, this only happens when demand is sufficiently high (e.g if demand is very low then nobody is delayed even at a low service speed). High demand is more likely in the good state, so whenever a small increase in service speed helps the seller in expectation, then a larger increase will help the seller even more, by revealing information for those demand realizations that are even more indicative of the good state. As such, the seller does best either to immediately serve all leaders or none at all.

We now turn to follower beliefs, in order to derive the purchase probabilities γ_G, γ_B, δ_G, and δ_B in the expression for seller profits.

Meeting a delayed leader provides positive information about the state, since a delayed leader must have received a good signal (otherwise they would not have ordered the product); and good signals are more likely in the good state. The question is whether this positive information is enough to convince a follower to buy after receiving a bad private signal. Suppose that the seller sets service speed $K = N$ with probability q and sets $K = 0$ with probability $(1 - q)$.[12] The expected number of delayed leaders in state ω is then

$$D_\omega(q) = (1 - q) \sum_{j=0}^{N} j Q_\omega(j) + q \varepsilon Q_\omega(N).$$

In particular, if the seller sets $K = 0$ for sure then all consumers are delayed. If the seller sets $K = N$ then one follower is delayed with probability ε, which is the probability that service speed $N - 1$ is implemented due to a tremble. Using Bayes' rule, the belief of a follower with $s = b$ who met a delayed leader is therefore

$$\mu(q, N) \equiv \frac{\beta(1-\alpha)D_G(q)}{\beta(1-\alpha)D_G(q) + (1-\beta)\alpha D_B(q)}. \tag{3}$$

The properties of these beliefs are described in the following Lemma.

Lemma 2. *In any equilibrium, the belief of a follower with a bad private signal who met a delayed leader satisfies the following properties:*

1. $\lim_{N\to\infty} \lim_{\varepsilon\to 0} \mu(1, N) = 1$.
2. *For any fixed* $q < 1$, $\lim_{N\to\infty} \lim_{\varepsilon\to 0} \mu(q, N) = \beta$.
3. *For any* $r > \beta$ *there exists* $\overline{N}(r)$ *such that for any* $N > \overline{N}(r)$ *there exists a unique* $q^*(N)$ *which solves* $\mu(q^*(N), N) = r$. *Moreover,* $\lim_{N\to\infty} \lim_{\varepsilon\to 0} q^*(N) = 1$.

The first part of Lemma 2 states that if followers expects the seller to immediately serve all leaders, then meeting a delayed leader could only result from the combination of high demand and a tremble, which convinces followers that the state is good.[13] The second part states that delaying leaders with any fixed probability, including playing a pure strategy $K = 0$, would lead followers with bad private signals who meet delayed leaders to revert to their prior beliefs, because the good private signal of the delayed leader effectively cancels the bad signal of the follower. Finally, the third part implies that there is a candidate equilibrium in which both the seller and followers may randomize, but as markets grow large, the seller's strategy converges to non-restricting service speed with probability one.

[12] Lemma 1 shows that we can restrict attention to service speeds $K = 0$ and $K = N$.
[13] This result in fact holds for any market size N, not just in the limit.

We now turn our attention to followers who did not meet any leaders, either served or delayed. Not meeting a leader is bad news, because there is higher demand in the first stage when $\omega = G$. Therefore, followers with bad private signals who remain unconnected prefer not to buy. Now consider the belief of an unconnected follower with a good private signal,

$$\nu(q, N) = \frac{\beta\alpha U_G(q, N)}{\beta\alpha U_G(q, N) + (1-\beta)(1-\alpha)U_B(q, N)},$$

where U_ω is the expected number of unconnected followers in state ω, given by

$$U_\omega(q, N) = (1-q)\left(nN - m_0\sum_{j=0}^{N} jQ_\omega(j)\right) + q\left(nN - m_1\sum_{j=0}^{N} jQ_\omega(j) - \varepsilon(m_0 - m_1)Q_\omega(N)\right).$$

Unconnected followers with good private signals will buy the product as long as $\nu(q, N) \geq r$. Solving for $\lim_{N\to\infty}\lim_{\varepsilon\to 0}\nu(q, N) = r$ allows us to formulate the following Lemma.

Lemma 3. *In a large market, a follower with a good private signal who did not meet a leader will buy the product if and only if* $\beta \geq \overline{\beta}_q(r, \alpha)$ *where*

$$\overline{\beta}_q(r, \alpha) \equiv \frac{(1-\alpha)\big(n - [(1-q)m_0 + qm_1](1-\alpha)\big)r}{(1-\alpha)\big(n - [(1-q)m_0 + qm_1](1-\alpha)\big)r + \alpha\big(n - [(1-q)m_0 + qm_1]\alpha\big)(1-r)}. \tag{4}$$

We can now also take the limit $N \to \infty$ in expression (2) to obtain the normalized profit function for large markets

$$\begin{aligned}
\pi(k) =& \beta\,[\alpha + \min\{\alpha, k\}m_1 + \max\{\alpha - k, 0\}m_0\gamma_G + \\
& (n - \min\{\alpha, k\}m_1 - \max\{\alpha - k, 0\}m_0)\delta_G] + \\
& (1-\beta)\,[1 - \alpha + \max\{1 - \alpha - k, 0\}m_0\gamma_B + \\
& (n - \min\{1-\alpha, k\}m_1 - \max\{1 - \alpha - k, 0\}m_0)\delta_B]\,,
\end{aligned} \tag{5}$$

where γ_ω is the purchase probability of a follower who met a delayed leader buys (consistent with Lemma 2), and δ_ω is the purchase probability that a follower is buys after remaining unconnected (consistent with Lemma 3), conditional on the state being ω.[14]

[14]Expression (5) also gives expected seller profits per consumer, taking into account the possibility of trembles. Given that the market is large we obtain $\lim_{N\to\infty}\tilde{\pi}(K) = \lim_{N\to\infty}\tilde{\pi}(K-1) = \lim_{N\to\infty}\pi(K)$.

Notice that although the limit profit function $\pi(k)$ is constant on $k \equiv \frac{K}{N} \in (\alpha, 1]$, setting service speed $K \in [1, N-1]$ is never optimal for any finite market size N. Taking the limit of finite markets, we either have that $\lim_{N\to\infty} \frac{1}{N} \arg\max \pi(K) = N$ or $\lim_{N\to\infty} \frac{1}{N} \arg\max \pi(K) = 0$, from Lemma 1. Thus we use $k = 1$ as the profit maximizer of expression (5) over $(\alpha, 1]$ in what follows.

3.1 Baseline case: $m_1 = m_0 = m$

Having derived the large-market profit function and described consumer beliefs, we are in a position to derive the seller's equilibrium choice of service speed.

We start with the baseline where $m_1 = m_0 = m$. All leaders who place orders then talk to the same number of followers, regardless of whether they themselves are immediately served or delayed. The seller's choice of service speed therefore does not affect the total level of word of mouth but only its informational content.

Before proceeding, observe that the critical value of β for an unconnected follower to follow their private signal simplifies to

$$\overline{\beta}_q(r, \alpha)|_{m_1=m_0=m} = \frac{(1-\alpha)(n - m(1-\alpha))r}{(1-\alpha)(n-m(1-\alpha))r + \alpha(n - m\alpha)(1-r)} \equiv \overline{\beta}(r, \alpha) \qquad (6)$$

in the baseline. Note that $\overline{\beta}(r, \alpha)$ now does not depend on q and that $\overline{\beta}(r, \alpha) < r$ for all values of $\alpha \in (1/2, 1)$.

Consider a candidate equilibrium with $k = 0$. From Lemma 2, the posterior belief of a follower with a bad private signal who met a delayed leader is just β. Therefore,

$$\gamma_G = \begin{cases} \alpha, & \beta < r \\ 1, & \beta \geq r \end{cases}, \quad \gamma_B = \begin{cases} 1-\alpha, & \beta < r \\ 1, & \beta \geq r \end{cases} \qquad (7)$$

From Lemma 3 and equation (6) we obtain the purchase probabilities of unconnected followers

$$\delta_G = \begin{cases} 0, & \beta < \overline{\beta}(r, \alpha) \\ \alpha, & \beta \geq \overline{\beta}(r, \alpha) \end{cases}, \quad \delta_B = \begin{cases} 0, & \beta < \overline{\beta}(r, \alpha) \\ 1-\alpha, & \beta \geq \overline{\beta}(r, \alpha). \end{cases} \qquad (8)$$

We can now directly substitute these purchase probabilities into the expression for seller profits. For $\beta \geq r$, expression (5) reduces to

$$\begin{aligned}\pi(k) =& \beta\left[\alpha(m+1) + (n-\alpha m)\alpha\right] + \\ & (1-\beta)\left[(1-\alpha) + \max\{1-\alpha-k, 0\}m + (n-(1-\alpha)m)(1-\alpha)\right],\end{aligned}$$

which is decreasing in k. Thus, $k = 0$ is an equilibrium when $\beta \geq r$.

