Essays on the Economics of Informative Advertising with Applications to the Television Industry

Levent Çelik
İzmir, Turkey

B.A., Boğaziçi University, 2000
M.A., University of Virginia, 2002

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
January, 2007

Abstract

This dissertation presents three essays on informative advertising in markets with differentiated products and consumer search. The first essay analyzes a single television station's choice of airing tune-ins (preview advertisements). I consider two consecutive programs located along a unit line. Potential viewers know the earlier program but are uncertain about the later one. They may learn its location through a tune-in if they watch the earlier program and the television station chose to air a tune-in, or by sampling it for a few minutes. If the sampling cost is sufficiently low, the unique perfect Bayesian equilibrium (PBE) exhibits no tune-ins. If it is sufficiently high, the unique PBE involves a tune-in whenever the two programs are similar enough. For all other values of the sampling cost, either PBE may arise. When the programs are also quality-differentiated, the willingness to air a tune-in, and thus to disclose location information, may be sufficient to signal high quality without any dissipative advertising.

The second essay extends the first one by including a second TV station. Now, each station's tune-in decision may also depend on the rival station's program, thereby revealing more information than the actual content of the tune-in. This happens only if the sampling cost is low enough. Otherwise, each station makes its tune-in decision independently of its rival's program. Thus, there may exist signaling via informative advertising. It is welfare improving to ban tune-ins in the latter case while this is not necessarily true in the former one.

The third essay analyzes informative advertising in a duopoly market with differentiated products when consumer search is costless. If consumers are fully rational, exposure to a single advertisement is sufficient for them to obtain complete market information. In this case, firms undersupply advertising compared to the social optimum because of free-riding. If consumers are not fully rational, they may ignore the existence of another firm when the only advertisement they receive quotes the monopoly price. In this case, both firms advertise the monopoly price, and the market may produce too much or too little advertising compared to the social optimum.

Acknowledgements

First and foremost, I would like to thank my primary advisor, Simon Anderson, for his continuous help, patience and support. Many of his comments and suggestions have been central in the writing of this dissertation. Most importantly, he taught me how to approach economic problems properly. I was very fortunate to have him as my primary advisor. I also owe special thanks to my second advisor, Maxim Engers, for his help over the years. His insightful suggestions have greatly improved the quality of this dissertation.

I would like to thank Federico Ciliberto and Albert Choi for agreeing to be in my dissertation committee and providing feedback during my defense. Thanks also go to Emily Blanchard and John McLaren for their input during my mock interview.

The main idea of the second chapter was partially developed in a term paper that I submitted to Andrew Cohen for his I.O. course. Without his stimulating comments, I may not have pursued this idea any further. I am very grateful for his encouragement in the subject.

The second chapter has benefited from the discussions I had with Bruno Jullien, Kieron Meagher, André de Palma and Régis Renault during their visits at the University of Virginia. Comments of several people at the Third Conference on Media Economics have also been valuable.

The years in graduate school have been the most demanding time of my life. I was lucky to have several good friends in such a tough period. Some of them also helped by reading earlier drafts of the second chapter and making useful suggestions. I would like to thank them, in particular Devrim Demirel, Maria Fitzpatrick, Başak Gönen, Jessica Howell, Adam Hyde, Bilgehan Karabay, Özer Kunaçav, Daniel Mackay, Esen Onur, Üstün Oral, Mehmet Orhan, Alper Özgit, Mary Simonson, Loren Smith and Özgür Tataroğlu.

Finally, I acknowledge the financial support of the Bankard Fund for Political Economy at the University of Virginia.

All errors are mine.

Table of Contents

Chapters

Dedicated to my parents, Fatma Göksel Özaytek and Ali Hasan Çelik, for their investment in my human capital, to my brother Ersel Çelik for being my best friend throughout my life, and to my aunt Leyla Özaytek for her endless support. Sizlerin varlığı ve desteği olmadan bulunduğum noktaya gelemezdim….

Chapter 1

Introduction

Product differentiation and advertising are two key elements of contemporary markets. According to Tirole (1988), firms differentiate their products in order to avoid price competition.[1] If consumers have sufficiently diversified, product differentiation tends to raise profits. However, as the range of differentiation gets larger, it becomes increasingly more difficult for consumers to acquire relevant product information. Thus, firms may engage in informative advertising in order to increase product awareness, and to inform consumers about relevant product characteristics. In this dissertation, I analyze directly informative advertising in markets with differentiated products and with consumers who actively search for product information.

I focus on differentiation whereby certain products are better suited for some customers than others. This type of differentiation is known as horizontal differentiation. Traditionally, spatial models have been used to analyze markets with this form of differentiation, and I continue the tradition. Accordingly, consumer tastes and product characteristics are represented by particular locations along a line a lá Hotelling (1929) or around a circle a lá Vickrey (1964), and consumers strictly prefer products that are closer to their ideal tastes.

Most studies of informative advertising in differentiated-products markets postulate

[1] D'Aspremont, Gabszewicz and Thisse (1979, 1983) are the first to demonstrate "the principle of (maximal) differentiation."

consumers as initially being unaware of the market's existence.[2] Thus, advertisements (henceforth, ads) inform them about product existence along with several other product characteristics. Since consumers are ex-ante unaware of the market structure, however, they do not make any inferences for the products about which they have not been informed through ads. Introducing consumer search into this context is problematic since one has to define a broad set of prior beliefs that consumers possess regarding the market structure. Therefore, most authors assume search costs are prohibitively high so that consumers never engage in search.

Many differentiated-products markets do not fit these specifications. In several consumer markets, product existence is common knowledge and/or consumers actively search for product information. If, for instance, consumers are aware of a product's existence but are not very well informed about its characteristics, then a firm's unwillingness to advertise may also be informative. If they also know the available number of products in a market, then a firm's willingness to advertise may reveal further information regarding the other products.[3] In markets that consumers are initially unaware of, on the other hand, information comes only from the content of the ads received if any. However, the assumption of high search costs in these markets restricts consumers from acquiring further information, and thus causes overprovision of advertising. In this dissertation, I present three essays that address these shortcomings of the existing literature's treatment of informative adverting, and that explicitly deal with them in simple, yet appealing, ways.

[2] For examples, see Grossman and Shapiro (1984) and Christou and Vettas (2003).

[3] This requires that firms have common knowledge about all products, and consumers know that firms have this information.

In Chapter 2, I analyze the provision of tune-ins in the market for television (henceforth, TV) broadcasting.[4] A distinctive feature of the TV market is that the existence of TV programs is common knowledge to everyone beforehand. Therefore, the choice of a broadcaster not to air any tune-ins may in fact carry a valuable amount of information. In order to capture this insight, I develop a simple spatial model with asymmetric information in which a single TV station broadcasts two consecutive, possibly different, programs. The TV station has complete information while potential viewers know the earlier program but are uncertain about the characteristics of the later one. So, depending on how similar they are, the TV station may or may not choose to promote the later program by airing a tune-in during the earlier one. I then allow for the possibility of sampling a program for a few minutes and show that it has important implications on the equilibrium provision of tune-ins. Finally, I introduce viewer uncertainty regarding the actual quality of the later program and compare my findings to those of the literature on quality signaling.

In Chapter 3, I extend the analysis of Chapter 2 by including a second TV station. A key element of the model is that TV stations have complete information about program characteristics. This is important because each TV station's tune-in strategy now depends on the characteristics of both stations' upcoming programs. Therefore, a TV station's tune-in decision does not only inform viewers about its own program but may also convey information about the other station's upcoming program. I demonstrate that this is actually possible. Viewers can deduce valuable information from all possible tune-in strategies of a TV station. This finding highlights a traditionally ignored aspect of signaling: it may also

[4]Tune-ins are preview ads for broadcasters' upcoming programs.

arise in horizontally differentiated markets.

In Chapter 4, I analyze the implications of active consumer search on informative advertising in a market that consumers are ex-ante unaware of. This, however, is not a trivial task. As described before, one needs to introduce a broad set of prior beliefs that consumers base their optimal search behavior on when they are informed about the existence of such a product market through advertising. In order to step away from this complication, I assume that search is costless. I then introduce "ignorant" consumers who do not search when the only ad they receive quotes the monopoly price. This behavioral assumption produces strikingly different results compared to the existing literature.

The market for TV advertising is huge: it attracted approximately $60 billion of a total of $246 billion spent on advertising in the United States in 2003.[5] In this regard, it constitutes the largest media channel for reaching consumers. Therefore, it is of central interest to economists.

Early literature on TV advertising mainly focused on the optimal choice of TV programming and the resulting viewer benefits. Steiner (1959) argues that TV stations air very similar programs because a larger proportion of viewers prefer them. So, popular programs tend to be aired too often while specialty programs do not. Others have extended Steiner's work. However, early literature's treatment of TV advertising is unsatisfactory because of its failure to take into account the two-sided nature of the TV industry. On the one side, there are the companies that are willing to purchase commercial spots in order

[5] Other important categories were: direct mail: $48.5 million, newspaper: $45 million, radio: $19 million, yellow pages: $14 million, consumer magazine: $11.5 million, internet: $5.5 million.

reach potential consumers. On the other side of the market, there are viewers who generally dislike watching commercials. The main role of a TV station is to find the right balance of delivering viewers to advertising companies.[6]

Anderson and Coate (2005) is the first formal study to acknowledge the two-sided role of the TV industry. They mainly analyze the welfare properties of TV advertising. They find that the market may fail to provide optimal advertising levels. Overprovision of advertising may occur when viewers' nuisance from an additional commercial exceeds advertisers' expected benefits from reaching one more viewer. Alternatively, underprovision may occur when the market is served by a single TV station and viewer nuisance cost is close to zero, or when programs are close substitutes.[7]

I focus on the provision of tune-ins in the next two chapters. Tune-ins constitute an important component of TV advertising. There are several TV programs which generate high audiences, and in turn, high advertising revenues. Nevertheless, TV stations continue to allocate a significant amount of time to tune-ins.[8] Although the main role of tune-ins is to convey program information to imperfectly informed viewers, they are different from most common forms of informative advertising. First, they do not impose direct costs to TV stations. Rather, TV stations incur opportunity costs for not having allocated the same

[6]Two-sided markets are platforms that bring together many agents with different interests. Agents' benefits usually depend on the number of other agents using the same platform. Examples include advertising supported media (such as newspapers, radio and websites), market for computer software (i.e. adware) and market for credit cards. The first formal studies of two-sided markets are Armstrong (2004), Caillaud and Jullien (2001, 2003), and Rochet and Tirole (2003, 2005).

[7]For a more comprehensive review of the literature on advertising in two-sided markets, see Anderson and Gabszewicz (2006).

[8]For instance, the number of minutes allocated to tune-ins during the Super Bowl in 2000 was 16.5 out of a total of 87 non-program minutes while the price of a 30-second commercial was $2 million.

time spot to a commercial. Therefore, the net costs of tune-ins depend on the audience sizes of the program during which they are aired and of the one which they promote. Second, TV stations are constrained to advertise their programs only to their viewers. This has an important implication; unlike in most other markets, viewers can indefinitely raise their likelihood of receiving tune-ins by simply watching a program. This is indeed one of the important findings of Chapter 2. I show that some viewers watch a program just to extract information about the next program although watching yields a negative utility. They do so because it may be more costly to learn the characteristics of the next program by simply sampling its few minutes.

These described features of tune-ins make them important for an economic analysis. However, the existing literature's treatment is unsatisfactory due to the lack of a theoretical study.[9] By presenting a careful analysis of tune-ins in the next two chapters, this dissertation also contributes to our understanding of information provision in the TV market.

In the remainder of this introduction, I provide a short overview of the literature on advertising. This also constitutes a general reference for the remaining chapters. Each chapter further discusses the relevant work in more detail. Interested readers should consult Bagwell (2005) for an excellent survey of advertising.

Butters (1977) is the first author to offer a formal model of informative advertising. He considers a homogenous goods market in which firms advertise the price of their products to a population of identical consumers. Because consumers have no knowledge of product existence prior to receiving an ad, advertising informs them of this as well. Each firm

[9] There are several empirical studies of tune-in some of which are discussed in Chapter 2.

sends (multiple) ads, which consumers receive randomly. A consumer who receives ads from different firms buys from the firm quoting the lowest price. Butters finds that the equilibrium exhibits price dispersion and that the ad level produced by the market is socially optimal.

