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

Franco Berbeglia*, Timothy Derdenger†, Kannan Srinivasan‡ & Joseph Xu§

September 2020

Abstract

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 all-at-once through a newly launched online advertisement based video on demand (AVOD) service (e.g. NBCUniversal's Peacock or ViacomCBS's Pluto TV), which lowers a consumers' viewing costs relative to the traditional TV channel 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. Particularly, we analyze whether the release timing of episodes signals show quality and moderates advertising levels. We show that by adding an AVOD channel that simultaneously (all-at-once) releases show episodes, there exists a separating equilibrium under which high and low quality shows choose different release timing (channel) strategies.

Key words: release timing of episodes, signaling quality, separating equilibrium, advertising as a signal, game theory, binge-watching, broadcast networks

*Franco Berbeglia is PhD student at Tepper School of Business, CMU. e-mail: fberbegl@andrew.cmu.edu.

†Tim Derdenger is Associate Professor in Marketing & Strategy, Tepper School of Business, CMU.

‡Kannan Srinivasan is H.J. Heinz II Professor of Management, Marketing and Business Technologies, Tepper School of Business, CMU.

§Joseph Xu is Assistant Professor of Operations Management, Tepper School of Business, CMU.

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 and NBCUniversal, 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 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 15 million monthly active users.[2] NBCUniversal is also entering the subscription free AVOD space with the launch of its streaming service Peacock.[3] But once such a service is created like Pluto TV or Peacock, how does a content provider determine what new and original shows should be released through the traditional broadcast television channel versus the new digital distribution AVOD service? Why did NBCUniversal decide to exclusively release reboots of "Saved By the Bell", "Battlestar Galactica" and "Punky Brewster" on Peacock and not through one of its NBCUniversal television networks?[4]

Currently, there is little understanding as to the strategic implications of the all-at-

[1] binge watching, the act of watching several episodes in one sitting
[2] https://www.viacbs.com/news/audience-insights/pluto-tv-and-the-avod-opportunity
[3] https://www.cnbc.com/2019/11/01/nbc-peacock-may-be-free-for-all-users.html
[4] https://tcrn.ch/2ukKx6M

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 paper studies the profitability impact of simultaneous releases for shows within a subscription free AVOD service in conjunction with a sequential release through a traditional television network. Our paper 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 ad revenue from over the top (OTT) TV is projected to hit 5 billion in 2020.[5]

With new advances in technology, especially the ability to watch on demand, binge consumption 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, which occur in different platforms (TV vs. AVOD) and 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

[5]https://www.adweek.com/tv-video/remember-free-tv-its-coming-back-in-a-big-way/

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 paper: 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, partially due to the advertising revenue difference between the two release timing options we are considering, but also due to the scarce resource of prime-time television hours. 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.

Our paper illustrates that adequate levels of advertising alone may signal quality, and for this to happen, the higher-quality show must incur a sizeable cost. We then show that by adding the simultaneous release timing AVOD channel, in equilibrium lower quality shows select this channel since it is more profitable than the sequential network channel. 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. Even though the higher quality show does not apply a simultaneous release strategy, by providing a more profitable channel 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 channel to release new content through 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 channels 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.

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 advertisments 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 paper, 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 paper 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 paper 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 paper 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 Model

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

Consider a traditional television based media company, such as NBCUniversal, that creates shows but also distributes its content through two different distribution channels, the traditional television channel and an advertising-based video on demand service. 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 that each map to a specific distribution channel.

These two options are: i) simultaneous releases that occur through the company's AVOD service, which we will denote by B , and ii) linear sequential releases that occurs through a traditional television channel, 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 sequential linear 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.[6] 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:

$$\pi^t(a, W) = \mathbb{E}[\# \text{ of views}|a, W \text{ and } t] - a \tag{1}$$

$$\pi^t(a, B) = \delta\mathbb{E}[\# \text{ of views}|a, B \text{ and } t] - a. \tag{2}$$

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 television show and channel, 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

[6] 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..."

