Optimal Sales Schemes for Network Goods *

Alexei Parakhonyak† and Nick Vikander‡

October 2015

Abstract

This paper explores how to best sequence sales in the presence of network effects. A monopolist sets a price for its product and also chooses whether to serve some consumers before others through its choice of sales scheme. We show that a firm with imperfect control over sequencing should serve consumers as sequentially as possible, with consumers in smaller groups served first, and that the optimal sales scheme is fully sequential. Under a fully sequential scheme, each consumer observes previous sales before choosing whether to buy himself, and independent-minded consumers can act as opinion leaders for those who follow.

JEL-codes: M31, D42, D82, L12

Key Words: Product launch, Network effects, Sequencing of sales

*This paper has benefited from presentations on Royal Economic Society Conference 2012, EARIE 2012, ... and from discussions with Maarten Janssen and Andre Veiga.

†University of Oxford, Department of Economics. E-mail: aleksei.parakhonyak@economics.ox.ac.uk

‡University of Copenhagen, Department of Economics. E-mail: nick.vikander@econ.ku.dk

1 Introduction

A wide variety of products exhibit network effects, where a consumer's benefit from buying is increasing in total sales. Network effects can arise through the presence of complementary products. For example, consumers who buy the Apple iPhone or Sony PlayStation will benefit if others do the same, since high sales will lead to more apps and game titles being released. They can also arise due to technological compatibility, for example with computer operating systems, such as Windows, Apple OS, and Google's Chrome OS, and with business solutions such as Microsoft Office 365 and Google Apps for Work. Yet another reason for network effects is consumer social concerns, for products with a fashion component or in settings where consumption is a social experience.[1] Regardless of their source, network effects push consumers to buy products they expect to be popular, and imply that the existing network of users can impact product adoption.

Previous research has looked at how firms can exploit network effects through pricing, advertising, seeding strategy, and release of limited-time or lower-quality versions of their products (freemium). But almost no attention has been paid to the sequencing of sales. That is, should a firm release its product simultaneously to all consumers, or instead follow a sequential strategy, serving some consumers before others? And if the firm does use sequential sales, which consumers should it serve first?

Intuition suggests that sequencing can be critical for sales dynamics. If a firm follows a sequential strategy, i.e. consumers observe the decisions of previous buyers, and initial sales are high, then success may breed success. Consumers who observe high initial sales will increase their own demand, both directly due to the observed installed base, and indirectly

[1]Customers may prefer attending nightclubs or watching movies alongside others, and purchasing clothes, books or music may associate the owner with the 'in thing', facilitating social interactions. For products such as multiplayer online games (from PlayStation, Xbox and others) these social interactions can be crucial in determining consumer payoffs.

as they expect high sales in the future. But by the same token, the risk of sequential sales is that failure can also breed failure, where low initial sales dissuade other consumers from buying. For this reason, it is not a priori clear how to best sequence sales in the presence of network effects, which is the question we explore in this paper.

Our key modelling assumption is that firms have at least some control over sequencing. This assumption is broadly plausible in a variety of settings. For example, in many business-to-business transactions, firms have almost complete control over how to sequence sales. Microsoft can release Office 365 simultaneously across the market, offer the product to one business before approaching another, or follow an intermediate strategy by first offering the product to businesses in specific sectors. The same is true for Google, which can offer its new Chromebook laptop to different school districts sequentially or simultaneously. Firms selling mass-market products may be unable to choose the precise order of sales, but they can still decide whether to release products simultaneously or sequentially across different markets. For example Sony's European Marketing director said the following regarding the launch of its gaming console:

> "We will launch [the PlayStation 4] this year. Exactly what regions, what timing, is being worked through. Which regions in 2013 - is it all of them, is it some of them? Is there some degree of phasing? We'll reveal that in more detail later, but we can't yet."[2]

Even firms with little direct control over sequencing can decide whether to make information on previous sales available to later consumers. For example, a restaurant can select a layout that makes it difficult for potential customers to see how many people are already inside. Doing so effectively makes decisions simultaneous by forcing customers to decide whether to enter the restaurant before observing how busy it is. A nightclub can place a conspicuous queue outside its entrance, to provide passers by with information about the

[2]See "PlayStation 4 launch in 'at least' one country in 2013", digitalspy.co.uk, February 22, 2013.

number of earlier arrivals, effectively making decisions sequential. In a similar spirit, when Apple launched the iPhone 5, it publicized pre-order sales figures prior to the official release, so that consumers looking to buy would know how many others had bought as well.[3]

Formally, we consider a setting where a firm sells a homogeneous good, and where each consumer's payoff from buying is the sum of two terms: an intrinsic payoff, and a network payoff that is increasing in total sales. Both a consumer's intrinsic payoff from buying and the weight he places on the network payoff are private information. The firm sets a price and chooses a sales scheme, which is a partition of consumers into different cohorts. Consumers in each cohort buy simultaneously, without observing one another's decisions, but consumers in different cohorts buy sequentially, having observed sales from all previous cohorts. The set of all potential sales schemes represents all possible ways of sequencing sales, from fully sequential (one consumer per cohort), to fully simultaneous (all consumers in one cohort). We show which sales scheme is optimal, and more broadly, rank a wide variety of schemes in terms of their expected profits. This ranking is relevant for situations where a firm's control over sequencing may be incomplete, so where it can only choose from a strict subset of all potential schemes.

Our first set of results considers simultaneous versus sequential sales. We show that for a firm with complete control over sequencing, the optimal scheme is fully sequential, with a single consumer per cohort. If a firm has incomplete control over sequencing, so that certain potential schemes are infeasible, it is best to choose a scheme that is as sequential as possible. That is, moving from any particular sales scheme to another that is more sequential, in a way that we make precise, will always increase expected profits. The former result suggests that a firm engaging in business-to-business transactions may benefit by approaching potential clients one by one. The latter suggests that a sequential product launch will tend to outperform a simultaneous launch in multiple markets, and that a firm

[3]See "iPhone 5 Pre-Orders Top Two Million in First 24 Hours", Apple.com, September 17, 2012.

in a given market may benefit from making available as much information about consumers' purchases as possible.

These results on sequential sales hold regardless of whether a firm is able to commit to its choice of sales scheme, so even if it can change to a different scheme after observing early sales. With the benefit of hindsight, a firm with low early sales may realize that a fully simultaneous scheme would have led to higher realized profits. But from then on, it will still prefer to serve the remaining consumers with a scheme that is as sequential as possible.

The economic mechanism supporting sequential sales relies crucially on consumers being rational and forward looking. Even before any sales are realized, a sequential scheme changes consumer expectations of later purchase behavior. An early consumer realizes that those in later cohorts will observe her purchase decision, and therefore takes into account how her own purchase will encourage others to buy themselves. This expectation of high later sales increase the consumer's own expected payoff from buying. Thus, while early failure can encourage later failure under a sequential scheme, the use of this scheme itself makes early failure less likely by encouraging early consumers to buy. The important point is not just that consumers are observed but that being observed makes consumers behave differently.

Our second set of results looks specifically at how to order sales, and does so in two different ways. Given a group of sales schemes that are equally sequential, in the sense of having the same total number of cohorts, and the same number of cohorts of each different size, the most profitable scheme will serve consumers in smaller cohorts first. That is, for a firm looking to launch its product across different markets, not only is it is best to release the product sequentially, but to start with smaller markets before moving on to larger ones. We also extend the analysis by assuming the firm can partially distinguish between different consumers, specifically when the weight consumers place on the network payoff is public information. The optimal sales scheme then serves consumers sequentially in increasing order of these weights, so that independent-minded consumers make purchase decisions first and can serve as opinion leaders for those who follow.

We then draw a parallel between sequential sales and consumer communication. The mechanism driving sequential sales relies on consumers learning about a product's popularity by observing previous sales, which will be less important if consumers can directly learn about one another's preferences via online forums and discussion boards. We establish a novel result showing that if consumers can exchange messages about their respective valuations before making purchases, then truthful communication can be incentive compatible, even if the firm may exploit this information to raise the price. Truthful communication is possible precisely because of network effects, even though they would seem to push consumers to exaggerate so as to convince others to buy. The practical implication is that the sequencing of sales will be less important for products about which consumers regularly communicate. The caveat is that consumer concern that the firm is monitoring their messages can potentially derail successful communication.

The vast literature on network goods, starting from seminal papers by Katz and Shapiro (1985) and Farrell and Saloner (1985, 1986), has largely assumed that the order of consumer entry is predetermined or endogenously chosen by consumers. These models often exhibit multiple equilibria, as consumer have self-fulfilling beliefs on how many others are going to enter (see, e.g. Dybvig and Splatt (1983), Cabral et al. (1999)). Our paper is the first to focus on firm's control over the timing of sales in a setting with network effects and rational, forward-looking consumers, and to examine the relative merits of different schemes. Although the main strategic concern of this paper has been pointed out in the literature[4], this is the first paper to examine how a monopolist can exploit this concern by sequencing sales. We avoid a problem of multiple equilibria by assuming that (i) the monopolist has

[4] "A dynamic adoption process, however, introduces a strategic consideration that is absent in the static game. Individuals who chose to enter early may influence the entry decisions of others who have not yet entered. This creates the possibility that early entrants may launch a domino chain reaction of widespread adoption", – Ochs and Park (2010).

control over the timing of potential entry and (ii) there are "extreme" types, whose decisions do not depend on their beliefs about other consumers' purchase behavior. Dou et al. (2011) consider a monopolist that can divide consumers into different segments and release the product to one segment after the other, but they assume consumers are myopic. Consumers' expectations therefore play no role in their analysis, whereas rational expectations lie at the heart of our mechanism supporting sequential sales.[5]

More broadly, earlier research has looked at a variety of ways that a firm can exploit network effects by adjusting different marketing variables. These include price and advertising (Kalish (1985), Dhebar and Oren (1985), Dockner and Jorgensen (1988), Xie and Sirbu (1995)), introduction of complementary goods (Basu et al., 2003), and release of clone products (Sun et al., 2004). Certain results in Padmanabhan et al. (1997) touch on sequencing, showing that a firm may want to first release a product to experts followed by a lower quality version for novices. However, their analysis focuses on how sequential quality provision can help the firm signal private information about the strength of network effects, something which plays no role in our setting. Our results add to this literature by showing more generally how a firm can exploit network effects through the sequencing of sales.

There is a small literature on optimal sequencing where consumers have private information about product quality, which superficially shares features with our work: a firm serves consumers simultaneously or sequentially, consumers can learn from observing previous sales, and they make one-off purchase decisions. But since this literature does not consider consumption externalities, it provides no insight into how to sequence sales in the presence of network effects.[6] In terms of results, Sgroi (2002) shows that simultaneously serving a group

[5]The focus of Dou et al. (2011) on optimal seeding with price discrimination, is also different from ours. Moreover, they assume that consumers always buy sequentially, even within each segment, which does not allow for a comparison between simultaneous and sequential sales.

[6]The economic mechanism at work with network effects also differs greatly from the case where consumers have private information about quality. There, it does not matter whether consumers are forward looking, and consumers have no incentive to influence one another.

of 'guinea pigs' can prevent an information cascade where all later consumers refrain from buying. Liu and Schiraldi (2012) show that the optimal scheme is often fully simultaneous when prior beliefs are low. Bhalla (2013) suggests instead using simultaneous sales when the firm's updated beliefs about quality are high, if it can adjust its price over time. Aoyagi (2010) argues that a seller should use sequential sales as a means to implement dynamic pricing.[7] These rather mixed conclusions contrast with our clear result supporting sequential sales, suggesting the importance of identifying the cause of interdependencies between consumers (e.g. quality uncertainty or network effects) in any particular setting.

The literature on dynamic platform competition (see recent papers by Cabral (2011), Halaburda et al. (2015)) looks at strategic considerations faced by firms in the presence of network effects, and generally focuses on pricing. Veiga (2015) considers a monopolistic platform in continuous time and, as the aforementioned papers, examines the trade-off between attracting new consumers and exploiting existing ones. Although this literature looks at the dynamic sales problem in the presence of network effects, as our paper does, to the best of our knowledge no earlier work considers firm control over the timing of sales. Thus, our paper adds the timing or sequencing dimension to the well-know price dimension from the analysis of dynamic platforms.

Aoyagi (2010) is the closest to our paper in terms of results, showing the optimality of sequential sales and targeting more independent consumers first. His results, however, apply in a very different setting, and are due to a very different economic mechanism.[8] There, the payoff from buying depends directly on the signals received by other consumers, but not on their actual purchase decisions. Thus, what matters is the quality of the good, not any

[7]These papers relate to a broader literature on how firms can influence social learning, through means such as pricing or product testing (see, e.g., Ottaviani and Prat (2001), Bar-Isaac (2003), Bose et al. (2006), Bose et al. (2008), Gill and Sgroi (2008), Gill and Sgroi (2012)).

[8]Our analysis also differs from Aoyagi (2010) in suggesting what sales scheme to use when a firm has incomplete control over sequencing, in particular whether to serve smaller or larger cohorts first, and in making the link with consumer communication.

network effect. Consumers there are essentially backwards looking: they observe the behavior of previous consumers, and use this information to attempt to infer these consumers' signals. In contrast, the payoff from buying in our paper does not depend directly on how much other consumers value the product, but instead on how many consumers chose to buy. The crucial point is that consumers are forward looking, and network effects give then an incentive to influence each others' behavior. Finally, unlike in Aoyagi (2010), dynamic pricing is not essential for obtaining the optimality of sequential sales in our setting.

The rest of the paper is organized as follows. Section 2 presents the model. Section 3 contains the main analysis, looking in turn at simultaneous versus sequential schemes, the ordering of sales, and the connection with direct consumer communication. Section 4 discusses issues of robustness, and Section 5 then concludes. Proofs of all Propositions and of technical lemmas can be found in the appendix.

2 Model

A seller of a network good faces a market of n consumers who each have unit demand. Consumers differ in their type (θ, λ), where subscript i denotes the type of consumer i. Both dimensions of type are drawn independently, θ_i from a uniform distribution $U \sim [\underline{\theta}, \overline{\theta}]$ and λ_i from a distribution F on $(0, \overline{\lambda})$, with $\overline{\lambda} \equiv \overline{\theta} - v_0$. We assume for the main analysis that both dimensions of type are private information, but later relax this assumption to explore the situation where values of λ are publicly known.