Several authors have extended Butters' model in different directions. Within homogeneous goods markets, Stegeman (1991) finds that the market undersupplies informative advertising when consumers have heterogeneous valuations. Similarly, Stahl (1994) obtains underprovision when consumers have downward-sloping demands. Roberts and Stahl (1993) extend Butters' model by introducing optimal consumer search. They characterize a unique symmetric equilibrium in which firms either charge a high price and do not advertise it – thus selling only to uninformed consumers – or randomly choose a price from an interval of lower prices and advertise it. In the latter case, firms who choose a lower price advertise with a higher intensity.

Grossman and Shapiro (1984) is the first to consider informative advertising in a market with differentiated products. Following Vickrey (1964), they use a circle model to represent product differentiation and consumer preferences. Advertising informs consumers not only about the existence and the price, but also the characteristics of a product. They find that firms oversupply advertising relative to the social optimum if consumers care about variety and the number of firms is sufficiently large.

Grossman and Shapiro only consider ads that convey full product information. But in general, a firm chooses how broadly to inform consumers, if at all. Anderson and Renault (2006) study the conditions under which a monopolist chooses to advertise price information and/or product match information. They find that a monopolist may publicize only price, only match, or both price and match information depending on the value of the search cost consumers face. They also find that a monopolist prefers to convey only limited product information rather than full information.

Advertising may also indirectly convey product information. Nelson (1974) argues that this role of advertising is especially important in markets for experience goods.[10] Since consumers cannot distinguish the actual quality of these goods before they purchase them, firms selling high-quality products may invest in advertising in order to convey the signal of high quality and as a result to attract repeat purchases. Nelson's original idea was later formalized by Kihlstrom and Riordan (1984) and by Milgrom and Roberts (1986). Kihlstrom and Riordan develop a model of dissipative advertising in which firms are price-takers. Each firm sells a high-quality or a low-quality product. When advertising does not cause repeat purchases, they find that it can signal high quality only if the marginal cost of production is nonincreasing in quality. When it causes repeat purchases, on the other hand, this assumption is no longer necessary; advertising can signal high quality even when marginal cost of production is increasing in quality.

Milgrom and Roberts consider a monopolist selling a single product whose quality is either high or low. In their model, quality may be signaled through price, advertising or both. Assuming that there are repeat purchases, they find that dissipative advertising can

[10] A distinction between search goods and experience goods was first made by Nelson (1970). Accordingly, a good is a search good if its quality is detectable before the purchase (possibly through a costly search), and is an experience good if its quality is not detectable before the purchase.

only act as a signal of high quality when it is used together with a high price. When there are no repeat purchases, however, only price is sufficient to signal high quality. In important extensions of Milgrom and Roberts' influential paper, Hertzendorf and Overgaard (2001) and Fluet and Garella (2002) show that similar results carry over to a duopoly setting. Both papers find that a positive level of advertising by a high-quality firm is necessary for separation when the difference between the possible quality levels is sufficiently small. Their result, however, does not require repeat purchases unlike the situation in Milgrom and Roberts (1986). So, price competition may be an explanation for dissipative advertising in static models.

There are many other views on advertising. According to Kaldor (1949), advertising misleads consumers and convinces them that they have different tastes. Becker and Murphy (1993) approach advertising as a consumption good that has a positive effect on the valuation of the product advertised. In a seminal paper, Dixit and Norman (1978) examine the welfare effects of advertising taking both pre- and post-advertising tastes as the true underlying preferences. They find that, regardless of the actual underlying preferences, there is more advertising in equilibrium than what is socially optimal. Finally, advertising may deter entry in another view which was initiated by Bain (1956). The underlying reason is that industries get more concentrated as the level of total advertising increases, and as a result, profitable entry becomes less likely.

Chapter 2

Monopoly Provision of Tune-ins in Broadcasting

2.1 Introduction

As described in the previous chapter, most models of informative advertising assume consumers are ex-ante uninformed about product existence. This assumption has at least two critical implications; first, consumers do not make any inferences about product attributes when they remain unreached by advertising, and second, they do not engage in any activities that may increase their chances of receiving ads. While it may be a reasonable assumption for certain product markets, it obviously does not fit all. A good example is the market for broadcasting.[11] Although many people may be uninformed about the attributes of TV programs, they are clearly aware of their existence. Therefore, a TV station's decision to advertise its upcoming programs must account for the possible inferences its current viewers may draw from the absence of a tune-in. In other words, the fact that a TV station does not advertise an upcoming program may constitute a valuable piece of information for viewers. In this chapter, I formulate and solve an equilibrium model of tune-in provision that accounts for such information transmission.

TV stations forgo about 20% of their advertising revenues to air tune-ins for their upcoming programs (Shachar and Anand (1998)).[12] This fact, on its own, suggests the im-

[11] Other examples include markets for most other types of media, mass transportation and real estate.

[12] Shachar and Anand (1998) report that in 1995, three major network stations in the U.S. devoted 2 of 12

portance of the incomplete information structure in the TV market, yet most of the related literature assumes that viewers possess full information about program characteristics. Although a person can acquire information about the attributes of a program through TV schedules that appear in magazines or through word-of-mouth, an important fraction of viewers remain uninformed due to the costs associated with gaining information. Furthermore, individuals have limited memories. Therefore, TV stations use tune-ins to communicate with their viewers. Had viewers already been fully informed about the upcoming programs, there would be no need for tune-ins.

Tune-ins often provide direct information about program characteristics.[13] The level of information they provide is quite high. Based on a detailed panel dataset on viewer choices, Emerson and Shachar (2000) report that about 65% of viewers continue to watch the same network broadcaster (including the times when a tune-in has not been aired). This observation demonstrates that tune-ins achieve their main goals: raising the audience sizes of the promoted programs. Tune-ins also depend on the revenues of the programs during which they are aired. Their opportunity costs are higher during programs whose audience sizes are higher. So, the equilibrium provision of tune-ins depends on a careful cost-benefit analysis.

Network TV stations and most cable TV stations do not charge nominal fees for their programs. They rather price their programs indirectly by bundling them with commerical

---

minutes of nonprogramming time to tune-ins. Since advertising revenues represent almost all of the revenues of a network, the share of revenues spent on tune-ins is proxied as 20%.

[13] It is necessary to distinguish between tune-ins for regular programs, such as everyweek sitcoms, and those for special programs, such as movies. The latter are expected to be more effective on ratings in the sense that people may possess little or no information about the timing and attributes of such programs.

ads. This enables a focus only on the interaction between program differentiation and provision of informative advertising.

Certain programs are advertised several times during an ongoing program. This, however, is not completely due to high revenues that TV stations expect to generate from the advertised programs. Although about 80% of the network commercial time is sold in the up-front market during May for the upcoming season and the price paid by advertisers depends on the expected audience size, TV stations are bound to make up for the difference between the expected and the actual audience sizes should the former exceed the latter. Therefore, TV stations' intention for airing several tune-ins for the same program may be to signal that program's high quality. I analyze this possibility as an extension of the main model which I describe next.

I first lay out a benchmark model that has a single TV station airing two consecutive programs. Potential viewers differ in their preferred program characteristics. Programs and viewer preferences are represented by locations along a unit interval a lá Hotelling. I assume that viewers know the location of the program to be aired in the first period, but are uncertain about the location of the program to be aired in the second period. The TV station has the option of placing a tune-in for the second program during the first one, and chooses to do so when the marginal advertising revenue resulting from the increase in the second-period audience size exceeds the opportunity cost of airing a tune-in. People may choose to watch or not to watch TV in each period. Once they choose to watch a program, they can do no better than watching it until the end even if it turns out a bad match. In making their viewing decisions in the first period, viewers consider the utility of the program itself and any informational benefits that may result from exposure to tune-ins. As a result of these informational benefits, some viewers watch TV in the first period who would otherwise choose not to watch.

The benchmark model produces a unique perfect Bayesian equilibrium (PBE) in which the TV station airs a tune-in as long as the advertising revenue generated by the viewers continuing to watch offsets or exceeds the cost of airing it. In other words, the TV station airs a tune-in whenever the two programs are similar enough. This finding offers a natural explanation for targeting of audiences which has been a popular topic in the press (especially with the invention of Ti-Vo's). In the absence of a tune-in, no viewer within the first-period audience keeps watching TV in this unique PBE.

I then introduce the possibility of program sampling. Hence, viewers have the option of sampling a few minutes of the second program and switching off if they do not like it. However, sampling is costly; a viewer incurs a utility loss if she switches off after sampling a program. The equilibrium now depends on the value of the sampling cost. For sufficiently small values of the sampling cost, the unique PBE involves no tune-ins. For high values of the sampling cost, the PBE in the benchmark case constitutes the unique PBE. Finally for moderate values of the sampling cost, both PBEs may arise.

In order to analyze the quality-signaling role of advertising, I extend the model by allowing TV programs to be differentiated along two dimensions; one horizontal, one vertical. The vertical dimension is interpreted as the quality of a program which, I assume, is either high or low. If the upcoming program is one of low quality, the TV station may try to mislead viewers so as to attract more viewers. The resulting equilibium depends on the location of the second program and there are both separating and pooling equilibria. Most importantly, airing the quality-certainty optimal number of tune-ins – which is one – may be sufficient to signal high quality in a separating equilibrium. There are program locations for which only a TV station with a high-quality program can afford to air one tune-in; that is these programs do not generate enough audiences to meet the cost of the tune-in when the upcoming program has a low quality since some viewers will switch off after realizing its actual quality.

2.2 Previous Literature

Directly informative advertising has been the topic of several previous studies. As mentioned in Chapter 1, Butters (1977) was the first to model the informative role of advertising. In his paper, products are homogenous. Advertising is the mechanism through which firms inform potential consumers about the price of their products. Because consumers have no knowledge of product existence prior to receiving an ad, the ad informs them of this as well. Grossman and Shapiro (1984) extended Butters' model by introducing differentiated product and heterogeneous consumers. Advertising informs consumers not only about the existence but also about the characteristics of the products. Common to both Butters (1977) and Grossman and Shapiro (1984) is the assumption that the advertising technology is exogenous. So, people cannot change their likelihood of receiving ads.

My model is similar to the one used in Grossman and Shapiro (1984) in that programs and viewer preferences are represented in a spatial framework. I depart from their work by introducing a two-period model and by assuming that program existence is common knowledge. Viewers also have the option of not watching TV, and sampling a program and switching off if they wish. Another important departure of my model is that people are not necessarily passive in receiving ads. More precisely, since tune-ins are always bundled with TV programs, a person receives a tune-in if and only if she chooses to watch the first program.[14]

A related recent paper is by Anderson and Renault (2006) who analyze a monopolist's choice of how much information to disclose in its ad. There is a single consumer who is uncertain about her match value with the monopolist's product. She can learn her match value and the price by conducting a costly search. The monopolist is also uncertain about the consumer's match value. The authors find that the monopolist may advertise only price, only match, or both price and match information depending on the search costs that consumers face. Furthermore, their results show that the monopolist prefers to convey only limited product information. Anderson and Renault use a random-utility model. The consumer's match value is a random draw from a known probability distribution which is common to both the monopolist and the consumer. Therefore, although product existence

[14]Previous work on advertising assumes that people cannot change their likelihood of receiving ads. However, in most real life situations, people can, and actually do, change their likelihood of receiving ads. Take the example of low fare alerts that one can receive in an email from Travelocity. Other examples are using a DVR to skip ads while watching TV, or subscribing to "Do Not Call List" to avoid calls by telemarketers. Although this chapter does not specifically model how people change their likelihood of receiving ads, it allows them to watch the first program even when it yields a negative utility.

is a priori known to the consumer in their model, the monopolist's choice of not advertising the match information is uninformative for her. In the model presented in this chapter, however, the broadcaster knows how its upcoming program matches each viewer's preference. Therefore, the broadcaster's choice of not airing a tune-in is informative for viewers.

