Ad-Skipping and Time-Shifting: A Theoretical Examination of the Digital Video Recorder

Sunit Nitin Shah
Huntington, NY

B. A., University of Virginia, 2001
M. A., University of Virginia, 2003

A Dissertation Presented to the Graduate
Faculty of the University of Virginia
in Candidacy for the Degree of
Doctor of Philosophy

Department of Economics

University of Virginia
May 2011

Acknowledgements

At the end of a journey as long as this, it is inevitable that one reflect on how much support he has received over the years to reach that end. It has been no different for me. This dissertation is not only the product of my own efforts, but it is also a result of the tremendous support I have received from numerous individuals throughout my time in graduate school.

I'd like to start my thanking my committee. Thanks to my main advisor, Simon Anderson, for all of his guidance throughout this project, and for always helping me stay focused on the bigger picture regardless of how deep I delved into the mathematical detail. Thanks also to my second advisor, Maxim Engers, who sharpened my ability to create theories, develop proofs, and communicate ideas concisely and effectively. I'd also like to thank my third reader, Nathan Larson, for his support honing the details contained within this document, and my fourth reader, Natasha Foutz, for bringing both an outside perspective and some much-appreciated enthusiasm to my committee as well. I am also very grateful to both the UVA Alumni Association Class of 1985 Fellowship and the Economics Department for all of the financial support I received throughout my time as a graduate student.

More than anything, I have been extremely fortunate to have received so much support from my friends and family throughout the years. Thank you to my fellow graduate students, Scott Davis and Cat Tyler, to Scott for all of his help creating this document, and for making sure I stayed distracted with Fantasy Football every fall, to Cat for all of her comments on my first essay, and to both for helping form the best Economics study group ever; to Ben Keane and Christian Del'Aune, for all of our POAs, including Saints games, the road trip to Breckenridge, and jaunts to Boston, New York, Chapel Hill, and Montreal, just to name a few; to Amanda Hardison, for her constant companionship, patience, and support throughout my time in graduate school, and for helping me learn so much about myself in the process; to Dominic Depasquale, for always keeping a spot open for me as a place to crash and feel at home, and for making sure the Super Nintendo was always handy as well; to Adam Namejko, for being so enthusiastic about everything in life and for always keeping me in high spirits regardless of the circumstances; to Dirk Koopmans, for expanding my horizons beyond the world of Economics, and for being there for me whenever I needed someone; to Caleb Doise, for always making time to help me kick back with some video games and blow off some steam; to Greg Herrington and Kennedy Prentiss, for keeping me fed and for countless trips to the farm; and to Maya, for her never-ending companionship and affection, especially throughout the last year of this endeavor.

Most importantly, I'd like to thank my parents, Nitin and Suhasini Shah, my sister, Vaishali Shah, my niece, Kirsi Shah Chinn, and all of the rest of my family for all of their patience and support throughout my time in graduate school. I love you all and would not be where I am today without you.

All remaining errors are my own.

Abstract

I examine the impact of the two main aspects of the DVR – ad-skipping and time-shifting – on the television industry and their effects on equilibrium outcomes. I begin by exploring ad-skipping in a monopoly model in which the network delivers content to viewers with decreasing, constant, or increasing marginal nuisance costs (MNCs) of advertising and faces competition only from a static outside option. With increasing MNCs, the DVR can increase network profit depending on the percentage of ads it filters out; otherwise, it reduces network welfare. I then adapt the constant MNC model to a duopoly setting in which two identical networks face each other rather than an outside option. In this context, the DVR unambiguously helps each network, both through the decreased sensitivity the most ad-averse consumers feel towards ads when using a DVR and through the element of price discrimination it affords.

Next, I examine the time-shifting capabilities of the DVR in a model in which two networks, each with a hit show and a mediocre one, compete for viewers over two timeslots. I start by exploring the lead-in effect, or the tendency for viewers to watch the same network over consecutive timeslots. When viewers are myopic, equilibrium lineups with either staggered hit shows or ones going head-to-head can exist. However, when viewers are foresighted, the lead-in effect essentially becomes bi-directional, and only staggered lineups can exist in equilibrium. I then examine the effect of the DVR, which allows viewers to watch any subset of the four shows. The

networks are then indifferent between their lineup choices, even as the equilibrium proportion of viewers that rent a DVR becomes arbitrarily small. Lastly, I examine what happens if the DVR also eliminates a proportion of network ad revenues derived from viewers who use the device, and I demonstrate that if this proportion is greater than one-half, network profits fall with the DVR, whereas if it is less than one-third, profits are increased by it.

Table of Contents

Overview

The television industry was forever changed with the advent of the commercially-available digital video recorder in 1999. Television distribution, like any other media, relies on the ability to bundle ads within the content the recipients would like to consume, which historically occurred by airing short spurts of ads every few minutes within a program. However, the DVR gives consumers the additional option of recording a television show for future consumption, at which point they have complete control over the flow of information within that program – ads and content both. Consequently, viewers can fast forward through the ads aired during recorded shows, receiving only a portion of the effective ad impressions. However, a viewer with a DVR can also record a show that airs when he or she might otherwise not be available for future consumption, expanding the roster of television programs at his or her disposal.

I examine the effects of each of these aspects of the DVR in turn. Part I explores the effect of ad-skipping on the television industry. First I explore a scenario in which a monopolist network faces a spectrum of ad-averse viewers who can choose between watching the network's show or performing the outside option. When viewers have decreasing or constant marginal nuisance costs, I demonstrate how the advent of the DVR necessarily lowers ad revenues. However, when viewers have increasing marginal nuisance costs, the last few ads seen cause the highest nuisance, and hence the DVR becomes even more valuable to them. Consequently, at a given filter rate, more users then convert from the outside option to watching television with a DVR. I demonstrate that the ad impressions gained from these viewers may outweigh those lost from individuals who switch from watching television to watching it with a DVR after the network optimizes under the new conditions.

I then modify the game so that each network now competes with an identical network rather than a static outside option. Each must therefore account for the dynamic nature of its competition rather than simply facing the invariable status of the outside option. I show that, in this model, the networks are unambiguously benefited by the DVR due to the extra market power it affords each in the form of modified price discrimination. Effectively, the DVR allows each network to "charge" mildly ad-averse individuals a high ad level while "charging" very ad-averse individuals a lower ad level due to the filtering mechanism of the device. Further, viewer welfare is not necessarily hurt by this result, since the ad burden shifts from those who are very ad-averse to those who are only mildly so.

In Part II, I develop a time-competition game in which two networks compete for viewers over two timeslots. Each network is endowed with one hit show and one mediocre show and needs to decide what to air in each timeslot. Potential viewers choose whether to watch one of the two networks or do something else; they are heterogenous across both programming preference and the value of their outside options. First I explore the "lead-in" effect – the tendency for those who watch one program to watch the subsequent program on the same channel – which is the main driver of networks' timing strategies in the traditional literature. I explore this effect, first when viewers are myopic and then again when they have perfect foresight. In the myopic treatment equilbrium, the networks could either go head-to-head with their hit shows or stagger their lineups, depending on the strength of the lead-in effect. When viewers are foresighted, however, the lead-in effect essentially becomes bidirectional, as viewers plan their behavior in advance based on their two-period payoffs over both timeslots. Consequently, the networks always opt to avoid direct competition in the foresighted viewer equilibrium, regardless of the lead-in effect's magnitude.

Next, I incorporate a digital video recorder (DVR), which, for a positive rental cost, allows viewers to "time-shift" programming - i.e., record a show and watch it another time. In equilibrium, DVR users comprise a larger mass of viewers with head-to-head lineups than with staggered ones. This effect cancels out any advantage the networks enjoyed with staggered lineups so that, for any parameter values, the networks' profits remain constant regardless of their lineup choices when DVR technology is available. This is true despite the fact that changes in the DVR's rental cost cause the mass of consumers who use DVRs under the various lineup scenarios to vary differently for different parameter values. Further, the critical mass of DVR users necessary to discipline the market can be arbitrarily small and the neutrality result still holds. Lastly, I add ad-filtering into the model, and find that if the DVR filters out fewer than one of every three ads, network profits increase with the DVR, whereas if the DVR filters out more than one out of every two ads, network profits necessarily fall.

Part I

Ad-Skipping

Chapter 1

Introduction

At the core of the television industry lies a strategic game for the networks. Each network acts as a middleman between advertisers and viewers: advertisers want to reach potential customers, but consumers dislike ads[1] and will watch a program only so long as its value exceeds that of the next-best option. Because network revenue depends on the number of times ad messages are seen, each network must choose its ad allotment carefully. They must also bundle the ads within the content in such a way that it would be difficult to avoid the former but still enjoy the latter. That the ads are effectively bundled is critical to the arrangement – viewers would avoid the ads entirely if they could, which would drive network revenue to zero and hence unravel the entire relationship. Historically, networks bundle ads by airing short spurts of thirty-second spots between small sections of the program. Each viewer then has but a few minutes to start and stop another activity, making it difficult to avoid the ads but still see the content in its entirety.

However, the digital video recorder (hereafter, DVR) has changed the game. A DVR allows a viewer to record programming, after which the user has complete control

[1]See Wilbur (2007) for empirical evidence of viewer aversion to ads.

over the flow of the content upon playback. At that point, the viewer can then fast-forward through whatever portions she chooses, including commercials. The user can even start playing the recording back before the program is complete, allowing one to delay the start of a show and then catch up by fast-forwarding through the commercials. To facilitate the practice, many DVRs even have a button to skip forward exactly thirty seconds, the length of a typical commercial. With these tools, a DVR user can effectively unbundle the content from the advertising. That these capabilities matter to viewers is evidenced by its popularity: Nielsen estimates that DVR penetration was 23% in 2008 (Eggerton, 2008), and Forrester Research expects that number to reach 50% by 2012 (Downey, 2007).

While the evidence is mixed, some studies have shown that DVR users skip through a significant portion of the ads in the programs they watch (Goetzl, 2006). Consequently, the potential impact on the industry is huge; here, I present two models to explore that impact.[2] I begin with a monopoly model in which one network faces a spectrum of consumers that choose whether or not to watch its show. Next, I incorporate the DVR, which allows consumers the additional option of paying a rental cost to bypass a fraction of the ads, and examine how the device affects the outcome. I then adapt the setting to a duopoly in which two identical networks compete with each other as opposed to an outside option, and explore the equilibria that exist first with and then without DVR technology.

The layout of the ad-skipping analysis is as follows. The subsequent chapter reviews the relevant literature. In Chapter 3, I develop a model in which a monopolist network competes with only the outside option, so that viewers must decide between watching the sole network or doing something else. In Chapter 4 then adapt the

[2]While I focus here on the television industry, one could apply the model to other media with bundled advertising and low duplication costs, such as radio or the internet.

model to a duopoly setting in which two identical networks face competition solely from each other. The dynamics of the game change significantly when each network faces a responsive opponent rather than a static alternative. Chapter 5 concludes.

Chapter 2

Literature Review

The models I use here trace their roots back to the famous line model from Hotelling (1929). Bowman and Farley (1972) and Bowman (1975) applied the idea to the television industry empirically. Gabszewicz, Laussel, and Sonnac (2000) first included ads as a nuisance with the Hotelling line representation of viewer heterogeneity; in their model, each viewer's location corresponds to her ideal programming mix. Since then, it has become standard to model ad-averse viewers who are distributed across a line that represents heterogeneity in their preferences for programming (see Peitz and Valletti (2004), Anderson and Coate (2005), or Ambrus and Reisinger (2006) for examples). However, none of these models incorporate DVR technology. Ad-aversion serves as a critical driver of viewer behavior in these models; hence, incorporating a product that allows users to avoid ads could have a large effect on equilibrium outcomes. Despite the DVR's impact on the industry, Anderson and Gans (2011) is the only paper thus far to model the technology. Tag (2009) also gives a theoretical treatment on a related topic.

Tag (2009) explores a question similar to the one here, but applied instead to the internet advertising industry. Consumers are heterogeneous in their preference for web site quality but homogeneous in their aversion to ads. The internet firm can provide a free web site with ads, a fee-based website that is ad-free, or provide both and allow each viewer to choose. In essence, the third option allows consumers to pay a subscription fee to avoid the ads, similar to a DVR's rental cost. Tag characterizes the conditions under which each option gives the firm the most profit. He finds that availability of the ad-avoidance option increases ad levels and harms consumers. While that analysis is of a different industry, the insights gleaned can still be compared to those here as both involve a media firm with sunk production costs and low duplication costs as well as consumers who have the option of paying to avoid advertising.

Anderson and Gans (2011) is more similar to the work presented here as it looks specifically at the question of how an ad-avoidance device affects two-sided markets. Consumers are heterogeneous in their personal aversions to ads as well as along their preferences for show quality. The authors examine how the equilibrium outcome of the game changes when the viewers obtain the option to purchase a technology that will eliminate all advertisements from the programming. They find that the availability of DVR technology causes ad levels to rise. Total welfare may decrease, and the least ad-averse individuals, who do not purchase the technology, bear a disproportionate burden of the advertising nuisance.

This model adds to those in Tag (2009) and Anderson and Gans (2011) in two key ways. First, the models in both Tag (2009) and Anderson and Gans (2011) involve consumers with constant marginal nuisance costs (hereafter, MNCs) for advertising. Here, I examine the constant MNCs case but also analyze how the results change if the viewers have decreasing or increasing MNCs. This represents an addition to the broader TV advertising literature as well, as most research in this area assumes constant MNCs without considering the alternatives. Also, the fee-based web site in

Tag (2009) and the ad-avoidance product in Anderson and Gans (2011) completely eliminate advertising, but the empirical evidence on the DVR's effectiveness is mixed. Some studies suggest DVR users can avoid almost all advertising (Goetzl, 2006), while others have found DVR users still watch a significant portion of the ads they could skip (Story, 2007). Moreover, data suggest that DVR users fast-forwarding through ads does not completely nullify the advertising message (Wilbur, 2008). In general, the magnitude of the DVR's effect on ad exposure is still under debate (Snedeker, 2007). Because of this uncertainty, I have parameterized the DVR's efficiency so that it can vary from eliminating all of the ads to eliminating none at all. This also allows one to examine how the equilibrium outcome will change over time as DVR technology evolves.

Each related theoretical work also differs individually from the models presented here. Because the media firm in Tag (2009) receives the fees spent on ad removal, whereas here the network and the DVR manufacturer are different entities, the incentives that the media firms face in each model differ. When an internet user switches from the ad-based version to the fee-based one, the internet company feels one negative and one positive effect on its bottom line; however, when a TV viewer decides to watch a show with a DVR rather than without one, all other things equal, there is only a negative effect on the network's revenues.

In Anderson and Gans (2011), the ad-avoidance product in the baseline model[1] is a durable good, so the consumers are bound by the purchasing decisions they make. This impacts the game's timing mechanism; the viewers move first, after which the networks react. However, DVR purchases have dropped dramatically, and most DVR users now rent the device. By the end of 2006, over 85% of users were renting (Boles,

[1] As an extension, the authors do consider the DVR as a rental item, though all equilibria in this case involve trivial DVR usage.

2007); the combined market share of DVR producers such as TiVo was 4% in 2007 and is expected to fall to 2% by 2010 (Edwards, 2007). As a result, I have modeled the DVR as a rental item. Because the viewer's rental decision is of shorter term than the process networks undergo to recruit advertisers, in these models, the network(s) choose the ad allotments first, and then the consumers react to those decisions.

In addition to developing an original model of DVR technology, I contribute to the existing literature in three keys ways. First, the monopoly model borrows aspects from both the standard product differentiation and TV advertising models; I tie the two together by demonstrating a conversion between them. Further, I expand the monopoly model to compare the effects of increasing, decreasing, and constant marginal nuisance costs (hereafter, MNCs). Lastly, in both models, I allow the DVR's pass-through rate to vary, which adds some flexibility to the model and also allows one to examine how the outcome will change as the technology evolves over time.

Chapter 3

Monopoly

3.1 Model

There is one television network and it airs one show. I assume production and other costs for this show are sunk, so the network will always air the show. The network also airs advertisements within the show, from which it receives ad revenue. Call an instance of a viewer watching an ad an "impression." Impressions provide a constant return to the network; the network therefore chooses its ad allotment, a, to maximize total impressions.

After the network makes and announces its advertising allotment decision, consumers have a choice between watching the television show and doing something else, which without loss of generality gives 0 utility. Viewers receive a utility V from watching the television show free of advertising but derive some displeasure from watching ads. They are heterogeneous in their aversion to ads, with aversion parameter γ_i. Specifically, consumer i's utility from watching the show with a advertisements is

$$U_i = V - \gamma_i a^s,$$

where s determines whether MNCs are decreasing ($s < 1$), constant ($s = 1$), or increasing ($s > 1$).

To solve for an explicit equilibrium, I specify a functional form for the distribution of γ_i:

$$F(\gamma) = 1 - \frac{1}{\gamma} \Longrightarrow f(\gamma) = \gamma^{-2} \text{ for } \gamma \in [1, \infty)$$

Note that this is equivalent to the model:

$$U_i = \theta_i \cdot V - a^s,\, \theta \sim U[0, 1]$$

with $\gamma_i = \frac{1}{\theta_i}$. In the latter, the heterogenous paramater is applied to V rather than to the nuisance costs, which follows the norm in the product differentiation literature (see, e.g., Mussa and Rosen (1978)). However, following the norm in the related advertising literature (see Peitz and Valletti (2004), Anderson and Coate (2005), Anderson and Gabszewicz (2006), or Ambrus and Reisinger (2006), among others, for examples) I apply the nuisance parameter directly to the cost of watching advertisements. Note that the two versions of the model above are equivalent[1]. While I will use the latter for the exposition, the reader should feel free to think of the model in terms of whichever version is most comfortable.[2]

The timing of the game is as follows:

1. The network chooses and announce its ad allotments, to which it is then committed,
2. Consumers decide whether to watch network 1 or exercise the outside option,
3. All parties receive their payoffs.

[1]See Section A.1 for a proof.

[2]This equivalence also allows a more direct comparison to Tag (2009) which specifies a model similar to the product differentiation version above.

3.2 Equilibrium

To solve for the equilibrium, I start in the final stage and work backwards. Immediately before parties receive their payoffs, consumers choose whether to watch the show or exercise the outside option given the networks' choice of ad allotment in stage 1. Conditional on an ad allotment a, the consumer for whom

$$V - \gamma_i a^s = 0 \Longrightarrow \gamma_i = \frac{V}{a^s} \equiv \hat{\gamma}$$

is indifferent between watching TV and exercising her outside option. Since utility is strictly decreasing in γ_i, all viewers to the left of $\hat{\gamma}$ watching television and those to the right do not. Figure 3.1 presents a graphical depiction of the viewers' choices.

Given consumers' optimal responses, the network can determine its eventual profit as a function of the ad allotment it chooses:

$$\begin{aligned} \pi &= a \cdot F\left(\frac{V}{a^s}\right) \\ \Longrightarrow &= a\left(1 - \frac{a^s}{V}\right) \\ \Longrightarrow &= a - \frac{a^{s+1}}{V} \end{aligned}$$

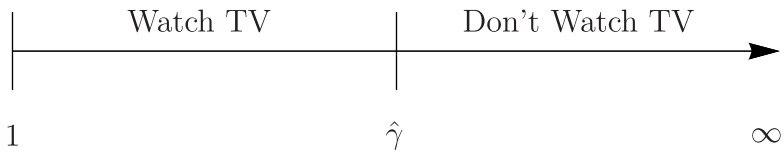


Figure 3.1: Viewer Optimal Response - No-DVR Case

so that the first-order condition is:

$$\begin{aligned} \frac{d\pi}{da} &= 0 \\ \Longrightarrow 1-(s+1)\frac{a^s}{V} &= 0 \\ \Longrightarrow a^* &= \left(\frac{V}{s+1}\right)^{\frac{1}{s}} \end{aligned}$$

Since

$$\frac{d^2\pi}{da^2} = -s(s+1)\frac{a^{s-1}}{V} < 0$$

the second-order condition is satisified, and hence the ad allotment above represents a maximum.

In response to this ad allotment, the viewer located at

$$\hat{\gamma}^* = \frac{V}{a^{*s}} = s+1$$

will be indifferent between watching television and doing something else; all those to the left will choose to watch the show, whereas all those to the right will choose the outside option. The following proposition summarizes this result.

Proposition 1 *In the ad-skipping monopoly model, the unique pure strategy equilib-*

rium outcome involves the network setting an ad allotment of

$$a = a^* = \left(\frac{V}{s+1}\right)^{\frac{1}{s}}$$

and the cutoff viewer being the one for whom

$$\gamma_i = \hat{\gamma} = s + 1$$

with all viewers to the left of $\hat{\gamma}$ *watching television and all those to the right exercising the outside option.*

In equilibrium, the mass of viewers that watch television is

$$F(\hat{\gamma}^*) = 1 - \frac{1}{\hat{\gamma}^*} = \frac{s}{s+1}$$

so that the network makes a profit of

$$\pi^* = a^* \cdot F(\hat{\gamma}^*) = \frac{s}{s+1} \cdot \left(\frac{V}{s+1}\right)^{\frac{1}{s}}$$

In the sole pure strategy equilibrium of this model, the network faces a tradeoff when setting its ad allotment. Consider a network planning to set an ad allotment of a. An increase to this ad allotment would result in a marginal reduction in the resulting viewership, for whom any potential ad impressions would be entirely lost. However, there would also be an increase in ad impressions from the entire mass of viewers that would watch in either case. This tradeoff varies with the planned level of a; the network aims to maximize impressions based on this tradeoff, which results in an interior solution regardless of parameter values. The sole pure strategy equilibrium in this case follows from this maximization.

3.3 The Digital Video Recorder

In this section, I adapt the previous model to include an ad-skipping device, the digital video recorder. When this technology is available, the model remains the same except that consumers have the additional option of renting a DVR device that, for a rental cost of r $(0 < r < V)$, allows the viewer to avoid a proportion of the advertising. I call the advertising pass-through rate α $(0 < \alpha < 1)$, and occassionally refer to the proportion of ads blocked, $(1 - \alpha)$, as the DVR's "efficiency". After the network makes and announces its advertising allotment decision, consumer i will then aim to maximize:[3]

$$U_i = \begin{cases} V - \gamma_i a^s, \text{if consumer } i \text{ watches television without a DVR} \\ V - \gamma_i (\alpha a)^s - r, \text{if consumer } i \text{ watches television with a DVR} \\ 0 \text{ otherwise} \end{cases} \tag{3.1}$$

To determine the equilibrium of the adapted model, I again begin with the viewer response to a given network ad allotment. Conditional on an ad level a, there will now be two cutoff points: one representing the viewer indifferent between watching TV without a DVR and watching it with one, and a second representing the viewer indifferent between watching TV with a DVR and not watching TV at all.

Call the former cutoff point $\hat{\gamma_1}$ and the latter $\hat{\gamma_2}$. Then:

$$\begin{aligned} V - \hat{\gamma_1} a^s &= V - \hat{\gamma_1}(\alpha a)^s - r \\ \Longrightarrow \hat{\gamma_1} &= \frac{r}{(1 - \alpha^s) a^s} \end{aligned}$$

[3]The product differentiation analog to this model involves the θ term applied as $\theta(V - r)$ for viewers using a DVR.

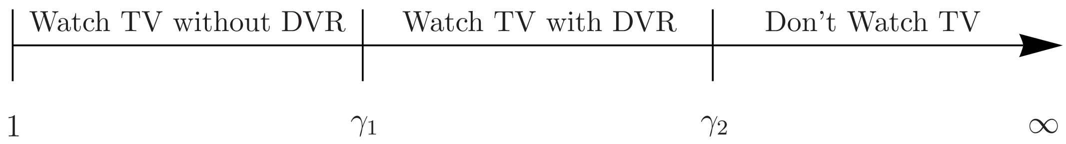


Figure 3.2: Viewer Strategy - DVR Available

and

$$\begin{aligned} V - \hat{\gamma}_2(\alpha a)^s - r &= 0 \\ \Longrightarrow \hat{\gamma}_2 &= \frac{V-r}{\alpha^s a^s} \end{aligned}$$

Assuming at least one viewer chooses each option in equilibrium, viewers for whom $\gamma_i \in [1, \hat{\gamma}_1)$ watch the show without a DVR, viewers for whom $\gamma_i \in (\hat{\gamma}_1, \hat{\gamma}_2)$ watch with a DVR, and viewers for whom $\gamma_i \in (\hat{\gamma}_2, \infty)$ do not watch at all. Figure 3.2 displays the viewers' optimal responses graphically.

Note that, if $\alpha > \left(\frac{V-r}{V}\right)^{\frac{1}{s}}$, $\hat{\gamma}_2 < \hat{\gamma}_1$, so that there effectively exists only one cutoff point and no one chooses to rent a DVR. In this case, the equilibrium simply reverts back to that of the previous section, so that the DVR has no impact on the equilibrium outcome. Essentially, the efficiency-cost combination of the DVR is not attractive enough for any viewer to find its rental worthwhile. To focus on the more interesting cases in which the DVR is worth its cost to some positive mass of viewers, I assume from here forward that $\alpha < \left(\frac{V-r}{V}\right)^{\frac{1}{s}} \equiv \bar{\alpha}$.

Given the viewers' optimal response function, the network's eventual profit as a function of its ad choice is:

$$\begin{aligned} \pi &= F(\hat{\gamma}_2) \cdot \alpha a + F(\hat{\gamma}_1) \cdot (1-\alpha) a \\ &= a - \frac{(\alpha a)^{s+1}}{V-r} - \frac{a^{s+1}(1-\alpha)(1-\alpha^s)}{r} \end{aligned}$$

This yields the first-order condition:

$$
\begin{aligned}
\frac{d\pi}{da} &= 0 \\
\Longrightarrow 0 &= 1 - \frac{(s+1)\alpha^{s+1}a^s}{V-r} - \frac{(s+1)a^s(1-\alpha)(1-\alpha^s)}{r} \\
\Longrightarrow 1 &= a^s\frac{r(s+1)\alpha^{s+1}a^s - (s+1)(V-r)(1-\alpha)(1-\alpha^s)}{r(V-r)} \\
\Longrightarrow a^* &= \left(\frac{r(V-r)}{(s+1)r\alpha^{s+1} + (s+1)(1-\alpha)(1-\alpha^s)(V-r)}\right)^{\frac{1}{s}}
\end{aligned}
$$

Since:

$$
\frac{d^2\pi}{da^2} = -sa^{s-1}\frac{r(s+1)\alpha^{s+1}a^s - (s+1)(V-r)(1-\alpha)(1-\alpha^s)}{r(V-r)} < 0,
$$

the second-order condition is satisfied so that the above solution is a maximum.

The resulting cutoff points in the viewers' optimal response are:

$$
\begin{aligned}
\Longrightarrow \hat{\gamma_1}^* &= \frac{r}{1-\alpha^s} \cdot \frac{(s+1)r\alpha^{s+1} + (s+1)(1-\alpha)(1-\alpha^s)(V-r)}{r(V-r)} \\
&= (s+1)\frac{r\alpha^{s+1} + (1-\alpha)(1-\alpha^s)(V-r)}{(1-\alpha^s)(V-r)}, \\
\hat{\gamma_2}^* &= \frac{V-r}{\alpha^s} \cdot \frac{(s+1)r\alpha^{s+1} + (s+1)(1-\alpha)(1-\alpha^s)(V-r)}{r(V-r)} \\
&= (s+1)\frac{r\alpha^{s+1} + (1-\alpha)(1-\alpha^s)(V-r)}{\alpha^s r}
\end{aligned}
$$

which represent the viewer indifferent between watching TV with a DVR and without one, and the viewer indifferent between watching TV with a DVR and not watching TV at all, respectively.[4]

[4]Depending on the parameter values, it is possible that $\hat{\gamma_1}^* \leq 1$. This would imply that no consumer watches TV without a DVR. Because empirical evidence shows that a significant number

This equilibrium is summarized in the following proposition.

Proposition 2 *The unique pure-strategy equilibrium in the ad-skipping monopoly model with DVR technology available involves the network setting an ad level of:*

$$a^* = \left(\frac{r(V-r)}{(s+1)r\alpha^{s+1} + (s+1)(1-\alpha)(1-\alpha^s)(V-r)} \right)^{\frac{1}{s}},$$

viewers for whom

$$\gamma_i < \hat{\gamma}_1 = (s+1)\frac{r\alpha^{s+1} + (1-\alpha)(1-\alpha^s)(V-r)}{(1-\alpha^s)(V-r)}$$

watching TV without a DVR, viewers for whom

$$\gamma_i > \hat{\gamma}_2 = (s+1)\frac{r\alpha^{s+1} + (1-\alpha)(1-\alpha^s)(V-r)}{\alpha^s r}$$

not watching TV at all, and viewers for whom

$$\hat{\gamma}_1 < \gamma_i < \hat{\gamma}_2$$

watch TV with a DVR.

3.4 Welfare

Now that the equilibrium in the model with DVR technology incorporated has been determined, I turn to the question of how the DVR affects the network and the viewers. Plugging the equilibrium outcome back into the original profit function yields an equilibrium profit of:

of consumers do watch the show without a DVR, I will assume from here on out that the parameters are such that ${\hat{\gamma}_1}^* > 1$.

$$\begin{aligned}
\pi_D^* &= a_D^*\left[1 - {a_D^*}^s\left(\frac{\alpha^{s+1}}{V-r} + \frac{(1-\alpha)(1-\alpha^s)}{r}\right)\right] \\
&= a_D^*\left[1 - \frac{1}{s+1}\left(\frac{r\alpha^{s+1}}{r\alpha^{s+1} + (1-\alpha)(1-\alpha^s)(V-r)}\right.\right. \\
&\qquad \left.\left. + \frac{(1-\alpha)(1-\alpha^s)(V-r)}{r\alpha^{s+1} + (1-\alpha)(1-\alpha^s)(V-r)}\right)\right] \\
&= a_D^*\left[1 - \frac{1}{s+1}\right] \\
&= \frac{s}{s+1}a_D^*
\end{aligned}$$

Note that, from the no-DVR case:

$$A_N^* = a_N^* \cdot F(\hat{\gamma}^*) = a_N^*\frac{s}{s+1},$$

from which the following proposition follows.

Proposition 3 *The digital video recorder does not affect the average percentage of ads seen by each consumer in equilibrium, regardless of its efficiency of availability.*

Once the DVR comes out, two effects take place to change the equilibrium. First, consumers change their response to any given ad level such that both some people that would have watched TV anyway as well as some that would not have watched previously now do watch television, but with a DVR. Anticipating this change to the response function, the network changes its ad allotment so as to maximize the new profit function. The extra percent of ads aired from those who change from not watching to watching with a DVR balances perfectly with the percent loss from those who change from watching the show anyway to now watching with the DVR so that, on net, the average percent of ads seen by each viewer remains the same.

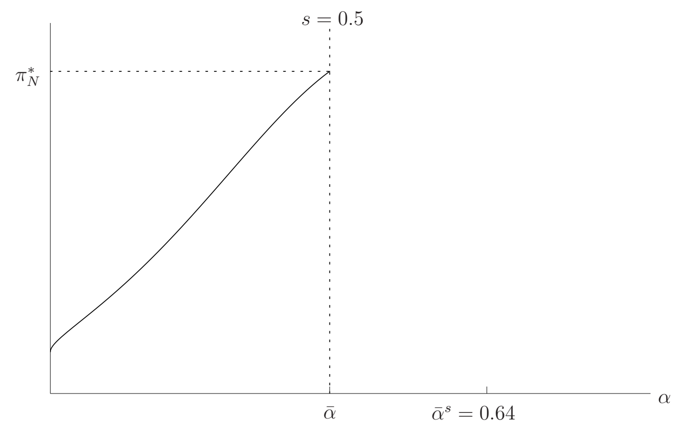
Consequently, profit will move in the same direction as the ad-level does; if the DVR causes the equilibrium ad-level to increase, profits also increase, whereas if the DVR's advent causes the network to drop its ad allotment, its profits will fall as well. One might expect that a device that allows consumers to filter out some of the ads aired would hurt the network, but that is not necessarily the case.

Proposition 4 *With decreasing or constant MNCs, the network's profit is lower with DVR technology than it is without. With increasing MNCs, the network's profit is lower with DVR technology for $\alpha \in (0, \bar{\alpha}^s)$ and higher with DVR technology for $\alpha \in (\bar{\alpha}^s, \bar{\alpha})$.*

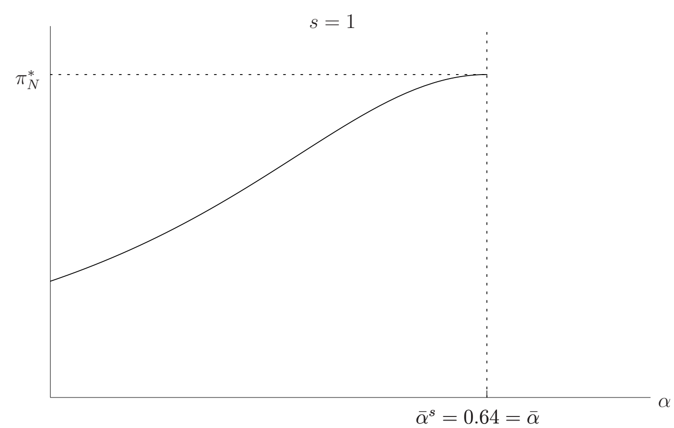
Proof See Section A.2 in the Appendix.

To demonstrate this effect graphically, Figure 3.3 displays an example of π_D^* versus the DVR's pass-through rate for $s = 0.5$, $s = 1$ and $s = 2$.

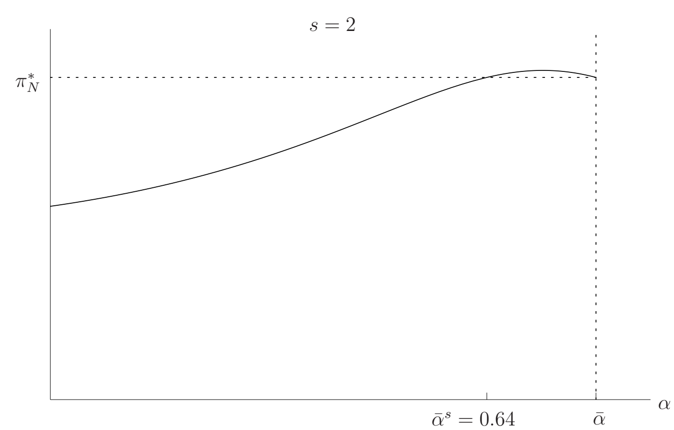
Intuitively, when MNCs are increasing, the DVR will increase profits while it is relatively inefficient for the following reasons. Each person who converts from not watching TV to watching with a DVR adds α to total impressions, whereas each person who converts from watching without a DVR to watching with one subtracts $1-\alpha$ from the total impressions. When α is large, the former effect can outweigh the latter effect, causing overall impressions to rise. However, as s drops, MNCs drop, so that the added cost to the viewer of enduring more commercials is smaller. As a result, the threshold pass-through rate such that a positive mass of consumers finds the DVR worth its rental cost decreases. When MNCs are constant, this threshold coincides with the efficiency level at which DVRs can increase total impressions and therefore profit. When MNCs are negative, the threshold level occurs before DVRs are inefficient enough to increase total impressions and profit. However, when MNCs are increasing, consumers find the last few ads they watch to be the costliest. Since



(a) Decreasing MNCs



(b) Constant MNCs



(c) Increasing MNCs

Figure 3.3: Profit versus the Pass-Through Rate for $r = 9$ and $V = 25$

the DVR eliminates the last fraction of ads from the viewing experience, the efficiency range for which the DVR is still valuable to some positive mass of consumers is larger when the consumers have increasing MNCs. As a result, there is a range of efficiency for which profits increase with the DVR while at least some consumers find the device to be worth the rental cost.

From Proposition 3, the ratio of the network's profit to its ad level in equilibrium is a function only of s. Consider a viewer who watches the show in the no-DVR case and also watches the show unfiltered when DVR technology is available. Whether or not the number of ads this viewer sees increases with DVR technology varies directly with the network's profit, so the same conditions from Proposition 4 apply to this question as well. However, it is not clear whether or not a DVR-user could end up watching more ads than each viewer in the no-DVR case since, even if profit – and therefore the ad level – increases with the DVR, the filtering mechanism of the device could cause these viewers to still watch fewer ads. As the following proposition states, this is indeed the case.

Proposition 5 *All other parameters equal, a DVR user will watch fewer commercials than a viewer would have seen in the no-DVR case.*

Proof See Section A.3 in the Appendix.

Essentially, even in the cases when the DVR is not very efficient and the network increases its ad level with the DVR, users of the device still watch fewer ads than were aired before the advent of the technology.

The DVR also has an effect on the various choices consumers make in equilibrium, which is summarized in the following two propositions.

Proposition 6 *If the DVR is perfectly efficient, its availability does not change the mass of consumers that watch TV without a DVR. Otherwise, its availability lowers non-DVR TV viewers.*

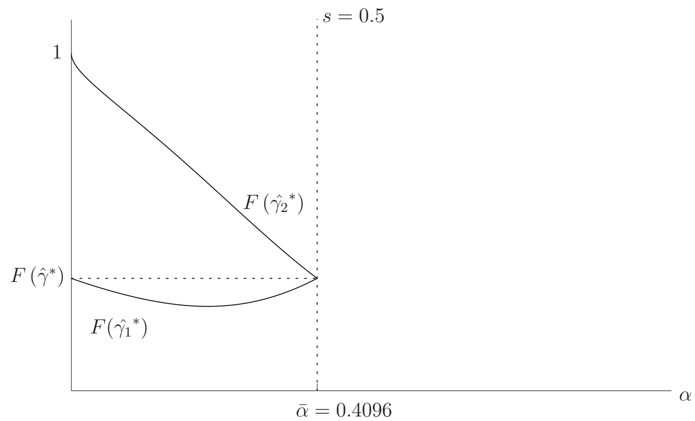
Proof See Section A.4 in the Appendix.

Proposition 7 *Total TV viewership increases with DVR efficiency.*

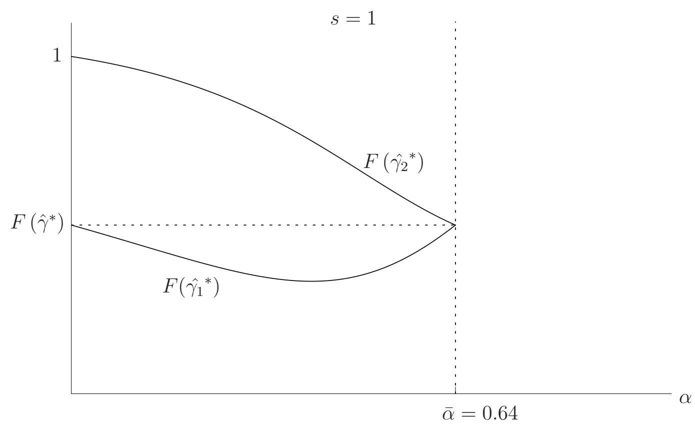
Proof See Section A.5 in the Appendix.

As one might expect, non-DVR television viewership will not increase when the DVR is introduced, regardless of its efficiency. At α's two endpoints, the non-DVR viewership mass is the same as the viewership in the no-DVR case. At the inefficient endpoint, the DVR is just inefficient enough that no consumer finds it worth its cost, and the no-DVR equilibrium prevails since the game is essentially equivalent to the no-DVR case. At the other endpoint, the DVR is perfectly efficient, so consumers can watch the show ad-free for a cost of r. r then essentially replaces V in the game as it then represents each consumer's willingness to "pay" for the show, where the ads are the "price" paid. As shown in section 3.2, TV viewership in the no-DVR case is $\frac{s}{s+1}$ and hence not dependent on V, so it will remain the same when the only change to the game structure is changing V to r.

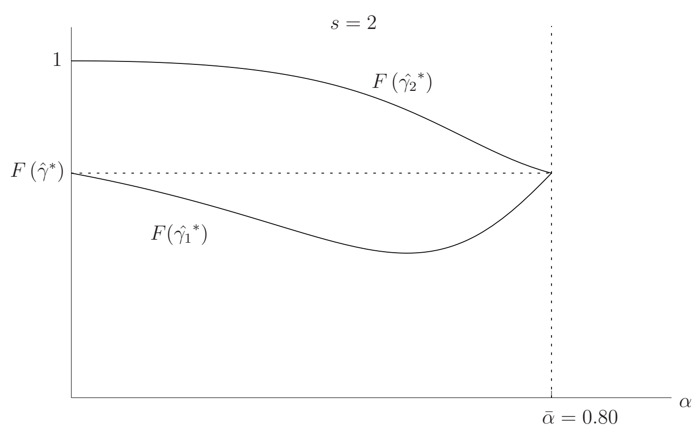
In addition to pulling consumers from the group of those that watch TV unfiltered, the DVR also pulls some viewers from the mass of those that would not watch at all in the absence of the technology. Figure 3.4 displays the unfiltered and total viewership under DVR technology as well as the mass of consumers that choose to watch TV in the no-DVR mode for the case $r = 9$ and $V = 25$.



(a) Decreasing MNCs



(b) Constant MNCs



(c) Increasing MNCs

Figure 3.4: Consumer Choice versus DVR Efficiency for $r = 9$ and $V = 25$

Chapter 4

Duopoly

4.1 Model

The previous chapter examined a network's choice of ad level when its only competition was the outside option. The game was analyzed to determine the network's optimal behavior both with and without DVR technology. In this chapter, the same game is analyzed, except that the network's competition is now another network exactly like itself. Consequently, each network must consider the dynamics of its rival's optimal response as opposied to simply a static outside option of equal value to all consumers.

In this setting, each network provides a different kind of programming. Let θ index programming type, with network 1 located at $\theta = 0$ and network 2 located at $\theta = \bar{\theta}$. The networks' types are fixed; there is no mechanism available for the networks to change their θs. Network j is to air one show of its given type and must choose its ad allotment for that show, a_j. The networks receive a constant payoff each time one viewer watches one ad; as a result, each aims to maximize the total ad impressions that result from airing its show.

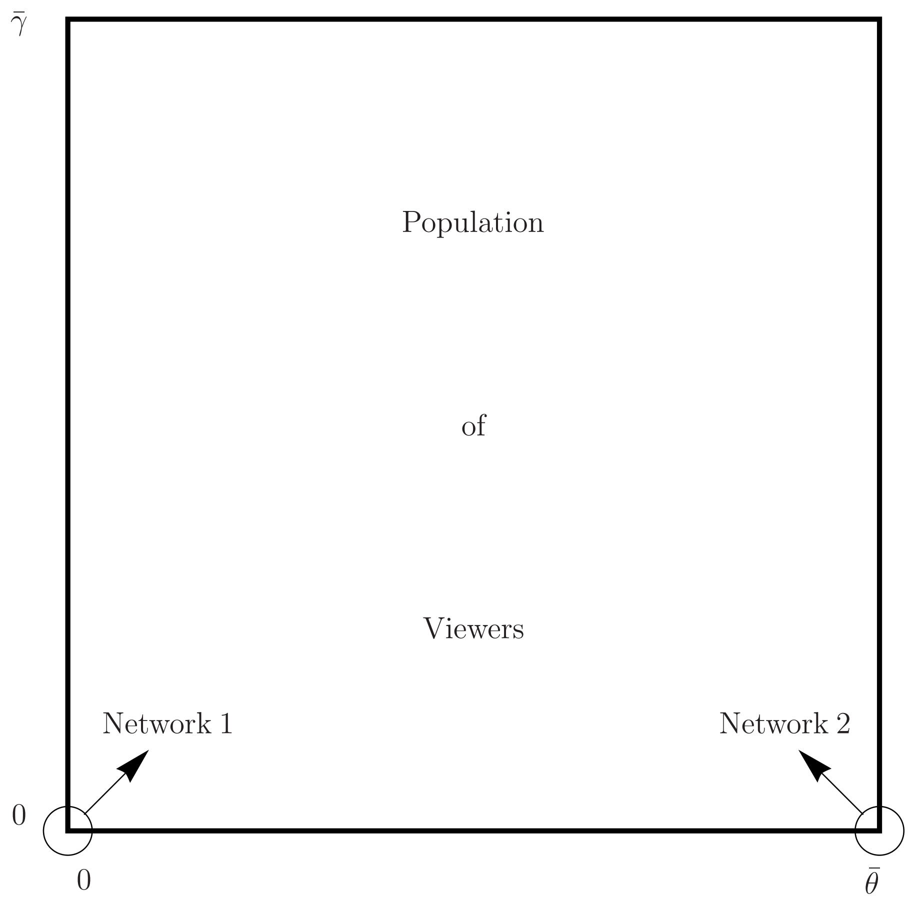


Figure 4.1: Population of Viewers and Network Locations

After ad allotments are announced, viewers decide which show they will watch. There is no outside option - each viewer must choose to either watch network 1 or watch network 2. Viewers are heterogeneous in their taste for programming, along which they are uniformly distributed: $\theta \sim U\left[0, \bar{\theta}\right]$. They vary uniformly along a second dimension of heterogeneity as well, $\gamma \sim U\left[0, \bar{\gamma}\right]$, representing their aversion to ads. As a normalization, I assume $\bar{\gamma} \cdot \bar{\theta} = 1$, so that each network's audience mass and market share are equivalent. Figure 4.1 displays the viewers and networks graphically.

Each viewer's utility is a function of the show she chooses to watch. She receives the same utility simply from the act of watching TV regardless of which network she watches, which I normalize to 0. However, a given viewer i receives disutility from any imperfect match between what she watches and her ideal type - i.e., θ_i if she watches network 1 and $\bar{\theta} - \theta_i$ if she watches network 2 - as well as disutility from the ads she sees in the amount of $\gamma_i a_j$.

Summarily, the utility viewer i receives from watching network j is:

$$U_i^j = -|\theta_i - \theta^a| - \gamma_i a_j$$

where s is the station viewer i chooses to watch and θ^s is the location of station s.

The timing of the game is as follows:

1. The two networks simultaneously choose and announce their ad allotments,
2. Viewers decide whether to watch network 1 or network 2,
3. All parties receive their payoffs.

4.2 Equilibrium

To determine the equilibrium of this game, I start at the final stage and work backwards. Immediately before parties receive their payoffs, viewers decide whether to watch network 1 or network 2 given the ad allotments to which each network has committed. Consider a scenario in which the networks have committed to $a_2 > a_1$, and the viewers must make their choices. Figure 4.2 demonstrates the viewers' optimal response map to this scenario.

The key to the figure is the set of individuals that are indifferent between the two stations. Note that, for an individual to be indifferent, her utility from each must be

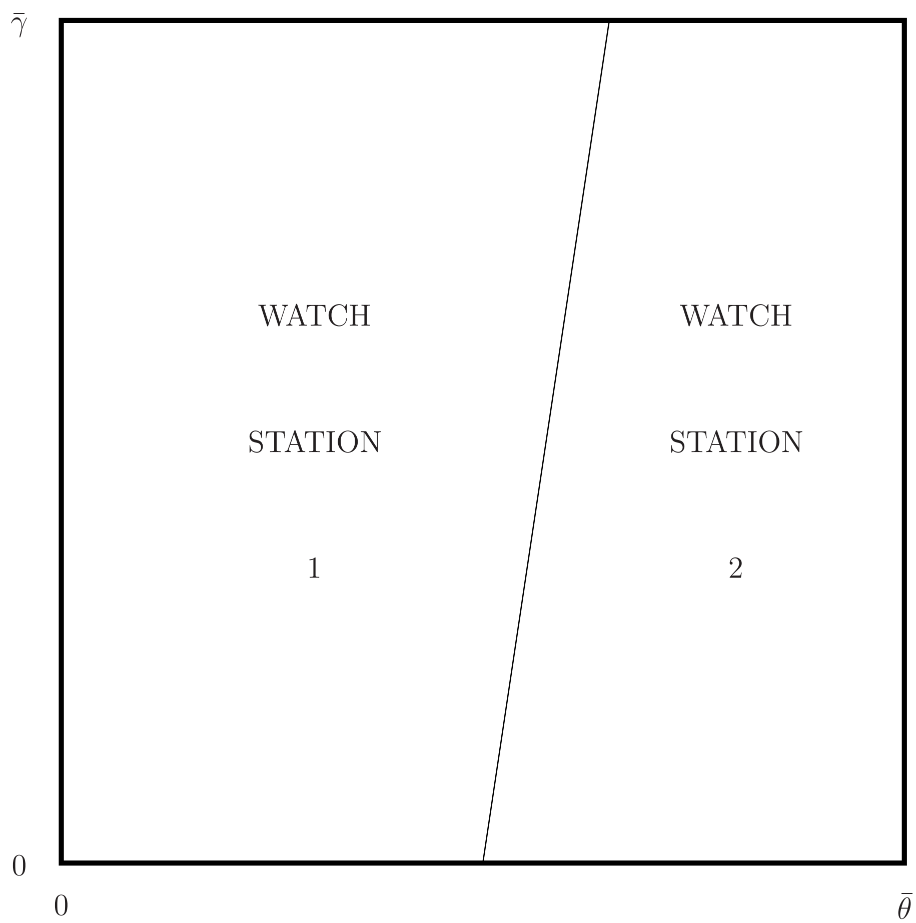


Figure 4.2: Optimal Viewer Response Example - No DVR

equivalent:

$$
\begin{aligned}
U_i^1 &= U_i^2 \Longrightarrow \\
-\theta_i - \gamma_i a_1 &= -\left(\bar{\theta} - \theta_i\right) - \gamma_i a_2 \Longrightarrow \\
\theta_i &= \frac{\bar{\theta}}{2} + \frac{\gamma_i}{2}\left(a_2 - a_1\right)
\end{aligned}
$$

The individual at $\left(\frac{\bar{\theta}}{2}, 0\right)$ will always be indifferent between the two networks since she has no preference for one network's type over the other and does not care about ads. From that point, the indifference line will slope up and to the left if network 1 airs more ads or up and to the right if network 2 does, owing to the tradeoff between how much an individuals cares about the extra ads on one station versus how strong her preference for that station is.

Once the indifference line has been determined, those to the left of the indifference line will watch station 1, and those to the right will watch station 2.

With the viewers' reaction functions determined, the next step is to determine the networks' optimal strategies. Given an action by a network and by its rival, the networks know what the viewers' response will be, and therefore what their resulting profits will be. Consequently, they can determine their optimal responses to their rivals' actions exactly.

Proposition 8 *Without DVR technology, each networks reaction function is:*

$$a_j^* = \begin{cases} \frac{a_k}{2} + \frac{\bar{\theta}}{\bar{\gamma}},\ 0 < a_k < 4\frac{\bar{\theta}}{\bar{\gamma}} \\ a_k - \frac{1}{2}\sqrt{\frac{\bar{\theta}}{\bar{\gamma}} a_k},\ a_k > 4\frac{\bar{\theta}}{\bar{\gamma}} \end{cases}$$

Proof See Appendix Section B.1.

The exact nature of the reaction function depends on whether the indifference line intersects the side of the box or its top. However, one item to note with the reaction function is that it is never optimal to flood the market with ads, whereas it is optimal to severely undercut your rival if he is flooding the market with ads. In other words, the optimal response for network 1 sometimes creates an indifference line that intersects the right side of the box so that network 1 receives the vast majority of the market share – this occurs only if network 2 has set a very high ad allotment. Conversely, the optimal response for network 1 is never such that the indifference line intersects the left side of the box, so that it has set a very high ad allotment itself while sacrificing the majority of the market share to its rival.

Overall, the reaction function admits but one equilibrium point.

Proposition 9 *The unique equilibrium in pure strategies involves both networks setting their ad levels at* $a_i = 2\frac{\bar{\theta}}{\bar{\gamma}}$.

Proof Note that $a_j = a_k = 2\frac{\bar{\theta}}{\bar{\gamma}}$ satisfies the reaction function for both networks, since

$$2\frac{\bar{\theta}}{\bar{\gamma}} = \frac{2\frac{\bar{\theta}}{\bar{\gamma}}}{2} + \frac{\bar{\theta}}{\bar{\gamma}}$$

Further, the reaction function has a slope of

$$\frac{da_j^*}{a_k} = \begin{cases} \frac{1}{2}, \ 0 < a_k < 4\frac{\bar{\theta}}{\bar{\gamma}} \\ 1 - \frac{1}{4}\sqrt{\frac{\bar{\theta}}{\bar{\gamma}}\frac{1}{a_k}}, \ a_k > 4\frac{\bar{\theta}}{\bar{\gamma}} \end{cases}$$

which is always less than 1. Therefore, there can only be a single crossing of the two reaction functions, completing the proof. ∎

Since the firms air the same level of ads, they split the market, resulting in an equilibrium profit of

$$\pi^* = \frac{\bar{\theta}}{\bar{\gamma}}$$

A couple of intuitive comparative static results also follow.