If consumer i buys a unit of the good, then his payoff consists of an intrinsic and a network component,

$$\theta_i + \frac{\lambda_i}{n-1} \sum_{j \neq i} x_j - p, \tag{1}$$

where $x_j = 1$ if consumer j buys and $x_j = 0$ if he does not, and where p is the price. The consumer's intrinsic payoff from buying is θ_i, and his network payoff from buying is proportional to the number of other consumers who buy. Thus, λ_i captures the weight consumer i places on this network effect, or equivalently his sensitivity to other consumers' purchases. If consumer i does not buy, then he obtains payoff v_0 from his outside option, where $v_0 < \overline{\theta}$. This constraint on v_0 implies that consumers with a sufficiently high intrinsic payoff will always choose to buy if the seller sets a sufficiently low price. We assume that consumers' purchase decisions are irreversible. An interpretation of irreversibility is that a consumer has an urgent need for the good or for a suitable alternative. The consumer either buys the good from the seller or exits the market by purchasing a default option, which gives a payoff of v_0, without the possibility to reenter the market in the short run. Another interpretation is that irreversible purchase decisions may be supported by an "exploding offers" strategy employed by firms (see, Armstrong and Zhou (2015)).[9]

The seller sets a price and selects a sales scheme, which determines the extent to which consumers buy simultaneously or sequentially. Specifically, at $t = 0$, the seller chooses a number of cohorts $m \leq n$ and how to partition the n consumers between the m different cohorts, $I = \{I_1, \dots, I_m\}$. We do not restrict a priori the set of all possible sales schemes. However, because type is unobservable, the seller cannot distinguish between different consumers. This means that the seller's choice of sales scheme I is effectively a choice of m (the number of cohorts) and the cardinality of $I_1, \dots, I_m$ (the size of each cohort).

We assume that the seller commits to its choice of sales scheme. Moreover, for the main analysis, we assume static pricing, where the seller fixes $p \geq 0$ at $t = 0$.[10] We also show that our conclusions remain unchanged if the seller is unable to commit, and that the mechanism driving our results will continue to function with dynamic pricing.

[9]This assumption is consistent with our desire to model situations where the seller has at least some control sequencing. Making the alternative assumption that purchase decisions are reversible would potentially expose us to the problem of multiple equilibria, as in Ochs and Park (2010), which would significantly complicate our analysis (see the discussion following Proposition 1).

[10]One possible reason for static pricing is consumer fairness concerns, in the sense that consumers may consider price changes to be unfair. For further discussion, see Dou et al. (2013) and the references therein.

Consumers make their purchase decisions as follows. At $t = 1$, all consumers in cohort I_1 simultaneously decide whether to buy a unit of the good. Similarly, for any period t with $2 \leq t \leq m$, all consumers in cohort I_t simultaneously decide whether to buy, having observed the choice of consumers in all previous cohorts $I_{t'}$ for $t' \leq t-1$. The game ends after consumers in cohort I_m make their purchase decisions.

For consumer $i \in I_t$, the relevant history is the number of consumers in cohorts $I_1, \ldots, I_{t-1}$ who bought the good. Denote this number by K_t. For a given sales scheme I such that $i \in I_t$, and a given price p, the strategy of consumer i is a decision rule that, for any K_t, specifies whether or not to buy, $x_i = 0$ or $x_i = 1$. The seller's strategy is a choice of p and I.

We look for a perfect bayesian equilibrium where the strategy of consumer i maximizes his expected payoff, for any history K_t. All expectations follow from Bayes' rule and other consumers' equilibrium strategies.[11] The seller's strategy maximizes expected profits, $p \sum_{1 \leq i \leq n} \mathbb{E}(x_i)$, where we focus on ranking different sales schemes and solving for the optimal I. We assume $\underline{\theta} + 1 < v_0 < \overline{\theta}$ to guarantee interior solutions as described below.

3 Analysis

To begin the analysis, we take some preliminary steps to describe consumers' incentives. Consider some consumer i who must decide whether or not to buy after observing sales from previous cohorts. Suppose that consumer i is in cohort I_t, that he observes K_t previous sales, and that the price is p. Let N_t denote the number of consumers who will buy in his own cohort I_t, and let $N_{t'}$ denote the number of consumers who will buy in a later cohort $I_{t'}$. Neither N_t nor $N_{t'}$ are known to consumer i, so his purchase decision will depend on how he expects other consumers to behave. Consumer i will find it optimal to buy himself if

[11] After any particular history, seller and consumer beliefs about the type of consumers who have yet to act are always given by the prior. Thus, our solution will closely resemble a subgame perfect equilibrium, where the role of unobservable type is to generate demand uncertainty.

$$\theta_i + \frac{\lambda_i}{n-1}\left(K_t + \mathbb{E}(N_t - x_i|K_t) + \sum_{t+1\leq t'\leq m} \mathbb{E}(N_{t'}|K_t, x_i = 1)\right) - p \geq v_0. \tag{2}$$

The left-hand-side of (2) gives consumer i's expected payoff from buying, which follows from (1). It consists of the intrinsic payoff from buying, θ_i, minus the price, plus the expected network payoff, which depends on three components: previous sales, K_t, expected sales from the current cohort, and expected sales from later cohorts. Consumer i's observation of previous sales will affect both his own behavior and the number of other consumers he expects to buy. By the same reasoning, all consumers in later cohorts will also observe whether consumer i chose to buy before making their own choice, which means that consumer i's action can influence their behavior. This is why the final expectation in (2) is conditional on consumer i's decision to buy, $x_i = 1$. Consumer i will find buying optimal if the left-hand-side of (2) exceeds v_0, the payoff from his outside option.

Expression (2) shows that the incentive for any consumer i to buy is increasing in his intrinsic payoff from buying. Substituting expected demand from each cohort into this expression and rearranging, the best response of consumer $i \in I_t$ after history K_t is to buy if and only if $\theta \in [\theta_i^*, \overline{\theta}]$, where

$$\theta_i^*(\lambda_i) = v_0 + p - \frac{\lambda_i}{n-1}\left(K_t + \sum_{j\in I_t\setminus\{i\}} \mathbb{E}(x_j|K_t) + \sum_{t'\geq t+1}\sum_{j\in I'_t} \mathbb{E}(x_j|K_t, x_i = 1)\right). \tag{3}$$

Consumer i uses a cut-off strategy, in the sense that he buys if θ exceeds a threshold value given by the right-hand-side of (3). This cutoff depends on the particular history he observes and on his value of λ. The consumer with $\theta = \theta_i^*(\lambda_i)$ earns exactly v_0 from buying which leaves him indifferent with his outside option.

The probability that consumer i will buy after history K_t, from the perspective of those who observe the history but are uncertain about his type, is

$$\mathbb{E}(x_i|K_t) = \frac{\overline{\theta} - \theta_i^*}{\overline{\theta} - \underline{\theta}}, \tag{4}$$

where $\theta_i^* \equiv \mathbb{E}_\lambda(\theta_i^*(\lambda))$ is the expectation of (3) taken with respect to λ. We now verify that $0 < \mathbb{E}(x_i|K_t) < 1$. This means that we have interior solutions where the probability of buying is always strictly positive but also strictly less than one. The parameter assumptions $\underline{\theta}+1 < v_0 < \overline{\theta}$ combined with (3) directly ensure that $\mathbb{E}(x_i|K_t) < 1$. To see that $\mathbb{E}(x_i|K_t) > 0$, the firm's optimal choice of p is bounded above by the price it would charge a consumer following the best possible history, where all other consumers have bought, who therefore has the highest possible willingness to pay. From (3) and (4), expected profits from this consumer are $\left(\frac{\overline{\theta}-v_0-p+\mathbb{E}(\lambda)}{\overline{\theta}-\underline{\theta}}\right)p$, yielding optimal price $p^* = \frac{\overline{\theta}-v_0+\mathbb{E}(\lambda)}{2}$, where $\lambda < \overline{\lambda}$ implies $p^* < \overline{\theta} - v_0$. The optimal price after any other history therefore satisfies $p \leq p^* < \overline{\theta} - v_0$, where (3) and (4) then yield $\mathbb{E}(x_i|K_t) > 0$.

From $\theta_i^* \equiv \mathbb{E}_\lambda(\theta_i^*(\lambda))$ and (3), write

$$\theta_i^* = u_0 - \frac{\mathbb{E}(\lambda)}{n-1}\left(K_t + \sum_{j\in I_t\setminus\{i\}}\mathbb{E}(x_j|K_t) + \sum_{t'\geq t+1}\sum_{j\in I'_t}\mathbb{E}(x_j|K_t, x_i = 1)\right), \tag{5}$$

where $u_0 \equiv v_0 + p$ denotes a consumer's effective outside option, taking into account the price. Once again (4) and (5) imply $0 < \mathbb{E}(x_i|K_t) < 1$.

From an ex ante perspective, the overall probability that consumer i will buy depends on his probability of buying after a particular history K_t and on the ex ante probability distribution over all possible histories. Our assumption that θ is uniformly distributed reduces the problem from analysing the whole distribution of relevant histories to just the expected number of consumers who will buy, $\mathbb{E}(K_t)$. This assumption makes the analysis tractable, and combined with $\lambda_i \leq \overline{\lambda}$, allows us to establish equilibrium existence and uniqueness.

Proposition 1. *For any sales scheme* I*, the game has a unique perfect bayesian equilibrium. That is, for any consumer $i \in I_t$ and history K_t, the cut-off function $\theta_i^*(\lambda_i)$ is uniquely defined.*

The fact that every seller strategy yields a unique value for expected profits is useful when addressing what scheme it should use. Since the equilibrium strategy profile of consumers is unique, given any choice of price and sales scheme, the seller can unambiguously rank different schemes based on these expected profits. If there were multiple equilibrium consumer strategy profiles consistent with a single scheme, then there would also be multiple values of expected profits consistent with that scheme. The ranking of different schemes could then be ambiguous and depend on factors outside of the formal modeling framework, which would prevent us from providing advice about the best way to sequence sales.

3.1 Sequential versus simultaneous sales

Given uniqueness, we are now in a position to examine whether the seller should employ simultaneous or sequential sales. In order to do so we make the following definition.

Definition 1. *A sales scheme* I' $= \{I_{1'}, \ldots, I_{t'}, \ldots, I_{m'}\}$ *is* more sequential *than another scheme* I $= \{I_1, \ldots, I_t, \ldots, I_m\} \neq$ I' *if any two consumers in the same cohort under* I' *are also in the same cohort under* I*:* $i \in I_{t'}$ *and* $j \in I_{t'}$ *implies* $i \in I_t$ *and* $j \in I_t$*, for some* $t \in \{1, \ldots, m\}$.

We say that a second sales scheme is more sequential than a first if all consumers who were served sequentially and at least some who were served simultaneously in the first scheme are served sequentially in the second. This is equivalent to saying that the first scheme can be transformed into the second by repeatedly breaking up cohorts, taking groups of consumers who were served simultaneously and instead serving some of these consumers before others. Alternatively, the second scheme can be transformed into the first by repeatedly combining together cohorts that, loosely put, lie next to one another.

Definition 1 allows us to make pairwise comparisons between many sales schemes in an intuitive way. Applying the definition, every possible scheme is more sequential than a scheme with all consumers in one cohort (fully simultaneous), and a scheme with a single consumer per cohort (fully sequential) is more sequential than every other possible scheme.

The following result says that the seller should use a sales scheme that is as sequential as possible.

Proposition 2. *Suppose that a sales scheme* I' *is more sequential than another scheme* I, *according to Definition 1. Then* I' *delivers strictly higher expected profits.*

This result supporting sequential sales is relevant for a variety of situations with imperfect control over sequencing. For example, when a firm launches a product across M different markets, it may be constrained to serve consumers simultaneously within each market. But a firm can still choose whether to launch the product sequentially across markets. The set of feasible schemes then corresponds to those partitions with $M' \leq M$ cohorts that place all consumers in each market within the same cohort. Proposition 2 says that the most profitable scheme has exactly M cohorts, each corresponding to a single market. Interpreted in this way, a firm should follow a waterfall launch strategy, releasing the product across markets sequentially. Proposition 2 holds regardless of whether the firm sets the optimal price for each scheme or simply sets the same price for both schemes.

A sequential scheme provides consumers with increased information about each others' purchases, essentially making their decisions visible to one another. This visibility can allow success to breed success. High sales from consumers who are served first can then encourage increased sales from consumers served later. The Proposition shows that sequential sales increase expected profits despite the fact that failure can also breed failure, where low early sales can depress sales from those who follow.

The intuition for the result is as follows. With sequential sales, consumers not only observe earlier purchases, but they also realize their own purchases will be observed by later consumers. The very fact of being observed makes buying more attractive, since consumers who are served early understand that those who see them buy will become more likely to buy themselves. The key formal point is that the expectations in (5) for consumers in cohorts $t' \geq t+1$ all condition on the purchase of consumer i in cohort t. It follows that a sequential strategy will tend to yield high initial sales, precisely because consumers are rational and forward looking, starting a virtuous cycle where early success is then compounded.

An immediate corollary of Proposition 2 is that a fully simultaneous scheme is the worst possible choice. Every sales scheme is more sequential than a fully simultaneous scheme, according to Definition 1, and will give strictly higher expected profits. Our mechanism suggests that if a firm has any influence at all over sequencing, no matter how small, then it should use this influence to ensure that at least some consumers are served before others.

Another consequence of Proposition 2 is that we can describe the optimal scheme.

Corollary 1. *The sales scheme* I *that maximizes expected profits has a single consumer per cohort.*

The optimal sales scheme is purely sequential with one consumer served after another. This result is relevant for a firm with perfect control over sequencing, for example one engaged in business-to-business transactions that can choose the precise order to approach potential clients. In this case, Corollary 1 says that the firm should approach clients one by one. Unlike the literature on seeding, the firm does not attempt to kickstart product adoption by giving away its product to some clients to strengthen network effects. The firm instead tries to strengthen network effects and promote early sales by exploiting consumers' strategic incentives to influence one another through their purchases.