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 advertisers [Nathanson, 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 all-at-once releases provide more flexibility to viewers than weekly releases. [7] 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 $i \geq 0$

[7] DVR lowers C_W but still cannot lower it below C_B. This is because viewers would still have to record and wait for the whole season to be released to be able to binge-watch it.

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

$$\mu_{\{i+1,a,X\}} = \frac{\mu_{\{1,a,X\}}(1-g)^{i-1}}{\mu_{\{1,a,X\}}(1-g)^{i-1} + (1-\mu_{\{1,a,X\}})b^{i-1}}, \tag{3}$$

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

$$u_{\{i,a,X\}} = \mu_{\{i,a,X\}}\mathbb{E}[Q_g] + (1 - \mu_{\{i,a,X\}})\mathbb{E}[Q_b] - C_X + \gamma a, \tag{4}$$

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

$$Q_g = \begin{cases} \theta_H & \text{w.p. } g \\ \theta_M & \text{w.p. } (1-g), \end{cases} \tag{5}$$

$$Q_b = \begin{cases} \theta_M & \text{w.p. } b \\ \theta_L & \text{w.p. } (1-b). \end{cases} \tag{6}$$

Then, $\mathbb{E}[Q_g] = g\theta_H + (1-g)\theta_M$ and $\mathbb{E}[Q_b] = b\theta_M + (1-b)\theta_L$.[8]

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 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.

3 Signaling Through Advertising Alone

To identify the impact of the simultaneous release timing, we first study the traditional TV setting in which shows are released sequentially. 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

[8] 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.

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. [9]

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

3.2 Complete Information With Multiple Episodes

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

$$\pi^g(a_g, W) = 3\mathbb{1}(\theta_H - C_W + \gamma a_g \geq 0) - a_g, \tag{7}$$

[9] 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.

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

$$\pi^b(a_b, W) = 3\mathbb{1}(\theta_L - C_W + \gamma a_b \geq 0) - a_b. \tag{8}$$

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

$$a^*_{gW} = \begin{cases} [(C_W - \theta_H)/\gamma]^+ & \text{if } \pi^g\left([(C_W - \theta_H)/\gamma]^+, W\right) \geq 0 \\ 0 & \text{otherwise} \end{cases}, \tag{9}$$

and

$$a^*_{bW} = \begin{cases} [(C_W - \theta_L)/\gamma]^+ & \text{if } \pi^b\left([(C_W - \theta_L)/\gamma]^+, W\right) \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

respectively.

If we consider the situation in which both firms need to advertise to attract viewers while making non-negative profit, we have that $3 \geq a^*_{gW} = (C_W - \theta_H)/\gamma > 0$ and $3 \geq a^*_{bW} = (C_W - \theta_L)/\gamma > 0$. 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.

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 equlibrium 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 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, expected equilibrium profits and the expected profit of each show type mimicking the other must be derived. 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 recieves 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:

$$\pi^g(a_{gW}, W) = 3 - a_{gW} \tag{11}$$

$$\pi^g(a_{bW}, W) = 3 - a_{bW} \tag{12}$$

$$\pi^b(a_{bW}, W) = 3 - a_{bW} \tag{13}$$

$$\pi^b(a_{gW}, W) = 1 - a_{gW} + 2\mathbb{1}(a_{gW} \geq a^*_{bW}) \tag{14}$$

Equations (11) and (13) represent the equilibrium profits for a high and low quality show respectively, each of them attracting 3 views. Equation (12) represents the case where the high quality show mimics the low quality show, which would lead to the same number of views. Equation (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:

$$\pi^g(a_{gW}, W) = 3 - a_{gW} \geq 3 - a_{bW} = \pi^g(a_{bW}, W) \text{ and} \tag{15}$$

$$\pi^b(a_{bW}, W) = 3 - a_{bW} \geq 1 - a_{gW} + 2\mathbb{1}(a_{gW} \geq a^*_{bW}) = \pi^b(a_{gW}, W). \tag{16}$$