To the best of my knowledge, there are no theoretical papers that analyze the role of tune-ins. There are, however, several empirical studies of the effects of tune-ins on viewing choices of individuals. Anand and Shachar (1998) estimate the differential effects of tune-ins on viewing decisions for regular and special shows. They use a novel dataset in their estimation which includes micro-level panel data on the TV viewing choices of a large sample of people and data on program attributes and the frequency of tune-ins. They find that a viewer's utility from a regular show is a positive concave function of the number of times she was exposed to its tune-ins. This indicates that tune-ins are effective but have diminishing returns. The authors also find a significant difference between the effectiveness of regular and special tune-ins, with special ones being less effective when there are few tune-ins and more effective when there are many. They also find the optimal number of tune-ins for a TV network using a very simple model.

In Anand and Shachar (2005), the content of tune-ins is modeled as a noisy signal of program attributes. Consumers are a priori uncertain about program attributes and exposure to tune-ins affect their information sets. Consumers have additional sources of information other than tune-ins, such as word-of-mouth and media coverage. Before each period starts, they update their beliefs based on the tune-ins they have been exposed to and the other information they have received, and then choose the program that maximizes their utility. They find that while exposure to advertising improves the matching of viewers and programs, in some cases it decreases a viewer's tendency to watch a program.

There are main differences between the model in this chapter and the two papers by Anand and Shachar. I improve upon their models by assuming forward-looking viewers rather than myopic. Therefore, viewers correctly anticipate the tune-in strategy of the TV station. Most importantly, they infer that unadvertised programs are not likely to offer a good match. Anand and Shachar only analyze the viewer behavior thereby ignoring the optimal tune-in choices of TV stations. However, tune-in choices of TV stations depend on the viewing decisions of people. By explicitly modeling the optimal TV station behavior, I offer a more thorough analysis of tune-ins and their effects on people's viewing choices.

2.3 Benchmark Model

In this section, I present a simple model with a single TV station and no possibility of program sampling. The analysis provides a benchmark that I build upon for the extensions. The TV station airs two consecutive programs (x_1, x_2), where x_t represents the location of the program in period t over the unit interval. The programs are of the same length. The production costs are assumed to be sunk and the same for both programs, and are set to zero for simplicity. There are A available non-program breaks during each program, where A is taken as exogenous.[15]There is a large number of advertisers, each willing to pay up to $\$p$

[15]The assumption that the number of non-program breaks is fixed is certainly restrictive. However, while U.S. broadcasters are free to choose the number of their non-program breaks, advertising ceilings are imposed on broadcasters in most European countries. Therefore, in most cases, especially in the prime-time, the

per viewer reached for placing a commercial during a program. The TV station may choose to devote one of the non-program breaks in the first period to a tune-in for the purpose of promoting the next program. Production of a tune-in does not entail any costs. I assume that the TV station cannot lie in a tune-in, i.e. the TV station is legally bound to advertise a preview of the actual program in the tune-in, and that the tune-in fully informative. Finally, the objective of the TV station is to maximize total advertising revenues.

On the other side of the market, there is a continuum of N potential viewers who are uniformly distributed along the unit interval with respect to their ideal programs. To each possible program location, there corresponds a viewer for whom that program is the ideal one. An individual derives v units of utility from watching her ideal program that carries A non-program breaks.[16] Formally, a viewer who is located at λ obtains a net viewing benefit $u(\lambda, x) = v - |\lambda - x|$ from watching a program that is located at x. Not watching TV yields zero benefits.[17] I will refer to a particular viewer as "she" when it is convenient.

In each period, viewers choose between watching or not watching TV. An individual's objective is to make the decision at each time that maximizes her total utility. Viewers are assumed to be uncertain only about the type of the program in the second period. When

---

number of non-program breaks that mazimizes a broadcaster's revenue falls below the imposed ceiling. There are also technical reasons for making this assumption. First, if TV stations were allowed to choose the number of non-program breaks, then people would rationally form priors about it. Second, and most importantly, the number of non-program breaks in the first period would possibly provide a signal for the location of the second program. Addressing these issues is beyond the scope of this chapter, since the main focus is on the role of informative advertising. Doing so is an excellent area for future research.

[16] The base utility v also captures how interruptions during a program affect the utility of a viewer. Specifically, the effect of an increase (a decrease) in the nuisance cost of a non-program break on a viewer's utility can be captured by lowering (raising) the base utility.

[17] A constant, t, can be put in front of d that measures the disutility associated with one unit of distance from the ideal program location. However, since the value of not watching TV is zero, utility can easily be expressed as $r - d$, where $r = \frac{v}{t}$.

making their viewing decisions in the first period, viewers not only consider their current utilities, but they also consider the expected informational benefits that they may obtain by seeing a tune-in for the second program. They base their decisions on their prior beliefs about the location of the second program and the equilibrium tune-in strategy of the TV station, which they rationally anticipate. Their unconditional prior belief for the location of the second program is summarized by a uniform density function over $[0, 1]$. Because of possible informational gains associated with watching the first program, some viewers with $\lambda > v$ may also watch the first program in expectations of a higher second period utility.

To simplify the analysis, I assume that the first program is located at 0, and that viewers possess this information.[18] This assumption implies that viewers with ideal program locations to the left of v watch the first program even if there were no tune-ins.

Assumption 2.1 *The first program is located at zero, i.e.* $x_1 = 0$.

I also assume that $\frac{1}{4} < v < \bar{v}$, where $\bar{v} < \frac{1}{2}$ will be derived in the analysis. This assumption places a restriction on the behavior of the viewers who do not watch the first program. To be more specific, when $v < \frac{1}{4}$, none of these viewers watch the second program, and when $v \geq \frac{1}{2}$, all of them do. Futhermore, when v is large (i.e. when $v \geq \bar{v}$), viewers always expect the TV station to air a tune-in in equilibrium and it is difficult to specify their off-the-equilibrium path beliefs when there are no tune-ins.

Assumption 2.2 $\frac{1}{4} < v < \bar{v}$, where $\bar{v} < \frac{1}{2}$.

[18] This assumption does not change any of the main results.

The timing of the game is as follows. First, viewers make their first-period decisions that maximize their expected two-period utilities. The first program starts, and during its progress, the TV station makes its tune-in decision. After the first program ends, if the TV station aired a tune-in, the first-period viewers learn the exact location of the second program. If the TV station did not air a tune-in, they only rationally update their beliefs. Finally, viewers make their second-period optimal decisions and payoffs are realized. As a tie breaking rule, I assume that the TV station airs a tune-in whenever it is indifferent between airing and not airing one, and that people watch TV whenever they are indifferent between watching and not watching.

The equilibrium concept used is perfect Bayesian equilibrium (PBE). That is, we require the TV station to make an optimal tune-in decision taking into account the inferences viewers make in the absence of a tune-in, and in turn, people to make optimal decisions (correctly) anticipating the TV station's strategy. In particular, people's inferences (or posterior beliefs) about the location of the second program following no tune-ins during the first program must be correct.

As a result of the tie-breaking rule, TV station's optimal tune-in strategy is airing a tune-in with certainty if the resulting advertising revenue is at least as high as the revenue that it would earn without airing any tune-ins. Because a tune-in is assumed to be fully informative, and viewers watch a program until the end, the TV station only airs one tune-in. Since the location of the second program is unknown to viewers prior to their first period decisions, they form beliefs about when the TV station would air a tune-in. These beliefs will be described by a set of points Ω such that viewers ex-ante anticipate to see a tune-in for the second program whenever $x_2 \in \Omega$.

To describe the optimal viewer decision in the first period, it is useful to consider an individual whose ideal program location, λ, is to the right of v, i.e. $\lambda > v$. If she watches the first program and sees a tune-in for the second program, she would watch the second program as well provided that its location is at most v units apart from her location. So, her ex-ante expected utility in this case is given by $\int_{\lambda-v}^{\lambda+v} u(\lambda, x)\, \mathbf{1}[x \in \Omega]\, dx$ where $\mathbf{1}[\cdot]$ is an indicator function that equals one when $x \in \Omega$. If she watches the first program and does not see a tune-in, she would keep watching TV provided that her updated expected utility is non-negative. So, her ex-ante expected utility in this case is $\max\left\{0, \int_0^1 u(\lambda, x)\, \mathbf{1}[x \notin \Omega]\, dx\right\}$. Finally, if she does not watch the first program, she would choose to watch the second program baed on her prior belief about its location, that is when $\int_0^1 u(\lambda, x)\, dx > 0$. Hence, the benefit, $B(\lambda)$, of watching the first program for this viewer can be expressed as

$$\begin{aligned} B(\lambda) &= \int_{\lambda-v}^{\lambda+v} u(\lambda, x)\, \mathbf{1}[x \in \Omega]\, dx + \max\left\{0, \int_0^1 u(\lambda, x)\, \mathbf{1}[x \notin \Omega]\, dx\right\} \\ &\quad - \max\left\{0, \int_0^1 u(\lambda, x)\, dx\right\} \end{aligned} \tag{2.1}$$

Without any potential information gains, this viewer would not watch the first program since her direct utility, $(v - \lambda)$ is negative. However, $B(\lambda)$ may be positive So, her optimal first-period decision is to watch TV when $B(\lambda) \geq \lambda - v$.

Ω is determined in equilibrium by viewers' anticipations for the TV station's tune-in strategy corresponding to every possible program location. The marginal benefit of airing a tune-in for the TV station is the marginal second-period advertising revenue as a result of a higher audience size. The cost is the forgone revenue of an additional commercial that the TV station would have earned in the first period without a tune-in. Let the binary variable $q \in \{0, 1\}$ represent the TV station's tune-in decision, where $q = 0$ when it does not air a tune-in, and $q = 1$ when it does. So, from viewers' point of view, the optimal tune-in strategy of the TV station as a function of x_2 is

$$q = \begin{cases} 1, & ApN\left[s_2\left(x_2 \mid q = 1\right) - s_2\left(x_2 \mid q = 0\right)\right] \geq pNs_1 \\ 0, & \text{otherwise} \end{cases}$$

where s_1 is the fraction of viewers watching the first program, and $s_2\left(x_2 \mid q\right)$ is the fraction of viewers watching the second program conditional on the realization of q. The advertising revenue that the TV station forgoes by airing a tune-in is pNs_1 which is on the right-hand side of inequality above. The left-hand side is the increase in the advertising revenue when the TV station airs a tune-in.

The following lemma establishes that there cannot be any discontinuities in viewers' beliefs as to the tune-in strategy of the TV station. This result proves very useful for the remaining of this chapter.

Lemma 2.1 *The beliefs about when the TV station would air a tune-in must be in the form* $\Omega = [x_L, x_H]$.

The proof of Lemma 2.1 (as well as all the remaining proofs) can be found in the Appendix section of this chapter. It argues that if viewers anticipate to see a tune-in for two distinct programs and these programs are advertised in equilibrium, then any program located between these two programs must also be advertised. Therefore, viewers anticipate to see a tune-in for an interval of programs. Given Lemma 2.1, the integrals in equation (2.1) can be further simplified, and accordingly, the benefit of watching the first program can be expressed as

$$\begin{aligned} B(\lambda) \;=\; & \int_{\max\{x_L,\lambda-v\}}^{\min\{x_H,\lambda+v\}} u(\lambda,x)\,dx + \max\left\{0, \left(\int_0^{x_L} u(\lambda,x)\,dx + \int_{x_H}^{1} u(\lambda,x)\,dx\right)\right\} \\ & - \max\left\{0, \int_0^1 u(\lambda,x)\,dx\right\} \end{aligned}$$

Note that all of the three terms in $B(\lambda)$ are continuous functions of λ. However, since the value of λ such that the second term or the third term equals zero may be an interior value (depending on whether these values of λ are less than or greater than v), $B(\lambda)$ may display kinks. Nevertheless, $(B(\lambda) - (\lambda - v))$ must be continuous in λ. Furthermore, it is monotonically decreasing in λ. This is because the increase in $(\lambda - v)$ as λ increases always exceeds the aggregate increase in the other terms. Therefore, if $B(\lambda) > (\lambda - v)$ when $\lambda = v$, then there exists a unique value of λ such that $B(\lambda) = (\lambda - v)$. Let this value of λ be denoted by $\hat{\lambda}$. This critical value of λ also represents the audience share in the first period.