Our results so far show that the ranking of sales schemes is independent of the exact distribution of λ, the weight consumers place on the network payoff. However, the distribution of λ does have a quantitative impact on the strength of the link between sequential sales and expected profits. Sequential sales help the seller by exploiting network effects between consumers. The stronger these effects, the larger the impact we would tend to expect from a sequential scheme. We now show simulation results that provide support for this idea. For all i, we set $\lambda_i = \lambda$ with probability one, and calculate expected profits under both a fully simultaneous and a fully sequential scheme for different values of λ.

Figure 1

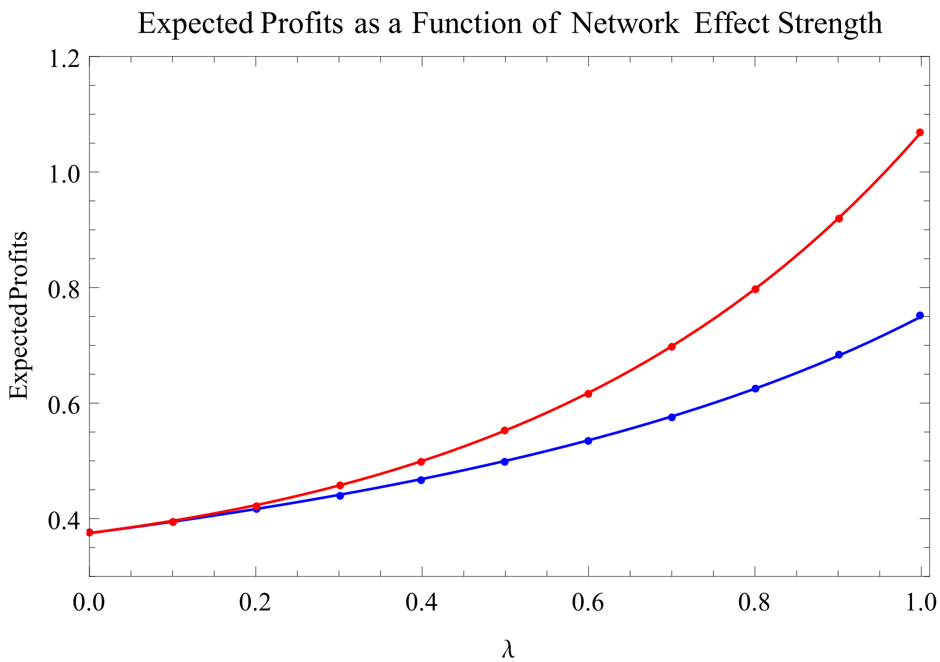

Figure 1 plots expected profits as a function of λ under the two schemes, with the simultaneous scheme represented in blue and the sequential scheme in red. As required by Corollary 1, expected profits are higher under a sequential scheme for all values of $\lambda > 0$. Figure 1 also shows that the difference in expected profits between the two schemes is increasing in λ. In particular, when $\lambda = 1$, a sequential scheme allows the seller to increase expected profits by approximately 40%.[12]

A fully sequential sales scheme will always maximize expected profits, but intuition suggests that it should also increase downside risk. Low early sales under a sequential scheme can be self reinforcing because they are observed by later consumers. This possibility that failure breeds failure might drive down realized profits if early consumers happen to have low willingness to pay. Even though a sequential scheme performs best on average, this potential for downside risk might concern a seller who is risk averse.

[12]The simulation uses parameter values $n = 5$, $\overline{\theta} = 2$, $\underline{\theta} = 0$, $u_0 = 1.85$, and $p = 1$ under both schemes.

Contrary to this intuition, we now present simulation results showing that a sequential scheme may actually be less risky than a scheme that is fully simultaneous. In fact, Figure 2 shows that for certain parameter values, the distribution of profits under a sequential scheme is unambiguously better for the seller than the distribution under a simultaneous scheme, in the sense of first order stochastic dominance.

Figure 2

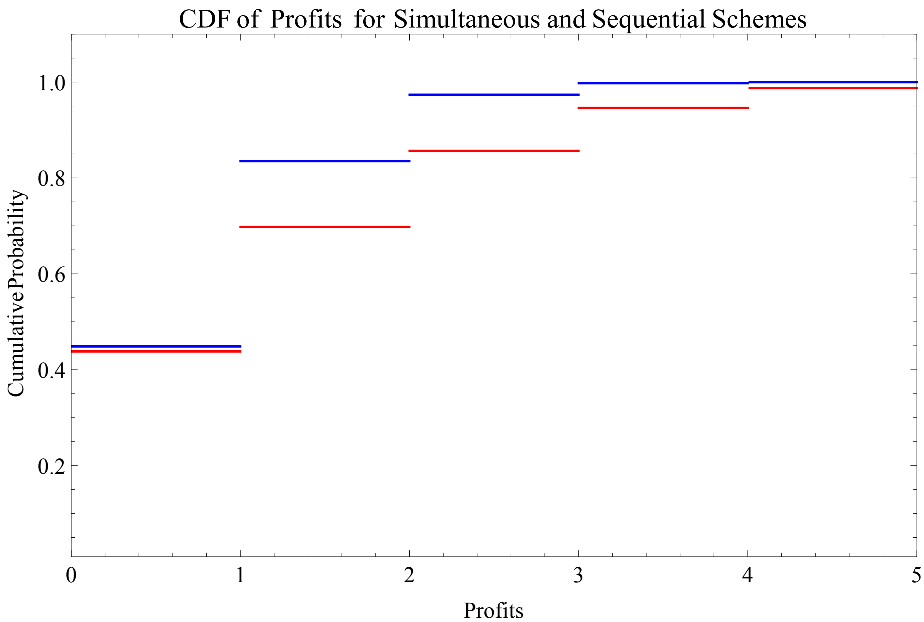

Figure 2 plots the cumulative distribution function of total profits under the two different schemes. The simultaneous scheme is represented in blue and the sequential scheme in red, where the former CDF lies entirely above the latter.[13] A sequential scheme here serves the dual purpose of increasing expected profits while decreasing the probability of a poor outcome where realized profits are very low. The positive incentive effect on early consumers is so strong that it outweighs any increased risk that might arise from low early sales.

[13]The simulation uses parameter values $n = 5$, $\overline{\theta} = 2$, $\underline{\theta} = 0$, $u_0 = 1.85$, $p = 1$, and $\lambda = 1$ under both schemes.

3.2 Ordering of sales

The analysis in Section 3.1 shows that a firm should use a sales scheme that is as sequential as possible. But if not all feasible schemes are comparable in the sense of Definition 1, then the question remains as to precisely which scheme is best. For example, Proposition 2 suggests that a firm should launch its product sequentially across markets, but markets may well differ in their size. Should the firm then first release the product in smaller markets or in larger ones? In our setting, this amounts to asking whether the firm should use a scheme that serves smaller or larger cohorts first.

Proposition 3. *Suppose* $\mathbb{E}(\lambda) < v_0 - \underline{\theta}$*. Consider sales schemes* $I = \{I_1, \ldots, I_t, I_{t+1}, \ldots I_m\}$ *and* $I' = \{I_1, \ldots, I'_t = I_{t+1}, I'_{t+1} = I_t, \ldots I_m\}$ *with* $|I_t| \equiv n_t > n_{t+1} \equiv |I_{t+1}|$*. Then* I' *yields strictly higher expected profits than* I*.*

This result provides additional insights into how to sequence sales when certain potential schemes are infeasible. An implication of Proposition 2 is that a firm should launch its product in one market after another. Any such sequential launch will yield higher expected profits than launching in different markets at the same time. Proposition 3 goes further by saying that the firm should carry out this sequential launch in increasing order of market size. The intuition behind this result is that if larger markets move later, more consumers possess valuable information about the decisions of others. Early movers understand that their actions influence the many consumers who follow and become more likely to buy themselves.

Moving to our second result about ordering, we will relax the assumption that consumer type is private information. Intuitively, if a firm can observe certain consumer characteristics, then it may well take this information into account when choosing its sales scheme. Corollary 1 showed that given complete control over sequencing, the optimal scheme is fully sequential. We now explore whether this remains the case when the seller can distinguish between different consumers, and in particular examine which consumers should be served first.

The literature on word-of-mouth communication in networks has examined a related question from the point of view of consumer influence. Typically in this literature, a firm initially informs certain consumers about its product, these consumers pass this information along to others, those consumers pass this information along in turn, and so on. The issue for the firm is who to initially inform, in particular if it can distinguish between consumers with different propensities to pass along information. This propensity captures the strength of a consumer's influence in the network.[14]

In our setting, all consumers are equally influential from an ex ante perspective, in the sense that network effects depend on total sales but not on the identity of the consumers who buy. Not all consumers however are equally easy to influence. Consumers with high values of λ place a high weight on the network payoff which makes them more sensitive to the purchase behavior of others. In contrast, consumers with low values of λ base their purchase decisions mainly on their intrinsic payoff from buying. To explore the issue of ordering and consumer influence, we now assume that each consumer's value of λ is public information. The following result shows that the optimal sales scheme then serves consumers in increasing order of λ, so in decreasing order of their sensitivity to other consumers' influence.

Proposition 4. *Suppose the weight consumers place on the network payoff, λ, is observable. Then the sales scheme* I *that maximizes expected profits has a single consumer per cohort, increasingly ordered in λ, i.e. $\lambda_1 \leq \ldots \leq \lambda_n$.*

The optimal sales scheme remains purely sequential, where the intuition for this result echoes that from Proposition 2 and Corollary 1. There, we argued that the qualitative advantage of a sequential scheme does not depend on the precise distribution of λ. In a similar way, this qualitative advantage does not depend on the exact realized values of these weights. As long as each consumer places a strictly positive weight on the network payoff,

[14] For example, Galeotti and Goyal (2009) suggest targeting influential consumers who will inform many friends, whereas Campbell (2013) show this may be suboptimal if these influential consumers are likely to already be informed via word of mouth.

then a sequential scheme will increase expected profits by increasing visibility, pushing early consumers to buy, and allowing success to breed success.

In addition, Proposition 4 derives the optimal ordering: consumers should be served in increasing order of the weight they place on the network payoff. This result complements those in the literature on word-of-mouth communication in networks stating that a firm should first serve consumers with the most influence. This result also echoes the notion that a firm launching a new product should target independent-minded consumers first, who can serve as opinion leaders for those who follow. These innovators (low λ) will decide whether or not to buy the product largely based on their own personal tastes. Their decision to buy can then encourage imitators (high λ) who care about their actions to jump on the bandwagon.

When values of λ are observable, the seller faces a new trade off. Serving consumers in increasing order of these weights means that later consumers (high λ) have a strong incentive to follow those who buy. In principle, doing so reinforces the benefits when early consumers buy and success breeds success. However, these benefits are limited by the fact that early consumers (low λ) do not become much more likely to buy just because they expect others to follow. Another way to understand the trade off is that consumers with high weights are likely to set a good example, but they are also more likely to follow a good example once it has been set. Proposition 4 shows that the second effect outweighs the first so the optimal order is increasing in these weights. The mechanism behind sequential sales is based on the idea that consumers want to influence one another, but the optimal scheme grants the largest visibility to consumers who care the least about this influence.

One surprising feature of the optimal scheme is that the seller first serves consumers believed to have the lowest willingness to pay. Another is that the optimal ordering's impact on expected profits can be non-monotonic in the difference between consumers' values of λ. To take a simple case, consider two consumers, with weights λ_1 and $\lambda_2 > \lambda_1$. Proposition 4 says that serving consumer 1 before consumer 2 will maximize expected profits. But as λ_1

approaches λ_2 from below, the consumers become increasingly similar, and expected profits in the limit are the same regardless of who is served first. If λ_1 instead approaches zero from above, then consumer 1 cares little about consumer 2's expected behavior, and expected profits in the limit do not depend on the firm's choice of sales scheme. The implication is that ordering can matter most if consumers are neither too similar nor too different.

3.3 Consumer communication

Typically, models of sequential decision making with private information assume that consumers cannot directly communicate, and all information transmission takes place indirectly via observing each others' purchases. For example, in the literature on quality uncertainty and social learning, consumers cannot directly share the private signal they receive about quality, and other consumers only update their beliefs about quality by observing the level of previous sales. We make a similar assumption in our analysis by assuming that consumers cannot directly communicate their willingness to pay. This assumption is reasonable in many situations where market interactions are anonymous.

But for a variety of products, including books, films, mobile phones, and computers, consumers do share information in online forums and communities (see, e.g., Godes et al. (2005) and the references therein). This information sharing can pertain to new products that consumers have purchased, but also to products that are unreleased. For example, there are various online sites where consumers engage in heated debate about the perceived merits of rumored Apple products that have yet to appear.[15]

Consumer communication is relevant in our setting because it may serve as a possible substitute for sequential sales. The whole purpose of sequential sales is to help consumers learn from one another about a product's popularity. However, if consumers successfully learn each others' willingness to pay through communication, then there is little scope for

[15] See for example www.9to5mac.com and www.appleinsider.com

future learning, and little need for the visibility of purchases provided by sequential sales. Consumers who successfully communicate would be able to correctly predict the good's popularity regardless of the seller's choice of sales scheme.

A crucial point for communication to be successful relates to credibility. If consumers read certain comments or reviews about a product, should they actually believe what they read? One concern here is the potential for firm manipulation. Previous work has explored how a firm strategically post positive reviews about its own products to influence consumer beliefs; if consumers realize this, it will naturally reduce the credibility of the information they receive (Dellarocas (2006), Mayzlin et al. (2014)). In what follows we take an alternative approach focusing more on consumers. Rather than looking at firm manipulation, we examine another potential obstacle to credible communication: possible incentives for consumers to misreport.

Specifically, we consider two reasons why consumers might want to misrepresent their willingness to pay to one another. Information that consumers share may be collected by the firm and used to adjust the price (Chen and Xie, 2008). A consumer who understates his willingness to pay may contribute to the impression that demand is low, leading the firm to reduce its price. On the other hand, a consumer who overstates his willingness to pay may convince other consumers to buy, resulting in a larger network payoff, which can also help that consumer. This reasoning suggests that network effects might push consumers to exaggerate whereas firm monitoring might generate countervailing incentives. We now show that despite these potential obstacles, consumers may be able to communicate truthfully.

