Optimal Sales Schemes for Network Goods*

Alexei Parakhonyak † and Nick Vikander‡

September 29, 2017

Abstract

This paper considers a monopolist's product-launch strategy in the presence of network effects, focusing on how to exploit these effects to the maximum possible extent. In our formal framework, the firm sets a price for its product and chooses a sales scheme, which effectively determines how consumers can learn from each other about the product's popularity. Our results on the profitability of different schemes provide insights on a variety of managerial issues that are of practical relevance for product launch. Specifically, to best exploit network effects, our results suggest releasing pre-order and sales information to consumers. They also suggest taking a sequential approach to multi-market product launch, launching first in smaller markets before larger ones. Moreover, they identify possible benefits of promoting pre-release consumer communication, and support the idea of targeting independent-minded consumers, so they can serve 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 at EARIE, EEA, Royal Economic Society, IIOC, Dusseldorf Institute for Competition Economics, the University of Edinburgh, the University of Copenhagen, the Toulouse School of Economics, University of Oxford, Tinbergen Institute Amsterdam, and Erasmus University Rotterdam, and from discussions with Maarten Janssen, Thomas Tregouet, Andre Veiga and Adrien Vigier.

†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 each consumer's benefit from buying is increasing in the total number of consumers who also buy. These effects can be indirect if high sales boost the availability of complementary products. Take, for example, DVD players and DVD movie titles, iPhone handsets and iPhone Apps, video game consoles such as PlayStation and XBox and their associated game titles, HDTV television sets and high-definition channels, or electric cars and charging stations.[1] Network effects can also be direct if buyers benefit from interacting with one another. Direct effects can arise from issues of technological compatibility, for example with different operating systems such as Linux, Google OS, MacOS, and Windows, or with business solutions such as Microsoft Office 365 and Google Apps for Work. They can also stem from social interactions, for example with massively multiplayer online video games such as World of Warcraft, or for books and movies, where discussing with others is part of the consumption experience.[2]

Network effects are relevant for managerial decision-making because of their significant impact on consumer behavior. In the presence of network effects, social influence can drive consumption decisions, since each consumer's incentive to buy the product depends on whether others do the same. Past product sales will matter if consumers can observe the size of the installed base, as will expectations about sales in the future. These expectations can be self-reinforcing and generate feedback loops, where high expected sales make buying more attractive, which in turn feeds back into increasing sales (Dickson et al. (2001)). For this reason, a crucial issue facing managers and marketers is how best to exploit network effects to promote product adoption, in particular by influencing consumer expectations (Lee and Colarelli O'Connor (2003), Hauser et al. (2006)).[3]

[1]Li et al. (2015) estimate network effects in the electric vehicle market and find them to be substantial, on both sides of the market. Drivers are more willing to buy electric cars if there are many public charging stations, and investors are more likely to build charging stations if there are many drivers with electric cars.

[2]As argued by Dickson et al. (2001): "Part of the utility derived from seeing a movie is talking about it with friends. The more a movie is seen by friends and talked about, the greater is the utility of seeing the movie so that you can participate in the discussion" (p.229).

[3]Godes et al. (2005) refer to social interactions between consumers as contexts where "consumer choice is influenced in a direct and meaningful way by the actions taken by others. These "actions" range from

To explore this issue, we develop a theoretical model of product launch in the presence of network effects, with two key features. The first is that we model consumer behavior at the individual level. The second is that we assume consumers are rational and forward-looking, and that they take all available information into account when forming expectations about each others' behavior. In contrast, the typical approach in the management and marketing literature on network effects has either been to assume that consumers follow a rule-of-thumb for decision-making, or to model demand at a more aggregate level. Our approach allows us to explicitly consider how different launch strategies (or 'sales schemes') can affect product sales, by influencing the expectations of strategic consumers. In so doing, the analysis generates insights into a number of issues of practical relevance for product launch.

A first issue is whether it makes sense to release sales figures or publish the number of product preorders. In practice, certain firms do just that. Tesla regularly published the number of preorders of its Model 3 Sedan, and Apple released information on iPhone preorders and early sales.[4] Movie studios systematically release weekly figures of box office receipts, which are then published online and in trade publications. Our analysis shows that releasing information generates a tradeoff in the presence of network effects. Success can breed success, as consumers who observe that the product is popular will become more willing to buy, in part because they also expect others to buy in the future. But by the same token, failure can breed failure. We show that, on average, the firm benefits from releasing information, which influences how consumers expect others to react to their own choices. The results suggest that releasing the maximum amount of preorder and sales information to consumers can help firms to exploit network effects to the greatest possible extent.

A related point is whether a firm may want to adjust its approach of releasing sales

face-to-face recommendations from a friend to the passive observation of what a stranger is wearing" (p.415). In this sense, network effects can be seen as one type of social interaction. The broader issue is how a firm can manage and leverage such social interactions, in particular when launching new products (Bloch (2016), Mayzlin (2016)).

[4]For Model 3 preorders, see https://twitter.com/elonmusk/status/716693951260938241?lang=en, accessed on September 1, 2017. For iPhone preorders, see 'Apple Announces Record Pre-orders for iPhone 6 & iPhone 6 Plus Top Four Million in First 24 Hours', on http://www.apple.com/pr/library/2014/, accessed on September 1, 2017.

information at an intermediate stage of the launch process. That is, if early released sales figures are disappointing, will the firm benefit from no longer releasing sales information in the future? Surprisingly, we show that the answer is no. The firm will certainly dislike the fact that the information released pushed down the willingness to pay of consumers who remain in the market. Nonetheless, the best way, on average, to exploit network effects amongst these consumers is to continue releasing as much information as possible about subsequent sales.

A second issue relates to multi-market entry, specifically whether a firm should launch its product simultaneously in multiple markets, or follow a more sequential approach, launching first in some markets and then in others. This is the classic dilemma of choosing between a sprinkler strategy (simultaneous launch) and a waterfall strategy (sequential launch). The stated purpose of sequential launch is often that of progressively building on early success. For example, Bruce Snyder, Twentieth Century Fox President of Distribution in 2006, stated that following a tiered release pattern for the movie Borat would allow them to build momentum.[5]

Our analysis shows that the tradeoff between a simultaneous and sequential launch strategy, in terms of exploiting network effects in our formal framework, is the same as that regarding the release of sales information to consumers. A sequential launch strategy provides consumers in some markets with information about the product's popularity in others, so that success can breed success. Failure can also breed failure, but on average the firm benefits, as a sequential strategy stimulates purchase behavior.

A third issue is what type of sequential product launch strategy to use when faced with markets that differ in size. Does the presence of network effects make it more attractive to launch in large markets before moving on to small ones, or vice versa? First launching in large markets reveals a good deal of information about product popularity, so that early sales have a particularly large impact on success in other markets. However, since these other markets are small, this impact affects the decision-making of relatively few consumers. In contrast, first launching in small markets implies that early sales reveal less information,

[5] See `http://www.firstshowing.net/2006/borat-a-limited-release-on-november-3rd/`, accessed on September 1, 2017.

but this information then influences more consumers.

We demonstrate that it is the later strategy, launching in small markets first, which generates higher sales. The implication is that network effects should push managers towards launching their product across markets in increasing order of market size. For example, this could correspond to first releasing a new movie in relatively small markets abroad before moving on to the larger home market. Supporting this idea, Marvel Studios President Kevin Feige reportedly stated that one reason to first release movies in selected foreign markets is to highlight impressive foreign box-office numbers when later releasing in the United States.[6]

A fourth issue is whether a firm can benefit from offering the product to certain specific consumers before others, and if so, who to target first. Similar questions arise in relation to viral marketing strategies, which look to leverage social interactions between consumers in the most effective possible way.[7] We address this issue in our setting by assuming that some consumers are more independent-minded than others, in the sense of placing less weight on the network effect. Targeting independent-minded consumers first reveals information to those who value it most, since those who then observe early sales are more susceptible. This means that the consumers targeted first can have a large impact on others. However, since these consumers are independent-minded, this impact has little effect on their own incentive to buy. We show that nonetheless, the best way to exploit network effects is to serve more independent-minded consumers first, so that they can serve as opinion leaders for those who follow.

A fifth and final issue relates to pre-release consumer communication. Consumers often communicate with one another and share views about new products prior to their launch, whether on blogs, discussion forums, or via entries on social media.[8] We consider whether

---

[6]See `http://www.cinemablend.com/new/Why-Marvel-Opens-Its-Superhero-Movies-Overseas-First-71409.html`, accessed on September 1, 2017.

