Strategies and Implications of Entertainment Media Consumption

Franco Berbeglia

April 2021

Tepper School of Business,
Carnegie Mellon University

Submitted to the Tepper School of Business in Partial Fulfillment of the Requirements for the Degree of Doctor in Operations Management

Dissertation Committee:

Tim Derdenger (Chair)

Param Vir Singh

Kannan Srinivasan

Sridhar Tayur

Pascal Van Hentenryck

Joseph Xu (Chair)

Abstract

Entertainment consumption has changed drastically in the last decade, presenting new problems and questions to the entertainment industry. My research addresses three different problems in this industry. First, we study how to rank products optimally in a trial-offer marketplace where consumers present heterogeneous preferences and are influenced by past purchases. Second, we study the strategic implications of binge consumption of entertainment for advertising video on demand platforms. And third, we study the optimal release timing of movies between theatrical and home video (DVD and Blu-ray) markets, as the quality of the latter approaches the quality of the former.

In the first chapter, we study optimal ranking policies in a stylized trial-offer marketplace model, in which a single firm offers multiple products and has consumers who express heterogeneous preferences. The platform owner needs to devise a ranking policy to display the products to maximize the number of purchases in the long run, and to decide whether to display the number of past purchases. We find that when past purchases are displayed, consumer heterogeneity makes buyers try sub-optimal products, reducing the overall sales rate. We then show that consumer heterogeneity makes the ranking problem NP-hard (under Turing reductions) and we analyze the benefits of market segmentation. We find tight bounds to the expected benefits of offering a distinct ranking to each consumer class and we show that the market segmentation strategy always benefits from social influence. The firm is better off using an aggregate ranking policy when the variety of consumer preference is limited, but it should perform a market segmentation policy when consumers are very heterogeneous. This result is robust to relatively small consumer classification mistakes; when these are large, an aggregate ranking is preferred.

In the second chapter, we study the strategic implications of binge consumption of entertainment for advertising video on demand (AVOD) services. As on-demand video streaming services succeed,

traditional television-based media companies have begun to look for new methods to reach viewers. One such method is for these companies to distribute show episodes through a newly launched online AVOD service (e.g. NBCUniversal's Peacock, ViacomCBS's Pluto TV, Roku Inc., or Fox Corp.'s Tubi), which provides a new release timing strategy for episodes of shows. Besides the traditional sequential releases (e.g. week to week), episodes can now be released simultaneously (all-at-once), which lowers a consumers' viewing costs relative to the sequential releases but at the cost of diminished responsiveness to advertising. In this paper, we study the impact of the introduction of this new release timing decision in an AVOD setting with a signaling model. Specifically, we analyze whether the release timing of episodes signals show quality and moderates advertising levels. We show that by using an AVOD service that allows for sequential and simultaneous releases of show episodes, there exists a separating equilibrium under which higher and lower quality shows choose different release timing strategies. Furthermore, the introduction of the simultaneous release timing reduces the advertising level that high quality shows need to incur in order to signal their quality through the traditional sequential release schedule. Meanwhile low quality shows select an all-at-once release timing. Although high quality shows do not release episodes simultaneously, with a more profitable alternative for lower quality shows, they are better-off.

In the third chapter, we study how innovations in the home video viewing experience affect the optimal release timing strategies of the movie industry. We specifically analyze versioning through home video release windows - the time between theater market exit and home video release - when there are two home video products with different technological qualities; DVDs and Blu-rays. We develop a dynamic discrete choice model for action movies in theaters and home videos. This model connects theater and home video markets through the home video window, theatrical performance, and the discounted value of waiting. Our model differentiates from the current literature in that consumers are forward looking and may postpone their purchases for higher expected utility periods in the future. Furthermore, all markets are connected to each other, so changing the release date for one technological quality home video will have an impact on the box office revenue, which will impact all other technological quality home videos. We estimate the model parameters using panel data on the weekly level for home videos and box office. We conduct a counterfactual analysis in which the home video windows for DVDs and Blu-rays are jointly optimized. We find that decreasing the home video window increases the home video demand by increasing freshness and the advertising spillover from the theatrical run. However, this change can simultaneously cannibalize some demand from theaters as consumers have a larger home video discounted value. This cannibalization of theater demand has further impact on home video demand because the box office revenue - a driver for home video demand - is reduced. We find that an immediate after-theater release of lower technological quality home videos (DVDs) combined with a 5-week delay on higher technological quality home videos (Blu-rays) is optimal. We attribute this result to consumer heterogeneity, where there is greater substitution between higher technological quality home videos and theaters.

Acknowledgments

First and foremost, I would like to thank my advisor, Tim Derdenger. Thank you for your invaluable support and patience with me. I am extremely grateful for your mentorship and for all the time you put into supporting my ideas, presentations, and papers. I have learned so much from you, and I will never forget all the great memories of this PhD experience.

The work of this dissertation has been made possible with the help of several co-authors. Thank you Tim for being an integral part of Chapters 2 and 3. I thank Pascal Van Hentenryck, my co-author of Chapter 1, for all of his support and encouragement even before I started my PhD. I feel extremely fortunate I am your co-author, and I want to thank you for always responding when I needed your help. I would also like to thank Gerardo Berbeglia, my co-author of Chapter 1, for all his help and energy put into this chapter. I've had countless late-night discussions with Gerardo on problems and ideas related to this thesis. Many thanks to Joseph Xu, my co-author of Chapter 2, for all your feedback about my research as well as how to improve my presentation skills. I also extend my gratitude to Kannan Srinivasan, for all of his feedback and invaluable ideas in the work of Chapter 2. Finally, I would like to thank Sridhar Tayur, my co-author of Chapter 3, for all of his input on this chapter and on how to present my work. Thank you for always being there when I needed your help, and for involving me so much with your teaching when I was your teaching assistant.

I feel extremely fortunate that I had the support of The Initiative for Digital Entertainment Analytics (IDEA) center during my last year. I would like to thank the directors, Mike Smith and Rahul Telang, for all the opportunities they have given me, either by providing me with access to data or connecting me with executives in the industry pertaining this thesis. Thank you for being there when I needed your help and for your unwavering support.

I also thank other Tepper faculty and staff that helped me through my PhD. Thank you, Alan Scheller-Wolf, for all the feedback you have given me on how to present my work. I also extend my gratitude to Nicola Secomandi for all your guidance during the last six years. Thank you Soo-Haeng Cho, for your feedback on my presentations, and for giving me the opportunity to be your teaching assistant and learn from your lectures. I am immensely thankful to Laila and Lawrence, who made every single problem so much easier. Thank you both for all your help during my job search, for your guidance, and for always being there when I needed your help.

I would like to thank my thesis committee, Tim Derdenger, Param Vir Singh, Kannan Srinivasan, Sridhar Tayur, Pascal Van Hentenryck, and Joseph Xu, for putting the time to review my work and for their valuable feedback.

During my PhD I was fortunate to have wonderful colleagues. I am extremely grateful to my office mates, Amin, Arash, Mehmet, and Bo. I cherish all the great memories we have created during the last few years, and I will always be thankful for how much I learned from each you. I would like to also thank Neda and Musa, for all the great discussions and for making teaching assistant jobs so much more fun. I also extend my gratitude to the Tepper alumni; Francisco, Siddharth, and Aleksadr, for all their guidance during my PhD and the invaluable research discussions we've had. Finally, I want to thank the CMU INFORMS student chapter members, for all the seminars, discussions, and meetings that contributed to my academic training, special thanks to Christian, Dabeen, Florian, Gerdus, Nam, Melda, Ozgun, Sagnik, Thiago, Violet, and Yuyan.

I want to thank my Pittsburgh South American friends; Santiago, Francisco, David, Ricardo, and Camilo, that made me learn so much from South America from far away. I will never forget all the great memories we have created during our time in Pittsburgh. I also want to thank all the support that my friends from Argentina provided me when I needed them, moving to a new country is never easy and I am extremely grateful to them. Thank you so much Gus, Mario, Mono, Lucas, Alexis, Juan, Martín, Sebi, Tomi, and Nacho.

Finally, I want to thank my family, this thesis is dedicated to you. I am forever in debt to my parents, Alberto and Inés, that always encouraged me to study and did so many sacrifices for my siblings and me. I want to thank my sister, Yanina, for all her support during this journey. I would also like to thank my brother, Gerardo, for showing me that the academic path was possible, and for all his guidance throughout my academic career. I want to also extend my heartfelt gratitude to my grandmother, Ana, as well as to my brother-in-law, Diego, to my niece, Nina, and to my cousins, aunts, and uncles. Finally, I want to thank my partner, Katie, without your love, support, and help, this would have been so much harder. Thank you for being there for me during this journey.

Table of Contents

List of Figures

List of Tables

Introduction

Entertainment consumption has changed progressively over the last few decades due to technological developments. The rise of the internet has created marketplaces for digital products as well as different video-on-demand platforms. With every newly created medium of consumption a new business model arises with its own complications. This dissertation studies three different problems that are consequences of changes in media/digital product consumption.

Technology allowed for the creation of digital markets which generate several additional challenges in comparison to traditional physical markets. First, in traditional physical markets, product positioning was a simpler decision in which all consumers were exposed to products in the same way. With digital markets this exposure changes, as the platform owners can devise different product positioning for different types of consumers. Second, digital markets allow for the use of social signals, such as ratings and total product purchases, in order to persuade customers to buy products. The effects of social influence on consumer behaviour have been observed in a wide range of settings (Salganik et al. 2006, Tucker and Zhang 2011, Viglia et al. 2014). Moreover, these popularity signals can be amplified through product visibilities (Craswell et al. 2008). The combination of popularity signals and product visibilities have been extensively studied in the past few years, both theoretically and experimentally: See, for instance, Abeliuk et al. (2016), Krumme et al. (2012), Abeliuk et al. (2015), Maldonado et al. (2018), Van Hentenryck et al. (2016a).

In the first chapter of this dissertation, we analyze a trial offer market for digital products motivated by Google Play, Steam, ITunes and the Apple Store among others.[1]. We analyze the effect that social influence, product positioning, and market segmentation strategies have on the overall purchase rate for the market platform owner. We add consumer heterogeneity to a simple

[1]The applications of this chapter also transfers to other online marketplaces for physical products which have free returns, such as Amazon.com or Nike.com

Multinomial Logit (MNL) trial-offer market studied in the past, and we analyze whether the results obtained in this simple model transfer to when consumers exhibit heterogeneous preferences. Contrary to the homogeneous case, we show that finding the ranking that maximizes the probability that the next customer will make a purchase is NP-Hard under Turing reduction. Furthermore, we show that popularity signals may be detrimental to the total number of purchases. We finally show that these negative results can be addressed by a simple segmentation strategy where customers are shown a quality ranking dedicated to their own consumer segment and only observe the popularity signal for their own market segment.

The second and third chapters of this dissertation analyze the impact that technological innovations have on optimal release timing strategies of media in two different contexts. Technology generates new ways to release content, and for traditional TV companies this meant expanding to advertising video on demand platforms. These platforms allow for the simultaneous release of episodes, which is unprecedented for traditional TV, and generate new viewing habits such as binge-watching behavior. In the movie industry, the technological improvement of home viewing technology (TV screen sizes, resolution, home audio systems, and home video quality) has made the home viewing experience closer to that of theaters. This technological improvement affects the interaction patterns that consumers have between the theatrical and the home video markets, so aligning the release timing strategies of home videos with technology updates is critical.

The development of internet-based technologies allowed for traditional broadcast networks to shift the way in which they release episodes of shows. Now they can release their episodes through advertising video on demand (AVOD) platforms (e.g. Pluto TV, Peacock, and Tubi), which still generate revenue through commercials but encourage a different type of viewing behavior. Using these platforms, the companies can decide how to release their episodes of shows, whether its a simultaneous release on premiere date, or a sequential (e.g. weekly) release that has been traditional for TV. The new problem is that these two release timing strategies encourage different viewing habits which affect consumer responsiveness to advertising. Simultaneous releases encourage binge-watching behavior, and several studies suggest that advertising becomes less efficient when such behavior is established (Schweidel and Moe 2016). The trade-off these companies face is whether they should give more flexibility to viewers by releasing their episodes simultaneously, at the cost of a lower advertising revenue, or release their episodes sequentially giving less flexibility to viewers but generating a larger advertising revenue.

In Chapter 2, we analyze the strategic implications of the simultaneous release strategy within an AVOD setting. We study an analytical model that captures the interactions between a single show with multiple episodes, and consumers. We exploit the asymmetry of information about the quality of the show using a signaling model in which show quality may be of two possible types, high or lower quality. We illustrate that adequate levels of advertising alone[2] may signal quality, and for this to happen, the higher-quality show must incur a sizable cost. We then show that the release timing strategy may act as a signal for show quality. In equilibrium, lower quality shows select the simultaneous release strategy as it is more profitable than the sequential strategy. We then determine that the introduction of the simultaneous release timing reduces the advertising level that higher quality shows need to incur in order to signal their quality (compared to having sequential releases alone).

The time between theatrical and home video releases has been decreasing over the last few decades, while at the same time, the improvement of home video viewing technology makes the home video viewing experience closer to that of theaters. This poses the natural question on whether decreasing the time between theatrical and home video release is optimal for studios, as it is not clear whether these markets are substitutes or complements. In Chapter 3 of this dissertation, we study how to optimally release home videos of different technological qualities (DVDs and Blu-rays). We develop a dynamic structural model for the box office market and the home video market, consisting of two technological qualities, DVDs and Blu-rays[3]. We use the framework of Gowrisankaran and Rysman (2012) to build the infinite horizon home video market, while we make an unknown-finite horizon modification for the box office market. We connect both markets through the value of waiting and the theatrical revenue which is known to be a driver for home video demand. We estimate the model using panel data for a set of 149 action movies. After estimation is completed, the model is able to quantify the change in revenue associated with advancing or delaying home video releases in a series of counterfactuals.

We find that without versioning (releasing DVDs and Blu-rays simultaneously), the time between theatrical and home video releases should be decreased in order to achieve a 4.47% increase

[2] This advertising is different from the one that generates revenue, it informs viewers about the show.

[3] This can be seen as a proxy for Full HD and 4K videos today. It is simple to add other home videos as the data becomes available

in studio revenue, while having insignificant impact on theater revenue. When versioning becomes available, DVDs should be released within a week of theatrical exit, while Blu-rays should be release about 5 weeks after exit. This strategy increases studio revenue by about 5.4% compared to the data. The implication of this result is that higher technological quality home videos, such as Blu-rays, are stronger substitutes to theaters than lower technological quality home video, such as DVDs. So studios should release the lower technological quality home videos as soon as possible, since these don't have a substantial impact on box office sales. Meanwhile they should hold-off the release of higher technological quality home videos for later, since these will impact box office ticket sales.

In Chapter 4, we summarize the contributions of this thesis and suggest future research opportunities.

Chapter 1

Market Segmentation in Online Platforms

1.1 Introduction

The effects of social influence on consumer behaviour have been observed in a wide range of settings (Salganik et al. 2006, Tucker and Zhang 2011, Viglia et al. 2014). Depending on the market and/or the marketing platform, social influence may be induced by different signals, including the number of past purchases, consumer ratings, and consumer recommendations. Moreover, these popularity signals can be amplified through product visibilities. Indeed, in digital markets, the impact of visibility on consumer behavior has been widely observed, including in internet advertisement where sophisticated mathematical models have been developed to determine the relative importance of the various product positions (Craswell et al. 2008). Positioning effects are also of high significance in online stores such as Expedia, Amazon, and iTunes, as well as physical retail stores (see, e.g., Lim et al. (2004)).

The combination of popularity signals and product visibilities have been extensively studied in the past few years, both theoretically and experimentally: See, for instance, Abeliuk et al. (2016), Krumme et al. (2012), Abeliuk et al. (2015), Maldonado et al. (2018), Van Hentenryck et al. (2016a). This research stream considered trial-offer markets in which consumers have the opportunity to try products or services for free and later decide whether to purchase them. Prominent examples of online trial-offer markets include the following: music markets such as iTunes where users can listen to songs prior to the purchase decision, phone apps stores such as Google Play[1] or the Apple Store where "apps" have a free trial-version (limited by the functionality or with an expiration deadline), video on demand platforms such as Netflix, where consumer try episodes of series and finish watching them if they like them, and game platforms, such as Steam, where games may be returned for free within 2 hours of usage. Furthermore, other examples of trial-offer markets include online stores offering physical products that have free shipping (sometimes for purchases above a threshold) and free returns, such as Amazon.com or Nike.com. Previous studies have shown that social influence[2], in conjunction with the optimization of product visibilities, can have significant benefits on market efficiency. Moreover, simple ranking policies, e.g., giving the most visibility to products with the best estimated qualities, are needed to realize these benefits. These positive results however assume that customer preferences are homogeneous and can be modeled by a multinomial logit.

This chapter studies a more realistic model with heterogeneous customers that follow a latent class multinomial logit model (McFadden and Train 2000, Rusmevichientong et al. 2014) and attempts to understand whether the results obtained earlier in simple models would still hold under this new setting. It presents both negative and positive results. First, contrary to the homogeneous case, this chapter shows that, in mixed Multinomial Logit (MNL) trial-offer markets, computing the ranking of products that maximizes the probability that the next customer will make a purchase is NP-Hard under Turing reductions. Moreover, the chapter shows that popularity signals, that are beneficial in the homogeneous case, may become detrimental to market efficiency in mixed MNL trial-offer markets. However, this chapter also shows that these negative results can be addressed by a simple segmentation strategy where customers are shown a quality ranking dedicated to their own consumer segment and only observe the popularity signal for their own market segment (i.e., the past purchases of customers of the same segment). This market segmentation strategy can be implemented easily by collecting information on customers and/or by providing customers different rankings based on various demographic features. Indeed, a recent analysis performed by the online travel agent Orbitz has shown that Mac users spend up to about 30% more in hotel bookings than their PC counterparts (Mattioli 2012), suggesting that it is beneficial to show different

[1] Google Play's refund policy allows consumers to get full refunds on app purchases if the claim is made within 48 hours.

[2] Almost all examples of trial-offer markets show some form of popularity signal (rating and/or number of purchases) of its products when consumers browse through them.

Figure 1.1: Rankings Averaging Scores of a Consumer Segment at Booking.com.

rankings to customers depending on the computer they use. Moreover, the major online travel agent Booking.com allows users to rank hotels according to the average consumers score of a particular segment such as couples, families, and solo travelers. See Figure 1.1 for an illustration of this feature. Although hotels are not a trial-offer market, the benefits of market segmentation extend to trial-offer markets, as Netflix, Google Play and Steam, among others, implement targeting strategies, giving different product recommendations based on past consumer behavior (India 2019, Loten 2020).

The contributions of this chapter can be summarized as follows:

1. The chapter constructs the first trial-offer market model with social influence and position biases in which consumers preferences can follow any finite mixture of multinomial logits.

2. The chapter shows that, in mixed MNL trial-offer markets, computing the ranking of products that maximizes the probability that the next customer will make a purchase is NP-Hard under Turing reductions.

3. The chapter shows that the popularity signal may, under some circumstances, decrease the expected market efficiency. In other words, the display of past purchases may reduce the number of sales by confusing consumers about which products to try.

4. The chapter studies the average quality ranking, which ranks the items in decreasing order of average quality. It shows that the average quality ranking converges to a unique equilibrium when consumers are shown the number of past purchases (the popularity signal). This proof is rather involved as it requires to show that the mixed MNL model can be seen as a special MNL model in which some parameters (appeal and quality) are no longer constants but

functions of the past purchases vector, and that these quantities can be upper and lower bounded in order to demonstrate convergence.

5. The chapter presents a simple segmentation strategy, where customers are shown a quality ranking dedicated to their own segment and only observe the popularity signal for their own market segment (i.e., the past purchases of customers of the same segment). The chapter quantifies the potential benefits in market efficiency of this strategy. Specifically, it proves that the expected purchases can increase up to a factor of K, where K is the number of segments of the mixed MNL model.

6. These theoretical results are complemented by a series of computational experiments which provide several managerial insights about trial-offer markets.

The remaining of this chapter is organized as follows. Section 1.2 reviews the literature most related to this work. Section 1.3 introduces the model of the dynamic trial-offer market. The most relevant ranking policies for this model are described in Section 1.4, which also presents the NP-hardness results for performance ranking in mixed multinomial logit models. Section 1.4.2 describes the convergence and the impact of social influence for the quality ranking in the same setting. Section 1.5 presents our segmentation strategy and its benefits. Section 1.6 presents results of computational experiments and Section 1.7 concludes the chapter. The proofs are deferred to the Appendix A.1. In Appendix A.2, we consider an extension of the model in which the platform owner makes mistakes during the customer classification process. Appendix A.3 shows that a local search heuristic for solving the ranking optimization problem does not bring much benefit with respect to the much less costly average quality ranking policy studied in Section 1.4.2. Appendix A.4 analyzes an extension in which the firm's objective is to maximize the expected revenue instead of purchases. Finally, Appendix A.5 analyzes an extension in which the platform owner may show a subset of products to all consumer segments.

1.2 Related literature

This work is related to the MusicLab experiment performed by Salganik et al. (2006). In that experiment, participants were presented a list of unknown songs from unknown bands, each song being described by its name and band. The participants were partitioned into two groups exposed to two different experimental conditions: the *independent* condition and the *social influence* condition. In the independent group, participants were shown the songs in a random order and they were allowed to listen to each of them and then download them if they wish. In the second group (social influence condition), participants were shown the songs in popularity order, i.e., allocating the most popular songs to the most visible positions. Moreover, these participants were also shown a popularity signal, i.e., the number of times each song was downloaded too. In order to investigate the impact of social influence, participants in the second group were distributed in eight "worlds" evolving completely independently. In particular, participants in one world had no visibility about the downloads and the rankings in the other worlds. The MusicLab is an ideal experimental example of a trial-offer market where each song represents a product, and listening and downloading a song represent trying and purchasing a product respectively. The results by Salganik et al. (2006) show that the different worlds evolve significantly differently from one another, providing evidence that social influence may introduce unpredictability in a market.

To explain these results, Krumme et al. (2012) proposed a framework in which consumer choices are captured by a multinomial logit model whose product utilities depend on songs appeal, position bias, and social influence. Abeliuk et al. (2015) provided a theoretical and experimental analysis of such trial-offer markets using different ranking policies following the framework of Krumme et al. (2012). They proved that social influence is beneficial in order to maximize the expected number of purchases when using a greedy heuristic known as *performance ranking*. The *performance ranking* selects the ranking that maximizes the expected number of purchases at the next time period, i.e. it maximizes the short-term market efficiency. Abeliuk et al. (2015) have also illustrated experimentally that the popularity ranking [3] is outperformed by the performance ranking in a variety of settings. Still based on the model of Krumme et al. (2012), Van Hentenryck et al. (2016a) have studied the performance of the *quality ranking* which ranks products by their intrinsic quality (the quality of a product is here defined as the probability that a consumer would purchase/download the product once she has tried the product out). They show that the quality ranking is in fact asymptotically optimal and has a considerably less unpredictability than the popularity ranking.

Maldonado et al. (2018) studied the impact of the popularity signal strength on a market with multiple products and social influence. In their model, the popularity signal strength is a an exogenous parameter $r > 0$. The authors provide a complete characterization of the long-term

[3]The popularity ranking ranks (dynamically) products by the number of purchases in decreasing order.

market share of each of the products and show that the market is completely predictable as long as $r \leq 1$.

The relative importance of different popularity signals have been recently investigated by Engström and Forsell (2018) and Viglia et al. (2014). The first paper focuses on how consumers choose apps in the GOOGLE PLAY platform, and the second one studies how people select hotels. Both experiments arrived to the same conclusion, namely that the popularity signal (i.e., the number of purchases) has a much stronger impact on consumer behavior than the average consumer rating signal.

Our work is related to the recent paper by Hu et al. (2015) who consider a monopolist facing a newsvendor problem with two substitutable products with the same quality in which consumer preferences are affected by past purchases. The authors showed that the market is unpredictable but it can become less so if one of products has an initial sales advantage (such as for example by providing an initial discount). Our model has considerable differences including the incorporation of position biases, highly richer consumer preferences (Mixed MNL), an arbitrary number of products, and the allowance of products to have difference qualities.

Ghose et al. (2012) proposed a ranking system for hotels which takes into account the economic value of different locations and service-based characteristics, as well as consumer heterogeneity. In a follow-up paper, Ghose et al. (2014) studied the effects of three ranking policies on consumer behaviour using archival data analysis and randomized experiments. The general idea of both papers is to build a simultaneous equations model of clickthrough, conversion (purchase decision), ranking (performed by the platform owner), and customer rating. In their model, the demand for the different hotels (i.e., the products) is independent between the different hotels (apart from the fact that each hotel is assigned a different position in the ranking) whereas, the model we study in this chapter, the different products are in direct competition to attract the demand.

In another related paper, Gopal et al. (2016) perform a quantitative study on how a firms can strategically alter malleable networks such as enterprise social networks (ESN) or consumer social networks (CSN) in order to increase the transmission of ideas, innovation, or other information. In our setting, we focus on comparing a complete network (consumers observe all purchases) versus a network where consumers are partitioned in K classes or clusters (and a consumer only observes purchases of individuals from their own cluster).

Vaccari et al. (2018) studied a model consumers where arrive in sequence and estimate the quality of products based on product reviews (likes and dislikes). In their model, consumers like a product if the product's quality exceeds their expectation (which is calculated based on past ratings). The authors provide conditions that allow consumers to learn the true quality of products in the long run. Our work is also related to recent studies of theoretical choice models that incorporate position biases. As in our model, they consider situations in which the probability of selecting a product does not only depend on the offer set but also on the way products are displayed (Abeliuk et al. 2016, Aouad and Segev 2015, Davis et al. 2013, Gallego et al. 2020).

Recently, Golrezaei et al. (2018) considered a similar framework to ours in which the firm needs to decide how to rank a set of products to sell to consumers. Similar to our setting, sorting products by utility in their model is not always optimal [4]. Moreover, finding the optimal ranking of products to maximize short-term sales (or revenue) can be found in polynomial time when there is a single consumer type, but the problem becomes NP-hard when the number of consumer types is more than one. In their work, the authors constructed the choice probabilities from a two-stage consumer search model based on a seminal work by Weitzman (1979) on the Pandora's problem. As a result, the resulting mathematical expression for the choice probabilities is different from the one we study. Another difference is that Golrezaei et al. (2018) considers the problem of maximizing welfare (not studied here) whereas a fundamental focus of this work is on segmentation strategies which are presented in Section 1.5.

Another recent paper that models the dynamics of customers influenced by social influence is due by Chen et al. (2019). An important difference to our work is that social influence signal are product ratings rather than purchases, and that their choice model is based on the MNL model rather than on any finite mixture of MNLs.

Other papers related to our work are Lee and Eun (2020), who uses sales transaction data to estimate the parameters of a Mixed MNL through which are able to identify heterogeneous consumers groups, Zhen et al. (2019), Capuano et al. (2017), and Molinero et al. (2015) who studied social influence models, and Lutz and Newlands (2018) and Anderson and Xie (2014) who focused on market segmentation of customers.

[4] Although the underlying reasons for it under each of the models are different.

1.3 The Model

Motivation We consider a firm running a marketplace that sells a set of products. Following Krumme et al. (2012) and Salganik et al. (2006), we focus our attention to trial-offer markets, i.e., markets in which consumers can try the product for free before deciding to make a purchase. Consumers are position-biased in the following sense: the likelihood of trying a specific product is affected by the position of the product, as well as the position of the other products in the market. We also consider that in this marketplace it is possible to display information about product popularity. In particular, we assume that the firm shows the total number of purchases for each product at each point in time.

Unlike Krumme et al. (2012), we consider that there are different segments of consumers. More precisely, the probability that a given consumer tries a product follows a Mixed Multinomial Logit (MMNL). Since McFadden and Train (2000) proved that every random utility model can be well approximated by a MMNL, our model of consumer preferences is indeed very general.

Formalization We now formally describe a dynamic model for this marketplace. Let $[N] = \{1, 2, \dots, N\}$ denote the set of items in the marketplace and S_N denote the set of the permutations of these items. At any point in time, the firm decides how to position the items in the market by selecting a permutation $\sigma \in S_N$ such that $\sigma(i) = j$ implies that item i is placed in position j ($j \in [N]$).

The consumer behavior can be described as follows. There are K different segments of consumers. At any point in time the probability that the next arriving customer belongs to segment $k \in [K]$ is given by w_k, the segment's *weight*[5]. Consumers from the same segment exhibit different purchase profiles due to idiosyncratic shocks. This is captured with the MNL model.

When consumer t enters the market, she observes all the items and a popularity vector $d^t = (d_1^t, d_2^t, \dots, d_N^t) \in \mathbb{N}^N$, where d_i^t is the number of times item i has been bought prior to her arrival at time t. When the consumer arrives, each of the N items have been given a position through a permutation $\sigma \in S_N$. The consumer selects an item to try and then decides whether to buy it. Following Krumme et al. (2012) and extending their model to multiple segments, if the consumer

[5]A special case of this is a Poisson arrival process with arrival rate λ_k for each consumer segment $k \in [K]$, such that $w_k = \frac{\lambda_k}{\sum_j \lambda_j}$.

belongs to segment k, the probability that she tries item i is given by

$$p_{i,k}(\sigma, d^t) = \frac{v_{\sigma(i)}(a_{i,k} + d_i^t)}{\sum_{j=1}^{N} v_{\sigma(j)}(a_{j,k} + d_j^t) + z_k} \tag{1.1}$$

where $z_k \in \mathbb{R}_{\geq 0}$ for all $k \in [K]$ is fixed to a constant for the duration of the process, $v_j \in \mathbb{R}_{\geq 0}$ represents the visibility of position $j \in [N]$ regardless of the consumer class (the higher the value v_j the more visible the item in that position is), and $a_{i,k} \in \mathbb{R}_{>0}$ captures the intrinsic appeal of item i for consumer segment k for all $i \in [N]$ (higher values correspond to more appealing items)[6]. The value z_k represents the outside option for those consumers of segment k that enter the market. As the total number of purchases increases, the fraction of consumers choosing the outside option will decrease.

If a consumer from segment k has selected item i for a trial, the probability that she would purchase the item is given by $q_{i,k} \in [0, 1]$. Observe that this probability is independent of both the appeal vector $(a_{1,k}, \ldots, a_{N,k})$ and the visibility vector v. Intuitively, this assumption which has been validated in the MusicLab experiment, captures the fact that it is more difficult to influence consumers after they have tested a product than before.

When the consumer decides to purchase item i, the popularity/sales vector d is increased by one in position i. To analyze this process, we divide time into discrete periods such that each new period begins when a new consumer arrives. Hence, the length of each time period is not constant.

The objective of the firm running this market is to maximize the total expected number of purchases. To achieve this, the key managerial decision of the firm is what is known as the ranking policy (Abeliuk et al. 2015), which consists in deciding at each point in time the permutation $\sigma \in S_N$ to display the items. The next section describes a number of relevant ranking policies for this model.

A key aspect of this chapter is to study the potential benefits of the popularity signal in terms of the rate of purchases (market efficiency) and compare the ranking policies with and without this signal. In this work, we always assume that the popularity signal is used as specified in Equation (1.1). When the popularity signal is not used, the probability of trying (or sampling) a product is obtained as if the popularity signal is simply the vector $\langle 0, \ldots, 0 \rangle$.

[6]Note that the expression in Equation (1.1) is the special case of a MMNL when the consumer utilities are logarithmic.

We conclude this section with several comments about the model described above. First, under the trial-offer market model, a customer who tries, but does not like, a product simply walks away without trying any other product. An alternative model would be to allow the customer to try another product with probability c_i every time she tried but did not like a product i. Perhaps surprisingly, one can show that under mild conditions[7], the resulting model is equivalent to the original trial-offer market with a single trial per consumer Van Hentenryck et al. (2016b) [8]. Second, although there are many different ways to model consumer trial and purchase probabilities as a function of the previous purchases, intrinsic appeal, etc; the way we model the problem in this chapter is a natural extension to an MNL model that has been empirically tested (Krumme et al. 2012): it has succeeded to fit the data from a large scale randomized experiment where over $14,000$ users listened and downloaded songs from an assortment 48 song pieces. The participation in the experiment was unpaid and voluntary and therefore, the setup is arguably closer to a genuine internet platform than if individuals were paid to participate (see Salganik et al. (2006) for more details). Third, the model studied here could potentially be useful in markets that do not have the trial-offer structure. Indeed, this model can be captured as a function whose inputs are (1) market observations (current product ranking and past product purchases) and (2) some parameters subject to estimation (product appeals, product qualities, and position visibility values). The function then returns each product purchasing probability. Thus, even for non-trial-offer markets, it is an open empirical question on whether this model is better or worse than other choice models that also consider product rankings and/or popularity signals such as Golrezaei et al. (2018) and Vaccari et al. (2018), in terms of predictive accuracy of the purchasing probabilities. Finally, it is worth observing that our model assumes there is an independence between trying a product and buying it. Although this assumption was not problematic to fit the large-scale experiment carried out in Salganik et al. (2006), a model extension that incorporates a correlation between the two steps would be interesting.

[7] The probability c_i is a polynomial function on the product quality q_i.
[8] The equivalence is based on a redefinition of product appeal and quality.

1.4 Ranking Policies

Consider without loss of generality that the N locations are sorted by their visibility such that $v_1 \geq v_2 \geq \ldots \geq v_N$. A ranking policy is a function $f : \mathbb{N}^N \to S_N$ which, given a vector of past purchases, returns a ranking of the items.

Ranking policies can be partitioned into two groups: *static* and *dynamic*. A ranking policy g is said to be *static* if the output ranking does not depend on the popularity signal, i.e., if $f(d) = f(d')$ for all $d, d' \in \mathbb{N}^N$. On the other hand, a *dynamic* ranking policy is one in which the output ranking depends on this signal.

1.4.1 Performance ranking

The *performance ranking* is a dynamic policy that greedily selects a ranking that maximizes the expected number of purchases in the following period. This strategy was first proposed by Abeliuk et al. (2015) for the special case with $K = 1$ (where they show it is asymptotically optimal) and we now generalize its definition for the more general model considered in this chapter. The probability that the next incoming consumer belongs to segment k is given by w_k, therefore the performance ranking at time period t consists of finding the permutation $\sigma^* \in S_N$ maximizing the probability of a purchase in the next time period, i.e.,

$$\sigma^* = \underset{\sigma \in S_N}{\arg\max} \sum_{k=1}^{K} w_k \cdot \sum_{i=1}^{N} p_{i,k}(\sigma, d^t) \cdot q_{i,k}. \tag{1.2}$$

The probability Π^{PR} of a purchase in the next time period is thus given by

$$\Pi^{PR} = \max_{\sigma \in S_N} \left\{ \sum_{k=1}^{K} \Big(w_k \cdot \sum_{i=1}^{N} (p_{i,k}(\sigma, d^t) \cdot q_{i,k}) \Big) \right\} \tag{1.3}$$

$$= \max_{\sigma \in S_N} \left\{ \sum_{k=1}^{K} \Big(w_k \cdot \sum_{i=1}^{N} \Big(\frac{v_{\sigma(i)}(a_{i,k} + d_i^t)}{\sum_{j=1}^{N} v_{\sigma(j)}(a_{j,k} + d_j^t) + z_k} \cdot q_{i,k} \Big) \Big) \right\}. \tag{1.4}$$

Abeliuk et al. (2015) showed that when $K = 1$, this greedy ranking policy can be computed efficiently, i.e., in strongly polynomial time (Theorem 1 in Abeliuk et al. (2015), they assumed $z = 0$ but their proof can be easily generalized for any $z \in \mathbb{R}_{\geq 0}$). Moreover, despite the myopic focus of the performance ranking, a series of computational experiments performed by Abeliuk et al. (2015) showed that, for the special case of $K = 1$, the performance ranking was superior than the standard popularity ranking both in terms of unpredictability as well as in terms of number of purchases. Unfortunately, the performance ranking cannot be computed efficiently when there are at least two classes of consumers. More precisely, we can show that the assortment problem under a 2-segment Mixed Multinomial Logit choice model which is known to be NP-hard (Rusmevichientong et al. 2014) can be reduced (under Turing reductions) to computing the performance ranking in our setting.

Theorem 1.1. *Computing the performance ranking is NP-hard under Turing reductions. This is true even when $K = 2$ and the product qualities are the same for all consumer classes.*

Proof. Presented in Appendix A.1.1

In order to deal with this negative result, this chapter explores two options. The first is to rank products based on their average quality, a strategy we called *average quality ranking* and is studied in the next section. The second avenue is to explore a greedy local search heuristic which worked as follows. In the first period, we set the average quality ranking as initial solution, and then, we evaluate possible ways to exchange the position of two products in the ranking (2-swaps) and select the first swap that improves the objective function. The local search process is then repeated until there are no more improvements. In each consecutive period, we use the previous period ranking as the starting solution. Notice that this heuristic will finish in a finite number of steps. This is because the objective function increases with each swap and there exist finitely many assortments of products. Our experimental results have shown that the additional benefit of the local search is minimal (see Appendix A.3), we thus used the average quality ranking as our approximate method which is several orders of magnitude faster.

1.4.2 Average quality ranking

The *average quality ranking* (or quality ranking for short) is a simple and natural static policy: it consists in ranking the products by the weighted average quality (among the different customer segments), ignoring the appeals and popularity signal. The quality ranking for the MMNL model thus consists in placing in position j the item with the j^{th} highest weighted average quality, where the weighted average quality of item $i \in [N]$ is

$$\bar{q}_i = \sum_{k=1}^{K} w_k q_{i,k}. \tag{1.5}$$

For the special case $K = 1$, the quality ranking is optimal asymptotically and always benefits from the popularity signal used in our model (Van Hentenryck et al. 2016a). The next section studies whether this continues to hold in richer contexts when $K > 1$. Note that, in the following, the ranking which orders the products by decreasing values of $\bar{q}_i$ is called the *average quality ranking.*

Before analyzing some fundamental properties of the MMNL it is important to first make the following definition.

Definition 1.1. *The MMNL model* goes to a monopoly *using a ranking policy f if, for each realization of the N random sequences $\{\phi_i^t\}_{t\in\mathbb{N}}$ ($i \in [N]$), there exists a product i^* such that the realized sequence $\{\phi_{i^*}^t\}_{t\in\mathbb{N}}$ converges almost surely to 1 as t goes to infinity. In this case, we also say that item i^** goes (predictably) to a monopoly*.*

We now study some fundamental properties of the quality ranking for the MMNL model. We first show that the average quality ranking converges to a monopoly (under weak conditions). We then study the benefits of displaying the popularity signal for the average quality ranking.

Given a ranking policy f, the random variable

$$\phi_i^t = \frac{d_i^t}{\sum_j d_j^t}$$

is known as the market share of item i at time t: It represents the ratio between the number of times that item i was purchased and the total number of purchases up to time t.

We can now show that the MMNL model goes to a monopoly when using the average quality ranking. The proof is quite technical: Its key idea is to show that the Mixed Multinomial Logit Model (MMNL) can be reduced to a generalized case of the Multinomial Logit Model (MNL) where the appeal and the quality of an item at time t depend on the popularity signal at t (This is shown in Theorem 1.2). Then these generalized appeals and qualities can be bounded to obtain the result. The proof relies on the following lemma that generalizes the convergence result of the quality ranking for the MNL model by Van Hentenryck et al. (2016a) to the case where the appeal and quality of an item depend on the popularity ranking provided that the resulting functions are bounded by above and below. In Theorem 1.2 we show that there exists a time period t^* after which the time dependent qualities and appeals of this modified MNL model are bounded as in 1.6 and that these bounds satisfy 1.7. Thus, by using Lemma 1.1, we can show that this modified MNL model goes to a monopoly which implies that the MMNL goes to a monopoly.

Lemma 1.1. *Consider a different Multinomial Logit Model, i.e., a setting with $K = 1$ where the appeal and quality of each item i are functions of the purchases vector d^t, i.e., $\widetilde{a}_i^t = \widetilde{a}_i(d^t)$ and $\widetilde{q}_i^t = \widetilde{q}_i(d^t)$ respectively. Suppose that there exists a time period t^* such that these two quantities are upper and lower bounded by constants for any period $t > t^*$ independent of the realizations of $\widetilde{a}_i^t$ and $\widetilde{q}_i^t$, i.e.,*

$$q_{i,min} \leq \widetilde{q}_i^t \leq q_{i,max} \text{ and } a_{i,min} \leq \widetilde{a}_i^t \leq a_{i,max} \quad \forall i \in [N], t > t^*. \tag{1.6}$$

Let $\sigma \in S_N$ denote a static ranking policy. If there exists an item i^* and an instant $\hat{t}$ such that

$$v_{\sigma(i^*)} q_{i^*,min} > v_{\sigma(i)} q_{i,max} \ \forall \ i \neq i^* \text{ and } \forall \ t > \hat{t}, \tag{1.7}$$

then item i^* goes to a monopoly when using the ranking policy σ.

Proof. Presented in Appendix A.1.2

The main result of this section is about the convergence to a monopoly of a large class of static ranking policies. For simplicity, we assume a weak condition to break potential ties between items.

Definition 1.2. *A static ranking policy σ is* tie-breaking *for a MMNL model if there exists a unique item i^* with the highest product of visibility and weighted average quality, i.e.,*

$$\exists \ i^* \in [N] : \bar{q}_{i^*} v_{\sigma(i^*)} > \bar{q}_i v_{\sigma(i)} \quad \forall i \in [N], i \neq i^*. \tag{1.8}$$

Note that this tie-breaking property is a very mild assumption: If the quality for each pair (consumer class, product) is a realization from a (different or the same) continuous probability distribution over some interval (regardless of how small), the probability that a ranking policy is tie-breaking is indeed 1. We are now ready to prove the main result of this section.

Theorem 1.2. *Consider a MMNL model $\mathcal{M}$ and a static, tie-breaking ranking policy $\sigma \in S_N$ for $\mathcal{M}$. Model $\mathcal{M}$ goes to a monopoly using σ and the item i^* that goes predictably to a monopoly using σ in $\mathcal{M}$ is given by*

$$i^* = \arg\max_{1 \le i \le N} v_{\sigma(i)} \bar{q}_i.$$

Proof. Presented in Appendix A.1.3

The following corollary asserts that the average quality ranking converges to a monopoly for the product of highest average quality. This result is particularly interesting since it shows that the average quality ranking is asymptotically optimal, and that it generalizes the quality ranking from the MNL to the MMNL model. By asymptotically optimal we mean that it is the global ranking with global social influence that maximizes the purchase probability as time goes to infinity, $\lim_{t \to \infty} P^t$, where P^t is the expected purchase probability at period t.

Corollary 1.1. *(Asymptotic Optimality of the Average Quality Ranking). Whenever the average quality ranking is used, a MMNL model goes to a monopoly for the product with the highest weighted average quality.*

Proof. Presented in Appendix A.1.4

We end this section with some comments about the implications of Theorem 1.2 and Corollary 1.1. In general, market monopolies are not a desirable outcome from a consumer perspective. A key aspect of Theorem 1.2 is that the monopoly outcome does not come as a result of a restriction on the product offer variety (every product is always offered in our model), but because the consumer's utility ratio between the most popular product and the other product utilities tends to infinity. Although such monopoly convergence is somewhat surprising, it naturally applies to the very long run dynamics of these trial-offer markets (i.e., when time tends to infinity). In practice, however, it takes a very long time to even get close to a monopoly. It is not simple to find the convergence rate to a monopoly, even for the case $K = 1$. This model can be seen as an unbalanced and irreducible Pólya Urn Process for which the convergence rate is an open problem except for a few special cases of replacement matrices (see remark 4.7 in Janson (2004), where they conjecture a convergence rate of $o(n^\gamma)$ with $\gamma = \min\{1/2, 3(1/2 - Re(\lambda_2/\lambda_1))\}$, with λ_1 and λ_2 being the greatest and second greatest eigenvalues of the replacement matrix respectively). As an illustration of the dynamics, we have performed a variety of computational experiments which are reported in Section 1.6. In all those experiments, the convergence to a monopoly of a single product was not yet achieved at the end of the simulation. From a practical point of view, the main insight obtained from Theorem 1.2 and Corollary 1.1 is that, as time goes by, consumers are more likely to purchase the product that has the greatest average quality, which is due to the effect of the popularity signal in the market dynamics. An interesting question that arises is whether the effect of the popularity signal is beneficial to the firm. Specifically, what is the effect of the popularity signal on the expected rate of products purchased? This question is addressed in the following subsection.