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 (15) and (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:

$$\max\{a^*_{gW}, a_{bW} - 2\} \leq a_{gW} < \min\{a_{bW}, a^*_{bW}\} \tag{17}$$

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 (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:

$$0 < a^*_{gW} < a^*_{bW} \leq 3. \tag{18}$$

According to constraint (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 1 shows that any separating equilibrium with positive advertising levels and viewership (i.e., satisfying constraint (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 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* $\mu_0 < \frac{(C_W - \theta_L - \gamma a_{gW})}{\theta_H - \theta_L} = \frac{2\gamma}{\theta_H - \theta_L}$.

In Proposition 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 multiperiod 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. [10]

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 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

[10] The results of this paper 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.

watching the show.

Lemma 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 $a_{\mu_0} \doteq \mu_0 a^*_{gW} + (1 - \mu_0)a^*_{bW}$.*

Depending on model parameters and out-of-equilibrium beliefs, there exist a continuum of pooling equilibria. In Proposition 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. *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 $\mu_0 > \frac{2\gamma}{\theta_H - \theta_L}$.*

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 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 1 and 2, $\frac{2\gamma}{\theta_H - \theta_L}$.

[Insert Figure 1 about here]

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 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 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 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 sequential TV setting, 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.

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:

$$\pi^g(a, W) = 3\mathbb{1}(\theta_H - C_W + \gamma a \geq 0) - a \tag{19}$$

$$\pi^g(a, B) = 3\delta\mathbb{1}(\theta_H - C_B + \gamma a \geq 0) - a \tag{20}$$

$$\pi^b(a, W) = 3\mathbb{1}(\theta_L - C_W + \gamma a \geq 0) - a \tag{21}$$

$$\pi^b(a, B) = 3\delta\mathbb{1}(\theta_L - C_B + \gamma a \geq 0) - a. \tag{22}$$

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

$$a_{tX} = \begin{cases} \left[\left(C_X - \mathbb{E}[Q_t]\right)/\gamma\right]^+ & \text{if } \pi^t\left(\left[\left(C_X - \mathbb{E}[Q_t]\right)/\gamma\right]^+, X\right) \geq 0 \\ 0 & \text{otherwise.} \end{cases} \tag{23}$$

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 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 shows the different regimes.

Lemma 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).*

[Insert Figure 2 about here]

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:

$$\pi^g(a_{gW}, W) < \pi^g(a_g^*, B) \text{ and} \tag{24}$$

$$\pi^b(a_{bW}, W) < \pi^b(a_b^*, B), \tag{25}$$

where a_{gW} and a_{bW} are defined according to Equation (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 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 20 and 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.

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 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 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).*

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 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 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* $a_g \geq a_W \doteq \frac{1}{\gamma}(C_W - \theta_H)$.

Through Lemmas 3 and 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

$$\pi^g(a_b, B) = 3\delta - a_b^* \tag{26}$$

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:

$$\pi^b(a_g, W) = 1 - a_g + 2\mathbb{1}(a_g \geq a_{bW}), \tag{27}$$

where $a_{bW} = \frac{1}{\gamma}(C_W - \theta_L)$ is defined in Equation (23). When $a_g \geq a_{bW}$ we are in case (A) and Equation 27 becomes $\pi^b(a_g, W) = 3 - a_g$, whereas when $a_g < a_{bW}$ we are in case (B) and Equation 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:

$$\pi^g(a_g, W) = 3 - a_g \geq 3\delta - a_b = \pi^g(a_b, W) \text{ and} \tag{28}$$

$$\pi^b(a_b, B) = 3\delta - a_b \geq 1 - a_g + 2\mathbb{1}(a_g \geq a_{bW}) = \pi^b(a_g, W)). \tag{29}$$