For $\beta \in [\overline{\beta}(r,\alpha), r)$, the profit function reduces to

$$\begin{aligned}\pi(k) =& \beta\left[\alpha + \min\{\alpha, k\}m + \max\{\alpha-k, 0\}m\alpha + (n-\alpha m)\alpha\right] + \\ & (1-\beta)\left[1-\alpha + \max\{1-\alpha-k, 0\}m(1-\alpha) + (n-(1-\alpha)m)(1-\alpha)\right].\end{aligned}$$

Now consider a deviation from $k = 0$ to $k = 1$.[15] This deviation is profitable if and only if

$$\beta\alpha^2 m + (1-\beta)(1-\alpha)^2 m < \beta\alpha m,$$

which simplifies to $\beta > 1-\alpha$. Thus, $k = 0$ is an equilibrium for $\beta \in [\overline{\beta}(r,\alpha), r)$ as long as $\beta < 1-\alpha$.

Finally, for $\beta < \overline{\beta}(r,\alpha)$, the profit function reduces to

$$\pi(k) = \beta\left[\alpha + \min\{\alpha, k\}m + \max\{\alpha-k, 0\}m\alpha\right] + (1-\beta)\left[1-\alpha + \max\{1-\alpha-k, 0\}m(1-\alpha)\right],$$

which again is increasing in k if an only if $\beta > 1-\alpha$. Taken together, we can we can conclude that $k = 0$ in an equilibrium whenever $\beta \geq r$, and whenever both $\beta < r$ and $\beta < 1-\alpha$.

Intuitively, restricting reducing service speed affects seller profits by influencing the informational content of consumer word of mouth. Delayed leaders cannot transmit hard information to followers that the state is good or bad, and instead just effectively reveal that they themselves had good private signals and could not get the product immediately. As such, delaying leaders maximizes the number of followers who receive positive information via word of mouth, but also limits how convincing this information will be.

It pays off to restrict service speed when the good state is likely (i.e. $\beta \geq r$) since meeting delayed leaders is then convincing enough to induce followers to buy regardless of the followers' private signals. In contrast, when the good state is less likely (i.e. $\beta < r$),

[15] We can restrict attention to this deviation by Lemma 1.

followers who meet delayed leaders just act on their own signal. Delaying leaders is then a double-edged sword: hiding hard information via delay helps the seller if the state turns out to be bad but hurts the seller if the state turns out to good. When signal precision is high, more (fewer) leaders want to buy in the good (bad) state, so reducing then service speed tends to hide hard information whose revelation would help (hurt) the seller.

Now consider a candidate equilibrium with $k = 1$. The purchase probabilities of unconnected followers, δ_G and δ_B, are still given by (8). From Lemma 1, the posterior belief of followers who meet delayed leaders equals to 1, so these followers will buy regardless of their private signal: $\gamma_G = \gamma_B = 1$.

For $\beta \geq \overline{\beta}(r, \alpha)$, substituting these purchase probabilities into expression (5) for seller profits gives

$$\begin{aligned}\pi(k) =& \beta \left[\alpha(m+1) + (n - \alpha m)\alpha\right] + \\ & (1-\beta)\left[(1-\alpha) + \max\{1-\alpha-k, 0\}m + (n-(1-\alpha)m)(1-\alpha)\right].\end{aligned}$$

Doing the same for $\beta < \overline{\beta}(r, \alpha)$ gives

$$\pi(k) = \beta \left[\alpha(m+1)\right] + (1-\beta)\left[(1-\alpha) + \max\{1-\alpha-k, 0\}m\right].$$

Since both these profit expressions are decreasing in k, we can rule out a pure strategy equilibrium with $k = 1$ for any parameter values.

Intuitively, if followers expect a high service speed, then nonetheless meeting a delayed leader is interpreted as a result of tremble and gives compelling evidence of high demand and hence the good state. This evidence will induce followers to buy even if the prior is low and even if they have bad private signals. Thus, the seller can profitably deviate by reducing service speed, to delay as many leaders as possible.

The results so far, together with Lemma 1, imply that no pure strategy equilibrium exists for $1 - \alpha < \beta < r$. We now derive a mixed strategy equilibrium, in which the seller randomises between $k = 0$ and $k = 1$, and in which followers who meet delayed leaders randomize between buying and not buying.

Lemma 2 showed that for any finite market size N large enough, there exists a mixed strategy of the seller $q^*(N)$ that leaves followers who meet delayed leader indifferent about buying. We now verify that whenever $1 - \alpha < \beta < r$, there exists a mixed strategy for these followers that leaves the seller is indifferent between setting $k = 0$ and $k = 1$. From

(5), the seller indifference condition is[16]

$$\begin{aligned}\pi(0) =& \beta[\alpha + \alpha m\gamma_G + (n - \alpha m)\delta_G] + (1-\beta)[1-\alpha + (1-\alpha)m\gamma_B + (n-(1-\alpha)m)\delta_B] \\ &= \beta[\alpha + \alpha m + (n - \alpha m)\delta_G] + (1-\beta)[1-\alpha + (n-(1-\alpha)m\delta_B] = \pi(1),\end{aligned}$$

which simplifies to

$$\beta\alpha\gamma_G + (1-\beta)(1-\alpha)\gamma_B = \beta\alpha.$$

Recall that γ_G and γ_B denote the ex ante purchase probabilities of followers who meet delayed leaders, depending on the state. These followers who receive good private signals will buy; meeting a delayed leader provides good news, so their posterior beliefs exceed $P(G|g) > r$. If these unconnected followers who receive bad private signals buy with probability $\chi \in [0, 1]$, then $\gamma_G = \alpha + (1-\alpha)\chi$ and $\gamma_B = (1-\alpha) + \alpha\chi$. Substituting into the seller indifference condition yields

$$\beta\alpha[\alpha + (1-\alpha)\chi] + (1-\beta)(1-\alpha)[(1-\alpha) + \alpha\chi] = \beta\alpha.$$

The right-hand side of this condition is independent of χ, whereas the left-hand side is increasing in χ and equals $\beta\alpha^2 + (1-\beta)(1-\alpha)^2$ when $\chi = 0$ and $\beta\alpha + (1-\beta)(1-\alpha) > \beta\alpha$ when $\chi = 1$. Thus, whenever $\beta\alpha^2 + (1-\beta)(1-\alpha)^2 < \beta\alpha$, or equivalently $\beta > 1-\alpha$, there exists a value $\chi^* \in (0, 1)$ that leaves the seller indifferent. If followers with bad signals who meet delayed leaders then buy with probability χ^*, then the seller is willing to mix between $k = 0$ and $k = 1$.

It follows that a mixed strategy equilibrium exists for $1-\alpha < \beta < r$. Since $q^*(N) \to 1$ by Lemma 2, we can conclude that the seller sets a service speed equal of $k = 1$ almost surely in this equilibrium, whereas followers who meet delayed leader randomize with a probability strictly greater than 0 and strictly less than 1.

The difference between a pure strategy equilibrium with $k = 1$, which never exists in the baseline, and a mixed strategy equilibrium with $q^*(N) \to 1$, which exists for $1-\alpha < \beta < r$, relates to follower beliefs. In the mixed strategy equilibrium, followers who meet delayed leaders understand that either the seller set $k = 0$, in which case meeting a delayed leader

[16] We work directly with profits expressions for the limiting case of large markets, $N \to \infty$. It is easy to show that for any finite N, profit expressions for both $\pi(K = 0)$ and $\pi(K = N)$ are continuous in γ_ω and hence there exists a unique pair $1-\alpha < \gamma_B < \alpha < \gamma_G$ that solves $\pi(K = 0) = \pi(K = 1)$. Thus, the mixed strategy equilibrium corresponding to the limit of this solution exists and is unique.

is relatively likely; or the seller set $k = 1$, in which case meeting a delayed leader is very unlikely (due to a combination of high demand and a tremble). Indeed, in the limit $N \to \infty$, the probability of meeting a delayed leader when the seller set $k = 1$ becomes vanishingly small. Followers would therefore update their beliefs with full weight on $k = 0$ (i.e. posterior equal to β), unless the seller's mixed strategy called on it to set $k = 1$ almost surely. In contrast, in a candidate pure strategy equilibrium with $k = 1$, followers update beliefs with full weight on $k = 1$ (i.e. posterior equal to 1), even though the equilibrium probability of meeting a delayed leader given $k = 1$ is vanishingly small.

We can summarize the results of this section in the following Proposition.

Proposition 1. *Suppose that $m_1 = m_0$, so that immediately-served leaders and delayed leaders engage in the same amount of word of mouth. Then:*

1. *If $\beta \leq 1 - \alpha$ or $\beta \geq r$, a unique equilibrium exists and is in pure strategies, where the seller sets $k = 0$ and delays all leaders.*

2. *If $1 - \alpha < \beta < r$, a unique equilibrium exists and is in mixed strategies, where the seller almost surely sets $k = 1$ immediately serves all leaders.*

We can represent these results in the following figure, for parameter values $r = 0.2$, $m = 3$, $n = 4$. Signal precision α is depicted on the horizontal axis and the prior β on the vertical axis.