The Impact of the Popularity Signal

In the previous subsection, we have shown that the average quality ranking for the MMNL model inherits the asymptotic convergence of the quality ranking for the MNL. Under the MNL model, the probability (in expectation) that the next individual purchases some product is always increasing if the popularity signal is used (Van Hentenryck et al. 2016a). In other words, the information of past purchases is helping consumers to make better choices on which products to try, meaning that they are more likely to purchase them. Unfortunately, this result does not always hold when consumers follow the more general MMNL model.

Theorem 1.3. *When using the average quality ranking, the MMNL model can perform (1) up to K times better if the popularity signal is not shown, where K is the number of classes; and (2) can perform arbitrarily worse without showing the popularity signal than by showing it.*

Proof. Presented in Appendix A.1.5

Theorem 1.3 provides an upper bound on how much the expected sales per period can be reduced due to the impact of the popularity signal (part (1)). In addition, it shows that sometimes, showing the social influence can be extremely beneficial (part (2)). For the latter, it is important to note that not showing the popularity signal can perform arbitrarily worse than showing it even when $z = 0$. This implies that this reduction is not caused by having consumers switching from the outside option to other products, but because consumers under social influence can sometimes try popular products that have low quality for them (relatively to other products). In Proposition 1.1 we show that the bounds in Theorem 3 are tight, and this happens when we choose $z = 0$ (however, for different values of z, not showing the popularity signal can still perform arbitrarily worse than by showing it).

Proposition 1.1. *The bounds in Theorem 1.3 are tight.*

Proof. Presented in Appendix A.1.6

This result shows that, depending on the specific parameters of the trial offer market (i.e. consumer's appeals, product qualities and weights of the different consumer segments), the popularity signal can enhance the number of sales or it can be detrimental to them. Specifically, the expected number of purchases under the average quality ranking policy can decrease by a factor of K if the popularity signal is used (for some settings) but they can also be increased by an arbitrarily large factor in other settings. In Proposition 1.1, the first tight bound occurs in a trial-offer market where each of the K different classes of consumers have a unique product in which they are interested; Moreover, this product is different for each consumer segment and there is a perfect alignment between product appeal and product quality for each class. In such a setting, the global popularity signal would be detrimental for the firm as well as for the consumers. The reason is that in the long run product 1 will become a monopoly (since its average quality ranking is higher than all the others). This means that in the long run, consumers will tend to try product 1. But this is problematic because, except for one consumer segment, all consumers segments do not like this product. In the limit, only a $1/K$ fraction of the consumers would purchase this product. In short, in this scenario a global popularity signal will persuade consumers to try products they do not like.

On the other hand, the second bound in Proposition 1.1 was obtained in a similar market setting but in which there is negative correlation between the products appeal and their quality. In that setting, the market converges slowly to a monopoly in the case the popularity signal is used. If no social influence is used, the rate of purchases tends to zero as the products have essentially no appeal. Thus, an heterogeneous set of customers complicates the managerial decisions in the marketplace: Whether or not social influence is beneficial in trial-offer markets depends on the particular structure of preferences among consumers. We recall that this is in sharp contrast to the results in (Van Hentenryck et al. 2016a) where it is shown that the popularity signal is always beneficial in the case consumers share preferences (i.e $K = 1$). We now quantify the benefit (or cost) of showing the popularity signal given the specific model parameters. The asymptotic purchase probability ratio between the average quality ranking with no social influence (AQNSI) and the average quality ranking with social influence (AQGSI) is (see the proof of Theorem 1.3):

$$\frac{\lim_{t\to\infty} P^t_{AQNSI}}{\lim_{t\to\infty} P^t_{AQGSI}} = \frac{\sum_{k=1}^{K} w_k \sum_{i=1}^{N} q_{i,k} \frac{v_{\sigma(i)} a_{i,k}}{\sum_{j=1}^{N} v_{\sigma(j)} a_{j,k} + z_k}}{\max_{1\le i\le N} \sum_{k=1}^{K} w_k q_{i,k}}. \tag{1.9}$$

If the expression in 1.9 is smaller (greater) than 1 showing the popularity signal under the average quality ranking is beneficial (detrimental). Proposition 1.1 shows that under different model parameters, the ratio in Equation (1.9) could be equal to zero, K, or any number between them. This is important for the platform owner and has to be decided particularly for each market settings, since social influence may hurt or benefit purchases in the long run.

1.5 Market Segmentation and its Benefits

In the previous section, we have shown a number of negative results for the MMNL model. In particular, we have shown that, in MMNL models, computing the performance ranking is intractable and that displaying the popularity signal to customers may significantly reduce the asymptotic market efficiency (i.e. the expected rate of purchases) of the average quality ranking. In this section, we show that the widely used marketing strategy known as *market segmentation* remedies these limitations, while retaining the original benefits of quality ranking for the Multinomial Logit Model.

The market segmentation considered here assumes that the firm has the ability to know the segment of each arriving consumer. This is a natural assumption in a number of online markets (e.g., Amazon, online retail stores, iTunes, Google Play and Netflix) where firms are able to learn information about their customers over time. Armed with this information, the firm will now propose item rankings dedicated to each customer segment. Moreover, and equally important, the popularity signal will be tailored to each segment. In other words, *the firm will only show the popularity signal derived from purchases of customers of the same segment as the incoming customer*, not the popularity obtained from the entire customer pool. As shown in Figure 1.1, websites such as `Booking.com` already give customers the option of selecting their peer groups to refine the site recommendations (although hotels bookings are not a trial-offer market, segmentation is an important factor in their revenue maximizing strategies). Under this new strategy where each consumer segment has its own quality ranking and observes the past purchases of its own segment only, the policy is called the *segmented quality ranking*. The firm uses K permutations $\sigma_k \in S_N (k \in [K])$, where σ_k sorts the products in decreasing order according to their quality for consumer segment k. In addition, the probability of trying item i for a customer of segment k is given by

$$p_{i,k}(\sigma, d_k^t)$$

where $d_k^t = (d_{1,k}^t, \ldots, d_{N,k}^t)$ and $d_{i,k}^t$ denotes the number of purchases of item i by customers from segment k up to time t.

We now study the benefits of this market segmentation. Observe first that each market segment can be viewed as evolving independently and hence directly inherits the original benefits identified for the quality ranking under the MNL model: The market share of the highest quality product converges asymptotically to 1. This observation will enable us to quantify the benefits of market segmentation. We begin by providing some key definitions that will be required later.

Definition 1.3. *The segmented quality ranking policy $\sigma_k \in S_N$ $(k \in [K])$ is* tie-breaking *if, for each segment k, there exists a unique item i_k^* with the highest quality:*

$$\forall k \in [K]\ \exists i_k^* \in [N]\ \forall j \in [N], j \neq i_k^* :\ q_{i_k^*,k} > q_{j,k}. \tag{1.10}$$

Assuming a MMNL model for which the average quality ranking and the segmented quality ranking are tie-breaking, we compare the probability of a purchase at time t in both settings. More precisely, we compare two quantities:

- P_{AQGSI}^t: the probability of a purchase at time t when the firm uses the average quality ranking and the "global" popularity signal d^t;
- P_{SQSSI}^t: the probability of a purchase at time t when the firm uses the segmented quality ranking with the consumer segment's popularity signal d_k^t.

The probabilities P_{AQGSI}^t and P_{SQSSI}^t concern the behavior of the consumer arriving to the market at time t independently from the customer class. Comparing P_{AQGSI}^t and P_{SQSSI}^t for any time t is a very challenging task. Instead, we compare both variables in the limit. The following theorem shows that the market segmentation strategy is always beneficial for the firm and the benefit it provides is upper bounded by a factor of K.

Theorem 1.4. *Assume that the average quality ranking and its segmented version are tie-breaking for a MMNL model. Then,*

$$1 \leq \lim_{t \to \infty} \frac{P^t_{SQSSI}}{P^t_{AQGSI}} \leq K. \tag{1.11}$$

Proof. Presented in Appendix A.1.7

The following proposition shows that there are settings in which the (multiplicative) benefits of market segmentation are indeed equal to the upper bound provided in Theorem 1.4.

Proposition 1.2. *The upper bound of Theorem 1.4 is tight.*

Proof. Presented in Appendix A.1.8

These results show that the segmented quality ranking always outperforms (or it is equal to) the average quality ranking in expectation, and that the improvement in market efficiency can be up to a factor of K. As it can be seen from the proof of Proposition 1.2, the market settings in which the segmentation strategy beats by the most the global ranking is when each consumer segment $k \in \{1, \ldots, K\}$ of consumers have a single product with a non-zero quality and it is pairwise different between any two classes. Thus, the different segments of the market have very distinct preferences, which is the setting where the strategy of market segmentation is generally used in practice (Dickson and Ginter 1987). Proposition 1.2 means that there are settings under which a segmented ranking performs the same, or up to K times better than a single ranking. This imposes a maximum cost rate to what the platform owner may be willing to pay for information about consumer segments, and deploying a segmented ranking. It is important to remark that these results hold for the very long run (i.e. asymptotic in nature). They don't necessarily imply that the segmented quality ranking is always better, since the popularity signal is weaker early in the market evolution. This will be illustrated in the computational experiments presented in the next section.

It is important to remark that the results of this section can be extended to the case where there are classification mistakes while performing the segmentation policy. This situation is analyzed in Appendix A.2. We formulate the problem using a mistake probability matrix and show that the system converges to a monopoly, analogous to Theorem 1.2 but incorporating the values of the error probability matrix. Assuming that classification mistakes occur with an exogenous probability β_0 evenly distributed among all products, we find an analogous bound as Theorem 1.4, which becomes $K(1-\beta_0)$. For an example, if we make classification errors 10% of the time with 5 consumer classes of equal weights, the maximum benefit of segmentation is $5(1-0.1) = 4.5$, while its maximum benefit is 5 with perfect classification. This extension generalizes our model to a more realistic setting where we can analyze the trade-off between segmenting the market and taking the risk of making classification errors versus being risk averse and showing an average quality ranking. We performed a numerical simulation to analyze the impact of having different mistake probabilities for the SQSSI policy and we find that the average quality ranking AQGSI outperforms the SQSSI policy when $\beta_0 > 10\%$ for a parameter set (see Figure A.1 in Appendix A.2).

We end this section with two additional comparisons. First, we compare how the average quality ranking without social influence (AQNSI) performs in comparison to the segmented quality segmented social influence ranking (SQSSI). It is no surprise that SQSSI outperforms AQNSI as the number of purchases goes to infinity: this is because under SQSSI, in the limit, each consumer segment will try the product that maximizes its purchase probability. Finally, we analyze how SQSSI performs in comparison to its counterpart without social influence (SQNSI). Again here, it should be no surprise that SQSSI outpeforms SQNSI. Both of these results are shown in the following theorem.

Theorem 1.5. *The ratio between the asymptotic purchase probability of the average quality or the segmented quality ranking without social influence (AQNSI or SQNSI), and its segmented quality ranking with segmented social influence (SQSSI) is always less than 1,*

$$\frac{\lim_{t\to\infty} P^t_{AQNSI}}{\lim_{t\to\infty} P^t_{SQSSI}} \leq 1 \textit{ and } \frac{\lim_{t\to\infty} P^t_{SQNSI}}{\lim_{t\to\infty} P^t_{SQSSI}} \leq 1.$$

Proof. Presented in Appendix A.1.9

We have shown that SQSSI outperforms AQGSI, AQNSI and SQNSI, and that AQGSI can outperform or underperform AQNSI. Figure 1.2 shows a pictogram with the ranking policies studied in this chapter.

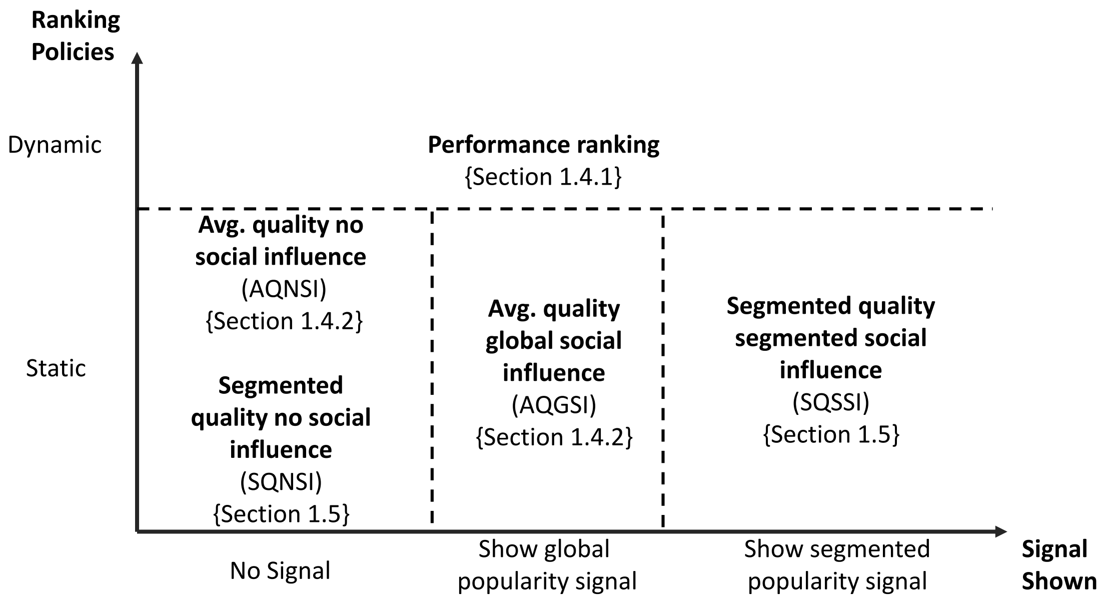

Figure 1.2: Pictogram of the ranking policies studied and where to find them.

1.6 Computational Experiments

This section presents the results of computational experiments to illustrate the theoretical results and complement them by depicting how the markets evolve over time for different types of rankings.

1.6.1 The Experimental Setting

The Agent-Based Simulation The experimental setting uses an agent-based simulation to emulate the MusicLab (Salganik et al. 2006). It generalizes prior results which simulated the MusicLab through the use of a MNL model (e.g., Krumme et al. (2012), Abeliuk et al. (2015)) to a MMNL model. Each simulation consists of T iterations and, at each iteration t ($1 \leq t \leq T$), the simulator

1. randomly selects a customer segment k according to the classes weights w_k;
2. randomly selects an item i for the incoming customer according to the probabilities $p_{i,k}(\sigma, d)$, where σ is the ranking proposed by the policy under evaluation and d is the popularity signal;
3. randomly determines, with probability $q_{i,k}$, whether the selected item i is purchased. In the case of a purchase, the simulator increases the popularity signal for item i, i.e., $d_{i,t+1} = d_{i,t}+1$. Otherwise, $d_{i,t+1} = d_{i,t}$.

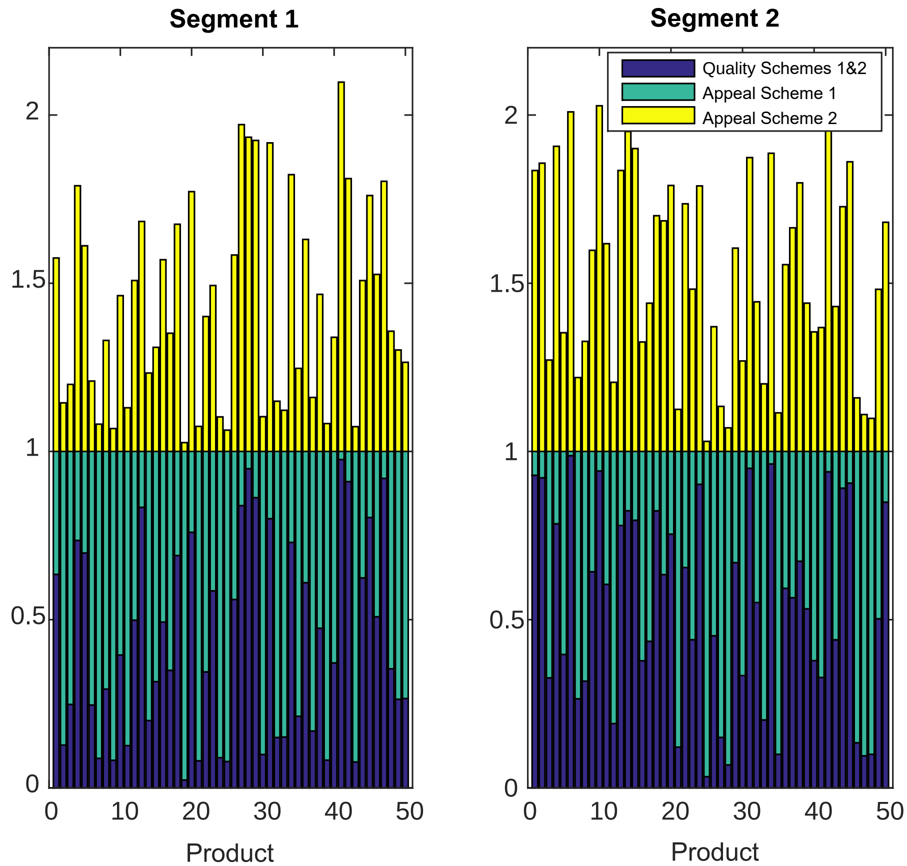

Figure 1.3: The quality q_i (blue) and appeal A_i (green and yellow) of product i for settings 1 and 2 for both classes of consumers. The settings only differ in the appeal of items, and not in the quality of items. In Setting 1, the appeals are negatively correlated with quality, so that the sum between them is always 1. In Setting 2, the appeal is correlated to the quality with a small noise.

The experimental setting aims at being close to the MusicLab experiments and it considers 50 items and simulations with $T = 200,000$ steps. The reported results in the graphs are the average of 400 simulations. The analysis in Krumme et al. (2012) indicated that participants are more likely to sample products with better ranking positions. More precisely, the visibility decreases with the ranking position, except for a slight increase at the bottom positions. To have a fair comparison between the settings with and without social influence we set $z = 0$ for all simulations, in that way the fraction of customers choosing the outside option does not change over time.

Qualities and Appeals To highlight and complement the theoretical results, we consider four different schemes, the schemes share the following characteristics: They have two customer classes with the same weight and they use 50 products. They differ in how the values for the item appeals

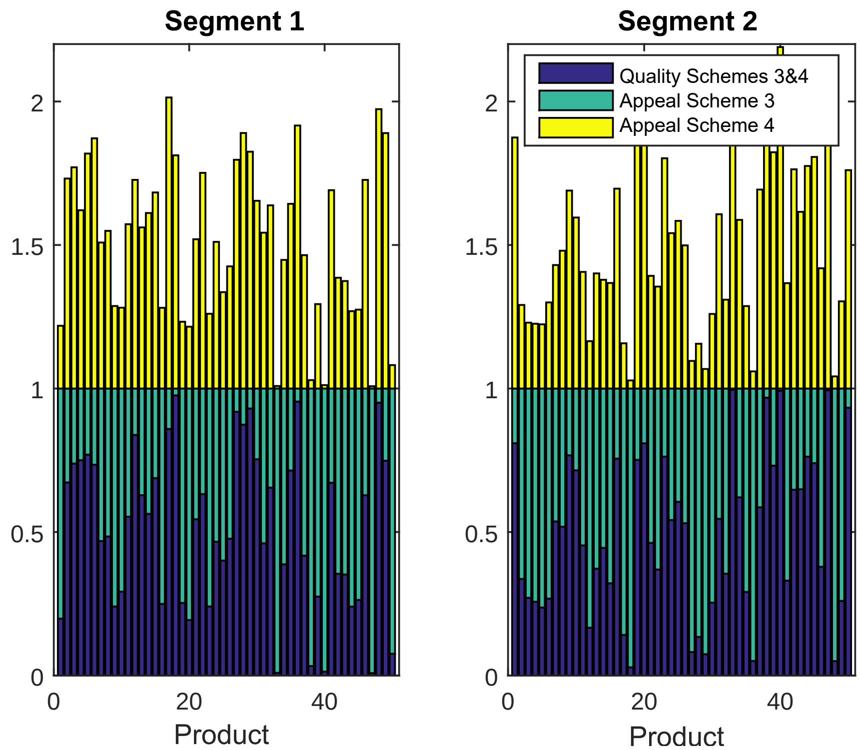

Figure 1.4: The quality q_i (blue) and appeal A_i (green and yellow) of product i in settings 3 and 4 for both classes of consumers. The settings only differ in the appeal of items, and not in the quality of items. In Setting 3, the appeals are negatively correlated with quality, so that the sum between them is always 1. In Setting 4, the appeal is correlated to the quality with a small noise. All appeals were then multiplied by a factor of 200 to use in the model.

and qualities are chosen. The schemes are depicted visually in Figures 1.3 and 1.4 and were obtained as follows:

1. Scheme 1: The product qualities for each consumer segment were chosen randomly with a standard uniform distribution ($q_{i,1}$ and $q_{i,2}$ are independent for all $i \in [N]$). Appeals were negatively correlated with quality, i.e., $a_{i,k} = 1 - q_{i,k}$ for all $i \in [N]$.

2. Scheme 2: Product qualities are similar to Scheme 1. Appeal vectors are now correlated with the quality vectors. More precisely, the appeal vector for each consumer segment was set to 0.8 times the quality plus a random uniform vector between -0.4 and 0.4, i.e., $a_{i,k} = q_{i,k}(0.8 + 0.4 * \epsilon_i)$ for all $i \in [N]$, where ϵ_i is a standard uniform random variable.

3. Scheme 3: The product quality for segment 1 is a random vector, while $q_{i,2} = 1 - q_{i,1} + 0.01 * \epsilon_i$

for all $i \in [N]$, where ϵ_i is a standard uniform random variable. Appeals are negatively correlated with quality, i.e., $a_{i,k} = 1 - q_{i,k}$ for all $i \in [N]$.

4. Scheme 4: The product qualities are the same as in Scheme 3 but the appeals are correlated with qualities, $a_{i,k} = q_{i,k}(0.8 + 0.4\ rand(1, 50))$ for all $i \in [N]$.

Observe that, in Schemes 3 and 4, customers in the two classes associate fundamentally different qualities with the products. For the simulations, the appeals vector were multiplied by a factor of 200.

The Policies The simulations compare the average and segmented quality rankings with and without the popularity signal. We use the following notations:

- `SQSSI`: Segmented quality ranking with segmented popularity signal;
- `SQNSI`: Segmented quality ranking without popularity signal;
- `AQGSI`: Average quality ranking with global popularity signal;
- `AQNSI`: Average quality ranking without popularity signal.

1.6.2 Market Efficiency

Figure 1.5 depicts the results for Schemes 1 and 2. For Scheme 1, the popularity signal is beneficial for both the segmented and average quality rankings. SQSSI is the most efficient ranking policy. It is also interesting to observe that AQGSI outperforms SQSSI early on before being overtaken as highlighted in Figure 1.6. Scheme 2 exhibits similar results but the benefit of the popularity signal is lower.

Figure 1.7 depicts the results for Schemes 3 and 4 and they are particularly interesting. Recall that, in Schemes 3 and 4, the two classes of customers have opposite preferences in terms of product qualities. For Scheme 3, the popularity signal is again beneficial for the segmented and average quality rankings. SQSSI is again the best ranking policy, and particularly is almost twice as efficient than AQGSI, nicely illustrating Theorem 1.4, since the improvement is close to the best possible ratio. Once again, AQNSI performs the worst. For Scheme 4, SQSSI is again the best ranking policy but the second best policy is SQNSI, the segmented quality ranking with no popularity signal. The

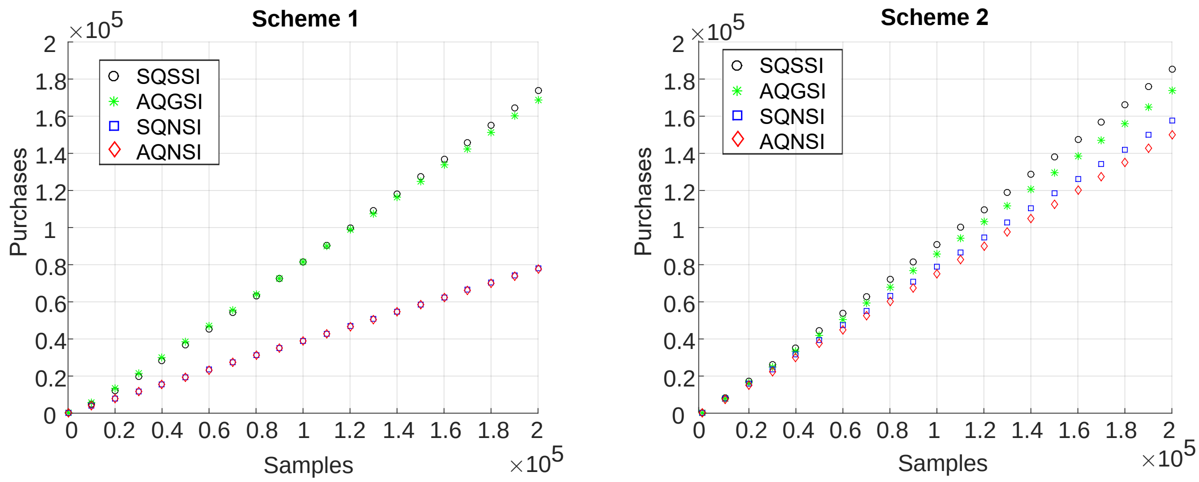

Figure 1.5: The Number of Purchases over Time for the Various Rankings. The x-axis represents the number of items tried and the y-axis represents the average number of purchases over all experiments. The left figure depicts the results for Scheme 1 and the right figure for Scheme 2.

worst policy is AQGSI, providing a compelling illustration of Theorem 1.3: The popularity signal may be detrimental to the average quality ranking.

These results can be summarized as follows:

1. SQSSI (segmentation with the popularity signal) is clearly the best policy and it dominates all other policies. Market segmentation with the popularity signal is very effective in these trial-offer markets.

2. The global popularity signal may be beneficial or detrimental to the average quality ranking. It is detrimental when the market has customers with very different product preferences.

1.6.3 Purchase Profiles

We now illustrate the customer and market behaviors for the SQSSI and AQGSI rankings, which exhibit some significant differences. For Scheme 4, the results are presented in Figures 1.8, 1.9, and 1.10. Figure 1.8 depicts separately the purchase profiles of customers of segments 1 and 2 for policy SQSSI. The products are sorted by increasing quality for each segment: i.e., the products of highest quality for customers of segment 1 (resp. segment 2) is in the rightmost position in the left (resp. right) picture. Since the market is segmented, the results are not surprising and consistent with past results: The number of purchases is strongly correlated with quality. Figure 1.9 is more

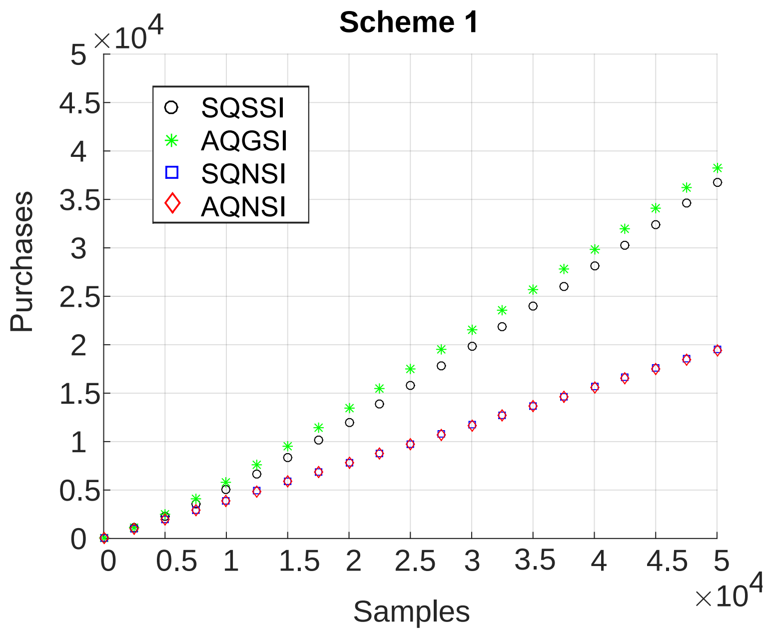

Figure 1.6: The number of purchases over time for the various rankings. The x-axis represents the number of items tried and the y-axis represents the average number of purchases over all experiments. The figure depicts the results for Scheme 1 in the early part of the simulation.

interesting and depicts the same information for policy AQGSI. Here the number of purchases is no longer correlated with quality for a specific customer segment. Figure 1.10 compares SQSSI and AQGSI over all customers and the products are sorted by average quality. The figure highlights a fundamental difference in market behavior between the two policies, with very different products emerging as the "best sellers".

Schemes 3 and 4 feature customer classes with opposite preferences. It is thus interesting to report the results on Scheme 2 where the product qualities were generated independently for the two classes. Figure 1.11 depicts these results. We already know from Figure 1.5 that policy SQSSI outperforms AQGSI but it is interesting to see how different the market behaves under these two policies. For AQGSI, as expected, the products of best average quality receives the most purchases: Asymptotically the market goes to a monopoly for that product. For SQSSI, the purchases at this stage of the market are distributed through a larger number of products, each of which have fewer purchases. Asymptotically, the market will go to a monopoly for two products (one for segment 1 and one for segment 2) but the popularity signal is weaker for SQSSI since it is spread across the two classes. It is interesting to observe that the segmentation policy SQSSI is still more efficient

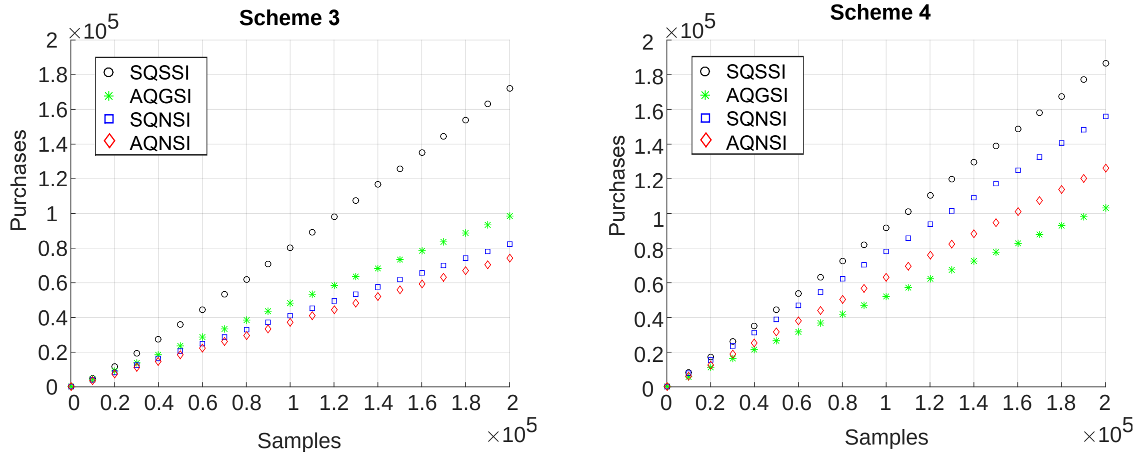

Figure 1.7: The number of purchases over Time for the various rankings. The x-axis represents the number of items tried and the y-axis represents the average number of purchases over all experiments. The left figure depicts the results for Scheme 3 and the right figure for Scheme 4.

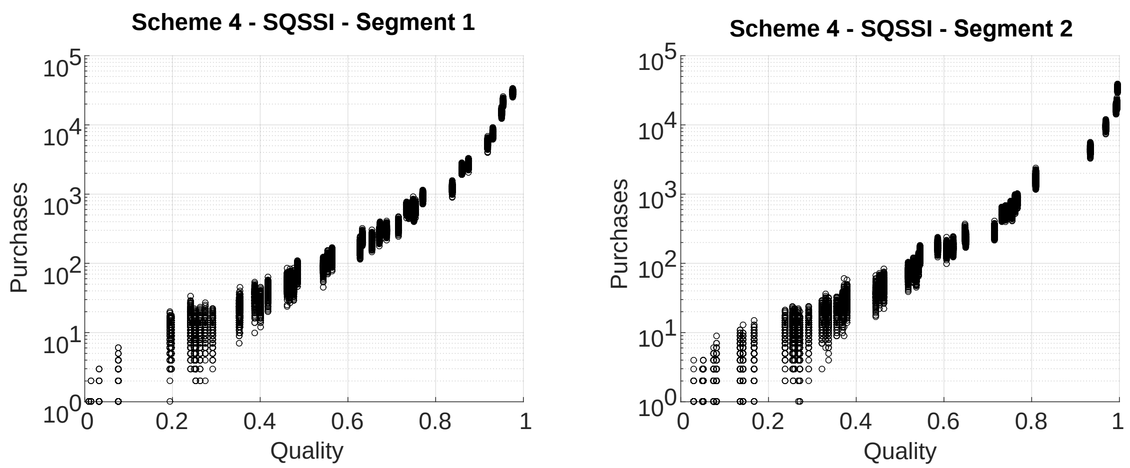

Figure 1.8: The purchase profiles of SQSSI on Scheme 4 for consumer segments 1 (left) and 2 (right).

than policy AQGSI despite this weaker popularity signal. Figure 1.12 depicts the profiles for policy SQSSI and nicely highlights that many products are receiving significant purchases.

We finish this section with a set of managerial insights that can be observed from the theoretical results and the computational experiments. First, whenever consumer segment information is available and the goal is to maximize long term purchases, it is optimal to segment the consumers and to rank the products within each segment according to their quality, while using a social

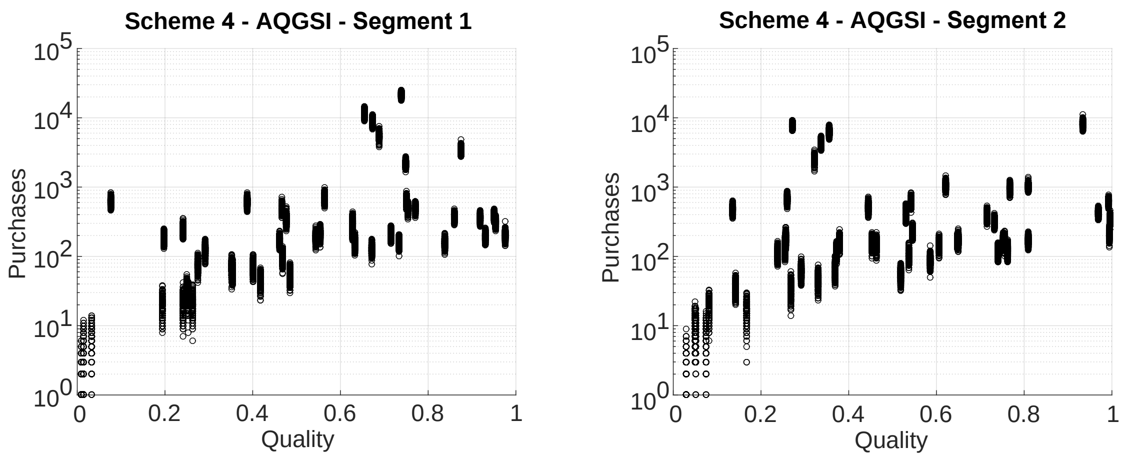

Figure 1.9: The purchase profiles of AQGSI on Scheme 4 for consumer segments 1 (left) and 2 (right).

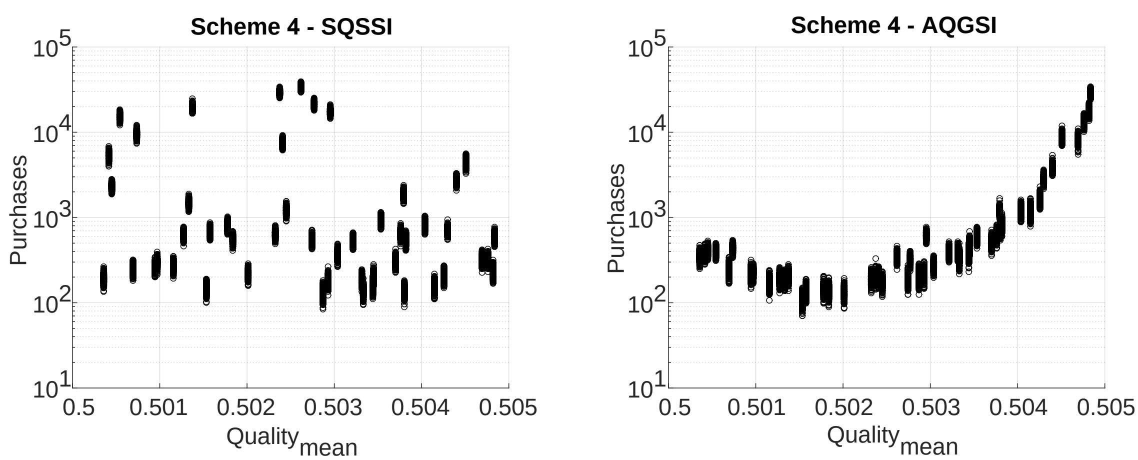

Figure 1.10: The purchase profiles of SQSSI and AQGSI on Scheme 4 for both segments of customers.

signal that is only shown across consumers within the same segment (Theorems 1.3 and 1.5). The first theorem shows that a segmented quality ranking with social influence within each segment (SQSSI) always outperforms the unique average quality ranking with the same social signal across all consumer segments (AQGSI). The second theorem shows that using social influence with a segmented quality ranking (SQSSI) always outperforms doing a segmented quality ranking without social influence (SQNSI). In the computational experiments we observe similar results. However, the platform owner may be better off in the short term using a unique average quality ranking

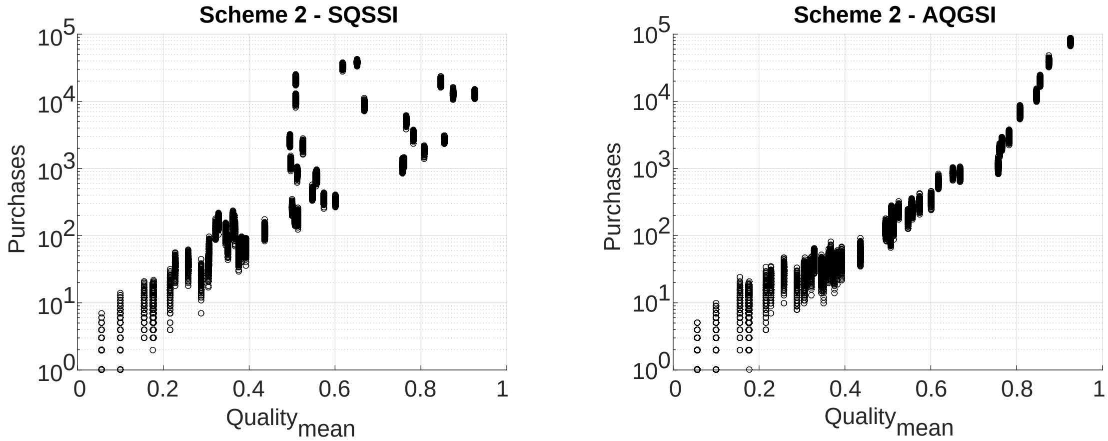

Figure 1.11: The purchase profiles of SQSSI and AQGSI on Scheme 2 for both Classes of Customers.

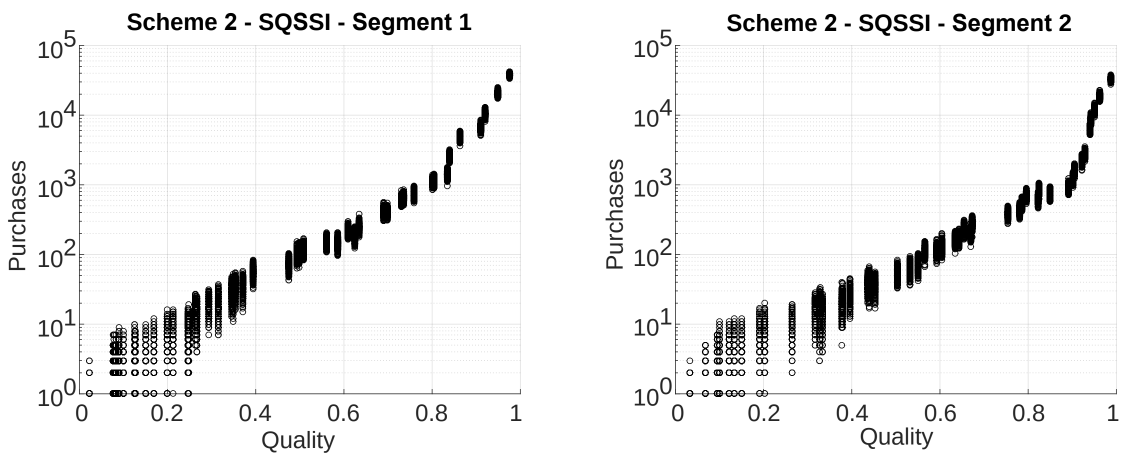

Figure 1.12: The purchase profiles of SQSSI on Scheme 2 for consumer segments 1 (left) and 2 (right).

with a global social signal (AQGSI) (Scheme 1, Figure 1.6). This is because the aggregate ranking AQGSI enhances the social signal instead of distributing it across different segments as seen with the segmented ranking SQSSI. But in the long run the segmented ranking is optimal whereas the aggregate ranking may lead to a monopoly of suboptimal products. Finally, when consumer segment information is not available, it is not always beneficial to show social influence. This is proved in Theorem 1.3 and nicely illustrated with the experiments, where the average quality ranking with a social signal (AQGSI) outperforms its counter-part without a social signal (AQNSI) in all schemes except for Scheme 4 (when consumer segments exhibit opposite preferences, and product appeals are positively correlated with qualities). On aggregate rankings, the platform owner should use social influence carefully, and mainly when consumer segments have similar product preferences, as otherwise social influence may confuse consumers on which products to try and reduce the rate of purchases.

1.7 Conclusions and Future Research

In this chapter, we studied a trial-offer market where consumers choices about which products to try are affected by the display of past product purchases and by product positions. Specifically, we focused on studying the case in which consumer choices follow a mixed multinomial logit model (MMNL), which generalizes the multinomial logit model proposed by Krumme et al. (2012) to explain the behavior in an online cultural market (Salganik et al. 2006). Unlike the case for the Multinomial Logit, we showed that finding the best way to rank products at every step in order to maximize the purchases is a computationally hard problem.

The chapter then studied the performance of a ranking policy (AQGSI) which ranks the products by the average quality (in decreasing order). Under such policy, we proved that the trial-offer market in the long run converges predictably to a monopoly by transforming the MMNL model into a traditional MNL model whose appeals and qualities depends on the popularity signal at each time step but are bounded from below and above. Unfortunately, this average quality ranking policy is no longer guaranteed to benefit from the popularity signal in all cases. In other words, the rate of product purchases can sometimes be reduced as time passes due to the fact that consumers are able to observe (and their decisions are affected by) the number past purchases for each product. This is in sharp contrast to the case where there is a unique consumer segment ($K = 1$) (Van Hentenryck et al. 2016a).