[7]As argued by Iyengar et al. (2011a): "Marketers are increasingly experimenting with various forms of network marketing. In the area of new product marketing, the rationale of many such efforts rests on three key assumptions: (1) social influence among customers is at work, (2) some customers' adoptions and opinions have a disproportionate influence on others' behavior, and (3) firms are able to identify and target those influentials or opinion leaders. These assumptions are quite reasonable and have been supported by prior research and experience" (p.17).

[8]For example, consumers engage in heated debate about the perceived merits of rumored Apple products

consumers can be willing to communicate truthfully and whether pre-release communication can influence the firm's subsequent launch strategy in the presence of network effects. The answer to both questions is yes in our setting. Specifically, pre-release consumer communication reduces the importance of other aspects of the launch decision (i.e. release of sales figures, or sequential vs. simultaneous launch), as it leaves less scope for post-launch consumer learning. Our results also suggest that such communication can boost sales, implying that managers may want to facilitate pre-release communication, for example by announcing the upcoming launch of products well in advance.

In terms of contribution, our paper presents a unified theoretical framework that allows us to explore the impact of network effects on a variety of managerial decisions related to product launch. By considering strategic, forward-looking consumers, and modeling their decisions at the individual level, we are able to analyze the key impact that expectations of future sales have on purchase decisions. These expectations, and feedback effects with current purchase behavior, are crucial in driving our results, in particular regarding release of sales information and multi-market entry. The key strategic issue in our formal framework is that a consumer's incentive to buy will differ depending on whether his purchase decision is observed by others. Specifically, if decisions are observed, then early sales can help encourage others to buy as well, which makes buying more attractive. This in turn makes it more likely that later consumers will observe high sales and buy themselves. The release of sales information and sequential multi-market entry can be a double-edged sword, in that success can breed success, and failure can also breed failure. But the use of such a sales scheme itself influences consumers in a way that makes failure less likely.

One interpretation of our mechanism driving sequential sales is that consumers explicitly reason how their own behavior can influence others who follow. Take the example of Google Chromebooks, with Google Apps, which from 2012 to 2017 became a major force in the American school system. A key reason for their success is that teachers and administrators in schools using the products actively tried to influence others to also use them.[9] The network

that have yet to appear (see, e.g., www.9to5mac.com and www.appleinsider.com). For movies, Liu (2006) documents high levels of word-of-mouth communication on the Yahoo Movies message board in the period prior to release.

[9]For further details, see https://www.nytimes.com/2017/05/13/technology/google-education-

benefits from increased use include better opportunities for students to collaborate, and for teachers to share best practices, along with increased efforts from Google to develop the technology, driven in part by its broad adoption.

An alternative interpretation, given our focus on equilibrium behavior, relates to the evolution of consumers' behavioral strategies. If consumers acted differently from what our analysis suggests, then some consumer would benefit on average from changing his behavior. As a result, learning to play the equilibrium over time, rather than explicit reasoning, would tend to lead to higher sales under a sequential scheme. It may also be that a consumer does not take into account his own influence, but still realizes that serving latter markets sequentially will tend to boost future sales. Consistent with others' equilibrium behavior, he may realize that high sales in one of these markets can push more consumers to buy in the other, and increase his own benefit from buying.

One may wonder whether this effect, where each consumer's choice helps to 'get the bandwagon rolling', will also be present in large markets. Related to this point, our formal analysis assumes that the network effect depends on the proportion of consumers who buy the product, rather than on total sales. This means that each consumer's purchase has only a small direct effect on any other individual when market size is large. That being said, there are three countervailing forces in large markets that support the relevance of sequential sales. First, in a large market, each consumer's purchase can directly influence many individuals. Second, each consumer's purchase may have a substantial indirect impact, since the fact that one consumer buys can push another consumer to buy, which can push another to buy, and so on. Third, even if a consumer does not consider the impact of his own purchase, he should realize there are many other consumers in the market who can push each other to buy under sequential sales. If the network effect depended on total sales, as might be reasonable in certain empirical settings, then an increase in market size would strengthen all three forces, and make buying even more attractive under sequential sales.[10]

`chromebooks-schools.html?mcubz=1`, accessed on September 1, 2017.

[10]For any fixed market size, assuming that network effects are proportional to total sales, rather than the proportion of consumers who buy, would effectively amount to a different normalization in our formal framework. Specifically, it would correspond to scaling our network-effect parameter by the number of consumers in the market.

2 Literature Review

There has been a long and sustained interest in studying network effects, both in economics and in management and marketing. However, this research has focused on a different set of issues than we do in our paper. Other work has examined many of the issues we consider here related to product launch, but has not explicitly incorporated network effects in the analysis.[11]

Research in marketing and management that explicitly models network effects has focused on how adjusting different marketing variables can boost sales or speed up product adoption. A common point of interest has been pricing (see, e.g., Dhebar and Oren (1985), Dockner and Jorgensen (1988)). Other work has combined this with different issues, such as compatibility with competitor products (Xie and Sirbu (1995)), and quality provision (Padmanabhan et al. (1997)). Still other work has considered the choice between offering product line extensions or licensing of compatible products (Sun et al. (2004)), and inefficient purchase delay, if consumers will only buy a product after a certain number of their neighbors have done the same (Goldenberg et al. (2010)).

While overall, the focus of these papers is very different than ours, certain results in Padmanabhan et al. (1997) do touch on sequencing. They show that a firm may first want to release a product to experts followed by a lower-quality version to novices, to help signal private information about demand to consumers through its price and quality choice. In contrast, we consider how sequential sales may help consumers learn from one another about the product's popularity.

Our analysis also differs in explicitly considering consumer expectations about future purchases. Dhebar and Oren (1985), Dockner and Jorgensen (1988), and Xie and Sirbu (1995) work with diffusion models, where potential demand and the hazard rate of adoption may depend on past adoption, but where there is no issue of consumer expectations. Expectation of future purchases also play no role in Goldenberg et al. (2010), where consumer decision-making follows a simple rule of thumb, in Sun et al. (2004), who considers a static setting,

[11] In our review of the research on network effects, we mainly consider theoretical work. Empirical work on network effects is also abundant (see Stremersch et al. (2007), with a focus on indirect network effects). Farrell and Klemperer (2007) also present an extensive review with a focus on competition.

or in Padmanabhan et al. (1997), where network effects only depend on current sales. While these models all capture the idea that past sales can drive current purchase behavior, they do not consider the interaction between current behavior and expectations about the future, which is crucial in our analysis.

Much of the research on network effects in industrial organization has also focused on pricing, often in conjunction with other issues related to the introduction of new technologies. Expectations here have often played a central role. These issues include whether self-fulfilling expectations can lead to multiple equilibria, and whether competitors will benefit from their technologies being compatible (see, e.g., Katz and Shapiro (1985)). They also include which of multiple incompatible technologies will tend to become dominant, and in particular whether it need be the most efficient technology (Farrell and Saloner (1985), Farrell and Saloner (1986), Katz and Shapiro (1986)). Other work has looked at advertising as a way to coordinate consumer expectations towards buying certain products (Bagwell and Ramey (1994), Pastine and Pastine (2002), Clark and Horstmann (2005)). The related literature on platforms and two-sided markets has focused on externalities between participants on one side of a market (say video game developers) and another (say potential players), where the role of the platform is often to get 'both sides of the market on board' (see Rochet and Tirole (2003), Rochet and Tirole (2006), and Armstrong (2006) for analysis in a static setting, and e.g. Cabral (2011), Veiga (2015) for analysis in a dynamic setting).[12]

Our modeling approach resembles, in broad strokes, that employed in much of this literature: we consider strategic players, and model decision-making at the individual level, where both past sales and expected future sales impact purchase behavior. The key strategic motive for consumers in our setting - that consumers have different incentives if their own choice can directly influence others' behavior - is also present in Farrell and Saloner (1985), along with related work on entry timing in network games.[13] Our paper differs in being

[12]An issue of interest in two-sided markets has been 'divide and conquer' strategies, where a platform's pricing policy leads participants on one side of the market to effectively subsidize those on the other. See, e.g., Jullien (2011).

[13]See Ochs and Park (2010): "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

the first to demonstrate the implications of this strategic motive of consumers for firm and managerial decision-making.