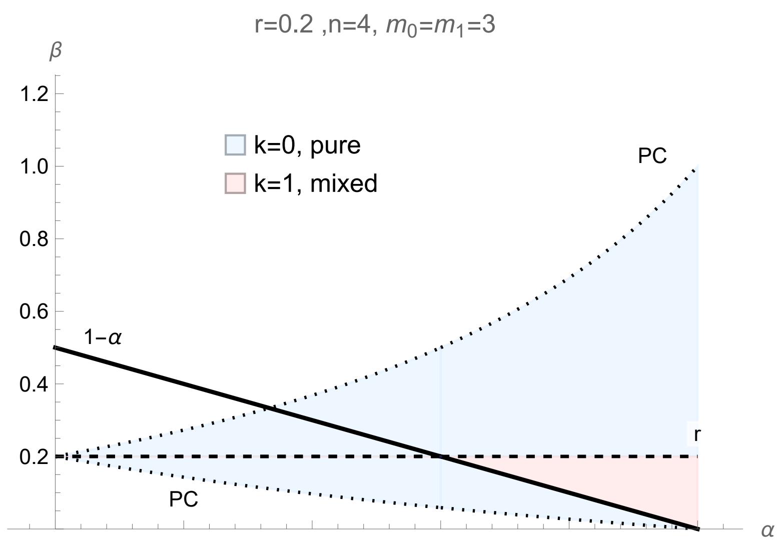


Figure 1: Optimal Service Speed

The region between the two dashed curves shows where leaders follow their own signals. The light blue part of this region shows where there is a pure strategy equilibrium with $k = 0$, whereas the light red triangle shows where there is a mixed strategy equilibrium where the seller sets $k = 1$ with probability 1.

We conclude this section by noting that the seller's equilibrium choice of service speed does not depend on the number of leaders relative to followers in the market (the parameter n) or on the number of followers that each leader connects with (the parameter m). That is, the fact that leaders engage in word of mouth matters for seller strategic behavior, but the exact number of followers who learn via word of mouth does not. An important reason is that service speed does not affect the total level of word mouth when all leaders who place orders communicate to the same extent. The seller may want to delay leaders, or immediately serve them, depending on whether it makes sense on balance to hide or reveal hard information to followers who end up connecting with leaders. More word of mouth just means that there are more of these followers who can be influenced in this way.

3.2 General Case

In this section, we consider a general case with $m_1 \geq m_0$, so where immediately-served leaders may talk to more followers than leaders who are delayed. We try to utilize the intuition from the analysis of our baseline. In particular, by Lemma 1, we can again restrict attention to candidate equilibria where the seller sets service speed $k = 0$, $k = 1$, or possibly randomizes between the two, and we can also restrict attention to $k = 0$ and $k = 1$ when looking at potential deviations.

We again start by considering a candidate equilibrium with $k = 0$. The purchase probabilities γ_G and γ_B, for a follower who meets a delayed leader, are still given by equation (7). The purchase probabilities δ_G and δ_B, for an unconnected follower, are similar to (8), but where m_0 now replaces m in the expression for the cutoff $\overline{\beta}(r, \alpha)$. Thus, we have

$$\gamma_G = \begin{cases} \alpha, & \beta < r \\ 1, & \beta \geq r \end{cases}, \quad \gamma_B = \begin{cases} 1-\alpha, & \beta < r \\ 1, & \beta \geq r \end{cases}$$

$$\delta_G = \begin{cases} 0, & \beta < \overline{\beta}_0(r, \alpha) \\ \alpha, & \beta \geq \overline{\beta}_0(r, \alpha) \end{cases}, \quad \delta_B = \begin{cases} 0, & \beta < \overline{\beta}_0(r, \alpha) \\ 1-\alpha, & \beta \geq \overline{\beta}_0(r, \alpha) \end{cases}$$

It is straightforward to verify that $n > m_0$ implies that $\overline{\beta}_0(r,\alpha) < r$, so we can sequentially consider three cases: when $\beta \geq r$, when $\beta \in [\overline{\beta}_0(r,\alpha), r)$ and when $\beta < \overline{\beta}_0(r,\alpha)$.

Suppose that $\beta \geq r$, which implies $\gamma_G = \gamma_B = 1$, $\delta_G = \alpha$, $\delta_B = 1-\alpha$.

Plugging these purchase probabilities into expression (5) for profits, then evaluating at $k = 0$, gives

$$\pi(0) = \beta\left[\alpha(m_0+1) + (n-\alpha m_0)\alpha\right] + (1-\beta)\left[(1-\alpha)(m_0+1) + (n-(1-\alpha)m_0)(1-\alpha)\right].$$

Doing the same but evaluating at $k = 1$ yields deviation profits

$$\pi(1) = \beta\left[\alpha(m_1+1) + (n-\alpha m_1)\alpha\right] + (1-\beta)\left[1-\alpha + (n-(1-\alpha)m_1)(1-\alpha)\right].$$

A direct comparison shows that we can rule out a profitable deviation to $k = 1$ when

$$\beta \leq \hat{\beta}(\alpha) \equiv 1 - \frac{m_1 - m_0}{m_1}\alpha. \tag{9}$$

This condition is trivially satisfied for all (α,β) when $m_1 = m_0$, consistent with our results in the baseline. If $m_1 > m_0$, then condition (9) reflects the following trade-off. Immediately serving all leaders helps the seller when the state turns out to be good, since more followers learn via word of mouth than if the seller had restricted service speed. These followers all buy but would have followed their private signals had they remained unconnected. In contrast, delaying all leaders helps the seller when the state turns out to be bad, as followers who meet delayed leaders do not learn the state and instead are all induced to buy. The seller therefore wants to immediately serve leaders and maximize the number of followers who learn when the good state is sufficiently likely.

Suppose now that $\beta \in [\overline{\beta}_0(r,\alpha), r)$, which implies $\gamma_G = \delta_G = \alpha$, $\gamma_B = \delta_B = 1-\alpha$. That is, followers who do not meet immediately-served leaders just follow their private signals. Profits from setting $k = 0$ are then

$$\pi(0) = \beta\left[\alpha + \alpha^2 m_0 + (n-\alpha m_0)\alpha\right] + (1-\beta)\left[1-\alpha + (1-\alpha)^2 m_0 + (n-(1-\alpha)m_0)(1-\alpha)\right].$$

whereas a deviation to $k = 1$ yields

$$\pi(1) = \beta\left[\alpha + \alpha m_1 + (n-\alpha m_1)\alpha\right] + (1-\beta)\left[1-\alpha + (n-(1-\alpha)m_1)(1-\alpha)\right].$$

Thus, we can rule out any profitable deviation when $\beta \leq 1-\alpha$.

Finally, suppose $\beta < \overline{\beta}_0(r,\alpha)$, which implies $\gamma_G = \alpha,\ \gamma_B = 1-\alpha$, and $\delta_G = \delta_B = 0$. Profits from setting $k=0$ are then

$$\pi(0) = \beta\left[\alpha + \alpha^2 m_0\right] + (1-\beta)\left[1-\alpha+(1-\alpha)^2 m_0\right],$$

while profits from a deviation to $k=1$ are

$$\pi(1) = \beta\left[\alpha + \alpha m_1\right] + (1-\beta)(1-\alpha).$$

It follows that we can rule out any profitable deviation whenever

$$\beta \leq \frac{(1-\alpha)^2 m_0}{(1-\alpha)m_0 + (m_1-m_0)\alpha}. \tag{10}$$

We denote

$$\tilde{\beta}(\alpha) = \frac{(1-\alpha)m_0}{(1-\alpha)m_0 + (m_1-m_0)\alpha}.$$

Using this notation, we sum up all these cases to get the following result.

Lemma 4. *An equilibrium with $k=0$ exists if and only if*

1. $\beta \in \left[r, \hat{\beta}(\alpha)\right]$
2. $\beta \in \left[\overline{\beta}_0(r,\alpha), \min\{r, 1-\alpha\}\right]$
3. $\beta \leq \min\left\{\overline{\beta}_0(r,\alpha), (1-\alpha)\tilde{\beta}(\alpha)\right\}$

We now provide some intuition on this result and how it compares with the baseline. Broadly speaking, the main effect in the baseline, by which reducing service speed affects the content of word of mouth, is still present. The difference is that now, restricting service speed also affects the level of word of mouth. Setting a low service speed means that fewer followers learn, in situations where immediately-served leaders talk more than leaders who are delayed.

This *level effect* of restricting service speed will reduce the range of parameter values for which the seller sets $k=0$ in equilibrium. This can be seen most clearly from Case 1 of Lemma 4. The relevant condition, $\beta \leq \hat{\beta}(\alpha)$, is trivially satisfied for all $\beta \in [r,1]$ in the baseline since $\hat{\beta}(\alpha) = 1$ when $m_1 = m_0$. If instead $m_1 < m_0$, the seller prefers to immediately serve leaders when the good state is sufficiently likely, to increase how many followers learn via word of mouth, as described after equation (9).

The seller's incentive to immediately serve leaders is also increasing in signal precision, i.e. $\hat{\beta}(\alpha)$ is decreasing in α. This is because the amount of extra word of mouth generated by a high service speed is increasing in the total number of orders placed by leaders. A higher signal precision implies more orders in the good state, when extra word of mouth helps, and fewer in the bad state, when extra word of mouth hurts.

Now we consider a candidate equilibrium with $k = 1$. As in the baseline, followers who meet immediately-served leaders will act based on hard information about the state, whereas followers who meet delayed leaders may either all buy, or follow their private signals, depending on the value of the prior.