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

$$3 \geq a_g = a_b + 3(1 - \delta) \geq a_{bW} \text{ and } a_g, a_b > 0, \tag{30}$$

where $a_b = \frac{1}{\gamma}(C_B - \theta_L)$ according to Lemma 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 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:

$$3(1-\delta) + a_b \geq a_g \geq \max\{1 - 3\delta + a_b, a_W\}, \quad a_{bW} > a_g \text{ and } 3\delta \geq a_b. \tag{31}$$

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 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 (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 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 1 does for traditional TV, but now for deviations that can switch release timing strategies (i.e., from simultaneous releases to sequential releases).

Lemma 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 $a_W \geq 1+a_b-3\delta$.*

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 6 we find the conditions under which it does.

Lemma 6. *Let $\bar{a}_g \doteq 3\delta - \pi^g(a_g, W) = a_g - 3(1-\delta)$. If both release timing channels 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*

$$\mu_0 < \gamma \frac{a_b^* - \bar{a}_g}{(\theta_H - \theta_L)} \text{ and } \delta - \bar{a}_g \geq 3\delta - a_b^*. \tag{32}$$

The intuition behind Lemma 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 5 performs the same task but for deviations to linear sequential releases. In Proposition 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 3. *When both release timing channels 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*

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

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 TV, 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 channel 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 channel 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 3.

[Insert Figure 3 about here]

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 $\delta - a \geq \delta - \bar{a}_g = 0.5364 > 3\delta - a_b^* = 0.3364$, low quality shows 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 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).

[Insert Figure 4 about here]

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.

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 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 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 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 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:*

$$\pi^b(a_g, W) = \begin{cases} 3 - a_g & \text{if } a_g \geq \frac{C_W - \mathbb{E}[Q_b]}{\gamma} \\ (1 + b + b^2) - a_g & \text{if } \frac{C_W - \mathbb{E}[Q_b]}{\gamma} > a_g \end{cases} \tag{33}$$

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 (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 from 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 3 we have that $a_b = a_b^*$ in order for these equilibriua 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 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 4 we find such conditions.

Proposition 4. *When both release timing channels 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* $a_W < 1 + a_b - 3\delta$.

Proposition 4 describes the parameter set under which we see separation in release timing 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 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 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

6 Extensions

6.1 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.

6.2 DVR Discussion

It is interesting to think about how the digital video recorder (DVR) might moderate these results. Although we do not formally model such, we do provide a discussion of its possible effects. DVRs allow viewers to record content aired at any time and then watch it at their own convenience. On the viewer side, DVRs decrease the viewing cost of sparse releases, moving it closer to the viewing cost of simultaneous releases. On the advertising side, they decrease the advertising responsiveness of ads, because viewers might now binge-watch a show and even skip the advertisements, making sequential linear TV act more like simultaneous. With these two effects in mind, it is not clear whether it will be easier to find separation through the release timing strategy; it depends on the strength of each effect. If advertising responsiveness was not affected by DVRs, we would find that separation is harder, since low-quality shows may now bare the reduced viewing costs of the weekly release timing strategy. On the other side, if the cost of viewing was not affected by DVRs, separation would now be more likely, since low quality shows now have lower incentives to choose weekly releases due to their reduced ad responsiveness.

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 sequential linear 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 channel 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 company, such as NBCUniversal, is tasked with selecting which channel to releases new content through 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 channels 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 companies and the online streaming channel 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.

References

[Abe, 1995] Abe, M. (1995). 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.

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

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

[Bagwell, 2007] Bagwell, K. (2007). The economic analysis of advertising. volume 3, chapter 28, pages 1701–1844. Elsevier, 1 edition.

[Bagwell and Riordan, 1991] Bagwell, K. and Riordan, M. H. (1991). High and declining prices signal product quality. *The American Economic Review*, pages 224–239.

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

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

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

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

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

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

[Cho and Kreps, 1987] Cho, I.-K. and Kreps, D. M. (1987). Signaling games and stable equilibria. *The Quarterly Journal of Economics*, 102(2):179–221.