Formally, we assume again that type is private information, but allow consumers to engage in cheap talk before making their purchase decisions. Consumers simultaneously send one another a message about their type, where the seller observes the set of messages with strictly positive probability. If the seller observes the messages then it can use this information when setting its price. The details of this price-setting process are not crucial for our results. The important point is just that the price be non-decreasing in the seller's updated beliefs about consumer willingness to pay, conditional on observing the messages.

Proposition 5. *Consider a simultaneous sales scheme, with all consumers in the same cohort. Suppose that before buying, consumers can simultaneously send a message* $m \in [\underline{\theta}, \overline{\theta}] \times (0, \overline{\lambda})$ *about their type which all other consumers observe, and where the seller observes* $M = (m_1, \ldots, m_n)$ *with probability* $q > 0$*. Furthermore suppose that the seller sets price* p^* *if it does not observe* M*, and sets price* $p(M)$ *if it does, where* $p(M)$ *is non-decreasing in* $\sum_{i=1}^{n} \mathbb{E}(x_i | p^*, M)$*. Then when* q *is sufficiently small, an equilibrium exists where communication is informative, in the sense that each consumer truthfully reveals to all others the minimum level of total sales required for him to buy himself at price* p^**. In the limit as* q *tends to zero, consumer purchase decisions approach those in a setting where consumers all observe each others' type,* (θ_i, λ_i) *for all* $i = 1, \ldots, n$*.*

Proposition 5 shows that potential incentive problems need not rule out successful communication, in that consumers may still truthfully reveal their planned purchase behavior to one another. However, for such communication to occur, consumers must believe it sufficiently unlikely that the seller is monitoring their messages. Curiously enough, successful communication is possible precisely because of network effects, even though they seemingly provide consumers with an incentive to exaggerate. If there were no network effects, and the seller's price was increasing in consumers' messages, then consumers would all claim low willingness to pay in the hopes of obtaining a price reduction.

Intuitively, a consumer who understates his willingness to pay can generate two effects. The first effect is that other consumers infer demand may be relatively low, making them less likely to buy themselves, which reduces aggregate demand at any given price. The resulting reduction in the network payoff means that consumers who buy are left worse off. The second effect of understating is that the seller may respond by charging a lower price. This price reduction would leave consumers better off but can only occur if the seller observes the messages. When the probability that the seller observes the messages is relatively small, the first effect dominates the second, and consumers have a strict incentive not to understate.

By a similar logic, overstating willingness to pay can lead to a higher price, but can also convince more consumers to buy and increase aggregate demand. The latter effect will lead to a larger network payoff for those who buy. However, the consumer who overstates will only benefit from this network payoff if he has a genuine incentive to buy himself. And if he has such a genuine incentive, then there was no reason to overstate willingness to pay.

The implication of Proposition 5 is that firms selling products where potential consumers regularly communicate may need to focus less on a sequential product launch than firms for which the opposite is true. Successful communication can reduce uncertainty and leave consumers with relatively little to learn from one another through sequential sales. It may be tempting to also conclude that firms should actively facilitate discussion and encourage consumer communication about new products, for example through an official online forum or discussion board. However, if the firm's involvement in this process leads consumers to suspect it is monitoring their messages, then this can derail successful communication, even absent any concern that the firm is strategically manipulating messages.

4 Discussion and Robustness

Our analysis has assumed that the seller commits to its choice of sales scheme, solutions are interior, and the intrinsic payoff θ is uniformly distributed. We now briefly comment on how each of these assumptions relates to our results supporting sequential sales. We then relax the assumption of static pricing, instead allowing the seller to adjust its price over time, and present a result suggesting the mechanism supporting sequential sales will continue to apply.

The fact that the seller can commit to a sales scheme is unimportant for the results. The analysis shows that for any cohort I_t with at least two consumers, given any history K_t, the seller always benefits by having some of these consumers act before the others. This means that a seller who chooses a sequential scheme at $t = 0$ has no incentive to change its mind after observing the actions of any number of consumers. If the first consumers do not buy, then the seller may well regret ex post using this scheme, but it will still prefer the remaining consumers to act sequentially.

Our assumption on parameter values ensures that after any history, the probability a consumer will buy lies strictly between zero and one, so that the equilibrium of the consumer game is interior and unique. Relaxing this assumption would mean that multiple values of expected profits could be consistent with each scheme, as discussed following Proposition 1. If network effects were sufficiently strong, then any sales scheme could generate both a good equilibrium outcome where all consumers buy and a bad equilibrium outcome where nobody buys. It would then be difficult to rank different schemes, but sequential sales might still be useful in helping with equilibrium selection, if observing an early purchase can coordinate the remaining consumers on the Pareto dominant outcome.

Assuming a uniform distribution of θ guarantees equilibrium existence and uniqueness, as discussed prior to Proposition 1. It also has an effect that relates to the variance of early sales. Intuitively, variance can be quite high under a sequential scheme, since consumers can observe and imitate one another. This reasoning suggests that in comparison with simultaneous sales, a sequential scheme may tend to generate more extreme histories.

The variance of early sales plays no role when θ is uniformly distributed. All that matters about early sales is their expected value, which is maximized under a sequential scheme. However, variance could potentially matter if θ followed a different distribution. For example, if many consumers had low θ, and only a very good history would persuade them to buy, then high variance could help by increasing the probability of such a history. If instead many consumers had high θ, so only a very bad history would dissuade them from buying, then high variance could hurt by the same reasoning. Our analysis would then underestimate the benefit of sequential sales in the first case but overestimate it in the second case.

We now turn to dynamic pricing and address whether sequential sales will remain attractive to a seller that can adjust its price over time. For example, the seller may increase the price from its initial level if early sales are high or decrease the price if early sales are low. A fully general analysis of dynamic pricing is complicated by the fact that the seller jointly chooses a price schedule and a sales scheme, and the preferred prices will vary across different schemes. For any given sales scheme, the analysis would involve considering all potential prices to charge each cohort, for every possible history, and then comparing the resulting profits across all possible schemes. An additional complication is that the optimal schedule will depend on whether the seller can commit to future prices. Commitment means that the seller fixes a price schedule at $t = 0$ so as to maximize expected profits from an ex ante perspective. No commitment means that the seller effectively makes a sequence of pricing decisions over time when facing each cohort, where the chosen price must maximize expected profits from that particular cohort and all later consumers, given observed sales. We analyze the former case in the following Proposition and leave the latter for further research.

Proposition 6. *Suppose the seller can commit to a dynamic pricing schedule, with price $p(K_t)$ for cohort t conditional on previous sales K_t. Then a fully sequential sales scheme (a single consumer per cohort), delivers higher expected profits than a fully simultaneous sales scheme (all consumers in a single cohort).*

Fully sequential sales remain more profitable than fully simultaneous sales under dynamic pricing, just as under static pricing. Dynamic pricing actually increases the difference in expected profits between these schemes because sequential sales now offer additional flexibility, allowing the seller to adjust the price depending on whether early sales were high or low. With dynamic pricing, the seller can always earn the same profits as under static pricing by maintaining its initial price, but can generally do better still by adjusting its price over time.

5 Conclusion

In a setting with network effects, consumers looking to buy a product will naturally take into account whether they expect others to buy as well. Consumers may be more willing to buy mobile phones or video games, movies tickets or books, if they believe that sales will be high. Firms may be more willing to adopt a new business solution or employ a new operating system if they expect others to follow suit. Put another way, buying becomes more attractive if the product in question will likely become a 'hit'. We consider a setting where consumers have precisely such concerns, and examine how a firm's choice of sales scheme, in particular between simultaneous and sequential sales, can help exploit network effects.

Our results show that a firm's choice of sales scheme can matter a great deal. Broadly speaking, the sequencing of sales affects how much information consumers receive about each others' behavior. The advantage of a sequential sales strategy is that success can breed success, but the disadvantage is that failure can breed failutre.

Despite this apparent trade-off, we show that a firm can always increase its expected profits by moving from one sales scheme to another that is more sequential. The key point is that consumers are rational and forward looking, and a sequential scheme affects their expectations about how others will behave in the future. A consumer knows that others who observe his purchase will become more likely to buy, which increases the incentive to buy himself. The use of a sequential scheme not only reveals to consumers whether or not the product is a hit, it also makes a hit more likely in the first place.

From a practical perspective, these results provide support for a sequential product-launch strategy, where a firm first releases its product in smaller markets before moving on to larger ones. They also suggest that a firm with full control of sequencing should approach potential clients in a way that is purely sequential, one after the other. Looking more closely at the economic mechanism behind these results, we argue that sequencing will tend to be less important if consumers can regularly communicate with each other, for example through online forums and discussion boards. We also present results showing that if possible, a firm should first serve more independent-minded consumers who are less sensitive to other consumers' behavior. Doing so is consistent with firms' targeting of opinion leaders whose initial take-up of a product can help generate later success.

6 Appendix*

6.1 Technical Lemmas*

Lemma A.1. *Suppose K_t consumers buy up until cohort I_t, and consider consumer $j \in I_{t'}$ with $t' \geq t+1$. Suppose a set of consumers $M \subseteq \cup_{l=t}^{t'-1} I_l$ choose to buy. Then*

$$\mathbb{E}(x_j|K_t, M) = \frac{\overline{\theta} - \mathbb{E}(\theta_j^*|K_t, M)}{\overline{\theta} - \underline{\theta}},$$

where

$$\mathbb{E}(\theta_j^*|K_t, M) = u_0 - \frac{\mathbb{E}(\lambda)}{n-1}\left(K_t + \sum_{t \leq l \leq t'} \sum_{i \in I_l \setminus \{j\}} \mathbb{E}(x_i|K_t, M) + \sum_{l \geq t'+1} \sum_{i \in I_l} \mathbb{E}(x_i|K_t, M, x_j = 1)\right).$$

Proof. By (5), for any $K_{t'}$, the relevant cutoff for consumer $j \in I_{t'}$ is

$$\theta_j^* = u_0 - \frac{\mathbb{E}(\lambda)}{n-1}\left(K_{t'} + \sum_{i \in I_{t'} \setminus \{j\}} \mathbb{E}(x_i|K_{t'}) + \sum_{l \geq t'+1} \sum_{i \in I_l} \mathbb{E}(x_i|K_{t'}, x_j = 1)\right), \tag{6}$$

so that

$$\mathbb{E}(x_j|K_{t'}) = \frac{\overline{\theta} - \theta_j^*}{\overline{\theta} - \underline{\theta}}. \tag{7}$$

We now work with (6) and (7) to obtain $\mathbb{E}(x_j|K_t, M)$. Let $\mathcal{K}$ be the set of all $K_{t'}$ consistent with (K_t, M). For each $K_{t'}$ we multiply (6) with $p(K_{t'}|K_t, M)$ and sum up over all $K_{t'} \in \mathcal{K}$. Since $\sum_{\mathcal{K}} p(K_{t'}|K_t, M) = 1$, we have that $\mathbb{E}(\theta_j^*|K_t, M)$ is equal to

$$u_0 - \frac{\mathbb{E}(\lambda)}{n-1}\left(\sum_{\mathcal{K}} K_{t'} p(K_{t'}|K_t, M) + \sum_{i \in I_{t'} \setminus \{j\}} \sum_{\mathcal{K}} p(K_{t'}|K_t, M)\mathbb{E}(x_i|K_{t'}) + \right.$$

$$\left. \sum_{l \geq t'+1} \sum_{i \in I_l} \sum_{\mathcal{K}} p(K_{t'}|K_t, M)\mathbb{E}(x_i|K_{t'}, x_j = 1)\right),$$

Note that

$$\mathbb{E}(x_j|K_t, M) = \sum_{\mathcal{K}} p(K_{t'}|K_t, M)\mathbb{E}(x_j|K_{t'}),$$

$$\mathbb{E}(x_i|K_t, M, x_j = 1) = \sum_{\mathcal{K}} p(K_{t'}|K_t, M)\mathbb{E}(x_j|K_{t'}, x_j = 1),$$

and $\sum_{\mathcal{K}} K_{t'} p(K_{t'}|K_t, M) = K_t + \sum_{t \leq l \leq t'-1} \sum_{i \in I_l} \mathbb{E}(x_i|K_t, M)$. Therefore,

$$\mathbb{E}(\theta_j^*|K_t, M) = u_0 - \frac{\mathbb{E}(\lambda)}{n-1}\left(K_t + \sum_{t \leq l \leq t'} \sum_{i \in I_l \setminus \{j\}} \mathbb{E}(x_i|K_t, M) + \sum_{l \geq t'+1} \sum_{i \in I_l} \mathbb{E}(x_i|K_t, M, x_j = 1)\right).$$

□

Lemma A.2. *For any consumer i in cohort I_t with history K_t,*

$$\frac{d\mathbb{E}(x_i|K_t)}{dK_t} > 0.$$

Proof. Write out $\mathbb{E}(x_i|K_t) = \frac{\bar{\theta} - \theta_i^*}{\bar{\theta} - \underline{\theta}}$ with

$$\theta_i^* = u_0 - \frac{\mathbb{E}(\lambda)}{n-1}\left(K_t + \sum_{j \in I_t \setminus \{i\}} \mathbb{E}(x_j|K_t) + \sum_{t' \geq t+1} \sum_{j \in I_{t'}} \mathbb{E}(x_j|K_t, x_i = 1)\right).$$

By Lemma A.1, write out each term in the second summation as $\mathbb{E}(x_j|K_t, x_i = 1) = \frac{\bar{\theta} - \mathbb{E}(\theta_j^*|K_t, x_i = 1)}{\bar{\theta} - \underline{\theta}}$ with

$$\mathbb{E}(\theta_j^*|K_t, x_i = 1) = u_0 - \frac{\mathbb{E}(\lambda)}{n-1}\left(K_t + 1 + \sum_{j' \in I_t \setminus \{i\}} \mathbb{E}(x_{j'}|K_t)\right.$$

$$\left. + \sum_{t+1 \leq l \leq t'} \sum_{j' \in I_l \setminus \{j\}} \mathbb{E}(x_{j'}|K_t, x_i = 1) + \sum_{l \geq t'+1} \sum_{j' \in I_l} \mathbb{E}(x_{j'}|K_t, x_i = 1, x_j = 1)\right).$$

Again by Lemma A.1, write out each term $\mathbb{E}(x_{j'}|K_t, x_i = 1, x_j = 1)$ in the last summation as $\mathbb{E}(x_{j'}|K_t, x_i = 1, x_j = 1) = \frac{\overline{\theta}-\mathbb{E}(\theta^*_{j'}|K_t, x_i=1, x_j=1)}{\overline{\theta}-\underline{\theta}}$, and so on. Consider a player in a cohort $k > t+1$. Let M_k be the subset of players such that (i) each player $i \in M_k$ decided to buy, (ii) for $i, j \in M_k$, $i \in I_{n_i}, j \in I_{n_j}$ $n_i \neq n_j$ and $n_i > t$. Let for $l \leq k$ $M_k^l \subseteq M_k : \forall i \in M_k^l, i \in I_{n_i} \Rightarrow n_i < l$. Then

$$\mathbb{E}(\theta^*_j|K_t, x_i = 1, M_k) = u_0 - \frac{\mathbb{E}(\lambda)}{n-1}\Bigg(K_t + 1 + \#M_k + \sum_{j' \in I_t \backslash \{i\}} \mathbb{E}(x_{j'}|K_t)$$

$$+ \sum_{t+1 \leq l \leq k} \sum_{j' \in I_l \backslash M_k} \mathbb{E}(x_{j'}|K_t, x_i = 1, M_k^l) + \sum_{l \geq k+1} \sum_{j' \in I_l} \mathbb{E}(x_{j'}|K_t, x_i = 1, M_k, x_j = 1)\Bigg).$$

Denoting the number of distinct equations for $\mathbb{E}(x_j|K_t, x_i = 1, M_k)$ by A, including terms with zero coefficient on the right-hand side of each equation, gives a system of A equations in A unknowns. As shown immediately after (4) in Section 3, consumers with θ sufficiently close to $\overline{\theta}$ have a dominant strategy to buy. This means any solution to this system must give $\mathbb{E}(x_i|K_t) > 0$, for any K_t.