The behaviour of unconnected consumers is also similar to before but with a different cut-off. That is, unconnected followers will act on their privates signals if $\beta \geq \overline{\beta}_1(r, \alpha)$ and otherwise all refuse to buy, whereas the condition in the case $k = 0$ was $\beta \geq \overline{\beta}_0(r, \alpha)$. Expression (3) shows that $\overline{\beta}_q(r, \alpha)$ is increasing in q when $m_1 > m_0$, so in particular $\overline{\beta}_0(r, \alpha) < \overline{\beta}_1(r, \alpha)$. That is, followers who do not learn anything via word of mouth become more pessimistic (and hence require all refuse to buy for a larger range of β) if they believe the seller sets a higher service speed. Since only leaders who place orders will talk, and immediately-served leaders talk more than leaders who are delayed, the relative probability of being unconnected depends more on the state when service speed is high. Thus, being unconnected brings relatively worse news in a candidate equilibrium with $k = 1$ compared to one with $k = 0$.

Note that if $k = 1$, then meeting a delayed leader is interpreted by consumers as a result of a tremble and therefore is decisively good news. Therefore, $\gamma_G = \gamma_B = 1$ for all β. The behaviour of unconnected followers may vary with parameters. We start our analysis with the case $\beta > \overline{\beta}_1(r, \alpha)$, which implies $\delta_G = \alpha$ and $\delta_B = 1 - \alpha$. In this case setting $k = 1$ yields profits

$$\pi(1) = \beta \left[\alpha + \alpha m_1 + (n - \alpha m_1)\alpha\right] + (1 - \beta)\left[1 - \alpha + (n - (1 - \alpha)m_1)(1 - \alpha)\right].$$

A deviation to $k = 0$ gives

$$\pi(0) = \beta \left[\alpha + \alpha m_0 + (n - \alpha m_0)\alpha\right] + (1 - \beta)\left[1 - \alpha + (1 - \alpha)m_0 + (n - (1 - \alpha)m_0)(1 - \alpha)\right].$$

Therefore setting $k = 1$ is optimal whenever inequality (9) is reversed, i.e.

$$\beta \geq \hat{\beta}(\alpha) = 1 - \frac{m_1 - m_0}{m_1}\alpha.$$

Now consider $\beta \leq \overline{\beta}_1(r, \alpha)$, so that purchase probabilities are $\gamma_\omega = 1, \delta_\omega = 0$. Followers who remain unconnected never buy, whereas followers who meet delayed leaders (off the equilibrium path) again always buy.

Profits from setting $k = 1$ are then

$$\pi(1) = \beta\left[\alpha + \alpha m_1\right] + (1-\beta)(1-\alpha),$$

whereas deviating to $k = 0$ gives

$$\pi(0) = \beta\left[\alpha + \alpha m_0\right] + (1-\beta)\left[1 - \alpha + (1-\alpha)m_0\right].$$

It follows that the deviation is unprofitable for

$$\beta \geq \frac{(1-\alpha)m_0}{(1-\alpha)m_0 + \alpha(m_1 - m_0)} = \tilde{\beta}(\alpha),$$

which is a weaker condition than $\beta \geq \hat{\beta}(\alpha)$. Intuitively, the seller is less tempted to deviate to $k = 0$ when unconnected followers never buy, because delaying leaders means that more followers end up unconnected.

These results can be gathered and restated as the following Lemma.

Lemma 5. *An equilibrium with $k = 1$ exists if and only if*

1. $\beta \geq \hat{\beta}(\alpha)$
2. $\beta \in \left[\tilde{\beta}(\alpha), \min\left\{\overline{\beta}_1(r, \alpha), \hat{\beta}(\alpha)\right\}\right]$

We now comment on three features of these results. First, unlike in the baseline, a pure strategy equilibrium will exist in which the seller immediately serves all leaders if the good state is sufficiently likely. The less delayed leaders talk, i.e. the lower the value of m_0, the larger the parameter region where this equilibrium exists. The intuition is as described after equation (9). When immediately-served leaders talk more than leaders who are delayed, setting a high service speed allows the seller to maximize the number of followers who learn via word of mouth, and these followers will all buy if the state turns out to be good.

Second, the fact that followers interpret meeting delayed leaders (off the equilibrium path) as conclusive good news about the state still limits the parameter range for which $k = 1$ is an equilibrium. It was these beliefs that ruled out a pure strategy equilibrium with $k = 1$ in the baseline. The seller can take advantage of these beliefs when deviating to $k = 0$ to systematically fool followers, since they cannot observe the seller's deviation.

Third, an equilibrium where the seller immediately serves all leaders is easier to sustain when unconnected consumers never buy, $\beta < \overline{\beta}_1(r, \alpha)$, than when they act on their private signals, $\beta \geq \overline{\beta}_1(r, \alpha)$: the equilibrium condition is $\beta \geq \tilde{\beta}(\alpha)$ in the first case and $\beta \geq \hat{\beta}(\alpha) > \tilde{\beta}(\alpha)$ in the second case. The reason is that increased word of mouth can hurt the seller if followers learn the state is bad, but this matters less if followers also would refuse to buy had they remained unconnected.

We now turn to mixed strategy equilibria. From Lemma 1, in any such equilibrium, the seller must randomize between $k = 1$ and $k = 0$. We start by looking for an equilibrium where a follower who meets a delayed leader and gets $s = b$ randomizes between buying and not buying. From Lemma 2, we know that for such equilibrium to exist it must be the case that $\beta < r$ and $q^*(N) \to 1$ as $N \to \infty$. That is, in large markets, the seller must set $q^* = 1$.

The seller's indifference condition can be written as

$$\begin{aligned}\pi(0) =& \beta[\alpha + \alpha m_0 \gamma_G + (n - \alpha m_0)\delta_G] + (1-\beta)[1 - \alpha + (1-\alpha) m_0 \gamma_B + (n - (1-\alpha) m_0)\delta_B] \\ =& \beta[\alpha + \alpha m_1 + (n - \alpha m_1)\delta_G] + (1-\beta)[1 - \alpha + (n - (1-\alpha) m_1 \delta_B] = \pi(1).\end{aligned}$$

Suppose that followers who meet a delayed leader but get $s = b$ buy with probability $\chi \in [0, 1]$. This implies $\gamma_G = \alpha + (1-\alpha)\chi$ and $\gamma_B = (1-\alpha) + \alpha\chi$. Since in this candidate equilibrium sellers sets $k = 1$ almost surely, the relevant cutoff for unconnected followers is $\overline{\beta}_1(r, \alpha)$, so we obtain

$$\delta_G = \begin{cases} 0, & \beta < \overline{\beta}_1(r, \alpha) \\ \alpha, & \beta \geq \overline{\beta}_1(r, \alpha) \end{cases}, \quad \delta_B = \begin{cases} 0, & \beta < \overline{\beta}_1(r, \alpha) \\ 1-\alpha, & \beta \geq \overline{\beta}_1(r, \alpha) \end{cases}$$

Recall that $\overline{\beta}_1(r, \alpha) < r$. If $\beta \in [\overline{\beta}_1(r, \alpha), r]$, then the seller's indifference condition reduces to

$$\begin{aligned}\pi(0,\chi) =&\beta[\alpha+\alpha m_0(\alpha+(1-\alpha)\chi)+(n-\alpha m_0)\alpha]+\\ &(1-\beta)[1-\alpha+(1-\alpha)m_0(1-\alpha+\alpha\chi)+(n-(1-\alpha)m_0)(1-\alpha)]\\ =&\beta[\alpha+\alpha m_1+(n-\alpha m_1)\alpha]+(1-\beta)[1-\alpha+(n-(1-\alpha)m_1)(1-\alpha)\\ =&\pi(1).\end{aligned}$$

Profits $\pi(0,\chi)$ are increasing in χ, so there will exist a value of $\chi \in (0,1)$ that makes the seller indifferent, $\pi(0,\chi)=\pi(1)$ if $\pi(0,\chi=0)<\pi(1)<\pi(0,\chi=1)$. It is straightforward to verify that this condition is equivalent to

$$1-\alpha<\beta<\hat{\beta}(\alpha).$$

If instead $\beta<\overline{\beta}_1(r,\alpha)$, the seller's indifference conditions reduces to

$$\begin{aligned}\pi(0,\chi) =&\beta[\alpha+\alpha m_0(\alpha+(1-\alpha)\chi)]+(1-\beta)[1-\alpha+(1-\alpha)m_0(1-\alpha+\alpha\chi)]\\ =&\beta(\alpha+\alpha m_1)+(1-\beta)(1-\alpha))\\ =&\pi(1)\end{aligned}$$

As above, there exists a value of $\chi \in (0,1)$ that makes the seller indifferent if $\pi(0,\chi=0)<\pi(1)<\pi(0,\chi=1)$. This condition is now equivalent to

$$(1-\alpha)\tilde{\beta}(\alpha)<\beta<\tilde{\beta}(\alpha).$$

We can summarize these results in the following lemma.

Lemma 6. *A mixed strategy equilibrium in which the seller almost surely sets $k=1$ and immediately serves all leaders, and where followers with bad signals who meet delayed leaders randomize, exists if and only if either*

1. $\max\{1-\alpha,\overline{\beta}_1(r,\alpha)\}<\beta<\min\{\hat{\beta}(\alpha),r\}$ *or*
2. $(1-\alpha)\tilde{\beta}(\alpha)<\beta<\min\{\tilde{\beta}(\alpha),\overline{\beta}_1(r,\alpha)\}$.