If an individual watched TV in the first period and did not see a tune-in for the second program, she infers that $x_2 \notin \Omega$. In this case, her second-period decision depends on her expected utility conditional on her updated belief. She watches the second program if $E\left[u\left(\lambda, x_2 \mid x_2 \notin \Omega\right)\right] \geq 0$, where

$$E\left[u\left(\lambda, x_2 \mid x_2 \notin \Omega\right)\right] = \int_0^{x_L} u\left(\lambda, x\right) \frac{dx}{1-\left(x_H - x_L\right)} + \int_{x_H}^{1} u\left(\lambda, x\right) \frac{dx}{1-\left(x_H - x_L\right)} \quad (2.2)$$

This function is clearly continuous in λ. Provided that x_H is high enough, the value of λ that equates this function to zero is unique. Using the relationship between $B\left(\lambda\right) - \left(\lambda - v\right)$ and $E\left[u\left(\lambda, x_2 \mid x_2 \notin \Omega\right)\right]$ as well as the equilibrium existence conditions, the next lemma shows that for Ω to be consistent with the TV station's optimal tune-in strategy, the equilibrium value of the lower bound of Ω must be zero. Thus, viewers' inferences from observing no tune-ins take a simple form. This is especially useful in characterizing the equilibrium.

Lemma 2.2 $x_L = 0$ *in equilibrium. That is* $\Omega = [0, x_H]$.

It directly follows from Lemma 2.2 that $x_H > \hat{\lambda}$. This is because the two programs, $x_2 = 0$ and $x_2 = \hat{\lambda}$, yield equal second-period audience sizes if advertised during the first program, i.e. $s_2\left(0 \mid q = 1\right) = s_2(\hat{\lambda} \mid q = 1)$.

Corollary 2.1 *The equilibrium value of* x_H *is greater than* $\hat{\lambda}$.

Now that the form of the PBE has been specified, it is possible to characterize the viewing decisions of individuals. Let λ^o denote the upper bound of the set of individuals who would watch the second program only based on their priors. Formally, it is given by

$$\lambda^o = \max\left\{\lambda \in [0,1] \mid \int_0^1 (v - |\lambda - x|)\, dx \geq 0\right\} \tag{2.3}$$

It is easy to show that $\int_0^1 (v - |\lambda - x|)\, dx \geq 0$ when $\lambda \in \left[\frac{1-\sqrt{4v-1}}{2}, \frac{1+\sqrt{4v-1}}{2}\right]$, so that $\lambda^o = \frac{1+\sqrt{4v-1}}{2}$.

Lemma 2.3 *In equilibrium,* $\hat{\lambda} < x_H < \hat{\lambda} + v$.

This lemma says that for the TV station to air a tune-in, the advertised program does not have to appeal to all the viewers who watch the first program. The final step is to determine when $E\left[u\left(\lambda, x_2 \mid x_2 \notin \Omega\right)\right] \geq 0$. Next lemma establishes this result.

Lemma 2.4 *No viewer from the first-period audience keeps watching TV unless the TV station airs a tune-in for its second program. That is,* $E\left[u\left(\lambda, x_2 \mid x_2 \notin \Omega\right)\right] < 0$ *for all* $\lambda \leq \hat{\lambda}$.

From earlier discussion, the necessary condition for the existence of a PBE is given by

$$s_2\left(x_2 \mid q=1\right) - s_2\left(x_2 \mid q=0\right) \begin{cases} \geq \dfrac{\hat{\lambda}}{A} & \text{for } \forall x_2 \in [0, x_H] \\ = 0 & \text{for } x_2 = x_H \end{cases} \tag{2.4}$$

where $\hat{\lambda}$ is given by the solution to $B(\hat{\lambda}) = \hat{\lambda} - v$, i.e. it is given by

$$\int_0^{x_H} (v - |\hat{\lambda} - x|)dx - \max\left\{0, \int_0^1 (v - |\hat{\lambda} - x|)dx\right\} = \hat{\lambda} - v \tag{2.5}$$

Note from equation (2.5) that if v is large enough so that $x_H = 1$, there is nothing to gain from watching the first program for an individual with location $\lambda > v$. That is, $\hat{\lambda} = v$ when $x_H = 1$. Therefore, we require that v is not too large. The upper limit of v, denoted by $\bar{v}$, is specified in the following proposition, which is the main result of this section.

Proposition 2.1 *Suppose that program sampling is sufficiently costly, so that people stay tuned until a program ends if they chose to watch it. Then there exists a PBE in which viewers with ideal program locations* $\lambda \leq \hat{\lambda}$ *watch the first program, where* $\hat{\lambda} > v$ *is given by*

$$\hat{\lambda} = \begin{cases} \left(\sqrt{A^2 + 2v\left(1+v\right)} - A\right) A & , v < \hat{v} \\ \dfrac{1 - \sqrt{1 - \left(v^2 + \frac{1}{2}\right)\left(1 - \frac{1}{2A^2}\right)}}{\left(1 - \frac{1}{2A^2}\right)} & , \hat{v} \leq v < \bar{v} \\ v & , v \geq \bar{v} \end{cases} \tag{2.6}$$

In this expression, $\hat{v}$ *is the value of* v *that solves* $\left(\sqrt{A^2 + 2v\left(1+v\right)} - A\right) A = \frac{1-\sqrt{4v-1}}{2}$, *and* $\bar{v}$ *is the value of* v *that solves* $\dfrac{1 - \sqrt{1 - \left(v^2 + \frac{1}{2}\right)\left(1 - \frac{1}{2A^2}\right)}}{\left(1 - \frac{1}{2A^2}\right)} = v$. *The TV station airs a tune-in for all* $x_2 \leq x_H^*$, *where* $x_H^* = v + \left(1 - \frac{1}{A}\right)\hat{\lambda}$. *This equilibrium is unique for a wide range of parameter values.*

The equilibrium value of $\hat{\lambda}$ is characterized by a fixed point which is the intersection point of the two best-reply functions:

$$x_H = v + \left(1 - \frac{1}{A}\right)\hat{\lambda}$$

$$\hat{\lambda} = v + \int_{0}^{x_H} (v - |\hat{\lambda} - x|)dx - \max\left\{0, \int_{0}^{1} (v - |\hat{\lambda} - x|)dx\right\}$$

At this fixed point, the TV station's tune-in strategy is correctly incorporated in the decision making of viewers, and the optimal viewing decisions of viewers are correctly incorporated in the decision making of the TV station. Uniqueness of the fixed point is proved in the Appendix using a graphical approach. This is depicted in the following figure.

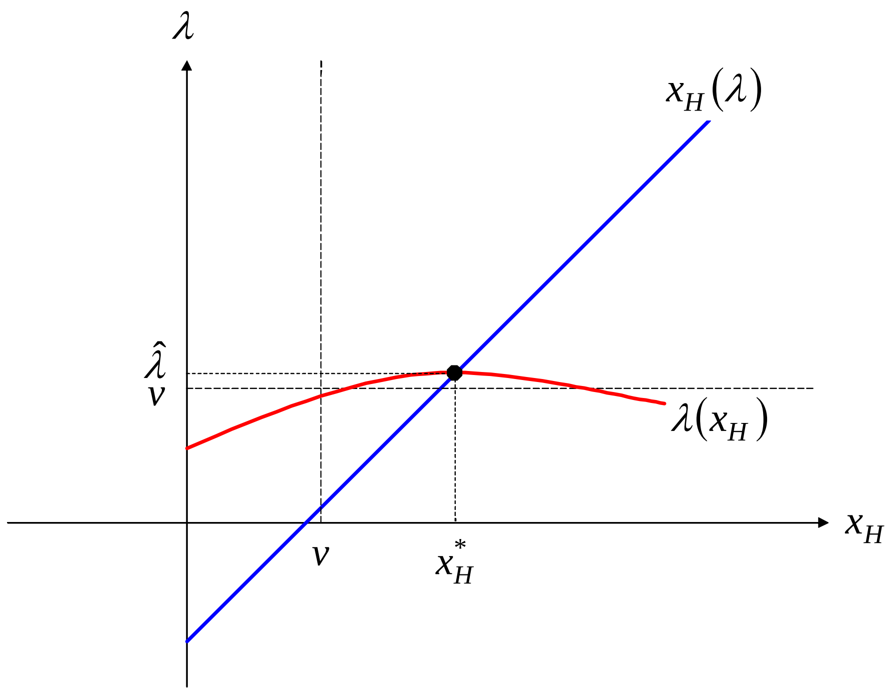

Figure 2.1 Determination of the equilibrium.

The unique PBE of the model is described by a binary tune-in strategy (air a tune-in or not) by the TV station that can be summarized by a unique variable x_H^*. The TV station airs a tune-in whenever the location of the second program is not too far from the location of the first one (which is taken to be zero). Before deciding to watch TV in the first period, viewers consider both their first period utilities and the associated informational benefits. In case there are no tune-ins during the first program, viewers who watched it correctly infer where the second program could possibly lie in and make their decisions accordingly. Knowing that viewers will correctly anticipate the resulting tune-in scheme, it never pays off for the TV station to deviate from this equilibrium decision rule. These results are valid for all values of v as specified in Proposition 2.1 and for all reasonable A. See the proof of Proposition 2.1 in the Appendix for a more detailed explanation regarding the restrictions on the parameter values.

2.4 Program Sampling

In this section, I introduce program sampling whereby people can sample the first few minutes of the second program, if they wish, before they make their final second period decisions. While this process fully reveals the true location of the program, it entails some cost, denoted by $c > 0$ and referred to as the “sampling cost”. This cost is incurred only if an individual opts out after sampling the program and thus enjoys the remaining part of the outside option. It should be interpreted as the amount of the forgone utility that an individual would have enjoyed had she chosen the outside option as the first thing, rather than sampling the program. Therefore, if an individual chooses to turn her TV off after sampling the second program, her net second period benefit would be $-c$.

The model is now slightly more complicated than the benchmark case because the viewers have more options. Given anticipations Ω for the optimal tune-in decision of the TV station, an individual with $\lambda > v$ makes a cost-benefit analysis to find her optimal first period decision. Take an individual whose ideal program location is between v and $v + c$. If she watches the first program and sees a tune-in, she would watch the second program to provided that the advertised program is at most v units apart from her location. If she watches the first program and does not see a tune-in, she would have to decide whether she should sample the second program or not. She would choose to sample if her posterior expected utility of doing so is nonnegative. Given that she chose to sample it, she continues to watch it until the end unless the program location turns out to be more than $v + c$ units apart from her location. If she does not watch the first program, she would base her decision to whether sample the second program or not on her prior beliefs. She would choose to sample if the expected benefit of doing so exceeds its cost, and would keep watching unless the program location is more than $v + c$ units apart from her location. So, we can express the benefits of watching the first program for an individual with location $\lambda > v$ as follows:

$$\begin{aligned} B(\lambda) &= \int_{\lambda-v}^{\lambda+v} u(\lambda, x)\,\mathbf{1}[x \in \Omega]\,dx \\ &\quad + \max\left\{0, \int_{0}^{\lambda+v+c} u(\lambda, x)\,\mathbf{1}[x \notin \Omega]\,dx - (1 - (\lambda + v + c))\,c\right\} \\ &\quad - \max\left\{0, \int_{0}^{\lambda+v+c} u(\lambda, x)\,dx - (1 - (\lambda + v + c))\,c\right\} \end{aligned} \tag{2.7}$$

Each term in $B(\lambda)$ is analogous to the terms given in equation (2.5). The only difference is that we now have additional terms accounting for program sampling. For instance, if this viewer watches the first program and does not see a tune-in, then she may choose to sample the second program. But the program may turn out to be more than $v + c$ units apart from her location in which case she would simply turn off and incur the sampling cost. This is captured by the term $(1 - (\lambda + v + c))\, c$.

Lemma 2.5 *Suppose c is small enough so that*

$$\int_{0}^{\lambda+v+c} u(\lambda, x)\, \mathbf{1}[x \notin \Omega]\, dx - (1 - (\lambda + v + c))\, c > 0 \tag{2.8}$$

Then no viewer with an ideal program location $\lambda > v$ watches TV in the first period.

Lemma 2.5 says that if the sampling cost is so small that a viewer located at $\lambda > v$ would still sample the second program even after watchingthe first program and seeing no tune-ins, then she would simply not watch the first program. So, for sufficiently low values of the sampling cost, absence of a tune-in does not anymore substitute for program sampling for some viewers.