Proposition 10 *As the diversity of programming increases, equlibrium ad level increases, whereas as the maximium aversion to ads increases, equilibrium ad level decreases.*

As one might expect, as the diversity of programming increases with all other factors remaining constant, each network has more market power due to the increasingly niche market that it fills. Consequently, a greater number of the viewers are more partial to a given network relative to their aversion for ads, allowing the networks to "charge" more for their shows (i.e., air more ads in their programs). Secondly, as the maximum level of ad aversion increases, the same mass of viewers is now stretched out over an expanse that includes even more ad-averse individuals, which are the individuals that are the most sensitive to the "price" that the networks are charging. Each network must therefore be more cautious with the number of ads it airs, lowering the ad allotment each airs in equilibrium.

4.3 The Digital Video Recorder

With the baseline model solved, I move on to the version of the game with a DVR. With its advent, users get the additional option of renting a DVR at a rental cost r, through which they can filter out $(1-\alpha)$ of the ads of the show they choose to watch. As in Chapter 3, I often refer to α as the pass-through rate and to $(1-\alpha)$ as the DVR's efficiency or the filter rate. I assume that $\bar{\theta} > \frac{r}{1-\alpha}$, i.e., that the price-efficiency combination of the DVR has at least a certain threshold attractiveness level relative to the other parameters.

The viewers now have four potential options:

$$\begin{aligned}
U_i^{1,N} &= -\theta_i - \gamma_i a_1, \\
U_i^{2,N} &= -(\bar{\theta} - \theta_i) - \gamma_i a_2, \\
U_i^{1,D} &= -\theta_i - \alpha\gamma_i a_1 - r, \text{ or} \\
U_i^{2,D} &= -(\bar{\theta} - \theta_i) - \alpha\gamma_i a_2 - r
\end{aligned}$$

where $U_i^{j,d}$ represents the utility received from watching station a either with a DVR $(d = D)$ or not $(d = N)$.

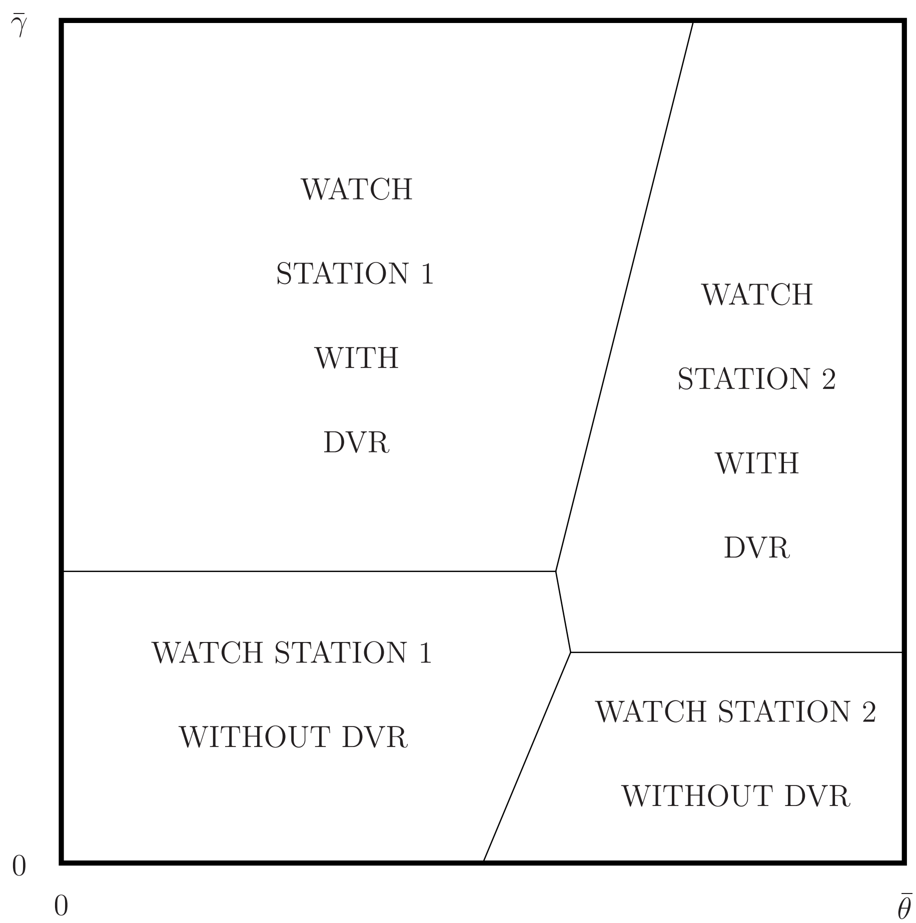


Figure 4.3: Viewer Optimal Response Map - DVR Available

To solve this game, I again begin at the last stage and work backwards. After each network has committed to its ad level, viewers have to choose among their four options. Figure 4.3 displays one example of the viewers' optimal response in a case where $a_2 > a_1$. The four options the viewers face create six indifference lines. Two run horizontally across the box, each representing those consumers that are indifferent between watching a given television station with a DVR and watching it without one. Two more originate from the point $\left(0, \frac{\bar{\theta}}{2}\right)$, the individual who is indifferent between both stations regardless of ad levels. One represents those indifferent between watching each station without a DVR, and the second represents those indifferent between watching each station with one. The former has a steeper slope than the latter, owing to the DVR's ability to filter out some of the ads, making DVR users less sensitive to the ad differential between the stations.

Lastly, there are two more indifference lines, each representing those that are indifferent between watching a particular station with a DVR and watching its rival without one. Only one is relevant given specific ad levels – the one representing those indifferent between watching the higher-ad station's show with a DVR and the lower-ad station's show without one.

These indifference lines section the box into the groups of consumers that perform each choice. Given these responses, each network knows their eventual profit as a function of its ad level and that of its rival station, and can therefore determine their optimal strategies accordingly.

Proposition 11 *In the ad game with the DVR available, the strategy*

$$a_1 = a_2 = a^* = \frac{1}{\alpha\bar{\gamma}}\left[\bar{\theta} + \sqrt{\bar{\theta}^2 - \frac{1+\alpha}{1-\alpha}r^2}\right]$$

constitutes an equilibrium.

Proof See Section B.2 in the Appendix.

While the equilibrium ad level after the DVR may seem very different than that before the DVR, they are actually quite similar. Re-arranging slightly gives:

$$\alpha \cdot a^* = \frac{\bar{\theta}}{\bar{\gamma}} + \sqrt{\left(\frac{\bar{\theta}}{\bar{\gamma}}\right)^2 - \frac{1+\alpha}{1-\alpha}\left(\frac{r}{\bar{\gamma}}\right)^2}$$

which represents the level of ads that viewers with a DVR see. Consider how this changes as α goes to 1, the scenario in which the DVR lets through all of the ads.

Holding $\bar{\theta}$ constant, we must also have r going to 0 due to the assumption $\bar{\theta} > \frac{r}{1-\alpha}$. This implies $\frac{r^2}{1-\alpha}$ must also go to 0. Consequently, the left-hand side simply represents the ad-level, while the right-hand side goes to $2\frac{\bar{\theta}}{\bar{\gamma}}$, so that the limiting case yields the same result we saw in the case without DVR technology.

4.4 Welfare

In this section I examine the welfare effects of the digital video recorder. I start with network welfare. One might expect that giving viewers the additional option of renting a device that allows them to filter out some percentage of the ads would hurt the networks, or at best be ambiguous as was the case when the network competed against only the outside option for viewers with increasing marginal nuisance costs. However, as summarized in the following proposition, this is not the case.

Proposition 12 *Each network's profit is increased with the advent of the DVR.*

Proof See Appendix Section B.3

As in the earlier case, the networks split the market. However, now a portion of their viewers see all of their ads, whereas those that dislike ads the most will rent a DVR and only see a fraction of them. In a sense, the DVR allows the networks to price discriminate, only in a very strict fashion. After the DVR arrives, when it sets its ad level a_j, it essentially charges a price a_j to those viewers that are less ad-averse, whereas more ad-averse viewers face a price $\alpha a_j + r$, of which the network receives αa_j.

The DVR also affects the strategic incentives for each network as it relates to the competitive aspects of the game. Previously, each network was very sensitive to the most ad-averse viewers, since a small change in ad levels would cause a large mass of these viewers to switch to the other network. However, after the advent of the DVR, this affect is dampened by a factor α, and consequently each network does not need to be as sensitive to these viewers since they now have the option of tuning down their ad-aversion for a rental rate r. On the whole, these two factors serve to give the networks more market power in a sense and hence increase their profits.

This of course implies that the average number of ads seen by each viewer increases from before the DVR to after. One might think that this implies viewer welfare has decreased, but this is not necessarily the case either.

Proposition 13 *The DVR's effect on viewer welfare is ambiguous.*

Proof See Appendix Section B.4

This might seem counter-intuitive since more ads are being watched. However, what matters is not just how many ads are seen, but also *who* is watching the ads. Before the DVR, all viewers watch the same number of ads, so that the ad burden falls on those most ad-averse, as would be expected. However, after the DVR arrives, more ads are seen by those less ad-averse, as the most ad-averse individuals get a DVR and therefore only watch a portion of the ads. Consequently, an interesting externality result arises – the advent of the DVR shifts the ad burden from those that are very ad-averse to those just under the ad-aversion threshold at which viewers rent a DVR. This is formalized in the following proposition.

Proposition 14 *Each DVR user sees fewer ads than each viewer does when there is no DVR, whereas non-DVR users in the DVR case see more ads than each viewer does when there is no DVR.*

Proof Each DVR users ad-exposure is αa_N^*, which is larger than a_D^* since:

$$\begin{aligned} 2\frac{\bar{\theta}}{\bar{\gamma}} &> \frac{\bar{\theta}}{\bar{\gamma}} + \sqrt{\left(\frac{\bar{\theta}}{\bar{\gamma}}\right)^2 - \frac{1+\alpha}{1-\alpha}\left(\frac{r}{\bar{\gamma}}\right)^2} \\ \Longrightarrow \alpha a_N^* &> a_D^* \end{aligned}$$

Further, from Proposition 12, network profits increase with the DVR, which implies that more ads are seen overall, or, equivalently, that the average viewer sees more ads after the advent of the DVR. Since DVR users actually see fewer ads, it must be the case that non-DVR users see more ads to bring the average above that of the non-DVR case. ∎

Additionally, the expected comparative statics hold.

Proposition 15 *The following comparative statics hold:*

1. $\frac{da_D^*}{d\bar{\theta}} > 0$, $\frac{d\pi_D^*}{d\bar{\theta}} > 0$
2. $\frac{da_D^*}{d\bar{\gamma}} < 0$, $\frac{d\pi_D^*}{d\bar{\gamma}} < 0$
3. $\frac{da_D^*}{d\alpha} < 0$, $\frac{d\pi_D^*}{d\alpha} < 0$
4. $\frac{da_D^*}{dr} < 0$

Proof See Appendix Section B.5.

As the networks are spaced farther apart along the taste dimension, they enjoy more market power over their niche of programming and therefore can air a higher number of ads as well as increase their profits. However, as the ad-aversion of the most ad-averse consumers increases, the networks have to be even more sensitive with thenumber of ads they air, decreasing both their optimal ad allotment and their profit.

The same holds for the proportion of ads that pass through the DVR, since as that increases, the ability of the network's to take advantage of the element of the price discrimination goes down, and the sensitivity of the most ad-averse people to the ad differential goes up. Lastly, as the DVR becomes more costly, it becomes less attractive and therefore the most ad-averse people are less likely to rent it, increasing the weighted-average sensitivity of the viewer population to ads. Essentially, since the DVR helps the networks, making it more attractive to the viewers – i.e., decreasing the rental rate or the number of ads that pass through it – allows each network to air more ads and increase its profit.

Chapter 5

Conclusion

I have presented two novel models for the television industry that incorporate the DVR's ad-skipping capabilities within. In the first model, a monopolist network and a spectrum of consumers play a strategic game in which the network must choose an ad allocation, after which the consumers must decide whether or not to watch TV. When DVR technology is available, consumers have the added option of renting a DVR in order to eliminate a percentage of the ads, dubbed the DVR's efficiency. I have analyzed the game both in the absence and in the presence of this technology and then compared the outcomes. I have also examined how these results change as the DVR's efficiency changes and as the consumers' MNCs (marginal nuisance costs) vary from decreasing to constant to increasing.

I have found that some aspects of the equilibrium outcome do not vary with the DVR's efficiency or with MNCs. For example, the DVR increases total viewership, but it will not increase the non-DVR viewing audience, regardless of its efficiency. In addition, I have found that total impressions per ad run stays constant with the advent of the DVR, regardless of its efficiency. Given constant returns from advertising for the network, this implies that the network's profit moves with its ad allotment in equilibrium.

However, the DVR's efficiency and MNCs do play a role in other aspects of the equilibrium outcome. The ad allotment can increase with the DVR, implying the DVR need not be a detriment to the network as it may increase the network's profit. If MNCs are increasing, the network can actually be made better off depending on the DVR's efficiency. If the DVR is inefficient enough, the viewers that switch from not watching to watching with a DVR will outweigh those that would have watched anyway but are now using a DVR so that total advertising impressions increase. In these cases, there will always be an interior maximum for the network's profit as a function of the DVR's efficiency. However, if MNCs are decreasing or constant, avoiding additional ads is not worth as much to the viewers and hence the efficiency level at which profit would increase is then too inefficient for anyone to find the DVR worth its rental cost.

I then expanded the constant marginal nuisance cost model to the realm of a duopoly, i.e., when the network's competition is another network rather than the static outside option. In this setting, the DVR benefit result is strengthened to the point where the networks' profits are unambiguously helped by the DVR. This implies that the average viewer watches more ads after the DVR comes out than before, which suggests that viewer welfare is harmed by the DVR. However, viewers may also be helped in total by the device, since the filtering mechanism of the DVR implies that those that are hurt the most by ads can pay a fixed rental cost to remove what to them is a large nuisance cost, shifting the burden of the ads from these individuals to those that do not mind ads as much. This externality applied from the most ad-averse individuals to those that are less ad-averse could act in such a way that viewer welfare improves as well.

Appendix A

Ad-Skipping / Monopoly

The following appendix contains the proofs of the results from Chapter 1.

A.1 Analogy to Production Differentiation Models

The model in this paper borrows aspects from the product differentiation literature, such as Mussa and Rosen (1978). In that model, when consumer i purchases product j, she receives a utility (assuming no unspent income) of

$$U_i = y + \theta_i \cdot q_j - P(q_j)$$

where y is the consumer's income (exogenous), q_j the quality of product j, $P(q_j)$ the price of a product of quality j, and $\theta_i > 0$ a heterogeneous taste parameter. As an example, (Mussa and Rosen, 1978, pp.312) assume $F(\theta) = \frac{\theta - \underline{\theta}}{\bar{\theta} - \underline{\theta}}$ (i.e., uniform) for

$\theta \in \left[\underline{\theta}, \overline{\theta}\right]$, $\underline{\theta} > 0$. Note that this model is equivalent to:

$$\tilde{U}_i = y + q_j - \tilde{\theta}_i \cdot P(q_j),\ F(\tilde{\theta}) = \frac{\bar{\theta} - \frac{1}{\tilde{\theta}}}{\bar{\theta} - \underline{\theta}} \text{ for } \tilde{\theta} \in \left[\frac{1}{\bar{\theta}}, \frac{1}{\underline{\theta}}\right].$$

Proof Consider $\tilde{U}_i = \frac{U_i}{\theta_i} + \left(1 - \frac{1}{\theta_i}\right) y$. Since this is a monotonic transformation of U_i, it represents the same preference relation. Simplifying, we get

$$\tilde{U}_i = y - \frac{1}{\theta_i} \cdot P(q_j) + q_j,\ \frac{1}{\theta_i} \in \left[\frac{1}{\bar{\theta}}, \frac{1}{\underline{\theta}}\right].$$

Define $\tilde{\theta}_i \equiv \frac{1}{\theta_i}$. The derivation of the distribution function for $\tilde{\theta}_i$, $\tilde{F}(\cdot)$, is all that is left:

$$\begin{aligned}
F(x) = \frac{x - \underline{\theta}}{\bar{\theta} - \underline{\theta}} \Longrightarrow \Pr\left(\theta < \frac{1}{x}\right) &= \frac{\frac{1}{x} - \underline{\theta}}{\bar{\theta} - \underline{\theta}} = \Pr\left(x < \frac{1}{\theta}\right) \text{ since } \theta > 0, \\
\Longrightarrow \Pr\left(\frac{1}{\theta} < x\right) &= 1 - \frac{\frac{1}{x} - \underline{\theta}}{\bar{\theta} - \underline{\theta}} \\
&= \frac{\bar{\theta} - \frac{1}{x}}{\bar{\theta} - \underline{\theta}} \\
&= \tilde{F}(x) \quad \blacksquare
\end{aligned}$$

The intuition is as follows: what drives each consumer's decision is the relative tradeoff between price and quality, so a consumer facing the utility function $U = y - P(q_j) + \theta \cdot q_j$ will make the same decisions as one facing $U = y - \frac{1}{\theta}P(q_j) + q_j$. Since θ varies across consumers, one only needs to adjust the distribution function to make sure that, facing the same price schedule in each case, the same proportion of consumers choose each product.

Following the same logic,

$$U_i = \theta_i \cdot V - a^s,\ \theta \sim U[0, 1]$$

and

$$U_i = V - \gamma_i \cdot a^s,\ F(\gamma) = 1 - \frac{1}{\gamma} \Longrightarrow f(\gamma) = \gamma^{-2} \text{ for } \gamma \in [1, \infty).$$

are equivalent. For DVR users, the analogy is between

$$U_i = \theta_i \cdot (V - r) - a^s,\ \theta \sim U[0, 1]$$

and

$$U_i = V - r - \gamma_i \cdot a^s,\ F(\gamma) = 1 - \frac{1}{\gamma} \Longrightarrow f(\gamma) = \gamma^{-2} \text{ for } \gamma \in [1, \infty).$$

A.2 DVR Effect on Profit

Proposition 4 *With decreasing or constant MNCs, the network's profit is lower with DVR technology than it is without. With increasing MNCs, the network's profit is lower with DVR technology for $\alpha \in (0, \bar{\alpha}^s)$ and higher with DVR technology for $\alpha \in (\bar{\alpha}^s, \bar{\alpha})$.*

Proof Recall that $\bar{\alpha} = \left(\frac{V-r}{V}\right)^{\frac{1}{s}} \Longrightarrow \bar{\alpha}^s = \frac{V-r}{V}$. Note that:

$$\begin{aligned}
\left(\alpha - \frac{V-r}{V}\right)\left(\alpha^s - \frac{V-r}{V}\right) &= \alpha^{s+1} - \frac{V-r}{V}\alpha^s - \frac{V-r}{V}\alpha + \left(\frac{V-r}{V}\right)^2 \\
&= \frac{1}{V^2}\left(V^2\alpha^{s+1} - \alpha^s V(V-r) - \alpha V(V-r) + (V-r)^2\right) \\
&= \frac{1}{V^2}\left[V\left(r\alpha^{s+1} + (1-\alpha)(1-\alpha^s)(V-r)\right) - r(V-r)\right],
\end{aligned}$$

which has the same sign as

$$\begin{aligned}
&= \left(\frac{V}{s+1} - \frac{r(V-r)}{(s+1)r\alpha^{s+1} + (s+1)(1-\alpha)(1-\alpha^s)(V-r)}\right) \\
&= \pi_N^* - \pi_D^*
\end{aligned}$$

Hence, the sign of $\left(\alpha - \frac{V-r}{V}\right)\left(\alpha^s - \frac{V-r}{V}\right)$ determines the sign of $\pi_N^* - \pi_D^*$. If $\alpha = \frac{V-r}{V}$, DVR technology does not change equilibrium profit. Otherwise, if

$$\min\left[\frac{V-r}{V}, \left(\frac{V-r}{V}\right)^{\frac{1}{s}}\right] < \alpha < \max\left[\frac{V-r}{V}, \left(\frac{V-r}{V}\right)^{\frac{1}{s}}\right]$$

profit increases with DVR technology; if

$$\alpha < \min\left[\frac{V-r}{V}, \left(\frac{V-r}{V}\right)^{\frac{1}{s}}\right] \text{ or } \alpha > \max\left[\frac{V-r}{V}, \left(\frac{V-r}{V}\right)^{\frac{1}{s}}\right]$$

it decreases. When $s \leq 1$, $\left(\frac{V-r}{V}\right)^{\frac{1}{s}} \leq \frac{V-r}{V}$, so any value of α that satisfies $\alpha < \bar{\alpha}$ also satisfies $\alpha < \min\left[\frac{V-r}{V}, \left(\frac{V-r}{V}\right)^{\frac{1}{s}}\right]$, implying that profit necessarily decreases. When $s > 1$, $\left(\frac{V-r}{V}\right)^{\frac{1}{s}} > \frac{V-r}{V} \Longrightarrow \min\left[\frac{V-r}{V}, \left(\frac{V-r}{V}\right)^{\frac{1}{s}}\right] = \frac{V-r}{V}$, $\max\left[\frac{V-r}{V}, \left(\frac{V-r}{V}\right)^{\frac{1}{s}}\right] = \left(\frac{V-r}{V}\right)^{\frac{1}{s}}$, so the comparison between π_D^* and π_N^* will depend on α: for $\alpha \leq \frac{V-r}{V}$, $\pi_D^* < \pi_N^*$, and for $\alpha \in \left(\frac{V-r}{V}, \left(\frac{V-r}{V}\right)^{\frac{1}{s}}\right)$, $\pi_D^* > \pi_N^*$.

To prove the rest of the proposition, from the equilibrium value of profit:

$$\begin{aligned}
\pi_D^* &= \frac{s}{s+1} \cdot a^* \Longrightarrow \\
\frac{d\pi_D^*}{d\alpha} &= -\frac{s+1^s}{sp} \cdot \frac{(\pi_D^*)^{s+1}}{sr(V-r)} \cdot \left((s+1)^2 r\alpha^2 + (s+1)(V-r)\left((s+1)\alpha^s - s\alpha^{s-1} - 1\right)\right) \\
&\overset{s}{=} \left[(s+1)(V-r)\left(1 + s\alpha^{s-1}\right) - (s+1)^2 V\alpha^2\right]
\end{aligned}$$

where $\overset{s}{=}$ represents sign-equivalence. Therefore, one only need find the sign of the last expression to determine the sign of $\frac{d\pi_D^*}{d\alpha}$. First note that:

$$\begin{aligned} h(\alpha) &\equiv \left[(s+1)(V-r)\left(1+s\alpha^{s-1}\right)-(s+1)^2 V\alpha^2\right] \Longrightarrow \qquad \text{(A.1)} \\ \frac{d(h(\alpha))}{dV} &= (s+1)+s(s+1)\alpha^{s-1}-(s+1)^2\alpha^s \\ &= (s+1)\left(1+s\alpha^{s-1}-(s+1)\alpha^s\right) \\ &= (s+1)\left(s\alpha^{s-1}(1-\alpha)+(1-\alpha^s)\right) > 0 \end{aligned}$$

Hence, showing that $h(\alpha) \geq 0$ at a particular value $V = \hat{V}$ implies that $h(\alpha) > 0$ for all $V > \hat{V}$. So, for $s \leq 1$, showing that $h(\alpha) \geq 0$ for $V = \frac{r}{1-\alpha^s}$ is enough to show that $h(\alpha) > 0$ for $V > \frac{r}{1-\alpha^s} \Longrightarrow \alpha < \left(\frac{V-r}{V}\right)^{\frac{1}{s}} = \bar{\alpha}$, and, for $s > 1$, showing that $h(\alpha) \geq 0$ for $V = \frac{r}{1-\alpha}$ is enough to show that $h(\alpha) > 0$ for $V > \frac{r}{1-\alpha} \Longrightarrow \alpha < \frac{V-r}{V}$. From (A.1) above,

$$h(\alpha) = (s+1)^2 V\left[\frac{V-r}{(s+1)V} + \frac{s(V-r)}{(s+1)V}\alpha^{s-1} - \alpha^s\right]$$

So, for $s \leq 1$:

$$\begin{aligned} V &= \frac{r}{1-\alpha^s} \Longrightarrow \\ h(\alpha) &= (s+1)^2\frac{r}{1-\alpha^s}\left[\frac{1}{s+1}\alpha^s + \frac{s}{s+1}\alpha^{2s-1} - \alpha^s\right] \\ &= (s+1)^2\frac{r\alpha^s}{1-\alpha^s}\left[\frac{s}{s+1}\left(\alpha^{s-1}-1\right)\right] > 0 \end{aligned}$$

$$\Longrightarrow \frac{d\pi_D^*}{d\alpha} > 0 \text{ for } \alpha < \bar{\alpha}$$

For $s > 1$:

$$\begin{aligned}
V &= \frac{r}{1-\alpha} \Longrightarrow \\
h(\alpha) &= (s+1)^2 \frac{r}{1-\alpha} \left[\frac{1}{s+1}\alpha + \frac{s}{s+1}\alpha^s - \alpha^s \right] \\
&= (s+1)^2 \frac{r\alpha^s}{1-\alpha} \left[\frac{1}{s+1} \left(1 - \alpha^{s-1}\right) \right] > 0, \Longrightarrow \frac{d\pi_D^*}{d\alpha} > 0 \text{ for } \alpha < \bar{\alpha}^s
\end{aligned}$$

Lastly, for $s > 1$, it was shown in section A.2 that $\pi_D^* = \pi_N^*$ at both $\alpha = \frac{V-r}{V}$ and $\alpha = \left(\frac{V-r}{V}\right)^{\frac{1}{s}}$. Since π_D^* is strictly increasing for $\alpha < \frac{V-r}{V}$, by continuity of π_D^*, profit must reach a maximum for some $\alpha \in \left(\frac{V-r}{V}, \left(\frac{V-r}{V}\right)^{\frac{1}{s}}\right)$. ∎

A.3 DVR Effect on Ad Impressions

Proposition 5 *All other parameters equal, a DVR user will watch fewer commercials than a viewer would have seen in the no-DVR case.*

Proof Since $\alpha < \left(\frac{V-r}{V}\right)^{\frac{1}{s}}$:

$$
\begin{aligned}
\alpha &< \left(\frac{V-r}{V}\right)^{\frac{1}{s}} \\
\Longrightarrow \alpha^s &< \frac{V-r}{V} \\
\Longrightarrow \frac{r}{V} &< (1-\alpha^s) \\
\Longrightarrow r &< V(1-\alpha^s) \\
\Longrightarrow (1-\alpha)r &< (1-\alpha)(1-\alpha^s)V \\
\Longrightarrow \alpha^s r &< (1-\alpha)(1-\alpha^s)V - (1-\alpha-\alpha^s)r \\
\Longrightarrow 1 &> \frac{\alpha^s r}{(1-\alpha)(1-\alpha^s)V - (1-\alpha-\alpha^s)r} \\
\Longrightarrow V &> \alpha^s \frac{r(V-r)}{(1-\alpha)(1-\alpha^s)V - (1-\alpha-\alpha^s)r} \\
\Longrightarrow \left(\frac{V-r}{V}\right)^{\frac{1}{s}} &> \alpha\left(\frac{r(V-r)}{(s+1)r\alpha^{s+1} + (s+1)(1-\alpha)(1-\alpha^s)(V-r)}\right)^{\frac{1}{s}} \\
\Longrightarrow \alpha a_D^* &< a_N^* \quad \blacksquare
\end{aligned}
$$

A.4 Relation Between Cases With and Without DVR

Proposition 6 *If the DVR is perfectly efficient, its availability does not change the mass of consumers that watch TV without a DVR. Otherwise, its availability lowers non-DVR TV viewers.*

Proof From the results in Section 3.2:

$$\begin{aligned}
\hat{\gamma_1}^* - \hat{\gamma}^* &= (s+1)\frac{r\alpha^{s+1} + (1-\alpha)(1-\alpha^s)(V-r)}{(1-\alpha^s)(V-r)} - (s+1) \\
&= \frac{s+1}{(1-\alpha^s)(V-r)}\left(r\alpha^{s+1} - \alpha(1-\alpha^s)(V-r)\right) \\
&= \frac{s+1}{(1-\alpha^s)(V-r)} \cdot \alpha V \cdot \left(\alpha^s - \frac{(V-r)}{V}\right)
\end{aligned}$$

When $\alpha = 0$, the expression equals 0, so $\hat{\gamma_1}^* = \hat{\gamma}^*$. Otherwise, the term inside the parentheses determines the sign; from the assumption $\alpha < \left(\frac{V-r}{V}\right)^{\frac{1}{s}}$, that term must be negative. Hence, for $\alpha > 0$, the expression is negative, and $\hat{\gamma_1}^* < \hat{\gamma}^*$. Since $F(\gamma) = 1 - \frac{1}{\gamma}$ is strictly increasing, the result follows. ∎

A.5 DVR Effect on Viewership Mass

Proposition 7 *Total TV viewership increases with DVR efficiency.*

Proof From the equilibrium cutoff value that separates those who watch TV with a DVR from those who do not watch TV at all:

$$\begin{aligned}
\hat{\gamma_2}^* &= \frac{(s+1)r\alpha^{s+1} + (s+1)(1-\alpha)(1-\alpha^s)(V-r)}{ra^s} \\
&= (s+1)\frac{V\alpha^{s+1} + (V-r)(1-\alpha-\alpha^s)}{ra^s} \\
\Longrightarrow \frac{F(\hat{\gamma_2}^*)}{d\alpha} &= 1 - \frac{1}{s+1}\frac{r\alpha^s}{[V\alpha^{s+1} + (V-r)(1-\alpha-\alpha^s)]} \\
\Longrightarrow \frac{dF(\hat{\gamma_2}^*)}{d\alpha} &= \frac{-1}{s+1}\left(\frac{sr\alpha^{s-1}}{[V\alpha^{s+1} + (V-r)(1-\alpha-\alpha^s)]}\right. \\
&\qquad \left. - \frac{r\alpha^s \cdot ((s+1)V\alpha^s - (V-r)(1+s\alpha^{s-1}))}{[V\alpha^{s+1} + (V-r)(1-\alpha-\alpha^s)]^2}\right) \\
&= \frac{r\alpha^{s-1}}{s+1}\left(\frac{V\alpha^{s+1} + (s-1)(V-r)\alpha - s(V-r)}{[V\alpha^{s+1} + (V-r)(1-\alpha-\alpha^s)]^2}\right)
\end{aligned}$$

The sign depends solely on the fraction's numerator since all the other terms are positive. To determine the sign, first note that, since $s, r > 0$ and $0 < \alpha < 1$:

$$\begin{aligned} 0 &> sr\alpha^s(\alpha - 1) \\ \Longrightarrow 0 &> r\alpha^{s+1} + r\alpha^s((s-1)\alpha - s) \\ \Longrightarrow 0 &> \frac{r}{1-\alpha^s}\alpha^{s+1} + \left(\frac{r}{1-\alpha^s} - r\right)((s-1)\alpha - s) \end{aligned}$$

so $g(\alpha) < 0$ when evaluated at $V = \frac{r}{1-\alpha^s}$. Further, $\alpha < 1$ implies that

$$\frac{dg(\alpha)}{dV} = \alpha^{s+1} + (s-1)\alpha - s = \alpha(\alpha^s - 1) + s(\alpha - 1) < 0,$$

so that $g(\alpha) < 0$ for any $V < \frac{r}{1-\alpha^s} \Longleftrightarrow \alpha < \left(\frac{V-r}{V}\right)^{\frac{1}{s}}$ as well. Hence, $\frac{dF(\hat{\gamma_2}^*)}{d\alpha} < 0$ for any $\alpha \in \left(0, \left(\frac{V-r}{V}\right)^{\frac{1}{s}}\right)$. ∎

Appendix B

Ad-Skipping / Duopoly

The following appendix contains the proofs of the results from Chapter 2.

B.1 Network Reaction Functions - No-DVR Case

Proposition 8 *Without DVR technology, each networks reaction function is:*

$$a_j^* = \begin{cases} \frac{a_k}{2} + \frac{\bar{\theta}}{\bar{\gamma}},\ 0 < a_k < 4\frac{\bar{\theta}}{\bar{\gamma}} \\ a_k - \frac{1}{2}\sqrt{\frac{\bar{\theta}}{\bar{\gamma}} a_k},\ a_k > 4\frac{\bar{\theta}}{\bar{\gamma}} \end{cases}$$

Proof As a result of the game's symmetry, each network's payoffs will be identical functions of the other's ad allotment; consequently, without loss of generality I focus solely on network 1's payoffs here.

In order to determine network 1's profits given an ad allotment for each network, one must first determine the viewer's optimal response to a given ad allotment. Note

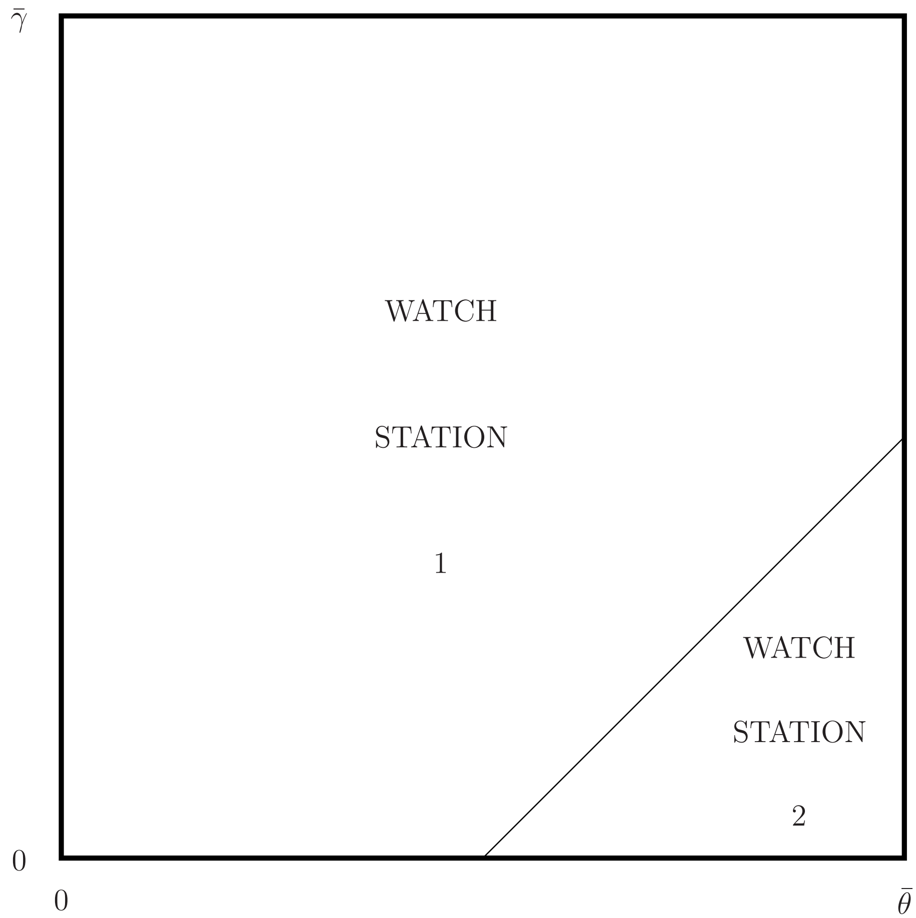


Figure B.1: Viewer Optimal Response - $a_1 < a_2 - \frac{\bar{\theta}}{\bar{\gamma}}$

that, if U_i^s represents viewer i's utility from watching station s:

$$\begin{aligned} \theta_i &< \frac{\bar{\theta}}{2} + \frac{\gamma_i}{2}(a_2 - a_1) \Longleftrightarrow \\ -\theta_i - \gamma_i a_1 &> -\left(\bar{\theta} - \theta_i\right) - \gamma_i a_2 \Longleftrightarrow \\ U_i^1 &> U_i^2 \end{aligned}$$

and vice versa for those for whom $U_i^1 < U_i^2$. Consequently, for given ad levels a_1 and a_2, viewers for whom $\theta_i < \frac{\bar{\theta}}{2} + \frac{\gamma_i}{2}(a_2 - a_1)$ watch network 1, and viewers for whom $\theta_i > \frac{\bar{\theta}}{2} + \frac{\gamma_i}{2}(a_2 - a_1)$ watch network 2.

Given an ad allotment for network 2, a_2, and the viewers' optimal response function, network 1 must choose the ad allotment that maximizes its payoffs. This depends on the level of ads it has chosen as well as the audience size that chooses to watch its show. Suppose network 1 starts at an ad allotment of $a_1 = 0$ and considers increasing its allotment. When network 1 airs only a few ads, the indifference curve will have a steep slope to the right and will intersect the right side of the box, and network 1 will capture most of the viewing market. Figure B.1 displays one example of this scenario. Network 1's audience mass comprises the entire box, less a triangle on the right edge of the box that represents network 2's audience, providing network 1 with a market share of:

$$s_1 = 1 - \frac{1}{\bar{\theta}\bar{\gamma}} \left[\frac{1}{2} \frac{\bar{\theta}}{2} \frac{\bar{\theta}}{a_2 - a_1} \right] = 1 + \frac{\bar{\theta}}{4\bar{\gamma}} \cdot \frac{1}{a_1 - a_2}$$

As network 1 increases its ad allotment, the indifference set separating the box into each network's market share will move accordingly. Note that the individual located halfway between the network's types and with no aversion to ads (i.e., viewer i for whom $\gamma_i = 0$ and $\theta_i = \frac{\bar{\theta}}{2}$) will always be indifferent between the two network's programs, regardless of their respective ad allotments. Consequently, the line's movement as network 1 increases its ad level will manifest itself as a rotation about this point.

As long as this indifference set intersects the right side of the box, the above expression will continue to represent network 1's market share. However, there will be a kink in the share function when this set intersects the corner of the box, or equivalently when both $\left(\frac{\bar{\theta}}{2}, 0\right)$ and $\left(\bar{\theta}, \bar{\gamma}\right)$ are in the indifference set, which occurs when $a_1 = a_2 - \frac{\bar{\theta}}{\bar{\gamma}}$. When the indifference set intersects the top of the box, each network's market share will be the area of the trapezoid formed by the set located

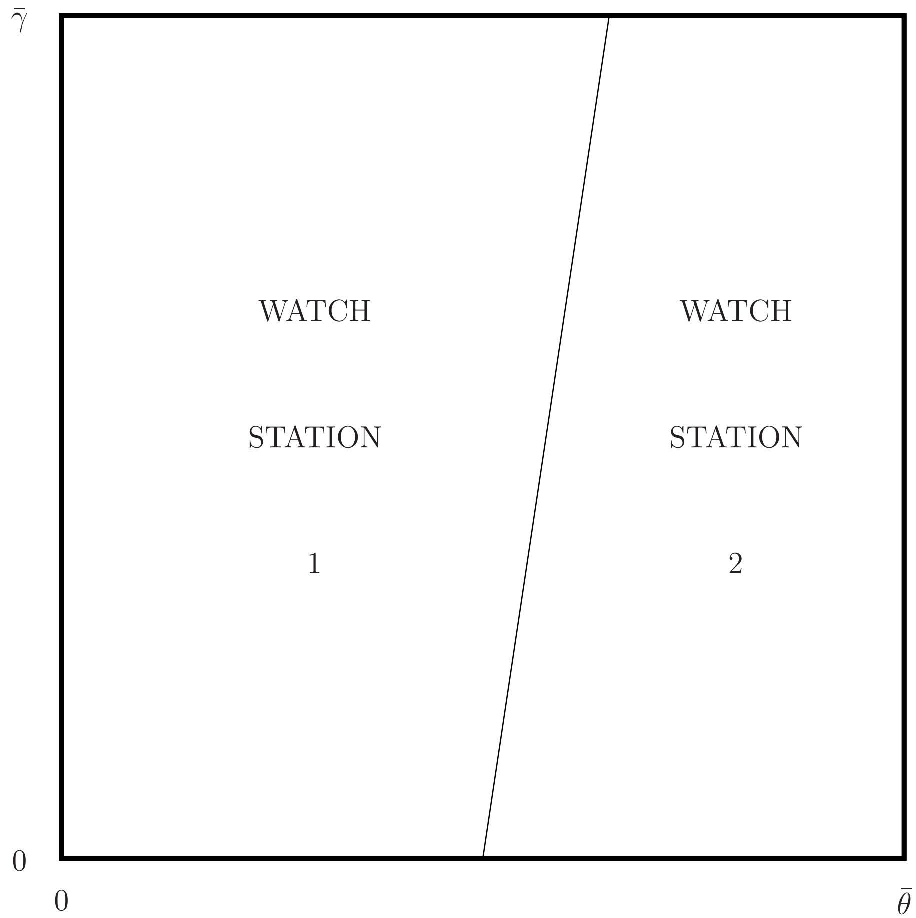


Figure B.2: Viewer Optimal Response - $|a_1 - a_2| < \frac{\bar{\theta}}{\bar{\gamma}}$

closer to its location. The share split for one such scenario is displayed in Figure B.2.

Network 1's audience mass in this case is:

$$\frac{1}{\bar{\theta}\bar{\gamma}} \cdot \frac{1}{2} \cdot \bar{\gamma} \cdot \left(\bar{\theta} + \frac{\bar{\gamma}}{2}(a_2 - a_1)\right) = \frac{1}{2} - \frac{\bar{\gamma}}{4\bar{\theta}}(a_1 - a_2)$$

As network 1 continues to increase its ad level from $a_1 = a_2 - \frac{\bar{\theta}}{\bar{\gamma}}$, the indifference set continues to rotate about $\left(\frac{\bar{\theta}}{2}, 0\right)$ and this expression for audience mass continues to hold. When network 1's allotment reaches network 2's, this indifference set is a vertical line that runs from $\left(\frac{\bar{\theta}}{2}, 0\right)$ to $\left(\bar{\gamma}, \frac{\bar{\theta}}{2}\right)$, splitting the box in half so that each

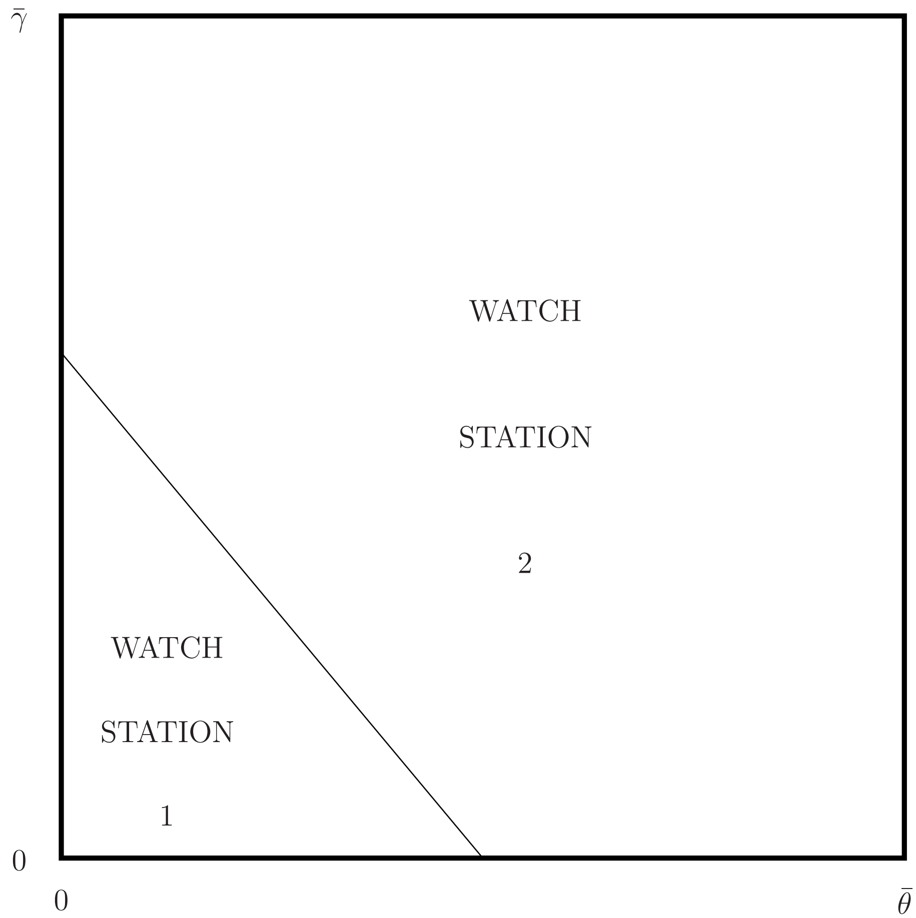


Figure B.3: Viewer Optimal Response - $a_1 > a_2 + \frac{\bar{\theta}}{\bar{\gamma}}$

network captures half the market. When network 1's ad level exceeds that of its rival, the indifference set continues to rotate past the halfway point so that network 1 captures less of the market than its rival does. The above expression continues to hold until the indifference set intersects the corner of the box, which occurs when both $\left(\frac{\bar{\theta}}{2}, 0\right)$ and $(0, \bar{\gamma})$ are in the indifference set, implying $a_1 = a_2 + \frac{\bar{\theta}}{\bar{\gamma}}$. After this point, network one's market share is simply the triangle sectioned off on the left side of the box by the indifference set. Figure B.3 displays such a scenario.

In this case, network 1's market share is:

$$\frac{1}{2} \cdot \frac{1}{\bar{\theta}\bar{\gamma}} \cdot \frac{\bar{\theta}}{2} \cdot \frac{\bar{\theta}}{a_2 - a_1} = \frac{\bar{\theta}}{4\bar{\gamma}} \cdot \frac{1}{a_1 - a_2}$$

Combining these three scenarios and utilizing the fact that each firm's profit is simply market share times its ad level, each firm's profit is:

$$\pi_j = \begin{cases} a_j + \frac{\bar{\theta}}{4\bar{\gamma}} \cdot \frac{a_j}{a_j - a_k}, \ a_j < a_k - \frac{\bar{\theta}}{\bar{\gamma}} \\ \frac{a_j}{2} - \frac{\bar{\gamma}}{4\bar{\theta}} \left(a_j^2 - a_j a_k \right), \ a_k - \frac{\bar{\theta}}{\bar{\gamma}} < a_j < a_k + \frac{\bar{\theta}}{\bar{\gamma}} \\ \frac{\bar{\theta}}{4\bar{\gamma}} \cdot \frac{a_j}{a_j - a_k}, \ a_j > a_k + \frac{\bar{\theta}}{\bar{\gamma}} \end{cases}$$

To determine each firm's optimal response function, I take the derivative of profit with respect to the choice variable for each case in turn. For $a_j < a_k - \frac{\bar{\theta}}{\bar{\gamma}}$:

$$\frac{d\pi_j}{da_j} = 1 + \frac{\bar{\theta}}{4\bar{\gamma}} \cdot \frac{a_j}{a_j - a_k} - a_j \left(\frac{\bar{\theta}}{4\bar{\gamma}} \cdot \frac{1}{(a_j - a_k)^2} \right)$$

Setting this equal to zero yields

$$\begin{aligned} 4\bar{\gamma}(a_j - a_k)^2 + \bar{\theta}(a_j - a_k) - \bar{\theta}a_j &= 0 \\ \implies 4\bar{\gamma}(a_j - a_k)^2 &= \bar{\theta}a_k \\ \implies |a_j - a_k| &= \sqrt{\frac{\bar{\theta}}{4\bar{\gamma}}a_k} \\ \implies a_j &= a_k - \frac{1}{2}\sqrt{\frac{\bar{\theta}}{\bar{\gamma}}a_k} \end{aligned}$$

where the last line utilizes the fact that $a_j < a_k - \frac{\bar{\theta}}{\bar{\gamma}}$. Since this only applies when

$a_j < a_k - \frac{\bar{\theta}}{\bar{\gamma}}$, for this to apply it must be the case that:

$$\begin{aligned} a_k - \sqrt{\frac{\bar{\theta}}{4\bar{\gamma}} a_k} &< a_k - \frac{\bar{\theta}}{\bar{\gamma}} \\ \Longrightarrow \sqrt{\frac{\bar{\theta}}{4\bar{\gamma}} a_k} &> \frac{\bar{\theta}}{\bar{\gamma}} \\ \Longrightarrow a_k &> \frac{4\bar{\theta}}{\bar{\gamma}} \end{aligned}$$

The second order condition is also satisfied for this case:

$$\begin{aligned} \frac{d\pi_j}{da_j} &= 1 + \frac{\bar{\theta}}{4\bar{\gamma}} \left(\frac{1}{a_j - a_k} \right) \left(1 - \frac{a_j}{a_j - a_k} \right) \\ &= 1 + \frac{\bar{\theta}}{4\bar{\gamma}} \left(\frac{-a_k}{(a_j - a_k)^2} \right) \\ \Longrightarrow \frac{d^2\pi_j}{da_j^2} &= \frac{\bar{\theta}}{2\bar{\gamma}} \left(\frac{a_k}{(a_j - a_k)^3} \right) \end{aligned}$$

which is negative since $a_j < a_k$.

For $a_k - \frac{\bar{\theta}}{\bar{\gamma}} < a_j < a_k + \frac{\bar{\theta}}{\bar{\gamma}}$:

$$\frac{d\pi_j}{da_j} = \frac{1}{2} - \frac{\bar{\gamma}}{4\bar{\theta}} (2a_j - a_k)$$

Setting this equal to zero yields:

$$\begin{aligned} \frac{1}{2} - \frac{\bar{\gamma}}{4\bar{\theta}} (2a_j - a_k) &= 0 \\ \Longrightarrow 2\frac{\bar{\theta}}{\bar{\gamma}} &= \frac{\bar{\gamma}}{4\bar{\theta}} (2a_j - a_k) \\ \Longrightarrow a_j &= \frac{a_k}{2} + \frac{\bar{\theta}}{\bar{\gamma}} \end{aligned}$$

The relevant constraints for a_j imply:

$$\begin{aligned} a_j &> a_k - \frac{\bar{\theta}}{\bar{\gamma}} \\ \Longrightarrow \frac{a_k}{2} + \frac{\bar{\theta}}{\bar{\gamma}} &> a_k - \frac{\bar{\theta}}{\bar{\gamma}} \\ \Longrightarrow a_k &< 4\frac{\bar{\theta}}{\bar{\gamma}} \end{aligned}$$

and

$$\begin{aligned} a_j &< a_k + \frac{\bar{\theta}}{\bar{\gamma}} \\ \Longrightarrow \frac{a_k}{2} + \frac{\bar{\theta}}{\bar{\gamma}} &< a_k + \frac{\bar{\theta}}{\bar{\gamma}} \\ \Longrightarrow a_k &> 0 \end{aligned}$$

The second order condition is also satisfied:

$$\frac{d^2\pi_j}{da_j^2} = -2\frac{\bar{\gamma}}{4\bar{\theta}}$$

which is negative since $\bar{\theta}, \bar{\gamma} > 0$. ∎

B.2 Establishment of Equilibrium Point - DVR Case

Proposition 11 *In the ad game with the DVR available, the strategy*

$$a_1 = a_2 = a^* = \frac{1}{\alpha\bar{\gamma}}\left[\bar{\theta} + \sqrt{\bar{\theta}^2 - \frac{1+\alpha}{1-\alpha}r^2}\right]$$

constitutes an equilibrium.

Proof To prove that $a_1 = a_2 = a^*$ constitutes an equilibrium, I first derive the viewers' optimal responses to a given set of ad levels by the networks, then I derive the profit function for each network, and lastly I show that, given the rival firm airs $a_k = a^*$ ads, neither network can achieve a higher profit by airing any other level of ads than $a_j = a^*$ itself.