We now look for a mixed strategy equilibrium where a follower who remains unconnected and gets $s = g$ randomizes.[17] An unconnected follower will be indifferent if $\overline{\beta}_q(r,\alpha) = \beta$, by Lemma 3. Since $\overline{\beta}_q(r,\alpha)$ is increasing in q, the necessary condition for the equilibrium to exist is $\overline{\beta}_0(r,\alpha) \leq \beta \leq \overline{\beta}_1(r,\alpha)$. This means that an unconnected follower with a bad signal would buy if they think all leaders were delayed, but would not buy if they think all leaders were immediately served. Moreover, for any interior $q \in (0,1)$, a follower who meets a delayed leader and gets $s = b$ retains their prior beliefs, by Lemma 2.

Suppose that an unconnected follower with $s = g$ buys with probability ξ. Then seller profits from setting $k = 0$ are

$$\pi(0,\xi) = \beta[\alpha + \alpha^2 m_0 + (n - m_0\alpha)\alpha\xi] + (1-\beta)[1-\alpha + (1-\alpha)^2 m_0 + (n - m_0(1-\alpha))(1-\alpha)\xi],$$

whereas profits from setting $k = 1$ are

$$\pi(1,\xi) = \beta[\alpha + \alpha m_1 + (n - m_1\alpha)\alpha\xi] + (1-\beta)[1-\alpha + (n - m_1(1-\alpha))(1-\alpha)\xi].$$

Since $\frac{\partial\pi(0,\xi)}{\partial\xi} > \frac{\partial\pi(1,\xi)}{\partial\xi} > 0$, by $m_0 < m_1$, there exists a unique ξ that satisfies the seller indifferent condition, $\pi(0,\xi) = \pi(1,\xi)$ whenever $\pi(0,\xi = 0) < \pi(1,\xi = 0)$ and $\pi(0,\xi = 1) > \pi(1,\xi = 1)$. The first inequality is equivalent to $\beta > (1-\alpha)\tilde{\beta}(\alpha)$ and the second is equivalent to $\beta < 1 - \alpha$, which implies the following.

Lemma 7. *A mixed strategy equilibrium in which the seller randomizes between $k = 0$ and $k = 1$ with $q \in (0,1)$, where unconnected followers with good signals randomize, exists if and only if*

$$\max\{\overline{\beta}_0(r,\alpha), (1-\alpha)\tilde{\beta}(\alpha)\} < \beta < \min\{\overline{\beta}_1(r,\alpha), 1-\alpha\}.$$

Having derived all pure and mixed strategy equilibria for the general case, we now combine Lemmas 4 - 7 to formulate an overall result.

Proposition 2. *In large markets the following equilibria exist:*

1. *If $\beta \geq r$ then*
   (a) *if $\beta \geq \hat{\beta}(\alpha)$ then $k = 1$;*
   (b) *if $\beta \leq \hat{\beta}(\alpha)$ then $k = 0$.*

[17] A follower who remains unconnected and gets $s = b$ will always have a strict incentive not to buy, because their posterior is less than $P(G|s = b) < r$.

2. *If* $\beta \in [\overline{\beta}_1(r,\alpha), r]$ *then*

   (a) *if* $\beta \geq \hat{\beta}(\alpha)$ *then* $k = 1$*;*

   (b) *if* $1-\alpha \leq \beta \leq \hat{\beta}(\alpha)$ *then there is a mixed strategy equilibrium with* $k = 1$ *almost surely;*

   (c) *if* $\beta < 1 - \alpha$ *then* $k = 0$

3. *If* $\beta \in [\overline{\beta}_0(r,\alpha), \overline{\beta}_1(r,\alpha)]$ *then*

   (a) *if* $\beta \geq \tilde{\beta}(\alpha)$ *then* $k = 1$*;*

   (b) *if* $(1-\alpha)\tilde{\beta}(\alpha) < \beta < 1-\alpha$ *then there is a mixed strategy equilibrium in which* $q \in (0, 1)$*.*

   (c) *if* $(1-\alpha)\tilde{\beta}(\alpha) \leq \beta \leq \tilde{\beta}(\alpha)$ *then there is a mixed strategy equilibrium with* $k = 1$ *almost surely;*

   (d) *if* $\beta \leq (1-\alpha)$ *then* $k = 0$

4. *If* $\beta < \overline{\beta}_0(r,\alpha)$ *then*

   (a) *if* $\beta \geq \tilde{\beta}(\alpha)$ *then* $k = 1$*;*

   (b) *if* $(1-\alpha)\tilde{\beta}(\alpha) \leq \beta \leq \tilde{\beta}(\alpha)$ *then there is a mixed strategy equilibrium with* $k = 1$ *almost surely;*

   (c) *if* $\beta \leq (1-\alpha)\tilde{\beta}(\alpha)$ *then* $k = 0$

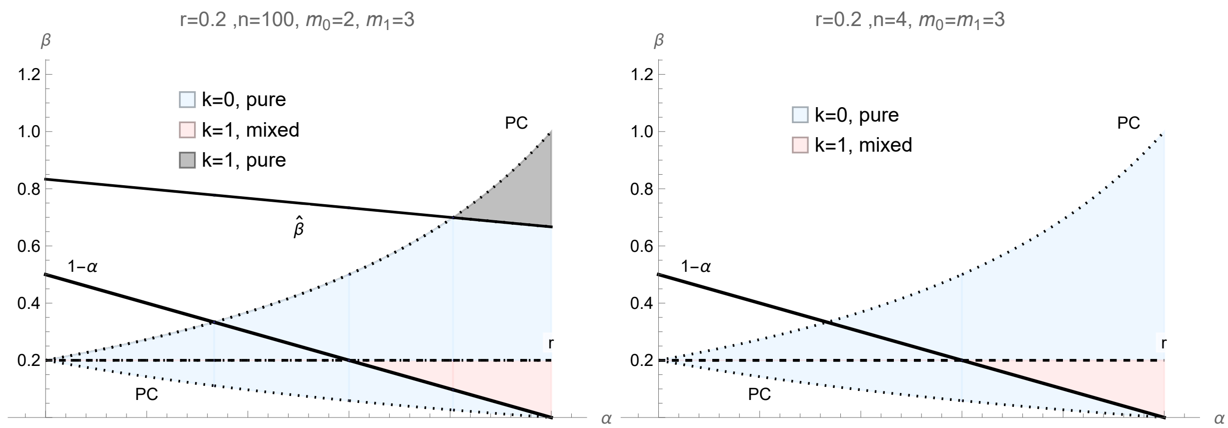
We now comment on two forces at play in this result that were absent in the baseline.

First consider a scenario where unconnected followers all act on their private signals, for all values of the prior and signal precision (that satisfy the participation constraint) and regardless of their beliefs about the seller's chosen service speed. This is in fact what will happen when there are many followers relative to leaders in the market, so when n is large. In the limit $n \to \infty$, most followers are unconnected in both states, so being unconnected is uninformative. That is,

$$\lim_{n\to\infty} \overline{\beta}_0(r,\alpha) = \lim_{n\to\infty} \overline{\beta}_1(r,\alpha) = \frac{(1-\alpha)r}{(1-\alpha)r + \alpha(1-r)},$$

where it then follows directly that $\beta \geq \lim_{n\to\infty} \overline{\beta}_0(r,\alpha)$ and $\beta \geq \lim_{n\to\infty} \overline{\beta}_1(r,\alpha)$ are equivalent to the participation constraint, $P(g|G) = \frac{\alpha\beta}{\alpha\beta+(1-\alpha)(1-\beta)} \geq r$.

Then only Case 1 and 2 of Proposition 2 are relevant, and we can reformulate in terms of this corollary:



(a) Large n, Corollary 1 (b) Baseline Case

Figure 2: Level effect for large n

Corollary 1. *In large markets, with many followers relative to leaders, $n \longrightarrow \infty$, so all followers act on their private signals, the following equilibria exist:*

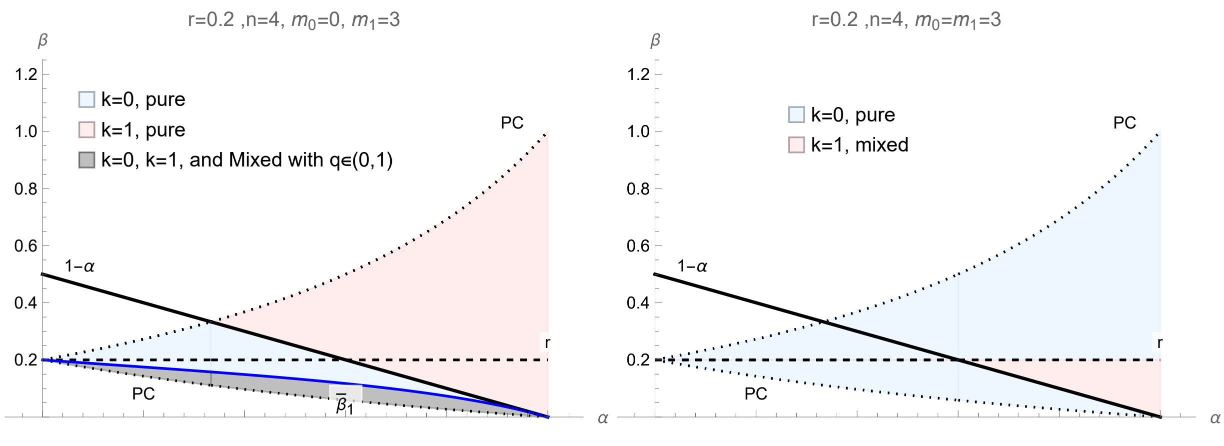
1. *If $\beta \leq 1 - \alpha$ or $r \leq \beta \leq \hat{\beta}(\alpha)$, a unique equilibrium exists and is in pure strategies, where the seller sets $k = 0$ and delays all leaders.*