It is possible that no tune-ins prevail in equilibrium if the sampling cost is sufficiently low. As discussed earlier, if the beliefs are such that Ω is nonempty, then it must be true that $v \in \Omega$. This directly follows from the specification of the model; $x_2 = v$ provides the TV station with the highest possible audience size. By Lemma 2.5, if some viewers with

$\lambda > v$ watch the first program, then it must be true that

$$\int_{0}^{\lambda+v+c} u(\lambda, x)\,\mathbf{1}[x \notin \Omega]\,dx - (1 - (\lambda + v + c))\,c < 0$$

This means that no one from the first period audience watches the second program unless they see a tune-in.

No viewer with $\lambda > v$ watch TV in the first period if the equilibrium does not involve a tune-in for any program type. Intuitively, this is because an equilibrium involving a tune-in for some program types must involve a tune in for $x_2 = v$, and if such an equilibrium existed, all of the first period audience would watch the second program.

To characterize the no tune-in equilibrium, suppose beliefs of people are given by $\Omega = (0, v)$. The second-period decisions of the first period audience conditional on no tune-ins is determined by the sign of

$$\int_{v}^{\lambda+v+c} (v - (x - \lambda))\,dx - (1 - (\lambda + v + c))\,c \tag{2.9}$$

Those for whom expression (2.9) is nonnegative watch the second program. It is easy to show that this condition is satisfied when $\lambda + c > \sqrt{2(1-v)\,c}$. So, conditional on $q = 0$, the total mass of viewers from the first-period audience who keep watching TV is given by $v - \left(\sqrt{2(1-v)\,c} - c\right)$.

If $x_2 \in [0, v]$ and the TV station aired a tune-in, then all of the first-period audience would watch the second program. If $s_2(x_2 \mid q = 1) - s_2(x_2 \mid q = 0) < \frac{s_1}{A}$ for every $x_2 \in$

$[0, v]$, there is no reason for people to believe that Ω is nonempty. This inequality is satisfied for every $x_2 \in [0, v]$ when c is such that $\sqrt{2(1-v)c} - c < \frac{v}{A}$. The left hand side, $\sqrt{2(1-v)c} - c$, is increasing in c for $c < \frac{1-v}{2}$. So, there is a cutoff value of c, denoted by c_1, such that there exists a unique PBE that is described by $q = 0$ for all $x_2 \in [0, 1]$, and $\Omega = \varnothing$, when $c < c_1$. Intuitively, when the sampling cost is sufficiently low, the TV station has no incentive to advertise its program since everyone will find it out anyhow.

Now suppose that $c \geq c_1$, and $\Omega = \varnothing$. The TV station would confirm this belief if condition (2.4) is not satisfied. Since $\Omega = \varnothing$, there are no possible informational gains associated with watching the first program. Hence, only the viewers with locations $\lambda \leq v$ watch it. What happens if they receive a tune-in for $x_2 = v$? They would simply keep watching although they would think that the TV station aired a tune-in by mistake. If they do not receive a tune-in, on the other hand, only those with ideal program locations satisfying the following condition would watch the second program

$$\int_0^{\lambda+v+c} (v - |\lambda - x|)\, dx - (1 - (\lambda + v + c))\, c \geq 0$$

This follows from the assertion that $\Omega = \varnothing$. This condition is satisfied when $\lambda \geq v + c - \sqrt{2\left((v+c)^2 - c\right)}$. So, the beliefs are confirmed when $v + c - \sqrt{2\left((v+c)^2 - c\right)} < \frac{v}{A}$, where the left-hand side is $s_2(x_2 \mid q = 1) - s_2(x_2 \mid q = 0)$, and the right-hand side is $\frac{s_1}{A}$. For values of $c \geq c_1$ that satisfy this inequality, $q = 0$ for all $x_2 \in [0, 1]$, and $\Omega = \varnothing$ constitute an equilibrium.

For some values of $c \geq c_1$, there exists another self-fulfilling equilibrium in which the TV station airs a tune-in for all $x_2 \leq x_H$, where $x_H > v$. Formally, watching the first program is more costly than just sampling the second program when $c_1 < c < c_2$, where c_2 is given by[19]

$$c_2 = 1 - \sqrt{(1-2v)^2 - \left(\frac{v}{A}\right)^2} - 2v \tag{2.10}$$

Therefore, no one with $\lambda > v$ watch the first program. However, some viewers choose to sample the second program in the absence of a tune-in. When $c \geq c_2$, on the other hand, the cost of sampling the second program conditional on not watching the first program exceeds the cost of watching the first program.

All of the discussion above is summarized in the following proposition. Its proof has been omitted since it is analogous to the derivation of Proposition 2.1.

Proposition 2.2 *Suppose that sampling a program is possible, but has a cost of c if an individual does not continue watching. Then, depending on the value of the sampling cost, the following constitute a PBE:*

(i) When $c < c_1$, no viewer with location $\lambda > v$ watches the first program, and no tune-in for the second program takes place. All viewers sample the second program.

(ii) When $c_1 \leq c < c_2$, no viewer with location $\lambda > v$ watches the first program. The TV station airs a tune-in for all $x_2 \leq x_H^*$, where $x_H^* = 1 - \frac{\left(c + \frac{v}{A}\right)^2}{2c}$. Some, not all, of the

[19] c_2 can be found by locating the marginal viewer from the first-period audience who keeps watching TV when there is no tune-in given that $\Omega = [0, x_H]$, and then imposing the condition that this value cannot be greater than v.

first-period viewers continues to watch the second program in the absence of a tune-in.

(iii) When $c_2 \leq c < c_3$, viewers with locations $\lambda < \hat{\lambda}$ watch the first program, where $\hat{\lambda} > v$ solves the equation

$$\left(1 - \frac{1}{A^2}\right)\lambda^2 - 2\left(1 + v + c\right)\lambda + \left(v^2 + 2v + \left(2 - 2v - c\right)c\right) = 0$$

The TV station airs a tune-in for all $x_2 \leq x_H^*$, where $x_H^* = v + \left(1 - \frac{1}{A}\right)\hat{\lambda}$. No viewer from the first-period audience continues to watch the second program in the absence of a tune-in.

(iv) When $c \geq c_3$, the equilibrium described in Proposition (1) prevails.

The equilibria in all cases, except for case (ii), are unique.

The cutoff value of the sampling cost that brings us back to the benchmark model can be found by equating the equilibrium value of $\hat{\lambda}$ in case (iii) to that stated in Proposition 2.1. Figure 2.2 below shows how the location of the marginal viewer in the first period and the equilibrium value of x_H^* evolve as a function of the sampling cost.

To summarize, in this extended model, two different equilibria may arise depending on the value of the sampling cost. If the sampling cost is sufficiently low, then the unique PBE exhibits no tune-ins. If it is sufficiently high, then the unique PBE involves a tune-in for the upcoming program unless the two programs are too dissimilar. An important implication in this case is that viewers rationally anticipate the range of programs for which the TV station would air a tune-in. Therefore, no one from the first period audience keeps watching TV unless she was exposed to a tune-in for the second program. For all other values of sampling cost, there are beliefs that support both of the equilibria.

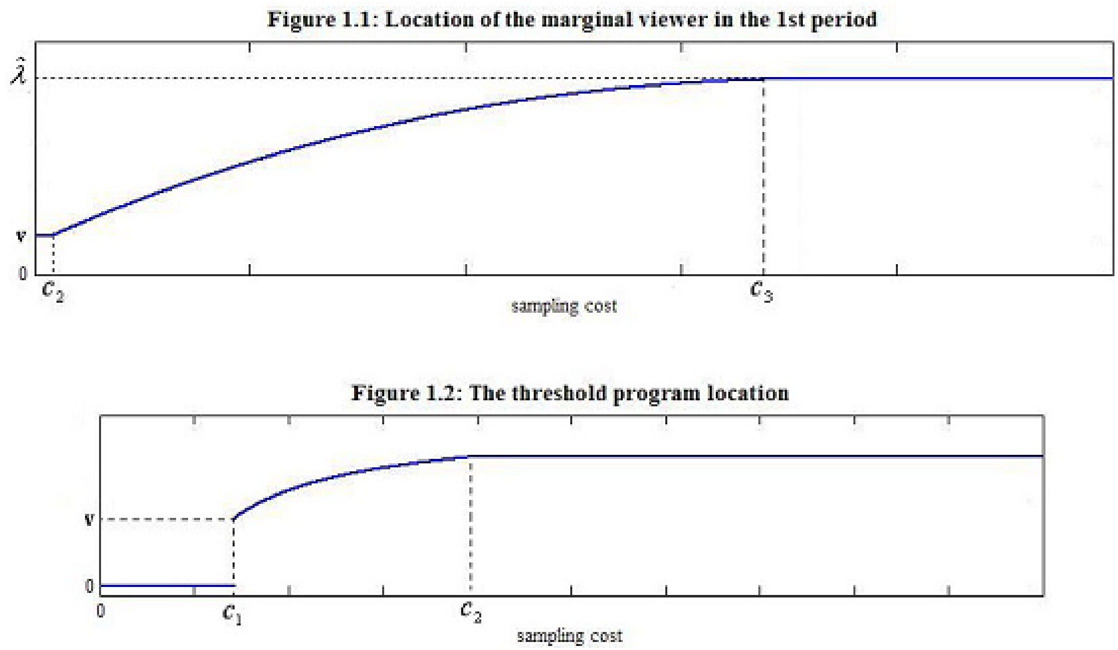

Figure 2.2 Evolution of $\hat{\lambda}$ and x_H^* as a function of c.

Comparing the results with the benchmark model, we see that the fear that an individual may end up watching a bad program until the end leads some people to gather early information by watching the first program. However, this is not necessarily good news for the TV station. Unless these viewers receive a tune-in, they will not keep watching TV. When it is possible to sample a program for a while, however, viewers are not as constrained because they do not have to watch a bad program until the end. Therefore, not as many viewers watch the first program just to alleviate their informational constraints.

2.5 Vertical Differentiation

The model with program sampling can be extended to include a second dimension of differentiation. Below, I sketch the main implications of adding a quality dimension about which viewers are uncertain beforehand. I assume that any direct information the TV station may provide about quality in a tune-in is not reliable. This is a common assumption in the literature. For simplicity, suppose that there are only two quality levels, high or low, and that viewers's utility function is given by $u(\lambda, x) = v_j - |\lambda - x|$, $j = H, L$, where $v_H > v_L$. Also suppose that the first program is of low quality. For ease of exposition, the TV station is referred as the "high-quality station" when its second program is of high quality, and as the "low-quality station" when it is of low quality.

The main findings of the previous section extend to this case, too. Namely, for values of c that are not very small – an assumption I maintain in this section – there exists a unique viewer for whom watching and not watching the first program yields the same expected utility. Let this viewer's location be denoted, as before, by $\hat{\lambda}$. Since the first program is assumed to be of low quality, the first-period audience comprises viewers with locations $\lambda < \hat{\lambda}$. Note that it is still true that $\hat{\lambda} - v_L < c$.

As before, it is optimal for the TV station to air a tune-in for the second program as long as it is similar enough with the first one. This only requires that the sampling cost is not too low. In the absence of a tune-in, no one from the first-period audience watch the second program, regardless of its quality. Therefore, there exists a unique program location x_L^* such that the low-quality station advertises its upcoming program when $x_2 \in [0, x_L^*]$.

Similarly, there exists a unique program location $x_H^* > x_L^*$ such that the high-quality station advertises its upcoming program when $x_2 \in [0, x_H^*]$.

In a separating equilibrium, we need the low-quality station to behave the same as it would behave if viewers knew with certainty that the second program had a low quality. Therefore, it must be true in a separating equilibrium that $x_L^* = v_L + \left(1 - \frac{1}{A}\right)\hat{\lambda}$.

The incentive for the low-quality station to act as if its upcoming program was of high quality comes from the fact that program sampling is costly. Suppose the low-quality station claims in its tune-in that the upcoming program has a high quality, and the viewers believe this statement. Those for whom watching a high-quality program yields nonnegative utility start to watch the second program. After a few minutes, the viewers realize that the TV station actually lied in the tune-in; the program was one of low quality. While some of these viewers switch off at this point, not all do the same. Viewers whose ex-post utilities are at least as high as $-c$ would keep watching since the cost of sampling has already been sunk. So, when the low-quality station lies in a separating equilibrium, the increase in its second-period audience size is at least c.