B.2.1 Viewer Optimal Response

As demonstrated in section 4.3, the viewers now have four potential viewing options:

$$\begin{aligned} U_i^{1,N} &= -\theta_i - \gamma_i a_1, \\ U_i^{2,N} &= -(\bar{\theta} - \theta_i) - \gamma_i a_2, \\ U_i^{1,D} &= -\theta_i - \alpha\gamma_i a_1 - r, \text{ or} \\ U_i^{2,D} &= -(\bar{\theta} - \theta_i) - \alpha\gamma_i a_2 - r \end{aligned}$$

This yields six points of intersection between each pair of indifference lines. The first two involve individuals indifferent between watching a station with a DVR and watching it without one:

$$\begin{aligned} U_i^{j,N} &= U_i^{j,D} \\ \Longrightarrow -|\theta^j - \theta_i| - \gamma_i a_j &= -|\theta^j - \theta_i| - \alpha\gamma_i a_j - r \\ \Longrightarrow -\gamma_i a_j &= -\alpha\gamma_i a_j - r \\ \Longrightarrow \gamma_i &= \frac{r}{(1-\alpha)a_j} \end{aligned}$$

The third involves those indifferent between watching each station without a DVR:

$$
\begin{aligned}
U_i^{1,N} &= U_i^{2,N} \\
\Longrightarrow -\theta_i - \gamma_i a_1 &= -(\bar{\theta} - \theta_i) - \gamma_i a_2 \\
\Longrightarrow \bar{\theta} &= 2\theta_i + \gamma_i(a_1 - a_2) \\
\Longrightarrow \gamma_i &= -\frac{\bar{\theta}}{a_2 - a_1} + \theta_i \frac{2}{a_2 - a_1}
\end{aligned}
$$

The fourth represents those indifferent between watching each station with one:

$$
\begin{aligned}
U_i^{1,D} &= U_i^{2,D} \\
\Longrightarrow -\theta_i - \alpha\gamma_i a_1 - r &= -(\bar{\theta} - \theta_i) - \alpha\gamma_i a_2 - r \\
\Longrightarrow \bar{\theta} &= 2\theta_i + \alpha\gamma_i(a_1 - a_2) \\
\Longrightarrow \gamma_i &= -\frac{\bar{\theta}}{\alpha(a_2 - a_1)} + \theta_i \frac{2}{\alpha(a_2 - a_1)}
\end{aligned}
$$

Lastly, the fifth and sixth represent those indifferent between watching one station with a DVR and the other without one.

$$\begin{aligned} U_i^{1,D} &= U_i^{2,N} \\ \implies -\theta_i - \alpha\gamma_i a_1 - r &= -(\bar{\theta} - \theta_i) - \gamma_i a_2 \\ \implies \bar{\theta} - r &= 2\theta_i + \gamma_i(\alpha a_1 - a_2) \\ \implies \gamma_i &= -\frac{\bar{\theta} - r}{(a_2 - \alpha a_1)} + \theta_i \frac{2}{(a_2 - \alpha a_1)}; \end{aligned}$$

$$\begin{aligned} U_i^{1,N} &= U_i^{2,D} \\ \implies -\theta_i - \gamma_i a_1 &= -(\bar{\theta} - \theta_i) - \alpha\gamma_i a_2 - r \\ \implies \bar{\theta} + r &= 2\theta_i + \gamma_i(a_1 - \alpha a_2) \\ \implies \gamma_i &= -\frac{\bar{\theta} + r}{(\alpha a_2 - a_1)} + \theta_i \frac{2}{(\alpha a_2 - a_1)} \end{aligned}$$

This yields five distinct intersection points between each pair of indifference lines. The first is a crossing between two unrelated indifference lines:

$$U_i^{1,D} = U_i^{2,N} \wedge U_i^{1,N} = U_i^{2,D} \implies \gamma_i = \frac{2r}{(1-\alpha)(a_1+a_2)}, \theta_i = \frac{\bar{\theta}}{2} + \frac{r}{2} \cdot \frac{1+\alpha}{1-\alpha}\left(\frac{a_2 - a_1}{a_2 + a_1}\right)$$

whereas the last four represent the four points of three-way indifference:

$$\begin{aligned} U_i^{1,N} = U_i^{2,N} = U_i^{1,D} &\implies \gamma_i = \frac{r}{a_1(1-\alpha)}, \theta_i = \frac{\bar{\theta}}{2} + \frac{r}{2} \cdot \frac{1}{1-\alpha}\left(\frac{a_2}{a_1} - 1\right) \\ U_i^{1,N} = U_i^{2,N} = U_i^{2,D} &\implies \gamma_i = \frac{r}{a_2(1-\alpha)}, \theta_i = \frac{\bar{\theta}}{2} + \frac{r}{2} \cdot \frac{1}{1-\alpha}\left(1 - \frac{a_1}{a_2}\right) \\ U_i^{1,N} = U_i^{1,D} = U_i^{2,D} &\implies \gamma_i = \frac{r}{a_1(1-\alpha)}, \theta_i = \frac{\bar{\theta}}{2} + \frac{r}{2} \cdot \frac{\alpha}{1-\alpha}\left(\frac{a_2}{a_1} - 1\right) \\ U_i^{1,D} = U_i^{2,N} = U_i^{2,D} &\implies \gamma_i = \frac{r}{a_2(1-\alpha)}, \theta_i = \frac{\bar{\theta}}{2} + \frac{r}{2} \cdot \frac{\alpha}{1-\alpha}\left(1 - \frac{a_1}{a_2}\right) \end{aligned}$$

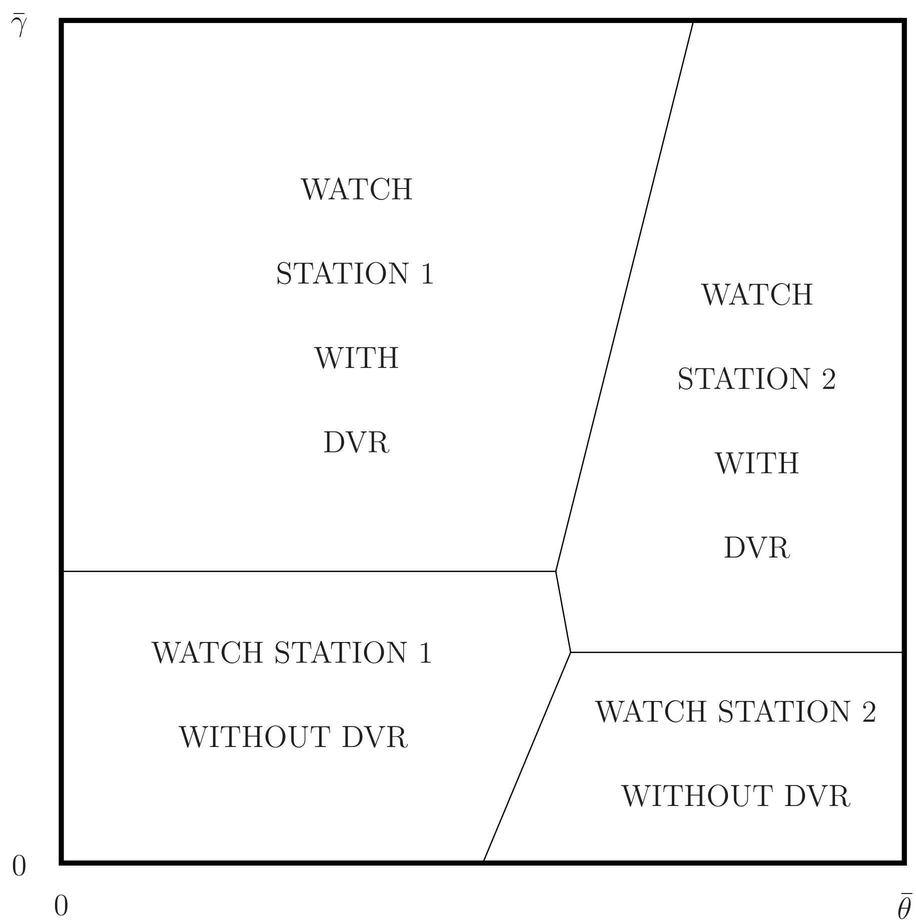


Figure B.4: Viewer Optimal Response Map - DVR Available

Figure B.4 displays an example of what these six indifference lines might look like. These indifference lines section off the viewers into the various decisions they make, as demonstrated in the figure.

B.2.2 Profit Function Derivation

Now that the viewer's response has been determined, the next step is to move on to the networks. I focus solely on network 1 here without loss of generality due to symmetry. Assume that network 2 has chosen to play $a_2 = a^*$, and that network 1 must determine its optimal response to this decision. When $a_1 = 0$, firm one would receive the entire market except for the trapezoid and triangle towards the bottom right of the box representing firm 2's market share, as displayed in Figure B.5. None of its viewers would be using a DVR, and the intersection of the two indifference lines that forms the trapezoid is guaranteed to exist inside the box since:

$$\begin{aligned}
\bar{\theta} &> \frac{r}{1-\alpha} \\
\implies \frac{\bar{\theta}}{2} + \frac{r}{2} \cdot \frac{1}{1-\alpha} &< \bar{\theta} \\
\implies \frac{\bar{\theta}}{2} + \frac{r}{2} \cdot \frac{1}{1-\alpha}\left(1 - \frac{a_1}{a^*}\right) &< \bar{\theta}
\end{aligned}$$

Note that since

$$\begin{aligned}
\bar{\theta} &> \frac{r}{1-\alpha} \\
\implies \frac{\bar{\theta}}{r} - \frac{\alpha}{1-\alpha} &> -\sqrt{\frac{\bar{\theta}^2}{r^2} - \frac{1+\alpha}{1-\alpha}} \\
\implies \frac{\bar{\theta}}{r} + \sqrt{\frac{\bar{\theta}^2}{r^2} - \frac{1+\alpha}{1-\alpha}} &> \frac{\alpha}{1-\alpha} \\
\implies \frac{1}{\alpha\bar{\gamma}}\left[\bar{\theta} + \sqrt{\bar{\theta}^2 - \frac{\alpha}{1-\alpha}}\right] &> \frac{r}{\bar{\gamma}(1-\alpha)} \\
\implies a^* &> \frac{r}{\bar{\gamma}(1-\alpha)} \\
\implies \bar{\gamma} &> \frac{r}{a^*(1-\alpha)},
\end{aligned}$$

the horizontal indifference set that represents those indifferent between watching station 2 with or without a DVR exists inside the box and remains there as firm one considers changing its ad level.

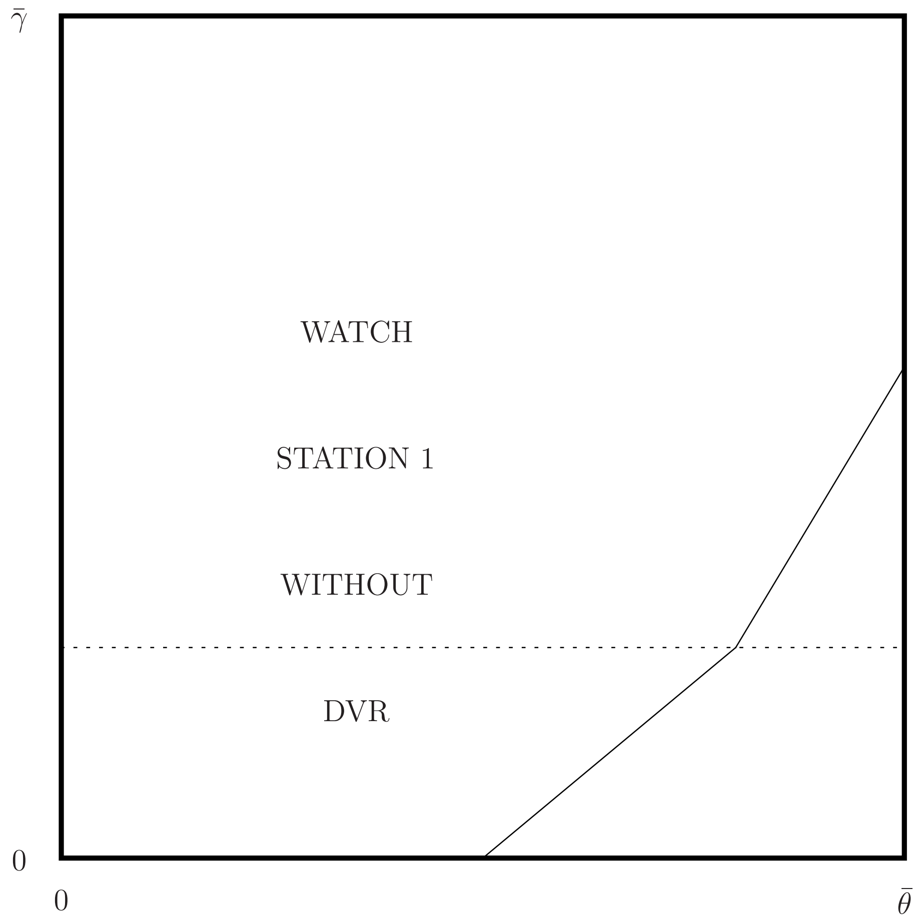


Figure B.5: Potential Viewer Response Map

The trapezoid has an area of:

$$\frac{r}{2a^*(1-\alpha)}\left(\bar{\theta}-\frac{r(a^*-a_1)}{2a^*(1-\alpha)}\right)=\frac{r\bar{\theta}}{2a^*(1-\alpha)}-\left(\frac{r}{2a^*(1-\alpha)}\right)^2(a^*-a_1)$$

The triangle has an area of:

$$\begin{aligned}&\frac{1}{2}\left(\frac{\bar{\theta}-r}{\alpha a^*-a_1}-\frac{r}{a_2(1-\alpha)}\right)\left(\frac{\bar{\theta}}{2}-\frac{r(a^*-a_1)}{2(1-\alpha)a^*}\right)\\&\quad=\frac{1}{4(\alpha a_1-a^*)}\left(\bar{\theta}-\frac{ra^*-\alpha ra^*+\alpha ra^*-ra_1}{a^*(1-\alpha)}\right)\\&\quad=\frac{1}{4(\alpha a_1-a^*)}\left(\bar{\theta}-\frac{r(a^*-a_1)}{a^*(1-\alpha)}\right)^2\end{aligned}$$

The area representing its set of viewers would therefore be:

$$
\begin{aligned}
&\bar{\theta}\bar{\gamma} - \frac{r\bar{\theta}}{2a^*(1-\alpha)} + \left(\frac{r}{2a^*(1-\alpha)}\right)^2 (a^* - a_1) - \frac{1}{4}\left(\frac{1}{\alpha a^* - a_1}\right)\left(\bar{\theta} - \frac{r(a^* - a_1)}{a^*(1-\alpha)}\right)^2 \\
&\quad = \bar{\theta}\bar{\gamma} - \frac{r\bar{\theta}}{2a^*(1-\alpha)} + \left(\frac{r}{2a^*(1-\alpha)}\right)^2 (a^* - a_1) \\
&\qquad - \frac{1}{4}\left(\frac{1}{\alpha a^* - a_1}\right)\left(\bar{\theta}^2 - 2r\bar{\theta}\frac{a^* - a_1}{a^*(1-\alpha)} + \frac{r^2\left(a^* - a_1\right)^2}{\left(a^*\left(1-\alpha\right)\right)^2}\right) \\
&\quad = \bar{\theta}\bar{\gamma} - \frac{r\bar{\theta}}{2a^*(1-\alpha)} + \left(\frac{r}{2a^*(1-\alpha)}\right)^2 (a^* - a_1) \\
&\qquad - \frac{1}{4}\left(\frac{1}{\alpha a^* - a_1}\right)\left(\left(\bar{\theta} - r\right)^2 - 2r\bar{\theta}\frac{\alpha a^* - a_1}{a^*(1-\alpha)} + \frac{(a^* - a_1)^2 - (a^*(1-\alpha))^2}{\left(a^*(1-\alpha)\right)^2} r^2\right) \\
&\quad = \bar{\theta}\bar{\gamma} - \frac{r\bar{\theta}}{2a^*(1-\alpha)} + \left(\frac{r}{2a^*(1-\alpha)}\right)^2 (a^* - a_1) \\
&\qquad - \frac{1}{4}\left(\frac{1}{\alpha a^* - a_1}\right)\left(\left(\bar{\theta} - r\right)^2 - 2r\bar{\theta}\frac{\alpha a^* - a_1}{a^*(1-\alpha)} + \frac{(a^*(2-\alpha) - a_1)(\alpha a^* - a_1)}{\left(a^*(1-\alpha)\right)^2} r^2\right) \\
&\quad = \bar{\theta}\bar{\gamma} - \frac{1}{\alpha a^* - a_1}\left(\frac{\bar{\theta} - r}{2}\right)^2 - \frac{r\bar{\theta}}{2a^*(1-\alpha)} + \left(\frac{r}{2a^*(1-\alpha)}\right)^2 (a^* - a_1) \\
&\qquad + \frac{r\bar{\theta}}{2a^*(1-\alpha)} + \left(\frac{r}{2a^*(1-\alpha)}\right)^2 (a^*(2-\alpha) - a_1) \\
&\quad = \bar{\theta}\bar{\gamma} - \frac{1}{\alpha a^* - a_1}\left(\frac{\bar{\theta} - r}{2}\right)^2 + \left(\frac{r}{2a^*(1-\alpha)}\right)^2 (-a^*(1-\alpha)) \\
&\quad = \bar{\theta}\bar{\gamma} - \frac{1}{\alpha a^* - a_1}\left(\frac{\bar{\theta} - r}{2}\right)^2 - \frac{r^2}{4a^*(1-\alpha)}
\end{aligned}
$$

and its profit would be that expression times $\frac{a_1}{\bar{\theta}\bar{\gamma}}$:

$$
a_1 - \frac{1}{\bar{\theta}\bar{\gamma}}\left(\frac{a_1}{\alpha a^* - a_1}\left(\frac{\bar{\theta} - r}{2}\right)^2 + \frac{r^2 a_1}{4a^*(1-\alpha)}\right)
$$

As a_1 continues to increase it ad level, eventually the horizontal indifference set that represents viewers indifferent between watching its show with a DVR and watching without one enters the box such that it gets a class of DVR-using viewers. This

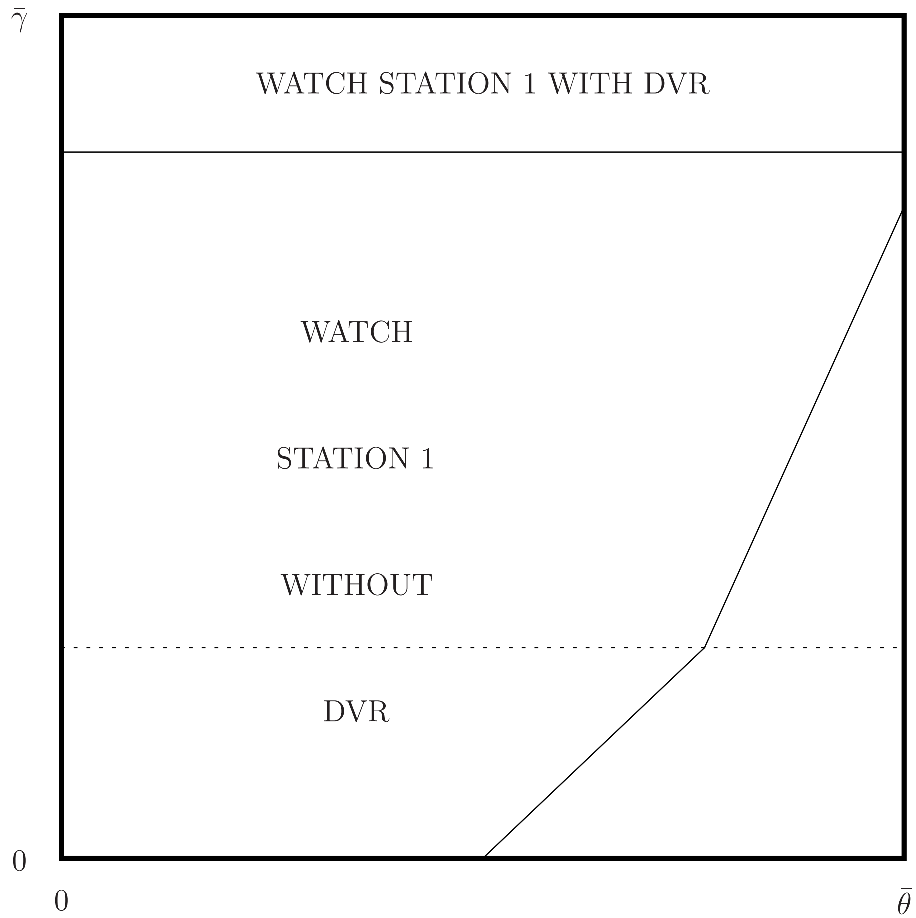


Figure B.6: Potential Viewer Response Map

occurs when:

$$\bar{\gamma} = \frac{r}{a_1(1-\alpha)} \Longrightarrow a_1 = \frac{r}{\bar{\gamma}(1-\alpha)}$$

Figure B.6 displays such a situation.

The only difference between this scenario and the last one is that the top box of DVR viewers now provide network 1 with only αa_1 impressions each instead of the

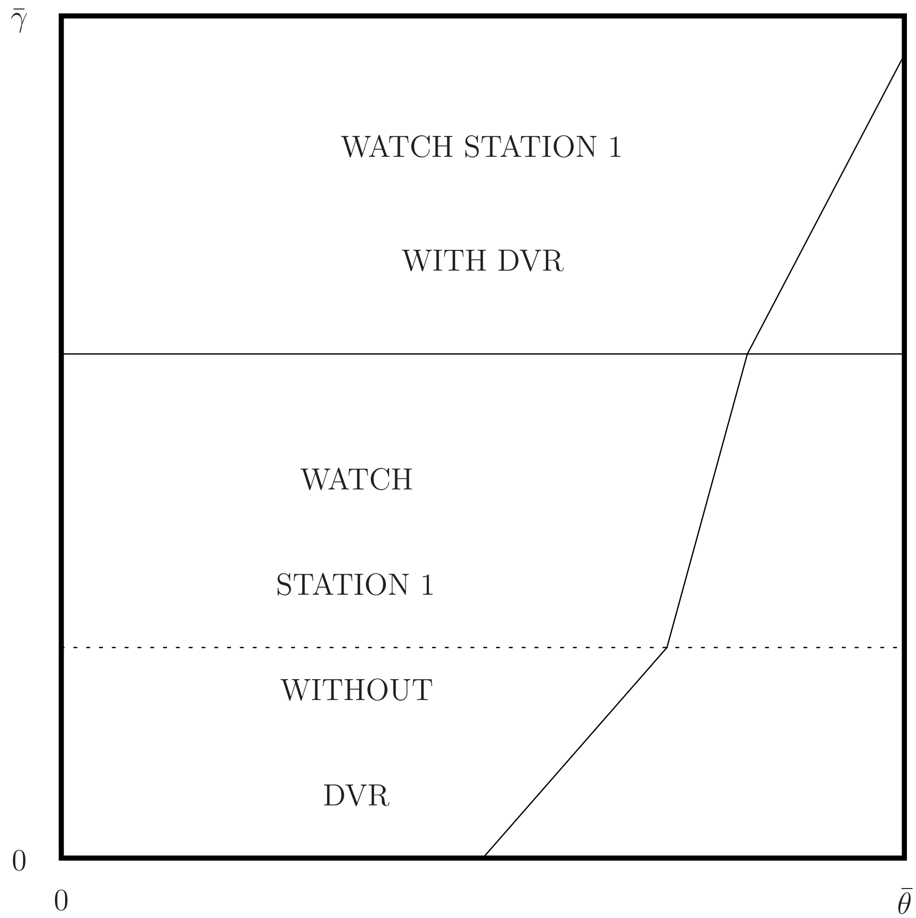


Figure B.7: Potential Viewer Response Map

full a_1 ad level. Therefore, the profit in this scenario is:

$$
\begin{aligned}
a_1 - \frac{1}{\bar{\theta}\bar{\gamma}}\left(\frac{a_1}{\alpha a^* - a_1}\left(\frac{\bar{\theta}-r}{2}\right)^2 + \frac{r^2 a_1}{4a^*(1-\alpha)}\right) - \left(\frac{a_1}{\bar{\theta}\bar{\gamma}}(1-\alpha)\cdot\bar{\theta}\left(\bar{\gamma} - \frac{r}{a_1(1-\alpha)}\right)\right) \\
= \alpha a_1 - \frac{1}{\bar{\theta}\bar{\gamma}}\left(\frac{a_1}{\alpha a^* - a_1}\left(\frac{\bar{\theta}-r}{2}\right)^2 + \frac{r^2 a_1}{4a^*(1-\alpha)} - r\bar{\theta}\right)
\end{aligned}
$$

The next kink in the profit function occurs when the indifference set representing those indifferent between watching station 1 with or without a DVR intersects the indifference line below it, as depicted in Figure B.7.

This occurs when:

$$
\begin{aligned}
\frac{\bar{\theta}}{2} + \frac{r}{2} \cdot \frac{\alpha}{1-\alpha} \cdot \left(\frac{a^*}{a_1} - 1 \right) &= \bar{\theta} \\
\implies \frac{\bar{\theta}}{r} \cdot \frac{1-\alpha}{\alpha} + 1 &= \frac{a^*}{a_1} \\
\implies a_1 &= \frac{r\alpha a^*}{\bar{\theta}(1-\alpha) + r\alpha}
\end{aligned}
$$

Solving for profit in this segment of the profit function is easier once the next segment has been solved for, so I move on to that one here and then subsequently come back to solve for profit along this segment.

The next scenario occurs when the topmost indifference set intersects the top of the box, which occurs when:

$$
\begin{aligned}
\frac{\bar{\theta}}{2} + \frac{\bar{\gamma}}{2} \left(\alpha \left(a^* - a_1 \right) \right) &= \bar{\theta} \\
\implies a^* - a_1 &= \frac{\bar{\theta}}{\alpha \bar{\gamma}} \\
\implies a_1 = a^* - \frac{\bar{\theta}}{\alpha \bar{\gamma}}
\end{aligned}
$$

Figure B.8 demonstrates an example of such a scenario. The area of network 1's DVR audience region is:

$$
\begin{aligned}
&\left(\bar{\gamma} - \frac{r}{a_1(1-\alpha)} \right) \left(\frac{1}{2} \left(\left(\frac{\bar{\theta}}{2} + \frac{r}{2} \cdot \frac{\alpha(a^* - a_1)}{a_1(1-\alpha)} \right) + \left(\frac{\bar{\theta}}{2} + \frac{\bar{\gamma}}{2} \alpha \left(a^* - a_1 \right) \right) \right) \right) \\
&= \bar{\theta} \left(\frac{\bar{\gamma}}{2} - \frac{r}{2a_1(1-\alpha)} \right) + \frac{\alpha(a^* - a_1)}{4} \left(\bar{\gamma}^2 - \frac{r^2}{a_1^2(1-\alpha)^2} \right) \equiv A
\end{aligned}
$$

The non-DVR audience region for network one is comprised of two trapezoids, one of

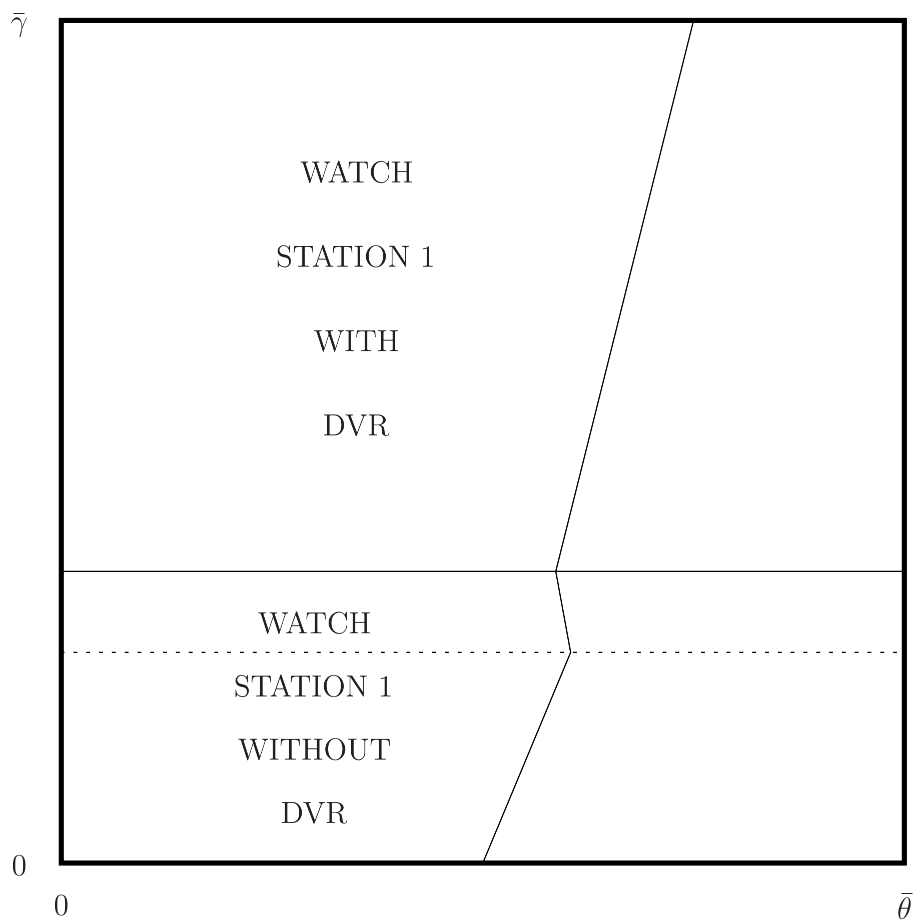


Figure B.8: Potential Viewer Response Map

area:

$$\left(\frac{r}{a_1(1-\alpha)} - \frac{r}{a^*(1-\alpha)}\right)\left(\frac{1}{2}\left(\bar{\theta} + \frac{r}{2}\frac{1}{1-\alpha}\left(\frac{\alpha(a^*-a_1)}{a_1} + \frac{a^*-a_1}{a^*}\right)\right)\right)$$
$$= \frac{r}{a_1 a^*(1-\alpha)}\left(a^*-a_1\right)\cdot\frac{1}{2}\left(\bar{\theta} + \frac{r}{2}\cdot\frac{1}{1-\alpha}\cdot\frac{\left(\alpha a^* + a_1\right)\left(a^*-a_1\right)}{a_1 a^*}\right)$$
$$= \frac{r(a^*-a_1)}{a_1 a^*(1-\alpha)}\frac{\bar{\theta}}{2} + \frac{r^2}{4}\frac{\left(\alpha a^* + a_1\right)\left(a^*-a_1\right)^2}{(a_1 a^*(1-\alpha))^2} \equiv B$$

and the other of area:

$$\frac{r}{a^*(1-\alpha)}\cdot\frac{1}{2}\left[\bar{\theta} + \frac{r}{2}\frac{1}{1-\alpha}\frac{a^*-a_1}{a^*}\right] = \frac{r\bar{\theta}}{2a^*(1-\alpha)} + \left(\frac{r}{2a^*(1-\alpha)}\right)^2(a^*-a_1) \equiv C$$

Consequently, network 1's profit is:

$$
\begin{aligned}
\pi_1 &= \frac{a_1}{\bar{\theta}\bar{\gamma}}\left[\alpha A + B + C\right] \\
&= \frac{a_1}{\bar{\theta}\bar{\gamma}}\left[\frac{\alpha\bar{\theta}\bar{\gamma}}{2} - \frac{\alpha r\bar{\theta}}{2a_1(1-\alpha)} + \frac{\alpha^2\bar{\gamma}^2(a^*-a_1)}{4} - \frac{\alpha^2 r^2 a^*}{4a_1^2(1-\alpha)^2} + \frac{\alpha^2 r^2}{4a_1(1-\alpha)^2}\right. \\
&\quad + \frac{r\bar{\theta}}{2a_1(1-\alpha)} - \frac{r\bar{\theta}}{2a^*(1-\alpha)} + \frac{r^2(\alpha a^* + a_1)(a^{*2} - 2a^*a_1 + a_1^2)}{4a_1^2 a^{*2}(1-\alpha)^2} \\
&\quad \left. + \frac{r\bar{\theta}}{2a^*(1-\alpha)} + \frac{r^2(a^*-a_1)}{4a^{*2}(1-\alpha)^2}\right] \\
&= \frac{a_1}{\bar{\theta}\bar{\gamma}}\left[\frac{\alpha\bar{\theta}\bar{\gamma}}{2} - \frac{\alpha r\bar{\theta}}{2a_1(1-\alpha)} + \frac{\alpha^2\bar{\gamma}^2 a^*}{4} - \frac{\alpha^2\bar{\gamma}^2 a_1}{4} - \frac{\alpha^2 r^2 a^*}{4a_1^2(1-\alpha)^2} + \frac{\alpha^2 r^2}{4a_1(1-\alpha)^2}\right. \\
&\quad + \frac{r\bar{\theta}}{2a_1(1-\alpha)} + \frac{\alpha r^2 a^*}{4a_1^2(1-\alpha)^2} - \frac{\alpha r^2}{2a_1(1-\alpha)^2} + \frac{\alpha r^2}{4a^*(1-\alpha)^2} + \frac{r^2}{4a_1(1-\alpha)^2} \\
&\quad \left. - \frac{r^2}{2a^*(1-\alpha)^2} + \frac{r^2 a_1}{4a^{*2}(1-\alpha)^2} + \frac{r^2}{4a^*(1-\alpha)^2} - \frac{r^2 a_1}{4a^{*2}(1-\alpha)^2}\right] \\
&= \frac{a_1}{\bar{\theta}\bar{\gamma}}\left[\frac{(1-\alpha)r\bar{\theta}}{2a_1(1-\alpha)} + \frac{(1-\alpha)\alpha r^2 a^*}{4a_1^2(1-\alpha)^2} + \frac{r^2(1-2\alpha+\alpha^2)}{4a_1(1-\alpha)^2} - \frac{(1-\alpha)r^2}{4a^*(1-\alpha)^2}\right. \\
&\quad \left. + \frac{\alpha^2\bar{\gamma}^2 a^*}{4} + \frac{\alpha\bar{\theta}\bar{\gamma}}{2} - \frac{\alpha^2\bar{\gamma}^2 a_1}{4}\right] \\
&= \frac{1}{\bar{\theta}\bar{\gamma}}\left[\frac{r^2 + 2r\bar{\theta}}{4} + \frac{\alpha r^2 a^*}{4(1-\alpha)a_1} + \frac{\alpha^2\bar{\gamma}^2 a_1 a^*}{4} + \frac{\alpha a_1\bar{\theta}\bar{\gamma}}{2} - \frac{r^2 a_1}{4a^*(1-\alpha)} - \frac{\alpha^2\bar{\gamma}^2 a_1^2}{4}\right]
\end{aligned}
\tag{B.1}
$$

Returning to the previous segment of the profit function, note that the only difference between that profit function and the one that applies here is that the previous function attributes an entire trapezoid at the top of the box to firm 1's DVR audience, whereas that trapezoid runs past the edge of the box and therefore needs to be truncated for that scenario. Consequently, one can just remove the excess audience attribution from the previous profit function to get the relevant profit function here.

The overspill of the trapezoid outside of the box has area:

$$
\begin{aligned}
&\frac{1}{2}\left(\bar{\theta} - \left(\frac{\bar{\theta}}{2} + \frac{\bar{\gamma}\alpha(a^* - a_1)}{2}\right)\right)\left(\bar{\gamma} - \frac{\bar{\theta}}{\alpha(a^* - a_1)}\right) \\
&\quad = \frac{1}{4}\left(\bar{\gamma}\left(\alpha(a^* - a_1) - \bar{\theta}\right)\right)\left(\bar{\gamma} - \frac{\bar{\theta}}{\alpha(a^* - a_1)}\right) \\
&\quad = \frac{1}{4\alpha(a^* - a_1)}\left(\bar{\gamma}\alpha(a^* - a_1)\right)^2
\end{aligned}
$$

which means the original profit function given in equation B.1 overstates the profit in this scenario by that amount times $\alpha\frac{a_1}{\bar{\theta}\bar{\gamma}}$. Consequently, profit in this scenario is:

$$
\begin{aligned}
&\frac{1}{\bar{\theta}\bar{\gamma}}\left[\frac{r^2 + 2r\bar{\theta}}{4} + \frac{\alpha r^2 a^*}{4(1-\alpha)a_1} + \frac{\alpha^2\bar{\gamma}^2 a_1 a^*}{4} + \frac{\alpha a_1\bar{\theta}\bar{\gamma}}{2} - \frac{r^2 a_1}{4a^*(1-\alpha)} - \frac{\alpha^2\bar{\gamma}^2 a_1^2}{4}\right. \\
&\qquad\left. - \frac{a_1}{4(a^* - a_1)}\left(\bar{\gamma}\alpha(a^* - a_1)\right)^2\right] \\
&= \frac{1}{\bar{\theta}\bar{\gamma}}\left[\frac{r^2}{4} + \frac{r\bar{\theta}}{2} + \frac{\alpha r^2 a^*}{4(1-\alpha)a_1} + \frac{\alpha^2\bar{\gamma}^2 a_1 a^*}{4} + \frac{\alpha a_1\bar{\theta}\bar{\gamma}}{2} - \frac{r^2 a_1}{4a^*(1-\alpha)} - \frac{\alpha^2\bar{\gamma}^2 a_1^2}{4}\right. \\
&\qquad\left. - \frac{\alpha^2\bar{\gamma}^2 a_1 a^*}{4} + \frac{\alpha^2\bar{\gamma}^2 a_1^2}{4} + \frac{\alpha a_1\bar{\theta}\bar{\gamma}}{2} - \frac{\bar{\theta}^2 a_1}{4(a^* - a_1)}\right] \\
&= \frac{1}{\bar{\theta}\bar{\gamma}}\left[\frac{r^2}{4} + \frac{r\bar{\theta}}{2} + \frac{\alpha r^2 a^*}{4(1-\alpha)a_1} + \alpha a_1\bar{\theta}\bar{\gamma} - \frac{r^2 a_1}{4a^*(1-\alpha)} - \frac{\bar{\theta}^2 a_1}{4(a^* - a_1)}\right]
\end{aligned}
$$

As a_1 continues to consider increasing its ad level for $a_1 = a^* - \frac{\bar{\theta}}{\alpha\bar{\gamma}}$, the next kink occurs when

$$
\frac{r}{a_1(1-\alpha)} = \frac{r}{a^*(1-\alpha)} \Longrightarrow a_1 = a^*
$$

After this, the viewers' responses to the networks' decisions changes slightly, as depicted in Figure B.9. Now, the horizontal indifference line representing those indifferent between watching network 1 with a DVR and watching it without one lies below the same line for those indifferent between watching network 2 either with or

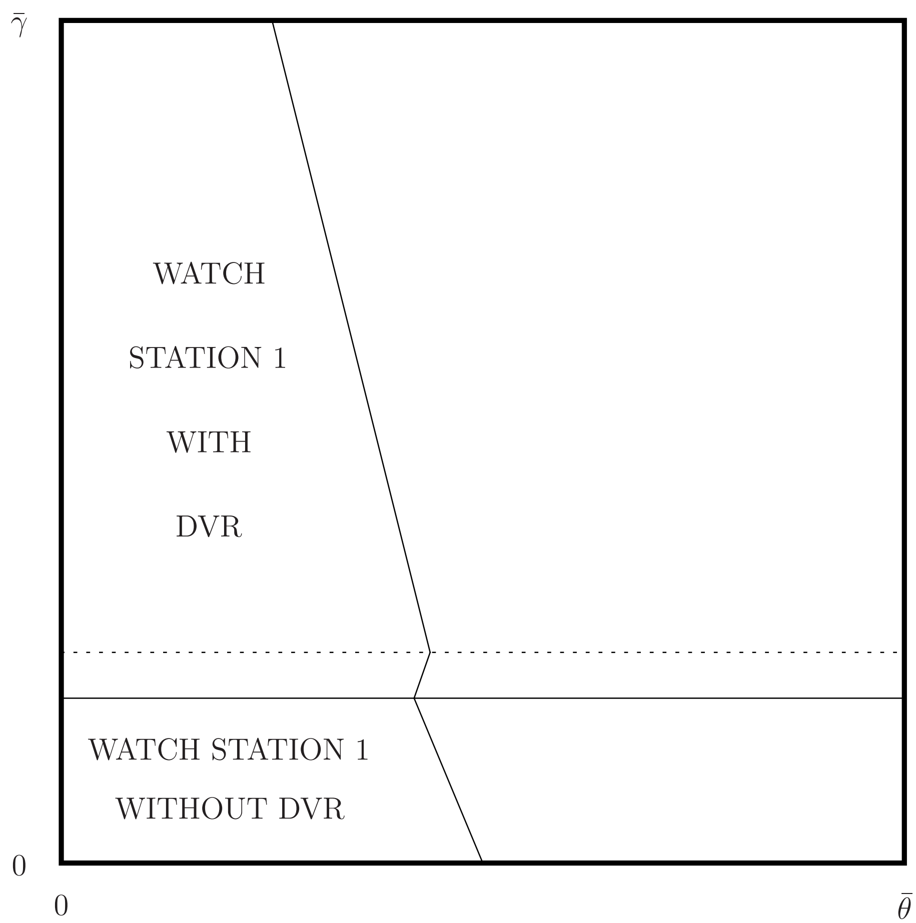


Figure B.9: Potential Viewer Response Map

without a DVR. Consequently, network 1's DVR audience region is composed of two trapezoids, one of area:

$$\left(\bar{\gamma} - \frac{r}{2a^*(1-\alpha)}\right)\left(\frac{1}{2}\left(\left(\frac{\bar{\theta}}{2} + \frac{r}{2} \cdot \frac{\alpha(a^* - a_1)}{a^*(1-\alpha)}\right) + \left(\frac{\bar{\theta}}{2} + \frac{\bar{\gamma}}{2}\alpha\left(a^* - a_1\right)\right)\right)\right)$$
$$= \bar{\theta}\left(\frac{\bar{\gamma}}{2} - \frac{r}{a^*(1-\alpha)}\right) + \frac{\alpha(a^* - a_1)}{4}\left(\bar{\gamma}^2 - \frac{r^2}{a^{*2}(1-\alpha)^2}\right) \equiv D$$

the other of area:

$$
\begin{aligned}
&\left(\frac{r}{a^*(1-\alpha)}-\frac{r}{a_1(1-\alpha)}\right)\left(\frac{1}{2}\left(\bar{\theta}+\frac{r}{2}\frac{1}{1-\alpha}\left(\frac{\alpha(a^*-a_1)}{a^*}+\frac{a^*-a_1}{a_1}\right)\right)\right)\\
&\quad=\frac{r}{a_1a^*(1-\alpha)}\left(a_1-a^*\right)\cdot\frac{1}{2}\left(\bar{\theta}+\frac{r}{2}\cdot\frac{1}{1-\alpha}\cdot\frac{\left(\alpha a_1+a^*\right)\left(a^*-a_1\right)}{a_1a^*}\right)\\
&\quad=\frac{r(a_1-a^*)}{a_1a^*(1-\alpha)}\frac{\bar{\theta}}{2}-\frac{r^2}{4}\frac{\left(\alpha a_1+a^*\right)\left(a_1-a^*\right)^2}{(a_1a^*(1-\alpha))^2}\equiv E
\end{aligned}
$$

with network 1's non-DVR region comprising an area:

$$
\frac{r}{a_1(1-\alpha)}\cdot\frac{1}{2}\left[\bar{\theta}+\frac{r}{2}\frac{1}{1-\alpha}\frac{a^*-a_1}{a_1}\right]=\frac{r\bar{\theta}}{2a_1(1-\alpha)}+\left(\frac{r}{2a_1(1-\alpha)}\right)^2(a^*-a_1)\equiv F
$$

Network 1's profit is therefore:

$$
\begin{aligned}
\pi_1&=\frac{a_1}{\bar{\theta}\bar{\gamma}}\left[\alpha(D+E)+F\right]\\
&=\frac{a_1}{\bar{\theta}\bar{\gamma}}\left[\frac{\alpha\bar{\theta}\bar{\gamma}}{2}-\frac{\alpha r\bar{\theta}}{2a^*(1-\alpha)}+\frac{\alpha^2\bar{\gamma}^2(a^*-a_1)}{4}-\frac{\alpha^2r^2}{4a^*(1-\alpha)^2}+\frac{\alpha^2r^2a_1}{4a^{*2}(1-\alpha)^2}\right.\\
&\qquad+\frac{\alpha r\bar{\theta}}{2a^*(1-\alpha)}-\frac{\alpha r\bar{\theta}}{2a_1(1-\alpha)}-\frac{\alpha r^2(\alpha a_1+a^*)(a_1^2-2a_1a^*+a^{*2})}{4a_1^2a^{*2}(1-\alpha)^2}\\
&\qquad\left.+\frac{r\bar{\theta}}{2a_1(1-\alpha)}+\frac{r^2(a^*-a_1)}{4a_1^2(1-\alpha)^2}\right]\\
&=\frac{a_1}{\bar{\theta}\bar{\gamma}}\left[\frac{\alpha\bar{\theta}\bar{\gamma}}{2}+\frac{\alpha^2\bar{\gamma}^2a^*}{4}-\frac{\alpha^2\bar{\gamma}^2a_1}{4}-\frac{\alpha^2r^2}{4a^*(1-\alpha)^2}+\frac{\alpha^2r^2a_1}{4a^{*2}(1-\alpha)^2}-\frac{\alpha r\bar{\theta}}{2a_1(1-\alpha)}\right.\\
&\qquad-\frac{\alpha^2r^2a_1}{4a^{*2}(1-\alpha)^2}+\frac{\alpha^2r^2}{2a^*(1-\alpha)^2}-\frac{\alpha^2r^2}{4a_1(1-\alpha)^2}-\frac{\alpha r^2}{4a^*(1-\alpha)^2}+\frac{\alpha r^2}{2a_1(1-\alpha)^2}\\
&\qquad\left.-\frac{\alpha r^2a^*}{4a_1^2(1-\alpha)^2}+\frac{r\bar{\theta}}{2a_1(1-\alpha)}+\frac{r^2a^*}{4a_1^2(1-\alpha)^2}-\frac{r^2}{4a_1(1-\alpha)^2}\right]\\
&=\frac{a_1}{\bar{\theta}\bar{\gamma}}\left[\frac{(1-\alpha)r\bar{\theta}}{2a_1(1-\alpha)}+\frac{(1-\alpha)r^2a^*}{4a_1^2(1-\alpha)^2}-\frac{r^2(1-2\alpha+\alpha^2)}{4a_1(1-\alpha)^2}-\frac{\alpha(1-\alpha)r^2}{4a^*(1-\alpha)^2}\right.\\
&\qquad\left.+\frac{\alpha\bar{\theta}\bar{\gamma}}{2}+\frac{\alpha^2\bar{\gamma}^2a^*}{4}-\frac{\alpha^2\bar{\gamma}^2a_1}{4}\right]\\
&=\frac{1}{\bar{\theta}\bar{\gamma}}\left[\frac{2r\bar{\theta}-r^2}{4}+\frac{r^2a^*}{4(1-\alpha)a_1}-\frac{\alpha r^2a_1}{4a^*(1-\alpha)}+\frac{\alpha a_1\bar{\theta}\bar{\gamma}}{2}+\frac{\alpha^2\bar{\gamma}^2a_1a^*}{4}-\frac{\alpha^2\bar{\gamma}^2a_1^2}{4}\right]
\end{aligned}
$$

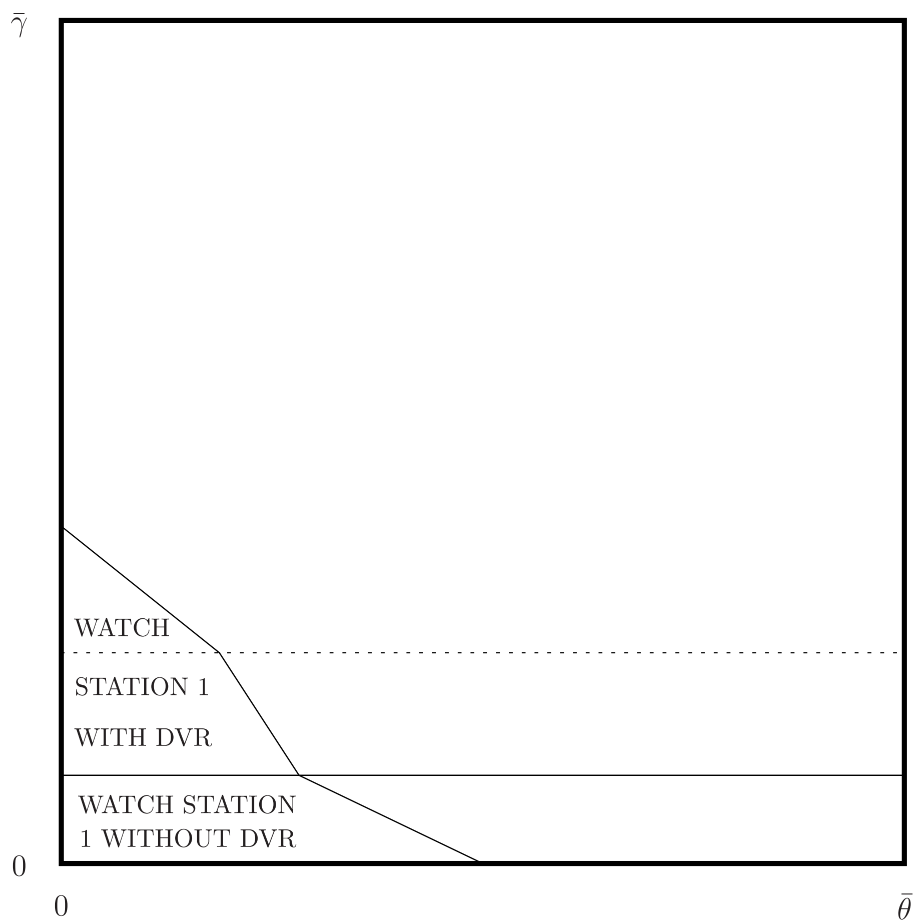


Figure B.10: Potential Viewer Response Map

As network 1 increases its ad level further, this expression continues to represent the firm's profit until the top indifference set intersects the top left corner of the box, which occurs when:

$$\begin{aligned}
\frac{\bar{\theta}}{2} + \frac{r}{2} \cdot \frac{\alpha}{1-\alpha} \cdot \left(\frac{a^* - a_1}{a^*}\right) &= 0 \\
\Longrightarrow \frac{\bar{\theta}}{r} \cdot \frac{1-\alpha}{\alpha} + 1 &= \frac{a_1}{a^*} \\
\Longrightarrow a_1 &= \frac{\bar{\theta}(1-\alpha) + \alpha r}{\alpha r} a^*
\end{aligned}$$

An example of such a situation is depicted in Figure B.10.

In this instance, the above expression underestimates the area of the DVR audience by the overspill of the top DVR-user trapezoid over the left edge of the box, or

$$\begin{aligned}
\frac{1}{2}\left(\frac{\bar{\gamma}}{2}\alpha(a_1-a^*)-\frac{\bar{\theta}}{2}\right)\left(\bar{\gamma}-\frac{\bar{\theta}}{\alpha(a_1-a^*)}\right) &= \frac{1}{4\alpha(a_1-a^*)}\left(\alpha\bar{\gamma}(a_1-a^*)-\bar{\theta}\right)^2 \\
&= \frac{\alpha\bar{\gamma}^2(a_1-a^*)}{4}-\frac{\bar{\gamma}\bar{\theta}}{2}+\frac{\bar{\theta}^2}{4\alpha(a_1-a^*)}
\end{aligned}$$

and therefore underestimates total impressions by this amount times $\alpha\frac{a_1}{\bar{\theta}\bar{\gamma}}$, or

$$\frac{1}{\bar{\theta}\bar{\gamma}}\left[\frac{\alpha^2\bar{\gamma}^2(a_1^2-a_1a^*)}{4}-\frac{\alpha\bar{\gamma}\bar{\theta}a_1}{2}+\frac{\bar{\theta}^2a_1}{4(a_1-a^*)}\right]$$

Adding this amount back to the above expression for profit yields a profit expression of:

$$\frac{1}{\bar{\theta}\bar{\gamma}}\left[\frac{2r\bar{\theta}-r^2}{4}+\frac{r^2a^*}{4(1-\alpha)a_1}-\frac{\alpha r^2a_1}{4a^*(1-\alpha)}+\frac{\bar{\theta}^2a_1}{4(a_1-a^*)}\right]$$

The next kink in the profit function occurs when network 1 increases its ad level to a point where its highest indifference set no longer intersects the indifference line representing those indifferent between watching network 2 with or without a DVR, which occurs when:

$$\begin{aligned}
\frac{\bar{\theta}}{2}+\frac{r}{2}\cdot\frac{\alpha}{1-\alpha}\cdot\left(1-\frac{a_1}{a^*}\right) &= 0 \\
\Longrightarrow \frac{\bar{\theta}}{r}\cdot\frac{1-\alpha}{\alpha}+1 &= \frac{a_1}{a^*} \\
\Longrightarrow a_1 &= \frac{\bar{\theta}(1-\alpha)+\alpha r}{\alpha r}a^*
\end{aligned}$$

An example is depicted in Figure B.11.