2. *If $1 - \alpha < \beta < \min\{r, \hat{\beta}(\alpha)\}$, a unique equilibrium exists and is in mixed strategies, where the seller almost surely sets $k = 1$ and immediately serves all leaders.*

3. *If $\beta \geq \hat{\beta}(\alpha)$, a unique equilibrium exists and is in pure strategies, where the seller sets $k = 1$ immediately serves all leaders.*

Corollary 1 is very similar to Proposition 1 from the baseline. For ease of comparison, Figure 2 depicts both scenarios. The only difference is that a pure strategy equilibrium with $k = 1$ now exists whenever $\beta \geq \hat{\beta}(\alpha)$. The smaller the value of m_0, the less that delayed leaders talk, and hence the less tempted the seller is to delay them. As such, the parameter region with the pure strategy equilibrium $k = 1$ grows larger as m_0 decreases. This is the mechanism described after expression (9) and Lemmas 4, by which increasing service speed leads to increased sales via extra word of mouth when the state turns out to be good, but reduced sales when the state turns out to be bad. In situations where all unconnected followers act according to their private signals, this is the only way that seller behavior in the general case differs from that in the baseline.

Now consider a situation where delayed leaders never talk, $m_0 = 0$. Cutoff values are then $\tilde{\beta}(\alpha) = 0$; $\overline{\beta}_0(r, \alpha) = \frac{(1-\alpha)r}{(1-\alpha)r+\alpha(1-r)}$, which is equivalent to the participation constraint;



(a) Delayed leaders do not talk, m_0 (b) Baseline Case

Figure 3: Level effect when delayed leaders do not talk

$\hat{\beta}(\alpha) = 1 - \alpha$.

Hence, only Case 1, Case 2, and parts (a) and (b) of Case 3 are relevant in Proposition 2. We can can reformulate in terms of this corollary:

Corollary 2. *In large markets, where delayed leaders do not engage in word of mouth, $m_0 = 0$, the following equilibria exist:*

1. *If $\overline{\beta}_1(r, \alpha) < \beta < 1 - \alpha$, a unique equilibrium exists and is in pure strategies, where the seller sets $k = 0$ and delays all leaders.*

2. *If $\beta \geq 1 - \alpha$, a unique equilibrium exists and is in pure strategies, where the seller sets $k = 1$ immediately serves all leaders.*

3. *If $\beta < \min\{\overline{\beta}_1(r, \alpha), 1 - \alpha\}$ then the following equilibria exist: a pure strategy equilibrium with $k = 0$, a pure strategy equilibrium with $k = 1$, and a mixed strategy equilibrium in which $q \in (0, 1)$.*

Taking $m_0 \to 0$ maximizes the strength of the effect that was isolated in Corollary 1. As leaders talk less and less, the region where the pure strategy equilibrium $k = 1$ exists gets larger, until it eventually covers all of $\beta \geq 1 - \alpha$ when $m_0 = 0$, as shown in Figure 3.

As m_0 drops, unconnected followers also become less and less pessimistic about the state, but only if they believe the seller has engaged in delay. Thus, unlike in the baseline, there is a region where the behavior of unconnected follower depends on their beliefs about the seller's chosen service speed, i.e. for $\overline{\beta}_0(r, \alpha) < \beta < \overline{\beta}_1(r, \alpha)$. The function

$\overline{\beta}_0(r, \alpha)$ is decreasing in m_0, and coincides with the participation constraint when $m_0 = 0$, so that the relevant constraint becomes $\beta < \overline{\beta}_1(r, \alpha)$, the blue line in the Figure.

The fact that the optimal behavior of unconnected followers depends on their beliefs about the service speed is what gives equilibrium multiplicty in the general case. If a high service speed is expected, then unconnected followers will not buy, which pushes the seller to increase service speed to minimize the number of unconnected followers. If instead a low speed is expected, then unconnected followers will buy if they receive good signals, which makes it more attractive for the seller to reduce service speed and leave more followers unconnected.

Notice, in particular, that equilibrium multiplicity is not an issue when $n \to \infty$ (see Corollary 1). There optimal follower behavior does not depend on their beliefs about the service speed which results in equilibrium uniqueness. Put another way, we expect equilibrium multiplicity to be most relevant in markets where there are many (or at least not too few) opinion leaders relative to followers.

4 Word of mouth and product awareness

How would seller strategic behavior change if unconnected followers were simply unable to buy? This could reasonable for situations where the seller launches a product that leaders are aware off, but where followers only become aware if they hear from leaders. Many of the works surveyed in the Introduction assume that word of mouth has precisely this role. Intuitively, we might expect that delay becomes less attractive if unconnected followers are unable to buy, because delaying leaders means more followers end up unconnected.

To model consumer awareness we assume that unconnected followers never buy, i.e. $\delta_G = \delta_B = 0$. Moreover, we assume that followers who meet delayed leaders then also get a private signal with accuracy α, which implies that γ_ω is defined in the same way as in the main analysis. The interpretation is that it may be straightforward to access noisy information about product quality, but only for consumers who first become aware of the product. At the end of the section we discuss how the absence of any follower private signals would affect the results.

We start our analysis with the case $\beta \geq r$, so followers who meet delayed leaders always buy, regardless of the service speed they expect. Profits from setting $k = 0$ are

$$\pi(0) = \beta\left[\alpha(m_0 + 1)\right] + (1-\beta)(1-\alpha)(m_0 + 1),$$

whereas profits from setting $k = 1$ are

$$\pi(1) = \beta\left[\alpha(m_1 + 1)\right] + (1-\beta)(1-\alpha).$$

Thus, we have $\pi(0) \geq \pi(1)$ if and only if

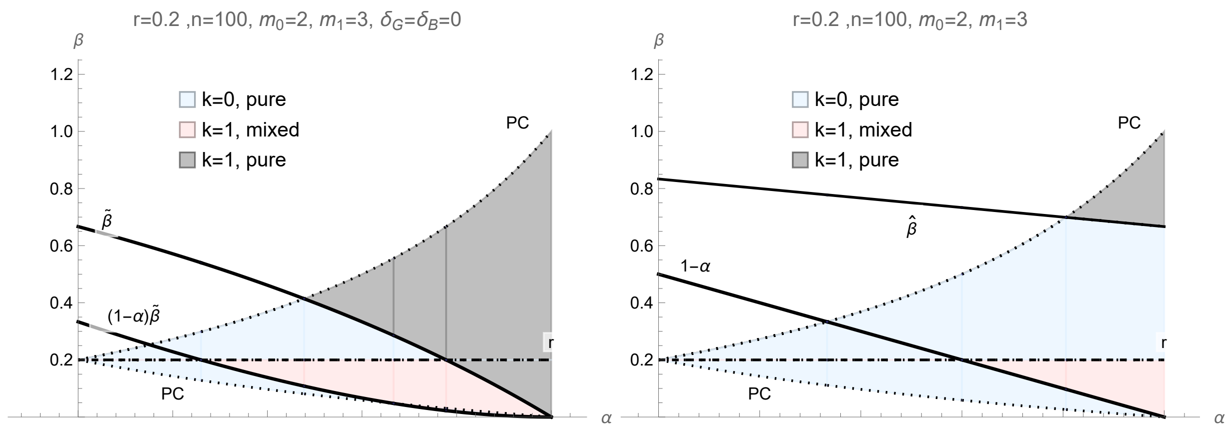
$$\beta \leq \frac{m_0(1-\alpha)}{m_0(1-\alpha) + (m_1 - m_0)\alpha} = \tilde{\beta}(\alpha).$$

Now consider $\beta < r$. Since unconnected followers do not buy, regardless of the candidate equilibrium in question, we can use our earlier results from Case 4 of Proposition 2. There we had $\beta < \beta_0(r, \alpha) \leq \beta_1(r, \alpha)$, so unconnected followers never found it optimal to buy. Taken together, we have the following.

Proposition 3. *In large markets, where unconnected followers cannot buy, the following equilibria exist:*

1. *If $\beta \geq r$ then*
   - *(a) if $\beta \geq \tilde{\beta}(\alpha)$ then $k = 1$;*
   - *(b) if $\beta \leq \tilde{\beta}(\alpha)$ then $k = 0$.*
2. *If $\beta < r$ then*
   - *(a) if $\beta \geq \tilde{\beta}(\alpha)$ then $k = 1$;*
   - *(b) if $(1-\alpha)\tilde{\beta}(\alpha) \leq \beta \leq \tilde{\beta}(\alpha)$ then there is a mixed strategy equilibrium with $k = 1$ almost surely;*
   - *(c) if $\beta \leq (1-\alpha)\tilde{\beta}(\alpha)$ then $k = 0$*

For the interpretation, first consider the case of $\beta > r$. It follows from $\tilde{\beta}(\alpha) \leq \hat{\beta}(\alpha)$ that the parameter region with $k = 1$ has increased in size, whereas the parameter region with $k = 0$ has decreased in size, compared to the main analysis. Intuitively, profits drop compared to our main analysis, but they drop to a larger extent when the seller delays



(a) Unmatched do not buy (b) Unmatched may buy, Large n

Figure 4: Unmatched followers do not buy

leaders, since then more followers end up unconnected. This makes it less attractive for the seller to use delay.