The chapter also studied a market segmentation policy for settings in which the firm is capable of detecting the consumer segment in advance. Under the segmentation policy, the products are presented to each consumer according to the quality ranking for their own class. In addition, the popularity signal displayed only aggregates the past purchases associated to the customers of the same class. The resulting policy (SQSSI) is optimal asymptotically in expectation and may improve the market efficiency up to a factor K over AQGSI, where K is the number of customer classes.

Computational experiments have been presented to illustrate our theoretical results and to show the market dynamics in the short term. These experiments, which were carried over four different settings, showed that there are settings in which the popularity signal is indeed detrimental to policy AQGSI and that the segmentation strategy produces the best possible improvement predicted by the theory.

Overall, the chapter shows that, in trial offer markets, the decision to display or not the popularity signal to consumers has to be analyzed very carefully. In markets where consumers have very different product preferences, we showed that the display of past purchases can be detrimental to the rate of purchases. Intuitively, this is because consumers watching past purchases become confused about which products to try. On the other hand, in markets where the preferences of the different consumer classes are not very distinct, the display of aggregated past purchases is beneficial. Nevertheless, our theoretical (asymptotic) results and our short-term simulation results indicate that the segmentation policy SQSSI outperforms all other policies. Moreover, these results highlight the fact that AQGSI and SQSSI produce very different market behavior, even in settings where the overall market efficiency is relatively close.

The chapter leaves some interesting questions for future research. The first is to study the market-share dynamics of the different ranking policies for more complex consumer choice models. One weakness of the current model is that the consumer choice model only depends on the displayed vector of past purchases and the current ranking (as well as the appeal of the products). However, a more sophisticated choice model could incorporate into the consumer choice, the type of past purchase information displayed: A consumer who observes a past purchase vector d might change the behavior depending on whether she/he knows that the vector d comes from all consumer purchases or if only comes from consumers of it own type or class. The second one is to extend the model to scenarios when the trial probabilities depend non-linearly on the past purchases, this is studied in Maldonado et al. (2018) for the special case of a unique consumer segment ($K = 1$). Another interesting research direction is to study the firm potential incentives to hide or mis-report some reviews and how it this affects the outcomes in terms of market share dynamics and consumer welfare.

Chapter 2

Strategic Implications of Binge Consumption of Entertainment Goods: an Analysis of AVOD Services

2.1 Introduction and Review

The success of streaming services like Netflix and Amazon has led to major changes in the television (TV) industry. Subscription-based video on-demand (SVOD) streaming services release episodes of their original content all-at-once allowing viewers to binge-watch a series without having to wait until the following weeks to watch additional episodes.[1] As a result, rival television based media companies, such as ViacomCBS, NBCUniversal and Fox, have begun to look for new methods to reach viewers because their traditional TV distribution channel restricts viewers to a sequential viewing schedule.

The difference in the level of viewer commitment/flexibility across SVODs and traditional TV may be one reason why subscription based video on-demand streaming services are successful. In response "networks are changing the way they develop and release new shows...as they seek to adapt to new TV viewing habits and profit from the 'binge-watching' made popular by video streaming services" Toonkel and Richwine (2016). One such method is for traditional television based media companies to mimic their digital streaming counterparts (e.g., Netflix and Amazon) and create

[1] binge watching, the act of watching several episodes in one sitting

a digital distribution service that would enable a content provider to also release new content all-at-once. Such an offering has been successful—ViacomCBS has a subscription free advertising-based video on demand (AVOD) service named Pluto TV which has more than 43 million monthly active users.[2] NBCUniversal also entered the subscription free AVOD space with the launch of its streaming service Peacock.[3] Fox also has a subscription free AVOD service named Tubi which made more than $300 million in ad revenue in 2020 and has more than 33 million users.[4]

The AVOD space has grown by offering repackaged content, but now they are in a position with enough capital base to fund their own original series. Tubi, Roku Inc., and Pluto TV started looking at developing originals.[5] [6] The AVOD business model is exhibiting similarities to that of SVOD's, where Netflix Inc. and Hulu started offering reruns and licensed programs, and now rely on original content to attract new customers.

But once such an AVOD service like Pluto TV, Peacock, Roku Inc. or Tubi creates original content[7], how does it determine which shows should be released sequentially (like in traditional TV) or all-at-once? Why did NBCUniversal decide to release reboots of "Saved By the Bell" and "Punky Brewster" on Peacock with an all-at-once release timing strategy, and not sequentially?[8]

Currently, there is little understanding as to the strategic implications of the all-at-once release strategy, particularly within an AVOD setting and more specifically of the type of shows (high or lower quality) that should implement such a strategy. This chapter studies the profitability impact of simultaneous and sequential releases for shows within a subscription free AVOD service. Our model exploits the asymmetry of information about show quality between the entertainment companies and viewers to analyze the optimal release strategy for high- and low-quality shows. We focus on AVOD distribution rather than a paid subscription service due to the close similarities an AVOD service has to the existing TV market and the fact that AVOD revenue is projected to hit 56 billion in 2024.[9]

With new advances in technology, especially the ability to watch on demand, binge consumption

[2] https://www.cordcuttersnews.com/pluto-tv-reaches-43-million-monthly-active-users/

[3] https://www.cnet.com/news/peacock-tv-everything-to-know-premium-free-app-plus-watch-wwe/

[4] https://www.fool.com/investing/2021/03/10/fox-is-understating-tubis-potential/

[5] https://www.msn.com/en-us/money/companies/fox-e2-80-99s-tubi-made-millions-with-reruns-now-it-wants-original-programming/ar-BB1ey3pQ

[6] https://www.theverge.com/2021/3/13/22329368/tubi-streaming-original-programming-fox

[7] https://www.cordcuttersnews.com/tubi-is-interested-in-creating-original-content/

[8] https://www.tomsguide.com/news/how-to-watch-punky-brewster-reboot-on-peacock-release-date-cast-trailer-and-more

[9] https://www.mediaplaynews.com/ad-supported-vod-revenue-to-reach-56-billion-by-2024-driven-by-u-s/

of entertainment has attracted a lot of attention. Also, millennials who are comfortable with technology are more likely to engage in this behavior. This is an emerging area of interest to researchers with limited analysis. As with any new substantive area of interest, any attempt to model the phenomenon immediately raises several interesting research possibilities and a desire to capture the richness of the phenomena. Early researchers, however, need to carefully carve out the issues to ensure tractability and gain meaningful insights. (In contrast, mature research projects have already addressed several issues and hence it becomes easy to focus on the remaining ones.) In that same spirit, we elect to focus on the issue in the context of one content and channel provider.

Traditional TV networks generate most of their profits through advertising. A 30-second spot during a hit TV show can cost as high as $\$400,000$ Nathanson (2013). It is reasonable to think that the cost per advertising spot would be different across the two release timing strategies that lead to different viewing behavior (binging vs. not-binging). Schweidel and Moe (2016) studies the impact of binge-watching on advertising through an empirical model based on Hulu user data. Their conclusion is that advertisements in a viewing session discourage binge-watching, and that binge-watchers are less responsive to advertisements compared to non-binge-watchers. Additionally, the satiation literature suggests that high rates of consumption lead to less enjoyment than consuming a good less often Galak et al. (2013). These results may have an interesting implication for the role of advertisements with different release timing strategies: in all-at-once releases, ads are consumed at a higher rate than in traditional weekly releases; thus, the efficacy of ads in all-at-once releases may be reduced compared to a traditional weekly release timing. These two lines of research support an assumption we hold throughout the chapter: users show less ad responsiveness in simultaneous releases compared to a sequential release timing. In our model, we translate this deterioration in advertisement responsiveness into a monetary value.

To study the impact a simultaneous release strategy could have on a traditional television based media company, we analyze the trade-off between the boost in ad efficiency via sequential (linear) releases and the increase in the cost of watching. For instance, watching a show that is released all-at-once gives viewers the possibility to binge-watch in a convenient way, which is not possible with a sequential release timing.

We assume viewers are unaware of the quality of a show until they watch it. One possible method in which an entertainment company can signal a show's quality to viewers is through advertising about its show. The show-specific advertising level is a decision for the firm, which can act as a signal of show quality to viewers. In addition to advertising, the release timing decision may also act as a signal about show quality to viewers due to the advertising revenue difference between the two release timing options we are considering.. Networks have more information about their shows than viewers, which influences their decisions on how to advertise and how to release its shows. Viewers see these decisions and experience imperfect information about the show quality. We therefore analyze this problem with a signaling model. It is important to note that in our analysis, quality measures are relative. The relative nature of our analysis is a direct result of the model and occurs in all signaling papers.

This chapter illustrates that adequate levels of advertising alone may signal quality, and for this to happen, the higher-quality show must incur a sizable cost. We then show that by adding the simultaneous release timing in a AVOD channel, in equilibrium lower quality shows select this release strategy since it is more profitable than the sequential strategy. Thus, we find that there exists a separating equilibrium under which release timing strategies signal quality. Furthermore, we determine that the introduction of the simultaneous release timing reduces the advertising level that higher quality shows need to incur in order to signal their quality (compared to having sequential releases alone). Even though the higher quality show does not apply a simultaneous release strategy, by providing a more profitable release strategy for low quality shows, they are better-off. This is because the incentive compatibility constraints are relaxed, allowing higher quality shows to reduce their equilibrium advertising level under sequential releases.

In our model, the traditional television based media company is tasked with selecting which release timing strategy to use for new content, and thereby determines what show quality types are binged and which are not. An interesting result of our analysis is that in equilibrium binge watching occurs with lower quality shows, not high quality. Fascinatingly, this result has direct parallels to work where consumers are found to binge low quality foods and/or beverages Boggiano et al. (2014). Our results also have important managerial implication; because release strategies may signal quality, it is beneficial for a firm to find a way in which to open new strategies that are profitable for its lower quality content. By doing this, the firm reduces the necessity of costly signaling through other mechanisms (i.e., advertising) for their higher quality content, and enjoys greater profits on all quality levels.

Literature Review

Our model contributes to the signaling literature by illustrating that the release timing of shows can be a signal of quality. To the best of our knowledge, there are no other papers that analyze the release timing decision of episodes as a signal of quality. This is an important result because it provides television networks with another mechanism to informatively separate their content. Furthermore, we show that release timing strategies reduce the advertising expenditure shows must incur in order to signal their quality. Other signals of quality that have been studied in the literature include price Bagwell and Riordan (1991), money-back guarantee Moorthy and Srinivasan (1995), umbrella branding Wernerfelt (1988), slotting allowances Lariviere and Padmanabhan (1997), advertising and price Linnemer (2002), Abe (1995), Milgrom and Roberts (1986), Desai (2000), Zhao (2000), Erdem et al. (2008), advertising frequency Erdem et al. (2008), warranty Lutz (1989), Gal-Or (1989), Balachander (2001), price image Simester (1995), product scarcity Stock and Balachander (2005), brand extension Moorthy (2012) and price discrimination Anderson and Simester (2001). Among these papers, the closest to our work is Moorthy and Srinivasan (1995). In this paper, the authors analyze how money-back guarantees can signal product quality. They show that money-back guarantees signal quality by exploiting the higher probability of returns for a lower quality product, and the attendant higher transaction costs. However, if the seller's transaction costs are too large, other mechanisms (like price) are needed to signal quality.

An important distinction between our model and most papers in the signaling through advertising literature is that we consider multiple consumption instances that lead to consumer learning, and the fact that advertisements can act as explicit and implicit provisions of information as well as generate prestige effects for consumers–similar to Ackerberg (2005), Stigler (1961), Butters (1977), and Grossman and Shapiro (1984) were some of the first papers to analyze the use of advertisements as explicit provisions of information. These papers analyzed the effect of firms explicitly informing their consumers of their brands' existence and observable characteristics through advertising. In our setting, such advertising would explicitly informs consumers of who the leading actors or actress are of the show, the genre of the show, and a synopsis of the show's plot. On the other hand, the initial literature of advertising of experience goods, led by Nelson (1974), Milgrom and Roberts (1986) and Kihlstrom and Riordan (1984), analyze the use of advertising as means to implicitly signal information to consumers about a brand's unobserved quality. They find that firms are able to signal unobserved quality via advertisement levels. Moraga-González (2000) analyzes quality signaling through informative advertising in an experience good in a one period model. He finds that no separating equilibrium exists in advertising. This contrasts with the result of this chapter, where we find a separating equilibrium in advertising. This is because we consider an experience good with multiple periods of consumption, including quality learning, which permits the feasibility of the incentive compatible constraints under a separating equilibrium.

Stigler and Becker (1977) and Becker and Murphy (1993) follow the earlier work of Stigler (1961), Butters (1977), and Grossman and Shapiro (1984) by examining models in which advertising levels interact with a consumer's utility function for a particular brand. Such an impact might occur through a prestige or image effect whereby consumers garner greater utility for a brand due to the content of the advertisement–e.g. a celebrity endorsement (Chung et al. (2013), Derdenger et al. (2018)) or highlighting the leading actors or actresses in a movie or TV show. Given this work and the relevance to our setting, it is important that we model this aspect of advertising by allowing it to impact a viewer's utility directly, independently of their beliefs of show quality.

Among signaling games, there exists an interesting difference between deterministic and stochastic signaling mechanisms. In the former, consumption gives complete information about the unknown feature to the responding agent, whereas in the latter, the responding agent is still not completely sure about the true value of the unknown feature. This work stands in the middle: depending on the experience draw a consumer receives from a particular episode, she will have perfect or imperfect information about the true show type. A similar model to the one in this chapter may be found in Jeitschko and Normann (2012). This paper contrasts a standard deterministic signaling game with a stochastic signaling mechanism. They find that in the stochastic setting, a unique equilibrium exists that separates agent types, whereas with a deterministic signaling mechanism, both pooling and separating equilibria exist. The main difference with this paper is that our model is suitable for a TV show-viewer interaction. We consider sequential trials (consumers receive a quality sample for each episode they watch), that the samples received may be deterministic or noisy depending on the realization of a random variable, and two different signaling mechanisms, advertising and release timing. It is also interesting to note that our equilibrium strategies are optimal in expectation, but there might be some sample paths in which they are not optimal. This is due to the randomness in the experienced episode quality; it could happen that a low quality show mimicking a high quality show gets lucky and does not perfectly inform certain viewers about its true quality, ending up better-off than in its equilibrium strategy.

Given our model setting, the area of research focused on binge behavior is quite relevant. Such an area is also steadily increasing in number of papers. Recent papers such as Trouleau et al. (2016) models user episode playback through a regression model on event counts that is contrasted with a dataset from a on-demand video streaming service. Their model includes several features that are key aspects of binge behavior, including episode data censorship (whether a viewer has watched all episodes available to her or not), deviations in the population, and external influences on consumption habits. They observe different types of binge behavior: that binge watchers often view certain content out-of-order, and that binge watching is not a consistent behavior among users. Lu et al. (2017) analyzes binge behavior in an educational setting through content in Coursera (one of the world's most popular online education platforms), in which they observe individual-level lecture and quiz consumption patterns across multiple courses. They generate a utility maximization decision process for individual consumers that features the contemporaneous utility of consumption and the long-run accumulation of knowledge. By examining consumption in certain courses, their model is able to predict consumption patterns in other courses, which has implications for new product launch, cross-selling, and bundling. Schweidel and Moe (2016) studies the impact of binge-watching on advertising through an empirical model based on Hulu user data. This is the closest paper to our work, since it analyzes the relationship between advertising effectiveness and binging-behavior. Their conclusion is that advertisements in a viewing session discourage binge-watching, and that binge-watchers are less responsive to advertisements compared to non-binge-watchers.

There is also extensive literature that studies the optimal release timing of media through same or different channels. Hennig-Thurau et al. (2007) present a model of revenue generation across four sequential distribution channels, combining choice-based conjoint data with other information. Under the conditions of the study, the authors find that the simultaneous release of movies in theaters and on rental home video generates maximum revenues for movie studios in the United States, but has devastating effects on other players, such as theater chains. Prasad et al. (2004) studies the entry time of goods in different channels. In their model, they include the discounting of future profits, the foresight of the firm, customers' expectations, and the possibility of cannibalization, and find an optimal closed form solution for the optimal sequential timing. Das (2008) analyzes the optimal release timing of movies in the context of piracy. Krider and Weinberg (1998) studies the release timing of movies as a competitive game and look for an equilibrium during high season. The remaining literature can be found in Chiou (2008), Gerchak et al. (2006), Frank (1994).

In addition to signaling, this work considers viewer learning. Previous papers that model consumer learning and advertising jointly are Erdem and Keane (1996) and Ackerberg (2005). As viewers watch episodes, they experience a random quality from each episode and update their beliefs about the show type, according to Bayesian update. Depending on the sample path, some viewers might take more or less time in identifying the true show type.

2.2 Model

As was discussed in the introduction, it is important for researchers analyzing new substantive areas to carefully carve out the issues to ensure tractability and gain meaningful insights. We believe the model presented to you below does exactly that.

Consider a traditional television based media company, such as NBCUniversal, Fox, or ViacomCBS, which used to have no alternative rather to release episodes sequentially through TV, but now has an advertising-based video on demand service which allows for two different release timing strategies (sequential vs. simultaneous). The "television" company in our model is assumed to produce only one show and knows its quality, but consumers do not. Similarly to the existing literature Bar-Isaac (2003), we assume show quality can come from two different types: "good" shows and "bad" shows. Additionally, the company decides its advertising expenditure ($a \geq 0$) and the release timing of its show to a homogeneous group of viewers. Without loss of generality, we normalize the size of the group to 1. For tractability and model simplicity we assume there are two release timing options: i) simultaneous releases, which we will denote by B, and ii) linear sequential releases, which we will denote by W. In simultaneous releases, all episodes from the show are offered to the viewer on premiere day (period 1), whereas in sequential releases, a new episode is released each subsequent period. The "television" company earns revenue from outside firms advertising within show episodes, the revenue per eyeball for sequential releases without loss of generality will be set to 1. The efficiency from advertisements is different between the two release timing strategies; simultaneous releases have an ad efficiency of δ compared to linear sequential releases. We see this as if the revenue per ad view for simultaneous shows is a fraction $\delta < 1$

of what it is for sequential shows.[10] Again, as highlighted in the introduction, this modeling assumption is also based on the literature of Schweidel and Moe (2016), Galak et al. (2013). This literature indicates that high rates of consumption lead to less enjoyment than consuming a good less often (Galak et al. (2013)), and that binge-watchers are less responsive to advertisements than to non-binge-watchers (Schweidel and Moe (2016)).

Additionally, the firm chooses its advertising expenditure for its own show. Firm revenue is dependent on how many people watch the show, the release timing choice, and its advertising expenditure; we can write a general expression for the expected profit of a type t (good or bad) show as follows:

$$\pi^t(a, W) = \mathbb{E}[\# \text{ of views}|a, W \text{ and } t] - a \tag{2.1}$$

$$\pi^t(a, B) = \delta\mathbb{E}[\# \text{ of views}|a, B \text{ and } t] - a. \tag{2.2}$$

At this point, it is important to remark that the advertising that generates revenue is completely different to the advertising level chosen by the firm. This latter advertising is specific to the show and release timing strategy, whereas the advertising revenue is generated from exposing viewers to different advertisements from independent and unrelated firms to viewers during the show. Additionally, as we already mentioned, the advertising revenue per eyeball is exogenous, independent of show quality and constant over time. This is because prices for advertising are set before the actual shows are finished and its quality is realized by the advertisersNathanson (2013).

Again, viewers do not know the show's type before watching episode 1, i.e., before watching they cannot determine the show's overall quality, and even after watching episode 1, there might be some residual uncertainty. Viewers have initial prior beliefs about show quality that will be updated directly after the firm announces its release timing strategy $X \in \{B, W\}$ and advertising level $a \geq 0$. Given the chosen strategy $X \in \{B, W\}$ and the advertising level $a \geq 0$, the viewers' priors before watching episode 1 are that with probability $\mu_{\{1,a,X\}}$ they are viewing a "good" show, and $(1 - \mu_{\{1,a,X\}})$ that they are viewing a "bad" show. Consumers have a cost C_X of viewing each episode that depends on the release timing strategy $X \in \{B, W\}$, where $C_B \leq C_W$ because

[10] A March 11, 2016 Fortune Magazine article on "How Network TV Figured Out Binge-Watching" indicates that "binging viewers are also less likely to watch ads..."

all-at-once releases provide more flexibility to viewers than weekly releases. The experience each viewer has about a particular episode is completely independent across viewers, i.e., two different viewers might have different perceptions about the same episode. Furthermore, the experience of a particular viewer across different episodes is completely independent as well. The probability a viewer perceives an episode from a "good" show with high quality (θ_H) is g, whereas the probability of experiencing medium quality (θ_M) is $(1-g)$. For a "bad" show, the probability a viewer perceives an episode as medium quality (θ_M) is b, whereas the probability of experiencing low quality (θ_L) is $(1-b)$. Note that $\theta_L < \theta_M < \theta_H$. We assume that the probabilities g and b, and the qualities θ_L, θ_M and θ_H are common knowledge, so that the uncertainty for viewers relies only on which type of show they are watching, and not on the quality distributions of those shows.

After viewing an episode, viewers update their beliefs about a show's type. If they have experienced a high quality episode (θ_H) in the past, they know they are watching a "good" show and will update their beliefs appropriately, whereas if they have experienced a low quality episode (θ_L) in the past, they know the show is "bad." On the other hand, if their past experiences about episode quality were all medium (θ_M, the degenerate quality), viewers update their beliefs, based on a Bayesian update. If the past $i \geq 0$ experienced episode quality was θ_M under release timing $X \in \{W, B\}$ and advertising level $a \geq 0$, then the probability they give to the show being "good" when deciding to watch episode $i+1$ is

$$\mu_{\{i+1,a,X\}} = \frac{\mu_{\{1,a,X\}}(1-g)^{i-1}}{\mu_{\{1,a,X\}}(1-g)^{i-1} + (1-\mu_{\{1,a,X\}})b^{i-1}}, \tag{2.3}$$

with probability $1 - \mu_{\{i+1,a,X\}}$ that the show is "bad."

Viewers decide to watch an episode based on the utility they expect from that episode which is a function of a show's advertising. This advertisement directly affects the viewer's utility of the show through a prestige effect and may also affect a consumer's expected utility through a signaling effect. The modeling of advertising in such a fashion is consistent with the complementary view where advertising may contain information as well as lead to social prestige that influences consumer behavior (Bagwell (2007)). We incorporate a prestige effect into the consumer's utility function due to a show's advertisements usually highlight its staring actors or actresses. The incorporation and interpretation of the prestige effect in our context is similar to the celebrity endorsement work of (Chung et al. (2013), Derdenger et al. (2018)) where consumers derive direct utility from a celebrity endorsing a given product. For simplicity we choose a linear form equal to γa, where $\gamma > 0$. If a viewer decides not to watch an episode, she quits the show and does not watch any of the following episodes. Consumer decisions on whether to watch an episode are based on their show quality perception, cost of watching, and advertising level. The viewer expected utility from watching episode i given release timing $X \in \{B, W\}$ and advertising level $a \geq 0$ is

$$u_{\{i,a,X\}} = \mu_{\{i,a,X\}}\mathbb{E}[Q_g] + (1 - \mu_{\{i,a,X\}})\mathbb{E}[Q_b] - C_X + \gamma a, \tag{2.4}$$

where Q_g and Q_b are the random qualities experienced in "good" and "bad" show respectively,

$$Q_g = \begin{cases} \theta_H & \text{w.p. } g \\ \theta_M & \text{w.p. } (1-g), \end{cases} \tag{2.5}$$

$$Q_b = \begin{cases} \theta_M & \text{w.p. } b \\ \theta_L & \text{w.p. } (1-b). \end{cases} \tag{2.6}$$

Then, $\mathbb{E}[Q_g] = g\theta_H + (1-g)\theta_M$ and $\mathbb{E}[Q_b] = b\theta_M + (1-b)\theta_L$.[11]

We can address this problem using a model of seller-consumer interaction as a sequential game of incomplete information and look for a modified Perfect Bayesian Equilibrium Gibbons (1992). The game would be as follows: nature chooses the firm's show type, with probability μ_0 it is "good" and with probability $1-\mu_0$ it is "bad." Then, the firm moves by choosing a release timing strategy, as well as an advertisement level for its show; the viewer moves second with her decision to watch or not to watch the first episode. The viewer's decision about watching the first episode will be based on her posterior assessment of the probability of dealing with a "good" show, after seeing the show's release timing signal and advertisement level, as that will determine the viewer's utility. If the addition of the expected quality and advertisement utility from the first episode is greater or equal to the cost of watching, the viewer will watch such an episode; otherwise not. If the viewer decides to watch the first episode, she performs a Bayesian update on her belief about the show

[11] Under the current model structure, if viewers included future show utility into their decision making process, the utility would simply be a rescaled amount of the presented utility in equation (4) due to the viewers expecting to receive a constant flow utility across each episode. For instance, in the infinite episode case, the expected utility associated with all future episodes would simply be a scaled utility of equation (4), specificially by a factor of 1/(1-β) where β is the discount factor of future utility.

being good, forms an expected utility for episode 2 and decides to watch it or not based on the same criterion. Viewers will stop watching the show when their expected utility from watching the next episode is negative.

2.3 Signaling Through Advertising Alone

To identify the impact of the simultaneous release timing, we first study the traditional release timing setting in which show episodes are released sequentially in an AVOD service. In this setting, the only way a show can signal its quality to viewers is through advertising. We make a simplification to the model by setting $g = 1$ and $b = 0$, which implies that $\mathbb{E}[Q_g] = \theta_H$ and $\mathbb{E}[Q_b] = \theta_L$. Viewers perceive episodes from "good" shows with high quality (θ_H), whereas they perceive episodes from "bad" shows with low quality (θ_L). The probability that a viewer receives the degenerate quality θ_M from an episode is zero, so by watching the first episode they will know the show's true type. We first analyze the setting in which we have a unique episode ($N = 1$). Then, we analyze the case in which we have three episodes ($N = 3$), which we believe is general enough. [12]

2.3.1 Single Episode Model

Our first exercise is to illustrate that signalling via advertising does not occur with only one period but rather multiple periods or episodes is required. A separating equilibrium in which a high quality show chooses an advertising level a_g, while a low quality show chooses an advertising level a_b, must satisfy the incentive compatibility constraints. In this scenario, the equilibrium profits from a high and a low quality show are $\pi^g(a_g, W) = 1 - a_g$ and $\pi^b(a_b, W) = 1 - a_b$ respectively. If a low (high) quality show were to mimic the equilibrium advertising level of the high (low) quality show, viewers would watch the first episode, after which they would find out the true quality show, but because there is only one episode to watch, this information does not affect the revenue. Thus, the mimicking profits are: $\pi^g(a_b, W) = 1 - a_b$ and $\pi^b(a_g, W) = 1 - a_g$.

The incentive compatibility constraints are $\pi^g(W, a_g) = 1 - a_g \geq \pi^g(W, a_b) = 1 - a_b$ and $\pi^b(a_b, W) = 1 - a_b \geq \pi^b(a_g, W) = 1 - a_g$, which may only be satisfied when $a_g = a_b$. Thus, we

[12] We analyze all possible equilibria in which both show types participate in the market and attract viewers; first under complete information, and then under incomplete information, where we find the parameter sets that lead to separating and pooling equilibria. We also include three episodes as to allow for an extension that incorporates a viewer's urge to close.

cannot have a separating equilibrium on the advertising level alone when there exists a unique episode.

2.3.2 Complete Information With Multiple Episodes

We now study the setting with three episodes. When information is symmetric, the game becomes simple. The profit from a high quality show with advertising expenditure a_g is

$$\pi^g(a_g, W) = 3\mathbb{1}(\theta_H - C_W + \gamma a_g \geq 0) - a_g, \tag{2.7}$$

while the profit that a low quality show earns advertising a_b is

$$\pi^b(a_b, W) = 3\mathbb{1}(\theta_L - C_W + \gamma a_b \geq 0) - a_b. \tag{2.8}$$

Each show type would choose the minimum advertising level that attracts viewers while generating a non-negative revenue, otherwise they don't participate in the market. We can represent the equilibrium advertising levels for a "good" and a "bad" show with

$$a^*_{gW} = \begin{cases} \left[\left(C_W - \theta_H\right)/\gamma\right]^+ & \text{if } \pi^g\left(\left[\left(C_W - \theta_H\right)/\gamma\right]^+, W\right) \geq 0 \\ 0 & \text{otherwise} \end{cases}, \tag{2.9}$$

and

$$a^*_{bW} = \begin{cases} \left[\left(C_W - \theta_L\right)/\gamma\right]^+ & \text{if } \pi^b\left(\left[\left(C_W - \theta_L\right)/\gamma\right]^+, W\right) \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{2.10}$$

respectively, where $f^+ = \max\{0, f\}\ \forall f \in \mathbb{R}$.

If we consider the situation in which both firms need to advertise to attract viewers while making non-negative profit, we have that $3 \geq a^*_{gW} = \left(C_W - \theta_H\right)/\gamma > 0$ and $3 \geq a^*_{bW} = \left(C_W - \theta_L\right)/\gamma > 0$. Thus, even in the complete information case, the firms need to advertise in order to enhance consumer utilities and encourage viewers to watch the show. In the next section we analyze the case of incomplete information.

2.3.3 Incomplete Information with Multiple Episodes

When there is asymmetry of information, high quality shows may want to separate from the first best strategies in order to inform viewers about its high quality by incurring a cost that the low quality show is not willing to afford. In the following two sections, we analyze the possible separating and pooling equilibria respectively.

Separating Equilibria

Under a separating equilibrium, as soon as the firm chooses an equilibrium strategy, viewers will have absolute beliefs about the show's type. We now look into a separating equilibrium in which both show types attract viewers; high quality shows choose an advertising level $a_{gW} > 0$ whereas low quality shows choose an advertising level $a_{bW} > 0$.

In order to attract viewers and obtain a non-negative profit in equilibrium, $3 \geq a_{gW} \geq a^*_{gW}$ and $3 \geq a_{bW} \geq a^*_{bW}$, where a^*_{gW} and a^*_{bW} are the first best solutions previously defined. To set the incentive compatibility constraints, we must derive the expected equilibrium profits as well as the expected profit of each show type mimicking the other. If a high quality show were to mimic the behavior of a low quality show, its profit would be the same as a low quality show receives in equilibrium $\pi^g(a_{bW}, W) = 3 - a_{bW}$. Whereas, if a low quality show were to mimic the behavior of a high quality show by choosing advertising level a_{gW}, viewers would watch the first episode, after which they would realize that the show is bad, and then there are two possibilities:

(A) The quality of the low-type show is good enough, so that viewers would still watch the remaining episodes even after determining that the show is of low quality. This happens when $\theta_L - C_W + \gamma a_{gW} \geq 0$.

(B) The quality of the low-type is not good enough, and viewers stop watching the show after watching episode 1 and determining that the show is of low quality. This happens when $\theta_L - C_W + \gamma a_{gW} < 0$.

We can write the payoffs as follows:

$$\pi^g(a_{gW}, W) = 3 - a_{gW} \tag{2.11}$$

$$\pi^g(a_{bW}, W) = 3 - a_{bW} \tag{2.12}$$

$$\pi^b(a_{bW}, W) = 3 - a_{bW} \tag{2.13}$$

$$\pi^b(a_{gW}, W) = 1 - a_{gW} + 2\mathbb{1}(a_{gW} \geq a^*_{bW}) \tag{2.14}$$

Equations (2.11) and (2.13) represent the equilibrium profits for a high and low quality show respectively, each of them attracting 3 views. Equation (2.12) represents the case where the high quality show mimics the low quality show, which would lead to the same number of views. Equation (2.14) represents the case where the low quality show mimics the high quality one; viewers would watch the first episode, find out that the show is of low quality, and then, depending on their expected utility from episode number 2, would continue watching the show or not. Thus, if a_{gW} is sufficiently large (greater than a^*_{bW}), viewers would watch all episodes, but if it is not, they will exit after watching episode 1.

The incentive compatibility constraints ensure that each show type is better-off by sticking with its equilibrium strategy rather than mimicking the other type's equilibrium strategy. We write these constraints as follows:

$$\pi^g(a_{gW}, W) = 3 - a_{gW} \geq 3 - a_{bW} = \pi^g(a_{bW}, W) \text{ and} \tag{2.15}$$

$$\pi^b(a_{bW}, W) = 3 - a_{bW} \geq 1 - a_{gW} + 2\mathbb{1}(a_{gW} \geq a^*_{bW}) = \pi^b(a_{gW}, W). \tag{2.16}$$

We may simplify them to $a_{bW} \geq a_{gW} \geq a_{bW} - 2\Big(1 - \mathbb{1}(a_{gW} \geq a^*_{bW})\Big)$. Note that if $a_{gW} \geq a^*_{bW}$ we have no separating equilibrium. In this case, the incentive compatibility constraints (2.15) and (2.16) lead to $a_{gW} = a_{bW}$, in which consumers cannot know the show quality by observing the advertising level. We restrict ourselves to the set where $a_{gW} < a^*_{bW}$. This means that if a low quality show were to mimic the behavior of a high quality show, then the low quality show would only get one view, since the advertising level is not large enough to attract viewers after determining the show is of low quality.

We may finally write the conditions for a separating equilibrium as follows:

$$\max\{a^*_{gW}, a_{bW} - 2\} \leq a_{gW} < \min\{a_{bW}, a^*_{bW}\} \tag{2.17}$$

with $0 < a^*_{bW} \leq a_{bW} \leq 3$ and $a^*_{gW} > 0$. Under these conditions, there exist out-of-equilibrium beliefs that allow for the existence of a separating equilibrium in which both show types advertise and attract viewers. This depends highly on the parameters, since the constraints are not feasible for all parameter values. From constraint (2.17) and the non-negative equilibrium profit conditions, we can find the parameter set that allows for the existence of this type of separating equilibrium; this set is:

$$0 < a^*_{gW} < a^*_{bW} \leq 3. \tag{2.18}$$

According to constraint (2.17) there can be infinitely many separating equilibria supported by different out-of-equilibrium beliefs. However, in our case, all such equilibria except one are ruled out or refined away by specifying what beliefs are "unreasonable" using the "intuitive criterion" by Cho and Kreps (1987). An in depth application of the "intuitive criterion" can be found in Jiang et al. (2011). Proposition 2.1 shows that any separating equilibrium with positive advertising levels and viewership (i.e., satisfying constraint (2.18)) must be such that $a_{bW} = a^*_{bW}$ and $a_{gW} = \max\{a^*_{gW}, a^*_{bW} - 2\}$ in order to be able to survive the intuitive criterion. This is because this equilibrium is the one that provides the minimum cost separation.

Proposition 2.1. *A separating equilibrium in which high quality shows advertise $a_{gW} > 0$ while low quality shows advertise $a_{bW} > 0$ survives the intuitive criterion as long as $a_{bW} = a^*_{bW}$ and either*

- $a_{gW} = a^*_{gW}$ *and* $a^*_{gW} \geq a^*_{bW} - 2$*, or*
- $a_{gW} = a^*_{bW} - 2$*,* $a^*_{gW} < a^*_{bW} - 2$ *and* $\mu_0 < \frac{(C_W - \theta_L - \gamma a_{gW})}{\theta_H - \theta_L} = \frac{2\gamma}{\theta_H - \theta_L}$.

Proof. Presented in Appendix B.1.1

In Proposition 2.1 we refine our separating equilibrium into a unique one, as long as μ_0 is low enough when $a^*_{gW} < a^*_{bW} - 2$. The intuitive criterion leads to the minimum cost separating equilibrium. We find that when $a^*_{gW} \geq a^*_{bW} - 2$, the only separating equilibrium that survives the intuitive criterion is $\{a^*_{gW}, a^*_{bW}\}$. Therefore no show types deviates from their first best solution, which is uninteresting. In the case that $a^*_{gW} < a^*_{bW} - 2$, the unique separating equilibrium that survives the intuitive criterion is $\{a^*_{bW} - 2, a^*_{bW}\}$ and it holds as long as $\mu_0 < \frac{2\gamma}{\theta_H - \theta_L}$. Compared to the first best solution, here the high quality show is incurring a cost in order to signal its quality to viewers; this cost is $a^*_{bW} - 2 - a^*_{gW} = \frac{\theta_H - \theta_L}{\gamma} - 2 > 0$. The intuition behind the upper bound of μ_0 comes from the following argument: there exist out-of-equilibrium beliefs under which the strategy a^*_{gW} dominates for both show types, thus the intuitive criterion sets a belief of μ_0 to such strategy. If μ_0 was large enough, shows might be willing to deviate to this strategy; by setting an upper bound on μ_0, the shows refrain from deviating.

It is important to note that we have separation because of the multi-period nature of this model. When a low quality show mimics the behavior of a high quality show, viewers would find out the true quality of the show after watching episode 1 (in this simple case when $\mathbb{E}[Q_g] = \theta_H$). So, from period 2 onwards, consumers expect episodes of low quality, $\mathbb{E}[Q_b]$, but since this show decided to choose a lower advertising level than their equilibrium strategy (by mimicking the high type), their expectations of the show are not strong enough to have viewers continue watching the show. The penalty that the low quality show receives from mimicking the high type comes from multiple periods of consumption. This separating equilibrium could never hold in a one-shot game. Moreover, we also find separation without having customer heterogeneity or differences in the marginal advertising costs of each firm type.

Finally, if we were to consider only implicit provisions of information through advertising (not explicit provision nor prestige effects), we would not find a separating equilibrium in advertising in which both firms participate. This is because under the first best solution, we would already have that both firms advertise 0 achieving 3 views; thus, there is no benefit in deviating under imperfect information. The separating equilibrium we could find is one in which the low quality show does not participate under the first best solution, while the high quality show advertises 0. When information becomes imperfect, the high quality show needs to increase their advertising level so that it is not profitable for the low quality show to enter the market and mimic its advertising level. Thus, we would still see that the high quality show separates to an advertising level that achieves the minimum cost separation with the low quality show not participating. [13]

[13] The results of this chapter differ with much of the existing literature that analyzes the effects of advertising as a signal of quality in absolute values, but in relative values (respect to the first best solution) they are consistent.

Pooling Equilibria

In this section we analyze the conditions under which we have a pooling equilibrium; that is, when the equilibrium advertising levels of both show types are equal to some a_p. When consumers see an advertising level a_p they form a belief $\mu(a_p) = \mu_0$, and both show types are worse-off deviating to any other strategy. We are interested in a pooling equilibrium where both types of shows get views with a positive advertising level ($a_p > 0$). Lemma 2.1 finds the lowest equilibrium advertising level at which both show types could pool. Any advertising level below $a_{\mu_0} \doteq \mu_0 a^*_{gW} + (1-\mu_0)a^*_{bW}$ would make viewer expected utilities from the first episode negative, so no consumer would start watching the show.

Lemma 2.1. *In a pooling equilibrium where both show types advertise and attract viewers, the equilibrium advertising level a_p is always greater or equal to $a_{\mu_0} \doteq \mu_0 a^*_{gW} + (1-\mu_0)a^*_{bW}$.*

Proof. Presented in Appendix B.1.2

Depending on model parameters and out-of-equilibrium beliefs, there exist a continuum of pooling equilibria. In Proposition 2.2, we refine these equilibria using the intuitive criterion. We find that the unique advertising level that may survive the intuitive criterion is $a_p = a_{\mu_0}$, and that happens as long as $\mu_0 > \frac{2\gamma}{\theta_H - \theta_L}$.

Proposition 2.2. *In a pooling equilibrium where both show types advertise and attract viewers, then $a_p = a_{\mu_0}$ is the unique equilibrium advertising level that survives the intuitive criterion, and that happens as long as $\mu_0 > \frac{2\gamma}{\theta_H - \theta_L}$.*

Proof. Presented in Appendix B.1.3

Through this Proposition we find that if a pooling equilibrium that survives the intuitive criterion exists, then it must have an equilibrium advertising level $a_p = a_{\mu_0}$. It is important to note that $a_p = a_{\mu_0}$ is between the first best equilibrium advertising levels a^*_{gW} and a^*_{bW}. Here the cost of separating for the high quality show is extremely high, and it's not worth the benefit of incurring such separation.

To illustrate how different parameter sets may lead to a separating equilibrium, to a pooling equilibrium or to none (no market entry or null advertising), we plot these equilibrium regions in a specific example. Figure 2.1 shows the equilibrium regions in $(\mu_0, \theta_H/\theta_L)$ under which we have a pooling or a separating equilibrium that survives the intuitive criterion when the other parameters are $C_W = 0.7$, $\theta_L = 0.26$ and $\gamma = 0.15$. The regions in blank represent parameter spaces in which the optimal advertising levels are not positive, or there is no market entry. As we see in the plot, there is a clear division between pooling and separation. We may find a pooling equilibrium when μ_0 is sufficiently large, and for lower values of μ_0 we find a separating equilibrium. This is because of the threshold we found in Propositions 2.1 and 2.2, $\frac{2\gamma}{\theta_H - \theta_L}$.

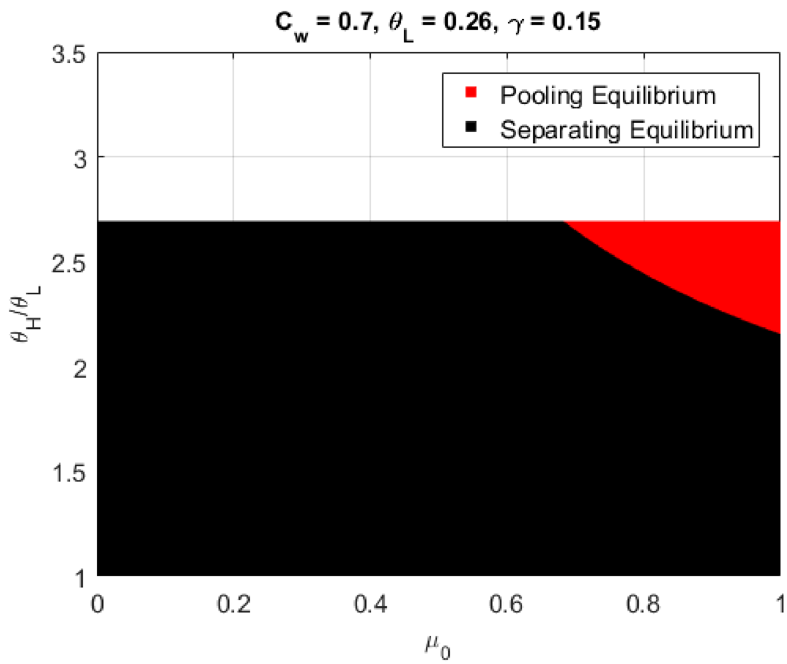

Figure 2.1: Regions in $(\mu_0, \theta_H/\theta_L)$ in which we find a separating or pooling equilibrium in the advertising level that survives the intuitive criterion when the only release timing possibility is sequential releases. The parameters are such that $C_W = 0.7$, $\theta_L = 0.26$ and $\gamma = 0.15$.

2.4 Multiple Release Timings: a Simple Case

In this section we allow the shows to change its release timing strategy by introducing the possibility of simultaneous releases (B). We continue to assume that $g = 1$ and $b = 0$, which implies that $\mathbb{E}[Q_g] = \theta_H$ and $\mathbb{E}[Q_b] = \theta_L$. In section 2.4 we analyze shows with a single episode, where we show that a separating equilibrium cannot exist. In all of the following sections, we analyze a setting with three episodes per show ($N = 3$). In section 2.4.1 we analyze a benchmark case of complete information, in which viewers know the true quality of the show before making the viewing decision. In that section we particularly focus on the scenario in which both show types pool under simultaneous releases. The reason behind this decision is that we find this scenario the most interesting and realistic. A pooling equilibrium on sequential releases as the first best solution would require a low value for the simultaneous advertising efficiency (δ) and the sequential viewing cost (C_W), which we know is unrealistic. In section 2.4.2 we analyze our focal case of asymmetric information, in which the high quality show signals its quality by deviating to sequential releases.