An important result follows from the discussion above; in a separating equilibrium, only one tune-in for a program located at $x_2 \in (x_L^* + c, x_H^*]$ suffices to signal high quality. Given that the low-quality station does not air any tune-ins for $x_2 > x_L^*$ in a separating equilibrium, it must not have any incentive to falsify viewers by airing a tune-in for $x_2 > x_L^* + c$. Therefore, separation occurs with no distortion in the tune-in strategy. This result is in contrast with the existing literature on quality signaling. A high-quality firm is generally required to engage in dissipative advertising – also referred to as "money burning" – in order to correctly signal its quality. In the current setup, however, it is possible to signal high quality with no distortion in the advertising strategy by simply providing the location of the product. When this information deters a sufficient number of viewers from continuing to watch, it is correctly understood that the program must have a high-quality.

For $x_2 \leq x_L^* + c$, we must have in a separating equilibrium that the high-quality station airs more than one tune-in, and that it is not optimal for the low-quality station to mimic this strategy. The high-quality station would be willing to separate itself as long as the cost of airing the extra tune-in(s) does not exceed the extra revenue it would enjoy by separation. Hence, the high-quality station would do so if the extra audience size generated by separation is at least as high as $k\frac{\hat{\lambda}}{A}$, where k is the number of extra tune-ins required for separation. If $c < \frac{\hat{\lambda}}{A}$, one extra tune-in would be sufficient for separation. If $\frac{\hat{\lambda}}{A} < c < 2\frac{\hat{\lambda}}{A}$, then two more tune-ins are required for separation. I will make the following assumption for the rest of the analysis.

Assumption 2.3 $v_H - v_L > \frac{\hat{\lambda}}{A} > c$.

Separation is not possible when the location of the second program is close to 0, i.e. when the two programs are more similar. This is because such a program would appeal to all of the first-period viewers regardless of its quality. Suppose the second program is also located at 0. In a separating equilibrium, no viewer with $v_L < \lambda \leq \hat{\lambda}$ would sample the second program if they inferred that it has a low quality. If, on the other hand, they inferred that the second program has a high quality, then all of them would continue to watch. However, the high-quality station has to be willing to air an additional tune-in in the first period to separate itself. By separation, it gains an extra audience size of at most $\hat{\lambda} - v_L$. Since $\hat{\lambda} - v_L < c$, the gain by separation falls short of its cost, and therefore, the high-quality station would choose to pool. This argument is valid for all $x_2 < v_L + \frac{\hat{\lambda}}{A}$.

To summarize, when the first program has low quality, we have the following results:

Proposition 2.3 *Under Assumption 2.3, the following constitutes a PBE:*

(i) When $x_2 \leq v_L + \dfrac{\hat{\lambda}}{A}$*, there is no separating equilibrium in quality. Each type airs 'one' tune-in.*

(ii) When $v_L + \dfrac{\hat{\lambda}}{A} < x_2 \leq \left(1 - \dfrac{1}{A}\right)\hat{\lambda} + (v_L + c)$*, the high-quality station airs two tune-ins to signal high quality. The low-quality station airs one tune-in for* $v_L + \dfrac{\hat{\lambda}}{A} < x_2 \leq \left(1 - \dfrac{1}{A}\right)\hat{\lambda} + v_L$ *and airs none for* $x_2 > \left(1 - \dfrac{1}{A}\right)\hat{\lambda} + v_L$*.*

(iii) When $\left(1 - \dfrac{1}{A}\right)\hat{\lambda} + (v_L + c) < x_2 \leq \left(1 - \dfrac{1}{A}\right)\hat{\lambda} + v_H$*, only the high-quality station airs one tune-in amd this is sufficient to signal high quality.*

(iv) When $x_2 > \left(1 - \dfrac{1}{A}\right)\hat{\lambda} + v_H$*, no tune-in takes place (so there is no separation).*

The strategies described in Proposition 2.3 satisfy individual rationality and incentive compatibility constraints for both station types. As mentioned before, the reason for why the high-quality station airs 'two' tune-ins for separation comes from the specification that there is an integer number of tune-ins, and the assumption that $\frac{\hat{\lambda}}{A} > c$. More generally, letting $c = k\frac{\hat{\lambda}}{A} - \varepsilon$, we would need k additional tune-ins by the high-quality station to signal quality. Figure 2.3 below displays the posssible equilibria.

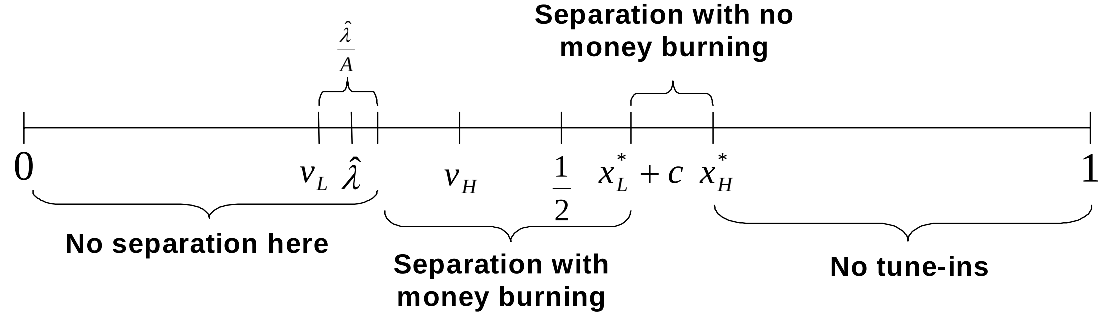

Figure 2.3 Quality signaling when the first program has low quality.

Similar results obtain when the first program has high quality. The only difference is that separation by airing more tune-ins is now also possible for program locations that are sufficiently close to 0. The reason is that there is now a higher number of viewers watching the first program, and therefore the high-quality station can gain enough by separation when $x_2 < \hat{\lambda} - \left(v_L + \frac{\hat{\lambda}}{A}\right)$. The possible equilibria are depicted in Figure 2.4 below.

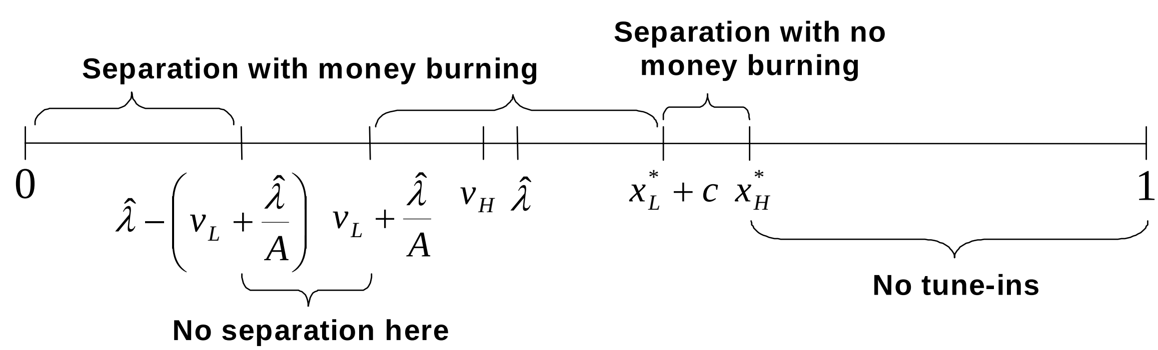

Figure 2.4 Quality signaling when the first program has high quality.

2.6 Conclusion

In this chapter, I have presented a model to analyze the incentives of a firm to provide information about its product. Rationality of people plays a crucial role in the derivation of the equilibrium. It implies that the decision of a firm not to provide information actually reveals useful information to people. This point has largely been ignored in the previous literature. Therefore, the findings in this chapter constitute an important step towards a more comprehensive understanding of the informative role of advertising. Analyzing the TV industry is especially suitable for such a purpose, since tune-ins directly inform people about program characteristics.

The main findings can be summarized as follows. When there is a single TV station offering two consecutive programs that are horizontally differentiated, two equilibria may arise depending on the value of the sampling cost viewers incur. If it is sufficiently low, the unique PBE exhibits no tune-ins. For higher values of the sampling cost, there exists a PBE in which the TV station airs a tune-in for its upcoming program as long as it is similar enough with the program during which it would air the tune-in. Moreover, this PBE is unique if the sampling cost is sufficiently high.

The benchmark model and its extension with program sampling also enable me to analyze quality signaling when viewers are uncertain about the actual quality of the upcoming program. The most important finding in this extension is that money burning is not always necessary for a TV station to signal high quality. This is because there are programs that only a TV station with a high-quality program can afford to advertise.

Some restrictive assumptions have been made in the analysis. First, viewers' priors for the location of the second program were assumed to be uniform. Even though uniform distribution is particularly relevant for describing priors when viewers are clueless for a program's characteristics, allowing for more general probability distributions serves as a useful extension. I believe, however, that similar results will prevail under continuous (quasi)concave distributions although algebra will get significantly more tedious.

It was also assumed that horizontal attributes of programs can be described by a single location. In reality, it is more probably that TV programs are differentiated along several horizontal dimensions. A useful extension may consider including more than one horizontal attribute and analyzing the incentives of TV stations to provide information on multiple dimensions. Moreover, such an extension would enable an application of the model to other industries, such as the market for real estate. There is usually a certain number of characteristics that may be advertised in real estate magazines. Therefore, the content of a particular ad plays a key role in shaping people's beliefs. If a person knows the population distribution of house preferences, then she can infer that the characteristics that are excluded in an ad, if any, have to be the ones that are unappealing to a majority of the recipients of that ad.

2.7 Appendix

Proof of Lemma 2.1 In a PBE, it must be true that $x_2 \in \Omega$ whenever $q = 1$. Suppose $q = 1$. Since people make their first period viewing decisions without seeing a tune-in, the first period audience size does not depend on x_2. This is also true for the audience size in the second period given that the station did not air a tune-in at $t = 1$, i.e. $s_2 (\cdot \mid q = 0)$ only depends on the updated beliefs about the second program. Therefore, $s_2 (x_2 \mid q = 0) + \frac{s_1}{A}$ is constant for all x_2.

$s_2 (x_2 \mid q = 1)$ can be found as follows. First note that the second-period audience comprises people who did not watch the first program. Therefore, these people base their decisions on their prior beliefs. Let $\Delta = \{\lambda > v \mid B(\lambda) \geq \lambda - v\}$ describe the set of people that watch the first program who would not if there were no tune-ins. Note that $\max(\Delta) \leq 2v$ since expected gains can never exceed the utility a person could obtain by watching her ideal program. Δ is determined by Ω in equilibrium. When $x_2 = 0$, only $\lambda \leq v$ from the first period audience watch the second program. Since $0 < x_2 < \max(\Delta) - v$, some people, but not all, from Δ also watch it. All of the first period audience watch the second program when $\max(\Delta) - v \leq x_2 \leq v$. As x_2 gets farther from v, some people will start dropping out, and eventually when $x_2 > \max(\Delta) + v$, no one from the first period audience watches the second program.

So, $s_2 (x_2 \mid q = 1)$ is a linear function of x_2 that monotonically rises for the values of x_2 from 0 to $(\max(\Delta) - v)$. It attains its maximum at $x_2 \in [(\max(\Delta) - v), v]$, and starts monotonically decreasing at $x_2 = v$. Hence, conditional on the existence of a PBE,

$s_2(x_2 \mid q=1)$ can intersect $s_2(x_2 \mid q=0)+\frac{s_1}{A}$ at a maximum of two points Denote these two points x_L and x_H. Then $s_2(x_2 \mid q=1) \geq s_2(x_2 \mid q=0)+\frac{s_1}{A}$ for all $x_2 \in [x_L, x_H]$ in a PBE, which implies that $q=1$ only if $x_2 \in [x_L, x_H]$. Therefore, $\Omega = [x_L, x_H]$. ■

Proof of Lemma 2.2 Assume on the contrary that $\Omega = \{[x_L, x_H] \mid x_L > 0\}$. Since $x_2 = 0$ does not lie in Ω, it must be true that $\hat{\lambda} \notin \Omega$. This follows from the fact that $s_2(0 \mid q=1) = s_2(\hat{\lambda} \mid q=1)$. For $\Omega = \{[x_L, x_H] \mid x_L > 0\}$ to be consistent with TV station's optimal tune-in decision, $[x_L, x_H]$ must satisfy condition (2.4). Since $s_2(x_2 \mid q=0)$ and $\frac{\hat{\lambda}}{A}$ do not depend on the actual value of x_2, it must be true that $s_2(x_L \mid q=1) = s_2(x_H \mid q=1)$ and $x_H = \hat{\lambda} - x_L$. Also note that $x_L < v < x_H$ since $s_2(x_2 \mid 1)$ attains its maximum for $x_2 \in [\hat{\lambda} - v, v]$. We are now ready to present the formal proof.