In terms of product launch, it has long been recognized that launch decisions can be crucial for product success. The question has then been to identify precisely which aspects are particularly important. For example, Di Benedetto (1999) points to launch timing and to firm marketing and technical skills. Just as Talke and Hultink (2010), he distinguishes between launch strategy (earlier decisions related to product type, innovation, choice of markets) and launch tactics (later decisions such as launch timing and management, as well as the marketing mix). One way to view our analysis is in terms of exploring the implication of network effects for a variety of tactical issues (information release to consumers, launch order, targeting individual consumers, favoring pre-release consumer communication), within a formal theoretical framework.

In so doing, our analysis relates to the qualitative work by Lee and Colarelli O'Connor (2003), who emphasize that network effects should be taken into account for product launch, and that particular marketing strategies designed to influence consumer expectations can be crucially important. It also relates to Hauser et al.'s (2006) call for research that helps shed light on the role of network effects in the take-off and growth of new products, and that develops normative tools that can help managers exploit network effects.

Turning to the more specific issues, there has been little formal analysis, to the best of our knowledge, of whether firms should release sales information about their product or publish pre-orders. Di Benedetto (1999) argues that information gathering from consumers and obtaining feedback is important, both during and after product launch. However, as in Xiong and Bharadwaj (2014), the stated purpose of such information gathering is to help the firm adjust its launch strategy, not to release this information to influence consumer behavior. Many firms do gather detailed information on sales figure, for example in the movie industry, where box office tracking companies provide studios with box office figures in close to real time.[14] Similarly, tools such as Nielsen BookScan and Soundscan provide detailed sales data

reaction of widespread adoption" (p.690).

[14]Tools for tracking such figures include 'Box Office Essentials' from comScore (formerly Rentrak). See `https://www.comscore.com/Products/Movies-Worldwide/Box-Office-Essentials`, accessed on September 1, 2017.

for industry use, suggesting that the potential scope for information release is quite large.

Regarding multi-market product launch, a large literature has considered the choice between launching simultaneously or sequentially across markets, so whether to follow a 'waterfall' or 'sprinkler' strategy. An important point here is the lead-lag effect, the extent to which sales in one market may influence sales in the other. This literature has identified a variety of factors that can push firms towards choosing a sequential launch. These include a strong lead-lag effect, low competitive pressures, high entry costs, and unfavorable conditions in the foreign (rather than the home) market (see, e.g., Ayal and Zif (1979), Kalish (1985), and Ganesh and Kumar (1996)). Libai et al. (2005) argue that business practice and the academic literature generally come out in favor of sequential entry, whereas their own analysis shows that a 'support-the-strong' strategy may be ineffective, providing evidence in the opposite direction. Ohmae (2000) suggests that increased information flows, rapidly changing technology, and competitive pressures have increased the importance for firms to simultaneously conquer multiple markets by following a sprinkler strategy.

Instead of explicitly comparing simultaneous and sequential launch strategies, other studies have focused on particular aspects of sequential launch. For example, Eliashberg and Helsen (1996) measure the size of the lead-lag effect in the market for VCRs. Bronnenberg and Mela (2004) consider interactions between manufacturers and retailers in their study of staggered, regional roll-out, which they argue is standard for many repeat-purchase products, whereas Elberse and Eliashberg (2003) consider sequential movie release, and document strong cross-country effects for box office performance. In terms of launch order and market size, Putsis et al. (1997) suggests targeting large markets before small ones to best exploit cross-market effects, based on estimation results in a diffusion model that show different-sized countries have different rates of external contact. This suggestion contrasts with our result on launch order, which concludes that the best way to exploit network effects is to serve smaller markets first.

That being said, none of this work on multi-market product launch explicitly considers network effects. The standard approach in this literature is to work with diffusion models, which has the advantage of tractability in modeling cross-market effects, but the disadvantage of modeling adoption decision at the aggregate level. Peres et al. (2010) argue that interactions in such models are usually interpreted as consumer word-of-mouth, but that a fascinating recent shift has been in considering whether they can also be interpreted as network effects or related social signals. Our analysis suggests that the answer is no, because consumer expectations are crucial in the presence of network effects, and diffusion models do not allow expectations to play any role. Peres et al. (2010) suggest that more research should focus on adoption decisions at the individual level, to explicitly incorporate different types of interactions (such as network effects) and individual heterogeneity, both of which we do in our analysis.

Our results on sequential versus simultaneous launch also relate to a small theoretical literature on optimal sequencing, in situations where consumers have private information about product quality. The key element in this literature is how sequential sales may allow consumers to infer something about quality by observing each others' purchase decisions. Sgroi (2002) shows that simultaneously serving a group of 'guinea pigs' can prevent an information cascade where all later consumers refrain from buying. Liu and Schiraldi (2012) show that it is often optimal to serve all consumers simultaneously when prior beliefs about quality 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.[15]

A key conceptual difference between our paper on network effects, and work on quality uncertainty and observational learning, is that consumers are essentially backward looking in the latter literature. That is, consumer expectations about future sales have no impact on their purchase decisions, unlike in our setting. In terms of results, certain papers on observational learning suggest that a firm may want to adjust its launch strategy (e.g. move from sequential to simultaneous) at an intermediate stage, if early sales are particularly high or low. That is not the case in a setting with network effects. We show that on average, the best way to exploit network effects is to always continue serving consumers sequentially.

[15]These papers relate to a broader literature on how firms can influence observational learning, for example through pricing, product testing, and marketing (see, e.g., Ottaviani and Prat 2001, Bar-Isaac 2003, Bose et al. 2006, Bose et al. 2008, Gill and Sgroi 2008, Kuksov and Xie 2010, Debo et al. 2012, Gill and Sgroi 2012, Miklós-Thal and Zhang 2013).

Our results on targeting independent-minded consumers relate to earlier work on seeding and opinion leadership.[16] Work on targeting has often considered how a person's influence can depend on her centrality in a particular network or her number of connections (Zenou (2016)). As described by Hinz et al. (2011), different empirical studies have offered support for a variety of strategies: targeting consumers with many connections (hubs), consumers with few connections (fringes), or consumers who link otherwise separate parts of the network (bridges). As for theoretical studies, Campbell (2013) explicitly considers information percolation and word-of-mouth communication within a network, and suggests targeting consumers with few connections. Galeotti and Goyal (2009) suggest that with network effects, it may instead make sense to target consumers with many connections, if these consumers are otherwise unlikely to purchase the product. Other work has employed agent-based modeling and used real-world data as inputs to their simulations (see Bloch (2016) for a review).

Much recent work has suggested that influence can also depend on personal characteristics (i.e. a person's 'type'). Iyengar et al. (2011b) present physician survey data showing that high-volume prescribers have more impact on their peers' decision of whether to prescribe a new drug. Libai et al. (2005), Trusov et al. (2010), and Hinz et al. (2011), in different settings, all suggest that people who are most persuasive, in the sense of effectiveness at influencing others' behavior, are not necessarily those who are most well connected. Aral and Walker (2012) present evidence from a field experiment with network effects suggesting that consumer type matters, both in terms of who is more influential, and who is more susceptible to others' influence. Making the link with our framework, we allow consumers to explicitly differ in their type (independent-minded versus susceptible to network effects). The firm's targeting decision corresponds to choosing each consumer's position within the network structure, in the sense of determining to what extent his purchase decision is observable to others, which is what affects that consumer's influence.

Finally, turning to direct consumer communication, most of the literature has focused on the situation after consumers have made their purchase decisions (see, e.g., Godes et al.

[16] The notion of innovators and imitators goes back to Bass (1969) and related diffusion models that involve both a coefficient of innovation and of imitation. But since such models do not explicitly consider individual consumer heterogeneity, there can be no question of which particular consumers to target.

(2005) and the references therein). Our analysis instead focuses on pre-release consumer communication. The distinction is relevant since, as argued by Xiong and Bharadwaj (2014), pre-release communication is more likely to reflect general consumer interest in the product, rather than product quality. Xiong and Bharadwaj (2014) describe a recent, growing, empirical literature on pre-release communication, suggesting that pre-sale 'buzz' is informative about subsequent sales, for products such as video games, music, and movies, either because buzz reflects consumer perception, or influences consumer perception. Both channels play a role in our setting: pre-release cheap talk reveals information about each consumer's willingness to pay, and the total information revealed then affects consumers' purchase decisions, because of network effects.