[Chung et al., 2013] Chung, K. Y., Derdenger, T. P., and Srinivasan, K. (2013). Economic value of celebrity endorsements: Tiger woods' impact on sales of nike golf balls. *Marketing Science*, 32(2):271–293.

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

[Derdenger et al., 2018] Derdenger, T. P., Li, H., and Srinivasan, K. (2018). 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.

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

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

[Erdem et al., 2008] Erdem, T., Keane, M. P., and Sun, B. (2008). A dynamic model of brand choice when price and advertising signal product quality. *Marketing Science*, 27(6):1111–1125.

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

[Gal-Or, 1989] Gal-Or, E. (1989). Warranties as a signal of quality. *Canadian Journal of Economics*, pages 50–61.

[Galak et al., 2013] Galak, J., Kruger, J., and Loewenstein, G. (2013). Slow down! insensitivity to rate of consumption leads to avoidable satiation. *Journal of Consumer Research*, 39(5):993–1009.

[Gerchak et al., 2006] Gerchak, Y., Cho, R. K., and Ray, S. (2006). Coordination of quantity and shelf-retention timing in the video movie rental industry. *IIE Transactions*, 38(7):525–536.

[Gibbons, 1992] Gibbons, R. (1992). *Game theory for applied economists*. Princeton University Press.

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

[Hennig-Thurau et al., 2007] Hennig-Thurau, T., Henning, V., Sattler, H., Eggers, F., and Houston, M. B. (2007). The last picture show? timing and order of movie distribution channels. *Journal of Marketing*, 71(4):63–83.

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

[Jiang et al., 2011] Jiang, B., Jerath, K., and Srinivasan, K. (2011). Firm strategies in the mid tail of platform-based retailing. *Marketing Science*, 30(5):757–775.

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

[Krider and Weinberg, 1998] Krider, R. E. and Weinberg, C. B. (1998). Competitive dynamics and the introduction of new products: The motion picture timing game. *Journal of Marketing Research*, pages 1–15.

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

[Linnemer, 2002] Linnemer, L. (2002). Price and advertising as signals of quality when some consumers are informed. *International Journal of Industrial Organization*, 20(7):931–947.

[Lu et al., 2017] Lu, T., Bradlow, E. T., and Hutchinson, J. W. (2017). Binge consumption of online content. In *Working Paper*.

[Lutz, 1989] Lutz, N. A. (1989). Warranties as signals under consumer moral hazard. *The Rand journal of economics*, pages 239–255.

[Milgrom and Roberts, 1986] Milgrom, P. and Roberts, J. (1986). Price and advertising signals of product quality. *Journal of political economy*, 94(4):796–821.

[Moorthy, 2012] Moorthy, S. (2012). Can brand extension signal product quality? *Marketing science*, 31(5):756–770.

[Moorthy and Srinivasan, 1995] Moorthy, S. and Srinivasan, K. (1995). Signaling quality with a money-back guarantee: The role of transaction costs. *Marketing Science*, 14(4):442–466.

[Moraga-González, 2000] Moraga-González, J. L. (2000). Quality uncertainty and informative advertising. *International Journal of Industrial Organization*, 18(4):615–640.

[Nathanson, 2013] Nathanson, J. (2013). The economics of a hit tv show. https://priceonomics.com/the-economics-of-a-hit-tv-show. Accessed: 2013-10-17.

[Nelson, 1974] Nelson, P. (1974). Advertising as information. *Journal of political economy*, 82(4):729–754.

[Prasad et al., 2004] Prasad, A., Bronnenberg, B., and Mahajan, V. (2004). Product entry timing in dual distribution channels: The case of the movie industry. *Review of Marketing Science*, 2(1).

[Schweidel and Moe, 2016] Schweidel, D. A. and Moe, W. W. (2016). Binge watching and advertising. *Journal of Marketing*, 80(5):1–19.