Step 1: $E[u(\lambda, x_2 \mid x_2 \notin \Omega)] \geq 0 \Leftrightarrow \int_0^1 u(\lambda, x)\,dx - \int_{x_L}^{x_H} u(\lambda, x)\,dx \geq 0$. We also have

$$\int_{x_L}^{x_H} u(\lambda, x)dx = \int_{x_L}^{x_H} (v - |\lambda - x|)dx = x_H - x_L \left(v - \frac{2\lambda - (x_H + x_L)}{2}\right) \text{ for } \lambda > x_H$$

Evaluating $E[u(\lambda, x_2 \mid x_2 \notin \Omega)]$ at $\lambda = \hat{\lambda}$, and using $x_H + x_L = \hat{\lambda}$, we have

$$E[u(\hat{\lambda}, x_2 \mid x_2 \notin \Omega)] = \int_0^1 (v - |\hat{\lambda} - x|)dx - (x_H - x_L)\,(v - \frac{\hat{\lambda}}{2})$$

The equilibrium value of $\hat{\lambda}$ cannot exceed $2v$, so $(x_H - x_L)\,(v - \frac{\hat{\lambda}}{2})$ is positive. This implies that if, in equilibrium, $E[u(\hat{\lambda}, x_2 \mid x_2 \notin \Omega)] \geq 0$, then $\int_0^1 u(\hat{\lambda}, x)dx$ must be positive, too.

Suppose $E[u(\hat{\lambda}, x_2 \mid x_2 \notin \Omega)] \geq 0$. Then

$$B(\hat{\lambda}) - (\hat{\lambda} - v) = \int_{\hat{\lambda}-v}^{x_H} (v - (\hat{\lambda} - x))dx - \int_{x_L}^{x_H} (v - (\hat{\lambda} - x))dx - (\hat{\lambda} - v) = 0$$

$(\hat{\lambda} - v)$ is always at least as large as x_L, since $x_H + x_L = \hat{\lambda}$ and $x_H \geq v$. Rearranging the above condition, we have

$$\frac{((\hat{\lambda} - v) - x_L)^2}{2} = \hat{\lambda} - v$$

The solution to this equation is $x_L = (\hat{\lambda} - v) - \sqrt{2(\hat{\lambda} - v)}$. However, $\sqrt{2(\hat{\lambda} - v)}$ is always bigger than $\hat{\lambda} - v$, since $0 < \hat{\lambda} - v < v$. This contradicts with the initial assumption that $\Omega = \{[x_L, x_H] \mid x_L > 0\}$ constitutes an equilibrium which is consistent with the TV station's optimal tune-in decision. Therefore, it must be true that $E[u(\hat{\lambda}, x_2 \mid x_2 \notin \Omega)] < 0$.

Step 2: Given that $x_H + x_L = \hat{\lambda}$, it can be shown for $\lambda < \hat{\lambda}$ that

$$\frac{dE\left[u\left(\lambda, x_2 \mid x_2 \notin \Omega\right)\right]}{d\lambda} > 0$$

It was shown that for $\Omega = \{[x_L, x_H] \mid x_L > 0\}$ to be consistent with TV station's optimal tune-in decision, $E[u(\hat{\lambda}, x_2 \mid x_2 \notin \Omega)]$ must be negative. This implies that $E\left[u\left(\lambda, x_2 \mid x_2 \notin \Omega\right)\right] < 0$ for every $\lambda < \hat{\lambda}$, which means that no one from the first-period audience watches the second program unless they were exposed to a tune-in. Depending on how larger $\hat{\lambda}$ is relative to v, $s_2\left(x_2 \mid q = 1\right) - s_2\left(x_2 \mid q = 0\right)$ can range between v and

$\hat{\lambda}$. So, from the equilibrium condition $s_2(x_2 \mid q=1) - s_2(x_2 \mid q=0) \geq \frac{\hat{\lambda}}{A}$, we have $v < \frac{\hat{\lambda}}{A} < \hat{\lambda}$, which cannot be true as long as $A \geq 2$. Therefore $\Omega = [x_L, x_H]$ cannot be consistent with the TV station's optimal tune-in decision unless $x_L = 0$. ■

Proof of Lemma 2.3 It will be enough to show that $x_2 = \hat{\lambda} + v$ does not belong to Ω in equilibrium. Note that

$$s_2(\hat{\lambda} + v \mid q = 1) = \lambda^o - \max\left\{\hat{\lambda}, 1 - \lambda^o\right\}$$

since only the people who did not watch the first program watch the second one. But these people's viewing decisions do not depend on q. Therefore, we have

$$s_2(\hat{\lambda} + v \mid q = 1) \leq s_2(\hat{\lambda} + v \mid q = 0)$$

So $x_H \geq \hat{\lambda} + v$ can never happen in equilibrium. ■

Proof of Lemma 2.4 Suppose, on the contrary, that some people keep watching. Then $\lambda = \hat{\lambda}$ has to be one of these viewers. If $\int_0^1 u(\hat{\lambda}, x)dx > 0$, then the condition $B(\hat{\lambda}) = (\hat{\lambda} - v)$ is expressed as

$$\int_{\hat{\lambda}-v}^{x_H} (v - |\hat{\lambda} - x|)dx - \int_0^{x_H} (v - |\hat{\lambda} - x|)dx = \hat{\lambda} - v$$

Rearranging the left-hand side, we have

$$-\int_0^{\hat{\lambda}-v}(v-(\hat{\lambda}-x))dx=\hat{\lambda}-v$$

The value of the integral on the lefthand side is $\frac{(\hat{\lambda}-v)^2}{2}$. But $(\hat{\lambda}-v)$ is always greater than $\frac{(\hat{\lambda}-v)^2}{2}$. So, if $\int_0^1 u(\hat{\lambda},x)dx>0$, then $E[u(\hat{\lambda},x_2 \mid x_2\notin\Omega)]$ cannot be positive. Now, if $\int_0^1(v-|\hat{\lambda}-x|)dx\leq 0$, then the condition $B(\hat{\lambda})=(\hat{\lambda}-v)$ becomes

$$\int_{\hat{\lambda}-v}^1(v-|\hat{\lambda}-x|)dx=\hat{\lambda}-v$$

This condition can be rearranged as

$$\int_0^1(v-|\hat{\lambda}-x|)dx=\int_0^{\hat{\lambda}-v}(v-|\hat{\lambda}-x|)dx+(\hat{\lambda}-v)$$

However, $\int_0^{\hat{\lambda}-v}(v-(\hat{\lambda},x))dx=\frac{-(\hat{\lambda}-v)^2}{2}$, so that $\int_0^{\hat{\lambda}-v}(v-|\hat{\lambda}-x|)dx+(\hat{\lambda}-v)>0$. This contradicts with $\int_0^1(v-|\hat{\lambda}-x|)dx\leq 0$. So, it has to be true that $E[u(\hat{\lambda},x_2 \mid x_2\notin\Omega)]<0$. Since $E[u(\hat{\lambda},x_2 \mid x_2\notin\Omega)]<0$ is increasing in λ for $\lambda\leq\hat{\lambda}$, no one with $\lambda\leq\hat{\lambda}$ keeps watching conditional on exposure to no tune-ins. ■

Proof of Proposition 2.1 Condition (2.4) evaluated at $x_2=x_H$ is $\hat{\lambda}-(x_H-v)=\frac{\hat{\lambda}}{A}$ since $s_2\left(x_H \mid q=1\right)=\hat{\lambda}-(x_H-v)$ and $s_2\left(x_H \mid q=0\right)=0$ by Lemma 2.4. Rearranging, we have $x_H=v+\left(1-\frac{1}{A}\right)\hat{\lambda}$. The value of $\hat{\lambda}$ is obtained by simultaneously solving condition (2.4) when $x_2=x_H$, and equation (2.5). Suppose the equilibrium value of $\hat{\lambda}$ turns out to be less than $\left(1-\lambda^{0}\right)$, in which case the second term in equation (2.5) equals zero. Then, $\hat{\lambda}$ is given by the solution to

$$\int_{\hat{\lambda}-v}^{v+\left(1-\frac{1}{A}\right)\hat{\lambda}}(v-|\hat{\lambda}-x|)dx=\hat{\lambda}-v$$

The integral on the lefthand side equals $(v^{2}-\frac{\hat{\lambda}^{2}}{2A^{2}})$. Solving the equation for $\hat{\lambda}$, we obtain

$$\hat{\lambda}=\left(\sqrt{A^{2}+2v\left(1+v\right)}-A\right)A$$

This constitutes an equilibrium as long as it is less than $(1-\lambda^{o})$, which is satisfied when $v<\hat{v}$. Similarly, if the equilibrium value of $\hat{\lambda}$ turns out to be greater than $(1-\lambda^{o})$, then $\hat{\lambda}$ is obtained by solving

$$\int_{\hat{\lambda}-v}^{v+\left(1-\frac{1}{A}\right)\hat{\lambda}}(v-|\hat{\lambda}-x|)dx-\int_{0}^{1}(v-|\hat{\lambda}-x|)dx=\hat{\lambda}-v$$

Solving this equation for $\hat{\lambda}$,we obtain

$$\hat{\lambda}=\frac{1-\sqrt{1-\left(v^{2}+\frac{1}{2}\right)\left(1-\frac{1}{2A^{2}}\right)}}{\left(1-\frac{1}{2A^{2}}\right)}$$

This value of $\hat{\lambda}$ constitutes an equilibrium as long as it is greater than $(1-\lambda^{o})$ which is satisfied when $v\geq\hat{v}$.

Uniqueness directly follows from the lemmas since there is only one possible form of equilibrium. Here, we only show the existence of a unique solution. From the equilibrium condition for the TV station we have $x_H^* = v + \left(1 - \frac{1}{A}\right)\hat{\lambda}$. Solving equation (2.4) for $\hat{\lambda}$ in terms of x_H, we obtain

$$\hat{\lambda} = v + (2 - x_H) - \sqrt{(2 - x_H)^2 - (1 - 4v + x_H)(1 - x_H)}$$

Denoting the right hand side with $h(x)$, we can state the problem as

$$x = v + \left(1 - \frac{1}{A}\right) h(x)$$

Taking the derivative of h, we obtain

$$\frac{dh}{dx} = \frac{2(1 + v - x)}{(h(x))^2} - 1$$

It can be shown that h is concave and attains its global maximum at $x^* = 1 + v - \sqrt{\frac{3}{2} - v^2}$, which is positive for all $v > \frac{1}{4}$. Note that $v + \left(1 - \frac{1}{A}\right) h(x)$ is less than x at $x = 1$, and is greater than x at $x = 4v - 1$ if $v < \dfrac{1}{2 - \frac{1}{A}}$. This is true for almost all values of v and A. Hence there must an interior solution to $x = v + \left(1 - \frac{1}{A}\right) h(x)$, and this is unique. Also note that v is less than $\frac{1}{2}$, so this equilibrium does not require any off-the-path beliefs. ■

Proof of Lemma 2.5 First observe that

$$\frac{\partial\left[\int_0^{\lambda+v+c}(v-|\lambda-x|)\,dx+(1-(\lambda+v+c))\,(-c)\right]}{\partial\lambda}>0 \text{ for } v<\lambda\leq v+c$$

When $\lambda=v$, the value of the integral in the derivative becomes $v^2-\frac{c^2}{2}-(1-(2v+c))\,c$, which can be rearranged as $\frac{1}{2}\left(c^2-2\,(1-2v)\,c+2v^2\right)$. This is positive for every $c<v$ when $v\geq\frac{1}{4}$. Therefore the last term in $B\,(\lambda)$ is positive for $v<\lambda\leq v+c$. Suppose c is small enough as stated in the lemma. Then, $B\,(\lambda)\geq(\lambda-v)$ when

$$\int_0^{\lambda-v}u\,(\lambda,x)\,(\mathbf{1}\,[x\notin\Omega]-1)\,dx+\int_{\lambda+v}^{\lambda+v+c}u\,(\lambda,x)\,(\mathbf{1}\,[x\notin\Omega]-1)\,dx\geq\lambda-v$$

Suppose there exists $v<\hat{\lambda}\leq v+c$ such that $B(\hat{\lambda})=\left(\hat{\lambda}-v\right)$. Then, $\mathbf{1}\,[x\notin\Omega]=1$ for $x>\hat{\lambda}+v$ since no one from the first-period audience keeps watching if the TV station advertises $x_2>\hat{\lambda}+v$. Now, the condition for $B(\lambda)=(\lambda-v)$ becomes

$$\int_0^{\lambda-v}u\,(\lambda,x)\,(\mathbf{1}\,[x\notin\Omega]-1)\,dx=\lambda-v$$

However, the left-hand side is at most $\frac{(\lambda-v)^2}{2}$ for any Ω. This is equal to $\lambda-v$ only when $\lambda=v$. So there is no $v<\lambda\leq v+c$ such that $B\,(\lambda)=(\lambda-v)$. People with $\lambda>v+c$ never watch the first program since their first-period disutility exceeds the sampling cost.