3 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 $[\underline{\lambda}, \overline{\lambda}]$, with $0 < \underline{\lambda} \leq \overline{\lambda} < \overline{\theta} - v_0$. 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.

The timing of the game is as follows. At $t = 0$, the seller sets a price $p \in \mathbb{R}_+$ and selects a sales scheme. The sales scheme determines the extent to which consumers can learn from each other about the product's popularity, by observing each others' purchase decisions. Specifically, the seller chooses a number of cohorts $m \leq n$ and how to partition the n consumers between the m different cohorts, $\mathcal{I} = \{I_1, \ldots, I_m\}$. Because type is unobservable, the seller cannot distinguish between different consumers, so the seller's choice of sales scheme $\mathcal{I}$ is effectively a choice of m (the number of cohorts) and the cardinality of $I_1, \ldots, I_m$ (the size of each cohort).

There are various ways that our formal notion of a 'sales scheme' can be interpreted in practice. One interpretation relates to the choice whether or not to release sales information to consumers. Another relates to the choice whether to launch in some markets before others. Yet another relates to the targeting of individual consumers. For some of these interpretations, it may be reasonable to think of a manager effectively choosing between a limited number of feasible schemes, the set of which might well depend on the precise interpretation in question. We do not restrict the set of possible sales schemes at this stage (i.e. we assume that all schemes are feasible) in order to maintain generality. That being said, we return to these issues in the analysis when presenting our formal results and discussing their interpretation.

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 $t' \leq t - 1$. After consumers in cohort I_m make their purchase decisions, payoffs accrue to all players, and the game ends. There is no discounting.

For the main analysis, we assume static pricing, where the seller fixes p at $t = 0$.[17] Moreover, we assume that the seller commits at $t = 0$ to its choice of sales scheme. We later show that our conclusions remain unchanged if the seller is unable to commit to its chosen scheme, and also address the issue of dynamic pricing.

We also assume, as in the work on quality uncertainty surveyed in the literature review, that consumers' purchase decisions are irreversible. An interpretation of irreversibility is

[17] 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.

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 2016).[18]

To describe the strategy of a consumer i in cohort I_t, note that 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 $\mathcal{I}$ 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 $\mathcal{I}$.

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.[19] 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 $\mathcal{I}$. We assume $\underline{\theta} + 1 < v_0 < \overline{\theta}$ to guarantee interior solutions as described below.

4 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

[18] This assumption is consistent with our desire to model situations where the seller has at least some control over 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).

[19] 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.

expects other consumers to behave. Consumer i will find it optimal to buy himself if

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

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 on average 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 $\mathbb{E}(\lambda) \leq \overline{\lambda}$ 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 $\mathcal{I}$, 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.*

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. That is to say, there is a unique equilibrium strategy profile of consumers, given any choice of price and sales scheme. As a result, every seller strategy can be associated with a unique value of expected profits. This uniqueness is useful for comparisons because it means that the seller can unambiguously rank different schemes in terms of profitability.[20]

Relaxing our assumption on parameter values would mean that multiple equilibrium consumer strategy profiles, and multiple values of expected profits, could be consistent with a single sales scheme. In particular, if network effects were sufficiently strong, then any given scheme could generate two diametrically opposed equilibrium outcomes: a good outcome where all consumers buy, and a bad outcome where nobody buys. This multiplicity would generate ambiguity in terms of how to rank different schemes, and such rankings are the focus of our analysis.[21]

4.1 Sequential versus simultaneous sales

Given uniqueness, we are now in a position to examine now the impact of network effects on consumer behavior depends on the firm's sequencing of sales. In particular, we are interested in whether or not the seller benefits from allowing consumers to learn about the product's popularity, by observing each others' purchases decisions. In order to do so we make the following definition.

Definition 1. *A sales scheme* $\mathcal{I}' = \{I_{1'}, \ldots, I_{t'}, \ldots, I_{m'}\}$ *is* more sequential *than another scheme* $\mathcal{I} = \{I_1, \ldots, I_t, \ldots, I_m\} \neq \mathcal{I}'$ *if any two consumers in the same cohort under* $\mathcal{I}'$ *are also in the same cohort under* $\mathcal{I}$*:* $i \in I_{t'}$ *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 first sales scheme is more sequential than a second one if all consumers who were served sequentially and at least some who were served simultaneously in the second scheme are served sequentially in the first. This is equivalent to saying that the second

[20] Equilibrium multiplicity could only occur in our setting if there were sufficiently strong strategic complementarities in consumers' purchase decisions. These strategic complementarities are a result of network effects. Thus, we are effectively assuming that network effects are not strong enough to completely overwhelm the intrinsic concern for the product displayed by all types.

[21] Our results will show that a sequential sales scheme is attractive for the seller, in terms of maximizing expected profits. In a setting with equilibrium multiplicity, sequential sales may also be attractive, but for a different reason. Intuitively, sequential sales might then help with equilibrium selection, if observing an early purchase can coordinate the remaining consumers on the Pareto dominant outcome.

scheme can be transformed into the first by repeatedly breaking up cohorts, taking groups of consumers who were served simultaneously and instead serving some of these consumers before others. Alternatively, the first scheme can be transformed into the second by repeatedly combining adjacent cohorts and serving these consumers simultaneously.

The fact that a sales scheme is more sequential than another means that it will unambiguously give consumers more information about the product's popularity before they make their own purchase decisions. Definition 1 allows us to make pairwise comparisons between a wide variety of 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). The later scheme minimizes the information available to consumers, as each consumer makes their purchase decision without observing the choices of any others. Similarly, a scheme with a single consumer per cohort (fully sequential) is more sequential than every other possible scheme, as it maximizes the amount of information available to consumers.

The following result says that the seller can best exploit network effects by using a sales scheme that is as sequential as possible.

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

A sequential scheme provides consumers with more information about the product's popularity, by allowing them to observe each others' purchase decisions. 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. Proposition 2 shows that sequential sales drive up expected profits despite the fact that failure can also breed failure, where low early sales can depress sales from those who follow.[22]

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

[22]Note that Proposition 2 holds regardless of whether the firm sets the optimal price for each scheme or simply sets the same price for both schemes.

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.

This virtuous cycle can involve multiple feedback loops between consumers' expectations of each others' purchase decisions. The fact that a consumer in an early cohort knows that his own purchase can help 'get the bandwagon rolling' makes it more attractive for him to buy. This consumer realizes that others in the same cohort will also be more likely to buy, by the same reasoning, which itself makes buying even more attractive. Moreover, these other consumers can also help 'get the bandwagon rolling', which further increases expectations about future sales, which increases each consumer's incentive to buy further still. Finally, the fact that early consumers are more willing to buy, by the above reasoning, makes it more likely that later consumers will actually observe high sales, and actually buy themselves.

We now relate Proposition 2 to the issue of whether to release pre-order and sales information to consumers. With this interpretation, a manager would progressively receive information about pre-orders or sales, by some prespecified process. For example, they might first learn about the purchases from a subset I_1 of consumers at $t = 1$, then about the purchases from another subset I_2 of consumers at $t = 2$, and finally about the purchases from a subset I_3 of consumers at $t = 3$, where $I_1 \cup I_2 \cup I_3 = n$. The subsets I_1, I_2, and I_3 are exogenous and implicitly determine the set of feasible schemes. This amounts to assuming that the extent to which a firm can access fine-grained sales information is fixed, at least in the short run. In practice, the information available to the firm may well depend on its longer-term strategic decisions. The tactical question considered here, in regards to product-launch, is whether the information that can be accessed should be disclosed to consumers.

In this example, a sales scheme would correspond to a rule that specifies whether or not to release previous sales information at the end of each period. There are four feasible schemes: $\{I_1, I_2, I_3\}$, which corresponds to releasing information after both $t = 1$ and $t = 2$; $\{I_1, I_2 \cup I_3\}$, which corresponds to only releasing information after $t = 1$; $\{I_1 \cup I_2, I_3\}$, which corresponds to only releasing information after $t = 2$; and $\{I_1 \cup I_2 \cup I_3\}$, which corresponds to never releasing information until all consumers have made their purchases. Proposition

2 says that the best way to exploit network effects is to choose the most sequential scheme, out of those that are feasible, in this case $\{I_1, I_2, I_3\}$. Doing so amounts to releasing pre-order and sales information to consumers at the rate at which the firm receives it, in order to inform consumers to the maximum possible extent. This is better than never releasing information, as with $\{I_1 \cup I_2 \cup I_3\}$, and is also better than any intermediate scheme.