In this case, the same expression for the area of the non-DVR region from earlier

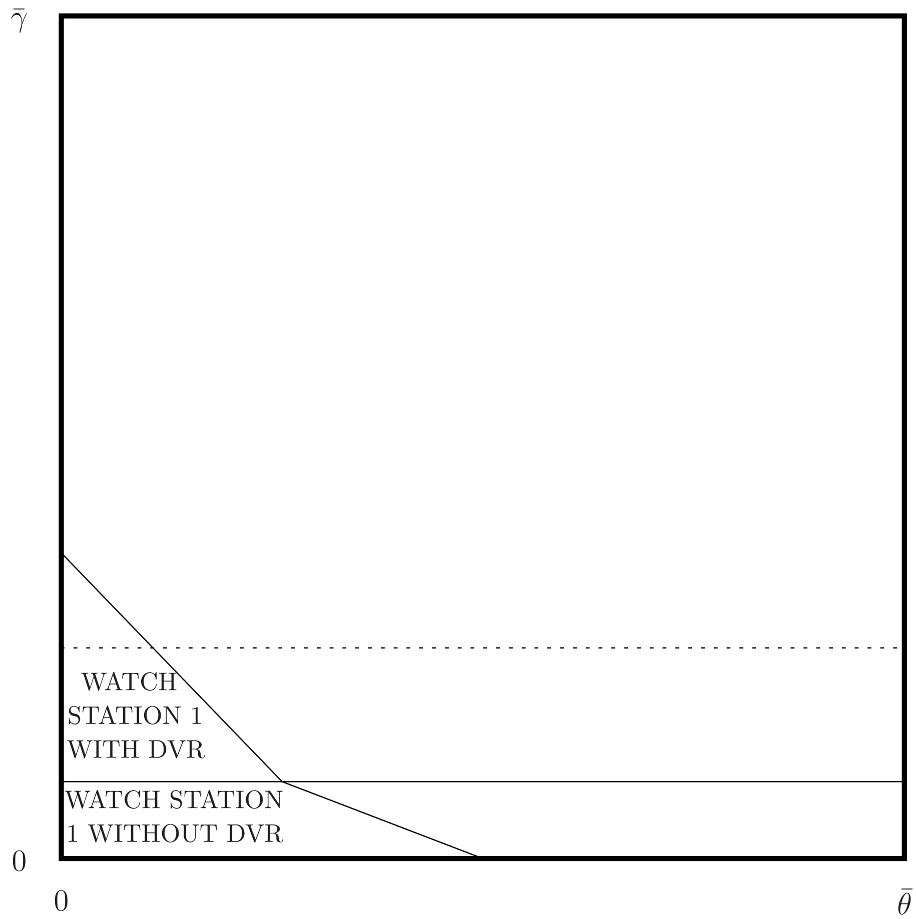


Figure B.11: Potential Viewer Response Map

still applies:

$$\frac{r\bar{\theta}}{2a_1(1-\alpha)} + \left(\frac{r}{2a_1(1-\alpha)}\right)^2 (a^* - a_1)$$

and the DVR audience region is now a triangle with area:

$$\begin{aligned}
&\frac{1}{2}\left(\frac{\bar{\theta}}{\alpha(a_1 - a^*)} - \frac{4}{a^*(1-\alpha)}\right)\left(\bar{\theta} - \frac{r}{2}\cdot\frac{1}{1-\alpha}\cdot\frac{a_1 - a^*}{a_1}\right) \\
&\quad = \frac{1}{4}\cdot\frac{1}{\alpha a_1 - a^*}\cdot\left(\bar{\theta} - \frac{ra_1 - \alpha r a_1 + \alpha r a_1 - ra^*}{a_1(1-\alpha)}\right)\left(\bar{\theta} - \frac{r}{1-\alpha}\left(\frac{a_1 - a^*}{a_1}\right)\right) \\
&\quad = \frac{1}{4}\cdot\frac{1}{\alpha a_1 - a^*}\cdot\left(\bar{\theta} - \frac{r}{1-\alpha}\left(\frac{a_1 - a^*}{a_1}\right)^2\right)
\end{aligned}$$

Total profit is therefore:

$$
\begin{aligned}
&\frac{a_1}{\bar{\theta}\bar{\gamma}}\left[\frac{r\bar{\theta}}{2a_1(1-\alpha)}+\left(\frac{r}{2a_1(1-\alpha)}\right)^2(a^*-a_1)+\frac{\alpha}{4}\cdot\frac{1}{\alpha a_1-a^*}\cdot\left(\bar{\theta}-\frac{r}{1-\alpha}\left(\frac{a_1-a^*}{a_1}\right)^2\right)\right]\\
&\quad=\frac{a_1}{\bar{\theta}\bar{\gamma}}\left[\frac{r\bar{\theta}}{2a_1(1-\alpha)}+\frac{r^2(a^*-a_1)}{4a_1^2(1-\alpha)^2}+\frac{\alpha}{4(\alpha a_1-a^*)}\right.\\
&\qquad\left.\cdot\left(\bar{\theta}^2-2\frac{r}{1-\alpha}\cdot\frac{a_1-a^*}{a_1}\bar{\theta}+\frac{r^2}{(1-\alpha)^2}\cdot\frac{(a_1-a^*)^2}{a_1^2}\right)\right]\\
&\quad=\frac{1}{\bar{\theta}\bar{\gamma}}\left[\frac{r\bar{\theta}}{2(1-\alpha)}+\frac{r^2a^*}{4a_1(1-\alpha)^2}-\frac{r^2}{4(1-\alpha)^2}+\frac{\alpha a_1}{4(\alpha a_1-a^*)}\right.\\
&\qquad\left.\cdot\left(\bar{\theta}^2-2\frac{r}{1-\alpha}\bar{\theta}+\frac{2ra^*}{(1-\alpha)a_1}\bar{\theta}+\frac{r^2}{(1-\alpha)^2}-\frac{2r^2a^*}{(1-\alpha)^2a_1}+\frac{r^2a^{*2}}{(1-\alpha)a_1^2}\right)\right]
\end{aligned}
$$

Note that

$$
\lim_{a_1\to\infty}\frac{\bar{\theta}}{2}+\frac{r}{2}\cdot\frac{1}{1-\alpha}\left(\frac{a^*}{a_1}-1\right)=\frac{\bar{\theta}}{2}-\frac{r}{2}\cdot\frac{1}{1-\alpha}
$$

which must be positive due to the assumption

$$
\bar{\theta}>\frac{r}{1-\alpha}
$$

This implies that the intersection of the two indifference lines that form the trapezoid occur inside the box regardless of how large network 1 sets its ad level. Consequently, the above expression for profit applies as a_1 gets arbitrarily large.

Putting all the cases together, this gives us the profit function:

$$\pi_{j|a_k=a^*} = \begin{cases} a_j - \dfrac{1}{\bar\theta\bar\gamma}\left(\dfrac{a_j}{\alpha a^* - a_j}\left(\dfrac{\bar\theta - r}{2}\right)^2 + \dfrac{r^2 a_j}{4a^*(1-\alpha)}\right) \equiv \pi_j^1 \\ \qquad\qquad \text{for } 0 \le a_j \le \frac{r}{\bar\gamma(1-\alpha)} \\ \alpha a_j - \dfrac{1}{\bar\theta\bar\gamma}\left(\dfrac{a_j}{\alpha a^* - a_j}\left(\dfrac{\bar\theta - r}{2}\right)^2 + \dfrac{r^2 a_j}{4a^*(1-\alpha)} - r\bar\theta\right) \equiv \pi_j^2 \\ \qquad\qquad \text{for } \frac{r}{\bar\gamma(1-\alpha)} \le a_j \le \frac{r\alpha a^*}{\bar\theta(1-\alpha)+r\alpha} \\ \dfrac{1}{\bar\theta\bar\gamma}\left[\dfrac{r^2}{4} + \dfrac{r\bar\theta}{2} + \dfrac{\alpha r^2 a^*}{4(1-\alpha)a_j} + \alpha a_j\bar\theta\bar\gamma - \dfrac{r^2 a_j}{4a^*(1-\alpha)} - \dfrac{\bar\theta^2 a_j}{4(a^* - a_j)}\right] \equiv \pi_j^3 \\ \qquad\qquad \text{for } \frac{r\alpha a^*}{\bar\theta(1-\alpha)+r\alpha} \le a_j \le a^* - \frac{\bar\theta}{\alpha\bar\gamma} \\ \dfrac{1}{\bar\theta\bar\gamma}\left[\dfrac{r^2 + 2r\bar\theta}{4} + \dfrac{\alpha r^2 a^*}{4(1-\alpha)a_j} + \dfrac{\alpha^2\bar\gamma^2 a_j a^*}{4} + \dfrac{\alpha a_j\bar\theta\bar\gamma}{2} - \dfrac{r^2 a_j}{4a^*(1-\alpha)} - \dfrac{\alpha^2\bar\gamma^2 a_j^2}{4}\right] \equiv \pi_j^4 \\ \qquad\qquad \text{for } a^* - \frac{\bar\theta}{\alpha\bar\gamma} \le a_j \le a^* \\ \dfrac{1}{\bar\theta\bar\gamma}\left[\dfrac{2r\bar\theta - r^2}{4} + \dfrac{r^2 a^*}{4(1-\alpha)a_j} + \dfrac{\alpha^2\bar\gamma^2 a_j a^*}{4} + \dfrac{\alpha a_j\bar\theta\bar\gamma}{2} - \dfrac{\alpha r^2 a_j}{4a^*(1-\alpha)} - \dfrac{\alpha^2\bar\gamma^2 a_j^2}{4}\right] \equiv \pi_j^5 \\ \qquad\qquad \text{for } a^* \le a_j \le a^* + \frac{\bar\theta}{\alpha\bar\gamma} \\ \dfrac{1}{\bar\theta\bar\gamma}\left[\dfrac{2r\bar\theta - r^2}{4} + \dfrac{r^2 a^*}{4(1-\alpha)a_j} - \dfrac{\alpha r^2 a_j}{4a^*(1-\alpha)} + \dfrac{\bar\theta^2 a_j}{4(a_j - a^*)}\right] \equiv \pi_j^6 \\ \qquad\qquad \text{for } a^* + \frac{\bar\theta}{\alpha\bar\gamma} \le a_j \le \frac{\bar\theta(1-\alpha)+\alpha r}{\alpha r}a^* \\ \dfrac{1}{\bar\theta\bar\gamma}\left[\dfrac{r\bar\theta}{2(1-\alpha)} + \dfrac{r^2 a^*}{4a_j(1-\alpha)^2} - \dfrac{r^2}{4(1-\alpha)^2} + \dfrac{\alpha a_j}{4(\alpha a_j - a^*)}\right. \\ \qquad\qquad \left.\cdot\left(\bar\theta^2 - 2\frac{r}{1-\alpha}\bar\theta + \frac{2ra^*}{(1-\alpha)a_j}\bar\theta + \frac{r^2}{(1-\alpha)^2} - \frac{2r^2a^*}{(1-\alpha)^2 a_j} + \frac{r^2 a^{*2}}{(1-\alpha)a_j^2}\right)\right] \equiv \pi_j^7 \\ \qquad\qquad \text{for } \frac{\bar\theta(1-\alpha)+\alpha r}{\alpha r}a^* < a_j \end{cases}$$

The only thing left to establish is that the kink points occur in the order described above. First, note that

$$r > 0, \alpha < 1 \Longrightarrow 0 < \frac{r}{\bar{\gamma}(1-\alpha)}$$

Secondly,

$$\begin{aligned}
\bar{\theta} &> \frac{r}{1-\alpha} \\
\Longrightarrow \bar{\theta}^2 &> \frac{\alpha^2}{(1-\alpha)^2}r^2 + \frac{1-\alpha^2}{(1-\alpha)^2}r^2 \\
\Longrightarrow \frac{\bar{\theta}^2}{\alpha^2\bar{\gamma}^2} &> \frac{r^2}{\bar{\gamma}^2(1-\alpha^2)} + \frac{(1+\alpha)}{(1-\alpha)}\frac{r^2}{\bar{\gamma}^2\alpha^2} \\
\Longrightarrow \sqrt{\left(\frac{\bar{\theta}}{\alpha\bar{\gamma}}\right)^2 - \frac{(1+\alpha)}{(1-\alpha)}\frac{r^2}{\bar{\gamma}^2\alpha^2}} &> \frac{r}{\bar{\gamma}(1-\alpha)} \\
\Longrightarrow a^* > \frac{\bar{\theta}}{\alpha\bar{\gamma}} + \frac{r}{\bar{\gamma}(1-\alpha)}
\end{aligned}$$

from which we know that

$$\begin{aligned}
\frac{\bar{\theta}}{\alpha\bar{\gamma}} + \frac{r}{\bar{\gamma}(1-\alpha)} &< a^* \\
\Longrightarrow \alpha\bar{\gamma}(1-\alpha)a^* &> \bar{\theta}(1-\alpha) + r\alpha \\
\Longrightarrow \frac{r}{\bar{\gamma}(1-\alpha)} &< \frac{r\alpha}{\bar{\theta}(1-\alpha) + r\alpha}a^*
\end{aligned}$$

and

$$\begin{aligned}
\frac{\bar{\theta}}{\alpha\bar{\gamma}} + \frac{r}{\bar{\gamma}(1-\alpha)} &< a^* \\
\Longrightarrow \frac{\bar{\theta}}{\alpha\bar{\gamma}} &< \frac{\bar{\theta}(1-\alpha)}{\bar{\theta}(1-\alpha)+\alpha r}a^* \\
\Longrightarrow \frac{\alpha r}{\bar{\theta}(1-\alpha)+\alpha r}a^* &< a^* - \frac{\bar{\theta}}{\alpha\bar{\gamma}} \\
\Longrightarrow \frac{r\alpha a^*}{\bar{\theta}(1-\alpha)+r\alpha} &< a^* - \frac{\bar{\theta}}{\alpha\bar{\gamma}}
\end{aligned}$$

Lastly,

$$\alpha, \bar{\theta}, \bar{\gamma} > 0 \Longrightarrow a^* - \frac{\bar{\theta}}{\alpha\bar{\gamma}} < a^* < a^* + \frac{\bar{\theta}}{\alpha\bar{\gamma}}$$

and

$$\begin{aligned}
\frac{r}{(1-\alpha)a^*} &< \bar{\gamma} \\
\Longrightarrow \frac{\bar{\theta}}{\alpha\bar{\gamma}} &< \frac{\bar{\theta}(1-\alpha)a^*}{\alpha r} \\
\Longrightarrow a^* + \frac{\bar{\theta}}{\alpha\bar{\gamma}} &< \frac{\bar{\theta}(1-\alpha)+\alpha r}{\alpha r}a^*
\end{aligned}$$

where the proof of the first line was demonstrated earlier in this section. Hence, all the kinked points occur in the correct order, so the function for profit given the rival network airs $a_k = a^*$ ads has been established.

B.2.3 Proof of Equilibrium

Proof Now that the profit function for one firm given the other plays $a_k = a^*$ has been derived, I demonstrate that it is not possible for a firm to achieve a profit greater than π^* by playing any other ad level itself. Note that $a_1 = a_2 = a^*$ occurs at the kink formed by the intersection of π_j^4 and π_j^5. First, I establish that this point constitutes a maximum for both of these functions over the ranges in which they are relevant.

For π_j^4, the first-order condition is:

$$\frac{d\pi_j^4}{da_j} = \frac{1}{\bar{\theta}\bar{\gamma}}\left[\frac{-\alpha r^2 a^*}{4(1-\alpha)a_j^2} + \frac{\alpha^2\bar{\gamma}^2 a^* + 2\alpha\bar{\theta}\bar{\gamma}}{4} - \frac{r^2}{4a^*(1-\alpha)} - \frac{\alpha^2\bar{\gamma}^2 a_j}{2}\right] = 0$$

$$\implies \frac{\alpha^2\bar{\gamma}^2(2a_j^3 - a^* a_j)}{2} - \frac{\alpha\bar{\theta}\bar{\gamma}a_j^2}{2} + \frac{r^2 a_j^2}{4a^*(1-\alpha)} + \frac{\alpha r^2 a^*}{4(1-\alpha)} = 0$$

This yields:

$$\begin{aligned} a^*\left[\frac{\alpha^2\bar{\gamma}^2(a^{*2})}{2} - \frac{\alpha\bar{\theta}\bar{\gamma}a^*}{2} + \frac{r^2 a^*}{4(1-\alpha)} + \frac{\alpha r^2}{4(1-\alpha)}\right] &= 0 \\ a^*\left[a^{*2} - \frac{2\bar{\theta}}{\alpha\bar{\gamma}}a^* + \frac{1+\alpha}{1-\alpha}\frac{r^2}{\alpha^2\bar{\gamma}^2}\right] &= 0 \end{aligned}$$

which yields the solution pair

$$a^* = \frac{1}{\alpha\bar{\gamma}}\left[\bar{\theta} \pm \sqrt{\bar{\theta}^2 - \frac{1+\alpha}{1-\alpha}r^2}\right]$$

since $a_j > 0$.

In order to confirm this is a maximum, one must check the second order condition:

$$\frac{d^2\pi_j^4}{da_j^2} = \frac{1}{\bar{\theta}\bar{\gamma}}\left[\frac{\alpha r^2 a^*}{2(1-\alpha)a_j^3} - \frac{\alpha^2\bar{\gamma}^2}{2}\right]$$

which, evaluated at $a_j = a^*$, yields the expression

$$\frac{1}{\bar{\theta}\bar{\gamma}}\left[\frac{\alpha r^2}{2(1-\alpha)a^{*2}} - \frac{\alpha^2\bar{\gamma}^2}{2}\right]$$

To show this holds for the candidate equilibrium:

$$
\begin{aligned}
\frac{1}{1-\alpha} &> \mp 2\sqrt{\frac{\alpha}{1-\alpha}\left(\frac{\bar{\theta}^2}{r^2} - \frac{1+\alpha}{1-\alpha}\right)} \\
\Longrightarrow \frac{\bar{\theta}^2}{r^2} &> \frac{\alpha}{1-\alpha} + \frac{\bar{\theta}^2}{r^2} - \frac{1+\alpha}{1-\alpha} \mp 2\sqrt{\frac{\alpha}{1-\alpha}\left(\frac{\bar{\theta}^2}{r^2} - \frac{1+\alpha}{1-\alpha}\right)} \\
\Longrightarrow \frac{\bar{\theta}}{r} &> \sqrt{\frac{\alpha}{1-\alpha}} \mp \sqrt{\frac{\bar{\theta}^2}{r^2} - \frac{1+\alpha}{1-\alpha}} \\
\Longrightarrow \frac{\bar{\theta}}{r} \pm \sqrt{\frac{\bar{\theta}^2}{r^2} - \frac{1+\alpha}{1-\alpha}} &> \sqrt{\frac{\alpha}{1-\alpha}} \\
\Longrightarrow \frac{1}{\alpha\bar{\gamma}}\left[\bar{\theta} \pm \sqrt{\bar{\theta}^2 - \frac{1+\alpha}{1-\alpha}r^2}\right] &> \frac{r}{\bar{\gamma}}\sqrt{\frac{1}{\alpha(1-\alpha)}} \\
\Longrightarrow a^* &> \frac{r}{\bar{\gamma}}\sqrt{\frac{1}{\alpha(1-\alpha)}} \\
\Longrightarrow a^{*2} &> \frac{r^2}{\alpha(1-\alpha)\bar{\gamma}^2} \\
\Longrightarrow \frac{\alpha^2\bar{\gamma}^2}{2} &> \frac{\alpha r^2}{2(1-\alpha)a^{*2}} \\
\Longrightarrow \frac{1}{\bar{\theta}\bar{\gamma}}\left[\frac{\alpha r^2}{2(1-\alpha)a^{*2}} - \frac{\alpha^2\bar{\gamma}^2}{2}\right] &< 0
\end{aligned}
$$

since the right-hand side of the penultimate expression must be greater than zero. This implies:

$$
\begin{aligned}
\Longrightarrow \frac{\alpha^2\bar{\gamma}^2}{2} &> \frac{\alpha r^2}{(1-\alpha)a^{*2}} \\
\Longrightarrow \frac{\alpha r^2}{(1-\alpha)a^{*2}} - \frac{\alpha^2\bar{\gamma}^2}{2} &< 0
\end{aligned}
$$

For our candidate equilibrium

$$
a^* = \frac{1}{\alpha\bar{\gamma}}\left[\bar{\theta} + \sqrt{\bar{\theta}^2 - \frac{1+\alpha}{1-\alpha}r^2}\right]
$$

the first line of the proof is automatically satisified, so that the SOC holds.[1]

For π_j^5:

$$\frac{d\pi_j^5}{da_j} = \frac{1}{\bar{\theta}\bar{\gamma}}\left[\frac{-r^2a^*}{4(1-\alpha)a_j^2} + \frac{\alpha^2\bar{\gamma}^2a^* + 2\alpha\bar{\theta}\bar{\gamma}}{4} - \frac{\alpha r^2}{4a^*(1-\alpha)} - \frac{\alpha^2\bar{\gamma}^2a_j}{2}\right] = 0$$
$$\implies \frac{\alpha^2\bar{\gamma}^2(2a_j^3 - a^*a_j)}{2} - \frac{\alpha\bar{\theta}\bar{\gamma}a_j^2}{2} + \frac{\alpha r^2a_j^2}{4a^*(1-\alpha)} + \frac{r^2a^*}{4(1-\alpha)} = 0$$

which, evaluated at $a_j = a^*$, yields the same expression as before:

$$a^*\left[\frac{\alpha^2\bar{\gamma}^2(a^{*2})}{2} - \frac{\alpha\bar{\theta}\bar{\gamma}a^*}{2} + \frac{r^2a^*}{4(1-\alpha)} + \frac{\alpha r^2}{4(1-\alpha)}\right] = 0$$

so the first-order condition is satisfied.

One must again check the second order condition:

$$\frac{d^2\pi_j^4}{da_j^2} = \frac{1}{\bar{\theta}\bar{\gamma}}\left[\frac{r^2a^*}{2(1-\alpha)a_j^3} - \frac{\alpha^2\bar{\gamma}^2}{2}\right] < 0$$

which, evaluated at $a_j = a^*$, yields

$$\frac{1}{\bar{\theta}\bar{\gamma}}\left[\frac{r^2}{2(1-\alpha)a^{*2}} - \frac{\alpha^2\bar{\gamma}^2}{2}\right] < 0$$

[1]Note that, for the root

$$a^* = \frac{1}{\alpha\bar{\gamma}}\left[\bar{\theta} - \sqrt{\bar{\theta}^2 - \frac{1+\alpha}{1-\alpha}r^2}\right]$$

further parameter restrictions would be necessary for the SOC to be satisfied. Indeed, the restrictions would need be such that r and α were large relative to θ, or in other words the DVR's price-efficiency combination would have to be at least a threshold level of unattractive.

To show this holds:

$$
\begin{aligned}
\frac{\alpha}{1-\alpha} &> \mp 2\sqrt{\frac{1}{1-\alpha}\left(\frac{\bar{\theta}^2}{r^2}-\frac{1+\alpha}{1-\alpha}\right)} \\
\implies \frac{\bar{\theta}^2}{r^2} &> \frac{1}{1-\alpha}+\frac{\bar{\theta}^2}{r^2}-\frac{1+\alpha}{1-\alpha}\mp 2\sqrt{\frac{1}{1-\alpha}\left(\frac{\bar{\theta}^2}{r^2}-\frac{1+\alpha}{1-\alpha}\right)} \\
\implies \frac{\bar{\theta}}{r} &> \sqrt{\frac{1}{1-\alpha}}\mp\sqrt{\frac{\bar{\theta}^2}{r^2}-\frac{1+\alpha}{1-\alpha}} \\
\implies \frac{\bar{\theta}}{r}\pm\sqrt{\frac{\bar{\theta}^2}{r^2}-\frac{1+\alpha}{1-\alpha}} &> \sqrt{\frac{1}{1-\alpha}} \\
\implies \frac{1}{\alpha\bar{\gamma}}\left[\bar{\theta}\pm\sqrt{\bar{\theta}^2-\frac{1+\alpha}{1-\alpha}r^2}\right] &> \frac{r}{\alpha\bar{\gamma}}\sqrt{\frac{1}{\alpha(1-\alpha)}} \\
\implies a^* &> \frac{r}{\alpha\bar{\gamma}}\sqrt{\frac{1}{(1-\alpha)}} \\
\implies a^{*2} &> \frac{r^2}{\alpha^2(1-\alpha)\bar{\gamma}^2} \\
\implies \frac{\alpha^2\bar{\gamma}^2}{2} &> \frac{r^2}{(1-\alpha)a^{*2}} \\
\implies \frac{r^2}{(1-\alpha)a^{*2}}-\frac{\alpha^2\bar{\gamma}^2}{2} &< 0
\end{aligned}
$$

This proves that the candidate equilibrium is a local maximum for both π_j^4 and π_j^5 at the intersection of the relevant ranges for each. To show it is a maximum for the entirity of $\left(a^*-\frac{\bar{\theta}}{\alpha\bar{\gamma}}, a^*+\frac{\bar{\theta}}{\alpha\bar{\gamma}}\right)$, note that both segments of the profit function have form

$$-A\cdot a_j^2+Ba_j+\frac{C}{a_j}+D$$

with $A, B, C, D > 0$. Consequently, these functions can have at most three critical points. Further, as $a_j \to \pm\infty$, both functions approach $-\infty$, and as $a_j \to 0$ from the left and right, both functions approach $-\infty$ and ∞ respectively. There must therefore be one critical point for some $a_j < 0$, and since we know $a_j = a^*$ constitutes a local

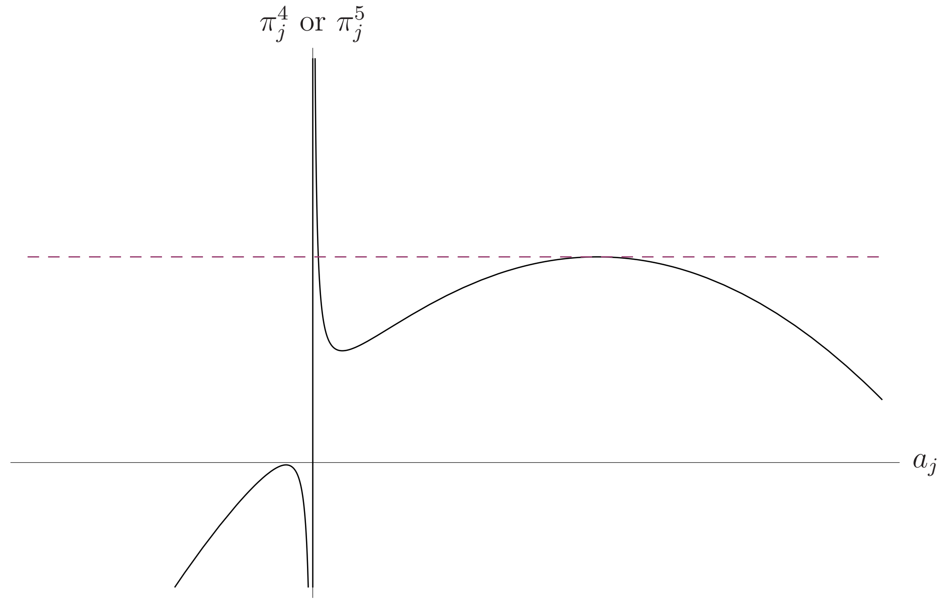


Figure B.12: Example Graph of π_j^4 or π_j^5

maximum, there can only one other critical point, which must be a local minimum for some $a_j < a^*$. Figure B.12 displays an example of what one of these functions might look like.

This implies that π_j^5 stays below π^* for all $a_j > a^*$, and hence the profit function does not exceed a^* for any a_j in π_j^5's relevant range. However, it is still possible that it is not a maximum over π_j^4's range, as, going leftwards from $a_j = a^*$, the π_j^4 function could reach a minimum within that range and come back up to exceed π^* before its relevant range is over.

Note, however, that the nature of the function's shape implies that if $\pi_j^4 < \pi^*$ for some $0 < a_j < a^*$, then there is no $a_j'' > a_j'$ such that $\pi_j^4(a_j'') > \pi^*$. Hence, all that is required is to find an a_j' below the relevant range of π_j^4 such that $\pi_j^4(a_j') < \pi^*$, and this is sufficient to conclude that the profit function does not exceed π^* within π_j^4's relevant range, either.

Consider

$$a'_j = \frac{\alpha r}{\bar{\theta}} a^*$$

which is actually below the start of π_j^3's relevant range, $\frac{\alpha r}{\bar{\theta}(1-\alpha)+r\alpha} a^*$, since $r < \bar{\theta}$. The difference between π^* and π_j^4 evaluated at this point is:

$$\begin{aligned}
\pi^* - \pi_j^4\left(a'_j\right) &= \frac{1}{\bar{\theta}\bar{\gamma}}\Bigg[\left(\frac{r\bar{\theta}}{2} + \frac{\alpha\bar{\theta}\bar{\gamma}a^*}{2}\right) \\
&\quad - \left(\frac{\alpha^2\bar{\gamma}^2}{4}a^{*2}\left(\frac{\alpha r}{\bar{\theta}} - \frac{\alpha^2 r^2}{\bar{\theta}}\right) + \frac{\alpha^2 r\bar{\gamma}}{2}a^* - \frac{\alpha r^3}{4\bar{\theta}(1-\alpha)} + \frac{r^2 + 2r\bar{\theta}}{4} + \frac{r\bar{\theta}}{4(1-\alpha)}\right)\Bigg] \\
&= \frac{1}{\bar{\theta}\bar{\gamma}} \cdot \frac{1}{4\bar{\theta}^2(1-\alpha)}\Big[\big(2r\bar{\theta}^3(1-\alpha) + 2\alpha(1-\alpha)\bar{\theta}^3\bar{\gamma}a^* - \alpha^3 r\bar{\gamma}^2(1-\alpha)a^{*2}(\bar{\theta} - r\alpha) \\
&\quad - 2\alpha^2(1-\alpha)r\bar{\gamma}\bar{\theta}^2 a^* + \alpha r^3\bar{\theta} - r^2\bar{\theta}^2(1-\alpha) - 2r\bar{\theta}^3(1-\alpha) - r\bar{\theta}^3\big)\Big] \\
&= \frac{1}{\bar{\theta}\bar{\gamma}} \cdot \frac{\bar{\theta} - \alpha r}{4\bar{\theta}^2(1-\alpha)}\left[-\alpha^3 r\bar{\gamma}^2(1-\alpha)a^{*2} + 2\alpha(1-\alpha)\bar{\theta}^2\bar{\gamma}a^* - (r^2\bar{\theta} + r\bar{\theta}^2)\right] \\
&= \frac{1}{\bar{\theta}\bar{\gamma}} \cdot \frac{\bar{\theta} - \alpha r}{4\bar{\theta}^2(1-\alpha)}\Big[-\alpha r(1-\alpha)\left(\alpha\bar{\gamma}a^*\right)^2 + 2\alpha(1-\alpha)^2\bar{\theta}^2(\alpha\bar{\gamma}a^*) \\
&\quad + 2(1-\alpha)(1-\alpha+\alpha^2)\bar{\theta}^2(\alpha\bar{\gamma}a^*) - (r^2\bar{\theta} + r\bar{\theta}^2)\Big] \\
&= \frac{1}{\bar{\theta}\bar{\gamma}} \cdot \frac{\bar{\theta} - \alpha r}{4\bar{\theta}^2(1-\alpha)}\Big[\alpha(1-\alpha)(\alpha\bar{\gamma}a^*)\left(2\bar{\theta}^2(1-\alpha) - r\alpha\bar{\gamma}a^*\right) \\
&\quad + 2(1-\alpha)(1-\alpha+\alpha^2)\bar{\theta}^2(\alpha\bar{\gamma}a^*) - (r^2\bar{\theta} + r\bar{\theta}^2)\Big]
\end{aligned}$$

Since $r < \bar{\theta}(1-\alpha)$ and $\alpha\bar{\gamma}a^* < 2\bar{\theta}^2$, the first term inside the bracket is positive.

Further:

$$
\begin{aligned}
&2(1-\alpha)(1-\alpha+\alpha^2)\bar{\theta}^2(\alpha\bar{\gamma}a^*) \\
&\quad = 2(1-\alpha)(1-\alpha+\alpha^2)\bar{\theta}^2(\bar{\theta}+\sqrt{\bar{\theta}^2-\frac{1+\alpha}{1-\alpha}r^2}) \\
&\quad > 2(1-\alpha)(1-\alpha+\alpha^2)\bar{\theta}^2\left(\bar{\theta}+\sqrt{\frac{1}{(1-\alpha)^2}r^2-\frac{1+\alpha}{1-\alpha}r^2}\right) \\
&\quad = 2(1-\alpha)(1-\alpha+\alpha^2)\bar{\theta}^2\left(\bar{\theta}+\frac{\alpha}{1-\alpha}r\right) \\
&\quad = 2(1-\alpha)(1-\alpha+\alpha^2)\bar{\theta}^3+2\alpha(1-\alpha+\alpha^2)\bar{\theta}^2 r \\
&\quad > 2(1-\alpha+\alpha^2)\bar{\theta}^2 r+2\alpha(1-\alpha+\alpha^2)\bar{\theta}^2 r \\
&\quad = 2(1+\alpha^3)\bar{\theta}^2 r \\
&\quad > r^2\bar{\theta}+r\bar{\theta}^2
\end{aligned}
$$

so that the second term in the parentheses is also positive. Since $\bar{\theta} > \alpha r$, the coefficients are also positive, and hence the term itself must be positive. This implies that $\pi^* > \pi_j^4(a_j')$ and hence that $\pi^* > \pi_j^4(a_j)\ \forall a_j \in \left(\frac{\alpha r}{\bar{\theta}}a^*\right)$.

Further, consider $\pi_j^4 - \pi_j^3$:

$$
\begin{aligned}
&\pi_j^4-\pi_j^3 \\
&\quad = \frac{1}{\bar{\theta}\bar{\gamma}}\left[\frac{r^2+2r\bar{\theta}}{4}+\frac{\alpha r^2 a^*}{4(1-\alpha)a_j}+\frac{\alpha^2\bar{\gamma}^2 a_j a^*}{4}+\frac{\alpha a_j\bar{\theta}\bar{\gamma}}{2}-\frac{r^2 a_j}{4a^*(1-\alpha)}-\frac{\alpha^2\bar{\gamma}^2 a_j^2}{4}\right] \\
&\qquad -\frac{1}{\bar{\theta}\bar{\gamma}}\left[\frac{r^2}{4}+\frac{r\bar{\theta}}{2}+\frac{\alpha r^2 a^*}{4(1-\alpha)a_j}+\alpha a_j\bar{\theta}\bar{\gamma}-\frac{r^2 a_j}{4a^*(1-\alpha)}-\frac{\bar{\theta}^2 a_j}{4(a^*-a_j)}\right] \\
&\quad = \frac{1}{\bar{\theta}\bar{\gamma}}\left[\frac{\alpha^2\bar{\gamma}^2 a_j}{4}(a^*-a_j)-\frac{\alpha a_j\bar{\theta}\bar{\gamma}}{2}+\frac{\bar{\theta}^2 a_j}{4(a^*-a_j)}\right] \\
&\quad = \frac{1}{4\bar{\gamma}\bar{\theta}(a^*-a_j)}\left[\alpha^2\bar{\gamma}^2 a_j(a^*-a_j)^2-2\alpha a_j\bar{\theta}\bar{\gamma}(a^*-a_j)+\bar{\theta}^2 a_j\right] \\
&\quad = \frac{a_j}{4\bar{\gamma}\bar{\theta}(a^*-a_j)}\left[\left(\bar{\theta}-\alpha\bar{\gamma}(a^*-a_j)\right)^2\right]
\end{aligned}
$$

which is positive for $a_j < a^*$. Therefore, evaluating profit as π_j^4 over the range in which π_j^3 is relevant actually overstates profit; since it was shown above that π_j^4 cannot exceed π^* for π_j^3's relevant range as well, it follows that the networks cannot do better than π^* by airing an ad level that falls within the relevant range of π_j^3 either.

So far, it has been established that a^* is a local maximum for both π_j^4 and π_j^5 that occurs at the intersection of the relevant ranges for each. It has also been shown that the resulting profit π^* cannot be exceeded by airing an ad level in the relevant ranges of either. In what follows, the same is proven for each of the remaining segments of the profit function in turn.

For π_j^1:

$$\pi_j^1 = a_j - \frac{1}{\bar{\theta}\bar{\gamma}}\left[\frac{a_j}{\alpha a^* - a_j}\left(\frac{\bar{\theta}-r}{2}\right)^2 + \frac{r^2 a_j}{4a^*(1-\alpha)}\right] \Longrightarrow \frac{d\pi_j^1}{da_j}$$
$$= \frac{1}{\bar{\theta}\bar{\gamma}}\left[\bar{\theta}\bar{\gamma} - \frac{\alpha a^*}{(\alpha a^* - a_j)^2}\left(\frac{\bar{\theta}-r}{2}\right)^2 - \frac{r^2}{4a^*(1-\alpha)}\right]$$

which is relevant on the range $a_j \le \frac{r}{\bar{\gamma}(1-\alpha)}$. Since $\alpha a^* > \frac{\bar{\theta}}{\bar{\gamma}}$ and $\bar{\theta} > \frac{r}{1-\alpha}$, $\alpha a^* > a_j$ for the entire range so that the derivative is decreasing in a_j. It will therefore reach its lowest point in this range for $a_j = \frac{r}{\bar{\gamma}(1-\alpha)}$.

Substituting this into the derivative gives us:

$$\frac{1}{\bar{\theta}\bar{\gamma}}\left[\bar{\theta}\bar{\gamma} - \frac{\alpha a^*}{(\alpha a^* - \frac{r}{\bar{\gamma}(1-\alpha)})^2}\left(\frac{\bar{\theta}-r}{2}\right)^2 - \frac{r^2}{4a^*(1-\alpha)}\right]$$
$$= \frac{1}{\bar{\theta}}\left[\bar{\theta} - \frac{\alpha\bar{\gamma}a^*}{(\alpha\bar{\gamma}a^* - \frac{r}{1-\alpha})^2}\left(\frac{\bar{\theta}-r}{2}\right)^2 - \frac{r^2}{4\bar{\gamma}a^*(1-\alpha)}\right]$$

Note that:

$$
\begin{aligned}
\bar{\theta} &> \frac{r}{1-\alpha} \\
\implies \bar{\theta}^2 &> \left(\frac{\alpha^2}{(1-\alpha)^2} + \frac{1+\alpha}{1-\alpha}\right) r^2 \\
\implies \sqrt{\bar{\theta}^2 - \frac{1+\alpha}{1-\alpha} r^2} &> \frac{\alpha}{1-\alpha} r \\
\implies \bar{\theta} + \sqrt{\bar{\theta}^2 - \frac{1+\alpha}{1-\alpha} - r^2} - \frac{r}{1-\alpha} &> \bar{\theta} - r \\
\implies \alpha a^* - \frac{r}{1-\alpha} &> \bar{\theta} - r
\end{aligned}
$$

so that

$$
\frac{1}{\bar{\bar{\theta}}}\left[\bar{\theta} - \frac{\alpha\bar{\gamma}a^*}{(\alpha\bar{\gamma}a^* - \frac{r}{1-\alpha})^2}\left(\frac{\bar{\theta}-r}{2}\right)^2 - \frac{r^2}{4\bar{\gamma}a^*(1-\alpha)}\right] > \frac{1}{\bar{\bar{\theta}}}\left[\bar{\theta} - \frac{\alpha\bar{\gamma}a^*}{4} - \frac{r^2}{4\bar{\gamma}a^*(1-\alpha)}\right]
$$

Lastly:

$$
\begin{aligned}
&\frac{2\bar{\theta}\left(\bar{\theta} + \sqrt{\bar{\theta}^2 - \frac{1+\alpha}{1-\alpha} - r^2}\right) - \left(\bar{\theta} + \sqrt{\bar{\theta}^2 - \frac{1+\alpha}{1-\alpha} - r^2}\right)^2}{4} \\
&\qquad + \left(\frac{2\bar{\theta}\left(\bar{\theta} + \sqrt{\bar{\theta}^2 - \frac{1+\alpha}{1-\alpha} - r^2}\right) - \frac{r}{1-\alpha} \cdot \alpha \cdot r}{4}\right) > 0 \\
&\implies \bar{\theta}\left(\bar{\theta} + \sqrt{\bar{\theta}^2 - \frac{1+\alpha}{1-\alpha} - r^2}\right) \\
&\qquad - \frac{\left(\bar{\theta} + \sqrt{\bar{\theta}^2 - \frac{1+\alpha}{1-\alpha} - r^2}\right)^2}{4} - \frac{\alpha r^2}{4(1-\alpha)} > 0 \\
&\implies \bar{\theta} - \frac{\left(\bar{\theta} + \sqrt{\bar{\theta}^2 - \frac{1+\alpha}{1-\alpha} - r^2}\right)}{4} - \frac{r^2}{4\bar{\gamma}a^*(1-\alpha)} > 0 \\
&\implies \frac{1}{\bar{\bar{\theta}}}\left[\bar{\theta} - \frac{\alpha\bar{\gamma}a^*}{4} - \frac{r^2}{4\bar{\gamma}a^*(1-\alpha)}\right] > 0
\end{aligned}
$$

Hence, the initial derivative is positive for the relevant range of $\frac{d\pi_j}{da_j}$, implying the network has an incentive to increase its ad level throughout this range.

Now consider the difference between π^* and π_j^2:

$$\begin{aligned}\pi^* - \pi_j^2 &= \frac{1}{\bar{\theta}\bar{\gamma}}\left[\frac{r\bar{\theta}}{2} + \frac{\alpha\bar{\theta}\bar{\gamma}a^*}{2} - \alpha\bar{\theta}\bar{\gamma}a_j - \bar{\theta}r\right. \\ &\qquad \left. + \left(\frac{a_j}{\alpha a^* - a_j}\right)\left(\frac{\bar{\theta} - r}{2}\right)^2 + \frac{r^2 a_j}{4(1-\alpha)a^*}\right] \\ &= \frac{1}{\bar{\theta}\bar{\gamma}}\left[-\frac{r\bar{\theta}}{2} + \alpha\bar{\theta}\bar{\gamma}\left(\frac{a^*}{2} - a_j\right)\right. \\ &\qquad \left. + \left(\frac{a_j}{\alpha a^* - a_j}\right)\left(\frac{\bar{\theta} - r}{2}\right)^2 + \frac{r^2 a_j}{4(1-\alpha)a^*}\right] \end{aligned} \tag{B.2}$$

which is only relevant over the range

$$\frac{r}{\bar{\gamma}(1-\alpha)} \leq a_j \leq \frac{r\alpha}{\bar{\theta}(1-\alpha) + r\alpha}a^*$$

which implies

$$a_j \leq \frac{r\alpha}{\bar{\theta}(1-\alpha) + r\alpha}a^* < \frac{r\alpha}{r + r\alpha}a^* < \alpha a^*$$

so that the final two terms in brackets in B.2 above both positive. Further,

$$
\begin{aligned}
&-\frac{r\bar{\theta}}{2} + \alpha\bar{\theta}\bar{\gamma}\left(\frac{a^*}{2} - a_j\right) \\
&\quad \leq -\frac{r\bar{\theta}}{2} + \alpha\bar{\theta}\bar{\gamma}\left(\frac{a^*}{2} - \frac{r\alpha}{\bar{\theta}(1-\alpha)+r\alpha}a^*\right) \\
&\quad = -\frac{r\bar{\theta}}{2} + \alpha\bar{\theta}\bar{\gamma}\left(\frac{\bar{\theta}(1-\alpha)-r\alpha}{2(\bar{\theta}(1-\alpha)+r\alpha)}a^*\right) \\
&\quad < -\frac{r\bar{\theta}}{2} + \alpha\bar{\theta}\bar{\gamma}\left(\frac{r(1-\alpha)}{2(\bar{\theta}(1-\alpha)+r\alpha)}a^*\right) \\
&\quad = \frac{r\bar{\theta}}{2}\left(\frac{\bar{\theta}(1-\alpha)\left(\bar{\theta}+\sqrt{\bar{\theta}^2-\frac{1+\alpha}{1-\alpha}r^2}\right)}{(\bar{\theta}(1-\alpha)+r\alpha)} - 1\right) \\
&\quad = \frac{r\bar{\theta}}{2(\bar{\theta}(1-\alpha)+r\alpha)}\left(\bar{\theta}(1-\alpha)+(1-\alpha)\sqrt{\bar{\theta}^2-\frac{1+\alpha}{1-\alpha}r^2}-\bar{\theta}(1-\alpha)-\alpha r\right) \\
&\quad = \frac{r\bar{\theta}}{2(\bar{\theta}(1-\alpha)+r\alpha)}\left((1-\alpha)\sqrt{\bar{\theta}^2-\frac{1+\alpha}{1-\alpha}r^2}-\alpha r\right)
\end{aligned}
$$

which must be positive, since:

$$
\begin{aligned}
\bar{\theta} &> \frac{r}{1-\alpha} \\
\Longrightarrow \bar{\theta}^2 - \frac{1+\alpha}{1-\alpha}r^2 &> r^2\frac{\alpha^2}{1-\alpha^2} \\
\Longrightarrow (1-\alpha)\sqrt{\bar{\theta}^2 - \frac{1+\alpha}{1-\alpha}r^2} &> \alpha r
\end{aligned}
$$

Therefore the initial profit difference expression must be positive, so the networks cannot do better than equilibrium profit by airing an ad level in this range.

For π_j^6.

$$\begin{aligned}
\pi_j^6 &= \frac{1}{\bar{\theta}\bar{\gamma}}\left[\frac{2r\bar{\theta}-r^2}{4}+\frac{r^2a^*}{4(1-\alpha)a_j}-\frac{\alpha r^2 a_j}{4a^*(1-\alpha)}+\frac{\bar{\theta}^2a_j}{4(a_j-a^*)}\right]\\
\Longrightarrow \frac{d\pi_j^6}{da_j} &= \frac{1}{\bar{\theta}\bar{\gamma}}\left[-\frac{r^2a^*}{4(1-\alpha)a_j^2}-\frac{\alpha r^2}{4a^*(1-\alpha)}-\frac{\bar{\theta}^2a^*}{4(a_j-a^*)}\right]
\end{aligned}$$

for which each term in the brackets is individually negative. Hence, any firm considering airing an ad level in this range would be better served decreasing it throughout that range.

Lastly, for π_j^7:

$$\begin{aligned}
\pi_j^7 = {}& \frac{1}{\bar{\theta}\bar{\gamma}}\left[\frac{r\bar{\theta}}{2(1-\alpha)}+\frac{r^2a^*}{4a_j(1-\alpha)^2}-\frac{r^2}{4(1-\alpha)^2}+\frac{\alpha a_j}{4(\alpha a_j-a^*)}\right.\\
&\left.\cdot\left(\bar{\theta}^2-2\frac{r}{1-\alpha}\bar{\theta}+\frac{2ra^*}{(1-\alpha)a_j}\bar{\theta}+\frac{r^2}{(1-\alpha)^2}-\frac{2r^2a^*}{(1-\alpha)^2a_j}+\frac{r^2a^{*2}}{(1-\alpha)a_j^2}\right)\right]\\
\Longrightarrow \frac{d\pi_j^7}{da_j} = {}& \frac{1}{\bar{\theta}\bar{\gamma}}\left[-\frac{r^2a^*}{4a_j^2(1-\alpha)^2}+\frac{\alpha a_j}{4(\alpha a_j-a^*)}\left(-\frac{2ra^*}{(1-\alpha)a_j^2}\bar{\theta}+\frac{2r^2a^*}{(1-\alpha)^2a_j^2}-\frac{2r^2a^*}{(1-\alpha)a_1^3}\right)\right.\\
&\left.+\left(\frac{\alpha}{4(\alpha a_j-a^*)}-\frac{\alpha^2a_j}{4(\alpha a_j-a^*)^2}\right)\left(\bar{\theta}-\frac{r}{1-\alpha}\frac{a_j-a^*}{a_j}\right)^2\right]\\
= {}& \frac{1}{\bar{\theta}\bar{\gamma}}\left[-\frac{r^2a^*}{4a_j^2(1-\alpha)^2}-\frac{\alpha a_j}{4(\alpha a_j-a^*)}\left(\frac{2ra^*(\bar{\theta}-r)}{(1-\alpha)a_j^2}+\frac{2r^2a^{*2}}{(1-\alpha)a_j^3}\right)\right.\\
&\left.-\frac{\alpha a^*}{4(\alpha a_j-a^*)^2}\left(\bar{\theta}-\frac{r}{1-\alpha}\frac{a_j-a^*}{a_j}\right)^2\right]<0
\end{aligned}$$

so that a firm considering airing an ad level at any point throughout this range should again decrease its allotment.

It has therefore been shown that, given that the rival station is airing an ad level a^*, the profit function cannot exceed π^* for any level of ads, establishing the strategy $a_1 = a_2 = a^*$ an equilibrium. ∎

B.3 DVR Effect on Network Welfare

Proposition 12 *Each network's profit is increased with the advent of the DVR.*

Proof Equilibrium profit in the DVR case is:

$$\pi_D^* = \frac{1}{2\bar{\gamma}} \cdot [r + \alpha\bar{\gamma}a^*] = \frac{1}{2\bar{\gamma}} \cdot \left[r + \bar{\theta} + \sqrt{\bar{\theta}^2 - \frac{1+\alpha}{1-\alpha}r^2} \right]$$

and from section 4.2, equilbrium profit without a DVR is

$$\pi_N^* = \frac{\bar{\theta}}{\bar{\gamma}}$$

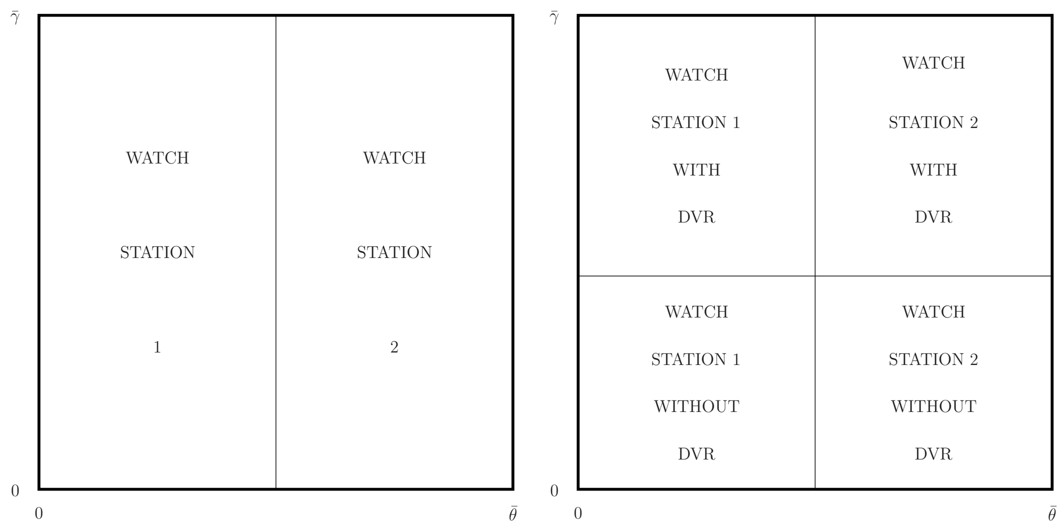
From the assumption $\bar{\theta} > \frac{r}{1-\alpha}$, we have:

$$\begin{aligned}
\bar{\theta} &> \frac{r}{1-\alpha} \\
\Longrightarrow 2\bar{\theta}r &> r^2\left(1 + \frac{1+\alpha}{1-\alpha}\right) \\
\Longrightarrow (\bar{\theta} - r)^2 &< \bar{\theta}^2 - \frac{1+a}{1-a}r^2 \\
\Longrightarrow \bar{\theta} - r &< \sqrt{\bar{\theta}^2 - \frac{1+a}{1-a}r^2} \\
\Longrightarrow 2\bar{\theta} &< r + \bar{\theta} + \sqrt{\bar{\theta}^2 - \frac{1+\alpha}{1-\alpha}r^2} \\
\Longrightarrow \frac{\bar{\theta}}{\bar{\gamma}} &< \frac{1}{2\bar{\gamma}} \cdot \left[r + \bar{\theta} + \sqrt{\bar{\theta}^2 - \frac{1+\alpha}{1-\alpha}r^2} \right] \\
\Longrightarrow \pi_N^* &< \pi_D^* \quad \blacksquare
\end{aligned}$$

B.4 DVR Effect on Viewer Welfare

Proposition 13 *The DVR's effect on viewer welfare is ambiguous.*

Proof More specifically, for $\alpha > \frac{1}{3}$ the viewers are helped by the DVR, whereas for $\alpha < \frac{2}{3}$, the viewers are helped by the DVR when $\bar{\theta} > \frac{5-3\alpha}{4(1-\alpha)}r$ and hurt by it otherwise. Figure B.13 displays the equilibrium outcomes both with and without DVR technology.