Single Episode Model

A separating equilibrium in which high quality shows choose an advertising level a_g and sequential releases, whereas low quality shows choose an advertising level a_b and simultaneous releases, must satisfy the incentive compatibility constraints. In this scenario, the equilibrium profits from a high and a low quality show are $\pi^g(a_g, W) = 1 - a_g$ and $\pi^b(a_b, B) = \delta - a_b$ respectively. If a low (high) quality show were to mimic the equilibrium advertising level and release timing strategy of the high (low) quality show, viewers would watch the first episode, after which they would find out the true quality show, but because there is only one episode to watch, this information does not affect the revenue. Thus, the mimicking profits are: $\pi^g(a_b, B) = \delta - a_b$ and $\pi^b(a_g, W) = 1 - a_g$.

The incentive compatibility constraints are $\pi^g(a_g, W) = 1 - a_g \geq \pi^g(a_b, B) = \delta - a_b$ and $\pi^b(a_b, B) = \delta - a_b \geq \pi^b(a_g, W) = 1 - a_g$, which may only be satisfied if $a_g = a_b$. Like in the traditional release timing setting where episodes can only be released sequentially, we cannot have a separating equilibrium when there is a unique episode. However, under a multiple period game, we find separation; this shows that the separation in this model is driven by the multi-period setting and not by some other means.

2.4.1 Complete Information With Multiple Episodes

Without information asymmetry, the game becomes simple to solve even with multiple episodes. For any release timing strategy and advertising level, viewers either watch all episodes or none, since they know the true quality of a show before they start watching. We can write firm profits as follows:

$$\pi^g(a, W) = 3\mathbb{1}(\theta_H - C_W + \gamma a \geq 0) - a \tag{2.19}$$

$$\pi^g(a, B) = 3\delta\mathbb{1}(\theta_H - C_B + \gamma a \geq 0) - a \tag{2.20}$$

$$\pi^b(a, W) = 3\mathbb{1}(\theta_L - C_W + \gamma a \geq 0) - a \tag{2.21}$$

$$\pi^b(a, B) = 3\delta\mathbb{1}(\theta_L - C_B + \gamma a \geq 0) - a. \tag{2.22}$$

The optimal advertising level for each show type and release timing strategy comes from choosing either $a = 0$ or the advertising level that that makes null the argument of the corresponding indicator function. Let a_{tX} be the optimal advertising level for a type t show ("good" or "bad") choosing release timing strategy X. Then

$$a_{tX} = \begin{cases} \left[\left(C_X - \mathbb{E}[Q_t]\right)/\gamma\right]^+ & \text{if } \pi^t\left(\left[\left(C_X - \mathbb{E}[Q_t]\right)/\gamma\right]^+, X\right) \geq 0 \\ 0 & \text{otherwise.} \end{cases} \tag{2.23}$$

These advertising levels maximize the expected profit for each show type and release timing strategy, because they are the minimum advertising levels that guarantee views. We restrict our parameter set such that both show types could attract viewers under both release timing strategies with a positive advertising level; $a_{gB} > 0$, $a_{bB} < 3\delta$, $a_{gW} > 0$ and $a_{bW} < 3$. Under this parameter set the profit functions simplify to $\pi^t(a_{tW}, W) = 3 - a_{tW}$ and $\pi^t(a_{tB}, B) = 3 - a_{tB}$ where $a_{tX} = \left(C_X - \mathbb{E}[Q_t]\right)/\gamma$ for $(t, X) \in \Big\{\{g, b\}, \{W, B\}\Big\}$. Lemma 2.2 establishes the regions under which we may see both firm types choosing sequential releases, simultaneous releases or both. In order to have pooling on simultaneous releases, the difference in the equilibrium advertising levels between simultaneous releases and sequential releases, $\frac{(C_W - C_B)}{\gamma}$, must be greater than the extra revenue obtained from the increased advertising efficacy in sequential releases, $3(1 - \delta)$. Depending on how these two terms compare to each other, we will be in a different regime; Figure 2.2 shows the different regimes.

Lemma 2.2. *Under complete information in a parameter set where both show types could attract viewers under both release timing strategies with a positive advertising level, then:*

- *If* $\frac{1}{\gamma}(C_W - C_B) < 3(1 - \delta)$ *both show types would choose sequential releases (W).*

- *If $\frac{1}{\gamma}(C_W - C_B) = 3(1-\delta)$ both show types are indifferent between choosing sequential releases (W) and simultaneous releases (B), so we may have separation or pooling in release timing.*

- *If $\frac{1}{\gamma}(C_W - C_B) > 3(1-\delta)$ both show types would choose simultaneous releases (B).*

Proof. Presented in Appendix B.1.4

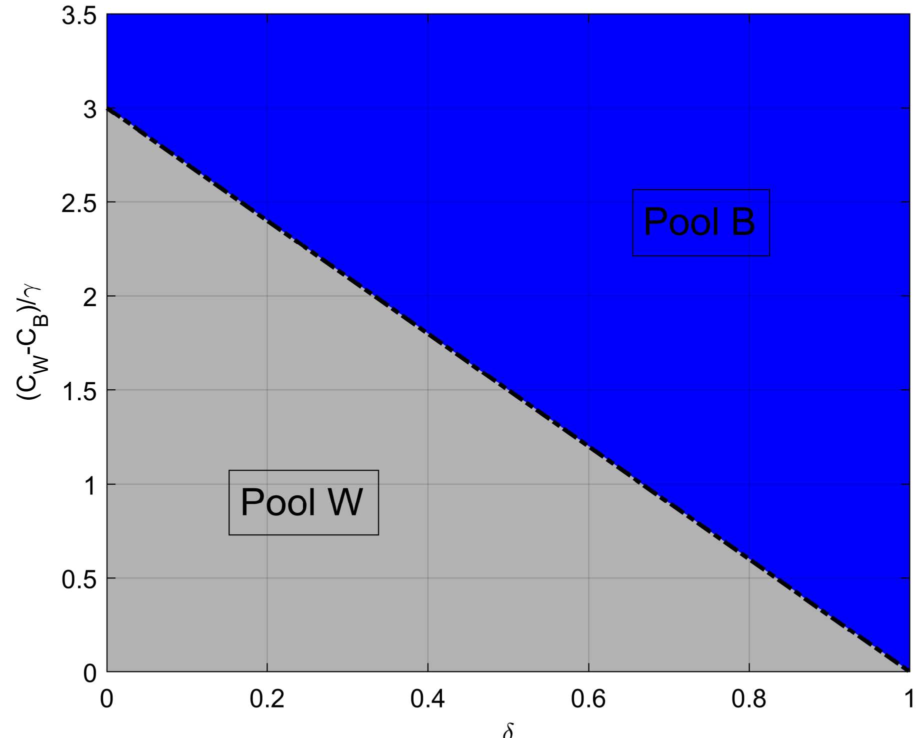

Figure 2.2: Regions in $(\delta, (C_W - C_B)/\gamma)$ with complete information under which we find pooling to sequential linear or simultaneous releases, as well as where we may find both or separating release timing strategies (the black line).

As the value of δ increases (compared to $(C_W - C_B)/\gamma)$), there is less cost to choosing simultaneous releases, and we see that shows switch from pooling on sequential releases to simultaneous releases. We now analyze pooling on simultaneous releases as our first best solution, and we find the parameter set under which both release timing strategies are profitable for both show types with positive advertising levels.

Pooling on Simultaneous Releases

We analyze the equilibria in which low quality shows choose simultaneous releases and advertising level $a_b^* \doteq a_{bB} > 0$, while high quality shows choose the same release timing strategy but advertising level $a_g^* \doteq a_{gB} > 0$. We do so by setting the incentive compatibility constraints to have a pooling equilibrium around Simultaneous releases:

$$\pi^g(a_{gW}, W) < \pi^g(a_g^*, B) \text{ and} \tag{2.24}$$

$$\pi^b(a_{bW}, W) < \pi^b(a_b^*, B), \tag{2.25}$$

where a_{gW} and a_{bW} are defined according to Equation (2.23).

Furthermore, in order to have positive advertising levels in equilibrium, we must have that both firms make non-negative revenue: $\pi^g(a_g^*, B) \geq 0$ and $\pi^b(a_b^*, B) \geq 0$. From Lemma 2.2 we can rewrite the incentive compatibility constraints as $\frac{1}{\gamma}(C_W - C_B) \geq 3(1-\delta)$.

As described previously, we are interested in the case where both show types would be able to attract viewers and make profit under any release timing strategy, then $a_{tX} = (C_X - \mathbb{E}[Q_t])/\gamma$ for $(t, X) \in \Big\{\{g, b\}, \{W, B\}\Big\}$ with $a_g^* > 0$, $a_b^* < 3\delta$, $a_{gW} > 0$ and $a_{bW} < 3$. We require that both firms could make profit under both release timing strategies with a positive advertising level, so that deviations from release timing strategies are still profitable and advertising is needed. We can finally write the parameter set that yields our first best solution as:

$$\frac{1}{\gamma}(C_W - C_B) \geq 3(1-\delta)$$

$$a_{gW} > 0 \text{ and } a_{bW} < 3$$

with $a_g^* = \frac{C_B - \theta_H}{\gamma} > 0$ and $a_b^* = \frac{C_B - \theta_L}{\gamma} \in (0, 3\delta]$ as equilibrium advertising levels with simultaneous releases for the high quality and low quality shows respectively. The equilibrium profit functions under this equilibrium are

$$\pi^g(a_g^*, B) = 3\delta - \frac{1}{\gamma}(C_B - \theta_H) \text{ and}$$

$$\pi^b(a_b^*, B) = 3\delta - \frac{1}{\gamma}(C_B - \theta_L)$$

according to Equations 2.20 and 2.22. Now that we have identified the parameter set that yields our first best solution, we may analyze what happens under incomplete information in the same parameter set, and particularly whether we may have deviations from these first best strategies.

2.4.2 Incomplete Information With Multiple Episodes

When we have information asymmetry, high quality shows may want to change their first best strategies in order to inform viewers about their high quality, incurring a cost that low quality shows may not afford. Sticking with the first best solution might not be a good strategy for high quality shows, because the low quality shows might be better off mimicking their behavior. The next subsection analyzes the separating equilibrium in which high quality shows separate to sequential releases, while low quality shows remain with their first best strategy. As discussed earlier, we focus on the parameter set that yields a pooling equilibrium on simultaneous releases as our first best solution.

Separating Equilibrium on Release Timing

Our focus is on separating equilibria, in such an equilibria the show's type is revealed by the chosen strategies. As soon as the firm chooses an equilibrium strategy, viewers will have absolute beliefs about the show's type. We look into equilibria where high quality shows choose a sequential release timing (W) with advertising level a_g, while low quality shows choose the all-at-once release timing (B) with equilibrium advertising level a_b. We say that the high quality show is separating from the first best solution (under which its equilibrium strategy is $\{a_g^*, B\}$), incurring a cost by choosing a strategy which would be not-optimal under perfect information. In Lemma 2.3 we show that in equilibrium the low quality show would not have any incentives to deviate from its first best solution with advertising level $a_b^* > 0$, so we have that $a_b = a_b^* = \frac{1}{\gamma}(C_B - \theta_L)$.

Lemma 2.3. *If the first best solution for a low quality show is (a_b^*, B), under incomplete information any equilibrium strategy in a separating equilibrium will be such that the low quality show still chooses its first best solution (a_b^*, B).*

Proof. Presented in Appendix B.1.5

In equilibrium both firm types advertise, attract viewers and make non-negative profit; thus, we have that $\pi^g(a_g, W) = 3 - a_g \geq 0$ and $\pi^b(a_b, B) = 3\delta - a_b \geq 0$. Lemma 2.4 shows that in order to have a separating equilibrium which attracts viewers, a_g must be greater or equal to $a_W \doteq \frac{1}{\gamma}(C_W - \theta_H)$. If a_g was less than a_W, then we would have that the viewers' expected utility from the first episode of a high quality show is negative, which generates no views and leads to a negative profit function. However, the strategy of choosing the equilibrium advertising level would be dominated by not advertising at all, which breaks the equilibrium conditions.

Lemma 2.4. *If there exists a separating equilibrium where high quality shows choose* (a_g, W) *with* $a_g > 0$ *while low quality shows choose their first best solution, then* $a_g \geq a_W \doteq \frac{1}{\gamma}(C_W - \theta_H)$.

Proof. Presented in Appendix B.1.6

Through Lemmas 2.3 and 2.4 we find that a separating equilibrium in release timing strategies which attracts viewers must satisfy $a_b = a_b^*$ and $a_g \geq a_W$. Now we look into the profit functions of each type mimicking the other to set up the incentive compatibility constraints.

If a high quality show were to mimic the behavior of a low quality show, their profit would be the same as a low quality show would obtain in equilibrium

$$\pi^g(a_b, B) = 3\delta - a_b^* \tag{2.26}$$

since a low quality show would already obtain 3 views. Whereas if a low quality show were to mimic the behavior of a high quality show, viewers would watch the first episode, after which they would realize that the show is bad and then there are two possibilities:

(A) The low quality is good enough, so that viewers would still watch the remaining episodes even after finding out that the show is of low quality. This happens when $\theta_L - C_W + \gamma a_g \geq 0$.

(B) The low quality is not good enough, and viewers stop watching the show after watching episode 1 and finding out that the show is of low quality. This happens when $\theta_L - C_W + \gamma a_g < 0$.

We can write the off-equilibrium profit of a low quality show mimicking the behavior of a high quality show as follows:

$$\pi^b(a_g, W) = 1 - a_g + 2\mathbb{1}(a_g \geq a_{bW}), \tag{2.27}$$

where $a_{bW} = \frac{1}{\gamma}(C_W - \theta_L)$ is defined in Equation (2.23). When $a_g \geq a_{bW}$ we are in case (A) and Equation 2.27 becomes $\pi^b(a_g, W) = 3 - a_g$, whereas when $a_g < a_{bW}$ we are in case (B) and Equation 2.27 becomes $\pi^b(a_g, W) = 1 - a_g$.

We may now write the incentive compatibility constraints, which removes incentives for each show type to mimic the equilibrium behavior of the other show's type:

$$\pi^g(a_g, W) = 3 - a_g \geq 3\delta - a_b = \pi^g(a_b, W) \text{ and} \tag{2.28}$$

$$\pi^b(a_b, B) = 3\delta - a_b \geq 1 - a_g + 2\mathbb{1}(a_g \geq a_{bW}) = \pi^b(a_g, W)). \tag{2.29}$$

We rewrite the incentive compatibility constraints with the following inequalities: $3 - a_g \geq 3\delta - a_b \geq 1 - a_g + 2\mathbb{1}(a_g \geq a_{bW})$ and condition on whether $a_g \geq a_{bW}$ in case A or $a_g < a_{bW}$ in case B.

Case A: $a_g \geq a_{bW}$, then the incentive compatibility constraints become an equality constraint in which both show types obtain the same profit

$$3 - a_g = 3\delta - a_b.$$

Then the conditions for a separating equilibrium can be written as

$$3 \geq a_g = a_b + 3(1 - \delta) \geq a_{bW} \text{ and } a_g, a_b > 0, \tag{2.30}$$

where $a_b = \frac{1}{\gamma}(C_B - \theta_L)$ according to Lemma 2.3. The first inequality ensures the equilibrium revenue from high quality shows is non-negative ($3 - a_g \geq 0$). The second inequality meets the incentive compatibility constraints at equality, which also implies non-negative revenue for the equilibrium strategies of the low quality show. The second inequality ensures we meet the conditions from Lemma 2.4 ($a_g \geq a_W$), because $a_{bW} > a_W$, and at the same time we are in the set corresponding to case A ($a_g \geq a_{bW}$). In this parameter set, both show types obtain the same equilibrium profit; this is because the low quality show obtains no penalty in mimicking the high quality one. We now show that this equilibrium cannot survive the intuitive criterion.

Take the first best solution $\{(a_g^*, B), (a_b^*, B)\}$ and a separating equilibrium $\{(a_g, W), (a_b^*, B)\}$ where $a_g \geq a_{bW}$. Then, according to this equilibrium $a_g = 3(1 - \delta) + a_b^*$. Suppose there is a deviation to $(a_b^* - \epsilon, B)$ with an $\epsilon > 0$. There exists $\epsilon > 0$ such that for no out-of-equilibrium belief, the low quality show would be better-off deviating strategy, because it would only obtain one view. On the other hand, for some out-of-equilibrium beliefs, the high quality show could be better-off when ϵ is sufficiently small and $a_b^* - \epsilon > a_g^*$, then

$3\delta - a_b^* + \epsilon = 3 - a_g + \epsilon > 3 - a_g$. Thus, consumers should not believe that such a deviation could come from the low quality show, but from the high quality show. Given this, the high quality show is better-off deviating, so this equilibrium does not survive the intuitive criterion.

Case B: $a_g < a_{bW}$, then the incentive compatibility constraints become

$$3 - a_g \geq 3\delta - a_b \geq 1 - a_g$$

This case leads to a continuum of advertising levels for which the high quality show could signal its quality. From the incentive compatibility constraints we get $3(1-\delta) + a_b \geq a_g \geq 1 - 3\delta + a_b$. Including constraints for positive advertising levels and non-negative profit functions, the set that leads to this separating equilibrium is:

$$3(1-\delta) + a_b \geq a_g \geq \max\{1 - 3\delta + a_b, a_W\}, \quad a_{bW} > a_g \text{ and } 3\delta \geq a_b. \tag{2.31}$$

The upper bound on a_g ensures that the incentive compatibility constraint for a "good" show holds, whereas $a_{bW} > a_g$ ensures that we are in case B. The lower bound on a_g satisfies the incentive compatibility conditions for a "bad" show and also ensures that Lemma 2.4 ($a_g \geq a_W$) is satisfied. Finally, the upper bound on $3\delta \geq a_b$ ensures that the equilibrium revenue from "bad" shows is non-negative, which, in conjunction with the other constraints, ensure non-negative revenue for the equilibrium strategy of a "good" show. Throughout this section we focus on this equilibrium.

From now on we focus on Case B in which $a_g < a_{bW}$. A continuum of perfect Bayesian equilibria depending on customer's out-of-equilibrium beliefs can be obtained from case B (Inequalities (2.31)). We show here that any separating equilibrium with sequential releases for the high quality show and advertising level $a_g \in \big(\max\{1 - 3\delta + a_b, a_W\},\, \min\{3(1-\delta) + a_b^*, a_{bW}\},\big]$ can be eliminated by the intuitive criterion Cho and Kreps (1987). Suppose that the high quality show deviates from (a_g, W) to some strategy (a, W) with $a \in \big[\max\{1 - 3\delta + a_b,\, a_W\}, a_g\big)$. This new advertising level a with release timing strategy W is equilibrium-dominated for the low quality show, regardless of what customers believe about the show's type. This is because this new advertising level a satisfies the incentive compatibility constraints. Therefore, customers should not believe that the show which voluntarily made such a deviation can be the low quality type with a positive probability. Consequently, the high quality show indeed prefers deviating to such an advertising level, as long as customers believe that such deviation cannot come from the low quality show. That is, the equilibria involving advertising level $a_g \in \big(\max\{1-3\delta+a_b, a_W\}, \min\{3(1-\delta)+a_b^*, a_{bW}\},\big]$ fails the intuitive criterion, leaving $a_g = \max\{1-3\delta+a_b, a_W\}$ as the only possible advertising level that may survive this refinement. In Lemma 2.5 we show that under which conditions $a_g = \max\{1-3\delta+a_b, a_W\}$ survives the intuitive criterion to deviations to sequential releases. This Lemma tries to do the same as Proposition 2.1 does for the traditional release timing setting, but now for deviations that can switch release timing strategies (i.e., from simultaneous releases to sequential releases).

Lemma 2.5. *A separating equilibrium in which high quality shows advertise $a_g > 0$ and choose sequential releases whereas low quality shows stick with their first best solution $(a_b^* > 0, B)$ survives the intuitive criterion for deviations to sequential releases, as long as $a_g = \max\{a_W, 1+a_b-3\delta\}$ and either; $\mu_0 < \gamma\frac{3\delta-1}{\theta_H-\theta_L}$ and $a_W < 1+a_b-3\delta$, or $a_W \geq 1+a_b-3\delta$.*

Proof. Presented in Appendix B.1.7

Clearly, among all separating equilibria $a_g = \max\{1-3\delta+a_b, a_W\}$ gives the greatest profit for the high quality show. We still need to check whether this equilibrium survives the intuitive criterion for any deviation to simultaneous releases (a, B); in Lemma 2.6 we find the conditions under which it does.

Lemma 2.6. *Let $\bar{a}_g \doteq 3\delta - \pi^g(a_g, W) = a_g - 3(1-\delta)$. If both release timing strategies are profitable for either show type, and we have a first best solution where both show types choose simultaneous releases with a positive advertising level, the separating equilibrium in which high quality shows choose strategy (a_g, W) while low quality shows choose strategy (a_b^*, B), with $a_g = \max\{1-3\delta+a_b^*, a_W\}$, survives the intuitive criterion for deviations to simultaneous releases as long as*

$$\mu_0 < \gamma\frac{a_b^* - \bar{a}_g}{(\theta_H - \theta_L)} \text{ and } \delta - \bar{a}_g \geq 3\delta - a_b^*. \tag{2.32}$$

Proof. Presented in Appendix B.1.8

The intuition behind Lemma 2.6 is that we need the high quality shows to separate from their first best solution described in the previous section. In order for such an equilibrium to survive the intuitive criterion, we need the existence of out-of-equilibrium beliefs that make the low quality show better-off for any strategy (a, B) under which the high quality show could be better-off. If there exists a strategy (a, B) that is equilibrium-dominated for the low quality show, but there exists off-equilibrium beliefs such that the high quality show is better off, then our equilibrium would not survive the intuitive criterion. This Lemma ensures that such strategies do not exist for simultaneous releases. Furthermore, it ensures that the belief viewers give to the strategies (a, B) in which both show types could be better off, μ_0, is sufficiently low so that viewers don't start watching the show in case either show deviates. Lemma 2.5 performs the same task but for deviations to linear sequential releases. In Proposition 2.3, we state the parameter set under which we have a separating equilibrium on release timing strategies that survives the intuitive criterion. It comes from merging the last two Lemmas.

Proposition 2.3. *When both release timing strategies are profitable for either show type, and we have a first best solution where both show types choose simultaneous releases with a positive advertising level, then there exists a separating equilibrium on release timing strategies* $\{(a_g, W), (a_b, B)\}$ *that survives the intuitive criterion as long as* $a_g = \max\{a_W, 1 + a_b - 3\delta\} < a_{bW}$, $a_b = a_b^*$, $a_g \leq 3 + a_b - 3\delta$, $\mu_0 < \gamma\frac{a_b^* - \bar{a}_g}{(\theta_H - \theta_L)}$, $a_b^* - \bar{a}_g \geq 2\delta$ *and*

$$\mu_0 < \gamma\frac{3\delta - 1}{(\theta_H - \theta_L)} \text{ if } a_W < 1 + a_b - 3\delta.$$

Proof. Presented in Appendix B.1.9

In order to provide intuition from the separating equilibrium and the impact of having a release timing decision, we provide the following example. Consider the case where $\theta_H = 1.85$, $\theta_L = 0.6$, $\gamma = 0.55$, $C_B = 1.9$, $C_W = 2.15$, $\delta = 0.9$, $\mu_0 = 0.1$, $g = 1$ and $b = 0$. First, we study what happens when the only release timing strategy possible is sequential releases (W). In this situation, the first best solution for a high and low quality show is $a_g^{*W} = 0.5455$ and $a_b^{*W} = 2.8182$ respectively. When information becomes incomplete, the high quality show separates to $a_g^W = 0.8182$ in order to satisfy the incentive compatibility constraints and signal its quality. The equilibrium profit for a high and a low quality show under incomplete information is 2.1818 and 0.1818 respectively. Now we analyze what happens when shows may also be released simultaneously. The parameter set is such that for the first best solution we have pooling on simultaneous releases, the high quality show chooses ($a_g^* = 0.0909, B$) whereas the low quality show chooses ($a_b^* = 2.3636, B$). When information is incomplete, the high quality show separates to ($a_g = 0.6636, W$), changing its release timing strategy to W. The equilibrium profit for the high quality show becomes 2.3364, whereas the low quality one obtains 0.3364. Under traditional sequential releases, the high quality show may separate by incurring a cost. When simultaneous releases are allowed, both firms types are better off. The low quality enjoys a release timing strategy for which they become more profitable; this relaxes the incentive compatibility constraint, and now the high quality show may reduce their advertising level and still signal its quality ($a_g^W = 0.8182$ vs $a_g = 0.6636$). The separating equilibrium comes from having the least cost separation under the linear sequential release timing. If the high quality show were to separate using simultaneous releases, their overall profit would be less. It is interesting to notice that a television media company with a high quality show has incentives to provide this highly profitable release timing strategy for lower quality shows, because this would reduce the advertising level threshold at which lower quality shows would be interested in mimicking them. We plot the different advertising parameters obtained from this model in Figure 2.3.

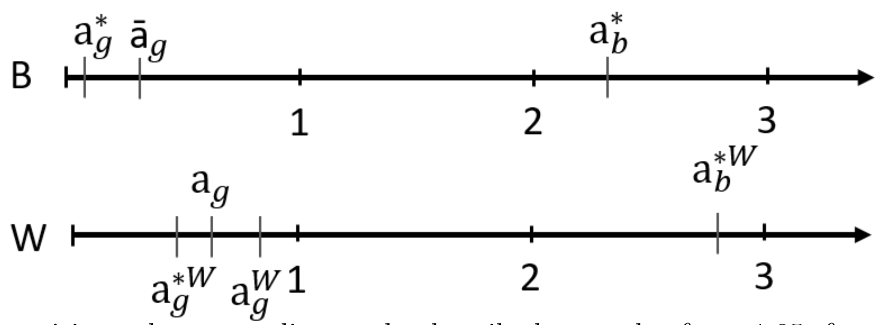

Figure 2.3: Advertising values according to the described example: $\theta_H = 1.85$, $\theta_L = 0.6$, $\gamma = 0.55$, $C_B = 1.9$, $C_W = 2.15$, $\delta = 0.9$, $\mu_0 = 0.1$, $g = 1$ and $b = 0$. $a_g^{*W} = 0.5455$, $a_b^{*W} = 2.8182$, $a_g^W = 0.8182$, $a_g^* = 0.0909$, $a_b^* = 2.3636$, $a_g = 0.6636$ and $\bar{a}_g = 0.3636$.

a_g^* is the first best advertising level for a high quality show. $\bar{a}_g$ is the maximum advertising level under simultaneous releases at which the high quality show could be better off by deviating from their separating equilibrium strategy (a_g, W). a_b^* is the first best and the separating equilibrium strategy for the low quality show, (a_b^*, W). a_g is the equilibrium advertising level for the high quality show. a_g^{*W} and a_b^{*W} are the first best solution for the high and low quality shows, respectively, in the traditional sequential releases setting. a_g^W is the advertising level at which the high quality show separates to when information is imperfect in the traditional sequential releases setting.

Note, if the high quality show were to deviate from (a_g, W), any advertising level a other than $a \in [a_g^*, \bar{a}_g]$ under simultaneous releases is dominated by its equilibrium strategy. If consumer beliefs were high enough, the high quality show would be better-off deviating to an $a \in [a_g^*, \bar{a}_g)$ in simultaneous releases, but given that $\delta - a \geq \delta - \bar{a}_g = 0.5364 > 3\delta - a_b^* = 0.3364$, the low quality show would also be better off. Thus, consumer beliefs for any advertising level in that region would be set to μ_0, and since $\mu_0 = 0.1 < \gamma\frac{a_b^* - \bar{a}_g}{(\theta_H - \theta_L)} = 0.8800$, viewers would not start watching the show.

To observe what happens when we perturb the parameters of the previous example, we show in Figure 2.4 the regions in $(\delta, \frac{C_W - C_B}{\gamma})$ under which we obtain separation on release timing strategies from the initial first best solution that pools on simultaneous releases. As one can see there exists a region under which we have separation to W (red), while on another portion separation may not exist (blank space), or it occurs through advertising on simultaneous releases (black).

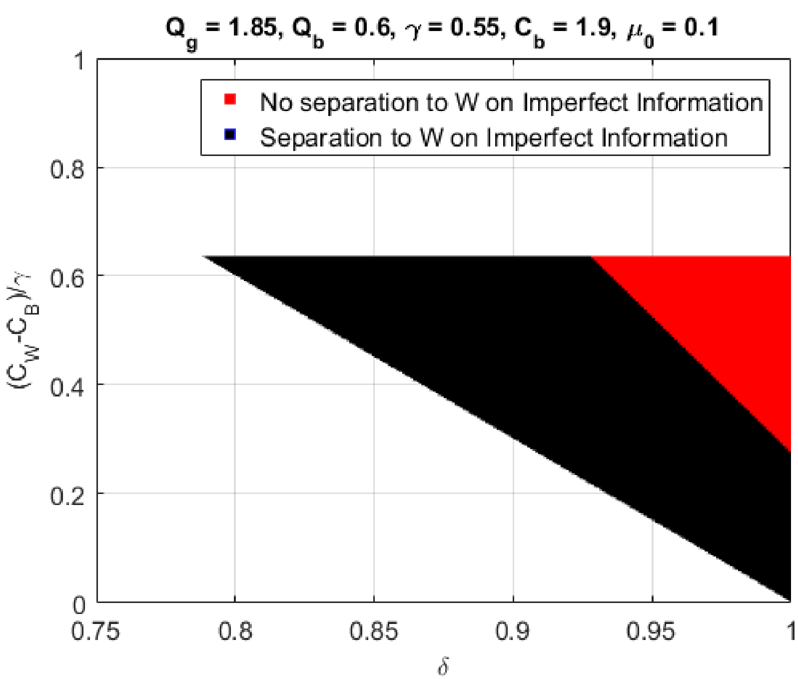

Figure 2.4: Regions in $(\delta, \frac{C_W - C_B}{\gamma})$ under which the high quality show separates to sequential releases or not, from a first best solution where both shows pool on simultaneous releases.

So far we have analyzed the simple case in which $g = 1$ and $b = 0$. In the next section we analyze the general case where we relax this assumption.

2.5 Multiple Release Timings: General Model

The next step is to analyze the general model, which incorporates the randomness in the perceived quality of each episode. Randomness gives more space for low quality shows to mimic the behavior of high quality shows, since viewers may now be uncertain about the show's true quality even after watching all episodes within the show. This would happen if the quality draw for a viewer for all three episodes was θ_M. We start with the same first best solution as in the previous section, where both shows pool on simultaneous releases. Then we analyze the existence of a separating equilibrium in which it is always better for high quality shows to choose sequential releases, and for low quality shows to choose simultaneous releases.

We look for separating equilibria in which high quality shows choose strategy (a_g, W) whereas low quality shows choose (a_b, B), but for the situation in which, under complete information, both show types pool on simultaneous releases, as described in section 2.4.1.

If a separating equilibrium of this type exists, the expected payoff of a low quality show mimicking the behavior of a high quality one can be obtained using Lemma 2.7. The intuition behind this lemma is that if consumers see a signal (a_g, W), their initial belief about the show being of high quality remains set to 1. However, if a viewer watches the first episode, she will obtain a random quality realized from Q_b, taking values θ_L or θ_M. If the realization is θ_M, then the belief about the show being good will remain absolute, due to the Bayesian update. However, if the realization is θ_L, then the belief about the show being of high quality would turn to zero, because θ_L is not a possible outcome for the quality of an episode of a high quality show. It could be the case that after realizing that the show mimicking the behavior of the high quality one is of low quality, the consumer keeps on watching. The different branches of $\pi^b(a, W)$ in Lemma 2.7 consider these two possibilities. i) If $a_g \geq \frac{C_W - \mathbb{E}[Q_b]}{\gamma}$ viewers would continue watching the show regardless of when they determine that the show is of low quality; ii) If $\frac{C_W - \mathbb{E}[Q_b]}{\gamma} > a_g$ viewers would stop watching as soon as they determine that the show is of low quality.

Lemma 2.7. *In a separating equilibrium in which high quality shows choose advertising level $a_g \geq a_W$ and sequential releases, whereas low quality shows choose simultaneous releases and advertising* *level a_b, the expected profit of a low quality show mimicking the behavior of a high quality one is:*

$$\pi^b(a_g, W) = \begin{cases} 3 - a_g & \text{if } a_g \geq \frac{C_W - \mathbb{E}[Q_b]}{\gamma} \\ (1 + b + b^2) - a_g & \text{if } \frac{C_W - \mathbb{E}[Q_b]}{\gamma} > a_g \end{cases} \tag{2.33}$$

Proof. Presented in Appendix B.1.10

We may now write the incentive compatibility constraints which remove incentives for each show type to mimic the equilibrium behavior of the other show type. We consider two separate cases that come from the two branches of Equation (2.33). For simplicity, we use the following notation $a_{bW} \doteq \frac{1}{\gamma}(C_W - \mathbb{E}[Q_b])$.

Case A: $a_g \geq a_{bW}$

$$\text{Good: } \pi^g(a_g, W) = 3 - a_g \geq 3\delta - a_b = \pi^g(a_b, B) \text{ and}$$

$$\text{Bad: } \pi^b(a_b, B) = 3\delta - a_b \geq 3 - a_g = \pi^b(a_g, W).$$

In case A the incentive compatibility constraints imply that the equilibrium profit for both show types is the same, $3 - a_g = 3\delta - a_b$, then $a_g = 3(1 - \delta) + a_b \geq a_{bW} > a_{gW}$. We now show that these equilibria does not survive the intuitive criterion.

From Lemma 2.3 we have that $a_b = a_b^*$ in order for these equilibria to survive the intuitive criterion. Then take the first best solution $\{(a_g^*, B), (a_b^*, B)\}$ and a separating equilibrium $\{(a_g, W), (a_b^*, B)\}$ where $a_g = 3(1-\delta) + a_b^* \geq a_{bW}$. Suppose there is a deviation to $(a_b^* - \epsilon, B)$ with an $\epsilon > 0$. There exists $\epsilon > 0$ such that for no out-of-equilibrium belief the low quality show would be better-off deviating to such a strategy. In the best case scenario, the expected profit of such deviation for the low quality show would be $(1 + b + b^2)\delta - a_b^* + \epsilon$, which is less than $3\delta - a_b^*$ for a sufficiently small ϵ. On the other hand, a high quality show could be better-off by performing such deviation. With a sufficiently small ϵ, we have that $a_b^* - \epsilon \geq a_g^*$, hence the expected profit would be $3\delta - a_b^* + \epsilon$, which is greater by ϵ compared to its equilibrium strategy. Consumers should not believe that the deviation to $(a_b^* - \epsilon, B)$ comes from a low quality show, because it is an equilibrium dominated strategy for that show type. Depending on out-of-equilibrium beliefs, the high quality show is better off with such deviation. Therefore, consumers will assess that the high quality show is the one performing the deviation, and then this equilibrium does not survive the intuitive criterion.

Case B: $a_{bW} > a_g$

$$\text{Good: } \pi^g(a_g, W) = 3 - a_g \geq 3\delta - a_b = \pi^g(a_b, B) \text{ and}$$
$$\text{Bad: } \pi^b(a_b, B) = 3\delta - a_b \geq 1 + b + b^2 - a_g = \pi^b(a_g, W).$$

The incentive compatibility constraint translates to $1 + b + b^2 - 3\delta + a_b \leq a_g \leq 3(1-\delta) + a_b$. We finally can set the conditions to have a separating equilibrium as:

$$3(1-\delta) + a_b \geq a_g \geq \max\{1 + b + b^2 - 3\delta + a_b, a_W\}, \quad a_{bW} > a_g \text{ and } 3\delta \geq a_b,$$

where $a_W = \frac{1}{\gamma}(C_W - \theta_H)$. The upper bound on a_g ensures the incentive compatibility constraint for a "good" show holds, whereas $a_{bW} > a_g$ ensures that we are in case B. The lower bound on a_g satisfies the incentive compatibility conditions for a low quality show and also ensures that Lemma 2.4 ($a_g \geq a_W$) is satisfied. Finally, the upper bound on $3\delta \geq a_b$ ensures that the equilibrium revenue from "bad" shows is non-negative, which in conjunction with the other constraints ensure non-negative revenue for the equilibrium strategy of a "good" show. Throughout this section we focus on this equilibrium.

We now focus exclusively on case B, and we refine it using the intuitive criterion. As in the previous section, we have that the unique equilibrium that may survive the intuitive criterion is $a_g = \max\{1 + b + b^2 - 3\delta + a_b, a_W\}$. If a_g was greater than $\max\{1 + b + b^2 - 3\delta + a_b, a_W\}$, that would leave the region $[\max\{1 + b + b^2 - 3\delta + a_b, a_W\}, a_g)$ equilibrium dominated for the low quality show, whereas it would dominate the equilibrium strategy for the high quality show. Then it would never be able to survive the intuitive criterion. Therefore, we need to find under which conditions $a_g = \max\{1 + b + b^2 - 3\delta + a_b, a_W\}$ survives the intuitive criterion; in Proposition 2.4 we find such conditions.

Proposition 2.4. *When both release timing strategies are profitable for either show type, and we have a first best solution where both show types choose simultaneous releases with a positive advertising level, then there exists a separating equilibrium on release timing strategies* $\{(a_g, W), (a_b, B)\}$

that survives the intuitive criterion as long as $a_g = \max\{a_W, 1 + b + b^2 + a_b - 3\delta\} < a_{bW}$, $a_b = a_b^*$, $a_g \leq 3 + a_b - 3\delta$, $\mu_0 < \gamma \frac{a_b^* - \bar{a}_g}{(\mathbb{E}[Q_g] - \mathbb{E}Q_b)}$, $a_b^* - \bar{a}_g \geq 2\delta - b - b^2$, *and* $\mu_0 < \gamma \frac{3\delta - 1}{(\mathbb{E}[Q_g] - \mathbb{E}Q_b)}$ *if* $a_W < 1 + a_b - 3\delta$.

Proof. Presented in Appendix B.1.11

Proposition 2.4 describes the parameter set under which we see separation in release timing strategies, surviving the intuitive criterion from the first best pooling equilibrium on simultaneous releases. In this equilibrium a high quality show signals its quality through sequential releases and a specific advertising level a_g, while a low quality show sticks with its first best strategy. An interesting aspect of this model is that there exist sample paths in which a low quality show is better off by mimicking the behavior of a high quality show, but in expectation it is better staying with its equilibrium strategy. The randomness in the episode experienced quality by viewers, creates different sample paths of quality draws per episode, in which it might make a viewer take too long to determine the show's true quality $(\theta_M, \theta_M, \theta_L)$, or not at all $(\theta_M, \theta_M, \theta_M)$. As such, this gives an incentive for a low quality show to behave like a high quality one, given that there exists the possibility that the low quality show can hide its quality from viewers. Consequently, under this setting, for separation to occur, we must have a greater advertising efficiency threshold, δ, than in the classical signaling game without randomness described in Section 2.4. Another way of seeing this is by checking the separation cost that the high quality show must incur in order to signal its high quality. In Section 2.4 we found that $a_g = \max\{a_W, 1 + a_b - 3\delta\} < a_{bW}$, whereas in this section we have that $a_g = \max\{a_W, 1 + b + b^2 + a_b - 3\delta\} < a_{bW}$. Therefore, the difference in separation costs could be as high as $b + b^2$. Consequently, with randomness, high quality shows must incur a higher seperation cost in order to make the actions of the low quality showing (mimicing the high quality show) less profitable than the low quality show's first best solution

2.6 Extension: Urge to close

We now analyze qualitatively the effects of adding a new term in the viewer's utility function, which increases as the episode consumed gets closer to the season finale. This term captures a viewer's urge to finish a season, and it enters the model as an addition to the base quality. Under complete information, we expect the advertising levels to be reduced compared to the original model. This is because the urge to close provides an extra base utility, which may be seen as an increase in quality. When we have imperfect information, lower quality shows have greater incentives to mimic the behavior of higher quality shows. This is because when a low quality show mimics a high quality one, viewers may stop watching the show if they realize its true quality early on, whereas they might keep on watching it if they realize its true quality later. This consumer behavior was not captured by the previous model, and it increases the expected profit low quality shows get from mimicking the behavior of the high quality shows. Given this, if separation were to exist, the separation costs for the higher quality content will be larger, similar to the case where randomness in show quality is introduced.

2.7 Conclusion

We have shown that under certain scenarios, high quality shows may use a sequential release timing in conjunction with a suitable advertising level to credibly inform consumers about their show quality. Furthermore, we analyze a situation in which, under complete information, both shows pool on simultaneous releases, whereas when information is private, the high quality show uses the linear sequential releases channel in order to achieve minimum cost separation. The release timing strategy affects consumers in their viewing cost, whereas the advertising level persuades viewers to watch the show. We find that there exists a separating equilibrium surviving the intuitive criterion under which release timing strategies signal quality. Furthermore, we find that the introduction of the simultaneous release timing reduces the advertising level high quality shows need to incur in order to signal their quality. Even though the high quality show does not use the simultaneous release timing, by providing a more profitable release timing strategy for low quality shows, they are better-off. This is because the incentive compatibility constraints are relaxed, and high quality shows may now reduce their equilibrium advertising level under sequential releases. Moreover, in our model the traditional television based media companies, such as NBCUniversal, Fox, and ViacomCBS, are tasked with selecting which release timing strategy to use for new content, and thereby determine what show quality types are most likely to be binged. An interesting result of our analysis is that in equilibrium binge watching occurs with lower quality shows, not high quality. Fascinatingly, this result has direct parallels to work where consumers are found to binge low quality foods and/or beverages Boggiano et al. (2014).

Our results also have important managerial implication; because release timing may signal quality, it is beneficial for a firm to find a way in which to open new channels that are profitable for its lower quality content. By doing this, the firm reduces the necessity of costly signaling through other mechanisms (i.e., advertising) for their higher quality content, and enjoys greater profits on all quality levels.

Finally, as we noted in the Introduction, research on binge consumption is in the nascent stage. A natural and important issue to study are the strategies associated with traditional television based media that now have online streaming channels, in the context of binge consumption. Static signaling models are challenging, with multi-period models even more complex. Thankfully, even with the added complexity, we have been able to achieve tractability and meaningful insights. Adding competition will enormously complicate our analysis and will remain an important but challenging future research topic.

Chapter 3

Optimal Timing of Home Video Releases: A Dynamic Model of Movie Distribution

3.1 Introduction

Home video release timing is arguably the most important decision for movie distribution; as a market segmentation strategy, it separates two different consumer experiences, theaters and home videos. For the last few decades the home video window, the time between theater market exit and home video release, has been decreasing. At the same time the technological quality of the home video experience has improved, from VHS in the 1990s to 4K Ultra HD today. This poses a question for the movie industry; is it optimal to decrease the home video window as the technological quality of home videos increase? Furthermore, as several technologies of home videos are available at the same time, e.g. DVD and Bluray during the 2010's, which technological quality home video should be released first in order to maximize studio revenue?

Besides release timing, advertising plays an important role in movie distribution. In general, 80% of the advertising budget is spent during theatrical release, while the remaining 20% is left for the home video release. This generates incentives for studios to release their home videos early, as the advertising expenditure in theaters will have a stronger effect on home video consumption. However, at the same time, some consumers may be willing to delay the box office purchase decision to the home video market, reducing the box office revenue. This impacts the home video market as well, as it is well known that the box office revenue serves a quality signal that drives home video demand. This trade-off clearly illustrates how difficult it is to choose an optimal home video window, which might differ between different technological quality home videos.