Differentiating each equation in the system with respect to K_t gives $\frac{d\mathbb{E}(x_i|K_t)}{dK_t} = \frac{\overline{\theta}-\frac{d\theta_i^*}{dK_t}}{\overline{\theta}-\underline{\theta}}$ with

$$\frac{d\mathbb{E}(\theta^*_j|K_t, x_i = 1, M_k)}{dK_t} = u_0 - \frac{\mathbb{E}(\lambda)}{n-1}\Bigg(1 + \sum_{j' \in I_t \backslash \{i\}} \frac{d\mathbb{E}(x_{j'}|K_t)}{dK_t}$$

$$+ \sum_{t+1 \leq l \leq k} \sum_{j' \in I_l \backslash M_k} \frac{d\mathbb{E}(x_{j'}|K_t, x_i = 1, M_k^l)}{dK_t} + \sum_{l \geq k+1} \sum_{j' \in I_l} \frac{d\mathbb{E}(x_{j'}|K_t, x_i = 1, M_k, x_j = 1)}{dK_t}\Bigg).$$

This system of A linear equations in A unknowns is identical to the first one, except that each conditional expectation is replaced by its derivative, and K_t has been set equal to 1. The associated matrix for this system has diagonal entries of 1 and off-diagonal entries of either 0 or $\frac{-1}{\overline{\theta}-\underline{\theta}} \frac{\mathbb{E}(\lambda)}{n-1} < 0$, where the number of non-zero off-diagonal entries in each row cannot exceed $\sum_{t' \geq t} \sum_{i \in I_{t'}} n_i - 1 \leq n-1$. By $\mathbb{E}(\lambda) \leq \overline{\lambda} = \overline{\theta} - v_o$ and $\underline{\theta} + 1 < v_o < \overline{\theta}$, the sum of the absolute values of off-diagonal entries in each row is therefore strictly less than one. Hence, this matrix is strictly diagonally dominant. By the Gershgorin theorem (1931), the system then has a unique solution, with $\frac{d\mathbb{E}(x_i|K_t)}{dK_t} > 0$.

$\square$

Lemma A.3. *For any consumer $j \in I_{t'}$, with $t' \geq t+1$, $\mathbb{E}(x_j|K_t)$ is strictly increasing in $\sum_{i \in I_t} \mathbb{E}(x_i|K_t)$.*

Proof. We proceed by induction. First, let $t' = t+1$. By Lemma 1, for any $x_j \in I_{t+1}$, write out $\mathbb{E}(x_j = 1|K_t) = \frac{\overline{\theta} - \mathbb{E}(\theta_j^*|K_t)}{\overline{\theta} - \underline{\theta}}$ with

$$\mathbb{E}(\theta_j^*|K_t) = u_0 - \frac{\mathbb{E}(\lambda)}{n-1}\left(\left[K_t + \sum_{i \in I_t} \mathbb{E}(x_i|K_t)\right] + \sum_{i \in I_{t+1} \setminus \{j\}} \mathbb{E}(x_i|K_t) + \sum_{l \geq t+2} \sum_{i \in I_l} \mathbb{E}(x_i|K_t, x_j = 1)\right).$$

Again by Lemma A.1, write out each expectation $\mathbb{E}(x_i|K_t, x_j = 1)$ in the last summation, and so on to generate a system of equations. Each of these equations will include the same expression in square brackets.

We can identify the expression in square brackets with a constant K_{t+1}. A strict increase in $\sum_{i \in I_t} \mathbb{E}(x_i|K_t))$ is then equivalent to a strict increase in K_{t+1}. Hence by Lemma A.2, $\mathbb{E}(x_j|K_t)$ must strictly increase.

Now let $t' \geq t+2$, and suppose the result holds for all cohorts $t+1, \ldots, t'-1$. We show that the result also holds for t'. For a consumer $j \in I_{t'}$, write

$$\mathbb{E}(\theta_j^*|K_t) = u_0 - \frac{\mathbb{E}(\lambda)}{n-1}\left(\left[K_t + \sum_{t \leq l \leq t'-1} \sum_{i \in I_l} \mathbb{E}(x_i|K_t)\right] + \sum_{i \in I_l \setminus \{j\}} \mathbb{E}(x_i|K_t) + \sum_{l \geq t'+1} \sum_{i \in I_l} \mathbb{E}(x_i|K_t, x_j = 1)\right).$$

Once again using Lemma A.1, write out each expectation $\mathbb{E}(x_i|K_t, x_j = 1)$ in the last summation, and so on to generate a system of equations which all include the same term in square brackets. By the induction hypothesis, the term in square brackets strictly increases, which is again equivalent to an increase in $K_{t'}$. By Lemma A.2, $\mathbb{E}(x_j|K_t)$ must strictly increase.

$\square$

Lemma A.4. *Let $t' \leq t-1$. Consider a history $K_{t'}$ and any consumer $i \in I_l$ with $l \geq t'$. Let $\mathcal{K}_{t'}$ be a set of histories K_t consistent with $K_{t'}$, and let $a \in \mathbb{R}$ be some parameter of arbitrary nature. Then if $\frac{d\sum_{j\in I_t}\mathbb{E}(x_j|K_t)}{da} > 0$ for all $K_t \in \mathcal{K}_{t'}$, then $\frac{\mathbb{E}(x_i|K_{t'})}{da} > 0$.*

Proof. First let $t' = t-1$, and consider a consumer $i \in I_{t-1}$, given history K_{t-1}. By (4) and (5), write $\mathbb{E}(x_i|K_{t-1}) = \frac{\bar{\theta}-\theta_i^*}{\bar{\theta}-\underline{\theta}}$ with

$$\theta_i^* = u_0 - \frac{\mathbb{E}(\lambda)}{n-1}\Bigg(K_{t-1} +$$

$$\sum_{j\in I_{t-1}\setminus\{i\}}\mathbb{E}(x_j|K_{t-1}) + \sum_{0\leq K'\leq \#I_{t-1}-1}\mathbb{P}(\sum_{j\in I_{t-1}\setminus\{i\}} x_j = K'|K_{t-1})\sum_{l\geq t}\sum_{j\in I_l}\mathbb{E}(x_j|K_{t-1}+1+K')\Bigg),$$

explicitly writing out all the possible histories K_t consistent with K_{t-1} and $x_i = 1$. Each such history corresponds to a value of $\sum_{j\in I_{t-1}\setminus\{i\}} x_j = K'$, with $K' = 0, \ldots, \#I_{t-1}-1$, representing the possible purchase decisions of the $\#I_{t-1}-1$ consumers in $I_{t-1}\setminus\{i\}$. Equivalently, consumer i will find it optimal to buy if and only if his expected payoff from buying,

$$\theta_i + \frac{\lambda_i}{n-1}\Bigg(K_{t-1} + \sum_{j\in I_{t-1}\setminus\{i\}}\mathbb{E}(x_j|K_{t-1}) + \sum_{0\leq K'\leq \#I_{t-1}-1}\mathbb{P}(\sum_{j\in I_{t-1}\setminus\{i\}} x_j = K'|K_{t-1})\sum_{l\geq t}\sum_{j\in I_l}\mathbb{E}(x_j|K_{t-1}+1+K')\Bigg), \tag{8}$$

exceeds that from his effective outside option, u_0.

Consider an increase in $\sum_{j\in I_t}\mathbb{E}(x_j|K_t)$ for every history K_t consistent with K_{t-1}. This implies an increase in $\sum_{j\in I_t}\mathbb{E}(x_j|K_{t-1}+1+K')$ for all $K' = 0, \ldots \#I_{t-1}-1$. Then by Lemma A.3, $\sum_{l\geq t+1}\sum_{j\in I_l}\mathbb{E}(x_j|K_{t-1}+1+K')$ must increase as well, for all such K'. Since $\lambda > 0$ for all consumers, the system of equations given by (8) defines a game with strategic complements between all consumers i in cohort I_{t-1} (increasing best-response functions). Therefore, if $\sum_{l\geq t}\sum_{j\in I_l}\mathbb{E}(x_j|K_{t-1}+1+K')$ increases, then for each $i \in I_{t-1}$, $\mathbb{E}(x_i|K_{t-1}) = \frac{\bar{\theta}-\theta_i^*}{\bar{\theta}-\underline{\theta}}$ must increase as well (see Vives (1990)). Given this increase in $\sum_{i\in I_{t-1}}\mathbb{E}(x_i|K_{t-1})$, Lemma A.3 implies that $\mathbb{E}(x_j|K_{t-1})$ must also increase, for any consumer j in cohort $l \geq t$.

Proceeding by induction for cohorts $t' = t-2, t-3, \ldots, 1$ completes the proof.

$\square$

6.2 Proofs of the Propositions*

Proposition 1. *For any sales scheme* I*, the game has a unique perfect bayesian equilibrium. That is, for any consumer $i \in I_t$ and history K_t, the cut-off function $\theta_i^*(\lambda_i)$ is uniquely defined.*

Proof. Consider a subgame starting with cohort I_t to act after a history summarized by K_t. By (4) and (5), for each consumer $i \in I_t$, write out $\mathbb{E}(x_i|K_t) = \frac{\overline{\theta}-\theta_i^*}{\overline{\theta}-\underline{\theta}}$, with

$$\theta_i^* = u_0 - \frac{\mathbb{E}(\lambda)}{n-1}\left(K_t + \sum_{j\in I_t\setminus\{i\}} \mathbb{E}(x_j|K_t) + \sum_{t'\geq t+1}\sum_{j\in I_{t'}} \mathbb{E}(x_j|K_t, x_i = 1)\right).$$

By Lemma A.1, write out each term in the second summation as $\mathbb{E}(x_j|K_t, x_i = 1) = \frac{\overline{\theta}-\mathbb{E}(\theta_j^*|K_t,x_i=1)}{\overline{\theta}-\underline{\theta}}$, with

$$\mathbb{E}(\theta_j^*|K_t, x_i = 1) = u_0 - \frac{\mathbb{E}(\lambda)}{n-1}\Bigg(K_t + 1 + \sum_{j'\in I_t\setminus\{i\}} \mathbb{E}(x_{j'}|K_t)$$

$$+ \sum_{t+1\leq l\leq t'}\sum_{j'\in I_l\setminus\{j\}} \mathbb{E}(x_{j'}|K_t, x_i = 1) + \sum_{l\geq t'+1}\sum_{j'\in I_l} \mathbb{E}(x_{j'}|K_t, x_i = 1, x_j = 1)\Bigg).$$

Again by Lemma A.1, write out each term $\mathbb{E}(x_{j'}|K_t, x_i = 1, x_j = 1)$ in the last summation as $\mathbb{E}(x_{j'}|K_t, x_i = 1, x_j = 1) = \frac{\overline{\theta}-\mathbb{E}(\theta_{j'}^*|K_t,x_i=1,x_j=1)}{\overline{\theta}-\underline{\theta}}$, and so on. Consider a player in a cohort $k > t+1$. Let M_k be the subset so players such that (i) each player $i \in M_k$ decided to buy, (ii) for $i, j \in M_k$, $i \in I_{n_i}, j \in I_{n_j}$ $n_i \neq n_j$ and $n_i > t$. Let for $l \leq k$ $M_k^l \subseteq M_k : \forall i \in M_k^l, i \in I_{n_i} \Rightarrow n_i < l$. Then

$$\mathbb{E}(\theta_j^*|K_t, x_i = 1, M_k) = u_0 - \frac{\mathbb{E}(\lambda)}{n-1}\Bigg(K_t + 1 + \#M_k + \sum_{j'\in I_t\setminus\{i\}} \mathbb{E}(x_{j'}|K_t)$$

$$+ \sum_{t+1\leq l\leq k}\sum_{j'\in I_l\setminus M_k} \mathbb{E}(x_{j'}|K_t, x_i = 1, M_k^l) + \sum_{l\geq k+1}\sum_{j'\in I_l} \mathbb{E}(x_{j'}|K_t, x_i = 1, M_k, x_j = 1)\Bigg).$$

Denoting the number of distinct equations for $\mathbb{E}(x_j|K_t, x_i = 1, M_k)$ by A, including terms with zero coefficient on the right-hand side of each equation, gives a system of A equations in A unknowns.