(a) Pre-DVR (b) Post-DVR

Figure B.13: Equilibrium Outcomes Before and After DVR Technology

I start by evaluating viewer welfare before the advent of the DVR. Since welfare is the same for the set of viewers watching network 1 as it is for those watching network

2, total viewer welfare is:

$$
\begin{aligned}
VW &= 2 \cdot \frac{1}{\bar{\theta}\bar{\gamma}} \int_0^{\bar{\gamma}} \int_0^{\frac{\bar{\theta}}{2}} \left(-\theta_i - \gamma_i \left(2\frac{\bar{\theta}}{\bar{\gamma}} \right) \right) d\theta_i d\gamma_i \\
&= 2 \cdot \frac{1}{\bar{\theta}\bar{\gamma}} \int_0^{\bar{\gamma}} \left(-\frac{1}{2}\theta_i^2 - 2\theta_i\gamma_i\frac{\bar{\theta}}{\bar{\gamma}} \right) \Big|_0^{\frac{\bar{\theta}}{2}} d\gamma_i \\
&= 2 \cdot \frac{1}{\bar{\theta}\bar{\gamma}} \int_0^{\bar{\gamma}} \left(-\frac{\theta_i^2}{8} - \frac{\bar{\theta}}{\bar{\gamma}}\bar{\theta}\gamma_i \right) d\gamma_i \\
&= 2 \cdot \frac{1}{\bar{\theta}\bar{\gamma}} \left(-\frac{\theta_i^2}{8} \cdot \gamma_i - \frac{1}{2}\frac{\bar{\theta}}{\bar{\gamma}}\bar{\theta}\gamma_i^2 \Big|_0^{\bar{\gamma}} \right) d\gamma_i \\
&= -\frac{5\bar{\theta}}{4}
\end{aligned}
$$

For viewer welfare after the advent of the DVR, I split the calculation into those who use a DVR (VW_D^1) and those who do not (VW_D^2):

$$
\begin{aligned}
VW_D^1 &= 2 \cdot \frac{1}{\bar{\theta}\bar{\gamma}} \int\limits_{\frac{r}{a^*(1-\alpha)}}^{\bar{\gamma}} \int\limits_{0}^{\frac{\bar{\theta}}{2}} \left(-\theta_i - \gamma_i \left(\alpha a^*\right) - r\right) d\theta_i d\gamma_i \\
&= 2 \cdot \frac{1}{\bar{\theta}\bar{\gamma}} \int\limits_{\frac{r}{a^*(1-\alpha)}}^{\bar{\gamma}} \left(-\frac{1}{2}\theta_i^2 - 2\alpha\theta_i\gamma_i a^* - r\theta_i\right)\Bigg|_0^{\frac{\bar{\theta}}{2}} d\gamma_i \\
&= 2 \cdot \frac{1}{\bar{\theta}\bar{\gamma}} \int\limits_{\frac{r}{a^*(1-\alpha)}}^{\bar{\gamma}} \left(-\frac{\bar{\theta}^2}{8} - \frac{\alpha a^* \bar{\theta}}{2}\gamma_i - \frac{r\bar{\theta}}{2}\right) d\gamma_i \\
&= 2 \cdot \frac{1}{\bar{\theta}\bar{\gamma}} \left(-\frac{\theta_i^2}{8}\gamma_i - \frac{r\bar{\theta}}{2}\gamma_i - \frac{\alpha a^* \bar{\theta}}{4}\gamma_i^2 \Bigg|_{\frac{r}{a^*(1-\alpha)}}^{\bar{\gamma}}\right) d\gamma_i \\
&= -\frac{1}{\bar{\theta}\bar{\gamma}} \left(\left(\frac{\bar{\theta}^2}{4} + r\bar{\theta}\right)\left(\bar{\gamma} - \frac{r\bar{\theta}}{2}\gamma_i\right) + \frac{\alpha a^* \bar{\theta}}{2}\left(\bar{\gamma}^2 - \frac{r^2}{a^{*2}(1-\alpha)^2}\right)\right)
\end{aligned}
$$

$$
\begin{aligned}
VW_D^2 &= 2 \cdot \frac{1}{\bar{\theta}\bar{\gamma}} \int\limits_{0}^{\frac{r}{a^*(1-\alpha)}} \int\limits_{0}^{\frac{\bar{\theta}}{2}} \left(-\theta_i - \gamma_i a^*\right) d\theta_i d\gamma_i \\
&= 2 \cdot \frac{1}{\bar{\theta}\bar{\gamma}} \int\limits_{0}^{\frac{r}{a^*(1-\alpha)}} \left(-\frac{1}{2}\theta_i^2 - \theta_i\gamma_i a^*\right)\Bigg|_0^{\frac{\bar{\theta}}{2}} d\gamma_i \\
&= 2 \cdot \frac{1}{\bar{\theta}\bar{\gamma}} \int\limits_{0}^{\frac{r}{a^*(1-\alpha)}} \left(-\frac{\theta_i^2}{8} - \frac{a^*}{2}\bar{\theta}\gamma_i\right) d\gamma_i \\
&= 2 \cdot \frac{1}{\bar{\theta}\bar{\gamma}} \left(-\frac{\theta_i^2}{8}\gamma_i - \frac{a^*}{4}\bar{\theta}\gamma_i^2 \Bigg|_{\frac{r}{a^*(1-\alpha)}}^{\bar{\gamma}}\right) d\gamma_i \\
&= -\frac{1}{\bar{\theta}\bar{\gamma}} \left(\frac{\bar{\theta}^2}{4a^*} \cdot \frac{r}{1-\alpha} + \frac{\bar{\theta}}{2a^*} \cdot \frac{r^2}{(1-\alpha)^2}\right)
\end{aligned}
$$

Combining the two yields total viewer welfare:

$$\begin{aligned}
VW_D &= -\frac{1}{\bar{\theta}\bar{\gamma}}\left(\frac{\bar{\theta}^2}{4}\bar{\gamma} + r\bar{\theta}\bar{\gamma} - \frac{r^2\bar{\theta}}{2a^*(1-\alpha)} + \frac{\alpha a^*\bar{\theta}\bar{\gamma}^2}{2} + \frac{\bar{\theta}r^2}{2a^*(1-\alpha)} - \frac{\alpha\bar{\theta}r^2}{2a^*(1-\alpha)}\right) \\
&= -\frac{1}{\bar{\theta}\bar{\gamma}}\left(\frac{\bar{\theta}^2\bar{\gamma}}{4} + r\bar{\theta}\bar{\gamma} + \frac{\alpha a^*\bar{\theta}\bar{\gamma}^2}{2}\right) \\
&= -\left(\frac{\bar{\theta}}{4} + r + \frac{\alpha\bar{\gamma}a^*}{2}\right)
\end{aligned}$$

Comparing viewer welfare before and after the DVR yields:

$$\begin{aligned}
VW_D &> VW_{ND} \\
\Longleftrightarrow -\left(\frac{\bar{\theta}}{4} + r + \frac{\alpha\bar{\gamma}a^*}{2}\right) &> -\frac{5\bar{\theta}}{4} \\
\Longleftrightarrow \frac{\bar{\theta}}{2} &> r + \frac{1}{2}\sqrt{\bar{\theta}^2 - \frac{1+\alpha}{1-\alpha}r^2} \\
\Longleftrightarrow \bar{\theta} - 2r &> \sqrt{\bar{\theta}^2 - \frac{1+\alpha}{1-\alpha}r^2} \\
\Longleftrightarrow \bar{\theta}^2 - 4r\bar{\theta} + 4r^2 &> \bar{\theta}^2 - \frac{1+\alpha}{1-\alpha}r^2 \\
\Longleftrightarrow \frac{5-3\alpha}{4(1-\alpha)}r &> \bar{\theta} \\
\Longleftrightarrow \left(\frac{1}{1-\alpha} + \frac{1-3\alpha}{4(1-\alpha)}\right)r &> \bar{\theta}
\end{aligned}$$

For $\alpha > \frac{1}{3}$, the last line is inconsistent with the assumption $\bar{\theta} > \frac{r}{1-\alpha}$ and hence the DVR hurts the viewers. However, for $\alpha < \frac{1}{3}$, the DVR can help the viewers so long as $\bar{\theta} < \frac{5-3\alpha}{4(1-\alpha)}r$. ∎

B.5 DVR Comparative Statics

Proposition 15 *The following comparative statics hold:*

1. $\frac{da_D^*}{d\bar{\theta}} > 0$, $\frac{d\pi_D^*}{d\bar{\theta}} > 0$

2. $\frac{da_D^*}{d\bar{\gamma}} < 0$, $\frac{d\pi_D^*}{d\bar{\gamma}} < 0$

3. $\frac{da_D^*}{d\alpha} < 0$, $\frac{d\pi_D^*}{d\alpha} < 0$

4. $\frac{da_D^*}{dr} < 0$

Proof From section 4.3, the equilibrium ad level is

$$a^* = \frac{1}{\alpha\bar{\gamma}}\left[\bar{\theta} + \sqrt{\bar{\theta}^2 - \frac{1+\alpha}{1-\alpha}r^2}\right]$$

and from section B.3, equilibrium profit is:

$$\pi_D^* = \frac{1}{2\bar{\gamma}} \cdot [r + \alpha\bar{\gamma}a^*] = \frac{1}{2\bar{\gamma}} \cdot \left[r + \bar{\theta} + \sqrt{\bar{\theta}^2 - \frac{1+\alpha}{1-\alpha}r^2}\right]$$

from which the comparative statics follow immediately. ∎

Part II

Time-Shifting

Chapter 1

Introduction

Television ads accounted for \$69.8 billion of the \$284.8 billion spent on advertising in 2008, or about 1 in every 4 dollars spent on advertising (Keane, 2008). These commercials constitute the vast majority of television networks' revenue streams; as a result, the networks have a strong incentive to make sure they are maximizing the ad dollars their programs receive. Once a network has produced its portfolio of shows, the main driver of audience size, and consequently total ad revenue, is the timing of when it airs its programs. Suboptimal lineup choices have been shown to have a significant effect on audience size. Horen (1980) examined lineups and ratings retrospectively and showed that rescheduling a network's programs to optimal lineups would have increased total viewership by over 10%, while Shachar and Emerson (2000) estimates this loss to be between 6% and 13% for each of the three major networks on select nights. As a result, networks spend considerable time and resources determining optimal lineup choices for a given portfolio of shows; Carter (1996) details the extent to which networks go to set their lineups optimally.

Traditionally, the main driver of program lineup strategies has been the so-called "lead-in effect," the positive effect a popular program has on the shows that follow.

Many studies have shown this effect to have a significant impact on viewing patterns.[1] Rust and Eechambadi (1989) found audience flow to be an even greater determinant of a show's ratings than program attractiveness. As a result, networks pay careful consideration to this effect in their lineup decisions. Horen (1980) discusses five timing strategies commonly used by networks, three of which depend on inheritance from the lead-in audience: "protecting newcomers," or airing a strong show before a new one to build its audience, "starting fast," or scheduling the best shows early in the evening to generate a lead-in boost to subsequent programs, and "homoegeneity," or putting similar programs on one night so people who watch one will stick around for the others as well.

However, the advent of the digital video recorder (hereafter, DVR) has complicated time competition for the networks. A DVR allows one to record shows for future consumption, effectively unraveling the timeslot structure imposed by the networks. Most DVRs can also record multiple shows at one time, enabling the user to watch programs that air simultaneously. Access to these abilities is commonplace across television viewers, with DVR penetration hitting 25% in 2008 (Vranica, 2008) and projected to reach 46% by the end of 2010 (Haugsted, 2007).

These capabilities have a profound effect on the dynamics of time competition across networks. The lead-in strategies described above rely on the networks having control over the timing of their programs; the digital video recorder takes this control out of the networks' hands and puts it into the users'. Additionally, the ability to record shows that air simultaneously eliminates another often-used programming strategy, "counterprogramming" (Horen, 1980), the strategy of airing a program of a different type than one's competitors' in a given timeslot. DVR users can record multiple shows on various stations throughout the week and then watch them at whatever

[1] See section 2 for a more thorough discussion on the empirical evidence of the lead-in effect.

time and in whatever order they prefer. As Harmon reported in the New York Times, "[DVR] owners often choose to watch what is on the machine rather than what is on TV. Ignoring the networks' painstakingly planned schedules, they watch prime-time programs late at night and late-night programs before dinner, often oblivious to the channel on which it originally appeared" (2002). Consequently, any programs that the networks air around those shows would lose the benefit they traditionally would have received from the lead-in audience.

To explore the programming strategies networks have historically employed as well as the DVR's effect on those strategies, I build a model of time competition across networks. Section 2 reviews the relevant literature, and section 3.1 presents the baseline model. Sections 3.2 and 3.3 examine the baseline model first with myopic viewers and then with foresighted ones, comparing the results of the two treatments. In Chapter 4, I incorporate DVR technology into the model and demonstrates how the device affects the networks' strategies and equilibrium outcomes. Chapter 5 concludes.

Chapter 2

Literature Review

The model presented here follows the well-established research stream of Hotelling (1929), applied to the two-sided market of television as in a number of papers, such as Anderson and Coate (2005) and Ambrus and Reisinger (2006), among others. These models, however, generally focus on competition among networks at one time rather than on the timing game across timeslots. As such, these models cannot incorporate any audience flow effects from a show of higher quality to those surrounding it.

The lead-in effect I add here is grounded in several empirical studies demonstrating its impact on television audiences. Early works such as Gensch and Shaman (1980) suggested incorporating the impact of lead-in audiences would improve predictive models for rating shares; contemporaneous research, such as Horen (1980) and Rust and Alpert (1984), verified this claim. More recent work has also resulted in similar findings. In Cooper (1993), lead-in and lead-out audience size explains 87% of the variation in a given program's ratings, overwhelming any other factor in the model, with lead-in ratings contributing the majority of the predictive value. Goettler and Shachar (2001) further supports the effect's impact on viewer behavior, finding that over 56% of a show's viewers watched the end of the preceeding show on the same network.

Generally, models of viewer behavior include the lead-in effect as a transaction or switching cost in the viewer's utility function (e.g., Rust and Alpert (1984), Rust and Eechambadi (1989), Shachar and Emerson (2000), Goettler and Shachar (2001), Liu et al. (2004)). The bulk of the related research, however, has been empirical in nature; few have presented theoretical work on the topic. One exception is Liu et al. (2004), which most closely resembles the research presented here. The authors begin with a model of network investment in quality under monopoly and duopoly conditions and then extend the analysis to a two-period duopoly game. Viewers are myopic, have heteregeneous tastes for programming, and undergo a lead-in effect represented by a cost to changing channels between periods. Networks make quality investments simultaneously, after which viewers decide what to watch. Under the assumptions that the market is fully covered and a Nash equilibrium in pure strategies exists, they solve for each network's equilibrium quality choices in each period, first with myopic viewers and then with foresighted ones. They find that, with myopic viewers, regardless of the effect's magnitude, its existence ensures that any equilibrium involves each network putting on a higher quality show in the first period than in the second. Further, the symmetry of the model dictates that the networks always air the same quality show in each timeslot, regardless of the types of viewers they face or the parameter values.

The work here adds to that of Liu et al. (2004) in several key ways. First, the market in this model need not be fully covered; instead, viewers also have an outside option. This serves as a critical driver of both network and viewer behavior, as the existence of an outside options allows the networks' strategies to affect the total audience size in either or both timeslots. Additionally, I include a second dimension of viewer heterogeneity along the value of the outside option. This introduces variance in behavior across any subset of viewers of a given taste level, as some will choose to watch television while others have another option they find more enriching.

The main contribution of this paper, however, is the incorporation of the digital video recorder (DVR). Some previous work has explored the DVR as a tool to avoid advertisements bundled within a program (see Anderson and Gans (2011) or Part I of this dissertation), but to-date none have modeled the time-shifting aspect of the DVR. The model described herein demonstrates how the DVR renders the timing game between the networks irrelevant. This result does not require all viewers to rent a DVR; in fact, it shows that the result can still hold even as the proportion of viewers that rent a DVR becomes arbitrarily small. Subsequently, ad-filtering is added back into the model, and conditions under which the DVR either helps or hurts the networks are identified. Given the prevalence of DVR technology among television-viewing households and the pace at which its use is growing, this represents the paper's most significant addition to the literature.

Chapter 3

Lead-In Effect

3.1 Model

Two television networks, or stations, compete for the attention of potential viewers over two consecutive hour-long time-slots. Each network provides a different kind of programming; e.g., one comedy and one drama, or one sports and one news, etc. $\theta \in \left[0, \bar{\theta}\right]$ represents the type of programming, with $\theta = 0$ corresponding to network 1's content and $\theta = \bar{\theta}$ corresponding to network 2's content. The networks' programming types are fixed; there is no mechanism for the networks to change their θs. To focus solely on the timing game, I assume ad levels per show are fixed for both networks. Each network receives a constant payoff anytime one viewer watches one ad; hence, each network aims to maximize total ad impressions across its two shows.

The networks are exogenously endowed with one hit show and one mediocre show. The only decision each network need make is what to air in each timeslot. The networks choose and announce their lineups simultaneously, to which they are then committed. Consequently, the game is symmetric across the networks.

After the lineups are announced, each potential viewer decides how to spend her

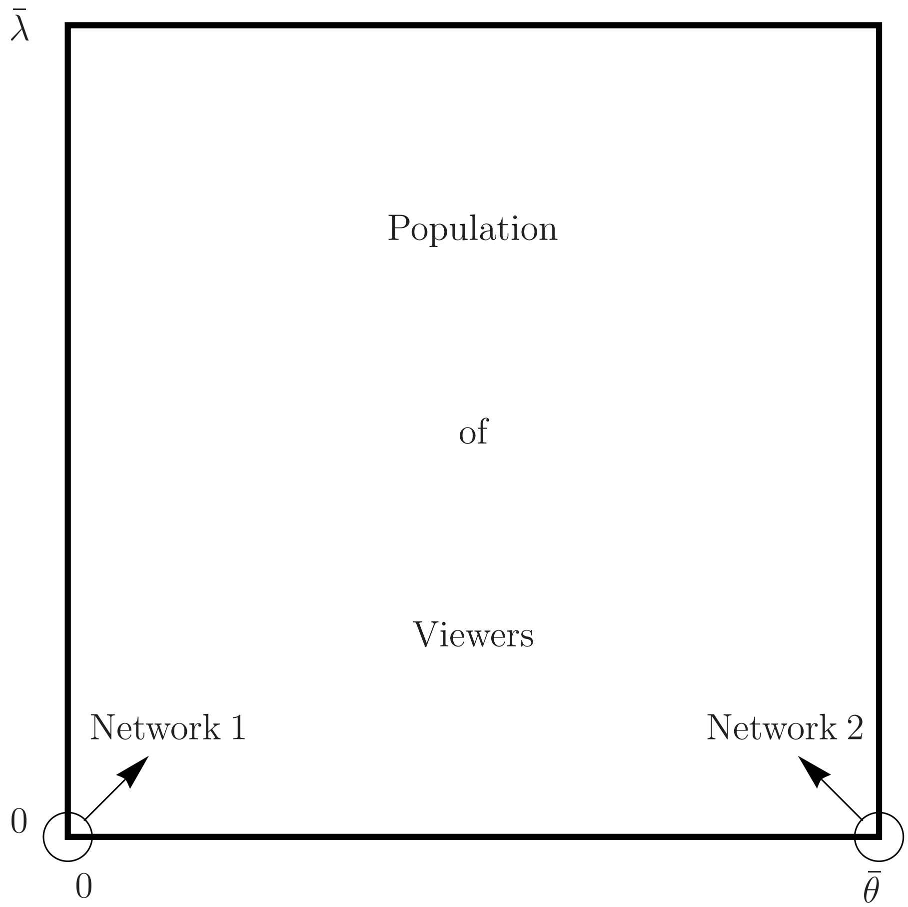


Figure 3.1: Population of Viewers and Network Locations

time. For succinctness and clarity, I refer to those considering their options as "individuals" and to those that have chosen to watch TV as "viewers." In each of the two hours, an individual may watch network 1, watch network 2, or utilize her outside option. The outside option imparts a benefit of $\lambda_i \sim U\left[0, \bar{\lambda}\right]$ to individual i. Individuals are heterogeneous over their ideal programming types as well: $\theta_i \sim U\left[0, \bar{\theta}\right]$. To focus solely on the impact of timing choice on the individuals' decisions, the individuals in this model have no aversion to ads. Figure 3.1 depicts the individuals and the networks graphically.

If an individual chooses to watch TV in a particular time slot, she receives V^H

utility for watching a hit show, or $V^M < V^H$ if she watches a mediocre show.[1] She also receives some disutility from any imperfect match between her ideal programming type and the show she views equal to the distance between her type and the location of the network she watches - i.e., θ_i for a network 1 show and $\bar{\theta} - \theta_i$ for a network 2 show. As the two time periods are only an hour apart, I apply no discount to second period utility. I assume that those with the highest opportunity cost of time prefer their outside options to watching a perfectly matched hit show $\left(\bar{\lambda} > V^H\right)$. I also assume that any individual with no value from her outside option chooses to watch television, regardless of what is on $\left(V^M > \frac{\bar{\theta}}{2}\right)$.

To examine the dynamics of audience flow, individuals that watch either station in period 1 also experience a "lead-in effect" that promotes the desire to watch the same station in period 2 by a factor $\delta > 0$. This effect represents the sum of all influences motivating the viewer to keep watching the same channel, including the cost of changing channels as well as any positive advertising for the second show that the network airs during the first. It could be considered a bonus to staying with the same station or a cost to changing channels or turning off the TV. In this sense, it is analogous to a traditional switching cost, except that there is no cost associated with switching from the outside option in the first period to a TV show in the second. I assume $\delta < \left(V^H - V^M\right) \equiv \bar{\delta}$, so that all viewers that watch a hit show do not automatically choose to watch a mediocre show that follows.

Incorporating all of the aforementioned effects, utility can be defined as follows:

$$U_i^{a,b} = \mathbf{1}\left(a \neq 0\right) \cdot \left[V^{\psi_a^1} - \left|\theta_i - \theta^a\right| - \lambda_i\right] + \mathbf{1}\left(b \neq 0\right) \cdot \left[V^{\psi_b^1} - \left|\theta_i - \theta^b\right| - \lambda_i\right] + \mathbf{1}\left(a = b \neq 0\right) \cdot \delta,$$

where $U_i^{a,b}$ represents the total utility individual i receives in excess of her outside

[1] Viewers receive this utility from watching a show only if they have not seen it before, so the networks have no incentive to simply air their hit shows in both timeslots.

option if she chooses station a in timeslot 1 and b in timeslot 2, with 0 representing the outside option; ψ_j^t is the type of show station j airs in timeslot t, with H representing a hit show and M representing a mediocre show; and θ^a represents station a's location.

The timing of the game is as follows:

1. The two networks simultaneously choose and announce their lineups,
2. Individuals decide which actions to perform,
3. All parties receive their payoffs.

In the sections that follow, I analyze equilibrium strategies of the game under various sets of conditions. I first consider equilibrium strategies when individuals are myopic in the sense that they make two separate one-period decisions at the beginning of each timeslot. Next, I examine the outcome when individuals are foresighted, meaning they instead make a joint two-period decision at the beginning of the first timeslot. Lastly, I incorporate the digital video recorder (DVR) which, for a positive rental cost, allows viewers to watch any subset of the four programs instead of just one show in each timeslot.[2] I solve for equilibrium strategies in each case and examine how the various treatments affect the equilibrium outcome.

3.2 Myopia

When the individuals are myopic, they make their decisions one period at a time. After the networks announce their lineups, individuals make their period 1 choices based on potential period 1 payoffs; at the beginning of period 2, they make their period 2 choices based on the period 2 payoffs conditional on what they did in period 1.

[2]See Section 4.1 for more detail on how the DVR is incorporated into the model.

To find equilibrium outcomes, I begin at the last stage and work backwards. Immediately before payouts are awarded, individuals make their decisions given the lineups announced by each network. To illustrate the individual choice problem, consider a period 1 example in which network 1 is airing its hit show and network 2 its mediocre show.[3] In this situation, the individual at $\left(0, V^H\right)$ will be indifferent between her outside option and watching network 1's hit show, whereas the one at $\left(\bar{\theta}, V^M\right)$ will be indifferent between her outside option and watching network 2. The set of individuals for whom these indifference relationships hold will form two lines that extend from these points inwards at 45° angles, representing the one-to-one utility tradeoff between programming match and the wage rate.

The assumption $V^M > \frac{\bar{\theta}}{2}$ ensures these lines will meet each other before hitting the θ-axis. This intersection represents the point of three-way indifference; the individual at this point receives equal value from watching either network or the outside option. The third and last indifference line extends downwards from this point, representing those that prefer watching TV to the outside option but are indifferent about which network to watch.

These three indifference lines segment the population of individuals into their optimal choices: individuals in the bottom left prefer network 1, those in the bottom right prefer network 2, and those above the indifference lines prefer the outside option. Figure 3.2 displays the individuals' period 1 optimal choices when network 1 is airing its hit show and network 2 is airing its mediocre show.

This demonstrates the general structure of optimal responses to a given lineup in period 1. Period 2 reactions evolve in the same way except that, due to the lead-in effect, individuals up to $\left(0, V^M + \delta\right)$ for network 1 and $\left(\bar{\theta}, V^M + \delta\right)$ for network 2 will watch a mediocre show if that station aired a hit show previously. The assumption

[3] For more detail on individual choice, see Section C.1 in the Appendix.

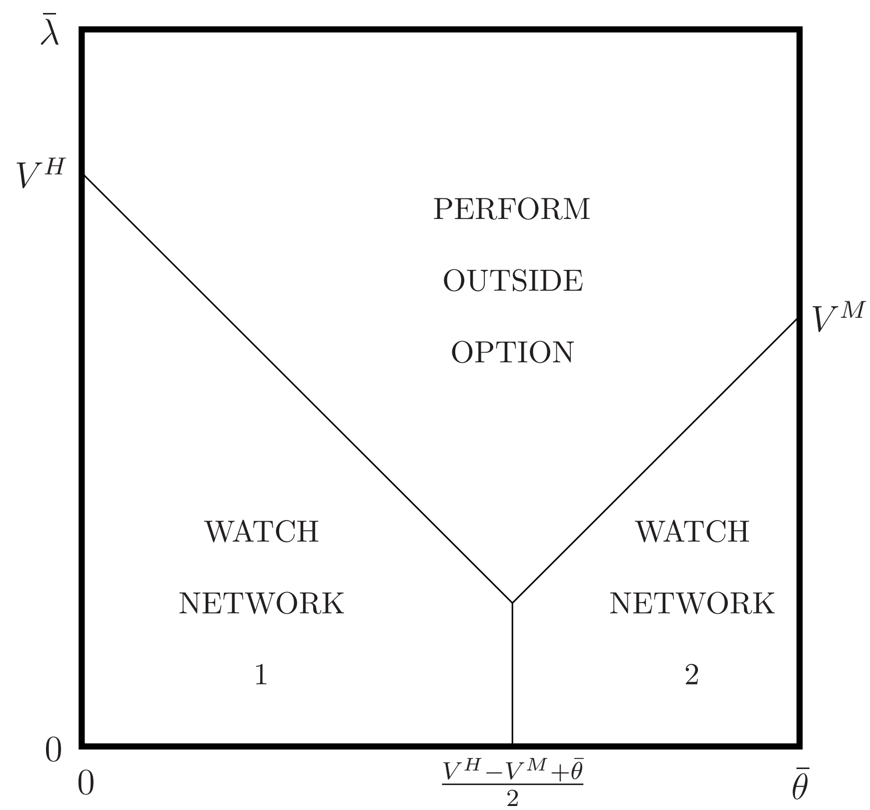


Figure 3.2: Individual Choice in Period 1 Facing One Hit Show

$V^H > V^M + \delta$ ensures that, in equilibrium, all of these individuals also watched that station's hit show in period 1.

Define the lineup outcome that results from the networks' decisions as $(\psi_1^1, \psi_1^2) \times (\psi_2^1, \psi_2^2)$, where ψ_j^t is the type of show station j airs in timeslot t. Individuals can face three potential lineup outcomes: there will either be zero, one, or two hit shows aired in the first period, with the remainder aired in the second. Since the networks are symmetric, I assume WLOG that if one hit show is aired early, it is aired by network 1. The potential lineup outcomes are therefore: $(H, M) \times (H, M)$, $(H, M) \times (M, H)$, and $(M, H) \times (M, H)$. I call the first outcome "head-to-head" lineups, the second "staggered" lineups, and the third "back-loaded" lineups.[4] Figure 3.3 displays the

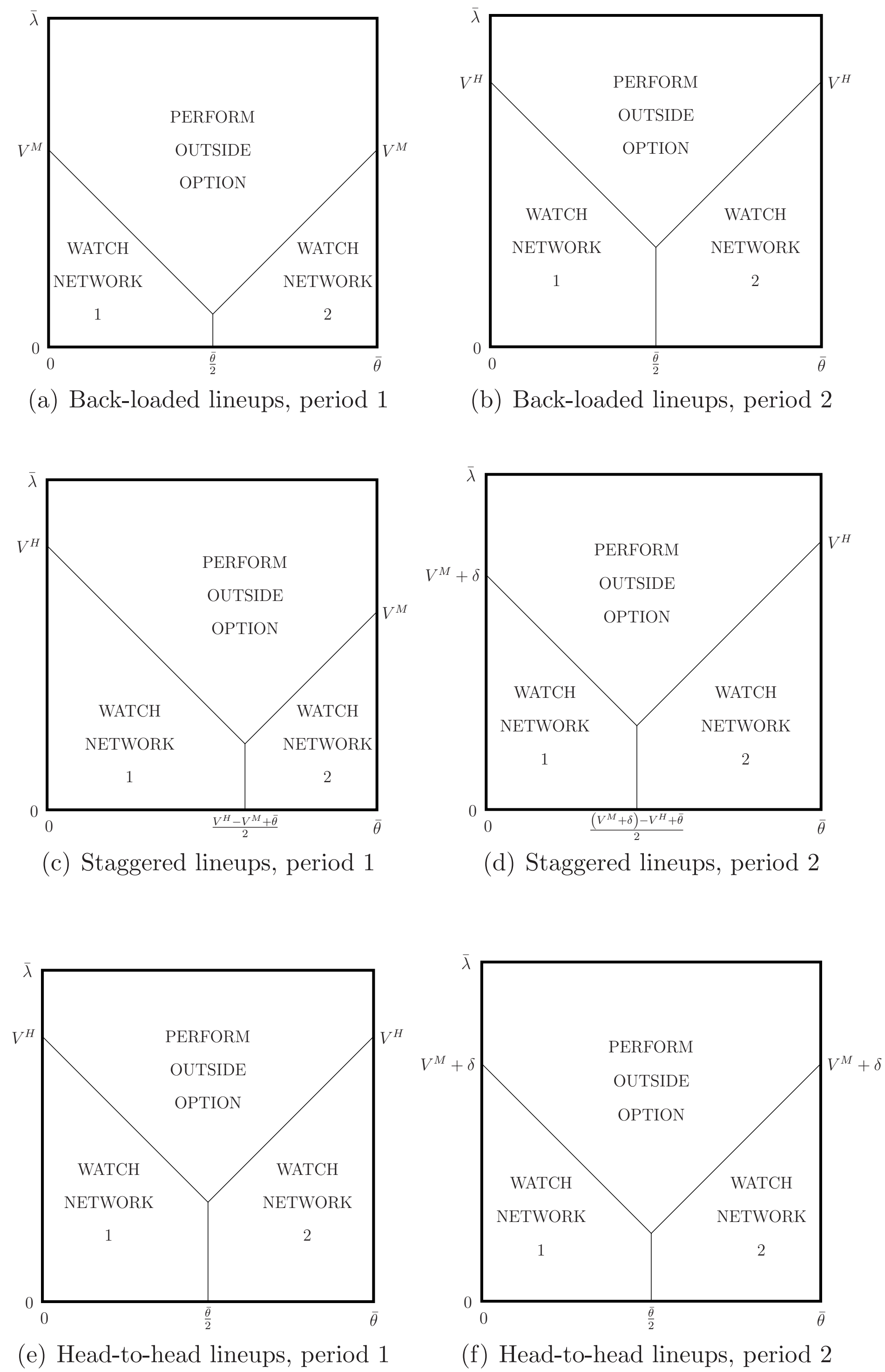
[4]Due to the symmetry of the game, any staggered lineup equilibrium would involve a second equilibrium where the roles are flipped.

individuals' decisions for each period under each possible lineup scenario.

The next step moving backwards in the game's timing is to determine each network's reaction function.[5] As Figure 3.3 demonstrates, the lead-in effect benefits a network if and only if that network airs its hit show first. An individual who normally would not watch a particular program will decide to do so if she is already watching that station and no other option provides her more than δ additional utility. This could only occur, however, if the hit show is aired first; any individual who watched a network's mediocre show in the first timeslot would only have done so if her location were such that she would watch that network's hit show in the second period anyway. Upon watching the hit show in the second period, she would still receive the δ bonus to her utility; however, her myopia precludes her from incorporating this into her first-period decision. Consequently, each network has an incentive to air its hit show first.

At the same time, each network has an incentive to air its hit show during the other network's mediocre show due to the presence of the outside option. Assume for a moment there were no lead-in effect. If there were also no outside option, neither network would have any incentive to avoid direct competition between the hit shows as the overall television audience across timeslots would be fixed. When the outside option is introduced, so too is the opportunity to increase total TV viewership across networks and timeslots by staggering the two hit shows. Essentially, the individuals with moderate λ_i's and θ_i's are only motivated to watch TV by hit shows. When these shows are pitted head-to-head, those individuals are forced to choose only one hit show in that timeslot; they then choose not to watch television at all in the other timeslot. When the hit shows are staggered, those individuals can watch one in each timeslot, increasing the total television audience across both periods.

[5] At least one mixed strategy equilibrium exists, but I focus solely on pure strategy ones here.



(a) Back-loaded lineups, period 1

(b) Back-loaded lineups, period 2

(c) Staggered lineups, period 1

(d) Staggered lineups, period 2

(e) Head-to-head lineups, period 1

(f) Head-to-head lineups, period 2

Figure 3.3: Optimal Viewing Decisions - Myopic Individuals

Consequently, two competing aspects of time competition exist. On the one hand, both networks have an incentive to put their hit shows first to capture the lead-in effect in the second timeslot. On the other hand, each network wants to air its hit show during the other's mediocre show so that the two hit shows are not competing for viewers simultaneously. The balance between these two forces determines each network's optimal response to a given lineup by its rival.

When the opposing station chooses (M, H), these two factors act in the same direction, so that one lineup response dominates the other. The following proposition follows.

Proposition 16 *When individuals are myopic, the optimal reaction to* (M, H) *is* (H, M).

Proof See Section C.2 in the Appendix.

If one network airs its mediocre show first, the other can take advantage of both incentives listed above by doing the opposite. By airing its hit show first in response, it captures the lead-in effect for its mediocre show in the latter timeslot, and it also avoids head-to-head competition between the two networks' hit shows.

However, the optimal reaction to (H, M) is not as straightforward, as the two factors work contrary to one anohter. Figure 3.4 depicts the tradeoff between these two forces when network 1 plays (H, M) and network 2 is contemplating its two options; it is essentially an overlay of Figures 3.3(c), 3.3(d), 3.3(e), and 3.3(f).

Assume network 2 has decided to play (H, M) in response but is considering deviating to (M, H). Its hit show's audience would increase in size due to the diminished quality of the competition it faces; this increase is represented by the trapezoid A in

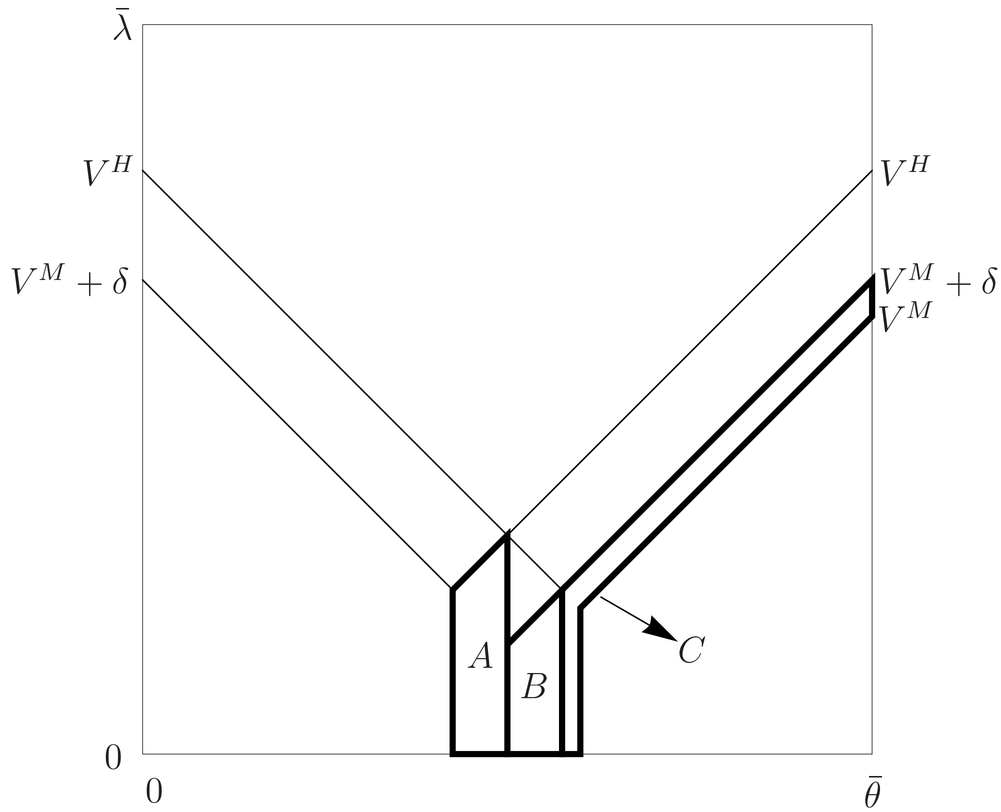


Figure 3.4: Tradeoff Between Lineup Choices Facing (H, M) with Myopic Individuals

the diagram. However, its mediocre show's audience will decrease in size, both by the trapezoid B, which represents the viewers who will defect to the other network in that time slot, and by C, which represents the loss from no longer capturing the lead-in effect. The comparative sizes of these two areas is not immediately clear; the following proposition clarifies this result.

Proposition 17 *There exists $\hat{\delta} \in \left[0, \bar{\delta}\right]$ such that, when individuals are myopic, the optimal response to (H, M) is (M, H) for $\delta < \hat{\delta}$ and (H, M) for $\delta > \hat{\delta}$.*

Proof See Appendix Section C.3 for a detailed proof. Briefly, the logic is as follows. Regardless of parameter values, $A > B$, but depending on C's magnitude, $B + C$ might be large enough to outweigh A and therefore prevent defection to (M, H). The size of all three shapes is affected by δ; A is monotonically decreasing in δ, while $B+C$

is monotonically increasing in δ. As one can ascertain from Figure 3.4, $B + C = 0$ when $\delta = 0$, $A = 0$ when $\delta = \bar{\delta}$, and all three shapes have positive area otherwise. Consequently, there must be a crossing point for some $\hat{\delta} \in \left[0, \bar{\delta}\right]$.

With the reaction function complete, equilibrium outcomes follow. When the lead-in effect is strong enough ($\delta > \hat{\delta}$), each network is better off leading with its hit show regardless of the other network's lineup. As a result, (H, M) is a dominant strategy for each, and $(H, M) \times (H, M)$ is the resulting outcome. However, when the lead-in effect is not as strong ($\delta < \hat{\delta}$), (H, M) and (M, H) are optimal responses to each other. As a result, for $\delta < \hat{\delta}$, any equilibrium involves staggered lineups.[6]

Proposition 18 *The timing game with myopic individuals has a unique subgame-perfect equilibrium. If $\delta < \hat{\delta}$, the networks play staggered lineups, and the equilibrium outcome is $(H, M) \times (M, H)$. For $\delta > \hat{\delta}$, the networks play head-to-head lineups, and the equilibrium outcome is $(H, M) \times (H, M)$.*

This equilibrium is not necessarily socially optimal, however. To measure social optimality, I examine network welfare, then viewer welfare, and then total welfare in turn. Since ad allotments are fixed and equal across shows, and viewers of a given show much see all ads within that show, network welfare in this case is simply the total viewership across both networks and both timeslots. The following proposition characterizes maximization of viewership, which may differ from the equilibrium outcome.

Proposition 19 *When individuals are myopic, there exists $\tilde{\delta} \in \left(\hat{\delta}, \bar{\delta}\right)$ such that, for $\delta < \tilde{\delta}$, network welfare is maximized when the networks choose staggered lineups, and for $\delta > \tilde{\delta}$, network welfare is maximized when the networks both choose (H, M).*

[6] See Sections C.2 and C.3 for a more rigorous explanation, including derivations of profit expressions in each scenario.

Proof See Section C.4 in the Appendix.

An implication of this result is that total viewership across the networks is maximized for $\delta < \hat{\delta}$ or $\delta > \tilde{\delta}$, but for $\hat{\delta} < \delta < \tilde{\delta}$, competition between the networks hurts their joint welfare. After network one chooses (H, M), network 2's decision depends on δ. For low δ, the lead-in effect is too small to overpower network 2's incentive not to compete head-to-head with network 1, causing it to choose (M, H). For high δ, the lead-in effect is large enough that network 2 competes head-to-head, and it is also high enough that it drives more additional individuals to tune in to a show than are lost to the outside option from direct competition between the networks' hit shows. Essentially, network 2 changing from head-to-head lineups to staggered ones has a positive externality on network 1's viewership, but network 2 does not take this into account when it makes its lineup decision. Therefore, moving downwards from $\bar{\delta}$, the δ-threshold for which overall viewership is no longer higher with direct competition occurs before the δ-threshold at which network 2 is better off avoiding direct competition. As a result, moderate levels of δ lead to socially inefficient outcomes.

Proposition 20 *From a first-best standpoint, viewer welfare is not maximized when $\delta < \hat{\delta}$, and consequently neither is total welfare. For $\delta > \hat{\delta}$, whether or not viewer welfare and total welfare are maximized is ambiguous.*

Proof See Section C.5 in the Appendix.

When $\delta < \hat{\delta}$, the networks play staggered lineups, and some viewers who should be watching the mediocre show that airs in the first timeslot do not do so because they do not rationally anticipate the additional utility they will get since they will also be watching the hit show in the second timeslot. Consequently, viewer welfare and, therefore, total welfare cannot be maximized in this scenario.

However, when $\delta > \hat{\delta}$, the networks play head-to-head lineups, after which all viewers automatically make the decisions that are best for them across timeslots, since the opportunity to watch a mediocre show in the first timeslot no longer exists. In this case, viewer and total welfare may or may not be maximized, depending on the values of the parameters.

Overall, myopic individuals make their decisions one period at a time, looking only at the payoffs that the very next period holds for them. This causes individuals to occassionally mistakenly choose an option that is suboptimal. Knowing this, the networks respond in-kind, and therefore make different lineup decisions then they might otherwise. If individuals instead made the correct decisions throughout the timing game, the outcome would contrast with that presented here; the subsequent section provides that contrast.

3.3 Foresight

When individuals are foresighted, they plan their actions for both periods at the beginning of period 1. As a result, the lead-in effect does not surprise them at the beginning of period 2; instead, they've already incorporated it into their joint two-period decision. Consequently, a viewer who will watch a mediocre show in period 2 will also plan around the fact that watching the preceding hit show on that station will yield an extra δ utility, as opposed to a myopic individual, who would be surprised by this utility bonus.

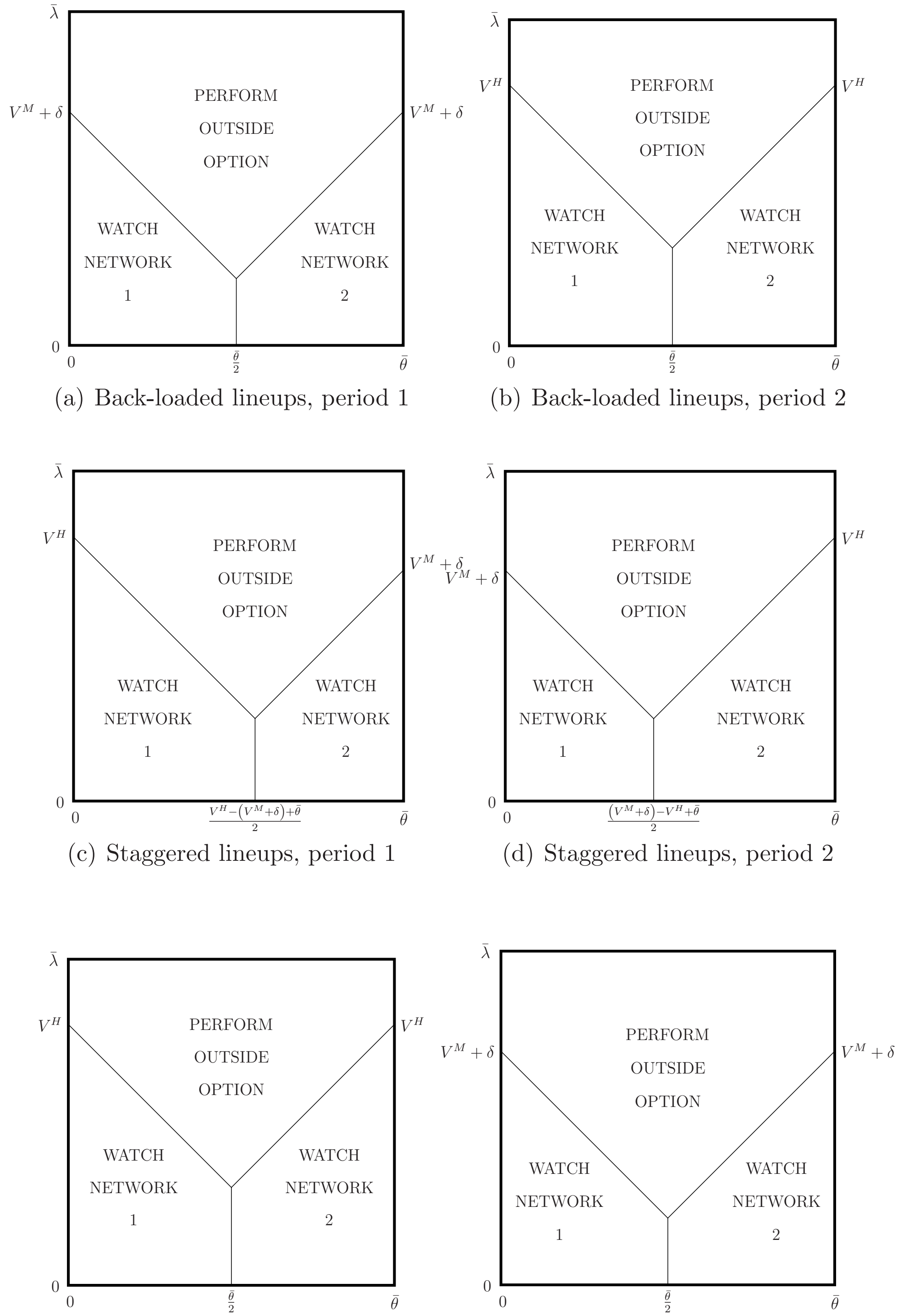
Figure 3.5 displays the choices of foresighted individuals for each of the potential lineups they may face. As Figure 3.5 demonstrates, when individuals are foresighted, the lead-in effect in essence becomes bidirectional. Individuals who plan to watch a hit show in the second period may be willing to take a small utility loss (as compared to the outside option) in the first period to capture the lead-in effect in the second. Consequently, the first of the two strategic incentives discussed in the myopic case disappears - each network captures the lead-in effect regardless of when it airs its hit show, leaving the desire not to go head-to-head as the only force driving in each network's decision.

Proposition 21 *With foresighted individuals,* (H, M) *and* (M, H) *are optimal responses to each other. As a result, only staggered lineups can exist in equilibrium with foresighted viewers.*

Proof See Section C.6 in the Appendix.

Appendix Section C.6 contains a formal proof as well as the expressions representing each firm's market share for each lineup decision. Briefly, however, the logic is as follows. Assume network 2 utilizes a lineup of (H, M). If network 1 chooses a lineup of (M, H), its mediocre show goes up against network 2's hit show in the first period and receives only a small market share. However, its hit show draws a large audience as it airs at the same time as its rival's mediocre show in the second period.

If network 1 considers deviating to (H, M), its mediocre show will air against network 2's mediocre show and therefore achieve a larger audience than it did previously. This gain is represented in Figure 3.6 by the trapezoid labeled B. However, its hit show will lose market share since it is now going against network 2's hit show as well. This loss is represented by triangle A plus trapezoid C in the graph. B and C are equivalent in area; therefore, the network would receive a net loss from deviating equal to the area of triangle A. This area is necessarily positive, with a magnitude of $\frac{\left(V^H - \left(V^M + \delta\right)\right)^2}{4}$, since triangle A is an isosceles right triangle with hypotenuse $\left(V^H - \left(V^M + \delta\right)\right)$.