In the past few years we have seen several efforts from distributors to shrink the home video window. With the start of the COVID-19 pandemic in December 2019 and the limited market power of theaters, these efforts expanded. In July 2020 AMC (the world's largest movie theater chain) and Universal struck an agreement to release home videos just 17 days after theatrical release (Watson 2020). Similar agreements were implemented by other studios, with Warner Bros. announcing they will release all of their year 2021 films simultaneously in theaters and on the HBO Max platform (Schwartzel and Flint 2020). Recently ViacomCBS launched their own streaming platform, Paramount+, which will release Paramount and some MGM titles just 45 days after theatrical release (Kit 2021). These changes were made in the midst of a market dominanted by home viewing; it remains uncertain whether theaters will regain their influence on the market.

This chapter develops a dynamic structural model for the box office market and the home video market, consisting of two technological qualities, DVDs and Blu-rays[1]. The model is built around release timing strategies, and it is able to quantify the change in revenue associated with advancing or delaying home video releases. In order to do this, modeling consumer forward looking behavior is critical. We model the home video market as an infinite horizon model for each movie and technological quality, where consumers decide whether to buy the disc, or to delay their purchase in each period. We model the box office market for each movie as a dynamic program with a finite horizon. During each period consumers face the decision whether to buy a ticket or wait for the next period. The terminal continuation value of the box office market is set to the discounted value of the home video market, which depends on whether the consumer owns a DVD or Blu-ray player.

Consumer forward looking behavior is critical to model distribution in the movie industry. The time a movie remains in theaters is not a decision the studios make, but one each theater decides on their own, depending on performance. Thus, consumers in the box office market do not know the horizon of such market, and have to form expectations on its duration in each period. Under our setting, consumers form dynamic expectations about the time a movie remains in theaters. These

[1] This can be seen as a proxy for Full HD and 4K videos today. It is simple to add other home videos as the data becomes available

expectations depend on movie performance and time since release. Consumers are also forward looking in the evolution of the movie quality over time, the home video market value, and the home video window, which determine the terminal continuation value of the box office market.

Once the consumer decision problem in terms of movie qualities is solved, we recover movie/medium fixed effects, and time dependent characteristics that depend on each market, such as age, advertising build-up, average price and seasonality. In a second stage, we regress the fixed effects on time independent characteristics, such as the home video window, theatrical revenue, production budget, distributor, release year and poster colors. We analyze the poster image for each movie and identify the percentage of pixels that belong to each color in a 12 color palette. We then analyze the correlation between using specific colors, and pairs of a color and its complement, with consumer preference for a movie.

With the demand estimates for the box office, DVD and Blu-ray markets, we are able to perform a series of counterfactuals. First, we estimate the impact of modifying the simultaneous release of DVDs and Blu-rays, which was the industry practice at the time of data collection. We leave everything constant, while adjusting the advertising build-up, (as having a shorter home video window increases the advertising build-up coming from theaters), and the box office revenue signal. We find that it is optimal to set the DVD and Blu-ray release 2.3 weeks after theatrical exit to achieve a 4.47% increase in studio revenue[2], while having little impact for theaters with respect to the data. Second, we estimate the impact of separately modifying the home video window for DVDs and Blu-rays. This allows the exploit of market segmentation strategies, as consumers express heterogeneous preferences depending on the technological quality of the home video player they possess. We find that the optimal strategy is to release DVDs within a week of theatrical exit, while to delay Blu-rays about 5.15 weeks from theatrical exit. This strategy increases studio revenue by 5.36% with respect to the data, and again, the impact to theaters is minimal. This result is driven by the different substitution patterns that different technological quality home videos present to theaters. Lower technological quality home videos, such as DVDs, have a low ex-ante value function for their market compared to the market of higher technological quality home videos, such as Blu-rays. This makes DVD owners less likely to delay the box office purchase decision compared to Blu-ray player owners when shortening the home video window. Thus, the studios can reap all

[2]Studio revenue is capturing all of the home video revenue, and a share of the box office revenue that comes from imposing a standard theater-distributor contract.

the benefits of the advertising spillover effect from theaters while posing minimal competition to theaters. The trade-off is more balanced for Blu-rays. Shortening their home video window will increase substitution from theater, and will have further impact on the DVD market, as the quality signal coming from box office revenue is reduced. This result captures consumer heterogeneity in the box office market; Blu-ray player owners value picture quality above everything else, while DVD player owners value timing and pricing above picture quality. This allows the studios to perform market segmentation strategies on releases to boost revenue.

The remainder of this chapter is organized as follows; Section 3.2 presents the related literature, Section 3.3 describes the industry and shows the main attributes of the dataset used, Section 3.4 presents the structural model including consumer utility specifications and the consumers' optimal purchase problem, Section 3.5 explains the estimation and identification procedure for the model parameters, Section 3.6 shows and describes the parameter estimates of the model, and Section 3.7 presents the counterfactual analysis on the home video windows.

3.2 Literature Review

In the past few decades, several researchers have studied the motion picture industry, which intersects several fields such as marketing, operations management, and economics. We separate the relevant literature for this chapter into two groups, demand estimation and release timing. Within each of these groups we discuss the papers related to the movie industry.

3.2.1 Demand Estimation

This chapter uses the frameworks of Nevo (2000), Gowrisankaran and Rysman (2012) and Berry et al. (1995) to present and estimate a dynamic discrete choice model for movie distribution that captures consumer forward looking behavior. Unlike these papers, we consider a setting with two different markets, box office and home video, that are connected to each other through the home video window and the home video market value. In a similar fashion to this chapter, Derdenger (2014) has analyzed the technological tying of software and hardware in the 128-bit video game industry.

Advertising plays a critical role in this chapter, as it drives a side of the home video window trade-off. We treat advertising as exogenous, where it enters the model as a covariate, following

Dubé et al. (2005) to analyze the long-term effects of advertising. In a different setting, Yan (2020) analyzes equilibrium advertising and theatrical releases while treating advertising as endogenous. Its setting is quite different from ours, as it finds equilibrium advertising levels and release dates in a competitive environment with consumers that are not forward looking, whereas in our chapter we focus in finding optimal release timing strategies in a dynamic monopolistic setting that captures consumer forward looking behavior in theater and home video markets.

There is an abundant number of papers that do an empirical analysis of demand in the movie industry. This includes Eliashberg and Shugan (1997), Eliashberg et al. (2000), Elberse (2007), Eliashberg et al. (2014) and Packard et al. (2016) studying the impact of critics' reviews, star actors, networks of cast and crew, and opening weekend box office on overall revenue performance. Lehmann and Weinberg (2000), Elberse and Anand (2007) and Rao et al. (2017) analyze the impact of advertising on box office revenue.

3.2.2 Release Timing

The problem of finding the optimal inter-release times of sequentially released products has been widely studied in several industries. Moorthy and Png (1992) analyze the optimal timing and quality of sequential product releases. Lehmann and Weinberg (2000) analyze the problem of demand cannibalization between box office and home videos while trying to reap gains as quickly as possible. This line of work is extended by Luan and Sudhir (2006) capturing box office consumer forward looking behavior. This is the closest line of research to this chapter, as they study the optimal inter-release time between box office and DVDs using box office sales, and DVD sales and rental data. However, their analysis only uses advertising expenditure and does not capture advertising build up over time, which lowers the incentives to shorten inter-release times. In our work, advertising build up is a major driver for shorter inter-release times. We specifically focus on how the optimal inter-release times change as the technological quality of home videos increases from DVDs to Blu-rays, while creating heterogeneous consumer preferences between these products. Another distinction with Luan and Sudhir (2006), is that in our model, consumers also form expectations about the evolution of movie quality during each market, as well as on the time a movie remains in theaters. We develop a discrete hazard model to asses the probability distribution of the remaining time in theaters of a movie, that is dependent on the time since release and performance.

Other relevant work on movie release timing may be found in the study of the trade-off between seasonality and freshness for DVDs, see Mukherjee and Kadiyali (2018). August et al. (2015) and Calzada and Valletti (2012) analyze, through theoretical models in different settings, the conditions under which day-and-date, direct-to-video, or delayed home video releases are optimal release timing strategies for the movie industry.

3.3 Industry Setting and Data

The motion picture industry consists of three stages: production, distribution and exhibition. The production stage consists of the development of a motion picture and is a creative process with important economic implications for the parties involved. The process usually begins with an idea, concept or true event, which a writer captures in a screenplay. If a producer is interested in the screenplay, it may sign an option agreement with the writer, which gives the producer the possibility of purchasing the complete screenplay, and provides an upfront payment for the writer. Substantial financing is needed to begin production, which is lowered when the producer is affiliated with a studio. Upon the signature of a studio contract, the producer gives up several rights, including sequels, spin-offs, and merchandising. At the same time, the producer increases its chances of obtaining bank loans and securing favorable distribution and exhibition deals. These contracts benefit the studios, as they provide a constant inflow of products from successful firms. Many producers face financing issues when they cannot reach a deal with a studio; in such case the studio must obtain financing from other sources, which is difficult when no distribution deals are guaranteed (Vogel 2014).

The distribution stage begins once a movie has completed production, and it includes the distribution to theaters, home video markets, as well as the marketing activities in each market where the movie is released. Distributors face a wide range of decisions in this stage, including when to release the movie in each channel and the advertising strategy for the motion picture. Among distributors there is a clear distinction between major and independent firms. The major distributors, usually referred to as *"The Big Six"*, include *Paramount, Sony, Twentieth Century Fox, Universal, Walt Disney* and *Warner Bros.*. These studios produce, finance and distribute their own movies. At the same time, they also finance and distribute films produced by independent film makers that are associated with the studio. Ensuring a strong US theatrical box-office gross is very important for these studios, because it is a performance metric to indicate sales potential in other distribution channels such as global theatrical, home video, and pay television (Eliashberg et al. 2006). Simultaneously, the growing importance of non-theatrical channels as a source of revenue is generating incentives for the studios to reduce the time between theatrical and non-theatrical releases. Figure 3.1 shows the evolution of the average DVD release windows, (time between theatrical and DVD releases), for major studios and years. This reduction in non-theatrical windows poses several questions, which we address in this chapter. One of them was proposed by Eliashberg et al. (2006), "To what extent are theatrical and nontheatrical windows substitutes or complements (i.e., either negatively or positively affecting each other's revenue potential)? For example, does the availability of DVDs deter people from going to the theater?". Building upon this question, we analyze the optimal non-theatrical window across different technological quality home videos, which present different substitution patters with theaters.

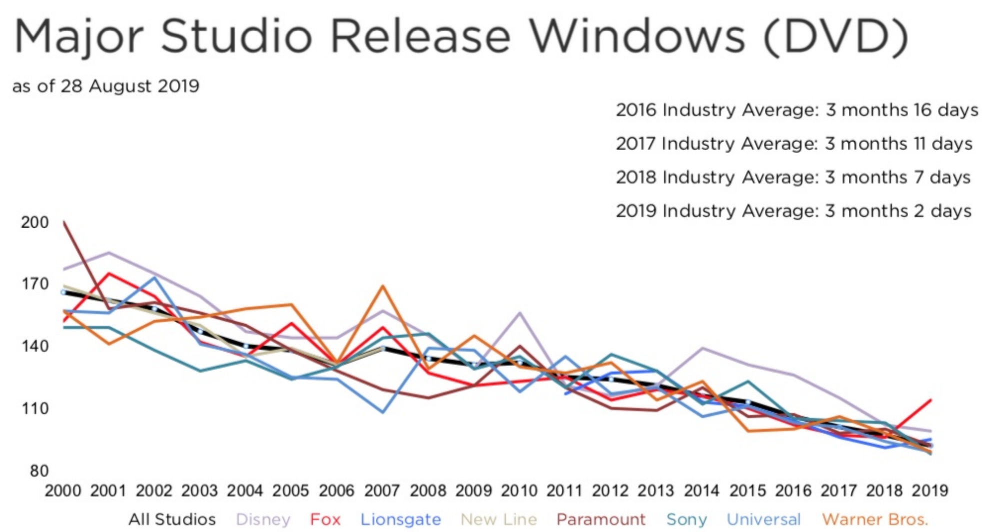

Figure 3.1: Evolution of the number of days between theatrical and DVD release for major studios. Source: https://www.natoonline.org.

The exhibition stage is the only stage in which major studios have limited control. The contractual agreements between exhibitors and distributors only involve a minimum playing time, as well as terms on how the box office revenue is shared between the parties involved. In the end, it is up to each individual theater to control the total playing time for each movie (beyond that set minimum). Studios generate a strong buzz prior to and during their theatrical release - combining advertising, word of mouth, and media attention - which drives demand for the motion picture in other distribution channels (Eliashberg et al. 2006).

3.3.1 Data Discussion

We estimate our model using panel data for box office, DVD and Blu-ray sales. The data is at the weekly level for all panels. For each week that a movie is in the box office, our data includes revenue ($), ticket sales, number of theaters, and other characteristics. For each movie and week in a home video environment, our data includes unit sales, revenue ($) and average price. We observe 149 movies, across box office, DVD and Bluray, with observations between the years 2009 and 2018. Out of those 149 movies, only 113 had Blu-ray releases, while all of them had a DVD release. We also obtain advertising expenditure data for each movie at the monthly level from Ad$pender. For each movie we also have information on its characteristics, such as distributor, domestic box office revenue, production budget, home window lengths and release dates. We also scraped poster images from www.themoviedb.org and applied color theory on them to extract hue information, the procedure is described in section 3.5.3. We create market shares by dividing the unit sales with the number of consumers, for the box office, or number of DVD/Blu-ray player owners for the home videos. We use data from the U.S. Census Bureau to get the number of consumers within each movie rating per year, and data from the Consumer Electronic Association to get information on DVD/Blu-ray player ownership. This data is interpolated to the monthly level, assuming a linear growth rate. We adjust inflation in all revenues to January 2019 dollars using data from the Bureau of Labor Statistics. To create our final dataset, we remove box office panel weeks with less than 100 theaters and we define the start of the home window as the week in which a movie is in less than 100 theaters. The final dataset consists of $1,797$ observations across 149 movies for the box office, $44,800$ for DVDs (across the same 149 movies), and $25,373$ observations for Blu-rays across a subset of 113 movies. Table 3.1 presents a summary statistics of the production budget, advertising expenditure, and revenue per medium of the final dataset, while Table 3.2 presents a summary of the home windows and time in theaters. It is important to note that the difference between the DVD and Blu-ray home video windows is due to the different sample sizes, and not due to different home video release dates for the same movie.

	Mean	**Median**	**Standard Deviation**
Production budget	$ 121,775,960	$99,979,420	$ 77,255,980
Domestic box office revenue	$ 122,169,272	$ 82,897,417	$ 106,867,804
Total DVD revenue	$ 37,480,615	$ 26,050,639	$ 42,840,121
Total Blu-ray revenue	$ 22,478,904	$ 16,729,886	$ 19,335,287
Total advertising expenditure	$ 27,612,363	$ 27,170,200	$ 11,091,510

Table 3.1: Data summary on revenues, advertising and budgets.

	Mean	**Median**	**S.D.**	**Maximum**	**Minimum**
DVD home window (weeks)	6.6951	6.2857	3.6906	16.2857	0.2857
Blu-ray home window	6.1871	6.2857	3.6453	16.2857	0.2857
Time in theaters (weeks))	11.4007	11.0000	3.1829	22.0000	6.0000

Table 3.2: Data summary on timing characteristics.

3.3.2 Data Analysis

In this section we present reduced form results that show the connection between relevant covariates and revenue. We retrieve estimates in a two stage process; with a first stage fixed effects regression that includes time dependent characteristics, and a second stage retrieval of movie characteristics from movie fixed effects.

Table 3.3 shows the box office reduced form using panel data on weekly box office ticket sales. It shows the first stage regression on top, while the fixed effect GLS (General Least Squares) retrieval of static movie characteristics at the bottom. The sign of the coefficients is quite intuitive; as a movie ages, the demand for it decreases, while an increase in advertising expenditure yields greater demand [3]. The average box office price per week is not significant, probably due to the low variation of price across both, time and titles[4]. The "lag 1st3weeks box revenue" covariate represents the total lagged box office revenue until the current period, or period 3 included, whichever comes first. This shows that if a movie performs well in the first few weeks since release, it will drive demand up for the consecutive weeks, but with diminishing returns as the quadratic component is negative. The fixed effects regressions shows that production budget and time in theaters are major drivers

[3]We create advertising covariates following Dubé et al. (2005), we describe the procedure in Section 3.5.3.
[4]To control for price endogeneity we used lagged prices as an instrument.

for box office demand. The home window contribution is smaller in magnitude, but presents the expected substitution pattern with a positive sign, which means that box office revenue increases by delaying home video releases.

Box office: two-stage reduced form

Time Dependent Variables	**Estimates**
Age (weeks)	-0.8184 (0.0161)***
Age^2 (weeks^2)	0.0223 (0.0008)***
Price ($)	-0.6720 (0.5636)
Advertising	0.50933 (0.1361)***
log(lag 1st3weeks box revenue)	0.23651 (0.0474)***
log(lag 1st3weeks box revenue) ^{2}	-0.0128 (0.0027)***
Movie fixed effects	
Month fixed effects	
N	1,797
Time Independent Variables	**Estimates**
Constant	7.3604 (1.3068)***
log(production budget)	0.2551 (0.0343)***
Time in theaters (week)	0.4945 (0.0297)***
Time in theaters^2 (week^2)	-0.0111 (0.0011)***
Home window (week)	0.0039 (0.0170)
Home window^2 (week^2)	0.0024 (0.0010)***
Release year fixed effects	
Distributor fixed effects	
Color fixed effects	
N	149

$^{***}p < .01$, $^{**}p < 0.05$, $^{*}p < .1$

Table 3.3: Reduced form estimates of a two stage fixed effects model on weekly box office ticket sales.

Table 3.4 shows the home video reduced form using panel data on home video sales. The top of the table shows the first stage regression in which DVD and Blu-ray weekly sales are regressed against time dependent covariates, and an interaction between them and a Blu-ray indicator variable with movie/medium fixed effects. We then run a fixed effects GLS regression for DVDs and Blu-rays separately on time independent characteristics. In the first stage regression we controlled for price endogeneity using lagged prices as instrumental variables. We can see that home video demand lowers with age and price. Advertising expenditure has a positive impact on home video sales, but this effect is lower for Blu-rays. The fixed effects GLS regression shows an increasing demand for home videos with an increase in box office opening revenue within the ranges of box office revenues observed. As for the home video window, DVD demand seems to be larger with lower windows, while Blu-rays exhibit concave demand shape with a maximum in around 8 weeks.

Home video: two-stage reduced form

Time Dependent Variables	**Estimates Home Video**	**Estimates Blu-ray Indicator**
Age (weeks)	-0.0119 (0.0035)***	-0.0866 (0.0263)***
Age^2 (weeks^2)	0.0000 (0.0000)***	-0.1119 (0.0365)***
Price ($)	-0.0182 (0.0012)***	-0.2212 (0.0499)***
Advertising	0.6462 (0.0078)***	-0.2832 (0.0647)***
Movie fixed effects	✓	✓
Month fixed effects	✓	✓
Year fixed effects	✓	✓
N	70,253	25,373
Time Independent Variables	**Estimates DVD Indicator**	**Estimates Blu-ray Indicator**
Constant	33.3590 (0.5997)***	51.2524 (0.9777)***
log(1st3week box revenue)	-3.3500 (0.0705)***	-5.6018 (0.1139)***
$\text{log(1st3week box revenue)}^2$	0.1030 (0.0021)***	0.1709 (0.0033)***
Home window (week)	-0.0233 (0.0053)***	0.0989 (0.0061)***
Home window^2 (week^2)	0.0015 (0.0003)***	-0.0056 (0.0004)***
Release year fixed effects	✓	✓
Distributor fixed effects	✓	✓
Color fixed effects	✓	✓
N	149	113

$^{***}p < .01$, $^{**}p < 0.05$, $^{*}p < .1$

Table 3.4: Reduced form estimates of a two stage fixed effects model on weekly home video sales.

With these reduced form results, we can see several features that are important to embed into a structural model. These include expectations about the home video window and time in theaters during the theatrical market, box office revenue as a signal of movie quality for the home video market, and advertising. Furthermore, the significant coefficients for age in both markets, show that consumers should be able to form expectations about movie quality evolution, which would be relevant in the case they decide to delay their purchase decision to later periods.

3.4 Demand model

In this section we discuss the structural model that captures the relationship between box office ticket sales, home video sales, and the forward-looking behavior of consumers. The box office and home video markets are linked through the home video window, which is the time between theatrical exit and home video release. Shrinking this window leads to higher freshness on the home video market and a greater spillover of advertising spent for the theatrical market, but it may lead to demand cannibalization from theaters. Consumers in the box office market form expectations about this home video window and decide whether to watch the movie in the current week, or delay their decision to the following one. Consumers form expectations about the home video window, as well as the evolution of prices, movie quality and the theatrical run time. Since we are interested in analyzing the aforementioned trade-off disregarding competition, we consider each movie to be a monopoly.

The timing of the model is as follows: Consumers own either a DVD or a Blu-ray player when a movie is released in theaters. In each week of the theatrical market, consumers decide whether or not to buy a box office ticket. They continue to make such a decision until they elect to watch the movie, or until the theatrical run time is over. After the theatrical runtime is over, we enter the home video window. During this period, consumers are not able to watch the movie in theaters, nor buy a home video for this movie. Then, the home video is released and we enter the video time. In each period (week) of the home video market, consumers decide whether or not they will purchase the their respective disc (DVD or Blu-ray). They continue to make such a decision in each period until they elect to purchase the home video. Figure 3.2 shows the described timeline of the model. In order to estimate the model involving the discussed forward-looking behavior, we use the frameworks of Gowrisankaran and Rysman (2012) and Derdenger (2014). The former paper embeds consumer expectations about price, movie characteristics and unobservable factors that might evolve over time. The later paper is used to link the utilities between the theatrical and home video markets.

Next, we present the utility specifications for the home video and theatrical markets, first discussing the associated utilities for home videos and then for box office tickets.

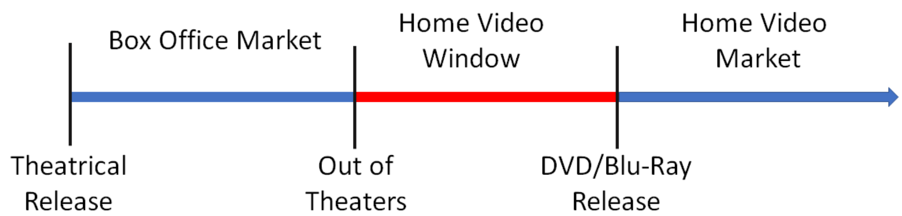

Figure 3.2: Timeline of movie distribution.

3.4.1 Home video utility

We now outline the utility model for home video consumption. The home video utility specification follows that of Gowrisankaran and Rysman (2012); it is an infinite horizon model. Our specification allows for model parameters to differ between DVDs and Blu-rays so that the expected value of the home market for users who hold a DVD or a Blu-ray player are different. This term plays an important role in the box office market, since it enters through the terminal continuation value of such finite horizon model.

In the home video market, each consumer decides in each time period t whether or not to purchase a home video for movie/medium j. If consumer i decides to purchase a home video for movie/medium j of technological quality $k \in$ {DVD, Blu-ray}[5] in time period t, she obtains a utility given by

$$u^h_{i,j,t} = f^h_{j,t} + \alpha^{p,h}_k p^h_{j,t} + \epsilon^h_{i,j,t} \tag{3.1}$$

where $f^h_{j,t}$ is the flow utility from the home video, $p^h_{j,t}$ is the price of the home video, and $\epsilon^h_{i,j,t}$ is an idiosyncratic shock which we assume to be the realization of a Type-1 Extreme Value distribution which is independent and identically distributed across consumers, products, and time periods.[6]

A consumer who does not purchase a movie/medium j of technological quality $k \in$ {DVD, Blu-ray} in period t receives

$$u^h_{0,j,t} = \beta \mathbb{E}[V^h_k(\Omega^h_{j,t+1})|\Omega^h_{j,t}] + \epsilon^h_{0,j,t}, \tag{3.2}$$

[5]Movie/medium j includes information about the technological quality of the home video, so we defer from adding subscript k to denote such technological quality when subscript j is present.

[6]Let superscript h and b denote utility specification for home video and box office respectively.

where $\Omega^h_{j,t}$, is industry state of movie/medium j at period t; involving the flow utility, price disutility, and all history factors that influence the home video's future attributes. Finally, $V^h_k(\Omega^h_{j,t})$ represents the value of having the purchase possibility of home video j of technological quality $k \in \{\text{DVD, Blu-ray}\}$, when the state of home video j is $\Omega^h_{j,t}$.

The home video flow utilities depend on movie characteristics with the following relation:

$$f^h_{j,t} = \alpha^{fe,h}_j + \alpha^{x,h}_k x^h_{j,t} + \xi^h_{j,t}, \tag{3.3}$$

where $x^h_{j,t}$ are time dependent observable movie characteristics (for DVD and Blu-ray), such as age, advertising expenditure, price, month and year, $\alpha^{fe,h}_j$ are movie/medium fixed effects, $\alpha^{x,h}_k$ are time dependent characteristics coefficients which depend on whether the technological quality of the home video is a DVD or a Blu-ray, and $\xi^h_{j,t}$ are unobservable characteristics that vary both over time and across movies. It is important to remark that we don't explicitly express the dependence of $f^h_{j,t}$ on k since j (movie/medium) already includes that information.

The estimation of the model parameters involves a second stage in which the movie medium fixed effects are regressed on time fixed movie characteristics following Nevo (2000). Examples of movie characteristics involve opening box office revenue, distributor, the home window length, poster colors and release year.

3.4.2 Box office utility

Unlike the home video market, the box office market has a finite horizon. The consumer type is denoted by the subscript $k \in \{\text{DVD, Blu-ray}\}$, which provides a source of heterogeneity that depends on the technological quality of the home video player owned. This heterogeneity enters exclusively in the outside option and not in the purchase utility, as we assume that the box office ticket purchase continuation value is zero for both consumer types. In each period t, consumer i considers whether or not to watch a particular movie j. Once a consumer watches movie j, she exits the box office market for movie j. If consumer i decides to purchase a box office ticket for movie j in time period t, she obtains utility given by

$$\begin{aligned} u^b_{i,j,t} &= \phi(\tilde{f}^b_{j,t} + \tilde{\alpha}^{p,b} p^b_{j,t}) + \epsilon^b_{i,j,t} \\ &= f^b_{j,t} + \alpha^{p,b} p^b_{j,t} + \epsilon^b_{i,j,t}, \end{aligned} \tag{3.4}$$

where $f^b_{j,t}$ is the flow utility from observable and unobservable box office characteristics, $p^b_{j,t}$ is the price of the box office ticket, $\epsilon^b_{i,j,t}$ is an idiosyncratic shock, and β is the weekly discount factor. The parameter ϕ is a scaling parameter which permits the comparison between the home video and box office markets, it serves as a utility normalization factor, since we forced the error terms for both markets to have the same variance.

A consumer who does not watch movie j in period t, and owns a home video player of technological quality $k \in \{\text{DVD, Blu-ray}\}$ receives

$$u^b_{0,k,j,t} = \beta \mathbb{E}[V^b_k(\Omega^b_{j,t+1})|\Omega^b_{j,t}] + \epsilon^b_{0,j,t}, \tag{3.5}$$

where $\Omega^b_{j,t}$, is industry state of movie j in theaters at period t; involving the flow utility, price disutility, and all history factors that influence the movie's future attributes (for instance, revenue and time since release, number of theaters available, etc. are important to form expectations about the last period of the market). Finally, $V^b_k(\Omega^b_{j,t})$ represents the value of having a box office ticket purchase possibility when the state of movie j is $\Omega^b_{j,t}$ and the consumer has a home video player of technological quality $k \in \{\text{DVD, Blu-ray}\}$. Note that the outside option depends on j since we are modeling each movie as a monopoly, and the continuation value of delaying the box office purchase decision is different for each movie.

The box office flow utilities depend on the movie features with the following relation:

$$f^b_{j,t} = \alpha^{fe,b}_j + \alpha^{x,b} x^b_{j,t} + \xi^b_{j,t}, \tag{3.6}$$

where $x^b_{j,t}$ are observable box office movie characteristics that vary over time, $\alpha^{x,b}$ are the coefficients of observable time dependent movie characteristics, $\alpha^{fe,b}_j$ are movie fixed effects and $\xi^b_{j,t}$ are the unobservable components of the flow utility that vary both over time, and across movies. Examples of time-dependent box office movie characteristics include advertising expenditure, average price, month, age, and performance up to current period. Similarly to the home video market, the estimation of the model parameters involves a second stage in which fixed effects are regressed on movie characteristics. Examples of movie characteristics involve distributor, production budget, and release year. Note that in (3.4) and (3.6) we assume that the consumer purchase utility specification about movies in theaters does not depend on the type of home video owned (DVD or

Blu-ray), but this heterogeneity enters in (3.5), the outside option.

This utility specification generates a challenge. The box office market has a finite horizon, thus the model has to embed consumer's expectations about the market's horizon, the home window length, and the value of the home video market in order to quantify continuation values. To embed consumer's expectations we must make assumptions on what factors affect the movie's industry state $\Omega^b_{j,t}$. This is analyzed within the consumer's problem, in section 3.4.3.

3.4.3 Consumer's problem

We now outline the consumer's decision process, which incorporates the forward looking behavior about the evolution of the movie industry states $\Omega^b_{j,t}$. We begin by describing the decision process for the home video market as it is independent of the value functions of the box office market. We then describe the finite horizon model of the box office market, where the home video value enters through the terminal continuation value.

Home video market

The home video market consists of two separate markets that differentiate on the technological quality, we distinguish between them with subscript $k \in \{\text{DVD, Blu-ray}\}$. Each of them can be seen as an optimal stopping problem with an infinite horizon. In each period, consumers have the possibility to purchase their respective home video disc, or to wait. Following Equations (3.1) and (3.2), the value function prior the realization of $\vec{\epsilon}^{\,h}_{j,t} \doteq (\epsilon^h_{i,j,t}, \epsilon^h_{0,j,t})$ for consumer that owns home video player $k \in \{\text{DVD, Blu-ray}\}$ can be written as

$$V^h_k(\Omega^h_{j,t}) = \int \max\left\{f^h_{j,t} + \alpha^{p,h}_k p^h_{j,t} + \epsilon^h_{i,j,t}, \beta\mathbb{E}[V^h_k(\Omega^h_{j,t+1})|\Omega^h_{j,t}] + \epsilon^h_{0,j,t}\right\} g_\epsilon(\vec{\epsilon}^{\,h}_{j,t}) d\vec{\epsilon}^{\,h}_{j,t}. \tag{3.7}$$

From (3.7), the first element of the max operator indicates the purchase utility, while the second indicates the expected discounted value of delaying the purchase decision to the next period.

We proceed by using the aggregation properties of the extreme value distribution to express (3.7) in a rather simpler form, and then we make assumptions on how consumers form expectations about future movie industry states. Specifically, we can write

$$V^h_k(\Omega^h_{j,t}) = \ln\Big(\exp(\delta^h_{j,t}) + \exp(\beta\mathbb{E}[V^h_k(\Omega^h_{j,t+1})|\Omega^h_{j,t}])\Big), \tag{3.8}$$

where $\delta^h_{j,t} = f^h_{j,t} + \alpha^{p,h} p^h_{j,t}$, the *logit inclusive value*, is defined as the ex-ante present discounted lifetime value of buying a home video at period t, as opposed to waiting for the next period. We make the assumption that consumers only take into account the current value of $\delta^h_{j,t}$ to form the expectations of the evolution of V^h to the next period. So the history of the past values of $\delta^h_{j,t}$ does not matter. Thus, we have

$$V^h_k(\delta^h_{j,t}) = \ln\Big(\exp(\delta^h_{j,t}) + \exp(\beta \mathbb{E}[V^h_k(\delta^h_{j,t+1})|\delta^h_{j,t}])\Big). \tag{3.9}$$

We employ rational expectations for future values of $\delta^h_{j,t}$ by imposing a simple linear autoregressive specification:

$$\delta^h_{j,t+1} = \nu^h_{1,k} + \nu^h_{2,k}\delta^h_{j,t} + \eta^h_{j,t+1,k}, \tag{3.10}$$

where $\eta^h_{j,t+1}$ is normally distributed with zero mean and unobserved at time t, while $\nu^h_{1,k}$ and $\nu^h_{2,k}$ are general parameters to be estimated for $k \in \{\text{DVD, Blu-ray}\}$. This assumption ensures consumers are on average correct about the movie quality evolution. It is important to remark that the optimal consumer decisions, given a movie state $\delta^h_{j,t}$, will depend on the joint solution of the Bellman equation (3.9) and the movie state regression (3.10).

Once these two equations are solved, we can obtain the value functions $V^h_k(\delta^h_{j,t})$, and we then use them to estimate the individual purchase probabilities. The movie/medium j purchasing probability for a consumer at period t is given as a function of $\delta^h_{j,t}$, and is

$$\hat{s}^h_{j,t}(\delta^h_t) = \frac{\exp(\delta^h_{j,t})}{\exp(V^h_k(\delta^h_{j,t}))}. \tag{3.11}$$

Box office market

The box office market is similar to the home video market, but with a few differences. First, the market has a finite horizon which is unknown by the consumers, and second, the discounted home video value enters the box office model through the terminal continuation value of this unknown horizon. This allows for consumers to substitute between the box office market and the home video market, if this terminal continuation value is large enough. Furthermore, the home video value differs between the technological quality of the home video, i.e. DVD or Blu-ray, which may drive different substitution patterns between the different technological quality markets and the box office.

In order to model the unknown finite horizon, we use the time and revenue since release as drivers for a distribution of possible horizons. We use the data in order to fit a discrete hazard model that gives the probability for each possible horizon. This captures the endogeneity of theatrical runtime, where each theater decides based on performance whether to keep a movie in theaters or not. This procedure is described in Section 3.4.3

Following Equations (3.4) and (3.5) and using the aggregation properties of the extreme value distribution, similarly to how we arrived to Equation (3.8) in the home video market, we obtain the following value function equation for $V^b_{k,|T}$, the box office value function at time period t conditional on having an horizon at T and owning a home video player of type $k \in \{\text{DVD, Blu-ray}\}$;

$$V^b_{t,k|T}(\Omega^b_{j,t}) = \ln\Big(\exp(\delta^b_{j,t}) + \exp(\beta\mathbb{E}[V^b_{t+1,k|T}(\Omega^b_{j,t+1})|\Omega^b_{j,t}])\Big) \quad \forall t = 1, \ldots, T-1, \tag{3.12}$$

where $\delta^b_{j,t} = f^b_{j,t} + \alpha^{p,b}_k p^b_{j,t}$, ($\Omega^b_{j,t}$ is the state of industry of movie j at time t, and T is the unknown horizon of the market, following a probability distribution that depends on t and the revenue since release. Finally, the terminal value function is set using expectations on the home video window (NT), and the home video value of movie j with technological quality k ($V^h_{k,j}$);

$$V^b_{T,k|T}(\Omega^b_{j,T}) = \ln\Big(\exp(\delta^b_{j,T}) + \exp(\mathbb{E}[\beta^{NT}V^h_{k,j}])\Big). \tag{3.13}$$

Since the horizon of the problem is unknown, consumers form a probability distribution over probable horizons, $g_T(\Omega^b_{j,t})$, which depends on the industry state of the movie $\Omega^b_{j,t}$ at time t. Then Equations (3.12) and (3.13) can be solved for each T and aggregated using the probability distribution over possible horizons for each time period. Finally we have

$$V^b_{t,k}(\Omega^b_{j,t}) = \sum_T V^b_{t,k|T}(\Omega^b_{j,T})g_T(\Omega^b_{j,t}) \quad \forall t = 1, \ldots, T, \text{ and } k \in \{\text{DVD, Blu-ray}\} \tag{3.14}$$

where the probability distribution over values of T is created from data using a hazard model that depends on the revenue and weeks since release. Subsection 3.4.3 provides details on this procedure.

We assume that consumers only take into account the current value of $\delta^b_{j,t}$, the time since release, t, and the revenue since release, r_t, to form expectations about the future values of purchase decisions. Then our state variables can be assumed to be (t, r_t, and $\delta^b_{j,t}$), since time (t) and revenue since release (r_t) are important to determine the probability distribution over possible horizons. Then, Equations (3.12), (3.13) and (3.14) can be rewritten as

$$V^b_{t,k|T}(\delta^b_{j,t}) = \ln\Big(\exp(\delta^b_{j,t}) + \exp(\beta\mathbb{E}[V^b_{t+1,k|T}(\delta^b_{j,t+1})|\delta^b_{j,t}]\Big) \tag{3.15}$$

$$\forall t = 1,\ldots,T-1, \forall k \in \{\text{DVD, Blu-ray}\},$$

$$V^b_{T,k|T}(\delta^b_{j,T}) = \ln\Big(\exp(\delta^b_{j,T}) + \exp(\mathbb{E}[\beta^{NT_{k,j}}V^h_j])\Big), \text{ and} \tag{3.16}$$

$$V^b_{t,k}(\delta^b_{j,t}, r_t) = \textstyle\sum_T V^b_{t,k|T}(\delta^b_{j,t})g_T(t,r_t) \quad \forall t = 1,\ldots,T, \text{ and } k \in \{\text{DVD, Blu-ray}\}. \tag{3.17}$$

Once again, we employ rational expectations for future values of δ^b, by imposing a linear autoregressive specification:

$$\delta^h_{j,t+1} = \nu^b_1 + \nu^b_2\delta^h_{j,t} + \eta^b_{j,t+1} \tag{3.18}$$

where $\eta^b_{j,t+1}$ is normally distributed with zero mean and unobserved at time t, while ν^b_1 and ν^b_2 are general parameters for all movies to be estimated. This assumption ensures consumers are on average correct about the movie quality evolution.

We assume that consumer expectations about the home video window and the home video value to quantify $\mathbb{E}[\beta^{NT_{k,j}}V^h_{k,j}]$ in (3.13), are based on perfect foresight. This is because if we did not use perfect foresight we would need to model consumer beliefs, and when we run a counterfactual analysis on these home video windows, they won't be consistent with the estimated beliefs. By using perfect foresight we ensure that estimated beliefs are consistent with our counterfactual home window lengths video windows.

Note that equations (3.12), (3.13), (3.14) and (3.18) must be solved jointly. This is because a change in the Bellman equation will yield different ν^b_1 and ν^b_2 coefficients, and that will impact the Bellman equations. Specifically, these equations need to be solved twice, one time for DVD player owners, and another time for Blu-ray player owners. The difference between these two comes in the terminal continuation values used in (3.13)[7]. The fixed point will depend on the continuation value used, but we refrain from using it as a variable for V^b_t. Note that both consumer types have the same quality vector, $\delta^b_{j,t}$, since the difference between types lies in the outside option.

[7]Both, the home video value as well as the home window length may differ between DVDs and Blu-rays.

Once these equations are solved, we can compute the individual purchase probabilities for each consumer segment. For simplicity, we now refrain from using k to identify consumer types and we use the suprascript $b - dvd$ for DVD player owners and $b - blu$ for Blu-ray player owners instead. The probabilities that a DVD and a Blu-ray player owner purchase a ticket for movie j in period t is given by

$$\hat{s}_{jt}^{b-dvd}(\delta_{j,t}^{b}, r_t) = \frac{\exp(\delta_{j,t}^{b})}{\exp(V_t^{b-dvd}(\delta_{j,t}^{b}, r_t))} \text{ and } \hat{s}_{jt}^{b-blu}(\delta_{j,t}^{b}, r_t) = \frac{\exp(\delta_{j,t}^{b})}{\exp(V_t^{b-blu}(\delta_{j,t}^{b}, r_t))}, \tag{3.19}$$

respectively. And based on the remaining weight of Blu-ray player owners for movie j at time period t, $w_{j,t}^{b-blu}$ we can compute the purchase probability of a random consumer

$$\hat{s}_{jt}^{b}(\delta_{jt}^{b}) = (1 - w_{j,t}^{b-blu})\hat{s}_{jt}^{b-dvd}(\delta_{jt}^{b}, r_t) + w_{jt}^{b-blu}\hat{s}_{jt}^{b-blu}(\delta_{jt}^{b}, r_t). \tag{3.20}$$

Discrete-Time Proportional Hazard Model for time in theaters

Theaters decide when to stop showing a movie, which is dependent on how the movie is performing. To capture this, we build a Discrete-Time Proportional Hazard Model following Cameron and Trivedi (2005) (section 17.10.1). This model gives a probability distribution for the remaining time in theaters, given the number of weeks the movie has been in theaters and the total revenue until then. Let T be the number of weeks a movie is in theaters, and $R(t)$ the box office total revenue by period t, we define the discrete time hazard function

$$\lambda^d(t|R_{t-1}) = Pr[T = t|T \geq t, R_{t-1}], \quad t = 1, \ldots, M,$$

which denotes the probability t is the last week this movie is in theaters, given it is in theaters in week t and the the total revenue until week $t-1$ is R_{t-1}. Then, the associated discrete-time survivor function is

$$S^d(t|R_{t-1}) = Pr[T \geq t|R_{t-1}] = \prod_{s=1}^{t-1} \left(1 - \lambda^d(s|R_{s-1})\right).$$

We can then specialize the continuous PH model to obtain the following expression for the discrete time hazard

$$\lambda^d(t|R_{t-1}) = 1 - \exp\Big(- \exp\left(\ln \lambda_{0t} + \beta \log(R_{t-1})\right)\Big),$$

where λ_{0t} for $t = 1, \ldots, M$ and β are parameters to be estimated. The associated discrete-time survivor function is

$$S^d(t|R_{t-1}) = \prod_{s=1}^{t-1} \exp\Big(- \exp\left(\ln \lambda_{0t} + \beta \log(R_{t-1})\right)\Big).$$

We can finally write the likelihood function as

$$L(\beta, \lambda_{01}, \ldots, \lambda_{0M}) = \prod_{i=1}^{N} \Big[\prod_{s=1}^{T_i - 1} \exp\Big(- \exp\left(\ln \lambda_{0s} + \beta \log(R_{s-1})\right)\Big)\Big] \times \Big(1 - \exp\Big(- \exp\left(\ln \lambda_{0T_i} + \beta \log(R_{T_i - 1})\right)\Big)\Big),$$

where i refers to each individual movie, and T_i for the time in theaters of such movie. We can the use the data to maximize this function over the parameters, and obtain our final Discrete-Time Proportional Hazard Model.

3.5 Estimation and identification

The estimation procedure to recover model parameters follows that of Gowrisankaran and Rysman (2012) and Derdenger (2014). Since the estimation of the box office market depends on the home video value, we estimate the home video market first and then we proceed with the box office market.

We will use $\vec{\alpha}^b$ and $\vec{\alpha}^h$ to denote $(\alpha^{fe,b}, \alpha^{x,b}, \alpha^{p,b})$ and $(\alpha^{fe,h}, \alpha^{x,h}, \alpha^{p,h})$, respectively, these are the vector of all fixed effects and observable characteristics coefficients. We now discuss the identification of the structural parameters $(\vec{\alpha}^h, \vec{\alpha}^b, \beta)$, which requires solving the home video market and use its results to solve the box office market. We do not attempt to estimate β, since it is well known that estimating the discount factor in dynamic decision models is a notoriously difficult task (Gowrisankaran and Rysman 2012, Magnac and Thesmar 2002). This is because consumer waiting can be explained by moderate preferences for movies, or by little discounting of the future. Thus, we set $\beta = 0.9995$ on a weekly level (equivalent to 0.974 yearly), leaving $(\vec{\alpha}^h, \vec{\alpha}^b, \phi)$ to estimate.

In order to allow for the comparison between box office and home video utilities, we must identify the scaling parameter ϕ. To do so, we must impose the following constraint in estimation.

$$\alpha^{p,b} = \phi\alpha^{p,dvd}. \tag{3.21}$$

With ϕ identified, we redefine $\vec{\alpha}^b \doteq (\alpha^{fe,h}, \alpha^{x,h})$, since $\alpha^{p,b}$ is not to be directly estimated.