A broadly similar story holds for $\beta < r$. As $\overline{\beta}_1(r, \alpha)$ is not relevant when unmatched followers never buy, the region that had multiple equilibria now only has $k = 0$. At the same time, the constraint $\beta < (1 - \alpha)$ is now replaced by $\beta < (1 - \alpha)\tilde{\beta}(\alpha)$, thus for intermediate β the seller now prefers to set $k = 1$ almost surely.

These ideas are illustrated in Figure 4. Panel (a) depicts the seller's equilibrium behavior for particular parameter values, assuming that unconnected followers cannot buy. For the sake of comparison, Panel (b) depicts seller behavior from Corollary 1 in the general analysis, where unconnected followers always act on their private signals, and where setting $k = 0$ is more attractive.

Now we consider two special cases. First, suppose that $m_0 = 0$, so delayed leaders don't talk at all. In this case $\tilde{\beta}(\alpha) = 0$ and so the seller always sets $k = 1$. Clearly, it does not pay off to delay leaders if they never talk to followers, who therefore can never buy.

Second, consider the case of $m_0 = m_1$, so immediately-served and delayed leaders talk to the same extent. Then $\tilde{\beta}(\alpha) = 1$, and we can verify that Proposition 3 reduces to Proposition 1, so we get our baseline results. The total level of word of mouth, and hence the total number of unconnected followers, is then independent of service speed, and so the seller's strategic behavior does not depend on whether unconnected follower do or do not buy.

We can conclude that if word of mouth spreads awareness, for example with the launch of a new product, and hence both shapes beliefs and expands the size of the market, the seller will use delay less often, but only when $m_0 < m_1$.

Finally, we explore what happens if leaders are the only source of information of the followers, i.e. if followers never receive any private signals. In this situation followers who meet delayed leaders will all buy, since meeting a delayed leader is just as convincing as receiving a good signal. It follows that Case 1 of Proposition 3 then applies for both $\beta \geq r$ and $\beta < r$. Looking at Figure 4, instead of setting $k = 1$ almost surely in the red region of panel (a), the seller would prefer to set $k = 0$, as meeting a delayed leader has a larger impact when followers cannot receive bad private signals. This implies that the comparison with Corollary 1 is not as clear-cut as before: if quality is ex ante high, then the inability of unconnected followers to buy will push the seller to set $k = 1$ more often, but the opposite is true if quality is ex ante lower, where setting $k = 0$ now becomes more attractive.

5 Conclusions

In this paper, we explored how a seller can strategically use product delay to influence consumer learning about the quality of its product through word of mouth. The effect of service speed on learning can be decomposed into the content effect and the level effect.

The content effect influences the behavior of consumers who encounter either served or delayed leaders. If expected product quality is high, then meeting a delayed leader is sufficiently good news for these followers to buy. The seller then prefers to conceal information about quality by delaying, rather than revealing quality to leaders who are served. If expected quality is not high, then meeting a delayed leader does not significantly impact followers' behavior. In this case the seller prefers to hide information by delaying leaders if signals are inaccurate, and otherwise to serve all leaders almost surely. Broadly, if the seller mainly cares about influencing those followers who encounter leaders, and if the total amount of word of mouth depends little on how many leaders are immediately served, then it is often optimal for the seller to delay.

The level effect plays a significant role in determining how many consumers receive information from leaders, and it also affects the behavior of consumers who do not hear from leaders at all. This effect becomes evident in situations where immediately-served leaders engage in more word-of-mouth communication than leaders who are delayed. The service delay reduces the likelihood that a consumer will hear from a leader, prompting the seller to increase service speed when expected product quality is high. Moreover, the seller may choose to set a high service speed even when expected quality is moderate, because consumers who expect this service speed will perceive the lack of communication from leaders as particularly negative news. Thus, the seller has an incentive to increase service speed to reduce the number of consumers who do not receive any news. If word-of-mouth communication is crucial for spreading product awareness, then the level effect is amplified. Nonetheless, the seller may still use delay to influence consumer word-of-mouth communication.

6 Appendix: Proofs

Proof of Lemma 1. Let us again denote the number of leaders served in state ω as $S_\omega(K)$, and the number of leaders delayed as $D_\omega(K)$, where

$$S_\omega(K) = \sum_{j=0}^{N} \min\{j, K\} Q_\omega(j), \quad D_\omega(K) = \sum_{j=0}^{N} \max\{j-K, 0\} Q_\omega(j).$$

For notational simplicity we denote $S_\omega(-1) = 0$ and $D_\omega(-1) = D_\omega(K) = \sum_{j=0}^{N} jQ_\omega(j)$, i.e. when the firm sets $K = 0$ downward tremble is not possible and all consumers are delayed.

Let γ_ω be the probability that a consumer who met a delayed leader buys the product, and let δ_ω be the probability that a consumer who did not meet anyone buys the product. Profits from implementing service speed K are then

$$\begin{aligned}
\pi(K) =& \frac{\beta}{N}\left[\sum_{j=0}^{N} jQ_G(j) + (1-\varepsilon)\left(m_1 S_G(K) + m_0\gamma_G D_G(K) + \delta_G\left(nN - m_1 S_G(K) - m_0 D_G(K)\right)\right)\right. \\
&+ \left.\varepsilon\left(m_1 S_G(K-1) + m_0\gamma_G D_G(K-1) + \delta_G\left(nN - m_1 S_G(K-1) - m_0 D_G(K-1)\right)\right)\right] \\
&+ \frac{1-\beta}{N}\left[\sum_{j=0}^{N} jQ_B(j) + (1-\varepsilon)\left(m_0\gamma_B D_B(K) + \delta_B\left(nN - m_1 S_B(K) - m_0 D_B(K)\right)\right)\right. \\
&+ \left.\varepsilon\left(m_0\gamma_B D_B(K-1) + \delta_B\left(nN - m_1 S_B(K-1) - m_0 D_B(K-1)\right)\right)\right].
\end{aligned}$$

We now show that the profit-maximizing service speed K cannot take on any value $1 \le K \le N-1$. Consider the difference $\pi(K+1) - \pi(K)$. Note that

$$S_\omega(K+1) - S_\omega(K) = \sum_{j=0}^{N} \left(\min\{j, K+1\} - \min\{j, K\}\right) Q_\omega(j) = \sum_{j=K+1}^{N} Q_\omega(j),$$

and

$$D_\omega(K+1) - D_\omega(K) = \sum_{j=0}^{N} \left(\max\{j-K-1, 0\} - \max\{j-K, 0\}\right) Q_\omega(j) = -\sum_{j=K+1}^{N} Q_\omega(j).$$

which implies

$$\Delta(K) \equiv \pi(K+1) - \pi(K) =$$
$$\frac{\beta}{N}[m_1(1-\delta_G) - m_0(\gamma_G - \delta_G)]\left(\sum_{j=K+1}^{N} Q_G(j) + \varepsilon Q_G(K)\right)$$
$$\frac{1-\beta}{N}[-m_1\delta_B - m_0(\gamma_B - \delta_B)]\left(\sum_{j=K+1}^{N} Q_B(j) + \varepsilon Q_B(K)\right),$$

where the expression in the second set of square brackets is non-positive, by $m_1 \geq 0$. Thus, we have that

$$\operatorname{sign}\Delta(K) = \operatorname{sign}\left(\frac{\beta[m_1(1-\delta_G) - m_0(\gamma_G - \delta_G)]}{(1-\beta)([m_1\delta_B + m_0(\gamma_B - \delta_B)])}\frac{\sum_{j=K+1}^{N} Q_G(j) + \varepsilon Q_G(K)}{\sum_{j=K+1}^{N} Q_B(j) + \varepsilon Q_B(K)} - 1\right)$$

Note, that if $\Delta(K_0) > 0$ holds for some K_0 it must be the case that $m_1(1-\delta_G) - m_0(\gamma_G - \delta_G) > 0$. As such, $\Delta(K_0) > 0$ must also hold for all $K > K_0$, since $Q_G(K)/Q_B(K)$ is increasing in K. Therefore, $\pi(K)$ attains its maximum either at $K = 0$ or at $K = N$. □

Proof of Lemma 2. Part 1. Note that

$$D_\omega(q) = (1-q)\sum_{j=0}^{N} jQ_\omega(j) + q\varepsilon Q_\omega(N)$$

$$\begin{aligned}\mu(1,N) &= \frac{\beta(1-\alpha)D_G(1)}{\beta(1-\alpha)D_G(1) + (1-\beta)\alpha D_B(1)} \\ &= \frac{\beta(1-\alpha)\varepsilon Q_G(N)}{\beta(1-\alpha)\varepsilon Q_G(N) + (1-\beta)\alpha\varepsilon Q_B(N)} = \frac{1}{1 + \frac{(1-\beta)(1-\alpha)}{\beta\alpha}\frac{Q_G(N)}{Q_B(N)}}\end{aligned}$$