A setting where the seller has access to very fine-grained sales information, which it can disclose in close to real time, would correspond to a situation where effectively all sales schemes are feasible. In such a situation, Proposition 2 immediately implies what is the optimal scheme.

Corollary 1. *The sales scheme $\mathcal{I}$ that maximizes expected profits has a single consumer per cohort.*

We can also relate Proposition 2 to the issue of multi-market product launch. With this interpretation, a manager would explicitly choose whether to launch the product sequentially or simultaneously across markets. For example, I_1 could refer to consumers in a first market, I_2 to consumers in a second market, and I_3 to consumers in a third market, where $I_1 \cup I_2 \cup I_3 = n$. The subsets I_1, I_2, and I_3 are again exogenous and determine the set of feasible schemes. Here, the idea is that the manager can launch the product sequentially across different markets, but not within any single market, and that the boundary of each market is fixed, at least in the short run. An interpretation is that longer-term firm strategic decisions may determine the set of available markets. The tactical choice at hand is how to sequence sales across the markets that are available.

With this interpretation there are thirteen feasible schemes: six schemes that correspond to launching sequentially across all three markets, in different possible orders, such as $\{I_1, I_2, I_3\}$ and $\{I_1, I_3, I_2\}$; three schemes that correspond to first launching in a single market and then simultaneously in the two remaining markets, such as $\{I_1, I_2 \cup I_3\}$ and $\{I_2, I_1 \cup I_3\}$; three schemes that correspond to first launching simultaneously in two markets and then in the remaining market, such as $\{I_1 \cup I_2, I_3\}$ and $\{I_1 \cup I_3, I_2\}$; and one scheme that corresponds to launching simultaneously across all markets, $\{I_1 \cup I_2 \cup I_3\}$. Proposition 2 suggests that a sequential launch across all three markets will allow a firm to exploit network effects to the maximum possible extent. Moreover, any of the six schemes that correspond to sequential launch will do better than a fully simultaneously scheme, providing support for a waterfall strategy, rather than a sprinkler strategy, for product launch.[23] We look at the question of how to choose among these six sequential schemes in the next section.

4.2 Ordering of sales

The analysis in Section 4.1 shows that the best way to exploit network effects is to use a sales scheme that is as sequential as possible. With the interpretation of information-release, this scheme corresponds to immediately releasing sales information to consumers. With the interpretation of multi-market product launch, it corresponds to launching sequentially across markets. However Proposition 2 says nothing specifically about ordering.[24] For example, when carrying out a sequential product launch, in which specific markets should the firm release the product first? The following result suggests how the optimal launch order will depend on market size.

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.

Holding the set of cohorts fixed, the firm benefits from serving consumers in smaller cohorts before consumers in larger ones. For multi-market product launch, this corresponds to launching the product in increasing order of market size. The intuition behind Proposition 3 is subtle, as it is not immediately clear which launch order maximizes the amount of information available to consumers. Serving large markets first means that relatively few consumers, those in smaller markets, have a great deal of information when making their purchase decisions. Serving small markets first means that relatively many consumers, those in larger markets, have a small amount of information. Proposition 3 shows that the former

[23]Stronger still, the results suggest that a simultaneous launch would be the worst possible choice in terms of exploiting network effects. In the above example, it would lead to lower expected profits than any of the twelve other feasible schemes.

[24]Formally, if not all feasible sales schemes are comparable in the sense of Definition 1, then Proposition 2 generates a partial order, rather than a total order, over the set of feasible schemes.

strategy is the best. Consumers in small markets understand that they have influence over the many consumers who follow, which makes them more likely to buy themselves.

The notion that a firm may benefit from launching its product sequentially in increasing order of market size stands in contrast to the suggestion made by Putsis et al. (1997), namely to target larger markets before smaller ones. Putsis et al. (1997) do not consider network effects, but instead estimate a flexible model of cross-country diffusion, where the relevant parameters for word-of-mouth communication and mixing patterns can vary both across products and across different countries. Their estimation results suggest that the large countries in their sample (such as Germany, France, Italy, and Spain) each tend to have a higher number of external contacts. The small countries in their sample (such as Denmark, Sweden, Netherlands, and Austria) instead each tend to have a high number of external contacts as a fraction of their total contacts. In this sense, Putsis et al. (1997)'s suggestion amounts to launching first in markets where adoption decisions are more influential, and latter in markets where adoption decisions are more easily influenced, in terms of cross-market effects. While Proposition 3 abstracts away from such issues of influence, we consider them in our next result on ordering, which focuses on the targeting of individual consumers.

Moving to this next result, we will relax the assumption that consumer type is private information. Intuitively, if a firm can acquire information about consumer characteristics, then it may well take this information into account in its sales strategy, and target particular consumers. In order to investigate the issue of targeting in more detail, we will focus on the situation where all sales schemes are feasible, and assume that the weight each consumer places on the network effect is public information.

As described in Section 2, analytical models of word-of-mouth communication in networks have 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, who 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.

In our setting, all consumers are equally influential from an ex ante perspective, in the sense that the network payoff depend on total sales but not on the identity of those 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 others' purchases. In contrast, consumers with low values of λ base their purchase decisions mainly on their intrinsic payoff from buying. The following result shows that the optimal sales scheme serves consumers in increasing order of λ, so in increasing 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 $\mathcal{I}$ 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, as in Section 4.1, where the intuition for this result echoes that from Proposition 2 and Corollary 1. Just as the qualitative advantage of a sequential scheme did not depend on the precise distribution of λ, it 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, 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 largely based on their own personal tastes. Buying 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.

The benefits of targeting more independent-minded consumers first, and serving these consumers before others, rely on consumers placing a strictly positive weight on the network effect. Consider the special case where certain consumers are completely insensitive to the network effect, corresponding to a parameter value $\lambda = 0$. Then the order in which these consumers are served will not have any effect on expected profits.[25] The seller's choice of ordering will affect the information available to these consumers, as well as the information they expect others to have about their purchases. However, this information has no impact on their own purchase behavior. More generally, there is a benefit of serving independent-minded consumers first, as long as these consumers are not completely independent minded.

Proposition 4 also relates to the broader economics literature on leadership, which has followed from Hermalin (1998). While precise definitions vary, this literature typically describes a leader as someone with potential followers, whose interests are at least partly aligned with his own and who can potentially be influenced by the leader's actions. A leader typically wants to induce followers to coordinate on the action he has chosen. In our setting, given a purely sequential sales scheme, the leader corresponds to the consumer in the first cohort. This consumer can influence all those in later cohorts through his own purchase decision. Interests are partly aligned because of network effects, where the first consumer considers how his own decision to buy can help push others to do so the same.

Unlike Hermalin (1998), Dewan and Myatt (2008), and Bolton et al. (2013), who model

[25] A similar conclusion would hold if some consumers were naive, and based their purchase decisions on a rule of thumb which was independent of past sales or future expectations. It would not matter when these particular consumers were served, but a sequential approach would still maximize sales from the remaining consumers.

leaders as having better information than followers, leadership in our setting come from a consumer's position in the sales scheme. Moreover, the seller explicitly selects the sales scheme that will maximize consumers' effective leadership, i.e. their ability to coordinate others on the seller's preferred action (to buy the product). While the first consumer can be viewed as the leader, a sequential sales scheme also allows later consumers to show leadership, to a lesser degree, by helping them influence those who come later still. Consumers are able to lead by example precisely because of the seller's choice of a sequential scheme, which allows them to observe each others' actions.

By serving consumers in increasing order of λ, the seller selects as leader the consumer who is least sensitive to others' behavior. This result echoes that from both empirical work (Kaplan et al. 2012) and theoretical work (Goel and Thakor 2008, Bolton et al. 2013) showing that 'resoluteness' can sometimes help with leadership. Resoluteness there is seen as a characteristic that makes one less susceptible to external influence, in the sense of often sticking with an early decision regardless of others' subsequent actions or of the arrival of additional information. Thus, broadly speaking, the leader in our setting, as selected by the seller, will be the consumer who is most resolute, and least influenced by the actions of others.