Following Berry (1995) and Gowrisankaran and Rysman (2012), we specify a generalized method of moments (GMM) function

$$G(\vec{\alpha}^h, \vec{\alpha}^b, \phi) = Z'\vec{\xi}(\vec{\alpha}^h, \vec{\alpha}^b, \phi), \tag{3.22}$$

where $\vec{\xi}(\vec{\alpha}^h, \vec{\alpha}^b, \phi)$ is the stacked vector unobserved characteristics of the box office market (ξ^b_{jt}), DVD market (ξ^{dvd}_{jt}) and Blu-ray market (ξ^{blu}_{jt}), for which the predicted shares equal the observed shares, and Z is a matrix of exogenous instrumental variables. These instrumental variables consist of lagged home video prices, and box office prices which control for price endogeneity. We estimate the parameters to satisfy

$$(\hat{\alpha}^h, \hat{\alpha}^b, \hat{\phi}) = \underset{(\vec{\alpha}^h, \vec{\alpha}^b, \phi)}{\arg\min} \left\{ G(\vec{\alpha}^h, \vec{\alpha}^b, \phi)' W G(\vec{\alpha}^h, \vec{\alpha}^b, \phi) \right\}, \tag{3.23}$$

where W is a weighing matrix. Thus, to estimate $(\vec{\alpha}^h, \vec{\alpha}^b, \phi)$ we must first solve for $\xi(\vec{\alpha}^h, \vec{\alpha}^b, \phi)$, which requires solving for the shares of all markets. We first discuss how to solve for home video shares and then for box office shares. In the following sections, we explain how to obtain $\hat{\alpha}^h, \hat{\alpha}^b(\phi)$ and $\vec{\xi}(\hat{\alpha}^h, \hat{\alpha}^b(\phi), \phi)$ based on an initial guess of ϕ. The optimal value $\hat{\alpha}^h$ is independent of ϕ and can be solved separately. Given a guess for ϕ, one can solve for the optimal $\hat{\alpha}^b(\phi)$ easily, which will depend on the chosen ϕ. Finally, the optimal solution for $\hat{\phi}$ can be obtained by solving a single variable optimization problem, that includes a subproblem that finds $\hat{\alpha}^b(\phi)$,

$$\hat{\phi} = \underset{\phi}{\arg\min} \left\{ G(\hat{\alpha}^h, \hat{\alpha}^b(\phi), \phi)' W G(\hat{\alpha}^h, \hat{\alpha}^b(\phi), \phi \right\}. \tag{3.24}$$

3.5.1 Home video shares

The consumer decision problem for the home video market is defined in Section 3.4.3 as the fixed point of the Bellman equation (3.9), and the market evolution equation (3.10). We stack the

DVD and Blu-ray panel data, to find the vector $\delta^h_{j,t}$ for which the predicted shares $(\hat{s}^h_{jt})$ equals the observed shares $(\hat{s}^h_{jt})$ for each movie and time period, namely

$$s^h_{j,t} = \hat{s}^h_{j,t}(\delta^h_{j,t}) \quad \forall j, t. \tag{3.25}$$

Following Gowrisankaran and Rysman (2012) and Berry et al. (1995) the solution to equation (3.25) can be solved using a fixed point iteration:

$$\delta^{h.new}_{j,t} = \delta^{h.old}_{j,t} + \psi^h \cdot \left(\log(s^h_{j,t}) - \log\left(\hat{s}^h_{j,t}(\delta^h_{j,t})\right)\right), \tag{3.26}$$

where ψ^h is a tuning parameter set to 0.6.

After finding the vector $\delta^h_{j,t}$ that satisfies the Bellman equation (3.9), the market evolution equation (3.10), and makes the predicted shares equal the observed shares, we compute the home video value for each movie j and medium k, DVD or Blu-ray, at the beginning of this market. We then use this value, $V^h_{k,j}$, as the terminal continuation value for the box office market as used in equation (3.13).

3.5.2 Box office shares

Once the home video market is solved, the home video values enter the terminal continuation values in the box office market and we seek a fixed point between equations (3.12), (3.13), (3.14) and (3.18). This is done for DVD and Blu-ray player owners separately, and then we use (3.19) and (3.20) to compute the predicted box office market shares.

Like in the home video market, we wish to find a vector $\delta^b_{j,t}$ for which the predicted shares equals the observed shares for each movie and time period, namely

$$s^b_{j,t} = \hat{s}^b_{jt}(\delta^b_{j,t}) \quad \forall j, t. \tag{3.27}$$

We solve (3.27) by iterating over

$$\delta^{b.new}_{jt} = \delta^{b.old}_{jt} + \psi^{box} \cdot \left(\log(s^b_{j,t}) - \log\left(\hat{s}^b_{j,t}(\delta^b_{j,t})\right)\right), \tag{3.28}$$

where ψ^h is a tuning parameter set to 0.6.

3.5.3 Recovery of $\vec{\xi}$, $\vec{\alpha}^h$ and $\vec{\alpha}^b$

We use the estimated δ^b and δ^h on a set of regressions involving different movie characteristics. We begin by exposing the retrieval of characteristic coefficients for the home video market, and we proceed with the box office market. By the end of this subsection we describe our procedure to generate advertising and poster color covariates.

Home video

To recover the unobserved characteristics ξ^{dvd} and ξ^{blu}, which are required to compute the GMM objective function (3.24), we regress δ_{hv} as the purchase utility from equations (3.1) and (3.3) on a set covariates. The covariates involve movie-technology specific dummy variables [8] ($\alpha_j^{fe,h}$), age and the squared age of the movie in weeks, goodwill advertising stock, price, and, month and year dummies. The formation of the goodwill advertising stock follows that of Dubé et al. (2005) and it is explained in detail in subsection 3.5.3. For each covariate, we create a new one that is multiplied by a Blu-ray dummy as shown in (3.3).

Like other studies of market power since Bresnahan (1981), we allow price to be endogenous to unobserved characteristics (ξ^h), but we assume that movie characteristics are exogenous. This assumption is justified when movie characteristics are determined in advance, independently of unobserved ones at the moment the home videos are sold. As it is common in the literature, we use lagged prices and price differences from the mean as instruments in a two stage least squares regression.

This first stage regression to identify the contribution of time-dependent characteristics, must be performed after every fixed point on $\delta_{j,t}^h$ is achieved. This is because we need $\alpha^{p,h}$ to obtain $\alpha^{p,b}$ according to equation (3.21) in order to obtain the flow utilities for the box office after finding a fixed point in such market.

A second stage regression of our model can be performed after estimating $\hat{\phi}$ to recover estimates of non-time varying characteristics. This involves regressing the fixed effects obtained in the first stage regression with movie specific characteristics such as logarithm of total first three weeks box

[8] A movie that is both on DVD and Bluray will have separate dummy variables for the panel rows that correspond to DVD and Bluray.

office revenue and its squared term, home window and its square term, logarithm of the production budget for the film, distributor dummies, and poster color dummies (subsection 3.5.3 explains in detail the creation of the poster color dummies). Following Nevo (2000), we perform a minimum distance procedure, let $y = (y_1, \ldots, y_J)'$ denote the $J \times 1$ vector of fixed movie-technology coefficients ($\alpha_j^{fe,h}$) from Equation (3.3), X be the $J \times K(K < J)$ matrix of movie characteristics, and $\xi^{fe,h}$ is the movie specific deviation of the unobserved characteristics. Then, we have

$$y = X\beta^h + \xi^{fe,h}. \tag{3.29}$$

We do not make any assumptions on the error variance covariance matrix (Ω) since we can compute it from our first stage regression, thus, instead of of using an Ordinary Least Squares (OLS) procedure we perform a Generalized Least Squares (GLS) one. The GLS estimator is defined as

$$\beta_{GLS}^h = \arg\min_b (y - Xb)'\Omega^{-1}(y - Xb), \tag{3.30}$$

which can be rewritten as $\arg\min_b[\Omega^{-1/2}(y - Xb)]'[\Omega^{-1/2}(y - Xb)]$. This can be seen as OLS objective function of $\tilde{y} = \tilde{X}b + \tilde{\xi}^{fe,h}$, with $\tilde{y} \doteq \Omega^{-1/2}y$, $\tilde{X} \doteq \Omega^{-1/2}X$ and $\tilde{\xi}^{fe,h} \doteq \Omega^{-1/2}\xi^{fe,h}$. Thus, the GLS estimator can be written as $\hat{\beta}_{GLS}^h = (\tilde{X}'\tilde{X})^{-1}\tilde{X}'\tilde{y} = (X'\Omega^{-1}X)^{-1}X\Omega^{-1}y$, with $\hat{\xi}^{fe,h} = \hat{y} - X\hat{\beta}_{GLS}^h$. Furthermore, we can write the variance of the GLS estimator as $VAR(\hat{\beta}_{GLS}^h) = (X'\Omega X)^{-1}$.

Box office

For each value of ϕ, we impose constraint (3.21) and regress $\hat{\delta}_{j,t}^b \doteq \delta_{j,t}^b - \phi\alpha^{p,dvd}p_{j,t}^b$ as the flow utilities in equation (3.6), which involves movie specific dummy variables, the logarithm of the first three week revenue[9], age and the squared age of the movie in weeks, goodwill advertising stock and current month. This yields $\xi^b(\phi)$ which allows for the computation of the GMM objective function (3.24). We refrain from using a two stage least squares regression as used in the home video market because the endogeneity of price has already been subtracted with the use of the utility scaling parameter.

We then perform a second stage regression that finds the taste components for the movie specific

[9]For the first week we set this covariate as 0, while for weeks 2 and 3 we add all revenue in previous weeks until then.

characteristics, i.e. release year, logarithm of production budget, distributor dummies, and poster color dummies. We use the same GLS regression as in the home video second stage regression.

The procedure to compute the GMM objective function and estimate ϕ may be summarized as follows:

1. Recover $\hat{\delta}^b_{j,t} = \delta^b_{j,t} - \phi \alpha^{p,dvd} p^b_{j,t}$ given ϕ.

2. Run a movie fixed effects regression of $\hat{\delta}^b_{j,t}$ on time dependent movie characteristics to estimate $\alpha^{fe,b}_j$ and $\alpha^{x,b}$ from equation. (3.6)

3. Compute $\xi^b_{j,t}(\vec{\alpha}^b(\phi)) = \hat{\delta}^b_{j,t} - \alpha^{fe,b}_j + \alpha^{x,b} x^b_{j,t}$.

4. Construct $\xi(\vec{\alpha}^h, \vec{\alpha}^b, \phi)$ and compute the objective function of equation (3.24).

Goodwill advertising stock

Following Dubé et al. (2005), we implement a simple advertising model that captures the "carry-over" of advertising to posterior periods. Let $A_{j,t}$, $g_{j,t}$ and $g^a_{j,t}$ denote the advertising expenditure, goodwill stock and augmented goodwill stock for movie j in period t. The augmented goodwill stock is what enters in the consumers' utility function, and it is increased by advertising over an already present goodwill stock:

$$g^a_{j,t} = g_{j,t} + \psi(A_{j,t}), \tag{3.31}$$

where ψ is the goodwill production function. Dubé et al. (2005) discuss some possibilities for ψ, we particularly assume that $\psi(x) = \log(1 + x)$. Augmented goodwill stock in period t depreciates over time, with a discount rate $\lambda \in (0, 1)$ per period, and becomes the beginning of goodwill stock in period $t + 1$:

$$g_{j,t+1} = \lambda g^a_{j,t}. \tag{3.32}$$

Our advertising data shows advertising expenditure per movie per month. Given the difficulty of estimating discount rates, we assume $\lambda = 0.75$. Once we find the augmented goodwill for a given month, we use that amount for all weeks that start during such month.

Poster Colors

In this section we describe the procedure to apply color theory in consumer preferences for movie posters. Several online articles discuss the importance of color choices in marketing, see Morton (2012), O'Grady (2019) and Hauff (2018), suggesting that color theory might be used as a persuasion mechanism to increase purchases. Generally, color theory suggests that the use of complementary colors is used to drive sales.

In order to apply color theory in our model, we downloaded poster images with a resolution of 500×750 for each movie in the data using the API at https://www.themoviedb.org/. Following Ivasic-Kos et al. (2014), we extract color information from each movie poster by transforming the RGB values of pixels to HSV (hue/saturation/value) and extracting its hue. Then, we transform the hue from each pixel to a color by using a discretized 12 color palette, and finally we obtain a 12-color spectrum for each movie poster. As an example, Figure 3.3 shows the hue spectrum obtained for the poster of the movie "Iron Man 2". We can see that this poster makes use of complementary colors - two hues positioned exactly six spaces away from each other - using orange and azure to attract viewers. This a common practice in action pictures - most of them make use of red/orange/yellow explosions and contrast it with some form cyan/azure/blue.

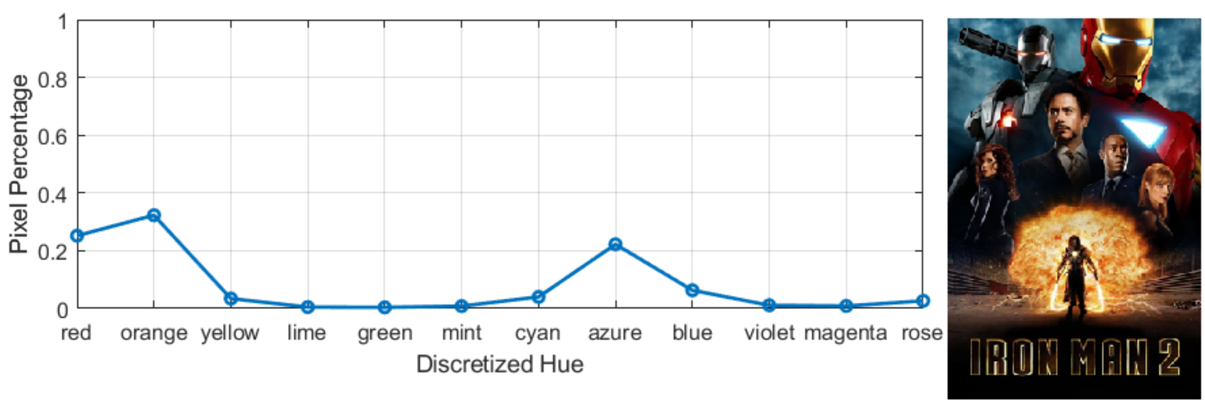

Figure 3.3: Discretized hue spectrum for the poster of the movie "Iron Man 2" seen on the right.

In order to identify consumer preference for different poster colors and to identify preference for the use of complementary colors, we first create a dummy variable for the color peak in each movie spectrum. We then create two different covariates:

1. peak color strength: we generate it by multiplying a peak color dummy with the percentage

of pixels that belong to this color in the poster.

2. complement color strength: we generate it by multiplying a peak color dummy with the percentage of pixels that belong to this color's complement in the poster.

3.6 Results

We present our time dependent parameter estimates in Table 3.5. For all markets, the average price coefficient is negative, which means consumers have marginal disutility towards price. Consumers dislike consuming older products, as seen with the negative values in the age coefficients. Advertising affects utility with a positive effect on all markets, and based on the coefficients it affects Blu-ray purchase utility the most, followed by DVDs and then box office tickets. The "lag 1st3weeks box revenue" variable represents the total box office lagged revenue until a given period or week 4, whichever comes first. The overall contribution of the linear and quadratic logarithmic terms of this variable is increasing with revenue. This means that movies that perform better during the first few weeks since release drive purchase utility up for the upcoming weeks.

Time Dependent Characteristics

Variable	**Box Office**	**Home Video**	**Blu-ray Indicator**
log(lag 1st3weeks box revenue)	0.2356 (0.0459)***	-	-
log(lag 1st3weeks box revenue)^2	-0.0126 (0.0026)***	-	-
Age (weeks)	-0.8158 (0.0157)***	-0.0154 (0.0042)***	-0.0004 (0.0069)
Age^2 (weeks^2)	0.0221 (0.0009)***	0.0000 (0.0000)***	-0.0000 (0.0000)***
Advertising	0.5048 (0.1334)***	0.6645 (0.0009)***	0.1043 (0.0159)***
Price	-0.0113 (-)	-0.0330 (0.002)***	-0.0194 (0.0037)***
ϕ	0.3421 (0.2705)	-	-
Month fixed effects	✓	✓	✓
Year fixed effects		✓	✓
Movie medium fixed effects	✓	✓	✓
N	1,797	70,253	25,373

$^{***}p < .01$, $^{**}p < .05$, $^{*}p < .1$

Table 3.5: Time dependent parameter estimates.

We present time independent parameter estimates in Table 3.6. The estimates for the logarithm opening box office revenue show a convex response for the utility, and specifically for the ranges of box office revenue dealt in the data, it is increasing for both DVDs and Blu-rays. This means that the better a movie performs in theaters, the better it will perform in the home video market. The home video window coefficients for DVD and Blu-ray show a concave response to utility. DVDs present a maximum at 5.2 weeks, while Blu-rays present a maximum 9.1 weeks. These non-zero maximums could be capturing the effects of word of mouth, and it is important to note that this is the case when everything else is held constant, as the advertising spillover effect from theaters will most likely shift these maximums to lower values. The production budget estimate for the box office and Blu-rays is positive, suggesting that higher production budget films generate greater revenue for the box office and Blu-rays. This effect is opposite for DVDs, showing that increasing the production budget reduces the DVD revenue.

Our results on movie colors are significant and present some interesting features. The use of yellow and its complement in a poster yields the largest contribution to purchase utility for both box office and Blu-rays. This is not surprising, as action movies exhibit explosions and fires in their posters, (which contribute to the yellow hue), and use a complementary color to make these stand out. What is interesting to note is that for DVDs, the contribution of yellow and complement is negative, which suggests that DVD consumer behavior is different from the behavior of box office and Blu-ray consumers.

3.7 Counterfactuals

For this section we use the 113 movies for which we have DVD and Blu-ray sales data. We modify the data by setting the DVD and Blu-ray home windows to a specific number of weeks NT_{dvd} and NT_{blu}, respectively. At the same time, we readjust the advertising goodwill by discounting through longer or shorter periods of time. For each market (box office and home video), we must reach an equilibrium described in Section 3.4.3, the Consumer's problem. As both markets depend on the box office revenue during the first three weeks of the theatrical run, and the box office market terminal continuation value depends on the home video value, this becomes a challenge. We begin with a guess of the box office revenue during the first three weeks of the theatrical run, and solve the home video market. We use that home video value and box office revenue to solve the box office market. We iterate between both markets until we find a fixed point in box office revenues and home video values. This must be done for each counterfactual set of values of the home video

Time Independent Characteristics

Variable	Box Office	DVD	Blu-ray
log(1st3weeks box revenue)	-	-3.8676 (0.0884)**	-5.2108 (0.1340)***
log(1st3weeks box revenue)^2	-	0.1226 (0.0027)***	0.1580 (0.0041)***
Home window (weeks)	-	0.0263 (0.0065)***	0.0752 (0.0076)***
Home window^2 (weeks^2)	-	-0.0025 (0.0004)***	-0.0042 (0.0005)***
log(production budget)	0.4902 (0.0324)***	-0.2251 (0.0122)***	0.1652 (0.0187)**
Red indicator × %	0.5099 (0.1055)***	0.1034 (0.0371)***	0.1120 (0.0614)**
Orange indicator × %	0.0635 (0.1146)	0.0797 (0.0384)**	-0.0345 (0.0729)
Yellow indicator × %	-0.3011 (0.2506)	0.2615 (0.0803)***	1.7793 (0.1225)***
Cyan indicator × %	0.5175 (0.1822)***	0.6823 (0.0807)***	1.5314 (0.0985)***
Azure indicator × %	0.2789 (0.0962)***	-0.0409 (0.0404)	0.2504 (0.0618)***
Red indicator × complement%	-2.8862 (0.7920)***	6.1720 (0.3358)***	-0.5325 (0.4289)
Orange indicator × complement%	-1.1732 (0.4195)***	1.0549 (0.1668)***	2.3883 (0.2000)***
Yellow indicator × complement%	14.1576 (3.2520)***	-3.3738 (1.0169)***	21.6641 (1.1852)***
Cyan indicator × complement%	-5.8510 (0.7074)***	-1.31377 (0.2467)***	-1.9363 (0.2581)***
Azure indicator × complement%	-2.6716 (0.3524)***	0.4767 (0.1268)***	0.9212 (0.1793)***
Distributor fixed effects	✓	✓	✓
Release year fixed effects	✓	✓	✓
N	149	149	113

$^{***}p < .01$, $^{**}p < 0.05$, $^{*}p < .1$

Table 3.6: Time dependent parameter estimates.

window.

We first analyze the case where both technological qualities, DVD and Blu-ray, are released simultaneously, and then we allow for versioning, where each of them may have different release strategies. We take the point of view of the studios and try to maximize their revenue from both box office and home videos. In order to split the box office revenue between studios and theaters we impose a standard contract (Vogel 2014). This contract involves a house nut of $2,000 per theater per week, and after subtracting the nut the split starts at 70% in favor of the studios, decreasing 10% every two weeks until it reaches 0%. These contracts generate incentives for theaters to exhibit movies for a longer period of time, as the share they keep from box office revenue is greater with the age of the movie.

For each movie, we find the optimal home video windows and then we compute the average across optimals.

3.7.1 Simultaneous DVD and Blu-ray Release

In this section we describe the counterfactual analysis in which the DVDs and Blu-rays are released simultaneously, as this was the current industry practice during the data time period. Figure 3.4 shows a histogram of the home video windows for the data, and for the studio optimals. As we see, for most movies it is optimal to shrink the home video window to a range between 0 to 4 weeks. This means that advertising spillover effect from theaters and the home video freshness are dominating over the increased competition to theaters that an early home video release may provide. Furthermore, the demand cannibalization of the box office, generated by this early release, reduces the box office signal which could impact home videos, but again the advertising spillover effect dominates.

Figure 3.5 shows the studio revenue from the box office and home videos for a particular movie as a function of the home video window. We see a steep increase in the home video revenue from shrinking the home video window, while the demand cannibalization to theaters is minimal. Table 3.7 shows the average home video window, studio revenue and theater revenue, under the optimized strategy and the data. We can see that the average optimal strategy shrinks the home video window to about 2.3 weeks, providing an increased studio revenue of 4.47% and an increase in theater revenue of 0.04%.

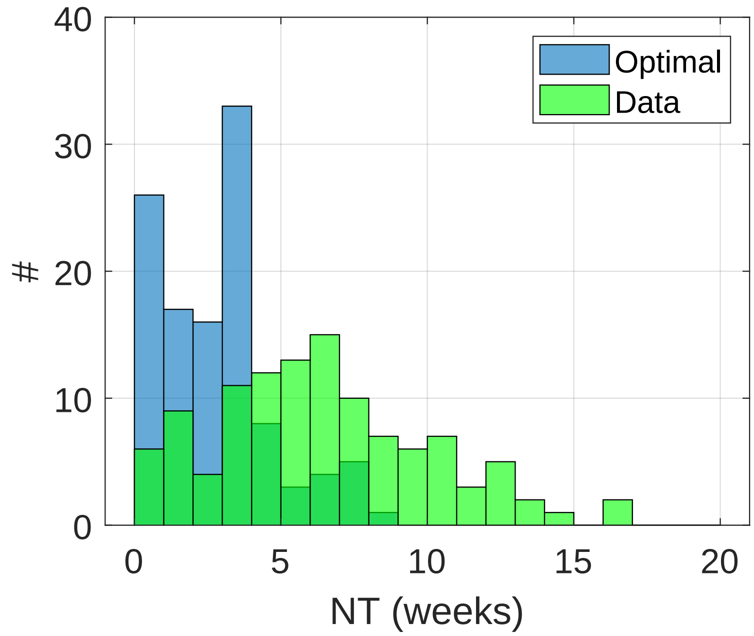

Figure 3.4: Histogram of the home window lengths, for the data and the optimal.

Simultaneous Release Revenues

Averages	**NT (weeks)**	**Studio Revenue ($)**	**Theater Revenue ($)**
Optimal Simultaneous	2.3186	124.34M (+4.47%)	64.59M (+0.08%)
Data	6.1960	119.02M	64.45M

Table 3.7: Home video windows and revenue information for the optimal simultaneous release strategy and the data. Studio revenue accounts for the home video window as well as the box office revenue studio share, while theater revenue accounts for the portion or box office revenue that corresponds to theaters.

In the following section we allow for separation between DVD and Blu-ray releases and analyze its benefits.

3.7.2 Versioning: Separate DVD and Blu-ray releases.

We now analyze the counterfactual in which the home video window may differ for different technological qualities. We denote NT_{dvd} and NT_{blu} as the home video window for DVD and Blu-ray respectively.

In order to ensure that Blu-ray player owners benefit from choosing a Blu-ray over a DVD for each movie, we must impose an incentive compatibility constraint. In general, the value of the Blu-

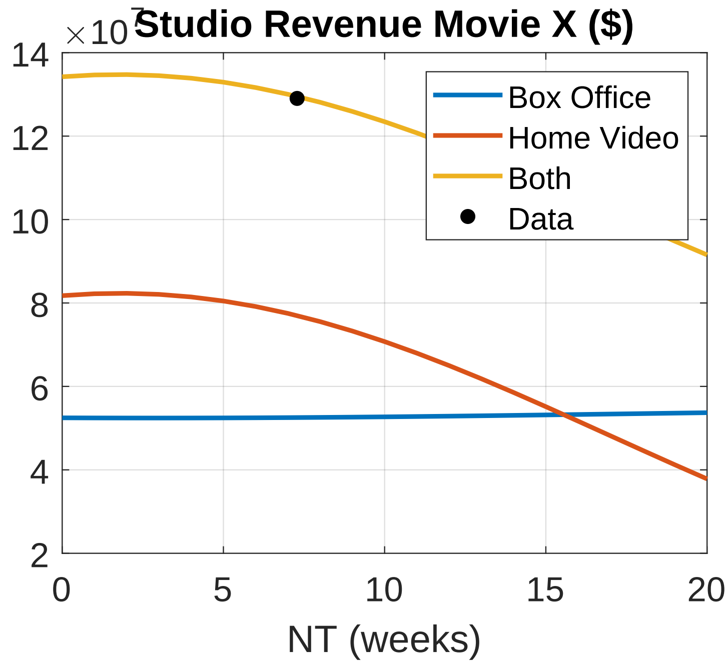

Figure 3.5: Studio revenue from the box office, home videos, and both, as a function of the home video window for a particular movie. The gray circle represents the studio revenue at the data home video window of 7 weeks.

ray is larger than value of the DVD, but if the Blu-ray release was to be delayed, some consumers might switch to the DVD. By imposing this constraint we ensure that our market size for Blu-rays would be better-off choosing Blu-rays over DVDs. This constraint is as follows:

$$\beta^{NT_{dvd}} V_{dvd} \leq \beta^{NT_{blu}} V_{blu}, \tag{3.33}$$

the discounted ex-ante value of the Blu-ray has to be greater than or equal to the discounted ex-ante value of the DVD. This is imposing an upper bound on the difference between home video windows. We can rewrite constraint (3.33) to

$$NT_{blu} - NT_{dvd} \leq \frac{\log\left(\frac{V_{blu}}{V_{dvd}}\right)}{|\log \beta|}. \tag{3.34}$$

We now search over (NT_{dvd}, NT_{blu}) to find the optimal studio revenue satisfying the incentive compatibility constraint (3.34). Figure 3.6 shows the histogram of home video windows for the data, the DVD optimal and the Blu-ray optimal. We can see a clear difference with the simultaneous release from Figure 3.4. Now it is optimal to have an immediate after theater release for DVDs, while it is optimal to delay the Blu-ray release to an average of about 5 weeks. We see that for DVDs, the advertising spillover from theaters offsets the increased competition. For Blu-rays, these two effects are more balanced. This happens because the ex-ante value function for DVDs is smaller than the one for Blu-rays, so shrinking the home video window for DVDs has very little effect on box office purchases. Then you can shrink the window and reap the benefits of the advertising spillover effect. This result supports the idea that higher technological quality home videos are closer substitutes to Blu-rays. Shrinking the home video window for Blu-rays will not only impact the box office demand, but it will reduce the box office revenue, reducing the movie quality signal for the home video market (DVD and Blu-ray).

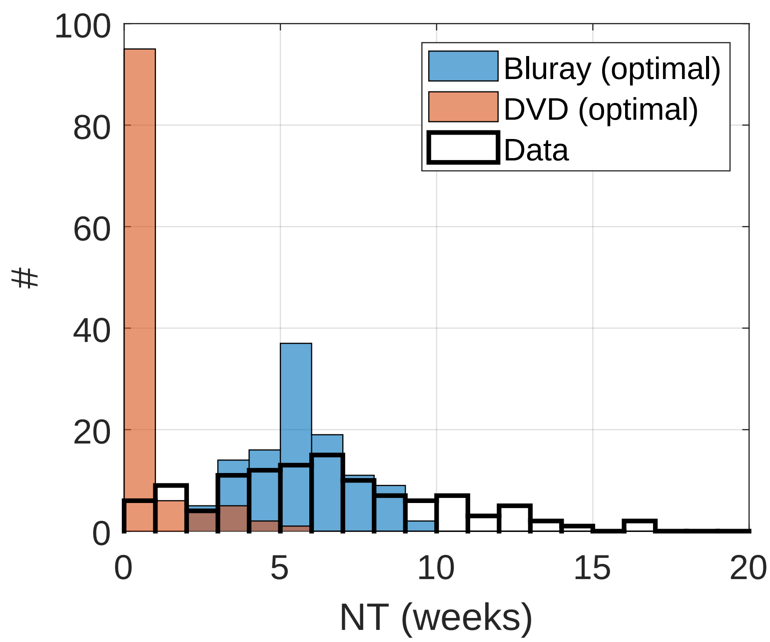

Figure 3.6: Histogram of the home window lengths, for the data, and the optimal ones for DVD and Blu-ray.

Figure 3.7 shows the optimal studio revenue as a function of the DVD and Blu-ray home video windows for the same movie as in Figure 3.5. We clearly see that there is a greater steepness in the DVD home video window axis compared to that of the Blu-ray one. This is because DVD and theaters experience very little competition between each other, and shortening the home video window allows DVD to benefit from the advertising spillover effect. While for Blu-rays, we see lower steepness in the revenue because the trade-off is more balanced. Table 3.8 illustrates the new average optimal release windows, studio revenues and theater revenues. We can see that the average optimal DVD and Blu-ray windows are about 0.4 and 5.2 weeks respectively. The studio revenue increases almost an extra 1% with respect to the simultaneous release in Table 3.7, while the theater revenue remains almost the same. This result highlights the benefit of exploiting market segmentation strategies in a consumer base that expresses heterogeneous preferences on home video technological qualities. Further improvements could be made by optimizing over advertising periods, and home video pricing, but that is out of the scope of this chapter.

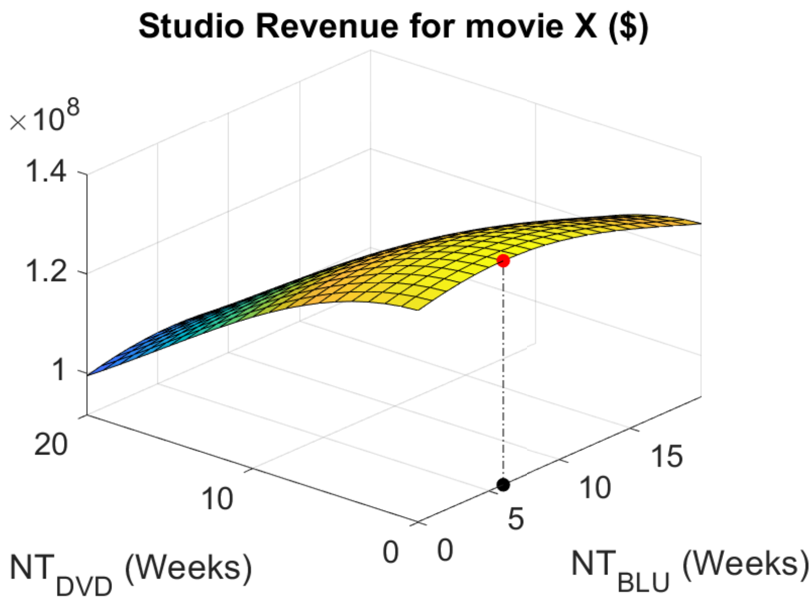

Figure 3.7: Studio revenue from the box office, home videos, and both, as a function of the home video window for a particular movie. The gray circle represents the studio revenue at the data home video window of 7 weeks.

Separate Release Revenues

Averages	**NT (weeks)**	**Studio Revenue ($)**	**Theater Revenue ($)**
Optimal (DVD,Blu-ray)	(0.3717, 5.1504)	125.40M (+5.36%)	64.56M (+0.03%)
Data	6.1960	119.02M	64.45M

Table 3.8: Home video windows and revenue information for the optimal separate release strategies and the data. Studio revenue accounts for the home video window as well as the box office revenue studio share, while theater revenue accounts for the portion or box office revenue that corresponds to theaters.

3.8 Conclusion

In this chapter we build and estimate a dynamic discrete choice model about movie distribution involving the box office, DVDs, and Blu-rays. We capture consumers value for delaying purchase decisions, allowing for substitution between box office tickets and home videos. We run counterfactuals on the home video windows and we find that releasing DVDs 0.4 weeks after the theatrical run, and Blu-rays about 5.2 weeks after, is optimal on average. This strategy achieves an average revenue increase of 5.36% for the studios with respect to the current practice, while having minimal impact on theaters. This analysis suggests that higher technological quality home videos are closer substitutes to theater, and their release balances advertising spillover effect from theaters with demand cannibalization. Releasing higher technological quality home videos early cannibalizes theater demand, which impacts theatrical revenue and further impacts all home video markets. For lower technological quality home videos such as DVDs, this is not the case, as they don't compete with theaters, one can release them early reaping all the benefits of the advertising spillover effect. These results highlight the benefit of exploiting market segmentation strategies in a consumer base that expresses heterogeneous preferences on home video technological qualities.

Chapter 4

Conclusion

In this dissertation, we study three problems that arise in the entertainment industry and digital markets due to technological innovations. We model and study the strategic interactions in the entertainment industry to better inform business practice, using tools from game theory, structural models, and optimization.

In Chapter 1, we analyze the impact of adding consumer heterogeneity to a simple trial offer market that follows a Multinomial Logit Model and has been empirically validated (Krumme et al. 2012). We detail several complexities that arise due to consumer heterogeneity, including the NP-hardness of obtaining the dynamic product ranking that maximizes the number of purchases in the next period, as well as the negative impacts of the use of social signals. We study a simple ranking which ranks products by the weighted average quality of products (AQGSI), and we show that with social influence it converges predictably to a monopoly for the product with the largest weighted average quality. We then analyze the impact of market segmentation, by showing a different ranking to each consumer segment, and we show that the platform owner can improve market efficiency by up to a factor K, where K is the number of consumer segments.

Our study reveals that in trial offer markets, the decision on whether or not to display the popularity signal to consumers has to be analyzed very carefully, as displaying past purchases may be detrimental to the rate of purchases. We also show that segmentation policies always improve market efficiency, and they should be used when possible. However, if consumer classification mistakes are likely, the platform owner should rethink whether market segmentation is the best strategy.

In Chapter 2, we analyze the impact that release timing choice may have on advertising video on demand platforms. We build a signaling model that involves a unique show and a group of consumers, where shows need to decide the release timing strategy for episodes as well as an advertising expenditure. The release timing strategy affects consumers in their viewing cost, whereas the advertising level persuades viewers to watch the show. The show generates revenue through advertising within the show that depends on the chosen release timing strategy, as binge watching behavior reduces the efficacy of advertising. We show that under certain scenarios, high quality shows may use a sequential release timing strategy in conjunction with a suitable advertising level to credibly inform consumers about their show quality. Furthermore, we analyze a situation in which, under complete information, both shows pool on simultaneous releases, whereas when information is private, the high quality show uses the sequential releases channel in order to achieve minimum cost separation. When the simultaneous release timing channel was unavailable, as seen in traditional TV, the minimum cost separation was achieved using advertising alone, and it was larger than the separation incurred when the simultaneous release strategy became available.

Our results have an important managerial implication; because release timing may signal quality, it is beneficial for a firm to find a way in which to open new channels that are profitable for its lower quality content. By doing this, the firm reduces the necessity of costly signaling through other mechanisms (i.e., advertising) for their higher quality content, and enjoys greater profits on all quality levels.

In Chapter 3, we analyze the optimal time between theatrical and home video releases, as the technological quality of the later increases. We build and estimate a dynamic discrete choice model about movie distribution involving the box office, DVD, and Blu-ray markets. We connect the box office and home video markets through the value of waiting and the theatrical revenue, which is known to be a driver for home video demand. We run counterfactuals on the home video windows, the time between theatrical exit and home video release, and we find that releasing DVDs 0.4 weeks after the theatrical run, and Blu-rays about 5.2 weeks after, is optimal on average. This strategy achieves an average revenue increase of 5.36% for the studios with respect to the current practice, while having minimal impact on theaters. This analysis suggests that higher technological quality home videos are closer substitutes to theater, and their release balances advertising spillover effect from theaters with demand cannibalization. Releasing higher technological quality home videos early cannibalizes theater demand, which impacts theatrical revenue and further impacts all home video markets. For lower technological quality home videos such as DVDs, this is not the case. As lower technological quality home videos do not compete with theaters, one can release them early reaping all the benefits of the advertising spillover effect. These results highlight the benefit of exploiting market segmentation strategies in a consumer base that expresses heterogeneous preferences on home video technological qualities.

This dissertation leaves room for further exploration. The main weakness of Chapter 1 is that the consumer choice model only depends on the displayed vector of past purchases and the current ranking (as well as the appeal of the products). A more sophisticated choice model could incorporate the type of past purchase information displayed: A consumer who observes a past purchase vector d might change their behavior depending on whether she/he knows that the vector d comes from all purchases or only from its own consumer segment. Another potential avenue for research is to analyze the same model when the trial probabilities depend non-linearly on past purchases; this is studied in Maldonado et al. (2018) for the special case of a unique consumer segment ($K = 1$). A final research avenue may involve studying the potential incentives to hide or mis-report some reviews, and how it this affects the outcomes in terms of market share dynamics and consumer welfare.

Chapter 2 is a good start for multi-period signaling models that incorporate binge watching behavior in a consumer-show interaction setting. A potential avenue for research, however complicated, is to study the effects of platform competition under the same setting. This is becoming particularly important in an environment where there are multiple advertising video on demand platforms available, such as Pluto TV, Peacock, Roku Inc. and Tubi.

Chapter 3, also has opportunity for further research. One could assess the impact that piracy has on optimal home video windows, as studios are trying to shrink windows due to the potential harm of pirate markets. One could also study the optimal time to video on demand services, as well as subscription based streaming platforms (e.g HBO Max).

Appendix A

Additional Material for Chapter 1

A.1 Proofs

A.1.1 Proof of Theorem 1.1

The proof uses the 2-Class Logit problem which is known to be NP-hard (Rusmevichientong et al. 2014). The inputs to a 2-Class Logit instance are N products, two sequences $V^1 = (V_1^1, V_2^1, \ldots, V_N^1)$ and $V^2 = (V_1^2, V_2^2, \ldots, V_N^2)$ with $V^1, V^2 \in \mathbb{Q}_+^N$, and a number $\alpha \in \mathbb{R}_{[0,1]}$. Each product i has a revenue $r_i \in \mathbb{Z}_+$. Each sequence V^i represents a realization of the product utilities under a multinomial logit model. Sequence V^1 (resp. V^2) has a realization probability of α (resp. $1 - \alpha$). The problem consists in finding a product assortment $S \subseteq [N]$ maximizing the expected revenue Π^{Logit}, i.e.,

$$\Pi^{Logit} = \max_{S \subseteq [N]} \alpha_1 \frac{\sum_{i \in S} r_i V_i^1}{1 + \sum_{i \in S} V_i^1} + (1 - \alpha) \frac{\sum_{i \in S} r_i V_i^2}{1 + \sum_{i \in S} V_i^2}.$$

The proof shows that, if there exists an oracle to compute the performance ranking for the MMNL with two classes of consumers (i.e., Equation (1.2) with $K = 2$), then the 2-Class logit problem can be solved in polynomial time.

Given an instance of the 2-Class Logit problem, the idea is to create N different instances of the performance-ranking problem in order to capture the various possible assortments. The N instances have a common core. Each of them has the same N items and two classes of consumers (i.e., $K = 2$). For each consumer segment $j \in \{0, 1\}$ and each item $i \in [N]$, we set the appeal of item i for segment j to satisfy $a_{i,j} = V_i^j$. Similarly, for each consumer segment $j \in \{0, 1\}$ and each item $i \in [N]$, we set the quality of i for segment j to satisfy $q_{i,j} = r_i$. Note that the quality of item i is the same for both classes. The weights of classes 0 and 1 are α and $1-\alpha$ respectively. We also set $z = 1$ and $t = 0$ which implies that $d^t = 0$. The N instances differ in the position visibilities. In instance i ($i \in [N]$), the visibility of position $j \in [N]$ is:

$$v_j = \begin{cases} 1 & \text{if } j \leq i \\ 0 & \text{otherwise.} \end{cases}$$

Let Π_i^{PR} denote the short-term optimal value of the performance ranking for problem instance i and let $\mathcal{S}_i$ denote the collection of all possible subsets of products whose size is i, i.e., $\mathcal{S}_i = \{S \subseteq [N] : |S| = i\}$. Define Π_i^{Logit} as the following optimization problem:

$$\Pi_i^{Logit} = \max_{S \in \mathcal{S}_i} \alpha_1 \frac{\sum_{i \in S} r_i V_i^1}{1 + \sum_{i \in S} V_i^1} + (1-\alpha) \frac{\sum_{i \in S} r_i V_i^2}{1 + \sum_{i \in S} V_i^2}.$$

It follows that

$$\Pi^{Logit} = \max_{i=1,\ldots,N} \Pi_i^{Logit}. \tag{A.1}$$

We now show that Π_i^{PR} is equal to Π_i^{Logit}.

$$\begin{aligned}
\Pi_i^{PR} &= \max_{\sigma \in S_n} \left\{ \sum_{c=1}^{2} \left(w_c \cdot \sum_{\ell=1}^{N} (p_i(\sigma, 0) \cdot q_{\ell,c}) \right) \right\}, && \text{(A.2)} \\
&= \max_{\sigma \in S_n} \left\{ \sum_{c=1}^{2} \left(w_c \cdot \sum_{\ell=1}^{N} \left(\frac{v_{\sigma(\ell)}(a_{\ell,k})}{\sum_{j=1}^{N} v_{\sigma(j)}(a_{j,c}) + 1} \cdot q_{\ell,k} \right) \right) \right\}. && \text{(A.3)} \\
&= \max_{\sigma \in S_n} \left\{ \alpha \cdot \sum_{\ell=1}^{N} \left(\frac{v_{\sigma(\ell)} V_\ell^1 r_\ell}{\sum_{j=1}^{N} v_{\sigma(j)} V_j^1 + 1} \right) + (1-\alpha) \cdot \sum_{\ell=1}^{N} \left(\frac{v_{\sigma(\ell)} V_\ell^2 r_\ell}{\sum_{j=1}^{N} v_{\sigma(j)} V_j^2 + 1} \right) \right\} && \text{(A.4)} \\
&= \max_{S \in \mathcal{S}_i} \left\{ \alpha \cdot \sum_{\ell \in S} \left(\frac{V_\ell^1 r_\ell}{\sum_{j \in S} V_j^1 + 1} \right) + (1-\alpha) \cdot \sum_{\ell \in S} \left(\frac{V_\ell^2 r_\ell}{\sum_{j \in S} V_j^2 + 1} \right) \right\} && \text{(A.5)} \\
&= \Pi_i^{Logit} && \text{(A.6)}
\end{aligned}$$

where the equivalence between (A.4) and (A.5) follows from the fact that the first i positions have visibility of 1 and the remaining ones have a visibility of 0 and therefore selecting a permutation $\sigma \in S_n$ reduces to deciding which i items should be assigned the top i positions. As a consequence, using (A.1), we have

$$\Pi^{Logit} = \max_{i=1,\dots,N} \Pi_i^{Logit} = \max_{i=1,\dots,N} \Pi_i^{PR}. \tag{A.7}$$

We have shown that, by using an oracle to solve N instances of the performance-ranking problem, it is possible to solve the original 2-class logit problem instance in polynomial time. Hence, the performance ranking is NP-hard under Turing reductions. □

A.1.2 Proof of Lemma 1.1

The market share of item i^* at any period of time $t > \hat{t}$ for this system would be underestimated by considering the following set of qualities and appeals:

$$q_{i,new} = \begin{cases} q_{i,min} & \text{if } i = i^* \\ q_{i,max} & \text{if } i \neq i^* \end{cases} \text{ and } a_{i,new} = \begin{cases} a_{i,min} & \text{if } i = i^* \\ a_{i,max} & \text{if } i \neq i^* \end{cases}.$$

If this new set of qualities satisfies that $v_{\sigma(i^*)} q_{i^*,new} > v_{\sigma(i)} q_{i,new}$ for all $i \in [N] \setminus \{i^*\}$, it follows from the convergence result in (Van Hentenryck et al. 2016a) (Theorem 4.3) that the system goes to a monopoly for item i^*. Therefore, the original system also goes to a monopoly for item i^*. □

A.1.3 Proof of Theorem 1.2

The proof first shows that the MMNL model can be reduced to a Multinomial Logit Model whose item appeals and qualities are functions of the vector of purchases at each time t. It then shows that these functions stay in the bounded range, so that it is possible to apply Lemma 1.1.