Now, due to $\lim_{N\to\infty} Q_G(N)/Q_B(N) = 0$ we get that $\lim_{N\to\infty} \mu(1,N) = 1$. *Part 2.* For any fixed q we get

$$\mu(q,N) = \frac{\beta(1-\alpha)\left((1-q)\sum_{j=0}^{N} jQ_G(j) + q\varepsilon Q_G(N)\right)}{\beta(1-\alpha)\left((1-q)\sum_{j=0}^{N} jQ_G(j) + q\varepsilon Q_G(N)\right) + (1-\beta)\alpha\left((1-q)\sum_{j=0}^{N} jQ_B(j) + q\varepsilon Q_B(N)\right)}$$

Thus, for $q < 1$ we get

$$\lim_{\varepsilon \to 0} \mu(q, N) = \frac{\beta(1-\alpha)\sum_{j=0}^{N} jQ_G(j)}{\beta(1-\alpha)\sum_{j=0}^{N} jQ_G(j) + (1-\beta)\alpha\sum_{j=0}^{N} jQ_B(j)}$$

and is independent of q. Using $\sum_{j=0}^{N} jQ_G(j) = \alpha N$ and $\sum_{j=0}^{N} jQ_G(j) = (1-\alpha)N$ we get that

$$\lim_{\varepsilon \to 0} \mu(q, N) = \frac{\beta(1-\alpha)\alpha}{\beta(1-\alpha)\alpha + (1-\beta)\alpha(1-\alpha)} = \beta$$

Part 3. We can rewrite consumer belief in the form

$$\mu(q, N) = \frac{\beta(1-\alpha)}{\beta(1-\alpha) + (1-\beta)\alpha\frac{(1-q)\sum_{j=0}^{N} jQ_B(j) + q\varepsilon Q_B(N)}{(1-q)\sum_{j=0}^{N} jQ_G(j) + q\varepsilon Q_G(N)}}$$

Note that

$$\frac{\partial}{\partial q}\left(\frac{(1-q)\sum_{j=0}^{N} jQ_B(j) + q\epsilon Q_B(N)}{(1-q)\sum_{j=0}^{N} jQ_G(j) + q\epsilon Q_G(N)}\right) = \frac{Q_B(N)\sum_{j=0}^{N} jQ_G(j) - Q_G(N)\sum_{j=0}^{N} jQ_B(j)}{\left[(1-q)\sum_{j=0}^{N} jQ_G(j) + q\epsilon Q_G(N)\right]^2}\varepsilon < 0$$

and therefore $\mu(q, N)$ is increasing in q. Moreover, $\mu(0, N) = \beta$ and $\lim N \to \infty \mu(1, N) = 1$. Thus, for any r, there exists $\overline{N}(r)$ such that $\mu(1, N) > r$ holds for all $N > \overline{N}(r)$. Thus, for any $r \in (\beta, 1)$, and any $N > \overline{N}(r)$, there exists a unique $q^*(N)$ such that $\mu(q^*(N), N) = r$.

Now we prove that $\lim_{N \to \infty} q^*(N) = 1$. To do so, we proceed by contradiction. Suppose that $q^*(N) < \overline{q}$ for all N. Then, using monotonicity of belief in q we get

$$\lim_{N \to \infty} \mu(q^*(N), N) < \lim_{N \to \infty} \mu(\overline{q}, N) = \beta < r$$

which is not possible as $\mu(q, N)$ is continuous in q. □

References

Ajorlou, A., Jadbabaie, A., and Kakhbod, A. (2018). Dynamic pricing in social networks: The word-of-mouth effect. *Management Science*, 64(2):971–979.

Aoyagi, M. (2010). Optimal sales schemes against interdependent buyers. *American Economic Journal: Microeconomics*, 2(1):150–182.

Bala, V. and Goyal, S. (1998). Learning from neighbours. *The review of economic studies*, 65(3):595–621.

Banerjee, A., Chandrasekhar, A. G., Duflo, E., and Jackson, M. O. (2013). The diffusion of microfinance. *Science*, 341(6144):1236498.

Banerjee, A. and Fudenberg, D. (2004). Word-of-mouth learning. *Games and economic behavior*, 46(1):1–22.

Bhalla, M. (2013). Waterfall versus sprinkler product launch strategy: Influencing the herd. *The Journal of Industrial Economics*, 61(1):138–165.

Bulow, J. and Roberts, J. (1989). The simple economics of optimal auctions. *Journal of political economy*, 97(5):1060–1090.

Campbell, A. (2013). Word-of-mouth communication and percolation in social networks. *American Economic Review*, 103(6):2466–2498.

Campbell, A. (2015). Word of mouth model of sales. *Economics Letters*, 133:45–50.

Campbell, A., Leister, C. M., and Zenou, Y. (2020). Word-of-mouth communication and search. *The RAND Journal of Economics*, 51(3):676–712.

Campbell, A., Mayzlin, D., and Shin, J. (2017). Managing buzz. *The RAND Journal of Economics*, 48(1):203–229.

Chevalier, J. A. and Mayzlin, D. (2006). The effect of word of mouth on sales: Online book reviews. *Journal of marketing research*, 43(3):345–354.

Debo, L. G., Parlour, C., and Rajan, U. (2012). Signaling quality via queues. *Management Science*, 58(5):876–891.

DeGraba, P. (1995). Buying frenzies and seller-induced excess demand. *The RAND Journal of Economics*, pages 331–342.

Duan, W., Gu, B., and Whinston, A. B. (2008). The dynamics of online word-of-mouth and product sales—an empirical investigation of the movie industry. *Journal of retailing*, 84(2):233–242.

Ellison, G. and Fudenberg, D. (1995). Word-of-mouth communication and social learning. *The Quarterly Journal of Economics*, 110(1):93–125.

Ferguson, D. G. (1994). Shortages, segmentation, and self-selection. *Canadian Journal of Economics*, pages 183–197.

Galeotti, A. (2010). Talking, searching, and pricing. *International Economic Review*, 51(4):1159–1174.

Galeotti, A., Golub, B., and Goyal, S. (2020). Targeting interventions in networks. *Econometrica*, 88(6):2445–2471.

Galeotti, A. and Goyal, S. (2009). Influencing the influencers: a theory of strategic diffusion. *The RAND Journal of Economics*, 40(3):509–532.

Godes, D. (2017). Product policy in markets with word-of-mouth communication. *Management Science*, 63(1):267–278.

Godes, D. and Mayzlin, D. (2004). Using online conversations to study word-of-mouth communication. *Marketing science*, 23(4):545–560.

Hauser, D. N. (2023). Censorship and reputation. *American Economic Journal: Microeconomics*, 15(1):497–528.

He, S., Hollenbeck, B., and Proserpio, D. (2022). The market for fake reviews. *Marketing Science*, 41(5):896–921.

Liu, T. and Schiraldi, P. (2012). New product launch: herd seeking or herd preventing? *Economic theory*, 51:627–648.

Liu, Y. (2006). Word of mouth for movies: Its dynamics and impact on box office revenue. *Journal of marketing*, 70(3):74–89.

Loertscher, S. and Muir, E. V. (2022). Monopoly pricing, optimal randomization, and resale. *Journal of Political Economy*, 130(3):566–635.

Luca, M. and Zervas, G. (2016). Fake it till you make it: Reputation, competition, and yelp review fraud. *Management Science*, 62(12):3412–3427.

Mayzlin, D. (2006). Promotional chat on the internet. *Marketing science*, 25(2):155–163.

Mayzlin, D., Dover, Y., and Chevalier, J. (2014). Promotional reviews: An empirical investigation of online review manipulation. *American Economic Review*, 104(8):2421–2455.

Mohr, I. (2017). Managing buzz marketing in the digital age. *Journal of Marketing Development and Competitiveness*, 11(2):10–16.

Möller, M. and Watanabe, M. (2010). Advance purchase discounts versus clearance sales. *The Economic Journal*, 120(547):1125–1148.

Nocke, V. and Peitz, M. (2007). A theory of clearance sales. *The Economic Journal*, 117(522):964–990.

Parakhonyak, A. and Vikander, N. (2019). Optimal sales schemes for network goods. *Management Science*, 65(2):819–841.

Parakhonyak, A. and Vikander, N. (2023). Information design through scarcity and social learning. *Journal of Economic Theory*, 207:105586.

Sgroi, D. (2002). Optimizing information in the herd: Guinea pigs, profits, and welfare. *Games and Economic Behavior*, 39(1):137–166.

Smirnov, A. and Starkov, E. (2022). Bad news turned good: reversal under censorship. *American Economic Journal: Microeconomics*, 14(2):506–560.

Stock, A. and Balachander, S. (2005). The making of a "hot product": A signaling explanation of marketers' scarcity strategy. *Management science*, 51(8):1181–1192.

Wilson, C. A. (1988). On the optimal pricing policy of a monopolist. *Journal of Political Economy*, 96(1):164–176.

Xiong, G. and Bharadwaj, S. (2014). Prerelease buzz evolution patterns and new product performance. *Marketing Science*, 33(3):401–421.