(a) Back-loaded lineups, period 1

(b) Back-loaded lineups, period 2

(c) Staggered lineups, period 1

(d) Staggered lineups, period 2

(e) Head-to-head lineups, period 1

(f) Head-to-head, period 2

Figure 3.5: Optimal Viewing Decisions - Foresighted Individuals

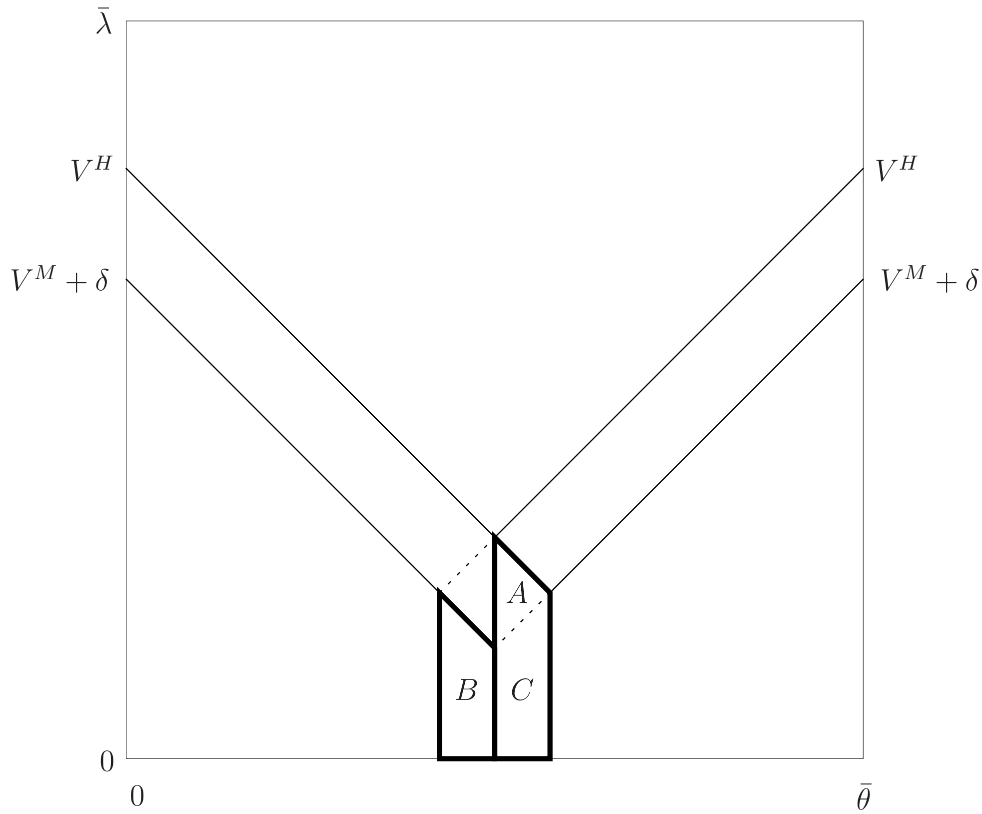


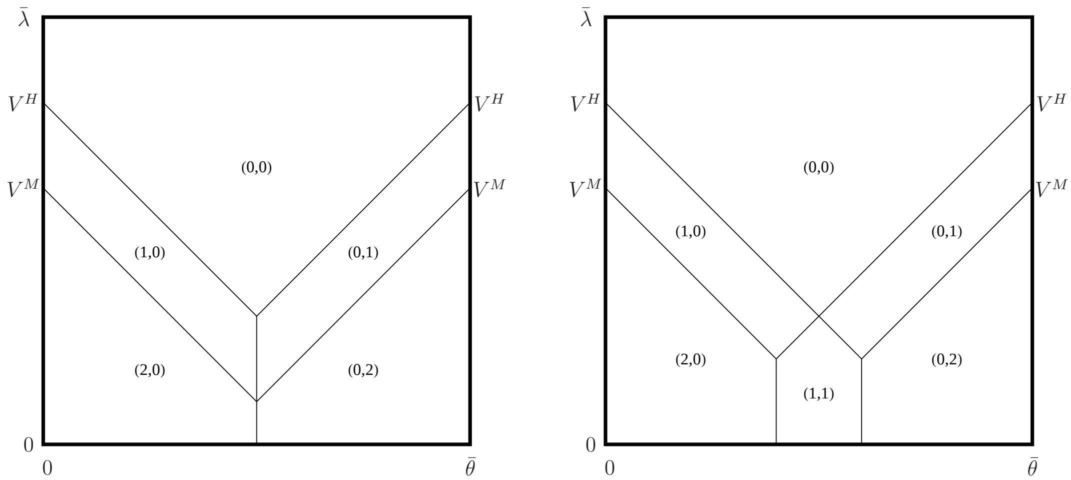
Figure 3.6: Reaction Payoffs vs. (H, M)

Since the staggered lineup outcome represents each network's individual audience maximum as well, the following proposition follows.

Proposition 22 *The SGP equilibrium with foresighted individuals maximizes total network welfare.*

It does not immediately follow, however, that this equilibrium is optimal for the viewers. To investigate viewer welfare, consider viewer choices in the staggered outcome compared to those of the head-to-head outcome. Figure 3.7 displays the equilibrium outcome for each scenario.

Comparing the two graphs, many viewers make the same decision in either lineup case, and therefore end up with the same welfare either way. Figure 3.8 displays a



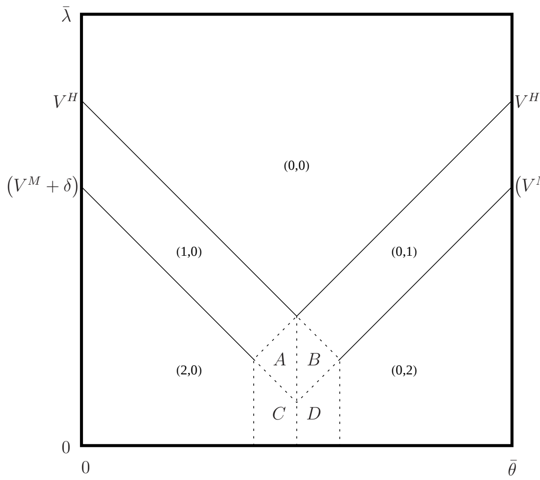
(a) Head-to-Head Lineups (b) Staggered Lineups

Figure 3.7: Foresighted Viewer Welfare - Head-to-Head vs. Staggered Lineups

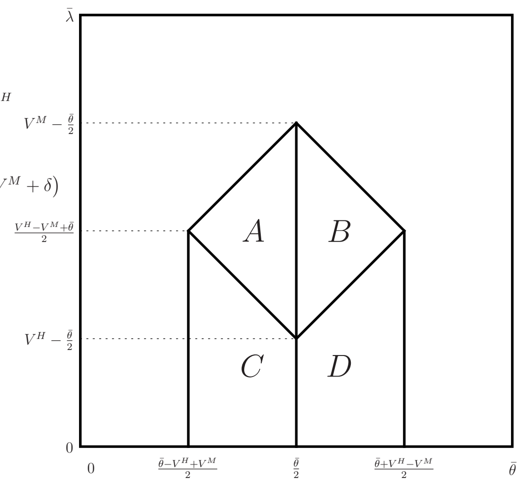
contrast of the viewers' decisions between the two lineup scenarios. Viewers outside the dotted area in figure 3.8(a) make the same decision in either lineup scenario, whereas those inside the dotted area make different decisions depending on the lineups they face. If starting in the head-to-head scenario and moving to the staggered one, the differences are as follows:

- Those in triangle A watch network 2's hit show instead of their outside option,
- Those in triangle B watch network 1's hit show instead of their outside option,
- Those in trapezoid C watch network 2's hit show instead of network 1's mediocre show, and
- Those in trapezoid D watch network 1's hit show instead of network 2's mediocre show.

Those in triangles A and B are of course unambiguously aided with staggered lineups instead of head-to-head ones, since they could perform their outside option with staggered lineups as well but choose not to. Each of those in trapezoids C



(a) Viewer Decision Comparison



(b) Close-up View

Figure 3.8: Foresighted Viewer Welfare - Head-to-Head vs. Staggered Lineups

and D, however, could be helped or hurt by the lineup change, since the quality of the show she watches with staggered lineups is improved, but it is no longer on her preferred network. Consequently, the net welfare difference is not immediately clear. The following proposition illuminates this issue.

Proposition 23 *When viewers are foresighted, staggered lineups maximize viewer welfare.*

Proof See section C.7 in the Appendix.

Because the advertiser market is perfectly competitive, adveriser surplus is zero regardless of the networks' or viewers' decisions, and total welfare therefore follows directly from network and viewer welfare. Since staggered lineups maximizes both, the following proposition immediately follows.

Proposition 24 *The unique SGP equilibrium with foresighted viewers results in staggered lineups. This equilibrium maximizes both network and viewer welfare, and therefore maximizes total welfare as well.*

Contrasting the models with and without perfect foresight, myopia allows both head-to-head lineups and staggered lineups to occur in equilibrium, depending on the magnitude of the lead-in effect. However, for any specific parameter values, only one of the two lineup outcomes can occur. Individual foresight, on the other hand, allows for only staggered lineups, regardless of the lead-in effect's strength. In essence, myopic individuals are surprised by the lead-in effect, so each network can only capture the associated benefit if it airs its hit show in the first timeslot. Conversely, foresighted viewers plan around the lead-in effect, effectively nullifying its impact on the networks' strategies.

Further, the outcome with myopic individuals could be socially inefficient, if δ is strong enough for network 2 to compete head-to-head but not strong enough to increase the overall television audience (i.e., not strong enough for network 2's lead-in effect to overpower the overall audience loss due to head-to-head competition). Foresighted individuals, on the other hand, incorporate the lead-in effect into their two-period decisions, so each network recognizes this benefit regardless of when it airs its hit show. Both networks consequently want to avoid direct competition and, in equilibrium, set their lineups to do so, which always leads to a socially efficient outcome.

Not surprisingly, the outcome with foresighted invididuals represents a Pareto-improvement over that with individual myopia. Each firm is made (weakly) better, as they avoid direct competition but still retain the lead-in benefit, while viewers are also made (weakly) better, since they maximize total utility over both periods instead of making decisions one period at a time. A practical implication of this is that all parties may be made better off through viewer education of channel lineups and advanced planning on viewing decisions. This suggests that promotional ads, in which a network uses its own ad space to promote a future show, may benefit more than just the shows they represent. Informing a viewer about a hit show airing in the future may make those viewers more likely to watch mediocre shows in adjacent time slots in addition to watching the hit show itself.

Note also that, for foresighted viewers, δ and V^M are not separately identifiable in equilibrium. This is because any viewer who watches a mediocre show also watches the hit show on the same network, and with foresighted viewers the lead-in effect is effectively bidirectional. Whether this will remain true in the DVR analysis that follows depends both on the exact cause of the lead-in effect and the nature of how one watches television when using a DVR, both of which remain uncertain. However, it is possible that the lead-in effect remains tied to the mediocre show's quality even with the DVR. For example, if the lead-in effect comes from promotional ads for the second show that air during the first, then the two factors are again inseparable as it remains true that anyone who watches a network's mediocre show still also watches the hit show as will be shown in the analysis that follows. Regardless, the goal of the DVR analysis is to examine the effect of the recording capabilities of such a device and of its expansion of the viewing possibilities for the viewers on their behavior and on that of the networks. As such, I drop the δ term in the subsequent analysis for both the baseline foresighted case and the DVR analysis to focus solely on the recording capability's effect on the game.

Chapter 4

The Digital Video Recorder

4.1 Model

To incorporate DVR technology into the timeslot game, after the networks choose their lineups, each individual has the additional option to rent a DVR at a cost $r > 0$ which allows her to record any to all of the four television shows aired by the two networks.[1] Since individuals have no aversion to ads, the ad-skipping functionality of the DVR is irrelevant here; consequently, any results arrived at herein result solely from the DVR's time-shifting ability.[2]

Hereafter, I refer to individuals who rent a DVR as "users." Once a show is

[1]This raises the question of why the networks were not already airing TV shows that the viewers would watch in the two extra timeslots they now have available to them. There are several possible interpretations in which this would arise. One in particular is that the networks have a significant drop-off in show quality after these two shows, and the first two time-slots are "prime-time" television slots, during which by far the largest percentage of all potential viewers are available, so that each network is best served putting on these two shows in these slots regardless of what its competition does. If a small subset of those viewers also have other time available to them, but the additional shows the network has are all below $\underline{\gamma}$ (normalized here to 0), they would only sacrifice their time if they could watch another mediocre or high quality show. The utility of the digital video recorder would be derived from this motivation to expand the choice set.

[2]For a version of the model that incorporates a modified aspect of ad-skipping, please see the end of this section.

recorded, a user can watch it at any time, forgoing one hour of her outside option for each show viewed. One implication is that a DVR user can watch shows from each of the competing networks, even if they are aired simultaneously. So, an individual may get a DVR to, e.g., watch two hit shows that are pitted head-to-head, or watch a mediocre show that is aired during the other network's hit show, etc.

I assume that r is small in comparison to the other parameters. Specifically, I assume that the DVR's rental cost is less than the quality difference in a station's shows $\left(r < V^H - V^M\right)$, and that any individual with no value from her outside option would be willing to rent a DVR to watch either quality TV show on her preferred network if that were the only way she could watch it $\left(r < V^M - \frac{\bar{\theta}}{2}\right)$. Further, as the decision to get a DVR requires advanced planning over the allocation of shows an individual plans to watch, I use the timeslot game with foresighted individuals as the benchmark here.

4.2 Viewer Response

Now that individuals can rent a DVR, their decisions cannot be categorized simply based on what they do in each timeslot. Instead, decisions are identified by which shows are watched and whether or not a DVR is rented. Note, however, that for any individual that rents a DVR, watching a station's mediocre show implies she also watches that station's hit show, since watching the mediocre show implies the value of doing so exceeds the individual's wage rate, which implies the value of that station's hit show exceeds her wage rate as well. Further, for individuals without a DVR, the same still holds true. If the hit shows are aired at different times, an individual who watches a station's mediocre show over its rival's hit show would also watch that station's hit show over its rival's mediocre show. If the lineups pit hit shows head-to-head, viewing the station's mediocre show implies its value is higher than the rival station's mediocre show for that individual, which implies she prefers that station's hit show to its rival's as well. As a result, an individual's decision set can be categorized simply by how many shows she watches on each station and on whether or not she chooses to rent a DVR.

To identify equilibrium strategies in this version of the game, I again start with the problem of individual choice in the stage immediately preceeding payouts and work backwards. To examine the choices individuals make when DVRs are available, I begin with the individuals' decisions before DVR technology and then see how these decisions change with the advent of the DVR.

The decisions an individual makes facing a given set of lineups before DVR technology reflect the tradeoff between each show's value to that viewer and the value of her outside option. The additional option of renting a DVR allows individuals to pay a rental cost to remove the restrictions of time slots, enabling the user to watch whatever subset of shows she may prefer. However, the DVR itself has no impact on the value of each show to the viewer or on the value of the outside option, so its availability could not cause any viewer to not watch a show that she would have watched otherwise. Further, an individual would only change her behavior if she actually chooses to rent a DVR. As a result, one can start with the pre-DVR optimal response mapping, determine which viewers would rent a DVR, and then simply add onto the original mapping the shows DVR technology would cause each user to additionally watch.

I start with the individual choice problem facing head-to-head lineups. Figure 4.1 demonstrates the individuals' optimal responses to head-to-head lineups without DVR technology derived in Section 3.3, with (a, b) representing the decision to watch a shows on network 1 and b shows on network 2.

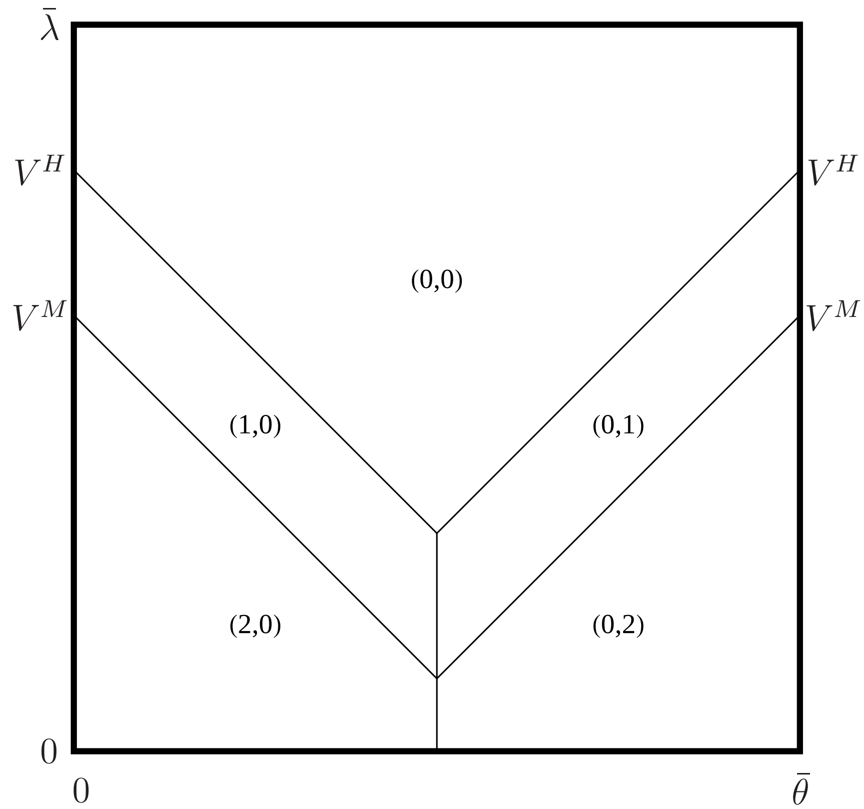


Figure 4.1: Individual Decisions with Head-to-Head Lineups - DVR Not Available

To quantify the effect of the DVR, I examine each subset of the individual population based on their varying decisions in turn. First, if an individual chooses not to watch television in either timeslot before the DVR, she must prefer her outside option to each of the four shows available to her. This implies she would not pay a rental cost for the DVR to watch television; consequently, any individual who chooses not to watch television when there is no DVR would not then rent a DVR once it becomes available.

As a result, we can restrict our attention to viewers only, i.e., those who watch at least one show before DVR technology. Consider the viewers of at least one network 1 show. None of these viewers would rent a DVR to watch network 2's mediocre show only. In other words, for network 1's viewers, the action of renting a DVR and watching network 2's hit show is a necessary condition for watching network 2's mediocre show. Consequently, we can determine those that rent a DVR and watch network 2's hit show, and then subsequently determine the subset of those viewers that also watch network 2's mediocre show.

A network 1 viewer will rent a DVR and watch network 2's hit show if the value of the show to her minus the rental cost exceeds her outside option. That is, for $\theta_i < \frac{\bar{\theta}}{2}$, the condition to rent a DVR and watch network 2's hit show is:

$$\begin{aligned} V^H - \bar{\theta} + \theta_i - r &> \lambda_i \Longrightarrow \\ \left(V^H - r\right) - \left(\bar{\theta} - \theta_i\right) &> \lambda_i \end{aligned}$$

Similarly, for $\theta_i > \frac{\bar{\theta}}{2}$, the condition to rent a DVR and watch network 1's hit show is:

$$\left(V^H - r\right) - \theta_i > \lambda_i$$

To determine which of these viewers also watch the mediocre show, note that there is no marginal rental cost associated with this decision since these individuals have already rented the DVR. As a result, the boundary conditions for a DVR user to watch the other network's mediocre show are simply extensions of those pre-DVR:

$$V^M - \left(\bar{\theta} - \theta_i\right) > \lambda_i$$

and

$$V^M - \theta_i > \lambda_i$$

for those closer to network 1 and those closer to network 2, respectively. Each of these boundary conditions is displayed in Figure 4.2.

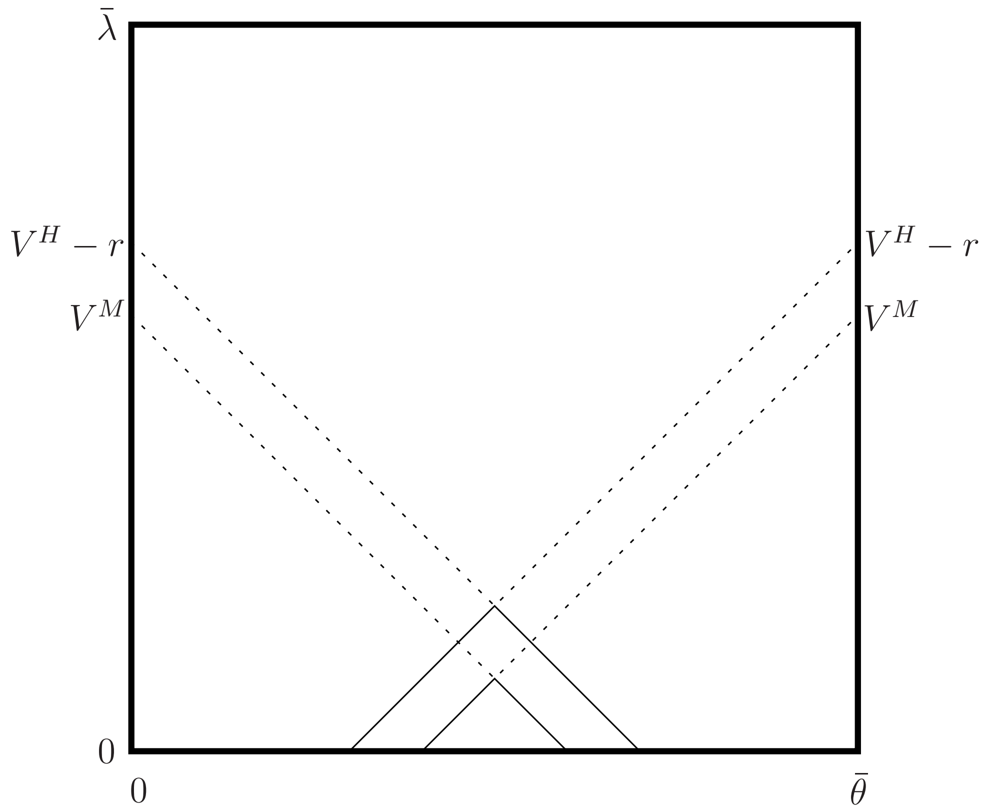


Figure 4.2: Effect of the DVR on Head-to-Head Lineup Outcome

Earlier I ruled out the possibility that a viewer rent the DVR and watch a station's mediocre show without watching its hit show as well. To complete the derivation, one must also rule out the possibility that there exist viewers who do not find enough value in either the hit or mediocre show alone to rent the DVR but do find enough value in both to rent it.

The condition for a network 1 viewer to value watching both shows on network 2 enough to rent a DVR is:

$$\begin{aligned} V^H + V^M - 2\left(\bar{\theta} - \theta_i\right) - r &> 2\lambda_i \\ \implies \left(\frac{V^H + V^M - r}{2}\right) - \left(\bar{\theta} - \theta_i\right) &> \lambda_i \end{aligned}$$

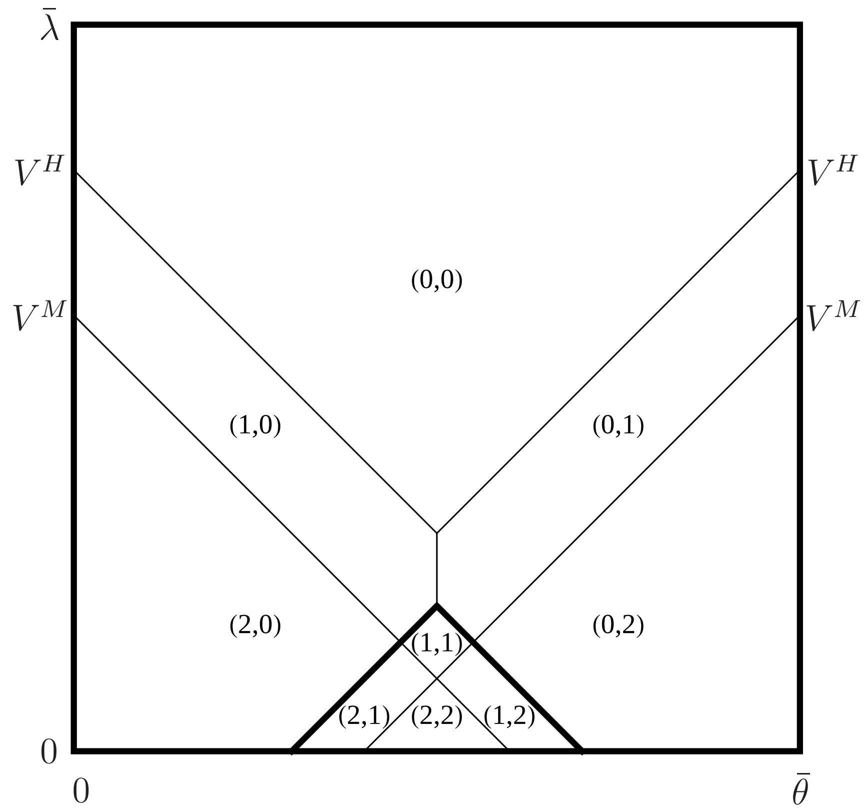


Figure 4.3: Individual Decisions with Head-to-Head Lineups - DVR Available

which is implied by the condition for a viewer to value the hit show enough by itself to rent the DVR:

$$\left(V^H - r\right) - \left(\bar{\theta} - \theta_i\right) > \lambda_i$$

due to the assumption $r < V^H - V^M$. Similarly, for network 2 viewers, the condition that must be satisfied for both network 1 shows to be worth a DVR rental is:

$$\begin{aligned} V^H + V^M - 2\theta_i - r &> 2\lambda_i \\ \Longrightarrow \left(\frac{V^H + V^M - r}{2}\right) - \theta_i &> \lambda_i \end{aligned}$$

which is implied by the condition to rent the DVR for network 2's hit show only:

$$\left(V^H - r\right) - \theta_i > \lambda_i$$

Therefore, the set of conditions displayed in Figure 4.2 characterizes all viewers for whom it is optimal to rent a DVR when facing head-to-head lineups.[3]

Combining the pre-DVR viewer decisions when facing head-to-head lineups with the DVR effect gives the post-DVR viewer decisions, displayed in Figure 4.3. Those who watch both hit shows or at least three shows do so with the help of a DVR; those who watch fewer than three shows do not rent a DVR. The set of DVR users is outlined in dark black in the figure.

Next, I examine the individual decision problem when facing staggered lineups. I follow a similar approach to that carried out in the head-to-head scenario above.[4] Figure 4.4 displays the individuals' decisions with staggered lineups when DVR technology is not available, with (a, b) again representing the decision to watch a shows on network 1 and b shows on network 2.

One category of DVR renters in the head-to-head lineup case, individuals who rent the DVR to watch both hit shows, are absent with staggered lineups since they can already watch both hit shows without a DVR. Instead, DVR renters will be comprised of individuals that would already watch a show in each timeslot without DVR technology and now prefer to watch more television. These viewers fall into three categories. The first contains those who watch both shows on one network pre-DVR and rent a DVR to watch one or both of the other network's shows. These individuals must find the other network's hit show worth renting a DVR on its own, as $V^M < V^H - r$ implies that a viewer wanting to watch a network's mediocre show

[3]For a more rigorous derivation, see Appendix Section D.1.

[4]See Appendix Section D.2 for a more rigorous derivation.

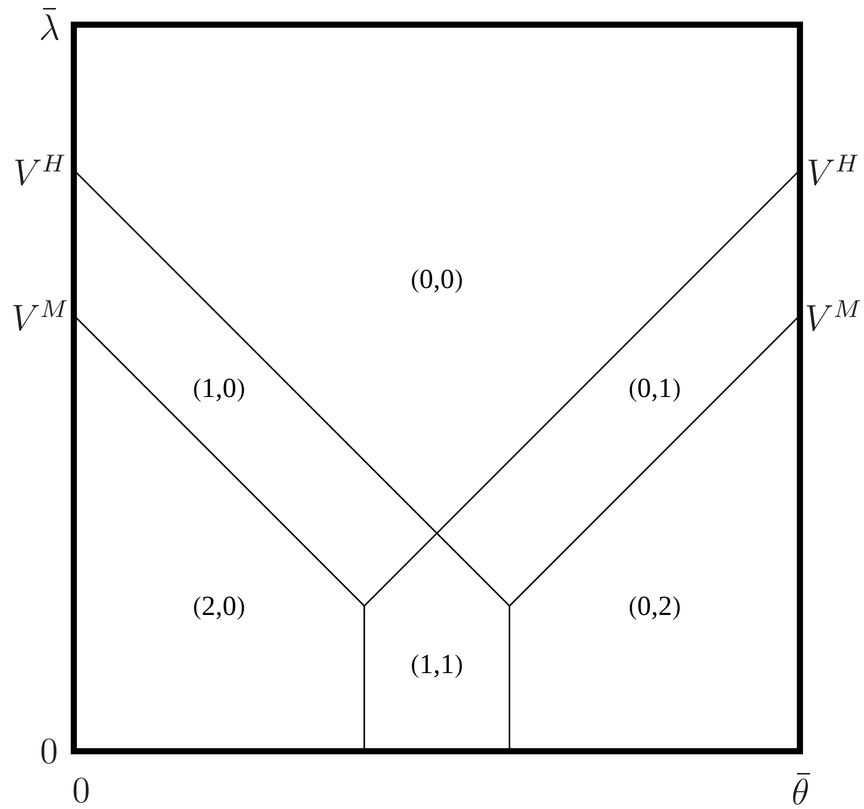


Figure 4.4: Individual Decisions with Staggered Lineups - No DVR Available

at all also finds its hit show worth the DVR's rental cost. The subset of these viewers that also find the other network's mediocre show worth watching on its own do so, as once the DVR is rented there is no marginal rental cost associated with viewing additional programs; the rest rent the DVR and watch the other network's hit show only.

The second category contains viewers who would watch both hit shows but also find at least one of the mediocre shows worth renting a DVR to watch. The indifference lines for the first and second categories are depicted in Figure 4.5. The first category is represented by the two higher dotted lines coming from the λ-intercepts of $V^H - r$ towards the θ-axis. The second category is represented by the two lower dotted lines coming from λ-intercepts of $V^M - r$.

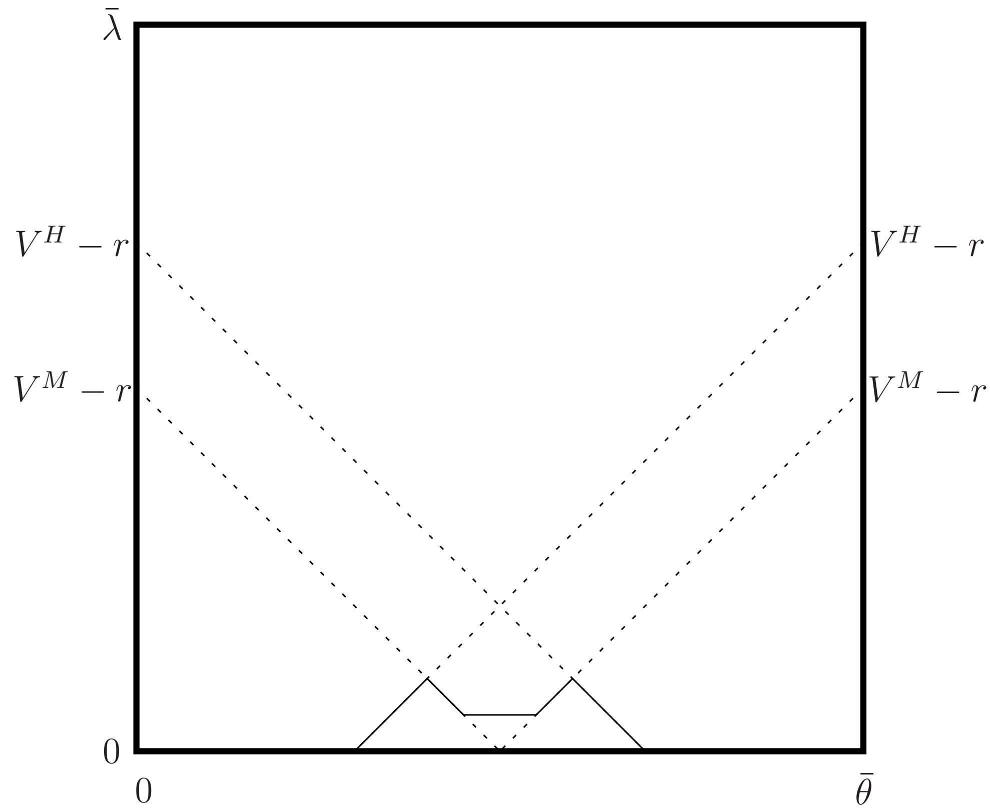


Figure 4.5: Effect of DVR on Staggered Lineup Outcome

There is a third category of viewers as well - those that would normally watch both hit shows and do not find either mediocre show worth renting the DVR for on its own, but do find the combination of both mediocre shows worth renting the DVR for. These individuals comprise the region below the horizontal solid line but above both dotted lines coming from the $\left(V^M - r\right)$-intercepts. The solid line is horizontal due to the nature of watching a show on each program - an individual's θ_i no longer matters as a local movement to the right or left takes the individual closer to one network by the same amount as she moves more distant from the other.

Overall, these three categories comprise the mass of DVR renters when lineups are staggered. Graphically, they consist of all individuals underneath the solid line in Figure 4.5. Adding this to Figure 4.4 yields the individuals' optimal response to

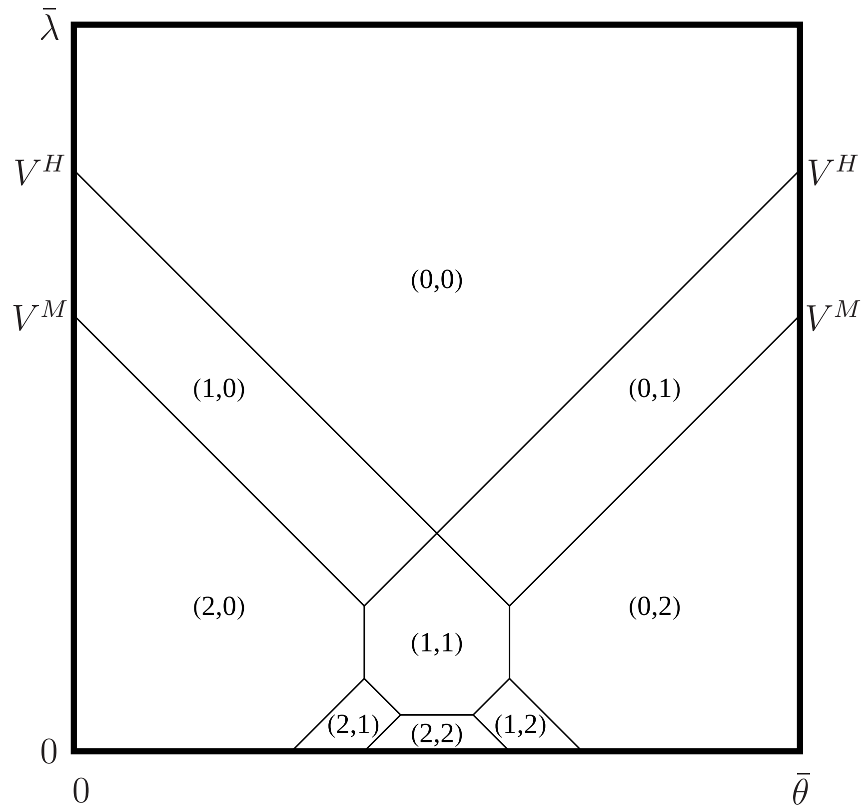


Figure 4.6: Individual Decisions with Staggered Lineups - DVR Available

staggered lineups when the DVR is available, displayed in Figure 4.6. Each individual watches at least as many shows as she did before the DVR was available, and each DVR user watches one or two additional shows depending on her location.

The dynamics of this game create the unique septagonal figure representing those who watch both hit shows only, labeled $(1, 1)$ in Figure 4.6. The evolution of this region captures the essence of the outcome. The top-left and top-right sides reflect the hit show indifference lines that remain from the outcome without DVR technology. Where these two lines intersect the two mediocre shows' indifference lines, there is a drop of length r, representing the DVR's cost — this forms the left and right sides of the septagon. The bottom-left and bottom-right sides are extensions of the mediocre shows' indifference lines after having been dropped down by r, sectioning off those who rent a DVR for the purpose of watching a mediocre show. Lastly, those that do not find watching either mediocre show worth the DVR's cost alone but do find watching both mediocre shows worth the cost appear at the bottom of the diagram, sectioned off from the rest of the septagon, thereby forming its base.

Comparing the optimal responses to each potential lineup outcome, one natural question is how the masses of DVR users compare in each scenario. Intuition suggests that the motivation to watch both hit shows would cause more individuals to rent a DVR with head-to-head lineups; the following proposition formalizes this result.

Proposition 25 *The mass of DVR users is larger with head-to-head lineups than with staggered ones.*

Proof See Section D.3 for a detailed proof. Briefly, as demonstrated in Figures 4.3 and 4.6, the viewers that rent a DVR when the lineups are staggered comprise a subset of those that do when the hit shows are pitted head-to-head. The number of DVR users in each scenario are

$$DVR_{HH} = \left(V^H - r - \frac{\bar{\theta}}{2}\right)^2$$

and

$$DVR_{ST} = \left(V^H - r - \frac{\bar{\theta}}{2}\right)^2 - \frac{\left(V^H - V^M\right)^2}{2} + \frac{r^2}{4}$$

for the head-to-head and staggered scenarios respectively. The assumption $r < V^H - V^M$ ensures the mass of DVR users in the staggered scenario is greater. ∎

4.3 Optimal Network Strategy

Continuing to work backwards in the game's timing, the next determination is each network's optimal response to the other's potential actions, which depends on the profit each network ultimately receives for each possible lineup outcome. As a result of Proposition 25, the lineup choices that provide the networks with the highest profits are not immediately clear. As shown in Section 3.3, when there is no DVR, the overall television audience is maximized when the networks stagger their lineups — each firm is better off having its hit show compete against the other firm's mediocre show. However, the DVR introduces a second effect in which more individuals rent a DVR when the hit shows are pitted head-to-head, many of whom end up watching the mediocre shows as well. In other words, the DVR causes a larger increase in overall viewership in the head-to-head scenario than in the staggered one. As a result, one needs to determine which factor outweighs the other.

Assume that network 1 and network 2 have both chosen (H, M), and network 1 is considering deviating to (M, H). Figure 4.7 displays the tradeoff involved with such a move. Switching from (H, M) to (M, H) would expand the audience for network 1's hit show by the parallelogram labeled A due to the lower quality competition it would face. At the same time, this move would also reduce its mediocre show's audience by the trapezoid labeled B due to higher quality competition in its timeslot as well as a lower incentive for viewers to rent a DVR since they can watch both hit shows without one. Whether or not deviation improves network 1's profit depends on the relative sizes of these two shapes; as it turns out, these two shapes are equal in size, regardless of the parameter values.

Proposition 26 *For any paramter values, when DVRs are available, each network's overall audience (equivalently, profit) is the same size for either lineup decision, re-*

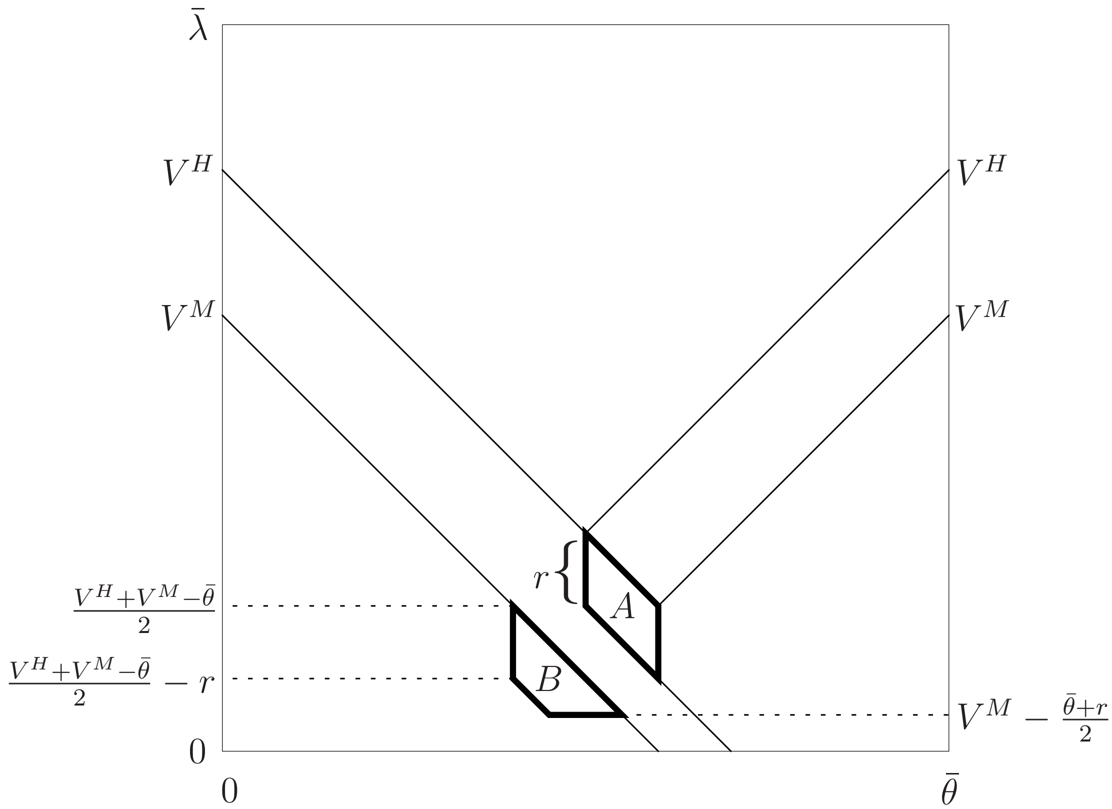


Figure 4.7: Tradeoff between (H, M) and (M, H) in Response to (H, M)

gardless of the lineup its rival chooses.

Proof Facing a lineup of (H, M), the profit difference between choosing (H, M) and (M, H) is represented by $A - B$ from Figure 4.7. The height and width of A are straightforward to determine from the graph, implying an area of:

$$A = r \cdot \left(\frac{V^H - V^M}{2} \right)$$

Viewing B as an isosceles right triangle with a smaller isosceles right triangle cut out of its base, its area is:

$$
\begin{aligned}
&\frac{1}{2}\left(\frac{V^H - V^M + r}{2}\right)^2 - \frac{1}{2}\left(\frac{V^H - V^M - r}{2}\right)^2 \\
&\quad = \frac{1}{2}\left[\left(\frac{V^H - V^M + r}{2} - \frac{V^H - V^M - r}{2}\right)\left(\frac{V^H - V^M + r}{2} + \frac{V^H - V^M - r}{2}\right)\right] \\
&\quad = \frac{1}{2} \cdot r \cdot \left(V^H - V^M\right),
\end{aligned}
$$

the same as A. Symmetry across timeslots implies there is also no profit difference between the two choices when facing a lineup of (M, H), and symmetry across networks then implies that each network's profit is the same in all scenarios. ∎

When there is no DVR, each network strictly prefers to avoid direct time competition to maximize the size of its audience. However, when a DVR is available, viewers have a stronger incentive to rent one with head-to-head lineups in order to watch both hit shows, after which there's no additional marginal cost to watching mediocre shows other than the individual's outside option. The latter incentive counterbalances the former exactly, so that profit for each network remains constant regardless of either network's decision.

At first glance, this result may not seem particularly surprising – after all, if everyone in the market were given a DVR, the networks' lineup choices would of course not affect their profits. Note, however, that this effect remains true for any rental cost $0 < r < \bar{r}$. As r moves, the mass of viewers that rent a DVR in each scenario also moves – and by different amounts, as demonstrated in the proof to Proposition 26. Yet changing r affects the sizes of A and B from Figure 4.7 by the same magnitude; in other words, it has as much of an effect on the change to a network's hit show's audience from switching lineups as it does on the change to the mediocre show's audience from doing so, so that overall, there is no net effect on the (lack of) desire to deviate. Consequently, every viewer need not rent a DVR in order for this result to hold.

In fact, the critical mass of DVR users necessary to discipline the market can be arbitrarily small, as demonstrated by the following proposition.

Proposition 27 *That each firm's profit remains the same regardless of lineup decisions only requires that some positive mass of viewers rent a DVR in equilibrium, and the mass of all viewers that rent a DVR in equilibrium can be an arbitrarily small proportion of the total mass of viewers.*

Proof See Section D.4 in the Appendix.

As a result of Proposition 26, one can use any lineup scenario to calculate equilibrium profit levels. From the $(H, M) \times (H, M)$ scenario, it is straightforward to show that each mediocre show's audience is

$$\pi_{H2H}^{DVR,M} = \frac{V^{M^2}}{2}$$

and the hit show's audience is

$$\pi_{H2H}^{DVR,H} = \frac{\bar{\theta}}{2}\left(V^H - \frac{\bar{\theta}}{4}\right) + \frac{\left(V^H - r - \frac{\bar{\theta}}{2}\right)^2}{2}$$

yielding a total profit of

$$\pi^{DVR} = \frac{1}{2}\left[\left(V^H - r - \frac{\bar{\theta}}{2}\right)^2 + V^{M^2} + \bar{\theta}\left(V^H - \frac{\bar{\theta}}{4}\right)\right]$$

Most of the expected comparative statics for DVR users also hold.

Proposition 28 *The mass of DVR users in both scenarios increases with V^H and decreases with r and $\bar{\theta}$. DVR users increase with V^M in the staggered scenario, whereas V^M has no effect on the DVR mass in the head-to-head scenario.*

Proof See section D.5 in the Appendix.

4.4 Incorporating Ad-Filtering

To this point, the model has assumed the DVR has no ad-filtering effect, only a time-shifting one. I add an ad-filtering effect to the model here. Now suppose that, if an individual watches a show with a DVR, the DVR filters out a proportion $0 < p < 1$ of the ads. Since individuals have no aversion to ads, this aspect of the DVR has no impact on their behavior. Note that this assumption is conservative in the sense that, if individuals were averse to ads, ad-filtering would cause them to watch more TV, so ignoring this effect would, if anything, underestimate the beneficial impact of the DVR on the networks.

Incorporating this new aspect of the DVR, the question of whether or not the DVR helps the networks depends on the value of f, its filtering capability. The previous section essentially explored the special case of this model for which $p = 0$, and found that both lineup scenarios yield the same profit, with staggered lineups yielding more DVR usage than head-to-head ones. Consequently, for any $p > 0$, staggered lineups remains as the only equilibrium since staggered lineups result in a smaller loss due to DVR usage from an equivalent starting point. Consequently, the DVR's impact in this new setting boils down to a comparison between total ad impressions for each network in the staggered scenario without a DVR to that with a DVR.

Proposition 29 *If the DVR eliminates a proportion p of the ad revenue a network* *receives for each DVR user that watches its show, then, for any values of the parameters, there exists* $\hat{p} \in \left(\frac{1}{2}, \frac{1}{3}\right)$ *such that* $\forall p < \hat{p}$*, each network's profit is higher with DVR technology, and* $\forall p > \hat{p}$*, each network's profit is higher without DVR technology for the entire parameter space.*

Since the results here reflect a conservative view of the effect of an ad-filtering DVR on network profits, DVR filter rates under $\frac{1}{3}$ should certainly help networks, whereas filter rates just above $\frac{1}{2}$ may still help them. The actual filter rate of the DVR is still hotly debated, with results ranging from reports that DVRs do not affect buying behavior at all (essentially $p = 0$) to DVRs eliminate the effect of ads entirely ($p = 1$). However, as more in-depth research narrows in on the effect of DVRs on ad impressions, one can then ascertain the effect of the device on network profits.

Chapter 5

Conclusion

Traditional explanations of timing strategies have included the effect of the lead-in audience as a critical driver of network behavior. However, the advent of DVR technology, which allows one to record shows for later consumption, could reduce or even eliminate this effect as its use becomes more popular. In this paper, I have presented a model of time competition across networks to explore how these two effects might impact networks' optimal timing strategies.

When viewers are myopic, equilibrium strategies can involve either head-to-head or staggered lineups. This reflects the networks' competing incentives both to air their hit shows in the first timeslot and to air them at a different time than the rival does, the former of which results from the viewers' myopia. The potential to have staggered lineups in equilibrium lies in contrast to Liu et al. (2004), which concludes that each network would air a higher quality show in the first period than in the second due to the lead-in effect. The existence of the outside option drives this difference, as it allows the overall television audience to change with the networks' strategies. If some viewers might not watch television, each network has an incentive to air its hit show at a different time than its rival to capture as many marginally interested individuals as possible in the timeslot in which its hit show airs.

In equilibrium, whether or not network welfare, viewer welfare, and total welfare are maximized is dependent on the strength of the lead-in effect. At least one network will lead its lineup with its hit show, after which high values of the lead-in effect motivate the other to lead with its hit show, while small values cause it to save its hit show for the second timeslot. Extreme values of the lead-in effect lead to socially optimal outcomes, whereas moderate levels of the lead-in effect engender socially inefficient results. In case of the latter, the effect is strong enough for network 2 to lead with its hit show but not strong enough to outweigh the losses to both stations from competition. Essentially, network 2 does not consider the positive externality network 1 will receive from the lack of direct competition, leading to a socially inefficient outcome when its own benefit from avoiding competition does not outweigh the cost of a smaller lead-in audience. This complicates the problem for a social planner, as the planner's preferences over the lead-in effect's magnitude are not monotonic.

In contrast to myopic viewers, foresighted viewers incorporate the lead-in effect into their first period decisions, essentially making the effect bidirectional. This removes the first of the two network incentives listed above, so that the networks avoid direct competition between their hit shows but both still capture the lead-in effect. Consequently, with foresighted viewers, the equilibrium outcome is socially optimal.

Comparing the forward-looking and myopic treatments, it is not surprising to see that no party is worse off in the foresighted case, while some are strictly better off. Some myopic viewers do not watch a show in the first period that they would be better off watching at a loss to receive the utility boost from the lead-in effect in the second period; this harms both those viewers and the networks they should watch. One implication is that events that educate viewers about future programs and prompt forward thinking, such as tune-in advertisements for other shows, may benefit all parties.

Finally, one might think adding DVR technology to the foresighted case would not affect the lineup outcomes, as all parties have equal access to the DVR at equal cost. However, the mass of DVR users, and therefore its impact on audience size, does vary depending on the lineups the networks play. When the hit shows are pitted head-to-head, there is a stronger incentive for some viewers to rent a DVR than when lineups are staggered, as the DVR is required for a viewer to watch both hit shows. Further, once the DVR is rented, each viewer is free to watch the mediocre shows for no additional rental cost. Consequently, the DVR has a larger impact on audience size with head-to-head lineups than it does with staggered ones.

Surprisingly, this effect cancels out exactly with the advantage staggered lineups had without DVR technology, so that neither network has any net incentive to air one lineup choice over the other. In other words, the availability of DVR technology itself makes each network indifferent between its two lineup choices, regardless of its rival's decision or the values of any parameters, including the rental cost. Essentially, the DVR eliminates the constraints of the timeslots and puts control of the timing in the viewers' hands so that the ordering of shows has no impact on what a DVR user watches or on how much utility she receives from any given show. This result holds in the face of changing DVR rental costs — in fact, the mass of DVR users necessary to discipline the market can be an arbitrarily small percentage of all TV viewers.

Lastly, ad-skipping was introduced to the model to ascertain the effect of changes in the DVR's efficiency at filtering ads on network profits. Under the conservative assumption that individuals are not averse to ads and therefore do not watch more TV due to the DVR's ad-filtering ability, it was shown that if the DVR filters out less than one out of every three ads, it results in a net benefit to the networks, whereas if it filters out more than one out of every two ads, it results in a net loss to the networks. The actual effect of the DVR on ad impressions is still hotly contested, but as research narrows in on its exact effect, one can use the research here to determine the net benefit or cost to the networks that will result once ad prices have adjusted accordingly.

Appendix C

Time-Shifting / Lead-In Effect

The following subsections contain additional discussion of Chapter 3, as well as proofs of the propositions contained therein.

C.1 Individual Decision Process

When there is no DVR available, each individual has three options to weigh:

1. Watching station 1, which provides a utility of $V_{i,1} - \theta_i$,
2. Watching station 2, which provides a utility of $V_{i,2} - \left(\bar{\theta} - \theta_i\right)$, or
3. Utilizing the outside option, which provides a utility of λ_i,

where $V_{i,j}$ represents the benefit to the individual i of watching station j's programming at that point in time, depending on the kind of show that is airing on that station, as well as any lead-in effect, if applicable. From these expressions one can derive the lines representing those indifferent between each pair of options. For those indifferent between station 1's show and the outside option, it must be true that

$$V_{i,1} - \theta_i = \lambda_i \Longrightarrow \lambda_i + \theta_i = V_{i,1},$$

for those indifferent between station 2 and the outside option

$$V_{i,2} - \left(\bar{\theta} - \theta_i\right) = \lambda_i \Longrightarrow \lambda_i - \theta_i = V_{i,2} - \bar{\theta},$$

and for those indifferent between the two stations

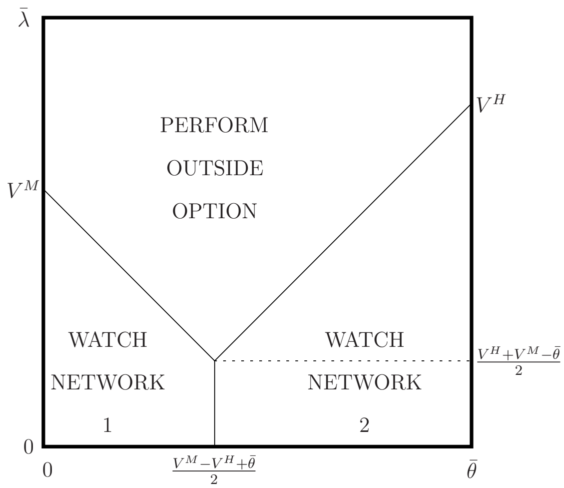
$$V_{i,1} - \theta_i = V_{i,2} - \left(\bar{\theta} - \theta_i\right) \Longrightarrow \theta_i = \frac{V_{i,1} - V_{i,2} + \bar{\theta}}{2}$$

Deriving the optimal choices from there is straightforward. Those above both of the first two indifference lines prefer the outside option, while those below at least one of them prefer to watch television. Of those that watch television, those to the left of the last indifference line watch station 1 and those to the right watch station 2. As an example, the indifference lines and optimal choices for individuals facing a lineup of $(M, H) \times (H, M)$ are demonstrated in Figure C.1.