[Simester, 1995] Simester, D. (1995). Signalling price image using advertised prices. *Marketing Science*, 14(2):166–188.

[Stigler, 1961] Stigler, G. J. (1961). The economics of information. *Journal of Political Economy*, 69(3):213–225.

[Stigler and Becker, 1977] Stigler, G. J. and Becker, G. S. (1977). De gustibus non est disputandum. *The American Economic Review*, 67(2):76–90.

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

[Toonkel and Richwine, 2016] Toonkel, J. and Richwine, L. (2016). How network tv figured out binge-watching. http://fortune.com/2016/03/11/netflix-changing-game-network-tv//. Accessed: 2018-12-05.

[Trouleau et al., 2016] Trouleau, W., Ashkan, A., Ding, W., and Eriksson, B. (2016). 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.

[Wernerfelt, 1988] Wernerfelt, B. (1988). 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.

[Zhao, 2000] Zhao, H. (2000). Raising awareness and signaling quality to uninformed consumers: A price-advertising model. *Marketing Science*, 19(4):390–396.

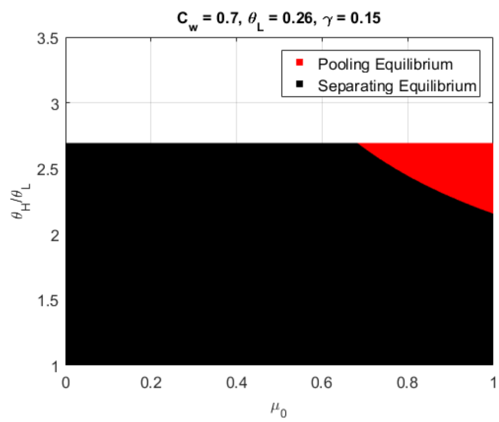

Figure 1: Regions in $(\mu_0, \theta_H/\theta_L)$ in which we find a separating or pooling equilibrium in the advertising level that survives the intuitive criterion when the only release timing possibility is sequential releases. The parameters are such that $C_W = 0.7$, $\theta_L = 0.26$ and $\gamma = 0.15$.

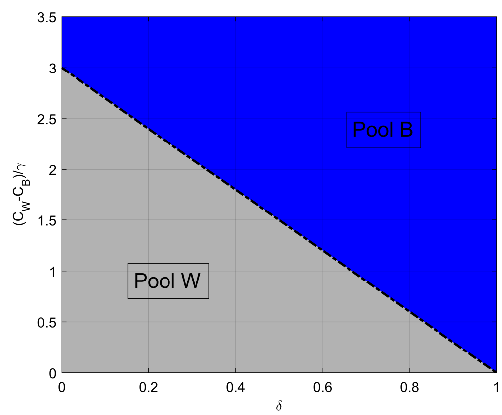

Figure 2: Regions in $(\delta, (C_W - C_B)/\gamma)$ with complete information under which we find pooling to sequential linear or simultaneous releases, as well as where we may find both or separating release timing strategies (the black line).

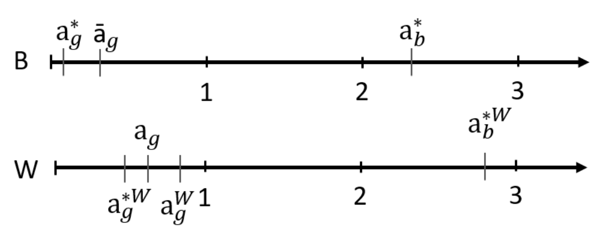

Figure 3: Advertising values according to the described example: $\theta_H = 1.85$, $\theta_L = 0.6$, $\gamma = 0.55$, $C_B = 1.9$, $C_W = 2.15$, $\delta = 0.9$, $\mu_0 = 0.1$, $g = 1$ and $b = 0$. $a_g^{*W} = 0.5455$, $a_b^{*W} = 2.8182$, $a_g^W = 0.8182$, $a_g^* = 0.0909$, $a_b^* = 2.3636$, $a_g = 0.6636$ and $\bar{a}_g = 0.3636$.