Given this interpretation, Proposition 4 provides a novel link between resoluteness and leadership. Bolton et al. (2013) show that in the face of a time-inconsistency problem, a resolute leader may be able to commit to a particular action, which can facilitate follower coordination. In contrast, in our setting, the seller selects the most resolute consumer as leader because he wants consumers who are more susceptible to be followers. It is not that having a resolute leader helps in and of itself; to the contrary, the principal would prefer that all consumers were easily influenced (high λ), which would generate larger bandwagon effects. Rather, for a given group of consumers, what helps is that the leader should be more resolute than the followers. This is what maximizes the extent to which consumers coordinate on buying the good, to the benefit of the seller.

4.3 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 observational 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.

Nonetheless, for a variety of products, including books, movies, 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. Indeed, as mentioned in Section 2, recent empirical work on pre-release communication suggests that the 'buzz' amongst consumer prior to product launch can reveal information about subsequent sales. We now consider such pre-release consumer communication. Our focus is on whether consumers can successfully communicate in equilibrium, and the implications of communication for the seller's choice of sales scheme.

A crucial point for communication to be successful relates to credibility. If consumers read certain comments 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 may 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)). 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. On the one hand, information that consumers share may be collected by the firm and used to adjust the price (see, e.g., Chen and Xie 2008).

A consumer who understates his willingness to pay may contribute to the impression that demand is low, leading to a price reduction. On the other hand, a consumer who overstates his willingness to pay may convince others to buy, and therefore himself benefit from a larger network payoff. This reasoning suggests that network effects might push consumers to overstate whereas firm monitoring might push them to understate. 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 (weakly) 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* $m \in [\underline{\theta}, \overline{\theta}] \times [\underline{\lambda}, \overline{\lambda}]$ *about their type which all other consumers observe, and where the seller observes* $M = (m_1, \dots, m_n)$ *with probability* $q \geq 0$*. 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^**. If* $q = 0$*, then consumer purchase decisions coincide with those in a setting where consumers all observe each others' type,* (θ_i, λ_i) *for all* $i = 1, \dots, n$*.*

To understand the intuition behind our result, consider the case where $q = 0$, so where the seller cannot observe consumers' communication. Consumers are then unafraid that sharing their enthusiasm for the product will contribute to a price increase. They actually have a weak incentive to overstate their willingness to pay, so that other consumers overestimate the network effect and become more likely to buy. However, a consumer can only strictly benefit from overstating if he is indeed willing to buy, as it is only then that others' behavior affects him. In other words, overstating can only matter for a consumer with sufficiently high willingness to pay, and such a consumer would also state they are willing to buy if they reported truthfully. The implication is that consumers obtain exactly the same payoff by truthfully reporting how many others must buy for them to buy themselves, as they would obtain by misrepresenting.

We now consider the case where $q > 0$. Proposition 5 shows that if q is small enough, potential incentive problems need not rule out successful communication, in that consumers may still truthfully reveal their willingness to pay to one another. Such communication can occur if consumers believe it sufficiently unlikely that the seller is monitoring their messages. Curiously enough, with $q > 0$, 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 expected price was increasing in the stated willingness to pay in consumers' messages, then consumers would all claim that their willingness to pay was very low, in the hopes of obtaining a price reduction.

As alluded to above, 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 hurts consumers who buy. The second effect of understating is that the seller may reduce its price. This price reduction helps consumers who buy but can only occur if the seller observes the messages. When the probability of observing 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 push other consumers to buy, increasing the network payoff, but can also lead to a higher price. However, the consumer who overstates will only benefit from the increased 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 in the first place.

Proposition 5 shows that pre-release consumer communication can be credible, in the sense of being incentive compatible for consumers. The managerial implication is that successful communication will decrease the importance of choosing any particular sales scheme when launching the product. In our framework, the seller's choice of sales scheme determines the extent to which consumers can learn from one another about product popularity by observing each others' purchase. However, if consumers successfully learn about each others' views on the product prior to launch, through direct communication, then there is little scope for future learning. Proposition 5 shows that consumers may effectively communicate all decision-relevant information about their type. This information can allow consumers to predict the product's popularity, which will no longer depend on the seller's choice of sales scheme. In particular, there is no longer need for the visibility of purchases provided by sequential sales.

A related issue is whether managers should encourage and promote pre-release consumer communication, for example through an official online forum or discussion board. That is, in our setting, does pre-release consumer communication increase seller profits, even in situations (captured by $q = 0$ in Proposition 5) where it cannot affect pricing? One complication is that successful communication can result in multiple values of sales being consistent with subsequent consumer behavior. That being said, if consumers coordinate on the most efficient equilibrium, with the highest sales, the seller may benefit from communication, in the sense of earning strictly higher profits than in a situation without communication but with a purely sequential scheme.[26] The intuition is that sequential sales benefit the seller by revealing information, but only from consumers who make their purchase decisions before others. Pre-release communication can go further by revealing information from all consumers, regardless of when they subsequently makes their purchases.

5 Discussion and Robustness

Our analysis has assumed that the seller commits to its choice of sales scheme, there is no discounting, 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 address the issue of dynamic, rather than static, pricing.

[26]We have analyzed a two-consumer example, which is available upon request, showing that consumer communication increases expected seller profits. A fully general analysis of this issue is beyond the scope of this paper.

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.

We have assumed that the firm does not discount cash flows from different cohorts of consumers. This assumption is compatible with our information-release story, as the choice of sequential scheme here is merely the choice of what consumers observe and does not have any impact on physical timing. However, when we speak about product launch in different markets, a sequential strategy inevitably delays some of the cash flows. Although the expected total cash flow is higher under a fully sequential scheme, this might not be true for the discounted cash flow, especially in situations where the discount factor is very low. In such situations, the firm might reject a fully sequential scheme in favour of a less sequential or even a fully simultaneous one. All the effects we describe in our analysis are still present, but they just might be outweighed by heavy discounting.

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). Dynamic pricing increases the difference in expected profits between these two schemes, relative to static pricing.*

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.

Finally, we conclude with a numerical example, under static pricing, to address whether a tradeoff arises between expected sales and risk from the seller's choice of sales schemes. Despite the notion that sequential sales may expose the seller to downside risk, where early failure breeds further failure, we now show that a sequential scheme may nonetheless dominate a simultaneous scheme, in the sense of first order stochastic dominance.

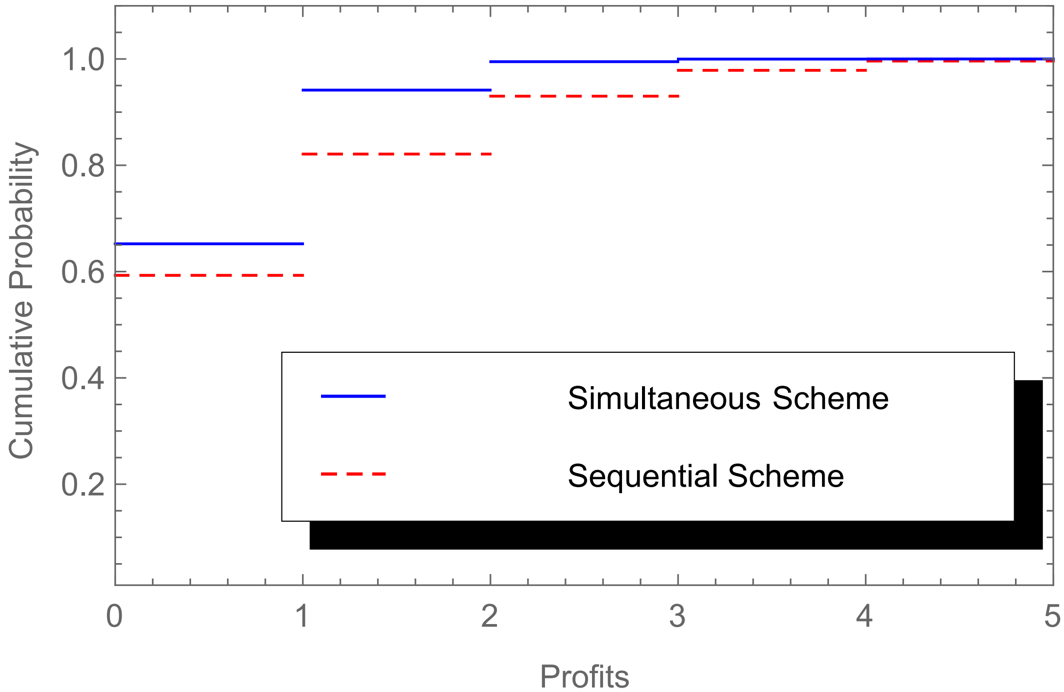

Figure 1: CDF of Profits for Simultaneous and Sequential Schemes

Figure 1 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.[27] 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. It is true that low early sales under a sequential scheme can dissuade later consumers from buying. However, the positive incentive effect of a sequential scheme on early consumers is so strong that it outweighs any increased risk that might arise from low early sales.