The associated matrix for this system has diagonal entries of 1 and off-diagonal entries of either 0 or $\frac{-1}{\overline{\theta}-\underline{\theta}}\frac{\mathbb{E}(\lambda)}{n-1} < 0$, where the number of non-zero off-diagonal entries in each row cannot exceed $\sum_{t'\geq t}\sum_{i\in I_{t'}} n_i - 1 \leq n-1$. By $\mathbb{E}(\lambda) \leq \overline{\lambda} = \overline{\theta} - v_o$ and $\underline{\theta} + 1 < v_o < \overline{\theta}$, the sum of the absolute values of off-diagonal entries in each row is therefore strictly less than one. Hence, this matrix is strictly diagonally dominant. By the Gershgorin theorem (1931), the system then has a unique solution.

In particular, this unique solution implies that $\mathbb{E}(x_j|K_t)$ for each consumer $j \neq i$ in cohort t, and $\mathbb{E}(x_j|K_t, x_i = 1)$ for each consumer j in cohort $t' \geq t+1$, are all uniquely defined. Hence, the cut-off function for consumer i,

$$\theta_i^*(\lambda) = u_0 - \frac{\lambda_i}{n-1}\left(K_t + \sum_{j\in I_t\setminus\{i\}} \mathbb{E}(x_j|K_t) + \sum_{t'\geq t+1}\sum_{j\in I_{t'}} \mathbb{E}(x_j|K_t, x_i = 1)\right),$$

given by (3) is uniquely defined as well.

□

Proposition 2. *Suppose that a sales scheme* I' *is more sequential than another scheme* I*, according to Definition 1. Then* I' *delivers strictly higher expected profits.*

Proof. We prove the following result, where repeated application given Definition 1 will immediately imply Proposition 2: *Consider sales schemes* I $= \{I_1, \ldots, I_{t-1}, I_t, I_{t+1}, \ldots I_m\}$ *and* I' $= \{I_1, \ldots, I_{t-1}, I_t', I_t'', I_{t+1}, \ldots I_m\}$*, where* $I_t = I_t' \cup I_t''$*. Then* I' *delivers strictly higher expected profits.*

Let p denote the optimal price under *I*. Suppose for now that the seller charges p under both schemes, so that both *I* and *I'* involve the same net outside option, $u_0 \equiv v_0 + p$.

Suppose that under *I'*, there are l consumers in cohort I_t'. Denote these consumers by subscript i, for $i = 1, \ldots, l$. Under *I*, these consumers are all members of cohort $I_t \supseteq I_t'$, and the probability that they will buy, given history K_t, is characterised by cut-off

$$\theta_i^* = u_0 - \frac{\mathbb{E}(\lambda)}{n-1}\left(K_t + \sum_{j \in I_t \setminus \{i\}} \mathbb{E}(x_j | K_t, \boldsymbol{\theta}) + \sum_{t' \geq t+1} \sum_{j \in I_{t'}} \mathbb{E}(x_j | K_t, x_i = 1, \boldsymbol{\theta})\right), \qquad (9)$$

where $\boldsymbol{\theta} = \{\theta_1^*, \ldots, \theta_l^*\}$ is the vector of cutoffs for these l consumers; $\mathbb{E}(\cdot | K_t, \boldsymbol{\theta})$ is the expectation conditional on history K_t and the fact that these l consumers have cut-offs $\boldsymbol{\theta}$. Due to Proposition 1, there exists a unique vector $\boldsymbol{\theta}$ resulting from consumer optimizing behavior, given K_t and I. In fact, (9) implies $\theta_1^* = \ldots = \theta_l^*$, but our notation allows for the fact that cutoffs will differ if λ is observable, in which case λ_i will replace $\mathbb{E}(\lambda)$ in (9).

Now, under I', I_t is split into two cohorts, I_t' and I_t''. For the l consumers in cohort I_t', the probability that they will buy, given history K_t, is characterised by cut-off

$$\begin{aligned} \theta_i^{*'} = u_0 - \frac{\mathbb{E}(\lambda)}{n-1}\Bigg(K_t &+ \sum_{j \in I_{t'} \setminus \{i\}} \mathbb{E}(x_j | K_t, \boldsymbol{\theta'}) + \sum_{j \in I_{t''}} \mathbb{E}(x_j | K_t, x_i = 1, \boldsymbol{\theta'}) + \\ &\sum_{t' \geq t+1} \sum_{j \in I_{t'}} \mathbb{E}(x_j | K_t, x_i = 1, \boldsymbol{\theta'})\Bigg), \end{aligned} \qquad (10)$$

where $\boldsymbol{\theta'} = \{\theta_1^{*'}, \ldots, \theta_l^{*'}\}$ is the vector of cutoffs for these l consumers; $\mathbb{E}(\cdot | K_t, \boldsymbol{\theta'})$ is the expectation conditional on history K_t and the fact that these l consumers have cut-offs $\boldsymbol{\theta'}$. Again due to Proposition 1, there exists a unique vector $\boldsymbol{\theta}$ resulting from consumer optimizing behavior, given K_t and I'.

We now use the Jacobi iterative method to show that $\boldsymbol{\theta'} < \boldsymbol{\theta}$; that is to say $\theta_i^*\boldsymbol{'} < \theta_i^*$ for $i = 1, \ldots, l$. This method consists of plugging an initial approximation for $\boldsymbol{\theta'}$ into the system of equations determining the cutoffs under I', solving for the cutoffs $\boldsymbol{\theta'}_1$ that are then implied by these equations, where $\boldsymbol{\theta'}_1$ may well differ from $\boldsymbol{\theta'}_0$, and repeating the process with $\boldsymbol{\theta'}_1$, $\boldsymbol{\theta'}_2, \ldots$. Recall from the proof of Proposition 1 that the system of equations determining the cutoffs is strictly diagonally dominant, which implies that given any initial approximation $\boldsymbol{\theta'}_0$, the iterations must converge to the unique fixed point $\boldsymbol{\theta'}$ of the system (see, e.g., Varga (1962)). Hence, to show $\boldsymbol{\theta'} < \boldsymbol{\theta}$, it is sufficient to find $\boldsymbol{\theta'}_0$ such that $\boldsymbol{\theta'}_n < \boldsymbol{\theta}$ holds for every iteration $n \geq 1$, and to show that the iterative process does not converge to exactly $\boldsymbol{\theta}$.

For each consumer k in cohort I''_t write $\mathbb{E}(x_k|K_t, x_i = 1, \boldsymbol{\theta'}) = \frac{\overline{\theta} - \mathbb{E}(\theta_k|K_t, x_i=1, \boldsymbol{\theta'})}{\overline{\theta} - \underline{\theta}}$, where

$$\mathbb{E}(\theta_k|K_t, x_i = 1, \boldsymbol{\theta'}) = u_0 - \frac{\mathbb{E}(\lambda)}{n-1}\left([K_t + 1] + \sum_{j \in I'_t \setminus \{i\}} \mathbb{E}(x_j|K_t, \boldsymbol{\theta'}) + \right.$$
$$\left. \sum_{j \in I''_t \setminus \{k\}} \mathbb{E}(x_j|K_t, \boldsymbol{\theta'}, x_i = 1) + \sum_{t' \geq t+1} \sum_{j \in I_{t'}} \mathbb{E}(x_j|K_t, x_i = 1, \boldsymbol{\theta'}, x_i = x_k = 1)\right). \tag{11}$$

Let $\boldsymbol{\theta'}_0 = \boldsymbol{\theta}$. By assumption, the behavior of consumers $i \in I'_t$ is then the same as under I. From (11), the decision problem of consumers $k \in I''_t$ is the same as under I, but with $\mathbb{E}(x_i|K_t, \boldsymbol{\theta}) < 1$ replaced by 1. Lemma A.2 and Lemma A.3 then imply $\mathbb{E}(x_j|K_t, x_i = 1, \boldsymbol{\theta'}_0) > \mathbb{E}(x_j|K_t, \boldsymbol{\theta})$ for all consumers in cohorts $I''_t, \ldots, I_m$. From (10), this in turn implies $\boldsymbol{\theta'}_1 \equiv R(\boldsymbol{\theta'}_0) < \boldsymbol{\theta'}_0$, hence $\boldsymbol{\theta'}_1 < \boldsymbol{\theta}$. It follows that the iterative process cannot converge to exactly $\boldsymbol{\theta}$, since that would require $\boldsymbol{\theta'}_1 = \boldsymbol{\theta}$.

Now assume $\boldsymbol{\theta'}_n < \boldsymbol{\theta}$ for some iteration $n \geq 1$, so that

$$\sum_{j \in I'_t \setminus \{i\}} \mathbb{E}(x_j|K_t, \boldsymbol{\theta'}_n) > \sum_{j \in I'_t \setminus \{i\}} \mathbb{E}(x_j|K_t, \boldsymbol{\theta}).$$

From (11), the decision problem of consumers $k \in I''_t$ is the same as under I, but with $\mathbb{E}(x_i|K_t, \boldsymbol{\theta}) < 1$ replaced by 1, and with $\sum_{j \in I'_t \setminus \{i\}} \mathbb{E}(x_j|K_t, \boldsymbol{\theta})$ replaced by $\sum_{j \in I'_t \setminus \{i\}} \mathbb{E}(x_j|K_t, \boldsymbol{\theta'}_n)$. Lemma A.2 and Lemma A.3 then imply $\mathbb{E}(x_j|K_t, x_i = 1, \boldsymbol{\theta'}_n) > \mathbb{E}(x_j|K_t, \boldsymbol{\theta})$ for all consumers in cohorts $I''_t, \ldots, I_m$. From (10), this in turn implies $\boldsymbol{\theta'}_{n+1} \equiv R(\boldsymbol{\theta'}_n) < \boldsymbol{\theta}$. It follows by induction that $\boldsymbol{\theta'}_n < \boldsymbol{\theta}$ holds for every iteration $n \geq 1$, as required.

The results so far show that conditional on any given history K_t, moving to I' will cause $\sum_{j \in I'_t} \mathbb{E}(x_j|K_t)$ to increase. For each consumer k in cohort I''_t write $\mathbb{E}(x_k|K_t, \boldsymbol{\theta'}) = \frac{\overline{\theta} - \mathbb{E}(\theta_k|K_t, \boldsymbol{\theta'})}{\overline{\theta} - \underline{\theta}}$, where $\mathbb{E}(\theta_k|K_t, \boldsymbol{\theta'}) =$

$$u_0 - \frac{\mathbb{E}(\lambda)}{n-1}\left([K_t + \sum_{j \in I'_t} \mathbb{E}(x_j|K_t, \boldsymbol{\theta'})] + \sum_{j \in I''_t \setminus \{k\}} \mathbb{E}(x_j|K_t, \boldsymbol{\theta'}) + \sum_{t' \geq t+1} \sum_{j \in I_{t'}} \mathbb{E}(x_j|K_t, \boldsymbol{\theta'}, x_k = 1)\right). \tag{12}$$

Looking at the term in square brackets, moving to I' is equivalent to replacing $\sum_{j \in I'_t} \mathbb{E}(x_j|K_t, \boldsymbol{\theta})$ by the strictly larger $\sum_{j \in I'_t} \mathbb{E}(x_j|K_t, \boldsymbol{\theta'})$. This is in turn equivalent to replacing history

$K_t + \sum_{j\in I'_t} \mathbb{E}(x_j|K_t,\boldsymbol{\theta})$ by the strictly larger $K_t + \sum_{j\in I'_t} \mathbb{E}(x_j|K_t,\boldsymbol{\theta'})$, so it follows from Lemma A.2 and Lemma A.3 that $\sum_{j\in I''_t} \mathbb{E}(x_j|K_t)$ will increase.

Hence, for any history K_t, moving to I' increases expected total sales from consumers previously in cohort t, from $\sum_{j\in I_t} \mathbb{E}(x_j|K_t)$ to $\sum_{j\in I'_t} \mathbb{E}(x_j|K_t) + \sum_{j\in I''_t} \mathbb{E}(x_j|K_t)$. Thus Lemma A.4 with $t' = 1$ implies that $\mathbb{E}(x_j)$ strictly increases for all consumers. Hence, ex ante expected profits, $p\sum_{j=1}^{n} \mathbb{E}(x_j)$ are strictly higher under I' than under I, given the assumption that the seller charges price p under both schemes. Let p' denote the optimal price under I'. By the optimality of this price, ex ante expected profits under I' at price p' must be strictly higher than expected profits under I at price p.

□

Proposition 3. *Suppose* $\mathbb{E}(\lambda) < v_0 - \underline{\theta}$*. Consider sales schemes* $I = \{I_1, \ldots, I_t, I_{t+1}, \ldots I_m\}$ *and* $I' = \{I_1, \ldots, I'_t = I_{t+1}, I'_{t+1} = I_t, \ldots I_m\}$ *with* $|I_t| \equiv n_t > n_{t+1} \equiv |I_{t+1}|$*. Then* I' *yields strictly higher expected profits than* I*.*

Proof. The proof is similar to that of Proposition 2. We first fix some history K_t and the expected actions of consumers in cohorts I_{t+2} and after. We then show that swapping I_t and I_{t+1} will strictly increase expected sales from these two cohorts, conditional on this history. Finally, direct application of Lemmas A.4 with $t' = 1$ implies that from an ex-ante perspective, expected sales from all cohorts will strictly increase. Let p denote the optimal price under I. Suppose for now that the seller charges p under both schemes, so that both I and I' involve the same net outside option, $u_0 \equiv v_0 + p$.