When the same ranking σ and popularity signals are shown to all consumers, the probability that item i is purchased in time period t is given by

$$P_i(\sigma, d^t) = \sum_{k=1}^{K} \left(w_k \cdot \left(v_{\sigma(i)} \frac{(a_{i,k} + d_i^t)}{\sum_{j=1}^{N} v_{\sigma(j)}(a_{j,k} + d_j^t) + z_k} \cdot q_{i,k} \right) \right). \tag{A.8}$$

By rearranging the previous expression, it comes

$$
\begin{aligned}
P_i(\sigma, d^t) &= \sum_{k=1}^{K} \frac{w_k q_{i,k} v_{\sigma(i)}}{\sum_{j=1}^{N} v_{\sigma(j)}(a_{j,k} + d_j^t) + z_k} a_{i,k} + \sum_{k=1}^{K} \frac{w_k \ q_{i,k} v_{\sigma(i)}}{\sum_{j=1}^{N} v_{\sigma(j)}(a_{j,k} + d_j^t) + z_k} d_i^t \\
&= \sum_{k=1}^{K} \frac{w_k q_{i,k} v_{\sigma(i)}}{\sum_{j=1}^{N} v_{\sigma(j)}(a_{j,k} + d_j^t) + z_k} a_{i,k} + d_i^t \left(\sum_{k=1}^{K} \frac{w_k \ q_{i,k} v_{\sigma(i)}}{\sum_{j=1}^{N} v_{\sigma(j)}(a_{j,k} + d_j^t) + z_k} \right) \\
&= \left(\sum_{k=1}^{K} \frac{w_k \ q_{i,k} v_{\sigma(i)}}{\sum_{j=1}^{N} v_{\sigma(j)}(a_{j,k} + d_j^t) + z_k} \right) \left(\frac{\left(\sum_{k=1}^{K} \frac{w_k q_{i,k} v_{\sigma(i)}}{\sum_{j=1}^{N} v_{\sigma(j)}(a_{j,k} + d_j^t) + z_k} a_{i,k} \right)}{\left(\sum_{k=1}^{K} \frac{w_k q_{i,k} v_{\sigma(i)}}{\sum_{j=1}^{N} v_{\sigma(j)}(a_{j,k} + d_j^t) + z_k} \right)} + d_i^t \right) \\
&= \left(\sum_{k=1}^{K} \frac{w_k \ q_{i,k} v_{\sigma(i)}}{\sum_{j=1}^{N} v_{\sigma(j)}(a_{j,k} + d_j^t) + z_k} \right) \left(\frac{v_{\sigma(i)} \left(\sum_{k=1}^{K} \frac{w_k q_{i,k} a_{i,k}}{\sum_{j=1}^{N} v_{\sigma(j)}(a_{j,k} + d_j^t) + z_k} \right)}{v_{\sigma(i)} \left(\sum_{k=1}^{K} \frac{w_k q_{i,k}}{\sum_{j=1}^{N} v_{\sigma(j)}(a_{j,k} + d_j^t) + z_k} \right)} + d_i^t \right) \\
&= \left(\sum_{k=1}^{K} \frac{w_k \ q_{i,k} v_{\sigma(i)}}{\sum_{j=1}^{N} v_{\sigma(j)}(a_{j,k} + d_j^t) + z_k} \right) \left(\frac{\left(\sum_{k=1}^{K} \frac{w_k q_{i,k} a_{i,k}}{\sum_{j=1}^{N} v_{\sigma(j)}(a_{j,k} + d_j^t) + z_k} \right)}{\left(\sum_{k=1}^{K} \frac{w_k q_{i,k}}{\sum_{j=1}^{N} v_{\sigma(j)}(a_{j,k} + d_j^t) + z_k} \right)} + d_i^t \right).
\end{aligned}
$$

Now, for each item i and each time period t, define the function

$$
\widetilde{a}_i(t) = \left(\sum_{k=1}^{K} \frac{w_k q_{i,k} a_{i,k}}{\sum_{j=1}^{N} v_{\sigma(j)}(a_{j,k} + d_j^t) + z_k} \right) \Bigg/ \left(\sum_{k=1}^{K} \frac{w_k q_{i,k}}{\sum_{j=1}^{N} v_{\sigma(j)}(a_{j,k} + d_j^t) + z_k} \right) \tag{A.9}
$$

which depends on the total number of purchases at time t. Using this definition, we have that:

$$
P_i(\sigma, d^t) = \left(\sum_{k=1}^{K} \frac{w_k \ q_{i,k} v_{\sigma(i)}}{\sum_{j=1}^{N} v_{\sigma(j)}(a_{j,k} + d_j^t) + z_k} \right) \left(\widetilde{a}_i(t) + d_i^t \right) = \left(\sum_{k=1}^{K} \frac{w_k \ q_{i,k}}{\sum_{j=1}^{N} v_{\sigma(j)}(a_{j,k} + d_j^t) + z_k} \right) v_{\sigma(i)} \left(\widetilde{a}_i(t) + d_i^t \right).
$$

By dividing and multiplying by $\sum_{j=1}^{N} v_{\sigma(j)}(\widetilde{a}_j(t) + d_j^t)$, $P_i(\sigma, d^t)$ becomes

$$
\left(\sum_{k=1}^{K} \frac{w_k \ q_{i,k}}{\sum_{j=1}^{N} v_{\sigma(j)}(a_{j,k} + d_j^t) + z_k} \right) \left(\sum_{j=1}^{N} v_{\sigma(j)}(\widetilde{a}_j(t) + d_j^t) \right) \left(\frac{v_{\sigma(i)} \left(\widetilde{a}_i(t) + d_i^t \right)}{\sum_{j=1}^{N} v_{\sigma(j)}(\widetilde{a}_j(t) + d_j^t)} \right).
$$

Now define the following function for each item i at each time period t:

$$
\widetilde{q}_i(t) = \left(\sum_{k=1}^{K} \frac{w_k \ q_{i,k}}{\sum_{j=1}^{N} v_{\sigma(j)}(a_{j,k} + d_j^t) + z_k} \right) \left(\sum_{j=1}^{N} v_{\sigma(j)}(\widetilde{a}_j(t) + d_j^t) \right). \tag{A.10}
$$

The probability of purchasing product i in the next iteration becomes:

$$P_i(\sigma, d^t) = \left(\frac{v_{\sigma(i)}(\widetilde{a}_i(t) + d_i^t)}{\sum_{j=1}^N v_{\sigma(j)}(\widetilde{a}_j(t) + d_j^t)} \right) \widetilde{q}_i(t).$$

This is almost a multinomial logit model, except that the quality and appeal vectors that depend on time. When the number of iterations t tends to infinity, the total number of purchases $\sum_{j=1}^N d_j^t$ also goes to infinity. Moreover, as t goes to infinity, the generalized appeal $(\widetilde{a}_i(t))$ and quality $(\widetilde{q}_i(t))$ for every item converges to

$$\bar{a}_i \doteq \lim_{t\to\infty} \widetilde{a}_i(t) = \frac{\sum_{k=1}^K w_k a_{i,k} q_{i,k}}{\sum_{k=1}^K w_k q_{i,k}} \quad \text{and} \quad \bar{q}_i \doteq \lim_{t\to\infty} \widetilde{q}_i(t) = \sum_{k=1}^K w_k q_{i,k}. \tag{A.11}$$

In addition, observe that $\widetilde{Q}_i(t) \doteq v_{\sigma(i)} \widetilde{q}_i(t)$ also converges when t goes to infinity:

$$\bar{Q}_i \doteq \lim_{t\to\infty} v_{\sigma(i)} \widetilde{q}_i(t) = v_{\sigma(i)} \bar{q}_i$$

The tie-breaking condition (Equation 1.8) guarantees that there exists only one item i^* such that $i^* = \arg\max_{i\in[N]} \bar{Q}_i$. Let i^{**} be the item with the second highest value $\bar{Q}_i$, i.e., $\bar{Q}_{i^{**}} \geq \bar{Q}_j$ for all $j \in [N], j \neq i^*$. Consider now the following difference $\Delta\bar{Q} = \bar{Q}_{i^*} - \bar{Q}_{i^{**}}$. Equation (A.9) can be seen as a weighted average on k for $a_{i,k}$ and hence

$$\min_{1\leq k\leq K} a_{i,k} \leq \widetilde{a}_i(t) \leq \max_{1\leq k\leq K} a_{i,k} \quad \forall i \in [N], t \in \mathbb{N} \tag{A.12}$$

Moreover, by applying this result to Equation (A.10), we obtain the following bounds for $\widetilde{q}_i$:

$$\begin{aligned}
\widetilde{q}_i(t) &\geq \sum_{k=1}^K \frac{w_k q_{i,k}}{\sum_{j=1}^N v_{\sigma(j)}(\max_{1\leq k\leq K} a_{j,k} + d_j^t) + z_k} \left(\sum_{j=1}^N v_{\sigma(j)}(-\max_{1\leq k\leq K} a_{j,k} + \max_{1\leq k\leq K} a_{j,k} + d_j^t) - z_k + z_k \right) \\
&= \sum_{k=1}^K \frac{w_k q_{i,k} \left(\sum_{j=1}^N v_{\sigma(j)}(\max_{1\leq k\leq K} a_{j,k} + d_j^t) + z_k \right) - w_k q_{i,k} \sum_{j=1}^N v_{\sigma(j)} \max_{1\leq k\leq K} a_{j,k} - w_k q_{i,k} z_k}{\sum_{j=1}^N v_{\sigma(j)}(\max_{1\leq k\leq K} a_{j,k} + d_j^t) + z_k} \\
&\geq \left(1 - \frac{\sum_{j=1}^N v_{\sigma(j)} \max_{1\leq k\leq K} a_{j,k} + \max_{1\leq k\leq K} z_k}{\sum_{j=1}^N v_{\sigma(j)} d_j^t} \right) \sum_{k=1}^K w_k q_{i,k} \\
&= \left(1 - \frac{\sum_{j=1}^N v_{\sigma(j)} \max_{1\leq k\leq K} a_{j,k} + \max_{1\leq k\leq K} z_k}{\sum_{j=1}^N v_{\sigma(j)} d_j^t} \right) \bar{q}_i \text{ and}
\end{aligned}$$

$$\widetilde{q}_i(t) \leq \sum_{k=1}^{K} \frac{w_k q_{i,k}}{\sum_{j=1}^{N} v_{\sigma(j)} d_j^t} \left(\sum_{j=1}^{N} v_{\sigma(j)} (\max_{1 \leq k \leq K} a_{i,k} + d_j^t) \right) \leq \left(1 + \frac{\sum_{j=1}^{N} v_{\sigma(j)} \max_{1 \leq k \leq K} a_{j,k}}{\sum_{j=1}^{N} v_{\sigma(j)} d_j^t} \right) \bar{q}_i.$$

As a result, the bounds for $\widetilde{Q}_i(t)$ $(\forall i \in [1, N], t \in \mathbb{N})$ are given by

$$\left(1 + \frac{\sum_{j=1}^{N} v_{\sigma(j)} \max_{1 \leq k \leq K} a_{j,k}}{\sum_{j=1}^{N} v_{\sigma(j)} d_j^t} \right) \bar{Q}_i \geq \widetilde{Q}_i(t) \geq \left(1 - \frac{\sum_{j=1}^{N} v_{\sigma(j)} \max_{1 \leq k \leq K} a_{j,k} + \max_{1 \leq k \leq K} z_k}{\sum_{j=1}^{N} v_{\sigma(j)} d_j^t} \right) \bar{Q}_i.$$

Notice that these bounds converge to $\bar{Q}_i$, so they could be arbitrary close to $\bar{Q}_i$ by choosing a sufficiently large d^t vector.

To conclude the proof, we need to estimate the total number of purchases $\hat{d}_{tot}$ that guarantees that

$$\forall t > t^* : \widetilde{Q}_{i^*}(t) > \widetilde{Q}_{i^{**}(t)}$$

where t^* is the time period in which the total number of purchases becomes $\hat{d}_{tot} = \sum_{i \in [N]} d_i^{t^*}$. The value $\hat{d}_{tot}$ and its associated vector of purchases d^{t^*} must satisfy the following condition

$$\Delta \bar{Q} > \frac{\sum_{j=1}^{N} v_{\sigma(j)} \max_{1 \leq k \leq K} a_{j,k} (\bar{Q}_{i^*} + \bar{Q}_{i^{**}}) + \bar{Q}_{i^*} \max_{1 \leq k \leq K} z_k}{\sum_{j=1}^{N} v_{\sigma(j)} d_j^{t^*}}. \tag{A.13}$$

To verify inequality (A.13), it suffices to choose $\hat{d}_{tot}$ to satisfy

$$\hat{d}_{tot} > \frac{\sum_{j=1}^{N} v_{\sigma(j)} \max_{1 \leq k \leq K} a_{j,k} (\bar{Q}_{i^*} + \bar{Q}_{i^{**}}) + \bar{Q}_{i^*} \max_{1 \leq k \leq K} z_k}{\max_j v_{\sigma(j)} \Delta \bar{Q}}. \tag{A.14}$$

Since $\hat{d}_{tot}$ can be as large as desired, with the previous condition we guarantee the validity of Eq. (A.13).

We have just shown that the conditions of Lemma 1.1 are satisfied, we can now apply it using ranking policy σ to prove that the model goes to a monopoly for item i^*, which maximizes the product of its visibility and its weighted average quality, i.e., $v_{\sigma(i)} \bar{q}_i$. □

A.1.4 Proof of Corollary 1.1

From Theorem 1.2, a MMNL model goes to a monopoly for the item i that maximizes $v_{\sigma(i)}\bar{q}_i$. When the quality ranking is used, the product i^* that goes to a monopoly is

$$i^* = \arg\max_i v_{\sigma(i)}\bar{q}_i = \arg\max_i(\bar{q}_i).$$

$\square$

A.1.5 Proof of Theorem 1.3

To prove this result, we need the following property. Let $A \in \mathbb{R}_{\geq 0}^{N\times K}$, then

$$\sum_{k=1}^{K} \max_{1\leq i\leq N} a_{i,k} \leq K \max_{1\leq i\leq N} \sum_{k=1}^{K} a_{i,k} \tag{A.15}$$

where $a_{i,k} \in A$. Its proof follows from the following argument; $K \max_{1\leq i\leq N} \sum_{k=1}^{K} a_{ik} = K \max_{1\leq i\leq N} \sum_{k=1}^{K} |a_{ik}|$ $= K\|A\|_\infty = K \sup_{v\in\mathbb{R}^{K\times 1}:\|v\|_\infty=1} \|Av\|_\infty \geq K \left\|A\mathbb{I}^{K\times 1}\frac{1}{K}\right\|_\infty = \left\|A\mathbb{I}^{K\times 1}\right\|_\infty = \sum_{i=1}^{N}\sum_{k=1}^{K} |a_{ik}| \geq \sum_{k=1}^{K} \max_{1\leq i\leq N} |a_{ik}| = \sum_{k=1}^{K} \max_{1\leq i\leq N} a_{ik}$, where $\mathbb{I}^{K\times 1}$ is an all-one vector of dimension $K \times 1$.

At the limit, the probability that an item is purchased under the average quality ranking with the popularity signal is given by

$$P_{AQGSI} = \max_{1\leq i\leq N} \bar{q}_i. \tag{A.16}$$

When no popularity signal is shown, this probability becomes

$$P_{AQNSI} = \sum_{k=1}^{K} w_k \sum_{i=1}^{N} q_{i,k} \frac{v_{\sigma(i)}a_{i,k}}{\sum_{j=1}^{N} v_{\sigma(j)}a_{j,k} + z_k}. \tag{A.17}$$

We can easily bound P_{AQNSI} as follows:

$$0 \leq \sum_{k=1}^{K} \min_{1\leq i\leq N}(w_k q_{i,k}) \frac{\sum_{i=1}^{N} v_{\sigma(i)}a_{i,k}}{\sum_{j=1}^{N} v_{\sigma(j)}a_{j,k} + z_k} \leq P_{AQNSI} \leq \sum_{k=1}^{K} \max_{1\leq i\leq N}(w_k q_{i,k}) \tag{A.18}$$

and hence, by Inequality (A.15),

$$0 \leq \frac{P_{AQNSI}}{P_{AQGSI}} \leq K. \tag{A.19}$$

$\square$

A.1.6 Proof of Proposition 1.1

Consider first the upper bound. Choose a MMNL model where $z = 0$, $K = N$, the quality matrix is diagonal with a value of 1 for the first element and $1 - \epsilon$ for all others, the appeal matrix is the identity, and the classes have the same weights $w_i = \frac{1}{K}$. Then,

$$P_{AQNSI} = \sum_{1 \leq i \leq N} \frac{1}{K} \left(1 - \epsilon(1 - \delta_{i1})\right) \text{ and } P_{AQGSI} = \frac{1}{K},$$

where δ_{ij} is the Kronecker delta, thus

$$\lim_{\epsilon \to 0} \frac{P_{AQNSI}}{P_{AQGSI}} = \lim_{\epsilon \to 0} \sum_{1 \leq i \leq N} (1 - \epsilon \delta_{i1}) = \lim_{\epsilon \to 0} (K - \epsilon(K-1)) = K. \tag{A.20}$$

Consider now the lower bound. Choose a MMNL model where $z = 0$, $K = N$ with the same quality matrix as before, the same weights, and an appeal matrix filled with ones except in its diagonal where each element has a value of ϵ_A. Then,

$$P_{AQNSI} = \sum_{1 \leq i \leq N} \frac{1}{K} \left(1 - \epsilon(1 - \delta_{i1})\right) \frac{v_{\sigma(i)} \epsilon_A}{\epsilon_A v_{\sigma(i)} + \sum_{j \neq i} v_{\sigma(j)}} \text{ and } P_{AQGSI} = \frac{1}{K},$$

thus

$$\lim_{\epsilon_A \to 0} \frac{P_{AQNSI}}{P_{AQGSI}} = \lim_{\epsilon_A \to 0} \sum_{1 \leq i \leq N} \left(1 - \epsilon(1 - \delta_{i1})\right) \frac{v_{\sigma(i)} \epsilon_A}{\epsilon_A v_{\sigma(i)} + \sum_{j \neq i} v_{\sigma(j)}} = 0. \tag{A.21}$$

□

A.1.7 Proof of Theorem 1.4

By Theorem 1.2, we have

$$P_{AQGSI} \doteq \lim_{t \to \infty} P^t_{AQGSI} = \max\{\bar{q}_1, \bar{q}_2, \ldots, \bar{q}_N\}.$$

As mentioned earlier, for the segmented quality ranking, each segment is independent from each other and all of them will converge to a monopoly for the product with the highest quality in that class. We have that

$$P_{SQSSI} \doteq \lim_{t\to\infty} P^t_{SQSSI} = \sum_{k=1}^{K} w_k \max_i q_{i,k}.$$

As a result,

$$\frac{P_{SQSSI}}{P_{AQGSI}} = \frac{\sum_{k=1}^{K} w_k \max_{1\le i\le N} q_{i,k}}{\max_{1\le i\le N}\sum_{k=1}^{K} w_k q_{i,k}} = \frac{\sum_{k=1}^{K} \max_{1\le i\le N} w_k q_{i,k}}{\max_{1\le i\le N}\sum_{k=1}^{K} w_k q_{i,k}}.$$

The lower bound is obviously valid and the upper bound follows from inequality (A.15). □

A.1.8 Proof of Proposition 1.2

Consider a model with K items and K consumer classes. Without loss of generality, let the segment 1 be the segment with the lowest weight, i.e., $w_1 \le w_k \forall k \in [K]$. Then, for any set of positive appeals in each class, define the elements $q_{i,k}$ as follows:

$$q_{i,k} = \begin{cases} \frac{\min_{j\in[K]} w_j}{w_1} & \text{if } i = k = 1 \\ \frac{\min_{j\in[K]} w_j}{w_k} - \epsilon & \text{if } i = k \neq 1 \\ 0 & \text{otherwise} \end{cases}$$

where ϵ is a positive number ensuring that the model is tie-breaking for the quality rankings. Then

$$\lim_{\epsilon\to 0} \frac{P_{SQSSI}}{P_{AQGSI}} = \lim_{\epsilon_A\to 0} \frac{K \min_{j\in[K]} w_j - \epsilon \sum_{k=2}^{K} w_k}{\min_{j\in[K]} w_j} = K.$$

□

A.1.9 Proof of Theorem 1.5

Using the result from Van Hentenryck et al. (2016a), each segment under the SQSSI ranking will go to a monopoly of the product that maximizes its quality. If this product is not unique, the market will be shared between such products. Thus, the ratio between the asymptotic purchase probability of the average quality ranking without social influence and its segmented version with segmented social influence is given by

$$\frac{\lim_{t\to\infty} P^t_{AQNSI}}{\lim_{t\to\infty} P^t_{SQSSI}} = \frac{\sum_{k=1}^{K} w_k \sum_{i=1}^{N} q_{i,k} \frac{v_{\sigma(i)} a_{i,k}}{\sum_{j=1}^{N} v_{\sigma(j)} a_{j,k} + z_k}}{\sum_{k=1}^{K} w_k \max_{1\le i\le N} q_{i,k}}.$$

Since $\sum_{i=1}^{N} q_{i,k} \frac{v_{\sigma(i)} a_{i,k}}{\sum_{j=1}^{N} v_{\sigma(j)} a_{j,k} + z_k} \leq \max_{1 \leq i \leq N} q_{i,k}$, we have that $\frac{\lim_{t\to\infty} P^t_{AQNSI}}{\lim_{t\to\infty} P^t_{SQSSI}} \leq 1$.

A similar argument can be made for the second comparison:

$$\frac{\lim_{t\to\infty} P^t_{SQNSI}}{\lim_{t\to\infty} P^t_{SQSSI}} = \frac{\sum_{k=1}^{K} w_k \sum_{i=1}^{N} q_{i,k} \frac{v_{\sigma_k(i)} a_{i,k}}{\sum_{j=1}^{N} v_{\sigma_k(j)} a_{j,k} + z_k}}{\sum_{k=1}^{K} w_k \max_{1 \leq i \leq N} q_{i,k}} \leq 1.$$

□

A.2 The impact of misclassification

In this appendix we consider scenarios in which the firm is not always able to identify correctly the corresponding consumer segment. Misclassification may occur, for example, when there is no historical data about the incoming consumer and it has to be assigned to one of the classes solely based on basic information provided by some internet marketing companies (e.g. keywords searched, geographical region, sex, age, etc). We analyze what is the impact of having some classification errors under some mild model assumptions. First, we provide some theoretical results about the convergence under our misclassification model. Finally, using a numerical experiment, we illustrate what is the impact in performance as a function on the misclassification rate.

We begin describing the model extension. Suppose that every time the system has an arriving customer of segment $l \in [K]$ (this occurs with probability w_l), there exists a probability α_{lk} that the consumer is recognized as segment $k \in [K]$. If a consumer is recognized by the system (or classifier) as segment k consumer (even if it is not really from segment k), it is said that the consumer is *observed* in segment k. Similarly to the policy `SQSSI` let `SQSSIM` denote the ranking policy of classifying consumers (but now misclassification errors occur) and then providing to the consumers observed as segment k the quality ranking of consumer segment k and updating the popularity signal locally in each segment.

Since `SQSSIM` has misclassification, the benefits of segmentations might be deteriorated depending on how often segments are mistakenly recognized. Under `SQSSIM`, which product will become the most popular in each of the segments? To answer this question, we rely on the important result obtained in Theorem 1.2. Namely, it provides the long term convergence of using a static ranking with global social influence among all consumer segments. The following corollary identifies the product in each segment that will become the most popular in the long term.

Corollary A.1. *Suppose the solution to* $\arg\max_{1\leq i\leq N} v_{\sigma_k(i)} \sum_{l=1}^{K} w_l \alpha_{lk} q_{i,l}$ *is unique for segment* k *where* $\sigma_k(\cdot)$ *denotes a static ranking associated to segment* k*. Then, under the* `SQSSIM` *policy, the product* i_k^* *that converges to a monopoly for the observed segment* k *is:*

$$i_k^* = \arg\max_{1\leq i\leq N} v_{\sigma_k(i)} \sum_{l=1}^{K} w_l \alpha_{lk} q_{i,l}. \tag{A.22}$$

Proof. Consumers observed in segment k may belong to K different segments due to classifications mistakes, thus, each observed segment k can be seen as a MMNL. Given a consumer is observed in segment k, the probability that this customer belongs to $l \in [K]$ is given by $w_l \alpha_{lk}$. Then, the purchase probability of product i for a consumer observed in segment k, with ranking σ_k and purchase vector $d^{t,k}$ is given by

$$P_i^k(\sigma_k, d^{t,k}) = \sum_{l=1}^{K} \Big(w_l \alpha_{lk} \cdot \Big(v_{\sigma_k(i)} \frac{(a_{i,l} + d_i^{t,k})}{\sum_{j=1}^{N} v_{\sigma_k(j)}(a_{j,l} + d_j^{t,k}) + z_l} \cdot q_{i,l} \Big) \Big).$$

This probability resembles Eq. (A.8) in the proof of Theorem 1.2. Thus, we know that each observed segment k of consumers converges to a monopoly for product

$$i_k^* = \arg\max_{1\leq i\leq N} v_{\sigma_k(i)} \sum_{l=1}^{K} w_l \alpha_{lk} q_{i,l}.$$

□

Now that we have found the long term behavior of the model with classifying errors, we analyze the impact of errors in market efficiency. From now on we assume that the probability of committing a mistake in classifying a segment is the same for every segment, and is equally likely to identify it as any other segment. Then we may define the mistake probabilities with two parameters, α_0 and β_0:

$$\alpha_{lk} = \begin{cases} \alpha_0 & \text{if } l = k \\ \frac{\beta_0}{K-1} & \text{otherwise} \end{cases}. \tag{A.23}$$

Equation (A.23) means that every customer has a probability α_0 of being recognized as their correct segment, while a probability β_0 that the consumer's segment is mistaken for another one. Naturally, $\alpha_0 = 1 - \beta_0$.

Similarly to the definition of P^t_{SQSSI}, let P^t_{SQSSIM} denote the probability of a purchase at time t when the firm applies the segmented quality ranking with the local popularity signal d^t_k under a classifier with errors. We are now ready to prove the following:

Theorem A.1. *Assume that the average quality ranking and its segmented version are tie-breaking for a MMNL model with mistake probability matrix α given by (A.23), and that $\alpha_0 > \frac{\beta_0}{K-1}$ (the probability than an observed segment is classified correctly is greater than the probability of misclassifying it with any other segment). Then,*

$$\frac{K}{K-1}\beta_0 \leq \lim_{t\to\infty} \frac{P^t_{SQSSIM}}{P^t_{AQGSI}} \leq K\alpha_0 \tag{A.24}$$

Proof. As mentioned earlier, when we have segmentation, each segment is independent from each other and every recognized segment k will converge to a monopoly for the product i^*_k given by Eq. (A.22). If each observed segment k is ranked according to the new qualities $\hat{q}_{i,k} \doteq \sum_{l=1}^K w_l \alpha_{lk} q_{i,l}$, then

$$\begin{aligned}
\lim_{t\to\infty} P^t_{SQSSIM} &= \sum_{l=1}^K \sum_{k=1}^K w_l \alpha_{lk} q_{i^*_k,l} = \sum_{k=1}^K \max_{1\leq i\leq N} \sum_{l=1}^K w_l \alpha_{lk} q_{i,l} \\
&\leq \sum_{k=1}^K [\max_{1\leq l\leq K} \alpha_{lk}] \max_{1\leq i\leq N} \sum_{l=1}^K w_l q_{i,l} = \sum_{k=1}^K \max\{\alpha_0, \frac{\beta_0}{K-1}\} \max_{1\leq i\leq N} \bar{q}_i \\
&= K\alpha_0 \max_{1\leq i\leq N} \bar{q}_i,
\end{aligned}$$

and

$$\begin{aligned}
\lim_{t\to\infty} P^t_{SQSSIM} &= \sum_{l=1}^K \sum_{k=1}^K w_l \alpha_{lk} q_{i^*_k,l} = \sum_{k=1}^K \max_{1\leq i\leq N} \sum_{l=1}^K w_l \alpha_{lk} q_{i,l} \\
&\geq \sum_{k=1}^K [\min_{1\leq l\leq K} \alpha_{lk}] \max_{1\leq i\leq N} \sum_{l=1}^K w_l q_{i,l} = \sum_{k=1}^K \min\{\alpha_0, \frac{\beta_0}{K-1}\} \max_{1\leq i\leq N} \bar{q}_i \\
&= K\frac{\beta_0}{K-1} \max_{1\leq i\leq N} \bar{q}_i.
\end{aligned}$$

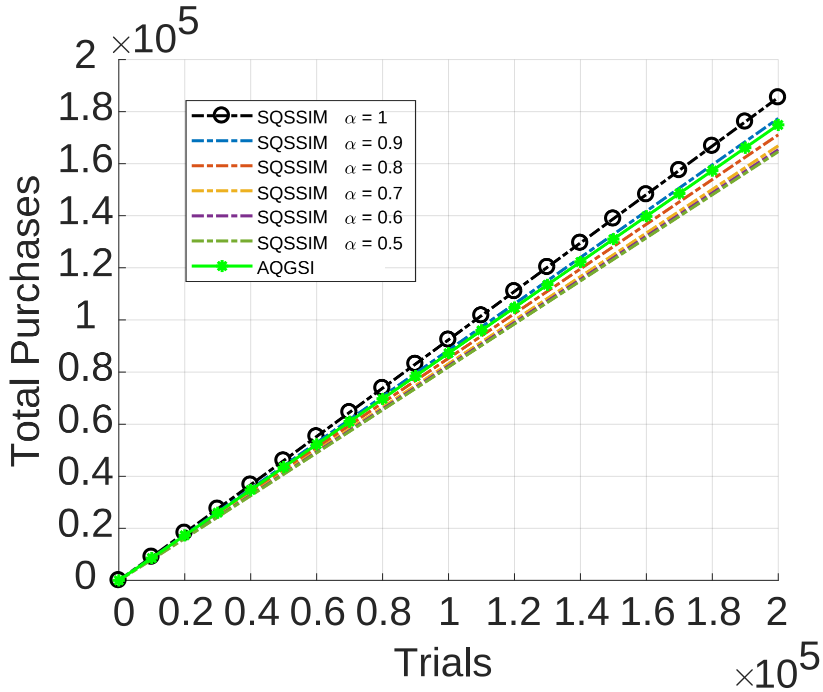

Figure A.1: The Number of Purchases over time for various SQSSIM rankings with different α_0 values, and the AQGSI ranking for Scheme 2. The x-axis represents the number of items tried and the y-axis represents the average number of purchases over all experiments. We can see that only when the classifications errors are small ($\alpha \geq 0.8$), the segmentation policy outperforms the global ranking policy (average quality ranking AQGSI).

Since $\lim_{t\to\infty} P^t_{AQGSI} = \max_{1\leq i\leq N} \bar{q}_i$, we finally have that

$$\frac{K}{K-1}\beta_0 \leq \lim_{t\to\infty} \frac{P^t_{SQSSIM}}{P^t_{AQGSI}} \leq K\alpha_0.$$

□

To illustrate the effects of classifications mistakes in the market, we perform a simulation for Scheme 2 under the ranking policy SQSSIM with different values of α_0. We also plot the ranking policy AQGSI in the same graph, the results are shown in Figure A.1. As expected, the performance of the ranking policy SQSSIM decreases as the percentage of correct consumer segment classification (α_0) decreases. Furthermore, it is interesting to see that the AQGSI ranking policy outperforms SQSSIM with an $\alpha = 0.8$ or less. The managerial insight is that segmentation in this setting is better than showing a single ranking, but only as long as the misclassification errors are relatively small. In cases where $\alpha = 0.8$ or lower, market segmentation is harmful.

A.3 2-Swap, a Performance Ranking Heuristic

In order to assess the average quality ranking policy (AQGSI), we performed computational experiments and compare it to a substantially more computationally expensive heuristic: the 2-swap heuristic. In all our experiments using the different schemes, the results obtained with the 2-swap heuristic were not significantly better than those obtained with the average quality ranking (AQGSI). Figure A.2 shows the average total purchases over 400 simulations as a function of the number of trials using the both methods. For completeness we also included the segmented ranking policy SQSSI.

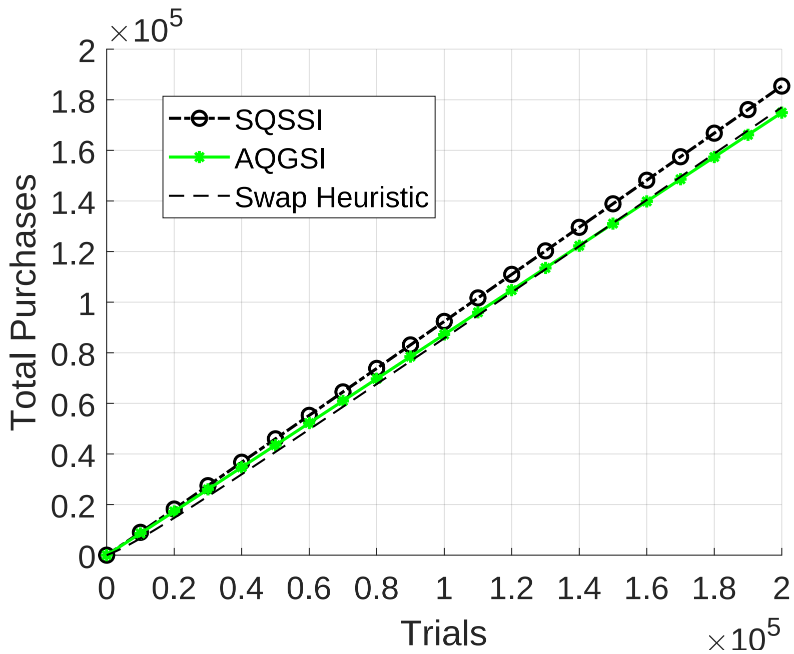

Figure A.2: Average total purchases vs trials of 400 simulations under SQSSI, SQSSI and the 2-swap heuristic with global social influence using Scheme 2.

A.4 Revenue Maximization

In this appendix we analyze an extension of the model in which each product $i \in [N]$ generates a profit r_i to the platform owner. Consumers are aware of the product prices and those are taken into account in the consumer choice probabilities of our model (Equation (1.1)).

Since now the firm is interested in maximizing profit, we need to modify the previous objective (Equation (1.2)) to the following one:

$$\Pi^{PR} = \max_{\sigma \in S_N} \left\{ \sum_{k=1}^{K} \left(w_k \cdot \sum_{i=1}^{N} \left(\frac{v_{\sigma(i)}(a_{i,k} + d_i^t)}{\sum_{j=1}^{N} v_{\sigma(j)}(a_{j,k} + d_j^t) + z_k} \cdot r_i q_{i,k} \right) \right) \right\}. \tag{A.25}$$

Observe that when $r_i = 1$ for all $i = 1, \dots, N$ the above problem reduces to the original setting in which the firm tries to maximize the expected number of purchases. This means that our NP-hardness result (Theorem 1.1) for the original model still holds for this new setting. Moreover, Theorem 1.2 (the monopoly convergence) 1.1 still hold, since consumer preferences under the Mixed MNL already account for the product's prices/revenues and the market dynamics is only dependent on those mixed MNL preferences (as previously). What is interesting to note is that in this model extension, the average quality ranking may lead to a monopoly for a product that doesn't yield the greatest revenue rate. This is because the product with the largest average quality, may not be the one that generates the largest expected revenue conditional on a trial. This lead us to analyze a new ranking heuristic, the "Average Expected Revenue Ranking", which ranks products based on their expected revenue conditional on a trial, namely $\sum_k w_k q_{i,k} \times r_i$. To obtain analytical results, we analyze this policy under an important assumption: the average quality ordering is the same as the revenue ordering. This assumption is reasonable in multiple settings: it says that higher quality products have a higher price. With this assumption, the average expected revenue ranking is the same as the average quality ranking, and thus, from Theorem 1.2, the market converges to monopoly for the product with the largest expected revenue conditional on a trial (asymptotic optimality). In addition, for the Average Expected Revenue Ranking, Theorem 1.3 still holds. The Average Expected Revenue Ranking with the same ordering between average product qualities and product revenues, can perform up to K (the number of consumer segments) times better, or arbitrarily worst, by not showing the popularity signal. The proof is straightforward from Theorem 1.3's proof, in which $q_{i,k}$ is exchanged by $q_{i,k} r_i$ each time.

For Theorem 1.4 to hold with the Average Expected Revenue Ranking, we need a stronger assumption. If the average quality ordering is just the same as the revenue ordering, it could happen that for some consumer segments product qualities and revenues are not ordered the same way. Thus, ranking by expected revenue conditional on a trial for each segment might not achieve asymptotic optimality in each consumer segment. In order to guarantee asymptotic optimality of each consumer segment we need that the quality ordering for each them is the same as the revenue ordering. Thus, the results from Theorem 1.4 hold only when all consumer segment product qualities follow the same ordering as the product revenue ordering.

A.5 Assortment Optimization

In this section we analyze the extension in which in each time period the platform owner chooses a subset $S \subseteq [N]$ of products to show to consumers, as well as a ranking σ_S among them. We particularly focus in the case where the platform owner has no information on consumers segments, and thus, needs to display the same assortment of products to all consumers. If there is only one consumer segment, this problem can be solved efficiently (see Abeliuk et al. (2016) and Sumida et al. (2019)). However, when there are at least two segments, the problem becomes a generalization of the classical assortment optimization under the latent class MNL model which is already NP-hard (see Bront et al. (2009) and Rusmevichientong et al. (2014)).

Given the impossibility of finding efficiently an optimal assortment, we study a heuristic which we called the *Average Quality Threshold Heuristic*. This heuristic first ranks products by their average quality, so let product i denote the product with the i^{th} highest average quality. Then, at each time period, it chooses how many products to show with the following condition: if it shows k products, those must be $\{1, 2 \ldots, k\}$ and they should be ranked in the same order: higher quality products are placed in positions with higher visibility. This heuristic is computationally more intensive than the standard average quality, since for each time period, we need to evaluate N different scenarios and choose the best one. Nevertheless, it is still computational practical since it takes at most $O(N^3)$ time.

Observe that the number of products shown by the Average Quality Threshold Heuristic is sensitive to the values associated to the outside option z_k's. In one extreme, very large values of the outside option (in comparison to the value associated to the products) means that the products offered by the platform owner face a rather weak cannibalization. In those scenarios, adding an extra product to an assortment is likely to be beneficial since it will probably increase overall sales. On the other extreme, when the outside option values are very small, products offered by the firm face a strong cannibalization: in these scenarios it is likely that offering only a few products is the optimal strategy.

We performed the same computational experiments as those in Section 1.6 on Scheme 1 and Scheme 2, but using the Average Quality Threshold Heuristic and varying the outside option value (z). A large outside option value means that at the early stages a higher number of consumers will decide not to try a product at all. In the long run, the effect of the outside option diminishes as some products become popular and their overall appeal increases. Figures A.3(a) and A.3(b) show the optimal number of products that are shown at each period (on average) for Schemes 1 and 2 respectively. As expected, observe that the optimal number of products shown increases when the outside option increases. However, as the number of purchases increase (and the social influence signal becomes stronger), the offer sets shown tend to reduce in size to exhibit only the products with the highest average quality.

In all our computational experiments, the offer sets offered by the average quality threshold heuristic are always reduced in size or they stay the same as times goes by. This suggests to us that even under the average quality threshold heuristic, which is a dynamic ranking policy, the market will also converge to a monopoly for the product with the highest weighted average quality. We leave this conjecture below.

Conjecture A.1. *(Asymptotic optimality of the Average Quality Threshold Heuristic) Whenever the average quality threshold heuristic is used, the market goes to a monopoly for the product with the highest weighted average quality.*

We conclude this section by assessing the performance of the Average Quality Threshold Heuristic with Global Social Influence (AQTGSI) and we compare against the performance of the Average Quality with Global Social Influence (AQGSI). Figures A.4(a) and A.4(b) display the average number of purchases under AQTGSI for Schemes 1 and 2 respectively. We can observe that more sales occur in settings where the outside option value is small, but this difference is reduced as time goes by. Figures A.5(a) and A.5(b) show the percentage improvement (or deterioration) that AQTGSI has over AQGSI for Schemes 1 and 2 respectively. Our experiments show that these two policies

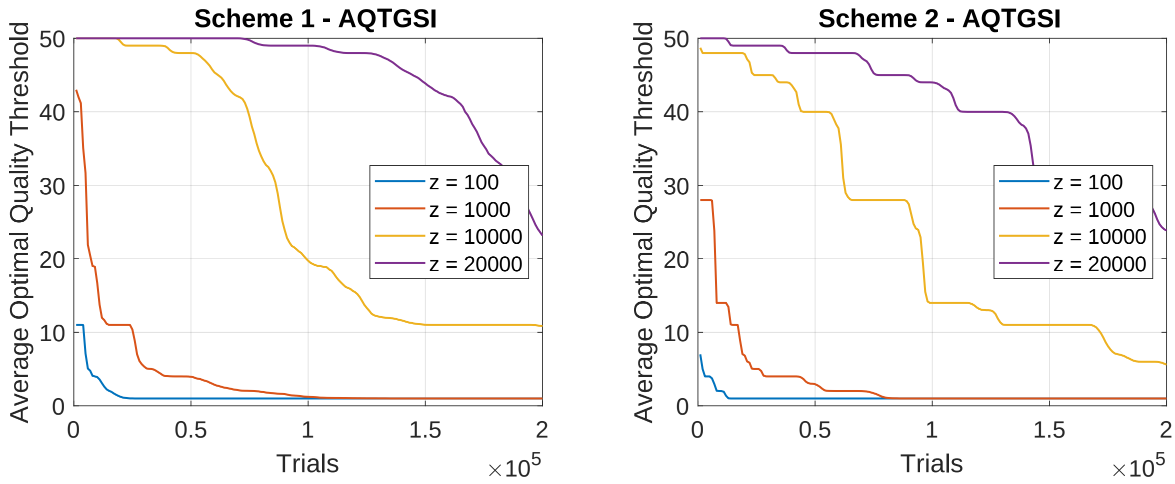

Figure A.3: Average threshold for the AQTGSI policy versus trials. On the left (a) we have it for Scheme 1, while on the right (b) we have it for Scheme 2.

have approximately the same performance for all the outside option values we tried, with less than 0.75% improvement in 200,000 trials.