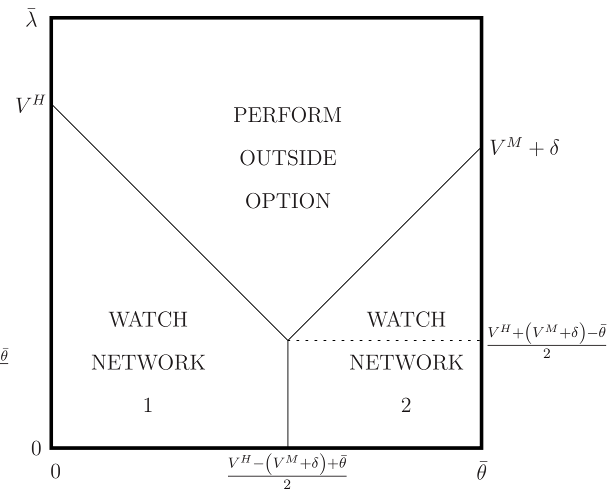
C.2 Optimal Lineup Response to (M, H) with Myopic Viewers

Proposition 16 *When individuals are myopic, the optimal reaction to* (M, H) *is* (H, M).

Proof Assume WLOG that network 1 has chosen (M, H). Network 2 has two options: follow suit with (M, H), or choose the opposite lineup, (H, M), each of which results



(a) Period 1



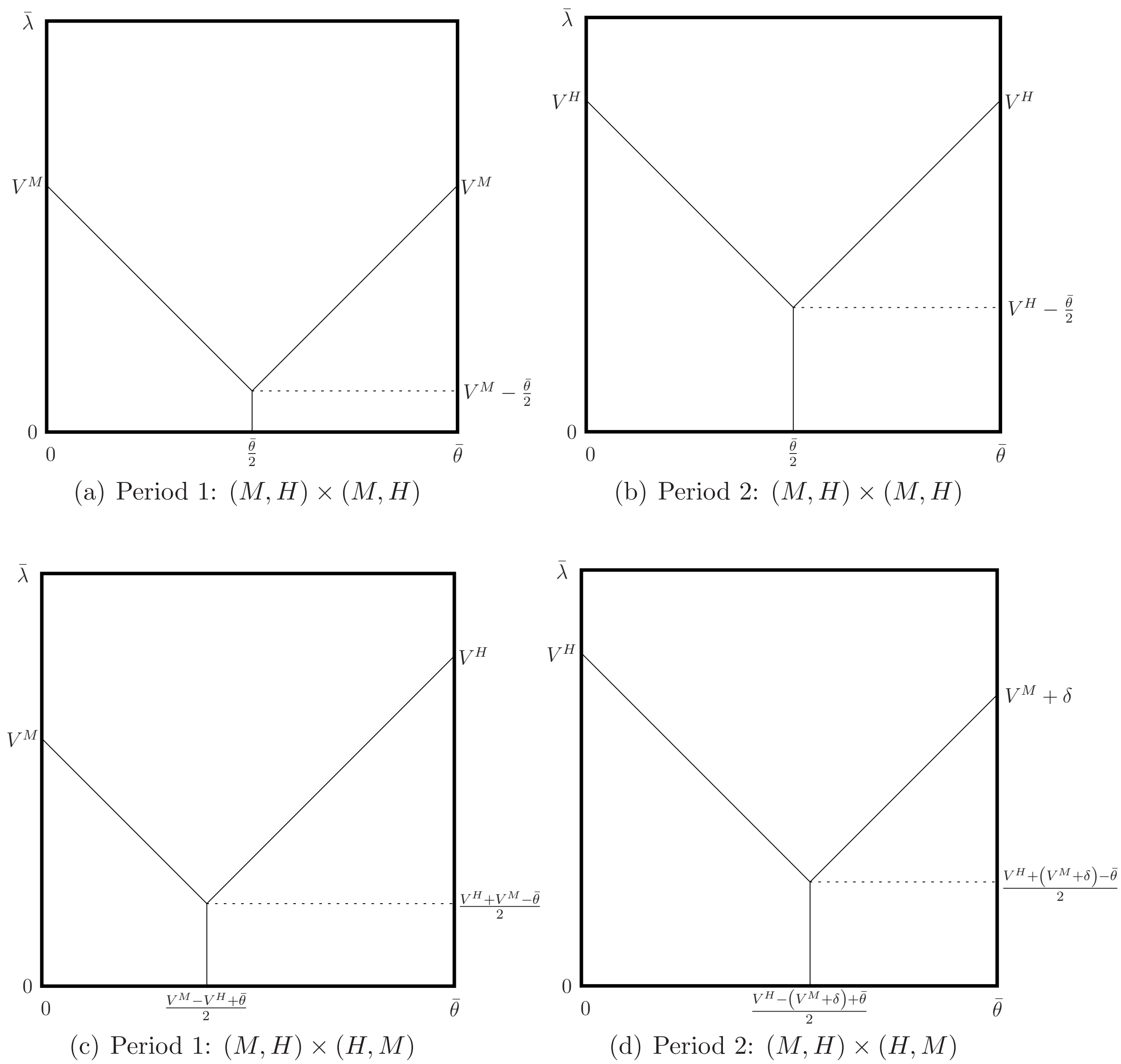
(b) Period 2

Figure C.1: Myopic Individuals' Choices Facing a Lineup of $(M,H)\times(H,M)$

in a different individual choice set in each time slot. Figure C.2 depicts the resulting audience allocation in each time slot for each response by network 2.

If network 2 responds with (M,H), it achieves a total audience of:

$$\begin{aligned}\frac{1}{2}\cdot\frac{\bar{\theta}}{2}\cdot\left(\left(V^M-\frac{\bar{\theta}}{2}\right)+V^M\right) \quad + & \quad \frac{1}{2}\cdot\frac{\bar{\theta}}{2}\cdot\left(\left(V^H-\frac{\bar{\theta}}{2}\right)+V^H\right)\\ = & \quad \frac{\bar{\theta}}{4}\left(2V^M-\frac{\bar{\theta}}{2}\right)+\frac{\bar{\theta}}{4}\left(2V^H-\frac{\bar{\theta}}{2}\right)\\ = & \quad \frac{1}{8}\left(4V^H\bar{\theta}+4V^M\bar{\theta}-2\bar{\theta}^2\right)\end{aligned}$$



(a) Period 1: $(M, H) \times (M, H)$

(b) Period 2: $(M, H) \times (M, H)$

(c) Period 1: $(M, H) \times (H, M)$

(d) Period 2: $(M, H) \times (H, M)$

Figure C.2: Audience Allocation - Potential Responses to (M, H)

If network 2 responds with (H, M), it achieves a total audience of:

$$\begin{aligned}
&\frac{1}{2}\left(V^H + \frac{V^H + V^M - \bar{\theta}}{2}\right)\left(\frac{\bar{\theta} + V^H - V^M}{2}\right) \\
&\quad + \frac{1}{2}\left(\left(V^M + \delta\right) + \frac{\left(V^M + \delta\right) + V^H - \bar{\theta}}{2}\right)\left(\frac{\bar{\theta} + \left(V^M + \delta\right) - V^H}{2}\right) \\
&= \frac{1}{8}\left[\left(3V^H + \left(V^M - \bar{\theta}\right)\right)\left(V^H - \left(V^M - \bar{\theta}\right)\right)\right] \\
&\quad + \frac{1}{8}\left[\left(3\left(V^M + \delta\right) + \left(V^H - \bar{\theta}\right)\right)\left(\left(V^M + \delta\right) - \left(V^H - \bar{\theta}\right)\right)\right] \\
&= \frac{1}{8}\left[3V^{H2} - 2V^H\left(V^M - \bar{\theta}\right) - \left(V^M - \bar{\theta}\right)^2\right] \\
&\quad + \left[3\left(V^M + \delta\right)^2 - 2\left(V^M + \delta\right)\left(V^H - \bar{\theta}\right) - \left(V^H - \bar{\theta}\right)^2\right] \\
&= \frac{1}{8}\left[3V^{H2} - 2V^HV^M + 2V^H\bar{\theta} - V^{M2} + 2V^M\bar{\theta} - \bar{\theta}^2\right] \\
&\quad + \frac{1}{8}\left[3\left(V^M + \delta\right)^2 - 2\left(V^M + \delta\right)V^H + 2\left(V^M + \delta\right)\bar{\theta} - V^{H2} + 2V^H\bar{\theta} - \bar{\theta}^2\right] \\
&= \frac{1}{8}\left[2V^{H2} - 2V^HV^M - V^{M2} + 3\left(V^M + \delta\right)^2 - 2\left(V^M + \delta\right)V^H + 2\delta\bar{\theta}\right] \\
&\quad + \frac{1}{8}\left(4V^H\bar{\theta} + 4V^M\bar{\theta} - 2\bar{\theta}^2\right) \\
&= \frac{1}{8}\left[\left(V^H - V^M\right)^2 + \left(V^H - \left(V^M + \delta\right)\right)^2 + 2\left(\left(V^M + \delta\right)^2 - V^{M2}\right) + 2\delta\bar{\theta}\right] \\
&\quad + \frac{1}{8}\left(4V^H\bar{\theta} + 4V^M\bar{\theta} - 2\bar{\theta}^2\right)
\end{aligned}$$

which is larger than the total audience achieved with a response of (M, H) due to the assumption $\delta, V^M > 0$. ∎

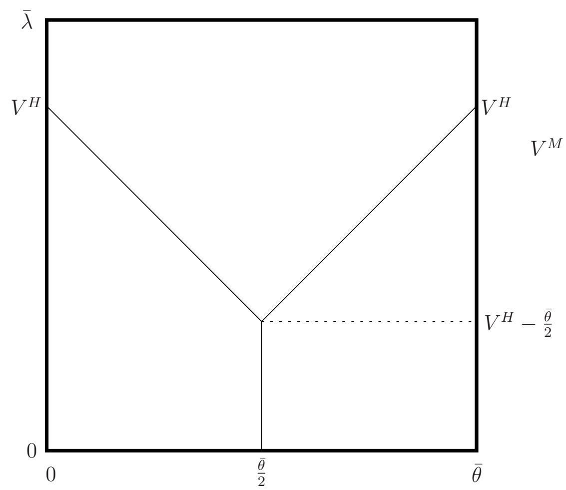
C.3 Optimal Lineup Response to (H, M) with Myopic Viewers

Proposition 17 *There exists $\hat{\delta} \in \left[0, \bar{\delta}\right]$ such that, when individuals are myopic, the optimal response to (H, M) is (M, H) for $\delta < \hat{\delta}$ and (H, M) for $\delta > \hat{\delta}$.*

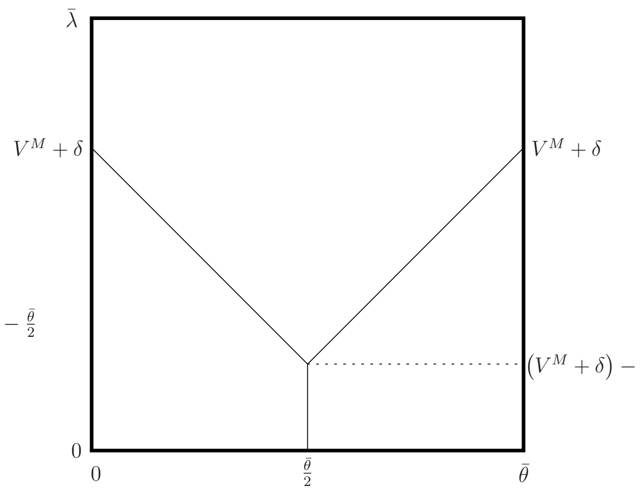
Proof Assume WLOG that network 1 has chosen (H, M). Network 2 has two options: go head-to-head with (H, M) or avoid direct competition by playing (M, H). Each option results in a different individual choice set in each time slot. Figure C.3 depicts the resulting audience allocation in each time slot for each response by network 2.

When network 2 responds with (H, M), it achieves a total audience of:

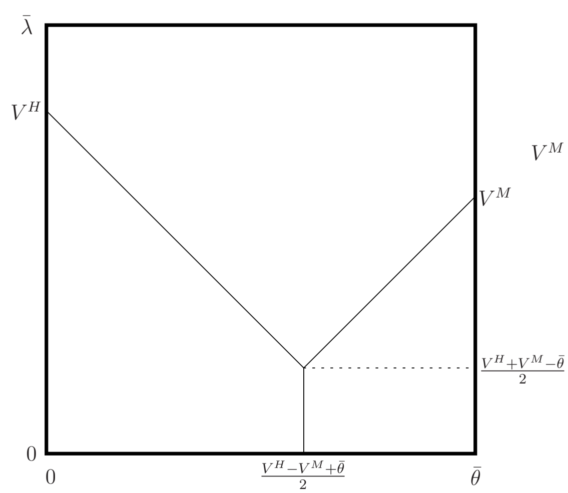
$$\begin{aligned}
\frac{1}{2} \cdot \frac{\bar{\theta}}{2} \cdot \left(\left(V^H - \frac{\bar{\theta}}{2} \right) + V^H \right) \quad + &\quad \frac{1}{2} \cdot \frac{\bar{\theta}}{2} \cdot \left(\left(\left(V^M + \delta \right) - \frac{\bar{\theta}}{2} \right) + \left(V^M + \delta \right) \right) \\
= &\quad \frac{\bar{\theta}}{4} \left(2V^H - \frac{\bar{\theta}}{2} \right) + \frac{\bar{\theta}}{4} \left(2 \left(V^M + \delta \right) - \frac{\bar{\theta}}{2} \right) \\
= &\quad \frac{1}{8} \left(4V^H \bar{\theta} + 4 \left(V^M + \delta \right) \bar{\theta} - 2\bar{\theta}^2 \right)
\end{aligned}$$



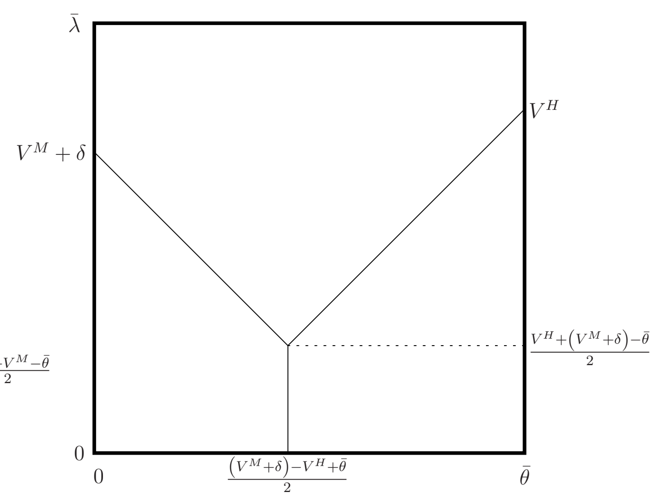
(a) Period 1: $(H, M) \times (H, M)$



(b) Period 2: $(H, M) \times (H, M)$



(c) Period 1: $(H, M) \times (M, H)$



(d) Period 2: $(H, M) \times (M, H)$

Figure C.3: Audience Allocation - Potential Responses to (H, M)

When network 2 responds with (M, H), it achieves a total audience of:

$$
\begin{aligned}
&\frac{1}{2}\left(V^M + \frac{V^H + V^M - \bar{\theta}}{2}\right)\left(\frac{\bar{\theta} + V^M - V^H}{2}\right) \\
&\quad + \frac{1}{2}\left(V^H + \frac{V^H + \left(V^M + \delta\right) - \bar{\theta}}{2}\right)\left(\frac{\bar{\theta} + V^H - \left(V^M + \delta\right)}{2}\right) \\
&= \frac{1}{8}\left[\left(3V^M + \left(V^H - \bar{\theta}\right)\right)\left(V^M - \left(V^H - \bar{\theta}\right)\right)\right] \\
&\quad + \frac{1}{8}\left[\left(3V^H + \left(\left(V^M + \delta\right) - \bar{\theta}\right)\right)\left(V^H - \left(\left(V^M + \delta\right) - \bar{\theta}\right)\right)\right] \\
&= \frac{1}{8}\left[3V^{M2} - 2V^M\left(V^H - \bar{\theta}\right) - \left(V^H - \bar{\theta}\right)^2\right] \\
&\quad + \left[3V^{H2} - 2V^H\left(\left(V^M + \delta\right) - \bar{\theta}\right) - \left(\left(V^M + \delta\right) - \bar{\theta}\right)^2\right] \\
&= \frac{1}{8}\left[3V^{M2} - 2V^H V^M + 2V^M\bar{\theta} - V^{H2} + 2V^H\bar{\theta} - \bar{\theta}^2\right] \\
&\quad + \frac{1}{8}\left[3V^{H2} - 2V^H\left(V^M + \delta\right) + 2V^H\bar{\theta} - \left(V^M + \delta\right)^2 + 2\left(V^M + \delta\right)\bar{\theta} - \bar{\theta}^2\right] \\
&= \frac{1}{8}\left[2V^{H2} - 2V^H\left(V^M + \delta\right) - \left(V^M + \delta\right)^2 + 3V^{M2} - 2V^M V^H - 2\delta\bar{\theta}\right] \\
&\quad + \frac{1}{8}\left(4V^H\bar{\theta} + 4\left(V^M + \delta\right)\bar{\theta} - 2\bar{\theta}^2\right) \\
&= \frac{1}{8}\left[\left(V^H - \left(V^M + \delta\right)\right)^2 + \left(V^H - V^M\right)^2 + 2\left(V^{M2} - \left(V^M + \delta\right)^2\right) - 2\delta\bar{\theta}\right] \\
&\quad + \frac{1}{8}\left(4V^H\bar{\theta} + 4\left(V^M + \delta\right)\bar{\theta} - 2\bar{\theta}^2\right)
\end{aligned}
$$

The first line of the last expression represents the profit difference between network 2's two lineup choices. Note that this is essentially the same as the expression that resulted from network 2's optimal reaction to (M, H) in Section C.2, only with V^M and $\left(V^M + \delta\right)$ having swapped places, and the sign of $2\delta\bar{\theta}$ now negative. As a result of these differences, this expression is not automatically greater than zero, so one must determine its roots. Rearranging terms and equating to zero:

$$\begin{aligned}
-\delta^2 - \delta\left(2\bar{\theta} + 2V^H + 2V^M\right) + \left(2{V^H}^2 - 4V^HV^M + 2{V^M}^2\right) &= 0 \\
\Longrightarrow \frac{\delta^2}{2} + \delta\left(\bar{\theta} + V^H + V^M\right) - \left(V^H - V^M\right)^2 &= 0 \\
\Longrightarrow \pm\sqrt{\left(\bar{\theta} + V^H + V^M\right)^2 + 2\left(V^H - V^M\right)^2} - \left(\bar{\theta} + V^H + V^M\right) &= \delta
\end{aligned}$$

One root is clearly less than zero leaving only the second root:

$$\hat{\delta} = \sqrt{\left(\bar{\theta} + V^H + V^M\right)^2 + 2\left(V^H - V^M\right)^2} - \left(\bar{\theta} + V^H + V^M\right)$$

Note that the original expression is decreasing in δ. As a result, $\delta < \hat{\delta}$ implies the expression is positive, making (M, H) the optimal response to (H, M), while $\delta > \hat{\delta}$ implies the expression is negative, in which case (H, M) is the best response to (H, M).

All that remains is to prove that $\hat{\delta} < \bar{\delta}$:

$$\begin{aligned}
&0 < V^M < V^H < \bar{\theta} \\
&\quad\Longrightarrow 0 < \left(V^H - V^M\right) < 2\left(\bar{\theta} + V^H + V^M\right) \\
&\quad\Longrightarrow \left(V^H - V^M\right)^2 < 2\left(\bar{\theta} + V^H + V^M\right)\left(V^H - V^M\right) \\
&\quad\Longrightarrow 2\left(V^H - V^M\right)^2 < \left(V^H - V^M\right)^2 + 2\left(\bar{\theta} + V^H + V^M\right)\left(V^H - V^M\right) \\
&\quad\Longrightarrow \sqrt{\left(\bar{\theta} + V^H + V^M\right)^2 + 2\left(V^H - V^M\right)^2} < \sqrt{\left(\left(\bar{\theta} + V^H + V^M\right) + \left(V^H - V^M\right)\right)^2} \\
&\quad\Longrightarrow \sqrt{\left(\bar{\theta} + V^H + V^M\right)^2 + 2\left(V^H - V^M\right)^2} < \left(\bar{\theta} + V^H + V^M\right) + \left(V^H - V^M\right) \\
&\quad\Longrightarrow \sqrt{\left(\bar{\theta} + V^H + V^M\right)^2 + 2\left(V^H - V^M\right)^2} - \left(\bar{\theta} + V^H + V^M\right) < \left(V^H - V^M\right) \\
&\quad\Longrightarrow \hat{\delta} < \bar{\delta} \qquad \blacksquare
\end{aligned}$$

C.4 Network Welfare with Myopic Viewers

Proposition 19 *When individuals are myopic, there exists* $\tilde{\delta} \in \left(\hat{\delta}, \bar{\delta}\right)$ *such that, for* $\delta < \tilde{\delta}$*, network welfare is maximized when the networks choose staggered lineups, and for* $\delta > \tilde{\delta}$*, network welfare is maximized when the networks both choose* (H, M).

Proof Clearly, from the standpoint of total viewership, the outcome $(H, M)\times(H, M)$ dominates $(M, H) \times (M, H)$ due to the lead-in effect, and the outcome $(H, M) \times (M, H)$ will provide the same total viewership as $(M, H) \times (H, M)$ due to symmetry. Therefore, comparing $(H, M)\times(H, M)$ to $(H, M)\times(M, H)$ will reveal the conditions under which total viewership is maximized.

Call total viewership in the head-to-head outcome A_{HH} and that in the staggered outcome A_{ST}. Each firm's viewership for each period in both outcomes were derived in Sections C.2 and C.3; from this, we know that

$$A_{HH} = \frac{1}{4}\left(4V^H\bar{\theta} + 4\left(V^M + \delta\right)\bar{\theta} - 2\bar{\theta}^2\right)$$

and that

$$\begin{aligned}
A_{ST} \;&=\; \frac{1}{8}\left[\left(V^H - V^M\right)^2 + \left(V^H - \left(V^M + \delta\right)\right)^2 + 2\left(\left(V^M + \delta\right)^2 - V^{M2}\right) + 2\delta\bar{\theta}\right] \\
&\quad + \frac{1}{8}\left(4V^H\bar{\theta} + 4V^M\bar{\theta} - 2\bar{\theta}^2\right) \\
&\quad + \frac{1}{8}\left[\left(V^H - \left(V^M + \delta\right)\right)^2 + \left(V^H - V^M\right)^2 + 2\left(V^{M2} - \left(V^M + \delta\right)^2\right)\right. \\
&\quad \left. - 2\delta\bar{\theta}\right] + \frac{1}{8}\left(4V^H\bar{\theta} + 4\left(V^M + \delta\right)\bar{\theta} - 2\bar{\theta}^2\right) \\
&= \frac{1}{4}\left[\left(V^H - \left(V^M + \delta\right)\right)^2 + \left(V^H - V^M\right)^2 - 2\delta\bar{\theta}\right] \\
&\quad + \frac{1}{4}\left(4V^H\bar{\theta} + 4\left(V^M + \delta\right)\bar{\theta} - 2\bar{\theta}^2\right) \\
&= \frac{1}{4}\left[\left(V^H - \left(V^M + \delta\right)\right)^2 + \left(V^H - V^M\right)^2 - 2\delta\bar{\theta}\right] + A_{HH}
\end{aligned}$$

The sign of the first term governs which outcome provides the larger overall television audience. To determine its roots:

$$\begin{aligned}
\frac{1}{4}\left[\left(V^H-\left(V^M+\delta\right)\right)^2+\left(V^H-V^M\right)^2-2\delta\bar{\theta}\right] &= 0\\
\Longrightarrow \left(V^H-\left(V^M+\delta\right)\right)^2+\left(V^H-V^M\right)^2 &= 2\delta\bar{\theta}\\
\Longrightarrow {V^H}^2-2V^H\left(V^M+\delta\right)+\left(V^M+\delta\right)^2+{V^H}^2-2V^HV^M+{V^M}^2 &= 2\delta\bar{\theta}\\
\Longrightarrow \delta^2+2V^M\delta-2V^H\delta-2\bar{\theta}\delta+2{V^H}^2-4V^HV^M+2{V^M}^2 &= 0\\
\Longrightarrow \frac{\delta^2}{2}+\delta\left(V^M-V^H-\bar{\theta}\right)+\left(V^H-V^M\right)^2 &= 0\\
\Longrightarrow \left(\bar{\theta}+V^H-V^M\right)\pm\sqrt{\left(\bar{\theta}+V^H-V^M\right)^2-2\left(V^H-V^M\right)^2} &= \delta
\end{aligned}$$

Since δ is bounded above by $\left(V^H-V^M\right)$, one of the two roots is eliminated, leaving only:

$$\tilde{\delta}=\left(\bar{\theta}+V^H-V^M\right)-\sqrt{\left(\bar{\theta}+V^H-V^M\right)^2-2\left(V^H-V^M\right)^2}$$

Since the argument of the square root is smaller than the outside expression squared,

$\tilde{\delta} > 0$. To show $\tilde{\delta} < \bar{\delta}$:

$$
\begin{aligned}
& V^H, V^M < \bar{\theta} \\
& \quad \Longrightarrow \left(V^H - V^M\right) < 2\bar{\theta} \\
& \quad \Longrightarrow \left(V^H - V^M\right)^2 < 2\left(V^H - V^M\right)\bar{\theta} \\
& \quad \Longrightarrow \bar{\theta}^2 < \bar{\theta}^2 + 2\left(V^H - V^M\right)\bar{\theta} - \left(V^H - V^M\right)^2 \\
& \quad \Longrightarrow \bar{\theta}^2 < \left(\bar{\theta} + V^H - V^M\right)^2 - 2\left(V^H - V^M\right)^2 \\
& \quad \Longrightarrow \bar{\theta} < \sqrt{\left(\bar{\theta} + V^H - V^M\right)^2 - 2\left(V^H - V^M\right)^2} \\
& \quad \Longrightarrow \left(\bar{\theta} + V^H - V^M\right) - \sqrt{\left(\bar{\theta} + V^H - V^M\right)^2 - 2\left(V^H - V^M\right)^2} < \left(V^H - V^M\right) \\
& \quad \Longrightarrow \tilde{\delta} < \bar{\delta}
\end{aligned}
$$

This demonstrates that the original expression, which represents $A_{ST} - A_{HH}$,

$$\frac{1}{4}\left[\left(V^H - \left(V^M + \delta\right)\right)^2 + \left(V^H - V^M\right)^2 - 2\delta\bar{\theta}\right]$$

equals 0 at exactly one point for $\delta \in \left[0, \bar{\delta}\right]$, that is, when $\delta = \tilde{\delta}$. Since $V^H > V^M + \delta$, the expression is decreasing in δ, and therefore positive for $\delta < \tilde{\delta}$, and negative for $\delta > \tilde{\delta}$. As a result, $\delta < \tilde{\delta} \Longrightarrow A_{ST} > A_{HH}$ and $\delta > \tilde{\delta} \Longrightarrow A_{ST} < A_{HH}$.

Demonstrating that $\hat{\delta} < \tilde{\delta}$ completes the proof. Since $0 < V^M < V^H < \bar{\theta}$, we know that:

$$
\begin{aligned}
&2\left(V^H - V^M\right)^2 \left[\left(\bar{\theta} + V^H + V^M\right)^2 - \left(\bar{\theta} + V^H - V^M\right)^2 + 2\left(V^H - V^M\right)^2\right] > 0 \\
&\Longrightarrow \left[\left(\left(\bar{\theta} + V^H\right) - V^M\right)\left(\left(\bar{\theta} + V^H\right) + V^M\right)\right]^2 > \\
&\qquad \left[\left(\left(\bar{\theta} + V^H\right) - V^M\right)\left(\left(\bar{\theta} + V^H\right) + V^M\right)\right]^2 - 2\left(V^H - V^M\right)^2 \\
&\qquad \cdot \left[\left(\left(\bar{\theta} + V^H\right) + V^M\right)^2 - \left(\left(\bar{\theta} + V^H\right) - V^M\right)^2\right] - 4\left(V^H - V^M\right)^4 \\
&\Longrightarrow \left[\left(\bar{\theta} + V^H\right)^2 - V^{M2}\right]^2 > \\
&\qquad \left[\left(\left(\bar{\theta} + V^H\right) - V^M\right)^2 - 2\left(V^H - V^M\right)^2\right]\left[\left(\left(\bar{\theta} + V^H\right) + V^M\right)^2 + 2\left(V^H - V^M\right)^2\right]
\end{aligned}
$$

Taking the square root of each side,

$$
\begin{aligned}
&\Longrightarrow \left[\left(\bar{\theta} + V^H\right)^2 - V^{M2}\right] > \\
&\qquad \sqrt{\left[\left(\bar{\theta} + V^H - V^M\right)^2 - 2\left(V^H - V^M\right)^2\right]\left[\left(\bar{\theta} + V^H + V^M\right)^2 + 2\left(V^H - V^M\right)^2\right]} \\
&\Longrightarrow 4\left(\bar{\theta} + V^H\right)^2 - \left[\left(\bar{\theta} + V^H - V^M\right)^2 + \left(\bar{\theta} + V^H + V^M\right)^2\right] > \\
&\qquad 2\sqrt{\left[\left(\bar{\theta} + V^H - V^M\right)^2 - 2\left(V^H - V^M\right)^2\right]\left[\left(\bar{\theta} + V^H + V^M\right)^2 + 2\left(V^H - V^M\right)^2\right]} \\
&\Longrightarrow 4\left(\bar{\theta} + V^H\right)^2 > \\
&\qquad \left(\bar{\theta} + V^H - V^M\right)^2 - 2\left(V^H - V^M\right)^2 + \left(\bar{\theta} + V^H + V^M\right)^2 + 2\left(V^H - V^M\right)^2 + \\
&\qquad 2\sqrt{\left[\left(\bar{\theta} + V^H - V^M\right)^2 - 2\left(V^H - V^M\right)^2\right]\left[\left(\bar{\theta} + V^H + V^M\right)^2 + 2\left(V^H - V^M\right)^2\right]}
\end{aligned}
$$

Again taking the square root of each side,

$$\implies 2\left(\bar{\theta}+V^H\right) >$$
$$\sqrt{\left(\bar{\theta}+V^H-V^M\right)^2-2\left(V^H-V^M\right)^2}+\sqrt{\left(\bar{\theta}+V^H+V^M\right)^2+2\left(V^H-V^M\right)^2}$$
$$\implies \left(\bar{\theta}+V^H-V^M\right)-\sqrt{\left(\bar{\theta}+V^H-V^M\right)^2-2\left(V^H-V^M\right)^2}$$
$$> \sqrt{\left(\bar{\theta}+V^H+V^M\right)^2+2\left(V^H-V^M\right)^2}-\left(\bar{\theta}+V^H+V^M\right)$$
$$\implies \tilde{\delta} > \hat{\delta} \quad \blacksquare$$

C.5 Myopic Viewer Welfare

Proposition 20 *When individuals are myopic, there exists $\delta' \in \left(\tilde{\delta}, \bar{\delta}\right)$ such that, for $\delta < \delta'$, staggered lineups maximize viewer welfare, and for $\delta > \delta'$, viewer welfare is maximized with head-to-head lineups.*

Proof In order to demonstrate that the change in total welfare is ambiguous, one must first derive an expression for the total welfare change and then provide an example of the change going in each direction. For ease of notation, define $V^D = \left(V^M + \delta\right)$. The total welfare change will equal the sum of the network welfare change and the viewer welfare change. The network welfare change was calculated in section C.3 to be

$$\frac{1}{8}\left[\left(V^H-\left(V^M+\delta\right)\right)^2+\left(V^H-V^M\right)^2+2\left(V^{M2}-\left(V^M+\delta\right)^2\right)-2\delta\bar{\theta}\right]$$

The derivation for the viewer welfare change follows the same calculation for foresighted viewers that appears in section C.7, only with a slightly smaller viewer welfare from staggered lineups due to the myopic behavior of those whose favorite station airs its mediocre show first.

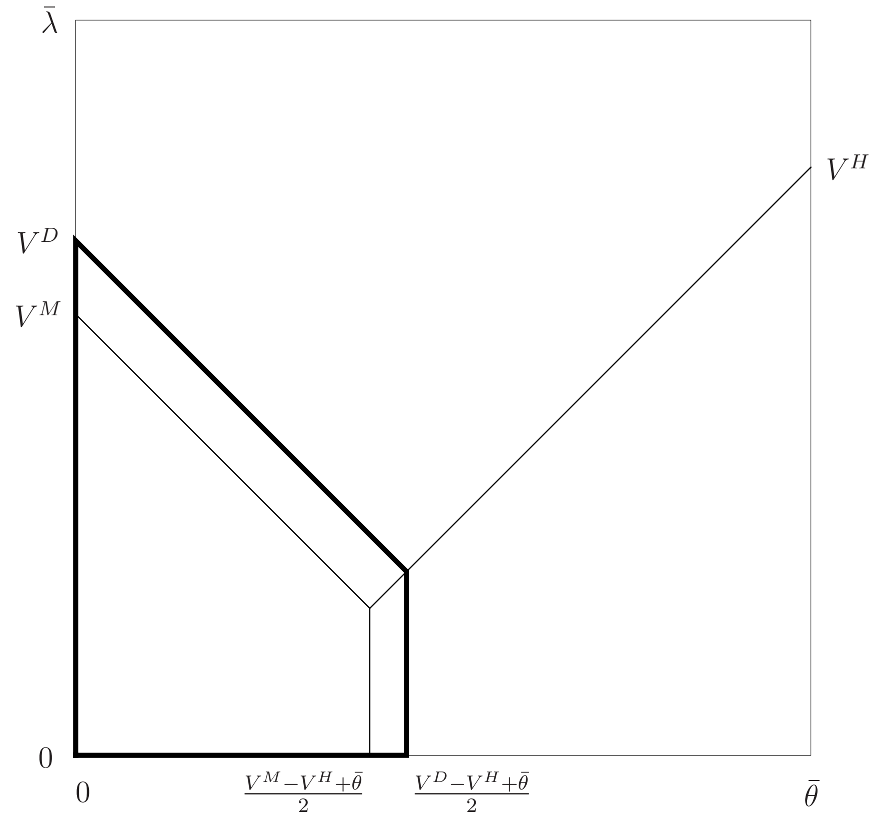


Figure C.4: Viewer Welfare Change from Myopic to Foresighted Viewers

The difference in the change in viewer welfare from staggered lineups is depicted in Figure C.4. When viewers are myopic instead of foresighted, they lose the welfare they would have gained from maximizing their utility over both periods simultaneously. Instead of all the viewers in the larger trapezoid outlined in bold in Figure C.4 receiving this positive utility, only those in the smaller trapezoid contained within do so. The viewer welfare from watching station 1's mediocre show for all those in the larger trapezoid is:

$$W^C = \int_0^{\frac{V^D - V^H + \bar{\theta}}{2}} \int_0^{V^D - \theta_i} \left(V^D - \lambda_i - \theta_i\right) d\lambda_i d\theta_i$$

which, borrowing from calculations in Section C.7, yields a total welfare of

$$\frac{\left(\theta_i - V^H\right)^3}{6}\Bigg|_{\frac{\bar{\theta}}{2}}^{\frac{V^H - V^D + \bar{\theta}}{2}} = \left(\frac{V^{D3}}{6} - \frac{\left(V^D + V^H - \bar{\theta}\right)^3}{48}\right)$$

Similarly for the smaller trapezoid, the total welfare gained is:

$$\begin{aligned}
W^C &= \int_0^{\frac{V^M - V^H + \bar{\theta}}{2}} \int_0^{V^M - \theta_i} \left(V^D - \lambda_i - \theta_i\right) d\lambda_i d\theta_i \\
&= \int_0^{\frac{V^M - V^H + \bar{\theta}}{2}} \left(V^D - \theta_i\right) \lambda_i - \frac{\lambda_i^2}{2}\Bigg|_0^{V^M - \theta_i} d\theta_i \\
&= \int_0^{\frac{V^M - V^H + \bar{\theta}}{2}} \frac{\left(V^M - \theta_i\right)^2}{2} + \delta(V^M - \theta_i) d\theta_i \\
&= \left(\frac{\left(\theta_i - V^M\right)^3}{6} - \frac{\delta(V^M - \theta_i)^2}{2}\right|_{\frac{V^D - V^H + \bar{\theta}}{2}}^{\frac{\bar{\theta}}{2}} \\
&= \left(\frac{V^{M3}}{6} - \frac{(V^M + V^H - \bar{\theta})^3}{48} + \frac{\delta V^{M2}}{2} - \frac{\delta(V^M + V^H - \bar{\theta})^2}{16}\right)
\end{aligned}$$

Subtracting the first expression from the second, doubling the sum, and adding the result to the addition in viewer welfare when moving from head-to-head lineups to staggered ones with foresighted viewers gives the change in viewer welfare for the same in the myopic case. Adding this to the change in network welfare listed above gives an expression for the change in total welfare when moving from head-to-head to staggered lineups.

Evaluating this expression at $V^M = 0.6$, $V^H = 0.8$, and $\bar{\theta} = 1$ gives a negative change if $\delta = 0.03$ but a positive change if $\delta = 0.003$. ∎

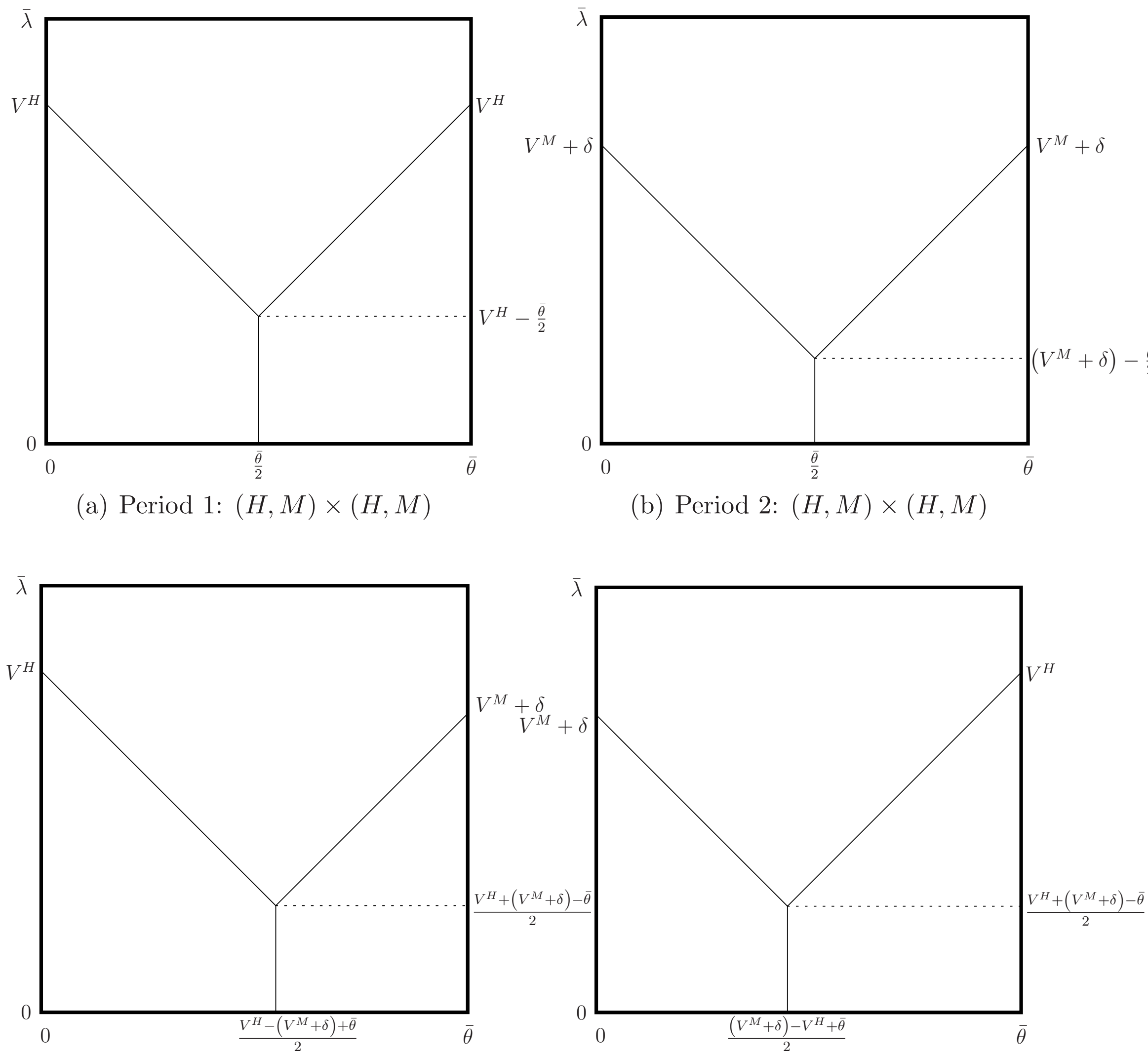
C.6 Foresighted Viewer Equilibrium

Proposition 21 *With foresighted individuals, (H, M) and (M, H) are optimal responses to each other. As a result, only staggered lineups can exist in equilibrium with foresighted viewers.*

Proof I follow an approach similar to that in Section C.3. Assume WLOG that Network 1 chooses (H, M). Network 2 can either go head-to-head (choose (H, M) as well) or avoid direct competition (choose (M, H)), with each option resulting in different individual decisions. Figure C.5 displays the audience allocation that results from each lineup decision.

When network 2 responds with (H, M), it achieves a total audience of:

$$
\begin{aligned}
\frac{1}{2}\cdot\frac{\bar{\theta}}{2}\cdot\left(\left(V^H-\frac{\bar{\theta}}{2}\right)+V^H\right) \quad &+ \quad \frac{1}{2}\cdot\frac{\bar{\theta}}{2}\cdot\left(\left(\left(V^M+\delta\right)-\frac{\bar{\theta}}{2}\right)+\left(V^M+\delta\right)\right) \\
&= \quad \frac{\bar{\theta}}{4}\left(2V^H-\frac{\bar{\theta}}{2}\right)+\frac{\bar{\theta}}{4}\left(2\left(V^M+\delta\right)-\frac{\bar{\theta}}{2}\right) \\
&= \quad \frac{1}{8}\left(4V^H\bar{\theta}+4\left(V^M+\delta\right)\bar{\theta}-2\bar{\theta}^2\right)
\end{aligned}
$$



(a) Period 1: $(H, M) \times (H, M)$

(b) Period 2: $(H, M) \times (H, M)$

(c) Period 1: $(H, M) \times (M, H)$

(d) Period 2: $(H, M) \times (M, H)$

Figure C.5: Audience Allocation for each response to (M, H)

whereas when it responds with (M, H), it achieves a total audience of:

$$
\begin{aligned}
&\frac{1}{2}\left(\left(V^M+\delta\right)+\frac{V^H+\left(V^M+\delta\right)-\bar{\theta}}{2}\right)\left(\frac{\bar{\theta}+\left(V^M+\delta\right)-V^H}{2}\right) \\
&\qquad +\frac{1}{2}\left(V^H+\frac{V^H+\left(V^M+\delta\right)-\bar{\theta}}{2}\right)\left(\frac{\bar{\theta}+V^H-\left(V^M+\delta\right)}{2}\right) \\
&= \frac{1}{8}\left[\left(3\left(V^M+\delta\right)+\left(V^H-\bar{\theta}\right)\right)\left(\left(V^M+\delta\right)-\left(V^H-\bar{\theta}\right)\right)\right] \\
&\quad +\frac{1}{8}\left[\left(3V^H+\left(\left(V^M+\delta\right)-\bar{\theta}\right)\right)\left(V^H-\left(\left(V^M+\delta\right)-\bar{\theta}\right)\right)\right] \\
&= \frac{1}{8}\left[3\left(V^M+\delta\right)^2-2\left(V^M+\delta\right)\left(V^H-\bar{\theta}\right)-\left(V^H-\bar{\theta}\right)^2\right] \\
&\quad +\left[3V^{H2}-2V^H\left(\left(V^M+\delta\right)-\bar{\theta}\right)-\left(\left(V^M+\delta\right)-\bar{\theta}\right)^2\right] \\
&= \frac{1}{8}\left[3\left(V^M+\delta\right)^2-2V^H\left(V^M+\delta\right)+2\left(V^M+\delta\right)\bar{\theta}-V^{H2}+2V^H\bar{\theta}-\bar{\theta}^2\right] \\
&\quad +\frac{1}{8}\left[3V^{H2}-2V^H\left(V^M+\delta\right)+2V^H\bar{\theta}-\left(V^M+\delta\right)^2+2\left(V^M+\delta\right)\bar{\theta}-\bar{\theta}^2\right] \\
&= \frac{1}{8}\left[2V^{H2}-4V^H\left(V^M+\delta\right)+2\left(V^M+\delta\right)^2\right] \\
&\quad +\frac{1}{8}\left[4V^H\bar{\theta}+4\left(V^M+\delta\right)\bar{\theta}-2\bar{\theta}^2\right] \\
&= \frac{\left(V^H-\left(V^M+\delta\right)\right)^2}{4}+\frac{1}{8}\left(4V^H\bar{\theta}+4\left(V^M+\delta\right)\bar{\theta}-2\bar{\theta}^2\right)
\end{aligned}
$$

which is the profit it receives from (H, M) plus the first term, which is strictly positive.

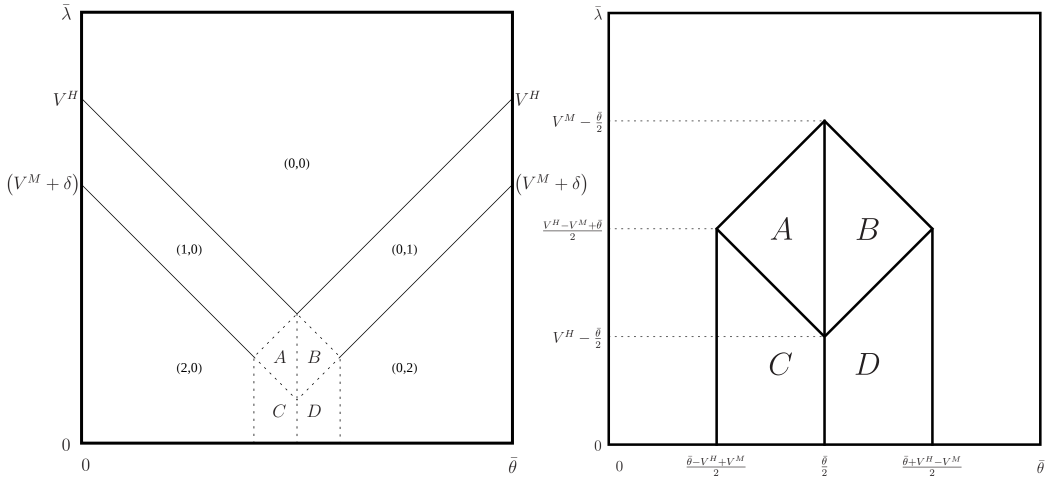
$\blacksquare$

C.7 Foresighted Viewer Welfare

Proposition 23 *When viewers are foresighted, staggered lineups maximize viewer welfare.*

Proof The related figures in the text are displayed again as Figure C.6.

The difference in viewer welfare when changing from staggered to head-to-head



(a) Viewer Decision Comparison

(b) Close-up View

Figure C.6: Foresighted Viewer Welfare - Head-to-Head vs. Staggered Lineups

lineups equals the welfare of B and D watching 1's hit show minus the welfare of C watching 1's mediocre show, plus the welfare of A and C watching 2's hit show minus the welfare of D watching 2's mediocre show. The effect from 1's shows are the same as those from 2's due to symmetry; therefore, I focus solely on the effect of 1's shows below.

For ease of notation, define $V^D \equiv \left(V^M + \delta\right)$.

$$W^{B+D} = \int_{\frac{\bar\theta}{2}}^{\frac{V^H - V^D + \bar\theta}{2}} \int_0^{V^H - \theta_i} \left(V^H - \lambda_i - \theta_i\right) d\lambda_i d\theta_i$$

$$= \int_{\frac{\bar\theta}{2}}^{\frac{V^H - V^D + \bar\theta}{2}} \left(V^H - \theta_i\right)\lambda_i - \frac{\lambda_i^2}{2}\bigg|_0^{V^H - \theta_i} d\theta_i$$

$$= \int_{\frac{\bar\theta}{2}}^{\frac{V^H - V^D + \bar\theta}{2}} \frac{\left(V^H - \theta_i\right)^2}{2} d\theta_i$$

$$= \frac{\left(\theta_i - V^H\right)^3}{6}\Bigg|_{\frac{\bar\theta}{2}}^{\frac{V^H - V^D + \bar\theta}{2}}$$

$$= \frac{1}{6}\left(\left(V^H - \frac{\bar\theta}{2}\right)^3 - \left(\frac{V^H - V^D + \bar\theta}{2}\right)^3\right)$$

$$= \frac{1}{6}\left(\frac{V^H - V^D}{2}\right) \cdot$$

$$\left[\left(V^H - \frac{\bar\theta}{2}\right)^2 + \left(V^H - \frac{\bar\theta}{2}\right)\left(\frac{V^H + V^D - \bar\theta}{2}\right) + \left(\frac{V^H + V^D - \bar\theta}{2}\right)^2\right]$$

$$\begin{aligned}
W^C &= \int_{\frac{V^D-V^H+\bar{\theta}}{2}}^{\frac{\bar{\theta}}{2}} \int_0^{V^D-\theta_i} \left(V^D-\lambda_i-\theta_i\right) d\lambda_i d\theta_i \\
&= \int_{\frac{V^D-V^H+\bar{\theta}}{2}}^{\frac{\bar{\theta}}{2}} \left(V^D-\theta_i\right)\lambda_i - \frac{\lambda_i^2}{2}\bigg|_0^{V^D-\theta_i} d\theta_i \\
&= \int_{\frac{V^D-V^H+\bar{\theta}}{2}}^{\frac{\bar{\theta}}{2}} \frac{\left(V^D-\theta_i\right)^2}{2} d\theta_i \\
&= \frac{\left(\theta_i-V^D\right)^3}{6}\Bigg|_{\frac{V^D-V^H+\bar{\theta}}{2}}^{\frac{\bar{\theta}}{2}} \\
&= \frac{1}{6}\left(\left(\frac{V^H+V^D-\bar{\theta}}{2}\right)^3 - \left(V^D-\frac{\bar{\theta}}{2}\right)^3\right) \\
&= \frac{1}{6}\left(\frac{V^H-V^D}{2}\right)\cdot \\
&\qquad \left[\left(\frac{V^H+V^D-\bar{\theta}}{2}\right)^2 + \left(V^D-\frac{\bar{\theta}}{2}\right)\left(\frac{V^H+V^D-\bar{\theta}}{2}\right) + \left(V^D-\frac{\bar{\theta}}{2}\right)^2\right]
\end{aligned}$$

The welfare change due to station 1's shows is $\left(W^{B+D}-W^C\right)$. Since the welfare change from station 2's shows is the same, the welfare change from both stations' shows put together is:

$$
\begin{aligned}
&2\left(W^{B+D}-W^{C}\right)\\
&= \frac{1}{3}\left(\frac{V^H-V^D}{2}\right)\left[\left(V^H-\frac{\bar{\theta}}{2}\right)^2-\left(V^D-\frac{\bar{\theta}}{2}\right)^2+\left(\frac{V^H+V^D-\bar{\theta}}{2}\right)\left(V^H-V^D\right)\right]\\
&= \left(\frac{V^H-V^D}{6}\right)\left[\left(V^H-V^D\right)\left(V^H+V^D-\bar{\theta}\right)+\left(\frac{V^H+V^D-\bar{\theta}}{2}\right)\left(V^H-V^D\right)\right]\\
&= \frac{\left(V^H-V^D\right)^2}{6}\cdot\frac{3}{2}\cdot\left(V^H+V^D-\bar{\theta}\right)\\
&= \frac{1}{4}\left(V^H-V^D\right)^2\left(V^H+V^D-\bar{\theta}\right)
\end{aligned}
$$

which is necessarily positive due to the assumption $V^H > V^D > \frac{\bar{\theta}}{2}$. ∎

Appendix D

Time-Shifting / Digital Video Recorder

The following subsections contain additional discussion of Chapter 4, as well as proofs of the propositions contained therein.