[27] The simulation uses parameter values $n = 5$, $\overline{\theta} = 2.2$, $\underline{\theta} = 0$, $v_0 = 1.1$, $p = 1$, and $\underline{\lambda} = \overline{\lambda} = 1$ under both schemes.

6 Conclusion

In a setting with network effects, consumers looking to buy a product will naturally take into account the expected purchase behavior of others. They may be more willing to buy an electric car or a mobile phone, movie tickets or books, if they believe that others will buy as well. For this reason, a firm selling a product that exhibits network effects would like consumers to believe it will likely become a 'hit'. We examine how the sequencing of sales can affect such beliefs, and allow consumers to lead by example, to best exploit network effects for the firm's own benefit.

In order to do so, we develop a theoretical framework where consumers are rational and forward-looking and where decisions are modelled at the individual level. We show that the firm's sequencing of sales can help consumers learn from each other about the product's popularity. This information also affects consumers' expectations about how others will behave in the future, which feeds back into their own purchase decisions. The key point in our framework is that each consumer realizes that those who observe high purchases will become more likely to buy, which increases his 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, by stimulating purchase behavior.

We explore the implications of our results for managerial decision-making related to product launch, for a firm looking to exploit network effects to the maximum possible extent. Our results suggest that a firm with access to sales information should release as much of this as possible to consumers, and that a sequential launch strategy across markets may be preferable to launching in all markets simultaneously. Moreover, the best way to launch sequentially is in increasing order of market size. In terms of targeting individual consumers, our results provide support for approaching independent-minded consumers first, so they can serve as opinion leaders for those who follow. Finally, pre-release consumer communication may be credible and may benefit the firm, but it will reduce the importance of the subsequent choice of sales scheme. All these different results are united by one common theme: the firm benefits from having large numbers of consumers base their purchase decisions on more information, in particular those consumers who value this information the most.

7 Appendix

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

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

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

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. As shown immediately after (4) in Section 4, 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

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

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

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 $\mathbb{E}(\lambda) \leq \overline{\lambda} < \overline{\theta} - v_0$ and $\underline{\theta} + 1 < v_0 < \overline{\theta}$, 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 $\frac{d\mathbb{E}(x_i|K_t)}{dK_t} > 0$.

□

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.

□

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

$$\theta_i^* = u_0 - \frac{\mathbb{E}(\lambda)}{n-1}\Bigg(K_{t-1} + \sum_{j \in I_{t-1} \setminus \{i\}} \mathbb{E}(x_j|K_{t-1}) + \sum_{0 \leq K' \leq \#I_{t-1}-1} \mathbb{P}\Big(\sum_{j \in I_{t-1} \setminus \{i\}} x_j = K'|K_{t-1}\Big) \sum_{l \geq t} \sum_{j \in I_l} \mathbb{E}(x_j|K_{t-1} + 1 + K')\Bigg),$$

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,

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

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 $K'=0,\dots\#I_{t-1}-1$. 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 $i\in I_{t-1}$, $\mathbb{E}(x_i|K_{t-1})=\frac{\overline{\theta}-\theta_i^*}{\overline{\theta}-\underline{\theta}}$ 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 $t'=t-2,t-3,\dots,1$ completes the proof.

□

7.2 Proofs of the Propositions

Proof of Proposition 1. 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

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

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

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 $\sum_{t' \ge t} \sum_{i \in I_{t'}} n_i - 1 \le n - 1$. By $\mathbb{E}(\lambda) \le \overline{\lambda} < \overline{\theta} - v_0$ and $\underline{\theta} + 1 < v_0 < \overline{\theta}$, 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 $t' \ge t+1$, are all uniquely defined. Hence, the cut-off function for consumer i,

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

given by (3) is uniquely defined as well.

$\square$

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

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

Suppose that under $\mathcal{I}'$, there are l consumers in cohort I'_t. Denote these consumers by subscript i, for $i = 1, \ldots, l$. Under $\mathcal{I}$, 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

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

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 $\mathcal{I}$. 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 $\mathcal{I}'$, 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

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

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 $\mathcal{I}'$.

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 $\mathcal{I}'$, 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 $\mathcal{I}$. From (11), the decision problem of consumers $k \in I''_t$ is the same as under $\mathcal{I}$, 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 $\mathcal{I}$, 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 $\mathcal{I}'$ 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 $\mathcal{I}'$ 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 $\mathcal{I}'$ 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 $\mathcal{I}'$ than under $\mathcal{I}$, given the assumption that the seller charges price p under both schemes. Let p' denote the optimal price under $\mathcal{I}'$. By the optimality of this price, ex ante expected profits under $\mathcal{I}'$ at price p' must be strictly higher than expected profits under $\mathcal{I}$ at price p.

□

Proof of Proposition 3. 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 $\mathcal{I}$. Suppose for now that the seller charges p under both schemes, so that both $\mathcal{I}$ and $\mathcal{I}'$ 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:

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

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

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

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 $\mathcal{I}'$ than under $\mathcal{I}$, if the seller charges price p under both schemes. Let p' denote the optimal price under $\mathcal{I}'$. By the optimality of this price, ex ante expected profits under $\mathcal{I}'$ at price p' must be strictly higher than expected profits under $\mathcal{I}$ at price p.

$\square$

Proof of Proposition 4. 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

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

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

$$S(\lambda_i, \lambda_{i+1}) \equiv \sum_{j=i}^{n} \mathbb{E}(x_j|K) =$$

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

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

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

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 $\sum_{1 \le i \le n} \mathbb{E}(x_i)$. It follows that ex ante expected profits, $p \sum_{1 \le i \le n} \mathbb{E}(x_i)$, 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.

□

Proof of Proposition 5. For consumer i, define N_i as the smallest value of N for which

$$\theta_i + \frac{\lambda_i}{n-1}N - p^* \geq v_0.$$

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 $(\theta, \lambda) \in [\underline{\theta}, \overline{\theta}] \times [\underline{\lambda}, \overline{\lambda}]$ 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 $p^* \in (\underline{\theta} - v_0, \overline{\theta} - v_0)$, 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 $m \in [\underline{\theta}, \overline{\theta}] \times [\underline{\lambda}, \overline{\lambda}]$ 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 $m \in [\underline{\theta}, \overline{\theta}] \times [\underline{\lambda}, \overline{\lambda}]$ 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 v_0. Each consumer who does not buy would receive a payoff strictly less than v_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

$$(1-q)\left(\mathbb{P}(N_{max} - \mathbb{1}_{m_i \leq N_{max}} < N_i)v_0 + \mathbb{P}(N_{max} - \mathbb{1}_{m_i \leq N_{max}} \geq N_i)\left(\theta_i + \lambda_i \frac{\mathbb{E}(N_{max}) - 1}{n-1}\right)\right) + qU_0,$$

where it is 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

$$(1-q)\left(\mathbb{P}(N'_{max} - \mathbb{1}_{m'_i \leq N'_{max}} < N_i)v_0 + \mathbb{P}(N'_{max} - \mathbb{1}_{m'_i \leq N'_{max}} \geq N_i)\left(\theta_i + \lambda_i \frac{\mathbb{E}(N'_{max}) - 1}{n-1}\right)\right) + qU_1,$$

where it is 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:

$$\Delta\mathbb{P}\left(\theta_i + \lambda_i \frac{\mathbb{E}(N'_{max}) - 1}{n-1} - v_0\right) > \frac{q}{1-q}(U_1 - U_0),$$

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 if $q = 0$, then the seller charges p^* with probability 1, which together with informative communication guarantees that consumers make the same purchase decisions as if they observed each others' types.

$\square$

Proof of Proposition 6. First consider a fully simultaneous scheme, with all consumers in a single cohort: $\mathcal{I} = \{I_1\}$, 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: $\mathcal{I}' = \{I'_1, \dots, I'_n\}$, with $n_t = 1$ for all $t = 1, \dots, n$. For each t, the seller's strategy specifies a price $p(K_t)$, for every possible value of previous sales $K_t = 0, \dots, t-1$. With slight abuse of notation, let $\pi(p(K_t)|\mathcal{I}')$ denote expected profits given this pricing schedule under this partition.