Denote $L = K_t + 1 + \sum_{l\geq t+2} \sum_{j\in I_l} \mathbb{E}(x_j|K_t)$. From the perspective of consumer i in cohort t, following history K_t, L is the expected number of total sales, ignoring the behavior of consumers in cohorts I_t and I_{t+t}. Its value will depend on the expected behavior of others in cohorts I_t and I_{t+1}, but for now this dependence is left implicit. Then, holding L constant, Lemma A.1 implies that the expected cut-off θ_{t+1} for consumers in cohort I_{t+1}, from the perspective of consumer i in I_t who buys, is determined by:

$$\theta_{t+1} + \frac{\mathbb{E}(\lambda)}{n-1}\Bigg(L + (n_t - 1)\mathbb{E}(x_j|K_t) + (n_{t+1} - 1)\mathbb{E}(x_j|K_t, x_i = 1)\Bigg) = u_0,$$

which can be rewritten as

$$\theta_{t+1} + \frac{\mathbb{E}(\lambda)}{n-1}\left(L + (n_t - 1)\frac{\overline{\theta} - \theta_t}{\overline{\theta} - \underline{\theta}} + (n_{t+1} - 1)\frac{\overline{\theta} - \theta_{t+1}}{\overline{\theta} - \underline{\theta}}\right) = u_0. \tag{13}$$

Expression (13) shows that θ_{t+1} depends on θ_t, which is the expected cut-off for a consumer in cohort I_t, conditional on K_t. This cutoff θ_t is determined in turn by

$$\theta_t + \frac{\mathbb{E}(\lambda)}{n-1}\left(L - 1 + (n_t - 1)\frac{\overline{\theta} - \theta_t}{\overline{\theta} - \underline{\theta}} + n_{t+1}\frac{\overline{\theta} - \theta_{t+1}}{\overline{\theta} - \underline{\theta}}\right) = u_0. \tag{14}$$

From the perspective of the seller, expected sales from cohorts I_t and I_t, conditional on history K_t, are then

$$S(n_t, n_{t+1}) = n_t\frac{\overline{\theta} - \theta_t}{\overline{\theta} - \underline{\theta}} + n_{t+1}\frac{\overline{\theta} - \theta'_{t+1}}{\overline{\theta} - \underline{\theta}},$$

where Lemma A.1 implies that θ'_{t+1} is defined by

$$\theta'_{t+1} + \frac{\mathbb{E}(\lambda)}{n-1}\left(L - 1 + n_t\frac{\overline{\theta} - \theta_t}{\overline{\theta} - \underline{\theta}} + (n_{t+1} - 1)\frac{\overline{\theta} - \theta'_{t+1}}{\overline{\theta} - \underline{\theta}}\right) = u_0. \tag{15}$$

That is, θ'_{t+1} is the expected cutoff for consumers in cohort I_{t+1}, from the perspective of the seller serving cohort I_t, conditional on history K_t. Solving the system (13)–(15) allows us to determine the cut-offs and compute $S(n_t, n_{t+1})$. Let $\Delta = S(n_t, n_{t+1}) - S(n_{t+1}, n_t)$. Then,

$$\Delta = \frac{(n_{t+1} - n_t)(\overline{\theta} - \underline{\theta})[(n-1)(\overline{\theta} - \underline{\theta}) + \mathbb{E}(\lambda)][(n-1)(u_0 - \underline{\theta}) - (L + n_t + n_{t+1} - 2)\mathbb{E}(\lambda)]}{\frac{1}{n_t n_{t+1}(n-1)(\mathbb{E}(\lambda))^3}G_1 \cdot G_2 \cdot G_3 \cdot G_4}.$$

where

$$G_1 = (n-1)(\overline{\theta} - \underline{\theta}) - (n_{t+1} - 1)\mathbb{E}(\lambda),$$

$$G_2 = (n-1)(\overline{\theta} - \underline{\theta}) - (n_t - 1)\mathbb{E}(\lambda),$$

$$G_3 = (n-1)^2(\overline{\theta} - \underline{\theta})^2 - (n-1)(n_t + n_{t+1} - 2)(\overline{\theta} - \underline{\theta})\mathbb{E}(\lambda) - (n_{t+1} - 1)(\mathbb{E}(\lambda))^2,$$

$$G_4 = (n-1)^2(\overline{\theta} - \underline{\theta})^2 - (n-1)(n_t + n_{t+1} - 2)(\overline{\theta} - \underline{\theta})\mathbb{E}(\lambda) - (n_t - 1)(\mathbb{E}(\lambda))^2.$$

Due to $\mathbb{E}(\lambda) \leq \overline{\theta} - v_0 < \overline{\theta} - \underline{\theta}$, G_3 can be rewritten as

$$G_3 = (n-1)(\overline{\theta} - \underline{\theta})\left[(n-1)(\overline{\theta} - \underline{\theta}) - (n_t + n_{t+1} - 2)\mathbb{E}(\lambda)\right] - (n_{t+1} - 1)(\mathbb{E}(\lambda))^2 \geq$$
$$(\overline{\theta} - \underline{\theta})^2[(n-1)(n-1-n_t-n_{t+1}+2) - n_{t+1} + 1] \geq (\overline{\theta} - \underline{\theta})^2(n-1-n_t+1) > 0.$$

In a similar fashion G_1, G_2 and G_4 are all positive. Note that as $u_0 - \underline{\theta} > v_0 - \underline{\theta} > \mathbb{E}(\lambda)$ holds by assumption, and $L - 1 + n_t + n_{t+1} \leq n$, the last term in the numerator of Δ is always positive, which implies that $\Delta > 0$ whenever $n_{t+1} > n_t$. Thus, compared with I, sales scheme I' yields strictly higher expected sales from cohorts I_t and I_{t+1}, conditional on K_t. Thus Lemma A.4 with $t' = 1$ implies that the ex ante probability of buying, $\mathbb{E}(x_j)$, strictly increases for all consumers. Hence, ex ante expected profits, $p\sum_{j=1}^{n}\mathbb{E}(x_j)$, are strictly higher under *I'* than under *I*, if the seller charges price p under both schemes. Let p' denote the optimal price under *I'*. By the optimality of this price, ex ante expected profits under *I'* at price p' must be strictly higher than expected profits under *I* at price p.

□

Proposition 4. *Suppose the weight consumers place on the network payoff, λ, is observable. Then the sales scheme* I *that maximizes expected profits has a single consumer per cohort, increasingly ordered in λ, i.e. $\lambda_1 \leq \ldots \leq \lambda_n$.*

Proof. Notice first that all previous results continue to hold when λ is observable. Proofs remain unchanged except that the relevant cutoff for any consumer i is now $\theta_i^*(\lambda_i)$ given by (3), rather than $\theta_i^* \equiv \mathbb{E}_\lambda(\theta_i^*(\lambda_i))$ given by (5). That is, the only difference is that λ_i replaces $\mathbb{E}(\lambda)$ in the expression for this cutoff. Hence, by Corollary 1, the sales scheme that maximizes expected profits still has a single consumer per cohort.

For the optimal ordering of these consumers, we prove the result directly. Consider a fully sequential partition with one consumer per cohort, and fix p at the optimal price for this partition. Consider two subsequent consumers: i and $i+1$. Suppose there where K

consumers who bought before consumer i. Then

$$\mathbb{E}(x_{i+1}|K, x_i = 1) = \frac{\overline{\theta} - u_0 + \frac{\lambda_{i+1}}{n-1}\left(K + 1 + \sum_{j=i+2}^{n} \mathbb{E}(x_j|K, x_i = x_{i+1} = 1)\right)}{\overline{\theta} - \underline{\theta}},$$

$$\mathbb{E}(x_{i+1}|K, x_i = 0) = \frac{\overline{\theta} - u_0 + \frac{\lambda_{i+1}}{n-1}\left(K + \sum_{j=i+2}^{n} \mathbb{E}(x_j|K, x_i = 0, x_{i+1} = 1)\right)}{\overline{\theta} - \underline{\theta}}.$$

Now we look at consumer i, where

$$\begin{aligned}\mathbb{E}(x_i|K) =& \frac{1}{\overline{\theta} - \underline{\theta}}\left(\overline{\theta} - u_0 + \frac{\lambda_i}{n-1}\left(K + \right.\right. \\ & \mathbb{P}(x_{i+1} = 1|K, x_i = 1)\left(1 + \sum_{j=i+2}^{n} \mathbb{E}(x_j|K, x_i = x_{i+1} = 1)\right) + \\ & \left.\left.\mathbb{P}(x_{i+1} = 0|K, x_i = 1)\sum_{j=i+2}^{n} \mathbb{E}(x_j|K, x_i = 1, x_{i+1} = 0)\right)\right).\end{aligned}$$

Clearly in our setting $\mathbb{P}(x_{i+1} = 1|K) = \mathbb{E}(x_{i+1}|K)$, for any history K. Now define:

$$\begin{aligned}S(\lambda_i, \lambda_{i+1}) \equiv & \sum_{j=i}^{n} \mathbb{E}(x_j|K) = \\ & \mathbb{P}(x_i = 1|K)\left(\mathbb{P}(x_{i+1} = 1|K, x_i = 1)\left(2 + \sum_{j=i+2}^{n} \mathbb{E}(x_j|K, x_i = x_{i+1} = 1)\right) + \right. \\ & \left.\mathbb{P}(x_{i+1} = 0|K, x_i = 1)\left(1 + \sum_{j=i+2}^{n} \mathbb{E}(x_j|K, x_i = 1, x_{i+1} = 0)\right)\right) + \\ & \mathbb{P}(x_i = 0|K)\left(\mathbb{P}(x_{i+1} = 1|K, x_i = 0)\left(1 + \sum_{j=i+2}^{n} \mathbb{E}(x_j|K, x_i = 0, x_{i+1} = 1)\right) + \right. \\ & \left.\mathbb{P}(x_{i+1} = 0|K, x_i = 0)\sum_{j=i+2}^{n} \mathbb{E}(x_j|K, x_i = 0, x_{i+1} = 0)\right).\end{aligned}$$

We now use the fact that all expectations are linear in prior sales: $\mathbb{E}(x_j|K, x_i = 0, x_{i+1} = 1) = \mathbb{E}(x_j|K, x_i = 1, x_{i+1} = 0)$ and $2\mathbb{E}(x_j|K, x_i = 1, x_{i+1} = 0) = \mathbb{E}(x_j|K, x_i = 0, x_{i+1} = 0) + \mathbb{E}(x_j|K, x_i = x_{i+1} = 1)$. Thus

$$S(\lambda_i, \lambda_{i+1}) - S(\lambda_{i+1}, \lambda_i) = -\frac{(2 + Q_2 - Q_0)^3 (Q_2 + K + 1)(\lambda_i - \lambda_{i+1})\lambda_i\lambda_{i+1}}{8(n-1)^3(\overline{\theta} - \underline{\theta})^3}, \tag{16}$$

where

$$Q_2 \equiv \sum_{j=i+2}^{n} \mathbb{E}(x_j | K, x_i = x_{i+1} = 1),$$

$$Q_0 \equiv \sum_{j=i+2}^{n} \mathbb{E}(x_j | K, x_i = x_{i+1} = 0).$$

Lemma A.2 implies that $Q_2 > Q_0$. Hence by (16), we have $S(\lambda_i, \lambda_{i+1}) - S(\lambda_{i+1}, \lambda_i) > 0$ if and only if $\lambda_i < \lambda_{i+1}$. If $\lambda_i > \lambda_{i+1}$, then allowing consumer $i+1$ to act before consumer i will strictly increase $\sum_{j=i}^{n} \mathbb{E}(x_j|K)$, for any history K. Hence, applying Lemma A.4 with $t' = 1$, allowing consumer $i+1$ to act before consumer i will also strictly increase $\sum_{1 \leq i \leq n} \mathbb{E}(x_i)$. It follows that ex ante expected profits, $p \sum_{1 \leq i \leq n} \mathbb{E}(x_i)$, are strictly higher under this new ordering, where p was the optimal price under the original ordering. Let p' denote the optimal price under the new ordering. Thus, by optimality of this price, ex ante expected profits under the new ordering at price p' must be strictly higher than expected profits under the original ordering at price p.

□

Proposition 5. *Consider a simultaneous sales scheme, with all consumers in the same cohort. Suppose that before buying, consumers can simultaneously send a message* $m \in [\underline{\theta}, \overline{\theta}] \times (0, \overline{\lambda})$ *about their type which all other consumers observe, and where the seller observes* $M = (m_1, \ldots, m_n)$ *with probability* $q > 0$*. Furthermore suppose that the seller sets price* p^* *if it does not observe* M*, and sets price* $p(M)$ *if it does, where* $p(M)$ *is non-decreasing in* $\sum_{i=1}^{n} \mathbb{E}(x_i | p^*, M)$*. Then when* q *is sufficiently small, an equilibrium exists where communication is informative, in the sense that each consumer truthfully reveals to all others the minimum level of total sales required for him to buy himself at price* p^**. In the limit as* q *tends to zero, consumer purchase decisions approach those in a setting where consumers all observe each others' type,* (θ_i, λ_i) *for all* $i = 1, \ldots, n$*.*

Proof. For consumer i, define N_i as the smallest value of N for which

$$\theta_i + \frac{\lambda_i}{n-1}N - p^* \geq u_0.$$

N_i is the minimum number of other consumers who must buy for consumer i to want to buy himself, given price p^*. For each $l = 0, 1, \ldots, n-1$, let B_l denote the set of all $(\theta, \lambda) \in [\underline{\theta}, \overline{\theta}] \times (0, \overline{\lambda})$ for which $N = l$. If $n-1$ consumers buying is insufficient to motivate consumer i to buy, then we write $N_i = n$.

Any consumer i with $\theta_i = \underline{\theta}$ has a strictly dominant strategy not to buy ($N_i = n$), regardless of the price. Let n' denote the value of N_i for a consumer with $\theta_i = \overline{\theta}$ and $\lambda_i = \overline{\lambda}$, where $n' \leq n-1$ in any situation of interest. Notice that $n' = 0$ if $p^* \in (\underline{\theta} - u_0, \overline{\theta} - u_0)$, since then $\theta_i = \overline{\theta}$ implies a strictly dominant strategy to buy. Willingness to pay is increasing in θ which has full support on $[\underline{\theta}, \overline{\theta}]$. Hence, from an ex ante perspective, for each consumer i, there is a strictly positive probability that $(\theta_i, \lambda_i) \in B_l$, for each $l = n', n'+1, \ldots n$.

Consider a candidate equilibrium where each consumer i plays a mixed strategy placing strictly positive probability on all messages $m \in B_{N_i}$ and zero probability on all $m \notin B_{N_i}$. Conditional on receiving any $m \in B_N$ from consumer i, all other consumers then infer that $N_i = N$. Notice that every $m \in [\underline{\theta}, \overline{\theta}] \times (0, \overline{\lambda})$ is on the equilibrium path, and corresponds to some $N \in \{n', n'+1, \ldots, n\}$.