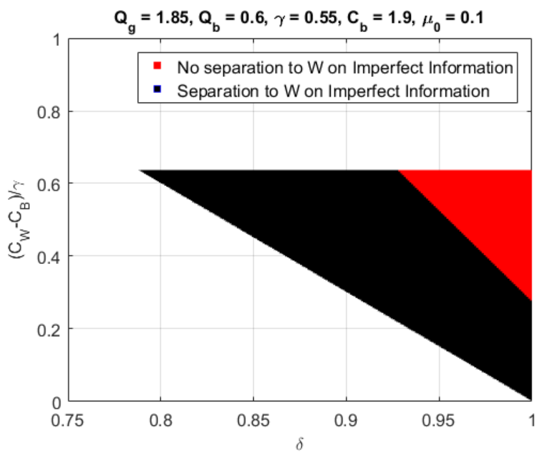

Figure 4: Regions in $(\delta, \frac{C_W - C_B}{\gamma})$ under which the high quality show separates to sequential releases or not, from a first best solution where both shows pool on simultaneous releases.

Appendix A

Proof of Proposition 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 (15) and (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 (15) and (16), which we can rewrite using constraint (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. □

Proof of Lemma 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 $a_p \geq a_{\mu_0}$. □

Proof of Proposition 2. From Lemma 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. □

Proof of Lemma 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.

□

Proof of Lemma 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). □

Proof of Lemma 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. □

Proof of Lemma 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 $a \in [a_W, a_g) \iff \mu_0 < \frac{(C_W - \theta_L - \frac{\alpha}{3} - \gamma a_g)}{\theta_H - \theta_L} = \frac{\gamma(3\delta - 1)}{\theta_H - \theta_L}$. 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. $\square$

Proof of Lemma 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 (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 $\mu_0 < \gamma \frac{a_b^* - \bar{a}_g}{\theta_H - \theta_L}$. $\square$

Proof of Proposition 3. This Proposition follows immediately from using Case B ($a_g < a_{bW}$), Lemmas 5 and 6 and the incentive compatibility constraints (31). $\square$

Proof of Lemma 7. The expected profit in this situation is given by

$$\pi^b(a, W) = \sum_{i=1}^{3} i\mathrm{P}(\text{watching exactly } i \text{ episodes}) - a. \tag{34}$$

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

$$\text{P(watching exactly 1 episode)} = \mathbb{1}(u_{\{1,a,W\}} \geq 0)(1-b)\Big[1 - \mathbb{1}\Big(\gamma a + \mathbb{E}[Q_b] \geq C_W\Big)\Big] \tag{35}$$

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

$$\text{P(watching exactly 2 episodes)} = \mathbb{1}(u_{\{1,a,W\}} \geq 0)b(1-b)\Big[1 - \mathbb{1}\Big(\gamma a + \mathbb{E}[Q_b] \geq C_W\Big)\Big]. \tag{36}$$

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

$$\begin{aligned}\text{P(watching exactly 3 episodes)} =& \mathbb{1}(u_{\{1,a,W\}} \geq 0)\Big[(1-b)\mathbb{1}\Big(\gamma a + \mathbb{E}[Q_b] \geq C_W\Big) \\ &+ b\Big[b + (1-b)\mathbb{1}\Big(\gamma a + \mathbb{E}[Q_b] \geq C_W\Big)\Big]\Big].\end{aligned} \tag{37}$$

Finally, using Equations (35), (36) and (37) in Equation (34), and knowing that in a separating equilibrium $u_{\{1,a_g,W\}} \geq 0$, we obtain the result we are looking for. □

Proof of Proposition 4. The proof of this Proposition is very similar to the ones of Lemmas 5 and 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. $\square$