D.1 Consumer Response - Head-to-Head Lineups

Since the decisions of individuals for whom $\theta_i > \frac{\bar{\theta}}{2}$ will simply be a reflection of those for whom $\theta_i < \frac{\bar{\theta}}{2}$ about the midline $\lambda = \frac{\bar{\theta}}{2}$, WLOG I focus solely on the latter set of viewers. Define (a, b) as the decision to watch a shows on network 1 and b shows on network 2 (note that if a viewer watches one show on a given network, it is the hit show). It is straightforward to see that these viewers would not watch a show on network 2 without also watching the equivalent-quality show on network 1 as well, so we can restrict the analysis to decisions of the form (a, b) for which $a \geq b$, or {(0,0), (1,0), (1,1), (2,0), (2,1), (2,2)}. A DVR rental is only required when the individual watches two shows that air simultaneously, namely, for (2,1) and (2,2). Using the definition of $U_i^{a,b}$ from section 3.1 — that is, the excess utility one receives from the decision (a, b) above (0,0), or the decision not to watch television at all — the utility for individual i for each decision is as follows:

$$\begin{aligned}
U_i^{0,0} &= 0, \\
U_i^{1,0} &= V^H - \theta_i - \lambda_i, \\
U_i^{1,1} &= \big(V^H - \theta_i - \lambda_i\big) + \big(V^H - \big(\bar{\theta} - \theta_i\big) - \lambda_i\big) \\
&= 2V^H - \bar{\theta} - 2\lambda_i, \\
U_i^{2,0} &= \big(V^H - \theta_i - \lambda_i\big) + \big(V^M - \theta_i - \lambda_i\big) \\
&= V^H + V^M - 2\theta_i - 2\lambda_i, \\
U_i^{2,1} &= \big(V^H - \theta_i - \lambda_i\big) + \big(V^M - \theta_i - \lambda_i\big) + \big(V^H - \big(\bar{\theta} - \theta_i\big) - \lambda_i\big) - r \\
&= 2V^H + V^M - \bar{\theta} - \theta_i - 3\lambda_i - r, \\
U_i^{2,2} &= \big(V^H - \theta_i - \lambda_i\big) + \big(V^M - \theta_i - \lambda_i\big) + \big(V^H - \big(\bar{\theta} - \theta_i\big) - \lambda_i\big) \\
&\qquad + \big(V^M - \big(\bar{\theta} - \theta_i\big) - \lambda_i\big) - r \\
&= 2V^H + 2V^M - 2\bar{\theta} - 4\lambda_i - r
\end{aligned}$$

The graph displaying the optimal choice for each individual (Figure 4.3 in the text) is repeated below as Figure D.1. To determine each individual's optimal response, I examine each choice in turn and, for each choice, I show that the set of individuals for whom $\theta_i < \frac{\bar{\theta}}{2}$ who find that choice optimal matches the set depicted in the graph.

Beginning with (0,0), note that $V^H < \theta_i + \lambda_i$ along with the assumptions $V^M < V^H$, $\theta_i < \frac{\bar{\theta}}{2}$, and $r > 0$ imply all other decisions yield negative net utility. Additionally, $V^H > \theta_i + \lambda_i$ implies (1,0) dominates (0,0). Consequently, individuals choose (0,0) if and only if $V^H < \theta_i + \lambda_i$; this corresponds to the (0,0) area in Figure D.1. All other individuals satisfy $V^H > \theta_i + \lambda_i$ and choose to watch at least one show; hence, all

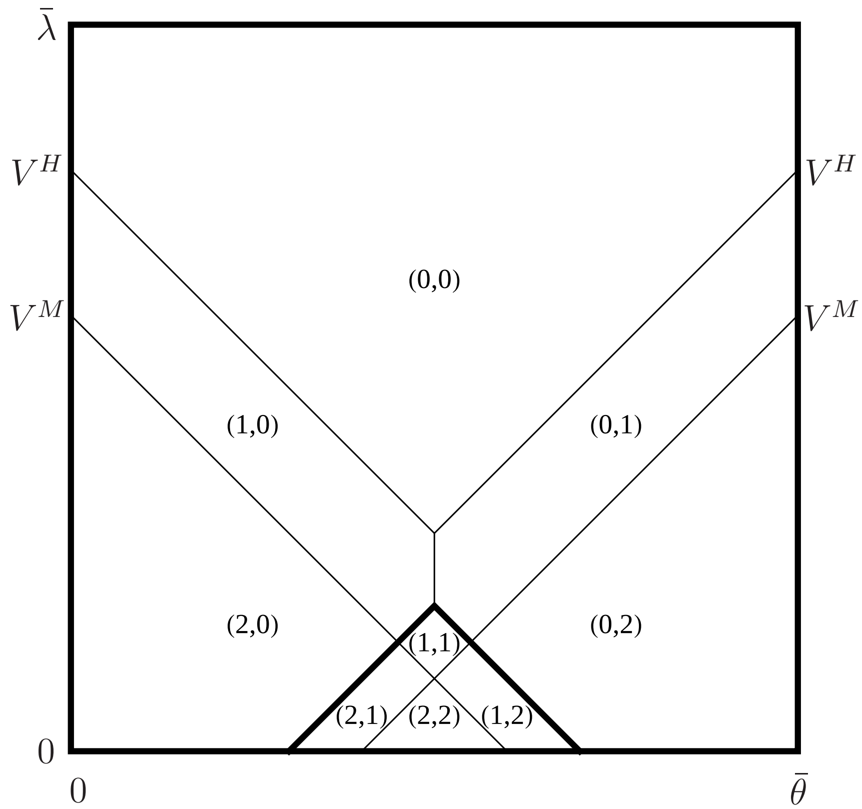


Figure D.1: Individual Decisions with Head-to-Head Lineups - DVR Available

remaining individuals are viewers.

Examining those for whom $V^H > \theta_i + \lambda_i$:

$$V^H < \lambda_i + \left(\bar{\theta} - \theta_i\right) + r \iff U_i^{1,0} - U_i^{1,1} = -V^H + \lambda_i + \left(\bar{\theta} - \theta_i\right) + r > 0,$$

$$V^M < \theta_i + \lambda_i \iff U_i^{1,0} - U_i^{2,0} = -V^M + \theta_i + \lambda_i > 0$$

and

$$\left(V^H < \lambda_i + \left(\bar{\theta} - \theta_i\right) + r\right) \wedge \left(V^M < \theta_i + \lambda_i\right) \Longrightarrow$$

$$U_i^{1,0} - U_i^{2,1} = -V^H - V^M + 2\lambda_i + \bar{\theta} + r > 0,$$

$$U_i^{1,0} - U_i^{2,2} = -V^H - 2V^M + 3\lambda_i + 2\bar{\theta} - \theta_i + r > 0$$

again due to the additional assumptions that $V^M < V^H$, $\theta_i < \frac{\bar{\theta}}{2}$, and $r > 0$. In other words, if both inequalities are satisified, (1,0) dominates all remaining options, and if either assumption is violated, another option dominates (1,0). Therefore, viewers find (1,0) optimal iff $V^H < \lambda_i + \left(\bar{\theta} - \theta_i\right)$ and $V^M < \theta_i + \lambda_i$, which defines the area of the graph labeled (1,0).

Next, consider viewers for whom $V^M > \lambda_i + \left(\bar{\theta} - \theta_i\right)$:

$$\begin{aligned} V^M > \lambda_i + \left(\bar{\theta} - \theta_i\right) \iff \quad & U_i^{2,2} - U_i^{2,1} = V^M - \left(\lambda_i + \left(\bar{\theta} - \theta_i\right)\right) > 0 \\ & U_i^{2,2} - U_i^{2,0} = V^H + V^M + 2\lambda_i - 2\left(\bar{\theta} - \theta_i\right) - r > 0, \\ & U_i^{2,2} - U_i^{1,1} = 2\left[V^M - \left(\lambda_i + \frac{\bar{\theta}}{2}\right)\right] > 0 \end{aligned}$$

since $\lambda_i > 0$, $\theta_i < \frac{\bar{\theta}}{2}$, and $V^H > \lambda_i + \left(\bar{\theta} - \theta_i\right) + r$ since $V^H - r > V^M$, which is implied by $V^M > \lambda_i + \left(\bar{\theta} - \theta_i\right)$. Therefore, viewers for whom $V^M > \lambda_i + \left(\bar{\theta} - \theta_i\right)$ find (2,2) optional, and all other viewers do not.

Now consider $V^H < \lambda_i + \left(\bar{\theta} - \theta_i\right) + r$ and $V^M > \theta_i + \lambda_i$:

$$V^H < \lambda_i + \left(\bar{\theta} - \theta_i\right) + r \iff U_i^{2,0} - U_i^{2,1} = -V^H + \lambda_i + \left(\bar{\theta} - \theta_i\right) + r > 0$$

and

$$\begin{gathered} \left(V^H < \lambda_i + \left(\bar{\theta} - \theta_i\right) + r\right) \wedge \left(V^M > \lambda_i + \theta_i\right) \Longrightarrow \\ U_i^{2,0} - U_i^{1,1} = -V^H + V^M + \bar{\theta} - 2\theta_i + r > 0, \end{gathered}$$

since $r > 0$. Therefore, those for whom $V^M > \lambda_i + \theta_i$ and $V^H < \lambda_i + \left(\bar{\theta} - \theta_i\right) + r$ prefer (2,0) to all other options, and those for whom the second inequality does not hold prefer another option. Earlier it was shown that $V^M < \lambda_i + \theta_i$ implies (1,0) provides more utility than (2,0). Together with the above results, this implies that viewers choose (2,0) if and only if $V^M > \lambda_i + \theta_i$ and $V^H < \lambda_i + \left(\bar{\theta} - \theta_i\right) + r$, which corresponds to the area labeled (2,0) in the graph.

This leaves only two potential options for the remaining viewers — (1,1) and (2,1), which are separated in Figure D.1 by the line $V^M = \theta_i + \lambda_i$. Since

$$V^M > \theta_i + \lambda_i \Longleftrightarrow U_i^{2,1} - U_i^{1,1} = V^M - \theta_i - \lambda_i > 0$$

the remaining viewers for whom $V^M > \theta_i + \lambda_i$ choose (2,1) and the rest choose (1,1), which corresponds to the remaining two areas in Figure D.1 and therefore completes the proof. ∎

D.2 Consumer Response - Staggered Lineups

This proof closely follows that in Section D.1. Using the same reasoning as the proof in that section, I consider only those for whom $\theta_i < \frac{\bar{\theta}}{2}$ WLOG and only decisions (a, b) such that $a \geq b$. Note that if a viewer watches only one show on a given network, it must be the hit show. $U_i^{a,b}$ represents the excess utility one receives from the decision (a, b) above what she would receive from (0,0), the decision not to watch television at all.

The utility for individual i for each decision is repeated again here:

$$
\begin{aligned}
U_i^{0,0} &= 0, \\
U_i^{1,0} &= V^H - \theta_i - \lambda_i, \\
U_i^{1,1} &= \big(V^H - \theta_i - \lambda_i\big) + \big(V^H - \big(\bar{\theta} - \theta_i\big) - \lambda_i\big) \\
&= 2V^H - \bar{\theta} - 2\lambda_i, \\
U_i^{2,0} &= \big(V^H - \theta_i - \lambda_i\big) + \big(V^M - \theta_i - \lambda_i\big) \\
&= V^H + V^M - 2\theta_i - 2\lambda_i, \\
U_i^{2,1} &= \big(V^H - \theta_i - \lambda_i\big) + \big(V^M - \theta_i - \lambda_i\big) + \big(V^H - \big(\bar{\theta} - \theta_i\big) - \lambda_i\big) - r \\
&= 2V^H + V^M - \bar{\theta} - \theta_i - 3\lambda_i - r, \\
U_i^{2,2} &= \big(V^H - \theta_i - \lambda_i\big) + \big(V^M - \theta_i - \lambda_i\big) + \big(V^H - \big(\bar{\theta} - \theta_i\big) - \lambda_i\big) \\
&\qquad + \big(V^M - \big(\bar{\theta} - \theta_i\big) - \lambda_i\big) - r \\
&= 2V^H + 2V^M - 2\bar{\theta} - 4\lambda_i - r
\end{aligned}
$$

The graph displaying the optimal choice for each individual (Figure 4.6 in the text) is repeated below as Figure D.2. I examine each choice in turn below and show that, for each choice, the set of individuals for whom $\theta_i < \frac{\bar{\theta}}{2}$ who find that choice optimal matches the set depicted in the graph.

Beginning with (0,0), note that $V^H < \theta_i + \lambda_i$ along with the assumptions $V^M < V^H$, $\theta_i < \frac{\bar{\theta}}{2}$, and $r > 0$ imply all other decisions yield negative net utility. Additionally, $V^H > \theta_i + \lambda_i$ implies (1,0) dominates (0,0). Consequently, individuals choose (0,0) if and only if $V^H < \theta_i + \lambda_i$; this corresponds to the area marked (0,0) in Figure D.2. All other individuals satisfy $V^H > \theta_i + \lambda_i$ and choose to watch at least one show; therefore, all remaning individuals are viewers.

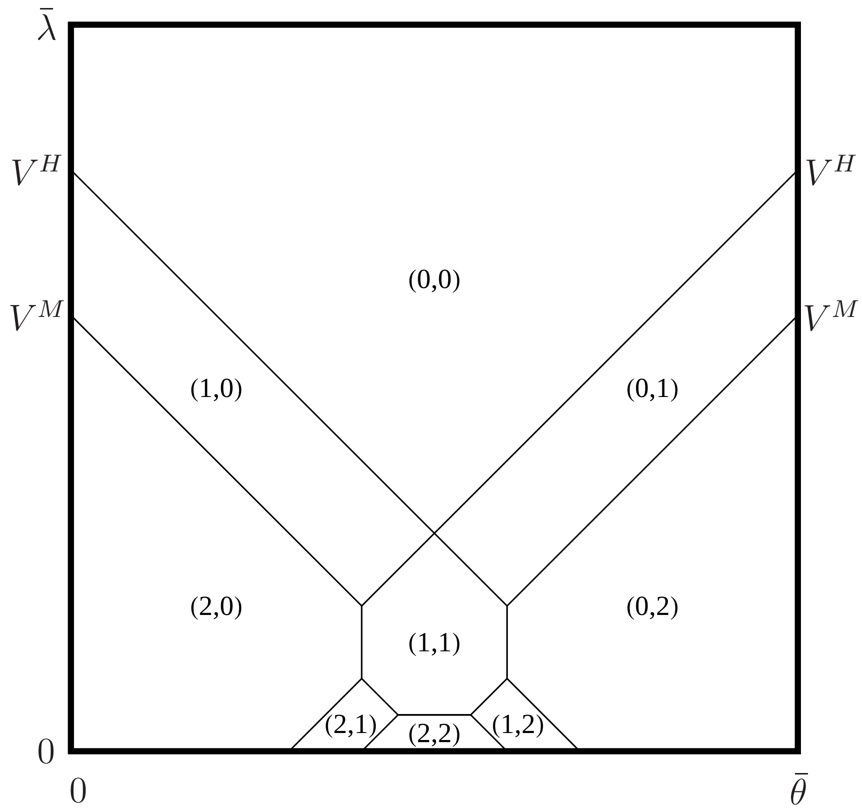


Figure D.2: Individual Decisions with Staggered Lineups - DVR Available

Examining those for whom $V^H > \theta_i + \lambda_i$:

$$V^H < \lambda_i + \left(\bar{\theta} - \theta_i\right) \iff U_i^{1,0} - U_i^{1,1} = -V^H + \left(\bar{\theta} - \theta_i\right) + \lambda_i > 0,$$

$$V^M < \theta_i + \lambda_i \iff U_i^{1,0} - U_i^{2,0} = -V^M + \theta_i + \lambda_i > 0$$

and

$$\left(V^H < \lambda_i + \left(\bar{\theta} - \theta_i\right)\right) \wedge \left(V^M < \theta_i + \lambda_i\right) \Longrightarrow$$

$$U_i^{1,0} - U_i^{2,1} = -V^H - V^M + \bar{\theta} + 2\lambda_i + r > 0,$$

$$U_i^{1,0} - U_i^{2,2} = -V^H - 2V^M + 2\bar{\theta} - \theta_i + 3\lambda_i + r > 0$$

again due to the additional assumptions that $V^M < V^H$, $\theta_i < \frac{\bar{\theta}}{2}$, and $r > 0$. In other words, if both inequalities are satisified, (1,0) dominates all remaining options, and if either assumption is violated, another option dominates (1,0). Therefore, viewers find (1,0) optimal iff $V^H < \lambda_i + \left(\bar{\theta} - \theta_i\right)$ and $V^M < \theta_i + \lambda_i$.

Next, consider viewers for whom $\theta_i < \frac{\bar{\theta} - \left(V^H - V^M\right)}{2}$ and $V^H - r < \bar{\theta} - \theta_i + \lambda_i$:

$$\begin{aligned}
\theta_i < \frac{\bar{\theta} - \left(V^H - V^M\right)}{2} \Longleftrightarrow U_i^{2,0} - U_i^{1,1} &= V^M - V^H - 2\theta_i + \bar{\theta} \\
&= \left(\bar{\theta} - \left(V^H - V^M\right)\right) - 2\theta_i \\
V^H - r < \bar{\theta} - \theta_i + \lambda_i \Longleftrightarrow U_i^{2,0} - U_i^{2,1} &= \left(\bar{\theta} - \theta_i + \lambda_i\right) - \left(V^H - r\right) > 0
\end{aligned}$$

and

$$\begin{aligned}
&V^H - r < \bar{\theta} - \theta_i + \lambda_i \Longrightarrow \\
&\qquad U_i^{2,0} - U_i^{2,2} = -V^H - V^M - 2\theta_i + 2\bar{\theta} + 2\lambda_i + r \\
&\qquad\qquad = \left[\bar{\theta} - \theta_i + \lambda_i - \left(V^H - r\right)\right] + \left[\bar{\theta} - \theta_i + \lambda_i - V^M\right] > 0
\end{aligned}$$

since $V^M < V^H - r$ by assumption. Therefore, viewers for whom $\theta_i < \frac{\bar{\theta} - \left(V^H - V^M\right)}{2}$ and $V^H - r < \bar{\theta} - \theta_i + \lambda_i$ find (2,0) optimal.

Now consider $\theta_i + \lambda_i < V^M - r$ and $\lambda_i - \theta_i > V^M - \bar{\theta}$:

$$\begin{aligned}
\theta_i + \lambda_i < V^M - r &\Longleftrightarrow U_i^{2,1} - U_i^{1,1} = V^M - r - \theta_i - \lambda_i > 0, \\
\lambda_i - \theta_i > V^M - \bar{\theta} &\Longleftrightarrow U_i^{2,1} - U_i^{2,2} = \lambda_i - \theta_i - V^M + \bar{\theta} > 0
\end{aligned}$$

This sets the (2,1) region as those for whom $V^H - r > \bar{\theta} - \theta_i + \lambda_i$, $\theta_i + \lambda_i < V^M - r$, and $\lambda_i - \theta_i > V^M - \bar{\theta}$.

Lastly, consider:

$$\lambda_i < V^M - \frac{\bar{\theta}}{2} - \frac{r}{2} \Longleftrightarrow U_i^{2,2} - U_i^{1,1} = 2V^M - \bar{\theta} - 2\lambda_i - r > 0$$

implying those for whom $\lambda_i - \theta_i < V^M - \bar{\theta}$ and $\lambda_i < V^M - \frac{\bar{\theta}}{2} - \frac{r}{2}$ find (2,2) optimal. As a result, the remainder of viewers for whom $\theta_i < \frac{\bar{\theta}}{2}$ find (1,1) optimal.

This covers all individuals for whom $\theta_i < \frac{\bar{\theta}}{2}$. With the decisions of all individuals closer to network 1 than network 2 mapped out, to determine the remaining individuals' decisions, one can simply reflect this graph over the midline $\lambda = \frac{\bar{\theta}}{2}$ to reproduce the graph in Figure D.2. ∎

D.3 DVR User Comparison Across Lineup Outcomes

Proposition 25 *The mass of DVR users is larger with head-to-head lineups than with staggered ones.*

Proof Figure D.3 displays the mass of DVR users in each lineup scenario. The dark-outlined area represents the mass of DVR users in the head-to-head scenario, the unshaded portion of that area reflects DVR users when lineups are staggered, and the shaded portion represents the difference between the two.

Define DVR_{HH} as the mass of DVR users in the head-to-head scenario, and DVR_{ST} the mass with staggered lineups. DVR_{HH} equals the area of the entire outlined figure, which is an isosceles right triangle with short diagonal of length $V^H - \frac{\bar{\theta}}{2} - r$. Therefore:

$$DVR_{HH} = \left(V^H - r - \frac{\bar{\theta}}{2}\right)^2$$

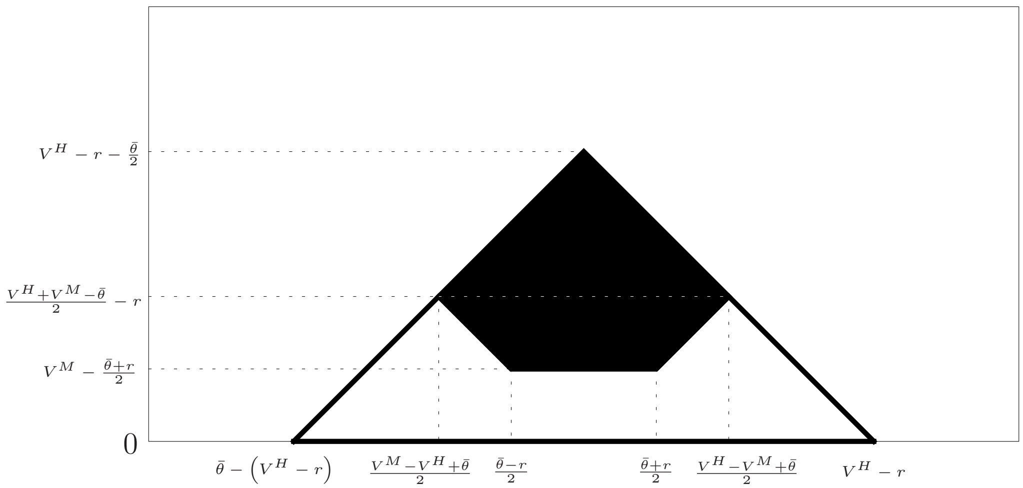


Figure D.3: Close-up View of DVR Users

To determine the mass of DVR_{ST}, I first find the area of the shaded region and then subtract that from DVR_{HH}. The white dotted line in Figure D.3 spans from $\frac{V^M-V^H+\bar{\theta}}{2}$ to $\frac{V^H-V^M+\bar{\theta}}{2}$ and consequently has length $V^H - V^M$. Therefore, the top of the shaded region is an isosceles right triangle with hypotenuse $V^H - V^M$, and hence it has area $\frac{\left(V^H-V^M\right)^2}{4}$. To determine the area of the trapezoid, note that it spans vertically from $\frac{V^H+V^M-\bar{\theta}}{2} - r$ to $V^M - \frac{\bar{\theta}+r}{2}$ and therefore has height $\frac{V^H-V^M-r}{2}$. Further, one base is the white dotted line and therefore has length $V^H - V^M$, while the other base spans from $\frac{\bar{\theta}-r}{2}$ to $\frac{\bar{\theta}+r}{2}$ and therefore has length r. Consequently:

$$
\begin{aligned}
&DVR_{ST}\\
&\quad= \left(V^H - r - \frac{\bar{\theta}}{2}\right)^2 - \left[\frac{\left(V^H - V^M\right)^2}{4} + \frac{1}{2}\left(V^H - V^M + r\right)\left(\frac{V^H - V^M - r}{2}\right)\right]\\
&\quad= \left(V^H - r - \frac{\bar{\theta}}{2}\right)^2 - \left[\frac{\left(V^H - V^M\right)^2}{4} + \frac{\left(\left(V^H - V^M\right) + r\right)\left(\left(V^H - V^M\right) - r\right)}{4}\right]\\
&\quad= \left(V^H - r - \frac{\bar{\theta}}{2}\right)^2 - \left[\frac{\left(V^H - V^M\right)^2}{4} + \frac{\left(\left(V^H - V^M\right)^2 - r^2\right)}{4}\right]\\
&\quad= \left(V^H - r - \frac{\bar{\theta}}{2}\right)^2 - \left[\frac{\left(V^H - V^M\right)^2}{2} - \frac{r^2}{4}\right]
\end{aligned}
$$

The first term of the last line is DVR_{HH}, so the last expression in square brackets represents the difference $DVR_{HH} - DVR_{ST}$. By assumption, $V^H - V^M > r$; this ensures the expression is positive. As a result, $DVR_{ST} < DVR_{HH}$.

This difference, represented by the shaded region in Figure D.3, equates to a diamond with a small area cut out of the bottom. The diamond represents those users that rent a DVR in the head-to-head scenario but not in the staggered scenario because it provides them additional access to a mediocre show only instead of to a hit show; this accounts for the $\frac{\left(V^H - V^M\right)^2}{2}$ portion of the difference. However, a slice is removed from the bottom of that diamond representing those users that rent a DVR to watch both mediocre shows even though one mediocre show alone does not provide enough value to do so, comprising an area of $\frac{r^2}{4}$ in the diagram. This difference must be positive, ensuring that $DVR_{ST} < DVR_{HH}$. ∎

D.4 DVR User Critical Mass

Proposition 27 *That each firm's profit remains the same regardless of lineup decisions only requires that some positive mass of viewers rent a DVR in equilibrium, and the mass of all viewers that rent a DVR in equilibrium can be an arbitrarily small proportion of the total mass of viewers.*

Proof The first portion of the proposition was proven in the proof of Proposition 26. For the second portion, note that the shaded region of Figure D.4 demonstrates the total viewer mass in either scenario. The total area of the two trapezoids comprising the viewer mass is:

$$\frac{\bar{\theta}}{2} \cdot \left(V^H + \left(V^H - \frac{\bar{\theta}}{2} \right) \right) = V^H \bar{\theta} - \left(\frac{\bar{\theta}}{2} \right)^2$$

As demonstrated in the proof to Proposition 25, the mass of DVR users in the head-to-head scenario is

$$DVR_{HH} = \left(V^H - r - \frac{\bar{\theta}}{2} \right)^2$$

The proportion of all viewers that use a DVR is therefore:

$$d \equiv \frac{\left(V^H - r - \frac{\bar{\theta}}{2} \right)^2}{V^H \bar{\theta} - \left(\frac{\bar{\theta}}{2} \right)^2}$$

Since the mass of DVR users in the head-to-head scenario is greater than the DVR mass in the staggered scenario, one need only prove the proposition for the head-to-head scenario to completely the proof.

For any given $\epsilon > 0$, one must find parameter values such that $d < \epsilon$ while the model's assumptions $0 < \frac{\bar{\theta}}{2} < V^M < V^H < \bar{\theta}$ and $r < \min \left[V^M - \frac{\bar{\theta}}{2}, V^H - V^M \right]$ still

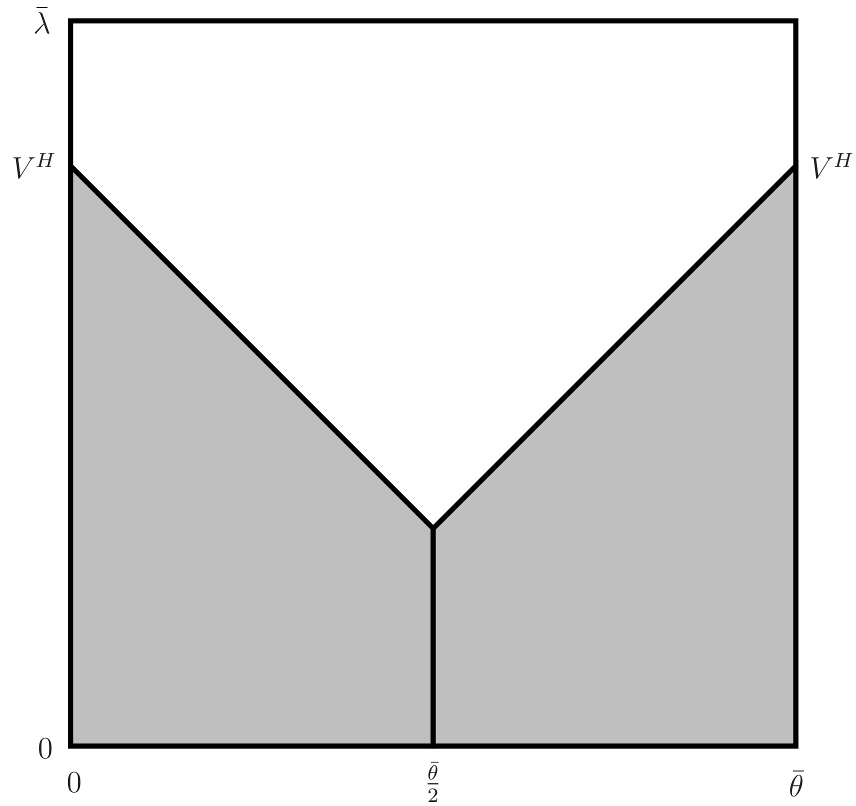


Figure D.4: Total Viewer Mass - Head-to-Head Lineups

hold. Consider the following parameter values:

$$
\begin{aligned}
r &= \eta \\
V^M &= \frac{\bar{\theta}}{2} + 2\eta \\
V^H &= \frac{\bar{\theta}}{2} + 5\eta \\
\bar{\theta} &> 10\eta
\end{aligned}
$$

The model's assumptions clearly hold. Further:

$$\begin{aligned} d &= \frac{\left(V^H - r - \frac{\bar{\theta}}{2}\right)^2}{V^H\bar{\theta} - \left(\frac{\bar{\theta}}{2}\right)^2} \\ &= \frac{\left(\left(\frac{\bar{\theta}}{2} + 5\eta\right) - \eta - \frac{\bar{\theta}}{2}\right)^2}{\left(\frac{\bar{\theta}}{2} + 5\eta\right)\bar{\theta} - \left(\frac{\bar{\theta}}{2}\right)^2} \\ &= \frac{16\eta^2}{\frac{\bar{\theta}^2}{4} + 5\eta\bar{\theta}} \end{aligned}$$

which can be made arbitrarily small by choosing large $\bar{\theta}$ and arbitrarily small η. ∎

D.5 DVR Comparative Statics

Proposition 28 *The mass of DVR users in both scenarios increases with V^H and decreases with r and $\bar{\theta}$. DVR users increase with V^M in the staggered scenario, whereas V^M has no effect on the DVR mass in the head-to-head scenario.*

Proof As demonstrated in the Proof to Proposition 25,

$$DVR_{HH} = \left(V^H - r - \frac{\bar{\theta}}{2}\right)^2$$

and

$$DVR_{ST} = \left(V^H - r - \frac{\bar{\theta}}{2}\right)^2 - \frac{\left(V^H - V^M\right)^2}{2} + \frac{r^2}{4}$$

I examine the comparative statics for each parameter in turn. For V^H:

$$\frac{\partial DVR_{HH}}{\partial V^H} = 2\left(V^H - r - \frac{\bar{\theta}}{2}\right) > 0$$

and

$$\frac{\partial DVR_{ST}}{\partial V^H} = 2\left(V^H - r - \frac{\bar{\theta}}{2}\right) - \left(V^H - V^M\right) = \left(V^H - r - \frac{\bar{\theta}}{2}\right) + \left(V^M - r - \frac{\bar{\theta}}{2}\right) > 0,$$

for V^M:

$$\frac{\partial DVR_{HH}}{\partial V^M} = 0$$

and

$$\frac{\partial DVR_{ST}}{\partial V^M} = \left(V^H - V^M\right) > 0,$$

and for $\bar{\theta}$:

$$\frac{\partial DVR_{HH}}{\partial \bar{\theta}} = \frac{\partial DVR_{ST}}{\partial \bar{\theta}} = -\left(V^H - r - \frac{\bar{\theta}}{2}\right) < 0$$

all due to the assumptions $V^M < V^H$ and $r < V^M - \frac{\bar{\theta}}{2}$.

For r:

$$\frac{\partial DVR_{HH}}{\partial r} = -2\left(V^H - r - \frac{\bar{\theta}}{2}\right) < 0$$

and

$$\begin{aligned} \frac{\partial DVR_{ST}}{\partial r} &= -2\left(V^H - r - \frac{\bar{\theta}}{2}\right) + \frac{r}{2} \\ &= -2\left(V^M - r - \frac{\bar{\theta}}{2}\right) - 2\left(V^H - V^M - \frac{r}{4}\right) \\ &< 0 \end{aligned}$$

due to the assumptions $r < V^H - V^M, V^M - \frac{\bar{\theta}}{2}$. ∎

D.6 Ad-Filtering Profit Conditions

Proposition 29 *If the DVR eliminates a proportion p of the ad impressions a DVR user receives, then for any values of the parameters, there exists $\hat{p} \in \left(\frac{1}{2}, \frac{1}{3}\right)$ such that $\forall p < \hat{p}$, each network's profit is higher with DVR technology, and $\forall p > \hat{p}$, each network's profit is higher without DVR technology for the entire parameter space.*

Proof The outcomes both with and without DVR technology are symmetric across networks, so WLOG I focus solely on network 1's profits. As shown in Appendix Section C.6, each network's profit in the sole equilibrium when there is no DVR technology is:[1]

$$A^{St}_{NoDVR} = \frac{V^H \bar{\theta}}{2} + \frac{V^M \bar{\theta}}{2} - \frac{\bar{\theta}^2}{4} + \frac{\left(V^H - V^M\right)^2}{4}$$

To calculate network 1's profit when DVR technology is available, I first derive the total audience size of each show and then determine the mass of viewers that use a DVR for each show. Figure D.5 displays the market shares for each of network 1's shows. Its hit show's audience forms an isoceles right triangle with side length $V^H - r$ plus the parallelogram labeled A, so that the total audience size for the hit show is:

$$\frac{1}{2}\left[\left(V^H - r\right)^2 + r\left(\bar{\theta} + V^H - V^M\right)\right]$$

Its mediocre show's audience forms an isoceles right triangle of side length V^M with the trapezoid labeled B removed. From the proof of Proposition 26 in section 4.3, the area of B is $\frac{1}{2} \cdot r \cdot \left(V^H - V^M\right)$, so that the mass of viewers that watch network 1's mediocre show is:

$$\frac{1}{2}\left[V^{M2} - r \cdot \left(V^H - V^M\right)\right]$$

[1]Note that δ was absorbed into V^M after it was shown that the two parameters are inseparable with foresighted viewers.

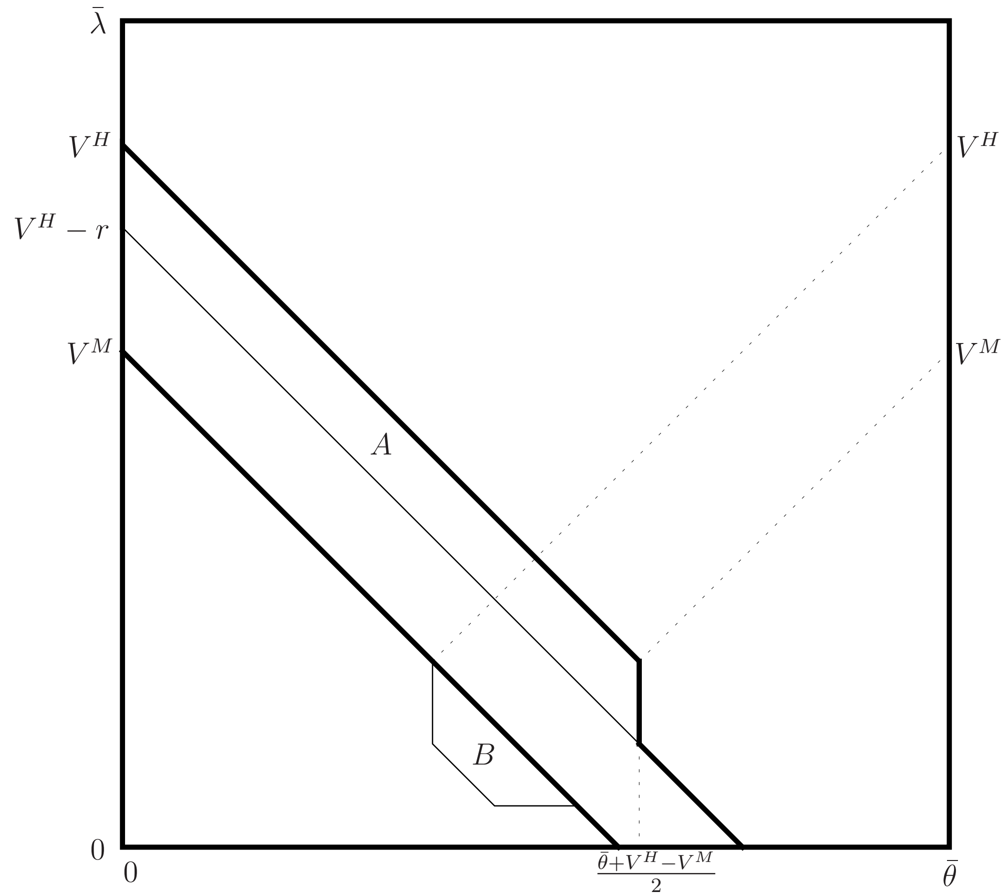


Figure D.5: Total Viewer Mass - Head-to-Head Lineups

The total of these two audience masses is therefore:

$$A_{DVR,1}^{St} = \frac{1}{2}\left[\left(V^H - r\right)^2 + r\bar{\theta} + {V^M}^2\right]$$

Calculating the audience mass that uses a DVR is more complicated. Figure D.6 shows a close-up of the DVR mass area of the graph to help illuminate the derivation. Viewers that use a DVR and watch station 1's mediocre show form the triangle C

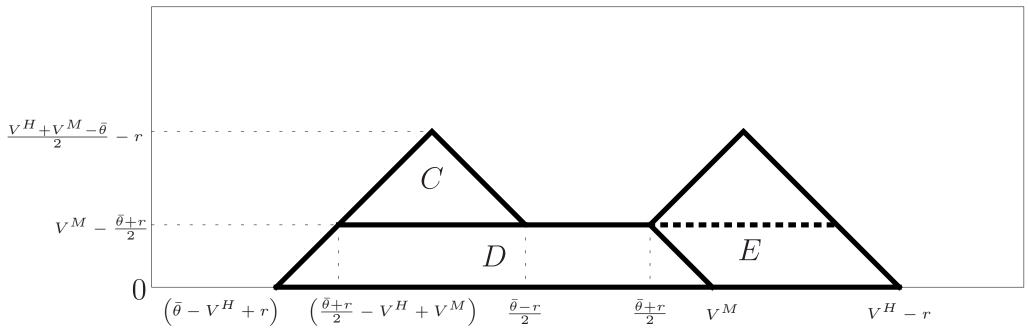


Figure D.6: Total Viewer Mass - Head-to-Head Lineups

plus the trapezoid D.

$$
\begin{aligned}
C &= \frac{1}{2}\left(V^H-V^M-r\right)\left(\left(\frac{V^H-V^M-\bar{\theta}}{2}-r\right)-\left(V^H-\frac{\bar{\theta}+r}{2}\right)\right) \\
&= \frac{1}{2}\left(V^H-V^M-r\right)\left(\frac{V^H-V^M-r}{2}\right) \\
&= \left(\frac{V^H-V^M-r}{2}\right)^2, \\
D &= \frac{1}{2}\left(V^M-\frac{\bar{\theta}+r}{2}\right)\left(\left(V^H+V^M-\bar{\theta}-r\right)+\left(V^H-V^M\right)\right) \\
&= \frac{1}{2}\left(2V^H-\bar{\theta}-r\right)\left(V^M-\frac{\bar{\theta}+r}{2}\right)
\end{aligned}
$$

Viewers that use a DVR and watch station 1's hit show comprise the same area plus the additional figure outlined in black, which is composed of a triangle equivalent to C plus the parallelogram E:

$$E=\left(V^H-V^M-r\right)\left(V^M-\frac{\bar{\theta}+r}{2}\right)$$

Therefore, the audience mass that uses a DVR and watches station 1's mediocre show is $C+D$, whereas that that uses a DVR and watches station 1's hit show is

$2C + D + E$. Consequently, the total of these two audience masses is:

$$\begin{aligned}
DVR_1^{St} &\equiv 3C + 2D + E \\
&= 3\left(\frac{V^H - V^M - r}{2}\right)^2 + \left(2V^H - \bar{\theta} - r\right)\left(V^M - \frac{\bar{\theta} + r}{2}\right) \\
&\quad + \left(V^H - V^M - r\right)\left(V^M - \frac{\bar{\theta} + r}{2}\right) \\
&= 3\left(\frac{V^H - V^M - r}{2}\right)^2 + \left(3V^H - V^M - \bar{\theta} - 2r\right)\left(V^M - \frac{\bar{\theta} + r}{2}\right)
\end{aligned}$$

The increase in ad impressions due to DVR technology is therefore:

$$
\begin{aligned}
\Delta\pi_1^{DVR} &\equiv A^{St}_{NoDVR,1} - \left(A^{St}_{DVR,1} - p \cdot DVR^{St}_1\right) \\
&= \left(\frac{V^H\bar{\theta}}{2} + \frac{V^M\bar{\theta}}{2} - \frac{\bar{\theta}^2}{4} + \frac{\left(V^H - V^M\right)^2}{4}\right) - \frac{1}{2}\left[\left(V^H - r\right)^2 + r\bar{\theta} + V^{M2}\right] \\
&\quad + \frac{3p}{2}\left(V^H - V^M - r\right)^2 + \frac{p}{2}\left(2V^M - \bar{\theta} - r\right)\left(3V^H - V^M - \bar{\theta} - 2r\right) \\
&= \frac{V^H\bar{\theta}}{2} + \frac{V^M\bar{\theta}}{2} - \frac{\bar{\theta}^2}{4} + \frac{V^{H2}}{4} - \frac{V^H V^M}{2} + \frac{V^{H2}}{4} - \frac{V^{H2}}{4} + V^H r - \frac{r^2}{4} - \frac{r\bar{\theta}}{4} \\
&\quad - \frac{V^{M2}}{4} + \frac{3p}{4}\left(V^H - V^M\right)^2 - \frac{3p}{2}\left(V^H - V^M\right)r + \frac{3p}{4}r^2 \\
&\quad + \frac{p}{2}\left(6V^H V^M - 2V^{M2} - 2V^M\bar{\theta} - 4V^M r\right) \\
&\quad + \frac{p}{2}\left(-3V^H\bar{\theta} + V^M\bar{\theta} + \bar{\theta}^2 + 2r\bar{\theta} - 3rV^H + rV^M + r\bar{\theta} - 2r^2\right) \\
&= -\frac{V^{H2}}{4} - \frac{V^H V^M}{2} + \frac{V^H\bar{\theta}}{2} + V^H r - \frac{V^{M2}}{4} + \frac{V^M\bar{\theta}}{2} - \frac{\bar{\theta}^2}{4} - \frac{r}{2} - \frac{r\bar{\theta}}{2} \\
&\quad + \frac{3p}{4}V^{H2} - \frac{3p}{2}V^H V^M + \frac{3p}{4}V^{M2} - \frac{3p}{2}V^H r + \frac{3p}{2}V^M r + \frac{3p}{4}r^2 + 3pV^H V^M \\
&\quad - pV^{M2} - pV^M\bar{\theta} - 2pV^M r - \frac{3p}{2}V^H\bar{\theta} + \frac{p}{2}V^M\bar{\theta} + \frac{p}{2}\bar{\theta}^2 + pr\bar{\theta} - \frac{3p}{2}rV^H \\
&\quad + \frac{p}{2}rV^M + \frac{p}{2}r\bar{\theta} + pr^2 \\
&= \frac{3p-1}{4}V^{H2} + \frac{3p-1}{2}V^H V^M - \frac{3p-1}{2}V^H\bar{\theta} - (3p-1)\,V^H r - \frac{p+1}{4}V^{M2} \\
&\quad + \frac{1-p}{2}V^M\bar{\theta} - \frac{1-2p}{4}\bar{\theta}^2 + \frac{7p-2}{4}r^2 + \frac{3p-1}{2}r\bar{\theta} \\
&= \frac{3p-1}{4}\left(V^{H2} + 2V^H V^M - 2V^H\bar{\theta} - 4V^H r + 2r\bar{\theta}\right) - \frac{p+1}{4}V^{M2} + \frac{1-p}{2}V^M\bar{\theta} \\
&\quad - \frac{1-2p}{4}\bar{\theta}^2 + \frac{7p-2}{4}r^2
\end{aligned}
$$

Now that a generalized expression for the increase in ad impressions due to DVR technology has been derived, explicit conditions under which network profits increase and decrease can be determined.

Clearly, the increase in ad impressions is increasing in p, so showing that it is unambiguously negative at $p = \frac{1}{3}$ and unambiguously positive at $p = \frac{1}{2}$ completes the proof. Evaluted at $p = \frac{1}{3}$:

$$\begin{aligned} \Delta\pi_1^{DVR} &= \frac{1}{8}\left(V^{H^2} + 2V^H V^M - 2V^H\bar{\theta} - 4V^H r + 2r\bar{\theta} - 3V^{M^2} + 2V^M\bar{\theta} + 3r^2\right) \\ &= \frac{1}{8}\left(V^H - V^M - r\right)\left(V^M + 3V^M - 2\bar{\theta} - 3r\right) \end{aligned}$$

which must be positive since

$$V^H - V^M > r \Longrightarrow V^H - V^M - r > 0$$

and

$$\begin{gathered} \left(V^H > \frac{\bar{\theta}}{2}\right) \wedge \left(V^M - \frac{\bar{\theta}}{2} > r\right) \Longrightarrow \\ V^H - 3V^M > 2\bar{\theta} + 3r \end{gathered}$$

At $p = \frac{1}{2}$:

$$\begin{aligned} \Delta\pi_1^{DVR} &= -\frac{1}{3}V^{M^2} + \frac{1}{3}V^M\bar{\theta} - \frac{1}{12}\bar{\theta}^2 + \frac{1}{12}r^2 \\ &= \frac{1}{12}\left(r^2 - 4\left(V^M - \frac{\bar{\theta}}{2}\right)^2\right) \end{aligned}$$

which, due to the assumption $V^M - \frac{\bar{\theta}}{2} > r > 0$, is stricly negative. ∎

Bibliography

Ambrus, A. and M. Reisinger (2006). Exclusive vs Overlapping Viewers in Media Markets. Discussion Paper.

Anderson, S. and S. Coate (2005). Market Provision of Broadcasting: A Welfare Analysis. *Review of Economic Studies 72*(4), 947–972.

Anderson, S. and J. Gans (2011). Platform Siphoning: Ad-Avoidance and Media Content. *American Economic Journal: Microeconomics*. Forthcoming.

Anderson, S. P. and J. J. Gabszewicz (2006, December). *The Media and Advertising: A Tale of Two-Sided Markets*, Volume 1 of *Handbook of the Economics of Art and Culture*, Chapter 18, pp. 567–614. Elsevier.

Boles, C. (2007). Rent or Own? The New Cable-TV Dilemma. *Wall Street Journal*, June 7, 2007.

Bowman, G. (1975). Consumer choice and television. *Applied Economics 7*(3), p175–184.

Bowman, G. W. and J. U. Farley (1972). TV viewing: application of formal choice model. *Applied Economics 4*(4), p245–259.

Carter, B. (1996). The Art and Craftiness of TV Scheduling. *The New York Times*, September 12, 1996.

Cooper, R. (1993). An Expanded, Integrated Model for Determining Audience Exposure to Television. *Journal of Broadcasting and Electronic Media 37*(4), 401–418.

Downey, K. (2007). More Folks Own DVRs Than You Think. *Media Life Research*, April 27, 2007.

Edwards, C. (2007). TiVo Targets the Mainstream. *Bloomberg Businessweek*, July 23, 2007.

Eggerton, J. (2008). Marketers: TV Advertiser Effectiveness Has Decreased. *Broadcasting & Cable*, February 20, 2008.

Gabszewicz, J., D. Laussel, and N. Sonnac (2000). TV-Broadcasting Competition and Advertising.

Gensch, D. and P. Shaman (1980). Models of Competitive Television Ratings. *Journal of Marketing Research 17*(3), 307–315.

Goettler, R. L. and R. Shachar (2001). Spatial Competition in the Network Television Industry. *The RAND Journal of Economics 32*(4), 624–656.

Goetzl, D. (2006). New Data Reveals Virtually No Viewers For Time-Shifted Spots. *MediaDailyNews*, April 6, 2006.

Harmon, A. (2002). Skip-the-Ads TV Has Madison Ave. Upset. *The New York Times*, May 23, 2002.

Haugsted, L. (2007). Cable Overtakes DBS in DVR Penetration. *Multichannel News*, July 17, 2007.

Horen, J. H. (1980). Scheduling of Network Television Programs. *Management Science 26*(4), 354–370.

Hotelling, H. (1929). Stability in Competition. *The Economic Journal 39*(153), 41–57.

Keane, M. (2008). Television Advertising To Shrink 4.6 Percent in 2009. *Wired*, November 18, 2008.

Liu, Y., D. S. Putler, and C. B. Weinberg (2004). Is Having More Channels Really Better? A Model of Competition among Commercial Television Broadcasters. *Marketing Science 23*(1), 120–133.

Mussa, M. and S. Rosen (1978). Monopoly and Product Quality. *Journal of Economic Theory 18*(2), 301–317.

Peitz, M. and T. Valletti (2004). Content and advertising in the media: Pay-tv versus free-to-air. CEPR Discussion Papers 4771, C.E.P.R. Discussion Papers.

Rust, R. T. and M. I. Alpert (1984). An Audience Flow Model of Television Viewing Choice. *Marketing Science 3*(2), 113–124.

Rust, R. T. and N. V. Eechambadi (1989). Scheduling Network Television Programs: A Heuristic Audience Flow Approach to Maximizing Audience Share. *Journal of Advertising 18*(2), 11–18.

Shachar, R. and J. W. Emerson (2000). Cast Demographics, Unobserved Segments, and Heterogeneous Switching Costs in a Television Viewing Choice Model. *Journal of Marketing Research 37*(2), 173–186.

Snedeker, L. (2007). Season's Big Unknown: The DVR Effect. *Media Life Research*, October 12, 2007.

Story, L. (2007). Viewers Fast-Forwarding Past Ads? Not Always. *The New York Times*, February 16, 2007.

Tag, J. (2009). Paying to Remove Advertisements. *Information Economics and Policy 21*(4), 245–252.

Vranica, S. (2008). Latest Pitch Is a Show Within an Ad. *The Wall Street Journal*, May 30, 2008.

Wilbur, K. C. (2007). A Two-Sided, Empirical Model of Television Advertising and Viewing Markets.

Wilbur, K. C. (2008). How the Digital Video Recorder Changes Traditional Television Advertising.