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

Finally, note that under a simultaneous scheme, dynamic and static pricing coincide. Moreover, for any given scheme, static pricing is equivalent to dynamic pricing with commitment which uses the same price in all periods. This implies that with a fully sequential scheme, expected profits with dynamic pricing can be no less than expected profits with static pricing. This proves the second statement of the proposition.

□

References

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

Aral, S. and Walker, D. (2012). Identifying influential and susceptible members of social networks. *Science*, 337(6092):337–341.

Armstrong, M. (2006). Competition in two-sided markets. *RAND Journal of Economics*, 37(3):668–691.

Armstrong, M. and Zhou, J. (2016). Search deterrence. *Review of Economic Studies*, 83(1):26–57.

Ayal, I. and Zif, J. (1979). Market expansion strategies in multinational marketing. *The Journal of Marketing*, pages 84–94.

Bagwell, K. and Ramey, G. (1994). Coordination economies, advertising, and search behavior in retail markets. *The American Economic Review*, pages 498–517.

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

Bass, F. M. (1969). A new product growth for model consumer durables. *Management Science*, 15(5):215–227.

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

Bloch, F. (2016). Targeting and pricing in social networks. In Bramoullé, Y., Galeotti, A., and Rogers, B., editors, *The Oxford Handbook of the Economics of Networks*.

Bolton, P., Brunnermeier, M. K., and Veldkamp, L. (2013). Leadership, coordination, and corporate culture. *Review of Economic Studies*, 80(2):512–537.

Bose, S., Orosel, G., Ottaviani, M., and Vesterlund, L. (2006). Dynamic monopoly pricing and herding. *RAND Journal of Economics*, 37(4):910–928.

Bose, S., Orosel, G., Ottaviani, M., and Vesterlund, L. (2008). Monopoly pricing in the binary herding model. *Economic Theory*, 37(2):203–241.

Bronnenberg, B. J. and Mela, C. F. (2004). Market roll-out and retailer adoption for new brands. *Marketing Science*, 23(4):500–518.

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

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

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

Clark, C. R. and Horstmann, I. J. (2005). Advertising and coordination in markets with consumption scale effects. *Journal of Economics & Management Strategy*, 14(2):377–401.

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

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

Dewan, T. and Myatt, D. P. (2008). The qualities of leadership: Direction, communication, and obfuscation. *American Political Science Review*, 102(03):351–368.

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

Di Benedetto, C. A. (1999). Identifying the key success factors in new product launch. *Journal of Product Innovation Management*, 16(6):530–544.

Dickson, P. R., Farris, P. W., and Verbeke, W. J. (2001). Dynamic strategic thinking. *Journal of the Academy of Marketing Science*, 29(3):216–237.

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

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

Elberse, A. and Eliashberg, J. (2003). Demand and supply dynamics for sequentially released products in international markets: The case of motion pictures. *Marketing Science*, 22(3):329–354.

Eliashberg, J. and Helsen, K. (1996). Modeling lead/lag phenomena in global marketing. In *The case of VCRs. Working paper*. The Wharton School, University of Pennsylvania Philadelphia, PA.

Farrell, J. and Klemperer, P. (2007). Coordination and lock-in: Competition with switching costs and network effects. *Handbook of Industrial Organization*, 3:1967–2072.

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

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

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

Ganesh, J. and Kumar, V. (1996). Capturing the cross-national learning effect: An analysis of an industrial technology diffusion. *Journal of the Academy of Marketing Science*, 24(4):328–337.

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

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

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

Godes, D., Mayzlin, D., Chen, Y., Das, S., Dellarocas, C., Pfeiffer, B., Libai, B., Sen, S., Shi, M., and Verlegh, P. (2005). The firm's management of social interactions. *Marketing Letters*, 16(3-4):415–428.

Goel, A. M. and Thakor, A. V. (2008). Overconfidence, CEO selection, and corporate governance. *Journal of Finance*, 63(6):2737–2784.

Goldenberg, J., Libai, B., and Muller, E. (2010). The chilling effects of network externalities. *International Journal of Research in Marketing*, 27(1):4–15.

Hauser, J., Tellis, G. J., and Griffin, A. (2006). Research on innovation: A review and agenda for marketing science. *Marketing Science*, 25(6):687–717.

Hermalin, B. E. (1998). Toward an economic theory of leadership: Leading by example. *American Economic Review*, pages 1188–1206.

Hinz, O., Skiera, B., Barrot, C., and Becker, J. U. (2011). Seeding strategies for viral marketing: An empirical comparison. *Journal of Marketing*, 75(6):55–71.

Iyengar, R., Van den Bulte, C., Eichert, J., and West, B. (2011a). How social network and opinion leaders affect the adoption of new products. *GfK Marketing Intelligence Review*, 3(1):16–25.

Iyengar, R., Van den Bulte, C., and Valente, T. W. (2011b). Opinion leadership and social contagion in new product diffusion. *Marketing Science*, 30(2):195–212.

Jullien, B. (2011). Competition in multi-sided markets: Divide and conquer. *American Economic Journal: Microeconomics*, 3(4):186–219.

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

Kaplan, S. N., Klebanov, M. M., and Sorensen, M. (2012). Which CEO characteristics and abilities matter? *Journal of Finance*, 67(3):973–1007.

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

Katz, M. L. and Shapiro, C. (1986). Technology adoption in the presence of network externalities. *Journal of Political Economy*, 94(4):822–841.

Kuksov, D. and Xie, Y. (2010). Pricing, frills, and customer ratings. *Marketing Science*, 29(5):925–943.

Lee, Y. and Colarelli O'Connor, G. (2003). The impact of communication strategy on launching new products: The moderating role of product innovativeness. *Journal of Product Innovation Management*, 20(1):4–21.

Li, S., Tong, L., Xing, J., and Zhou, Y. (2015). The market for electric vehicles: Indirect network effects and policy design. *Available at SSRN 2515037*.

Libai, B., Muller, E., and Peres, R. (2005). The role of seeding in multi-market entry. *International Journal of Research in Marketing*, 22(4):375–393.

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

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

Mayzlin, D. (2016). Marketing and networks.

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

Miklós-Thal, J. and Zhang, J. (2013). (de) marketing to manage consumer quality inferences. *Journal of Marketing Research*, 50(1):55–69.

Ochs, J. and Park, I.-U. (2010). Overcoming the coordination problem: Dynamic formation of networks. *Journal of Economic Theory*, 145(2):689–720.

Ohmae, K. (2000). *The Invisible Continent: Global Strategy in the New Economy*. Harper-Information.

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

Padmanabhan, V., Rajiv, S., and Srinivasan, K. (1997). New products, upgrades, and new releases: A rationale for sequential product introduction. *Journal of Marketing Research*, pages 456–472.

Pastine, I. and Pastine, T. (2002). Consumption externalities, coordination, and advertising. *International Economic Review*, 43(3):919–943.

Peres, R., Muller, E., and Mahajan, V. (2010). Innovation diffusion and new product growth models: A critical review and research directions. *International Journal of Research in Marketing*, 27(2):91–106.

Putsis, W. P., Balasubramanian, S., Kaplan, E. H., and Sen, S. K. (1997). Mixing behavior in cross-country diffusion. *Marketing Science*, 16(4):354–369.

Rochet, J.-C. and Tirole, J. (2003). Platform competition in two-sided markets. *Journal of the European Economic Association*, 1(4):990–1029.

Rochet, J.-C. and Tirole, J. (2006). Two-sided markets: a progress report. *RAND Journal of Economics*, 37(3):645–667.

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

Stremersch, S., Tellis, G. J., Franses, P. H., and Binken, J. L. (2007). Indirect network effects in new product growth. *Journal of Marketing*, 71(3):52–74.

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

Talke, K. and Hultink, E. J. (2010). Managing diffusion barriers when launching new products. *Journal of Product Innovation Management*, 27(4):537–553.

Trusov, M., Bodapati, A. V., and Bucklin, R. E. (2010). Determining influential users in internet social networks. *Journal of Marketing Research*, 47(4):643–658.

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

Veiga, A. (2015). Dynamic platform design.

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

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

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

Zenou, Y. (2016). Key players. In Bramoullé, Y., Galeotti, A., and Rogers, B., editors, *The Oxford Handbook of the Economics of Networks*.