Define X_l as the number of messages $m \in B_l$ in this candidate equilibrium, for each $l = n', n'+1, \ldots, n$. Define N_{max} as be the maximum value of $j+1$ such that $\sum_{l=0}^{j} X_l \geq j+1$; if no such $j+1$ exists, then define $N_{max} \equiv 0$. Then given price p^*, consumer i's strategy in this candidate equilibrium has him buy if and only if $(N_{max} - \mathbb{1}_{m_i \leq N_{max}}) \geq N_i$. Thus, N_{max} gives total sales at price p^*, conditional on messages $M = \{m_1, \ldots, m_n\}$. Since all $m \in [\underline{\theta}, \overline{\theta}] \times (0, \overline{\lambda})$ are on the equilibrium path, there is a strictly positive probability that N_{max} takes on each value $0, n'+1, n'+2, \ldots, n$.

Given the messages of other consumers, any message $m_i \in B_N$ leads to the same updated beliefs about consumer i's type, $(\theta_i, \lambda_i) \in B_N$, the same value of N_{max} and the same purchase behavior at price p^*. Thus, any such message must also lead to the same price p and the same purchase behavior if the seller observes the messages. This is the case for $N = n', n'+1, \ldots, n$. It follows that for each N, consumers are indifferent between all messages $m \in B_N$, so without loss of generality we can write $m \in \{n', n'+1, \ldots, n)$. That is, each consumer i's message is effectively an integer N, where the candidate equilibrium prescribes $m_i = N_i$. The incentive to buy at a given price depends only on the number of other consumers expected to also buy. Hence, if the seller sets price p^*, and each consumer i sends message $m_i \in N_i$, then consumers will make the same purchase decisions as if they all observed each others' type.

To establish our result, we need to show that for q sufficiently close to zero, no consumer has a profitable deviation. First consider the case where the seller does not observe the messages so consumers face price p^*. By $m_i = N_i$ for all $i = 1, \ldots, n$ and the definition of N_{max}, each consumer who buys receives a payoff of at least u_0. Each consumer who does not buy would receive a payoff strictly less than u_0 if he did buy. Hence, given price p^*, a deviation from consumer i can only be profitable if it involves a change of message, to some $m_i' = N_k \neq N_i$. Let X_l' be the number of messages $m = l$ following this deviation, for each $l = n', n'+1, \ldots, n$. We have $X'_{N_i} = X_{N_i} - 1$, $X'_{N_k} = X_{N_k} + 1$, and $X_l' = X_l$ for all $l \neq N_i, N_k$. Define N'_{max} as the maximum value of $j+1$ such that $\sum_{l=0}^{j} X_l' \geq j+1$; if no such $j+1$ exists, then define $N'_{max} \equiv 0$.

Suppose $N_k > N_i$, with $N_i \leq n-1$, so consumer i understates his willingness to pay. Then $\sum_{l=0}^{j} X_l' = \sum_{l=0}^{j} X_l$ for all $j = 0, \ldots, N_{i-1}$ and for all $j = N_k, \ldots, n$, whereas $\sum_{l=0}^{j} X_l' = \sum_{l=0}^{j} X_l - 1$ for all $j = N_i, \ldots, N_{k-1}$. This implies $N'_{max} - \mathbb{1}_{m_i' \leq N'_{max}} \leq N_{max} - \mathbb{1}_{m_i \leq N_{max}}$. Moreover, since all messages are on the equilibrium path, there is a strictly positive probability that $N'_{max} - \mathbb{1}_{m_i' \leq N'_{max}} < N_{max} - \mathbb{1}_{m_i \leq N_{max}}$, for any realized value of $N_{max} \in \{0, n'+1, n'+2, \ldots, n\}$. The payoff of reporting N_i is given by

$$(1-q)\left(\mathbb{P}(N_{max} - \mathbb{1}_{m_i \leq N_{max}} < N_i)u_0 + \mathbb{P}(N_{max} - \mathbb{1}_{m_i \leq N_{max}} \geq N_i)\left(\theta_i + \lambda_i \frac{\mathbb{E}(N_{max}) - 1}{n-1}\right)\right) + qU_0,$$

where it understood that the term $\mathbb{E}(N_{max})$ is conditional on $\mathbb{P}(N_{max} - \mathbb{1}_{m_i \leq N_{max}} \geq N_i)$, and where U_0 is the expected payoff if the seller observes the messages. Meanwhile the payoff from deviating to N_k is

$$(1-q)\left(\mathbb{P}(N'_{max} - \mathbb{1}_{m'_i \leq N'_{max}} < N_i)u_0 + \mathbb{P}(N'_{max} - \mathbb{1}_{m'_i \leq N'_{max}} \geq N_i)\left(\theta_i + \lambda_i \frac{\mathbb{E}(N'_{max}) - 1}{n-1}\right)\right) + qU_1,$$

where it understood that the term $\mathbb{E}(N'_{max})$ is conditional on $\mathbb{P}(N'_{max} - \mathbb{1}_{m'_i \leq N'_{max}} \geq N_i)$, and where U_1 is the expected payoff obtained by consumer i if the seller observes these messages which include m'_i. Taking the difference between the two payoffs and using the fact that $\mathbb{E}(N_{max}) > \mathbb{E}(N'_{max})$, the deviation is not profitable if:

$$\Delta\mathbb{P}\left(\theta_i + \lambda_i \frac{\mathbb{E}(N'_{max}) - 1}{n-1} - u_0\right) > \frac{q}{1-q}(U_1 - U_0),$$

where $\Delta\mathbb{P} \equiv \mathbb{P}(N_{max} - \mathbb{1}_{m_i \leq N_{max}} \geq N_i) - \mathbb{P}(N'_{max} - \mathbb{1}_{m'_i \leq N'_{max}} \geq N_i) > 0$. This inequality holds for sufficiently small q, thus underreporting willingness to pay is not profitable.

Now suppose $N_k < N_i$, with $N_i \geq 1$, so consumer i overstates his willingness to pay. Again consider the case where the seller does not observe messages, so consumers face price p^*. Then $\sum_{l=0}^{j} X'_l = \sum_{l=0}^{j} X_l$ for all $j \geq N_i$. Hence, $N'_{max} \geq N_{max}$ holds, but $N'_{max} > N_{max}$ can only hold if $N'_{max} < N_i$. The condition $N'_{max} > N_{max}$ is necessary for the deviation to increase consumer i's payoff, since the number of consumers other than i who buy must increase. But $N'_{max} < N_i$ implies that consumer i will not buy himself following the deviation, so the deviation will not increase his payoff.

Continue to suppose $N_k < N_i$ but consider the case where the seller does observe messages. Then $N'_{max} \geq N_{max}$ implies that the deviation leads to a weakly higher price: $p(M') \geq p(M)$, where M denotes the equilibrium messages, and M' denotes messages given the deviation. From (2), consumer best-response functions when simultaneously making purchase decisions are upward-sloping (strategic complements), where a price increase reduces the net payoff from buying. Thus, $p(M') \geq p(M)$ implies $\mathbb{E}(x_j|M') \leq \mathbb{E}(x_j|M)$ for each consumer j (see Vives (1990)), so the deviation will not increase consumer i's payoff. Hence, overreporting willingness to pay is not profitable.

Finally, note that as q approaches zero, the probability that the seller charges p^* approaches 1, which together with informative communication guarantees that consumers almost surely make the same purchase decisions as if they observed each others' types.

□

Proposition 5. *Suppose the seller can commit to a dynamic pricing schedule, with price $p(K_t)$ for cohort t conditional on previous sales K_t. Then fully sequential sales (a single consumer per cohort), delivers higher expected profits than fully simultaneous sales (all consumers in a single cohort).*

Proof. First consider a fully simultaneous scheme, with all consumers in a single cohort: $I = \{I_1\}$, with $n_1 = n$. Then dynamic pricing is equivalent to static pricing; both simply specify a single value of p. Let $\pi(p^*|I)$ denote expected profits given the optimal static price p^* under this partition.

Now consider a fully sequential scheme, with a single consumer per cohort: $I' = \{I'_1, \ldots, I'_n\}$, with $n_t = 1$ for all $t = 1, \ldots, n$. For each t, the seller's strategy specifies a price $p(K_t)$, for every possible value of previous sales $K_t = 0, \ldots, t-1$. With slight abuse of notation, let $\pi(p(K_t)|I')$ denote expected profits given this pricing schedule under this partition.

Suppose that for each $t = 1, \ldots, n$, the seller sets $p(K_t) = p^*$ for all $K_t = 0, \ldots, t-1$. Expected profits are then $\pi(p^*|I')$. Let $\pi(p(K_t)^*|I')$ denote expected profits given the optimal dynamic pricing schedule $p(K_t)^*$. Then optimality implies $\pi(p(K_t)^*|I') \geq \pi(p^*|I')$. Corollary 1 shows that $\pi(p^*|I') > \pi(p^*|I)$, which in turn implies $\pi(p(K_t)^*|I') > \pi(p^*|I)$.

□

References

Masaki Aoyagi. Optimal sales schemes against interdependent buyers. *American Economic Journal: Microeconomics*, 2(1):150–82, February 2010.

Mark Armstrong and Jidong Zhou. Search deterrence. *Review of Economic Studies*, forthcoming, 2015.

Heski Bar-Isaac. Reputation and survival: Learning in a dynamic signalling model. *Review of Economic Studies*, 70(2):231–251, April 2003.

Amiya Basu, Tridib Mazumdar, and S. P. Raj. Indirect network externality effects on product attributes. *Marketing Science*, 22(2):pp. 209–221, 2003.

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

Subir Bose, Gerhard Orosel, Marco Ottaviani, and Lise Vesterlund. Dynamic monopoly pricing and herding. *The RAND Journal of Economics*, 37(4):910–928, 2006.

Subir Bose, Gerhard Orosel, Marco Ottaviani, and Lise Vesterlund. Monopoly pricing in the binary herding model. *Economic Theory*, 37(2):203–241, November 2008.

Antonio Cabral. Dynamic price competiton with network effects. *Review of Economic Studies*, 78:pp. 83–111, 2011.

Luis M. B. Cabral, David J. Salant, and Glenn A. Woroch. Monopoly pricing with network externalities. *International Journal of Industrial Organization*, 17(2):199–214, February 1999.

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

Yubo Chen and Jinhong Xie. Online consumer review: Word-of-mouth as a new element of marketing communication mix. *Management Science*, 54(3):477–491, 2008.

Chrysanthos Dellarocas. Strategic manipulation of internet opinion forums: Implications for consumers and firms. *Management Science*, 52(10):1577–1593, 2006.

Anirudh Dhebar and Shmuel S. Oren. Optimal dynamic pricing for expanding networks. *Marketing Science*, 4(4):pp. 336–351, 1985.

Engelbert Dockner and Steffen Jorgensen. Optimal advertising policies for diffusion models of new product innovation in monopolistic situations. *Management Science*, 34(1):pp. 119–130, 1988.

Yifan Dou, Marius Florin Niculescu, and D. J. Wu. Optimal consumer network structure formation under network effects: Seeds controllability and visibility. Technical report, NET Institute Working Paper No. 11-07, October 2011.

Yifan Dou, Marius F. Niculescu, and D. J. Wu. Engineering optimal network effects via social media features and seeding in markets for digital goods and services. *Information Systems Research*, 24(1):164–185, 2013.

Philip H. Dybvig and Chester.S Splatt. Adoption externalities and public goods. *Journal of Public Economics*, 20:pp. 231–247, 1983.

Joseph Farrell and Garth Saloner. Standardization, compatibility, and innovation. *RAND Journal of Economics*, 16(1):70–83, Spring 1985.

Joseph Farrell and Garth Saloner. Installed base and compatibility: Innovation, product preannouncements, and predation. *The American Economic Review*, 76(5):940–955, 1986.

Semyon Gershgorin. Uber die abgrenzung der eigenwerte einer matrix. *Izv. Akad. Nauk SSSR Ser. Mat. 1 7*, 1(1):749–755, 1931.

David Gill and Daniel Sgroi. Sequential decisions with tests. *Games and Economic Behavior*, 63(2):663–678, July 2008.

David Gill and Daniel Sgroi. The optimal choice of pre-launch reviewer. *Journal of Economic Theory*, 147(3):1247–1260, 2012.

David Godes, Dina Mayzlin, Yubo Chen, Sanjiv Das, Chrysanthos Dellarocas, Bruce Pfeiffer, Barak Libai, Subrata Sen, Mengze Shi, and Peeter Verlegh. The firm's management of social interactions. *Marketing Letters*, 16(3-4):415–428, 2005.

Hanna Halaburda, Bruno Jullien, and Yaron Yehezkel. Dynamic competition with network externalities: Why history matters. 2015.

Shlomo Kalish. A new product adoption model with price, advertising, and uncertainty. *Management Science*, 31(12):pp. 1569–1585, 1985.

Michael L Katz and Carl Shapiro. Network externalities, competition, and compatibility. *American Economic Review*, 75(3):424–440, June 1985.

Ting Liu and Pasquale Schiraldi. New product launch: herd seeking or herd preventing? *Economic Theory*, 51(3):627–648, November 2012.

Dina Mayzlin, Yaniv Dover, and Judy Chevalier. Promotional reviews: An empirical investigation of online review manipulation. *American Economic Review*, 104(8):2421–55, 2014.

Jack Ochs and In-Uck Park. Overcoming the coordination problem: Dynamic formation of networks. *Journal of Economic Theory*, 145(2):689–720, March 2010.

Marco Ottaviani and Andrea Prat. The value of public information in monopoly. *Econometrica*, 69(6):1673–1683, November 2001.

V. Padmanabhan, Surendra Rajiv, and Kannan Srinivasan. New products, upgrades, and new releases: A rationale for sequential product introduction. *Journal of Marketing Research*, 34(4):pp. 456–472, 1997.

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

Baohong Sun, Jinhong Xie, and H. Henry Cao. Product strategy for innovators in markets with network effects. *Marketing Science*, 23(2):pp. 243–254, 2004.

Richard S. Varga. *Matrix Iterative Analysis,*. Prentice-Hall, Inc., Englewood Cliffs, N.J., 1962.

Andre Veiga. Dynamic platform design. 2015.

Xavier Vives. Nash equilibrium with strategic complementarities. *Journal of Mathematical Economics*, 19(3):305–321, 1990.

Jinhong Xie and Marvin Sirbu. Price competition and compatibility in the presence of positive demand externalities. *Management Science*, 41(5):pp. 909–926, 1995.