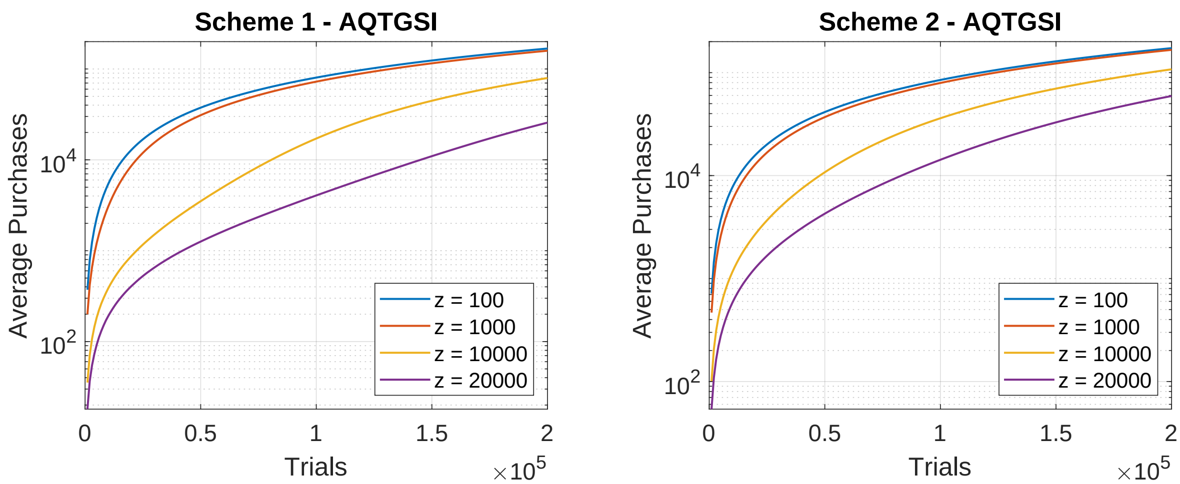

Figure A.4: Number of purchases versus trials using the AQTGSI policy for scheme 1 on the left (a) and Scheme 2 on the right (b).

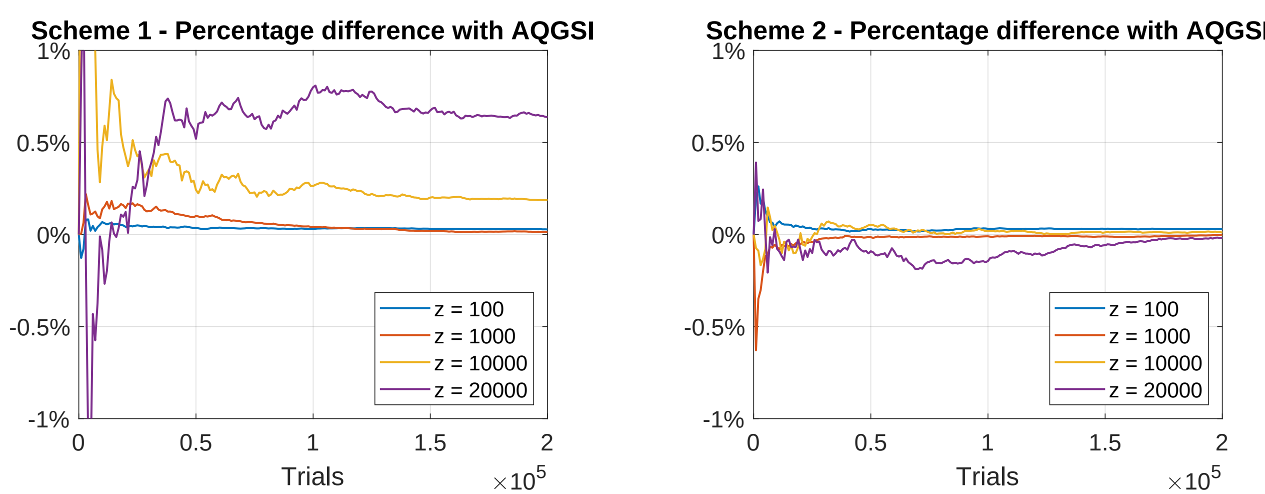

Figure A.5: Percentage difference in the number of purchases between AQTGSI and AQGSI for Scheme 1 on the left (a) and Scheme 2 on the right (b).

Appendix B

Additional Material for Chapter 2

B.1 Proofs

B.1.1 Proof of Proposition 2.1

First we prove that $a_{bW} = a^*_{bW}$ is the only advertising level for the low quality show that may survive the intuitive criterion. Let $\{a_{gW}, a_{bW}\}$ be a separating equilibrium satisfying the incentive compatibility constraints (2.15) and (2.16) with $a_{bW} \neq a^*_{bW}$. If $a_{bW} < a^*_{bW}$ then the equilibrium revenue for the low quality show would be $\pi^b(a_{bW}) = -a_{bW}$. Because $a^*_{bW} \leq 3$, and with such an advertising level it would obtain three views, the low quality show is better-off choosing strategy a^*_{bW}, achieving a profit of $3 - a^*_{bW} \geq 0$. If $a_{bW} > a^*_{bW}$, we have that $\pi^b(a_{bW}) = 3 - a_{bW} < \pi^b(a^*_{bW}) = 3 - a^*_{bW}$, so the low quality show is again better off deviating to a^*_{bW}. This is a contradiction; any separating equilibrium must be of the form $\{a_{gW}, a^*_{bW}\}$ in order to be able to survive the intuitive criterion.

Now we show that $a_{gW} = \max\{a^*_{gW}, a^*_{bW} - 2\}$ is the only advertising level for the high quality show that may survive the intuitive criterion. Let $\{a_{gW}, a^*_{bW}\}$ be a separating equilibrium satisfying the incentive compatibility constraints (2.15) and (2.16), which we can rewrite using constraint (2.17) as $\max\{a^*_{gW}, a_{bW} - 2\} \leq a_{gW} < a^*_{bW}$. Suppose $a_{gW} \in \Big(\max\{a^*_{gW}, a_{bW} - 2\}, a^*_{bW}\Big)$, then the equilibrium profit for the high quality show is $\pi^g(a_{gW}, W) = 3 - a_{gW}$. There always exists $\epsilon > 0$ such that $a_{gW} - \epsilon$ satisfies the incentive compatibility constraints, $a_{gW} - \epsilon \in \Big(\max\{a^*_{gW}, a_{bW} - 2\}, a^*_{bW}\Big)$ and achieves a greater profit for the high quality show. Since the low quality show would be worse off deviating to such an advertising level, by observing a deviation to $a_{gW} - \epsilon$, consumers' beliefs about show quality will be high. Thus, any equilibrium advertising level $a_{gW} \in \left(\max\{a^*_{gW}, a_{bW} - 2\}, a^*_{bW}\right)$ cannot survive the intuitive criterion.

Now we show that the separating equilibrium $\{\max\{a^*_{gW}, a^*_{bW} - 2\}, a^*_{bW}\}$ indeed survives the intuitive criterion. A high quality show would never deviate to $a > \max\{a^*_{gW}, a^*_{bW} - 2\}$ because, under its equilibrium strategy, it already obtain the maximum number of views, and at a lower advertising level than any advertising level in such set.

We now consider two cases: If $a^*_{gW} < a^*_{bW} - 2$ then $a_{gW} = a^*_{bW} - 2$. Any advertising level below a^*_{gW} is dominated for both show types, thus no show would deviate to such advertising levels. Any advertising level $a \in [a^*_{gW}, a^*_{bW} - 2)$ dominates both show types (i.e., there exist beliefs such that both shows would be better-off by deviating to such advertising levels). Thus $\mu(a) = \mu_0$ for all $a \in [a^*_{gW}, a^*_{bW} - 2)$, so in order to have no interest in deviations μ_0 must be low enough. We need that any deviation to $a_{gW} - \epsilon$ with any small $\epsilon > 0$ must satisfy $u_{\{1, a_{gW} - \epsilon\}} < 0$, which can be rewritten as $\mu_0 \theta_H + (1 - \mu_0)\theta_L - C_W + \gamma a_{gW} < 0 \iff \mu_0 < \frac{(C_W - \theta_L - \gamma a^*_{bW} + 2\gamma)}{\theta_H - \theta_L} = \frac{2\gamma}{\theta_H - \theta_L}$. Thus, in this way this equilibrium survives the intuitive criterion.

If $a^*_{gW} \geq a^*_{bW} - 2$ then $a_{gW} = a^*_{gW}$. All deviations for the high quality show are dominated by the equilibrium strategy, thus $\mu(a) = 0$ for any $a \neq a_{gW}$. The low quality show would always be better off choosing its first best solution a^*_{bW}; therefore this equilibrium survives the intuitive criterion. □

B.1.2 Proof of Lemma 2.1

Suppose there exists a pooling equilibrium on a_p such that $0 < a_p < a_{\mu_0}$. Then $u_{\{1, a_p\}} = \mu_0 \theta_H + (1 - \mu_0)\theta_L - C_W + \gamma a_p < \mu_0 \theta_H + (1 - \mu_0)\theta_L - C_W + \gamma a_{\mu_0} = \mu_0 \theta_H + (1 - \mu_0)\theta_L - C_W + \gamma \left[\mu_0 a^*_{gW} + (1 - \mu_0) a^*_{bW}\right] = 0$. Nobody would start watching the show, thus, both show types are always better-off by choosing no advertising level at all. Therefore, if both shows advertise and attract viewers we must have that $a_p \geq a_{\mu_0}$. □

B.1.3 Proof of Proposition 2.2

From Lemma 2.1 we have that $a_p \geq a_{\mu_0}$. Suppose we have a pooling equilibrium on $a_p > a_{\mu_0}$ that survives the intuitive criterion. For both show types, there exist out-of-equilibrium beliefs under which they are better of by deviating to a_{μ_0}, thus, consumer belief about an advertising level of

a_{μ_0} would be μ_0. Then both show types would be better-off by deviating to a_{μ_0}, so any pooling equilibrium on $a_p > a_{\mu_0}$ does not survive the intuitive criterion. This shows that a_{μ_0} is the unique equilibrium advertising level that could survive the intuitive criterion.

To show that $a_p = a_{\mu_0}$ can survive the intuitive criterion, we first show that for any deviation under which the high quality show could be better-off, the low quality show is also be better-off. Depending on out-of-equilibrium beliefs, the high quality show is only better-off deviating to any advertising level in $[a^*_{gW}, a_{\mu_0})$ (any advertising level below a^*_{gW}, regardless of out-of-equilibrium beliefs, would achieve zero views). Depending on out-of-equilibrium beliefs, low quality shows are also better off with such deviations. Thus, consumer belief for any deviation to $[a^*_{gW}, a_{\mu_0})$ would be set to μ_0 generating no views. Second, we must find the conditions under which the low-quality show is better off staying at its equilibrium a_{μ_0} and obtaining 1 view, over deviating to its first best solution a^*_{bW} and achieving 3 views. This happens as long as $1 - a_{\mu_0} > 3 - a^*_{bW}$, which may be reduced to $a^*_{gW} > \frac{2}{\mu_0} + a^*_{gW}$ and finally to $\mu_0 > \frac{2\gamma}{\theta_H - \theta_L}$. Now both show types don't have any incentives to deviate from a_{μ_0}, generating a pooling equilibrium. □

B.1.4 Proof of Lemma 2.2

The parameter sets are such that $\pi^t(a_{tW}, W) = 3 - a_{tW}$ and $\pi^t(a_{tB}, B) = 3 - a_{tB}$ where $a_{tX} = (C_X - \mathbb{E}[Q_t])/\gamma$ for $(t, X) \in \Big\{\{g, b\}, \{W, B\}\Big\}$.

- If $\frac{1}{\gamma}(C_W - C_B) < 3(1-\delta)$ then $a_{gW} - a_{gB} = a_{bW} - a_{bB} = \frac{1}{\gamma}(C_W - C_B) < 3(1-\delta)$. So $3 - a_{gW} = \pi^g(a_{gW}, W) > 3\delta - a_{gB} = \pi^g(a_{gB}, B)$ and $3 - a_{bW} = \pi^b(a_{bW}, W) > 3\delta - a_{bB} = \pi^b(a_{bB}, B)$. Both shows prefer linear releases to non-linear releases.

- If $\frac{1}{\gamma}(C_W - C_B) = 3(1-\delta)$ then $a_{gW} - a_{gB} = a_{bW} - a_{bB} = \frac{1}{\gamma}(C_W - C_B) = 3(1-\delta)$. So $3 - a_{gW} = \pi^g(a_{gW}, W) = 3\delta - a_{gB} = \pi^g(a_{gB}, B)$ and $3 - a_{bW} = \pi^b(a_{bW}, W) = 3\delta - a_{bB} = \pi^b(a_{bB}, B)$. Both shows are indifferent between linear releases and non-linear releases.

- If $\frac{1}{\gamma}(C_W - C_B) > 3(1-\delta)$ then $a_{gW} - a_{gB} = a_{bW} - a_{bB} = \frac{1}{\gamma}(C_W - C_B) > 3(1-\delta)$. So $3 - a_{gW} = \pi^g(a_{gW}, W) < 3\delta - a_{gB} = \pi^g(a_{gB}, B)$ and $3 - a_{bW} = \pi^b(a_{bW}, W) < 3\delta - a_{bB} = \pi^b(a_{bB}, B)$. Both shows prefer non-linear releases to linear releases. □

B.1.5 Proof of Lemma 2.3

Under B release timing, any advertising level below a_b^* would attract no viewers, whereas any advertising level above a_b^* will generate lower profits. The best they could do by having W as their release timing strategy is to choose an advertising level a_{gW}, but we already know from the first best solution argument that this strategy is dominated by (a_b^*, B).

B.1.6 Proof of Lemma 2.4

We prove it by contradiction. Suppose there exist a separating equilibrium in which high quality shows choose (a_g, W) with a positive advertising level $a_g < a_W$ and low quality shows choose their first best solution. Then we have that the expected utility for a high quality show choosing (a_g, W) is $u_{\{1,a_g,W\}} < u_{\{1,a_W,W\}} \leq 0$. Thus, the equilibrium profit of the high quality show is $-a_g$, so they are better-off deviating to a lower advertising level. Therefore, $a_g < a_W$ does not satisfy the equilibrium conditions.

B.1.7 Proof of Lemma 2.5

Suppose $a_g > \max\{1 - 3\delta + a_b, a_W\}$. Then any deviation to some strategy (a, W) with $a \in \Big(\max\{1 - 3\delta + a_b, a_W\}, a_g\Big)$ would be equilibrium-dominated for the "bad" show, regardless of what customers believe about the show's type. This is because this new advertising level a satisfies the incentive compatibility constraints. Therefore, customers should not believe that the show, which voluntarily made such a deviation, can be the low quality type with a positive probability. Consequently, the high quality show indeed prefers deviating to such an advertising level, as long as customers believe that such deviation cannot come from the low quality show. That is, the equilibrium involving advertising level $a > \max\{1 - 3\delta + a_b, a_W\}$ fails the intuitive criterion. So, the only value for a_g that might survive the intuitive criterion to deviations to W is $\max\{1 - 3\delta + a_b, a_W\}$. Let us show under which conditions $a_g = \max\{1 - 3\delta + a_b, a_W\}$ survives the intuitive criterion to deviations around linear releases (W).

If $1-3\delta+a_b > a_W$ then $a_g = 1-3\delta+a_b$, any deviation to (a, W) with $a \in [a_W, a_g)$ dominates both show types, thus the belief consumers would give to the show being good would be $\mu_{\{1,a,W\}} = \mu_0$. Then, in order for this equilibrium to survive the intuitive criterion, viewers need to refrain from watching. We impose $u_{\{1,a,W\}} < 0$ for all $a \in [a_W, a_g) \iff \mu_0 < \frac{\left(C_W - \theta_L - \frac{\alpha}{3} - \gamma a_g\right)}{\theta_H - \theta_L} = \frac{\gamma(3\delta - 1)}{\theta_H - \theta_L}$.

If $1 - 3\delta + a_b \leq a_W$ then $a_g = a_W$, then any deviation to any other (a, W) strategy would be equilibrium dominated for both show types. Any advertising level below a_W obtains no views, whereas the high quality show has no incentives in increasing their advertising level. Thus, any advertising level above a_W has null belief, so that the low quality show has no incentives in deviating from their first best solution.

B.1.8 Proof of Lemma 2.6

First note that $\bar{a}_g = 3\delta - \pi^g(a_g, W)$ is the advertising level under non-linear releases that would achieve the same revenue as the equilibrium strategy for a high quality show. Since the first best solution from the high quality show is (a_g^*, B) achieving a revenue of $3\delta - a_g^*$, from Inequality (2.24) we have that $3\delta - a_g^* > 3 - a_g = 3\delta - \bar{a}_g$, which implies that $\bar{a}_g > a_g^*$.

Also note that the incentive compatibility constraints ensure that $\bar{a}_g \leq a_b^*$. Thus, there exist consumer beliefs in which a high quality show choosing strategies (a, B) with $a \in [a_g^*, \bar{a}_g)$ would make greater profit than its equilibrium revenue $\pi^g(a_g, W) = 3 - a_g$ (this holds because under perfect information, the equilibrium strategy for the high quality show is (a_g^*, B)). If a low quality show chooses strategy (a, B) with $a < a_b^*$, viewers would watch a maximum of one episode (with the most optimistic belief) making at most $\delta - a$ of profit.

If we had that $\delta - \bar{a}_g < 3\delta - a_b$, then there exist off-equilibrium strategies (a, B) with $a \in [a_g^*, \bar{a}_g)$ that are dominated for the low quality show, while making the high quality show better off (for certain beliefs); thus, this equilibrium does not satisfy the intuitive criterion. If $\delta - \bar{a}_g \geq 3\delta - a_b$, there exist off-equilibrium beliefs that make the strategy (a, B) with $a \in [a_g^*, \bar{a}_g)$ dominate the low quality show's equilibrium; thus, the belief for such advertising levels would be μ_0. Therefore, in order to dissuade the shows to deviate, we must have that $u_{\{1,\bar{a}_g,B\}} < 0$, which translates to $\mu_0 < \gamma \frac{a_b^* - \bar{a}_g}{\theta_H - \theta_L}$.

B.1.9 Proof of Proposition 2.3

This Proposition follows immediately from using Case B $(a_g < a_{bW})$, Lemmas 2.5 and 2.6 and the incentive compatibility constraints (2.31).

B.1.10 Proof of Lemma 2.7

The expected profit in this situation is given by

$$\pi^b(a, W) = \sum_{i=1}^{3} i\text{P}(\text{watching exactly } i \text{ episodes}) - a. \tag{B.1}$$

A viewer would watch exactly 1 episode, if she watches episode 1 and does not watch episode 2. The only way for this to happen is if the quality draw from episode 1 is θ_L, then consumer beliefs would be updated to $\mu_{\{2,a,W\}} = 0$ (since the viewer now knows the show is *bad*), and if the expected utility from episode 2 is negative. Therefore, we have that

$$\text{P}(\text{watching exactly 1 episode}) = \mathbb{1}(u_{\{1,a,W\}} \geq 0)(1-b)\left[1 - \mathbb{1}\Big(\gamma a + \mathbb{E}[Q_b] \geq C_W\Big)\right] \tag{B.2}$$

A viewer would watch exactly 2 episodes, if she watches episode 1, then 2 and leaves. If the draw from episode 1 is θ_L, a viewer that watches episode 2, would also watch episode 3. In this case, the draw from episode 1 must be θ_M. In order for the viewer to stop watching after episode 2, we need her to draw a θ_L from this episode, and have a negative expected utility from episode 3. Then

$$\text{P}(\text{watching exactly 2 episodes}) = \mathbb{1}(u_{\{1,a,W\}} \geq 0)b(1-b)\left[1 - \mathbb{1}\Big(\gamma a + \mathbb{E}[Q_b] \geq C_W\Big)\right]. \tag{B.3}$$

A viewer would watch all 3 episodes if either of these scenarios holds: the draw from episode 1 is θ_L and the expected utility from episode 2 is non-negative, the draw from episode 1 is θ_M, the draw from episode 2 is θ_L and the expected utility from episode 3 is non-negative, or the draws from episodes 1 and 2 is θ_M. Then

$$\begin{aligned}\text{P}(\text{watching exactly 3 episodes}) =& \mathbb{1}(u_{\{1,a,W\}} \geq 0)\bigg[(1-b)\mathbb{1}\Big(\gamma a + \mathbb{E}[Q_b] \geq C_W\Big) \\ &+ b\Big[b + (1-b)\mathbb{1}\Big(\gamma a + \mathbb{E}[Q_b] \geq C_W\Big)\Big]\bigg].\end{aligned} \tag{B.4}$$

Finally, using Equations (B.2), (B.3) and (B.4) in Equation (B.1), and knowing that in a separating equilibrium $u_{\{1,a_g,W\}} \geq 0$, we obtain the result we are looking for.

B.1.11 Proof of Proposition 2.4

The proof of this Proposition is very similar to the ones of Lemmas 2.5 and 2.6. We want to show that a separating equilibrium $\{(a_g, W), (a_b, B)\}$ may survive the intuitive criterion when $a_g = \max\{a_W, 1 + b + b^2 + a_b - 3\delta\} < a_{bW}$ and $a_b = a_b^*$. The first needed are the conditions for a separating equilibrium, so we need the other incentive compatibility constraint $a_g \leq 3 + a_b - 3\delta$.

First, we analyze deviations to non-linear releases. The advertising levels that dominate the high quality show are in the region $[a_g^*, \bar{a}_g)$. If these advertising levels did not dominate the low quality show as well, this equilibrium would be ruled out by the intuitive criterion. Thus, we need to impose $\delta(1 + b + b^2) - \bar{a}_g > 3\delta - a_b^*$. This inequality is necessary for the existence of out-of-equilibrium beliefs that would make the low quality show better off (in expectation) by deviating to any strategy in non linear releases with advertising level in $[a_g^*, \bar{a}_g)$. Given this set dominates, for some out-of-equilibrium beliefs, the equilibrium strategies for both show types, consumers will believe that a deviation to such set is from the high quality show with probability μ_0. In order to ensure that viewers don't start watching the show, we set the expected utility of the first episode to be negative, $\mu_0 \mathbb{E}[Q_g] + (1 - \mu_0)\mathbb{E}[Q_b] - C_B + \gamma \bar{a}_g < 0 \iff \mu_0 < \gamma \frac{a_b^* - \bar{a}_g}{\mathbb{E}[Q_g] - \mathbb{E}[Q_b]}$.

Second, we analyze deviations to linear releases. If $a_W < 1 + a_b - 3\delta$, there exist out-of-equilibrium beliefs for which the advertising levels in the region $[a_W, 1 + a_b - 3\delta)$ dominate the equilibrium strategies for both show types. This is because this region does not satisfy the incentive compatibility constraints for the low quality show, whereas for the high quality show, it reduces the advertising level. Thus, when seeing such deviation, consumers would give a probability of μ_0 to the show being of high quality. In order to prevent both show types from performing such deviation, we must make the utility from watching the first episode negative, $\mu_0 \mathbb{E}[Q_g] + (1 - \mu_0)\mathbb{E}[Q_b] - C_W + \gamma(1 + a_b - 3\delta) < 0 \iff \mu_0 < \gamma \frac{3\delta - 1}{\mathbb{E}[Q_g] - \mathbb{E}[Q_b]}$. If $a_W \geq 1 + a_b - 3\delta$, there is no other advertising level on linear releases which could dominate (depending on out-of-equilibrium beliefs) the high quality show.

References

Makoto Abe. Price and advertising strategy of a national brand against its private-label clone: A signaling game approach. *Journal of Business Research*, 33(3):241–250, 1995.

Andrés Abeliuk, Gerardo Berbeglia, Manuel Cebrian, and Pascal Van Hentenryck. The benefits of social influence in optimized cultural markets. *Plos One*, 10:e0121934, 2015.

Andrés Abeliuk, Gerardo Berbeglia, Manuel Cebrian, and Pascal Van Hentenryck. Assortment optimization under a multinomial logit model with position bias and social influence. *4OR*, 14:57–75, 2016.

Daniel A. Ackerberg. Advertising, learning, and consumer choice in experience good markets: an empirical examination*. *International Economic Review*, 44(3):1007–1040, 2005.

Chris K Anderson and Xiaoqing Xie. Pricing and market segmentation using opaque selling mechanisms. *European Journal of Operational Research*, 233:263–272, 2014.

Eric T Anderson and Duncan I Simester. Price discrimination as an adverse signal: Why an offer to spread payments may hurt demand. *Marketing Science*, 20(3):315–327, 2001.

Ali Aouad and Danny Segev. Display optimization for vertically differentiated locations under multinomial logit choice preferences. *Available at SSRN: https://ssrn.com/abstract=2709652 or http://dx.doi.org/10.2139/ssrn.2709652*, 2015.

Terrence August, Duy Dao, and Hyoduk Shin. Optimal timing of sequential distribution: The impact of congestion externalities and day-and-date strategies. *Marketing Science*, 34(5):755–774, 2015.

Kyle Bagwell. The economic analysis of advertising. *Handbook of industrial organization*, 3:1701–1844, 2007.

Kyle Bagwell and Michael H Riordan. High and declining prices signal product quality. *The American Economic Review*, pages 224–239, 1991.

Subramanian Balachander. Warranty signalling and reputation. *Management Science*, 47(9):1282–1289, 2001.

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

Gary S. Becker and Kevin M. Murphy. A simple theory of advertising as a good or bad. *The Quarterly Journal of Economics*, 108(4):941–964, 1993.

Leonard L Berry. Relationship marketing of services—growing interest, emerging perspectives. *Journal of the Academy of marketing science*, 23(4):236–245, 1995.

Steven Berry, James Levinsohn, and Ariel Pakes. Automobile prices in market equilibrium. *Econometrica: Journal of the Econometric Society*, pages 841–890, 1995.

E. Boggiano, E. Burgess, B. Turan, T. Soleymani, S. Daniel, L. Vinson, K. Lokken, B. Wingo, and A. Morse. Motives for eating tasty foods associated with binge-eating. results from a student and a weight-loss seeking population. *Appetite*, 83:160–166, December 2014.

Timothy F Bresnahan. Departures from marginal-cost pricing in the american automobile industry: Estimates for 1977–1978. *Journal of Econometrics*, 17(2):201–227, 1981.

Juan José Miranda Bront, Isabel Méndez-Díaz, and Gustavo Vulcano. A column generation algorithm for choice-based network revenue management. *Operations research*, 57(3):769–784, 2009.

Gerard R. Butters. Equilibrium distributions of sales and advertising prices. *The Review of Economic Studies*, 44(3):465–491, 1977.

Joan Calzada and Tommaso M Valletti. Intertemporal movie distribution: Versioning when customers can buy both versions. *Marketing Science*, 31(4):649–667, 2012.

A Colin Cameron and Pravin K Trivedi. *Microeconometrics: methods and applications*. Cambridge university press, 2005.

Nicola Capuano, Francisco Chiclana, Hamido Fujita, Enrique Herrera-Viedma, and Vincenzo Loia. Fuzzy group decision making with incomplete information guided by social influence. *IEEE Transactions on Fuzzy Systems*, 26:1704–1718, 2017.

Ningyuan Chen, Anran Li, and Kalyan Talluri. Reviews and self-selection bias with operational implications. *Available at SSRN: https://ssrn.com/abstract=3407477 or http://dx.doi.org/10.2139/ssrn.3407477*, 2019.

Lesley Chiou. The timing of movie releases: Evidence from the home video industry. *International Journal of Industrial Organization*, 26(5):1059–1073, 2008.

In-Koo Cho and David M Kreps. Signaling games and stable equilibria. *The Quarterly Journal of Economics*, 102(2):179–221, 1987.

Kevin YC Chung, Timothy P. Derdenger, and Kannan Srinivasan. Economic value of celebrity endorsements: Tiger woods' impact on sales of nike golf balls. *Marketing Science*, 32(2):271–293, March 2013. ISSN 1526-548X.

Nick Craswell, Onno Zoeter, Michael Taylor, and Bill Ramsey. An experimental comparison of click position-bias models. In *Proceedings of the 2008 International Conference on Web Search and Data Mining*, pages 87–94. ACM, 2008.

Sanjukta Das. Timing movie release on the internet in the context of piracy. *Journal of Organizational Computing and Electronic Commerce*, 18(4):307–332, 2008.

James Davis, Guillermo Gallego, and Huseyin Topaloglu. Assortment planning under the multinomial logit model with totally unimodular constraint structures. *Technical report*, 2013.

Timothy Derdenger. Technological tying and the intensity of price competition: An empirical analysis of the video game industry. *Quantitative Marketing and Economics*, 12(2):127–165, 2014.

Timothy P. Derdenger, Hui Li, and Kannan Srinivasan. Firms' strategic leverage of unplanned exposure and planned advertising: An analysis in the context of celebrity endorsements. *Journal of Marketing Research*, 55(1):14–34, 2018.

Preyas S Desai. Multiple messages to retain retailers: Signaling new product demand. *Marketing Science*, 19(4):381–389, 2000.

Peter R Dickson and James L Ginter. Market segmentation, product differentiation, and marketing strategy. *Journal of Marketing*, 51:1–10, 1987.

Jean-Pierre Dubé, Günter J Hitsch, and Puneet Manchanda. An empirical model of advertising dynamics. *Quantitative marketing and economics*, 3(2):107–144, 2005.

Anita Elberse. The power of stars: Do star actors drive the success of movies? *Journal of marketing*, 71(4): 102–120, 2007.

Anita Elberse and Bharat Anand. The effectiveness of pre-release advertising for motion pictures: An empirical investigation using a simulated market. *Information economics and policy*, 19(3-4):319–343, 2007.

Jehoshua Eliashberg and Steven M Shugan. Film critics: Influencers or predictors? *Journal of marketing*, 61(2):68–78, 1997.

Jehoshua Eliashberg, Jedid-Jah Jonker, Mohanbir S Sawhney, and Berend Wierenga. Moviemod: An implementable decision-support system for prerelease market evaluation of motion pictures. *Marketing Science*, 19(3):226–243, 2000.

Jehoshua Eliashberg, Anita Elberse, and Mark AAM Leenders. The motion picture industry: Critical issues in practice, current research, and new research directions. *Marketing science*, 25(6):638–661, 2006.

Jehoshua Eliashberg, Sam K Hui, and Z John Zhang. Assessing box office performance using movie scripts: A kernel-based approach. *IEEE Transactions on Knowledge and Data Engineering*, 26(11):2639–2648, 2014.

Per Engström and Eskil Forsell. Demand effects of consumers' stated and revealed preferences. *Journal of Economic Behavior & Organization*, 150:43–61, 2018.

Tulin Erdem and Michael P. Keane. Decision-making under uncertainty: Capturing dynamic brand choice processes in turbulent consumer goods markets. *Marketing Science*, 15(1):1–20, 1996.

Tülin Erdem, Michael P Keane, and Baohong Sun. A dynamic model of brand choice when price and advertising signal product quality. *Marketing Science*, 27(6):1111–1125, 2008.

Björn Frank. Optimal timing of movie releases in ancillary markets: The case of video releases. *Journal of Cultural Economics*, 18(2):125–133, 1994.

Esther Gal-Or. Warranties as a signal of quality. *Canadian Journal of Economics*, pages 50–61, 1989.

Jeff Galak, Justin Kruger, and George Loewenstein. Slow down! insensitivity to rate of consumption leads to avoidable satiation. *Journal of Consumer Research*, 39(5):993–1009, 2013.

Guillermo Gallego, Anran Li, Van-Anh Truong, and Xinshang Wang. Approximation algorithms for product framing and pricing. *Operations Research*, 68:1–307, 2020.

Yigal Gerchak, Richard K Cho, and Saibal Ray. Coordination of quantity and shelf-retention timing in the video movie rental industry. *IIE Transactions*, 38(7):525–536, 2006.

Anindya Ghose, Panagiotis G Ipeirotis, and Beibei Li. Designing ranking systems for hotels on travel search engines by mining user-generated and crowdsourced content. *Marketing Science*, 31:493–520, 2012.

Anindya Ghose, Panagiotis G Ipeirotis, and Beibei Li. Examining the impact of ranking on consumer behavior and search engine revenue. *Management Science*, 60:1632–1654, 2014.

Robert Gibbons. *Game theory for applied economists*. Princeton University Press, 1992.

Negin Golrezaei, Vahideh Manshadi, and Vahab Mirrokni. Two-stage pandora's box for product ranking. *Available at SSRN: https://ssrn.com/abstract=3130378 or http://dx.doi.org/10.2139/ssrn.3130378*, 2018.

Ram Gopal, Hooman Hidaji, Raymond A Patterson, Erik Rolland, and Dmitry Zhdanov. Design improvements for message propagation in malleable social networks. *Production and Operations Management*, 25:993–1005, 2016.

Gautam Gowrisankaran and Marc Rysman. Dynamics of consumer demand for new durable goods. *Journal of political Economy*, 120(6):1173–1219, 2012.

Gene M. Grossman and Carl Shapiro. Informative advertising with differentiated products. *The Review of Economic Studies*, 51(1):63–81, 1984.

Ashton Hauff. The know it all guide to color psychology in marketing + the best hex chart. https://coschedule.com/blog/color-psychology-marketing/, 2018.

Thorsten Hennig-Thurau, Victor Henning, Henrik Sattler, Felix Eggers, and Mark B Houston. The last picture show? timing and order of movie distribution channels. *Journal of Marketing*, 71(4):63–83, 2007.

Ming Hu, Joseph Milner, and Jiahua Wu. Liking and following and the newsvendor: Operations and marketing policies under social influence. *Management Science*, 62:867–879, 2015.

Springboard India. How netflix's recommendation engine works? *Medium*, 2019.

Marina Ivasic-Kos, Miran Pobar, and Luka Mikec. Movie posters classification into genres based on low-level features. In *2014 37th international convention on information and communication technology, electronics and microelectronics (MIPRO)*, pages 1198–1203. IEEE, 2014.

Svante Janson. Functional limit theorems for multitype branching processes and generalized Pólya urns. *Stochastic Processes and their Applications*, 110:177–245, 2004.

Thomas D Jeitschko and Hans-Theo Normann. Signaling in deterministic and stochastic settings. *Journal of Economic Behavior & Organization*, 82(1):39–55, 2012.

Baojun Jiang, Kinshuk Jerath, and Kannan Srinivasan. Firm strategies in the mid tail of platform-based retailing. *Marketing Science*, 30(5):757–775, 2011.

Richard E Kihlstrom and Michael H Riordan. Advertising as a signal. *journal of Political Economy*, 92(3): 427–450, 1984.

Borys Kit. 'a quiet place part ii,' 'mission: Impossible 7' to debut on paramount+ 45 days after theatrical launch. https://https://www.hollywoodreporter.com/news/paranormal-activity-pet-sematary-movies-in-the-works-for-paramount, 2021.

Robert E Krider and Charles B Weinberg. Competitive dynamics and the introduction of new products: The motion picture timing game. *Journal of Marketing Research*, pages 1–15, 1998.

Coco Krumme, Manuel Cebrian, Galen Pickard, and Sandy Pentland. Quantifying social influence in an online cultural market. *Plos One*, 7:e33785, 2012.

Martin A Lariviere and V Padmanabhan. Slotting allowances and new product introductions. *Marketing Science*, 16(2):112–128, 1997.

Haengju Lee and Yongsoon Eun. Discovering heterogeneous consumer groups from sales transaction data. *European Journal of Operational Research*, 280:338–350, 2020.

Donald R Lehmann and Charles B Weinberg. Sales through sequential distribution channels: An application to movies and videos. *Journal of Marketing*, 64(3):18–33, 2000.

Andrew Lim, Brian Rodrigues, and Xingwen Zhang. Metaheuristics with local search techniques for retail shelf-space optimization. *Management Science*, 50:117–131, 2004.

Laurent Linnemer. Price and advertising as signals of quality when some consumers are informed. *International Journal of Industrial Organization*, 20(7):931–947, 2002.

Angus Loten. Retailers use ai to improve online recommendations for Shoppers. *Wall Street Journal*, 2020.

Tong Lu, Eric T. Bradlow, and J. Wesley Hutchinson. Binge consumption of online content. In *Working Paper*, 2017.

Jackie Y Luan and K Sudhir. Optimal inter-release time between sequentially released products. Technical report, Working Paper. Yale University, New Haven, CT, 2006.

Christoph Lutz and Gemma Newlands. Consumer segmentation within the sharing economy: The case of Airbnb. *Journal of Business Research*, 88:187–196, 2018.

Nancy A Lutz. Warranties as signals under consumer moral hazard. *The Rand journal of economics*, pages 239–255, 1989.

Thierry Magnac and David Thesmar. Identifying dynamic discrete decision processes. *Econometrica*, 70(2): 801–816, 2002.

Felipe Maldonado, Pascal Van Hentenryck, Gerardo Berbeglia, and Franco Berbeglia. Popularity signals in trial-offer markets with social influence and position bias. *European Journal of Operational Research*, 266:775–793, 2018.

Dana Mattioli. On Orbitz, Mac users steered to pricier hotels. *The Wall Street Journal*, 2012.

Daniel McFadden and Kenneth Train. Mixed MNL models for discrete response. *Journal of Applied Econometrics*, 15:447–470, 2000.

Paul Milgrom and John Roberts. Price and advertising signals of product quality. *Journal of political economy*, 94(4):796–821, 1986.

Xavier Molinero, Fabián Riquelme, and Maria Serna. Cooperation through social influence. *European Journal of Operational Research*, 242:960–974, 2015.

K Sridhar Moorthy and Ivan PL Png. Market segmentation, cannibalization, and the timing of product introductions. *Management science*, 38(3):345–359, 1992.

Sridhar Moorthy. Can brand extension signal product quality? *Marketing science*, 31(5):756–770, 2012.

Sridhar Moorthy and Kannan Srinivasan. Signaling quality with a money-back guarantee: The role of transaction costs. *Marketing Science*, 14(4):442–466, 1995.

José Luis Moraga-González. Quality uncertainty and informative advertising. *International Journal of Industrial Organization*, 18(4):615–640, 2000.

Jill Morton. Color & branding. https://www.colormatters.com/color-and-marketing/color-and-branding, 2012.

Anirban Mukherjee and Vrinda Kadiyali. The competitive dynamics of new dvd releases. *Management Science*, 64(8):3536–3553, 2018.

Jon Nathanson. The economics of a hit tv show. `https://priceonomics.com/the-economics-of-a-hit-tv-show`, 2013. Accessed: 2013-10-17.

Phillip Nelson. Advertising as information. *Journal of political economy*, 82(4):729–754, 1974.

Aviv Nevo. A practitioner's guide to estimation of random-coefficients logit models of demand. *Journal of economics & management strategy*, 9(4):513–548, 2000.

Natalie O'Grady. Why and how to apply color theory in social media marketing. https://www.prdaily.com/why-and-how-to-apply-color-theory-in-social-media-marketing/, 2019.

Grant Packard, Anocha Aribarg, Jehoshua Eliashberg, and Natasha Z Foutz. The role of network embeddedness in film success. *International Journal of Research in Marketing*, 33(2):328–342, 2016.

Ashutosh Prasad, Bart Bronnenberg, and Vijay Mahajan. Product entry timing in dual distribution channels: The case of the movie industry. *Review of Marketing Science*, 2(1), 2004.

Vithala R Rao, S Abraham Avri Ravid, Richard T Gretz, Jialie Chen, and Suman Basuroy. The impact of advertising content on movie revenues. *Marketing Letters*, 28(3):341–355, 2017.

Paat Rusmevichientong, David Shmoys, Chaoxu Tong, and Huseyin Topaloglu. Assortment optimization under the multinomial logit model with random choice parameters. *Production and Operations Management*, 23(11):2023–2039, 2014.

Matthew J Salganik, Peter Sheridan Dodds, and Duncan J Watts. Experimental study of inequality and unpredictability in an artificial cultural market. *Science*, 311:854–856, 2006.

Eric Schwartzel and Joe Flint. Warner bros. to release all 2021 films on hbo max and in theaters simultaneously. https://www.wsj.com/articles/warner-bros-to-release-all-2021-films-on-hbo-max-in-theaters-simultaneously-11607020941, 2020.

David A Schweidel and Wendy W Moe. Binge watching and advertising. *Journal of Marketing*, 80(5):1–19, 2016.

Duncan Simester. Signalling price image using advertised prices. *Marketing Science*, 14(2):166–188, 1995.

George J. Stigler. The economics of information. *Journal of Political Economy*, 69(3):213–225, 1961.

George J. Stigler and Gary S. Becker. De gustibus non est disputandum. *The American Economic Review*, 67(2):76–90, 1977.

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

Mika Sumida, Guillermo Gallego, Paat Rusmevichientong, Huseyin Topaloglu, and James Davis. Revenue-utility tradeoff in assortment optimization under the multinomial logit model with totally unimodular constraints. Technical report, Tech. rep., Cornell University, Ithaca, NY, 2019.

Jessica Toonkel and Lisa Richwine. How network tv figured out binge-watching. http://fortune.com/2016/03/11/netflix-changing-game-network-tv//, 2016. Accessed: 2018-12-05.

William Trouleau, Azin Ashkan, Weicong Ding, and Brian Eriksson. Just one more: Modeling binge watching behavior. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 1215–1224. ACM, 2016.

Catherine Tucker and Juanjuan Zhang. How does popularity information affect choices? A field experiment. *Management Science*, 57:828–842, 2011.

Stefano Vaccari, Costis Maglaras, and Marco Scarsini. Social learning from online reviews with product choice. *Available at SSRN*, 2018.

Pascal Van Hentenryck, Andrés Abeliuk, Franco Berbeglia, Felipe Maldonado, and Gerardo Berbeglia. Aligning popularity and quality in online cultural markets. In *Tenth International AAAI Conference on Web and Social Media (ICWSM-16)*, pages 398–407, 2016a.

Pascal Van Hentenryck, Alvaro Flores, and Gerardo Berbeglia. Trial-offer markets with continuation. *arXiv preprint arXiv:1610.01929*, 2016b.

Giampaolo Viglia, Roberto Furlan, and Antonio Ladrón-de Guevara. Please, talk about it! when hotel popularity boosts preferences. *International Journal of Hospitality Management*, 42:155–164, 2014.

Harold L Vogel. *Entertainment industry economics: A guide for financial analysis.* Cambridge University Press, 2014.

R.T. Watson. Amc, universal agree to trim theatrical window before movies go online. https://www.wsj.com/articles/amc-universal-agree-to-trim-theatrical-window-before-movies-go-online-11595968517, 2020.

Martin L Weitzman. Optimal search for the best alternative. *Econometrica: Journal of the Econometric Society*, pages 641–654, 1979.

Birger Wernerfelt. Umbrella branding as a signal of new product quality: An example of signalling by posting a bond. *The RAND Journal of Economics*, pages 458–466, 1988.

Weichen Yan. Strategic advertising and release in the movie industry. *Working paper*, 2020.

Hao Zhao. Raising awareness and signaling quality to uninformed consumers: A price-advertising model. *Marketing Science*, 19(4):390–396, 2000.

Xueping Zhen, Gangshu George Cai, Reo Song, and Sungha Jang. The effects of herding and word of mouth in a two-period advertising signaling model. *European Journal of Operational Research*, 275:361–373, 2019.