■

Chapter 3

Two Broadcasters and Strategic Provision of Tune-ins

3.1 Introduction

This chapter presents a duopoly extension of the model presented in the previous chapter.I maintain the assumption that viewers know the locations of the first programs at both stations beforehand. They are uncertain about their match values with the programs in the second period. A crucial element of the model is that the TV stations know how their own as well as their rival's program fits viewers' preferences. This information is also known to viewers. It is shown that this information structure leads to an equilibrium in which the decision of a TV station to advertise its own program may strategically depend on its rival's program, and viewers correctly incorporate this strategic interaction into their decision-making.

The previous chapter analyzed the extent to which a single TV station is willing to air a tune-in. In order to extend that analysis to include a second station, one has to introduce the possibility of switching from one station to the other. In fact, Section 2.4 introduced as an extension the possibility of switching off after sampling a few minutes of a program. In this chapter, I move one step further by incorporating viewers' switching behavior into the same setup when there are two TV stations. I assume that the amount of time required for learning the actual location of a program is fixed and the same for all programs and viewers. However, this process entails an opportunity cost if an individual does not continue to watch the program she chose to sample.

The process of costly sampling plays a crucial role for two reasons. First, for an equilibrium that involves the use of tune-ins to exist, sampling cost (or equally switching cost) has to be positive. Had it been zero, viewers could costlessly learn the programs at both stations and make their decisions without any uncertainty. Therefore, there would be no need for tune-ins. Second, a positive sampling cost may create an incentive for a station to choose not to air any tune-ins. This is because of the fact that the cost of sampling becomes sunk once a viewer chooses to engage in sampling. That is, when there is costly sampling, some individuals may end up watching a program that they would not choose to watch with complete information. By the same token, an individual's final decision may not be the one that maximizes her utility with complete information. That a positive sampling cost is necessary for the existence of tune-ins is easily confirmed by the statistics given earlier. It is harder to establish empirical support for the second role of sampling because of limited individual-level data. However, Emerson and Shachar's finding mentioned in the previous chapter is supportive of such a role.

When the TV stations are informed about each other's program content, their decision to air a tune-in may transmit information about not only their own program but also their rival's. Since sampling is costly, a station is relatively more inclined to air a tune-in in order to lock-in its current viewers when its rival has a more similar program. However, this may signal to the recipients of that tune-in that the program at the other station is more likely to be a good match than what they thought before. Similarly, if a station does not advertise its upcoming program, it does not necessarily mean that it is a bad match for that station's viewers. It could rather be the case that it is a better match than the other program. In this chapter, I am primarily interested in exploring the nature of such strategic behavior, and ultimately in finding out if an equilibrium in which viewers' priors are changed at an interim stage exists. I show that such an equilibrium exists although it is not unique. Without any restriction on viewers' beliefs, there is another equilibrium in which viewers' beliefs about either program are unchanged regardless of the tune-in decisions of the stations.

Signaling has traditionally been investigated within the context of vertically differentiated products. When consumers are uninformed about the actual quality of an experience good, it has been shown that a high-quality seller can credibly signal this information by setting a high price or by spending a nontrivial amount of money on uninformative advertising. The findings in this chapter suggest that signaling is also possible in horizontally differentiated markets. Signaling occurs regardless of whether a station chooses to or not to air a tune-in. In the former case, only information about the other program is signalled. In the latter, information about both programs is signalled. However, a fully separating equilibrium does not arise in the current setting; viewers cannot locate the programs with certainty.

For certain program locations, the model can also be interpreted as one with vertical differentiation. To be more specific, when a station's upcoming program is better suited to all of its current viewers than the other station's upcoming program, the two programs are effectively vertically differentiated for those viewers. In such a situation, I find that the former station does not air a tune-in. Although this result is strikingly different than what traditional models of signaling predict, a direct comparison may be misleading since TV programs are not experience goods. Neverthless, it is interesting to note that signaling is possible even in the absence of advertising.

I also analyze the welfare effects of a possible ban on the use of tune-ins. I find that when it is not a credible strategy for a station to behave strategically, it may be welfare improving to ban tune-ins. In such a situation, the stations advertise their programs more often. Although viewers enjoy a higher surplus as a result of improved information, social welfare is reduced because the decrease in revenues of the TV stations overweights the increase in consumer surplus.

3.2 Previous Literature

Confining attention to the literature on informative advertising in horizontally differentiated oligopolies, a related paper is Meurer and Stahl (1994). They analyze the welfare properties of informative advertising in a duopoly where a fraction of buyers are uninformed about the product characteristics. There are two types of buyers. One type is ideally matched with one firm and the other type is ideally matched with the other firm. As in Butters (1977) and Grossman and Shapiro (1984), a firm chooses its advertising intensity and a random fraction of consumers receive the ad. Advertising informs a buyer of her best match. Firms choose their prices after advertising takes place. They treat product information as a public good, which implies that information about one product provides information about the others as well. They characterize a unique subgame perfect Nash equilibrium in which the level of advertising provided may be more or less than socially optimal. While advertising improves the match between consumers and products, it gives firms a higher market power by increasing brand loyalty.

Within the same strand of literature, another related paper is Anand and Shachar (2006). They use the same setup with that of Meurer and Stahl (1994) with three major differences. First, a firm can only advertise through one or both of the two available media channels, and consumer preferences over product attributes are perfectly correlated with their choice of media channel. So, for instance, if consumers of media channel 1 are ideally matched with product 1, then firm 1 can target these consumers by advertising through media channel 1. Second, advertising messages are noisy in the sense that consumers may get the wrong idea from a firm's informative ad. Therefore, firms advertise more than once. Finally, firms do not choose prices, which are therefore suppressed in the analysis. In such a setting, Anand and Shachar (2006) characterize a separating equilibrium in which a firm advertises only to those consumers for whom that product is the ideal one. As long as the ads are not completely noisy – in which case the ads would equally be interpreted right or wrong – there exists a threshold amount of advertising which ascertains a consumer that the advertised product is her best match. Thus, regardless of the content of the ad, each consumer purchases the product that she was advertised to.

There are major differences between my model and those of Meurer and Stahl (1994) and Anand and Shachar (2006). First, in both of these papers, products are experience goods, so consumers do not have the option of obtaining product information by a costly search. I treat TV programs as search goods since program sampling is a common practice in real life. If I rather treated them as experience goods, the unique symmetric equilibrium would involve no strategic behavior by the stations. Therefore, sampling plays a crucial role for the results presented in this chapter. Second, there are only two distinct types of consumers in both papers, one ideally matched with one product and the other with the other product. In Meurer and Stahl (1994), this assumption implies that an informative ad by one firm necessarily informs the recipient about the other firm's product, and therefore plays a critical role for their results. In Anand and Shachar (2006), it is a necessary assumption for perfect separation. In my model, on the other side, there is a continuum of people who may or may not be ideally matched with either program. Therefore, the tune-in decision of a station is a function of the program location of the other station. Third, advertising in my model is purely informative unlike as in Anand and Shachar (2006), and reaches a nonrandom group of consumers unlike as in Meurer and Stahl (1994). In this sense, I use a different advertising technology.

The findings in this chapter are also (weakly) related to the literature on quality signaling with multiple senders when firms have common knowledge of product qualities. As described in Chapter 1, Hertzendorf and Overgaard (2001a) and Fluet and Garella (2002) constitute two important duopoly extensions of Milgrom and Roberts (1986). Hertzendorf and Overgaard consider a static duopoly in which nature selects one firm as the high-quality producer and the other as the low-quality producer. Both firms observe their qualities before they make their price-advertising decisions. Thus, one firm's price-advertising choice also reveals information about the quality of the other firm's product .Assuming that variable costs are independent of quality, they identify a unique separating equilibrium (after applying two equilibrium refinements) in which the high-quality producer uses dissipative advertising to signal its quality.when the quality difference is sufficiently small. For higher quality differentials, a high price is sufficient to signal quality without any advertising. Similarly, Fluet and Garella consider a static duopoly model in which each firm is informed about the quality of both products (which may be either high or low). Assuming that variable costs are increasing in quality, they similarly find that a positive level of advertising by a high-quality firm is necessary for separation when the quality difference is sufficiently small.

Bontems and Meunier (2005) also consider a two-sender duopoly model of quality signaling when the products are both vertically and horizontally differentiated. As in Hertzendorf and Overgaard (2001a), nature assigns only one of the firms as the high-quality producer. However, this assignment occurs after firms choose their locations. In contrast to Hertzendorf and Overgaard (2001a) and Fluet and Garella (2002), the authors find that a positive level of advertising is necessary for separation regardless of the degree of vertical differentiation. Firms choose maximal horizontal differentiation when the the quality difference is small and minimal horizontal differentiation when it is sufficiently high.

Matthews & Fertig (1990) consider an incumbent-entrant setup where product quality of the entrant is known to both firms while that of the incumbent is known by everyone in the market. Prices are exogenous and the incumbent may advertise in order to inform consumers about the product quality of the entrant; i.e. it may counteract misleading attempts by a low-quality entrant. In sharp contrast to the existing models of quality signaling, they show that a high-quality entrant can successfully signal its quality by spending an infinitesimal amount on advertising.[20]

3.3 The Model

The basic setup follows from the previous chapter with the exception that there are now two TV stations, station Y and station Z, each airing two programs in two consecutive time periods. The programs are characterized by their locations on the unit line. They are of the same length and have zero production costs. There are A available non-program breaks during each program in each period. There is a large number of firms that are willing to pay up to $\$p$ per viewer reached for placing an advertisement during a program in each period.

On the other side of the market, there is a continuum of N potential viewers who are uniformly distributed on the unit line with respect to their ideal program types. To each possible program type on the unit line, there corresponds a viewer for whom that program is the ideal one. Individuals have the same program preferences during both periods. An

[20] For other studies of signaling when there are multiple senders with common information, see Bagwell and Ramey (1991), de Bijl (1997) and Hertzendorf & Overgaard (2001b).

individual derives v units of utility from watching her ideal program that carries A non-program breaks. Formally, a viewer who is located at a distance of d units from a program obtains a net viewing benefit $v - d$. Not watching TV yields zero benefits. The parameter λ will be used to represent the location of an individual, and a particular individual will be referred to as “she” when it is convenient.

The locations of the first programs is assumed to be common knowledge. Although people know that each station offers two consecutive program, they do not know where on the unit line the second programs are located. Denoting the location of second program of station Y with y and that of Z with z, I assume that viewers’ priors are given by $y, z \in \left\{0, \frac{1}{2}, 1\right\}$. From their perspective, each of these three locations is equally likely to be the actual location of each program. The stations know the location of their own as well as their rival’s program, and people know that the stations have this information. They may devote one of the non-program breaks in the first period to a tune-in.[21] A tune-in may only include information about the actual location of the upcoming program at the same station. I assume that the TV stations cannot lie, i.e. they are legally bound to advertise a preview of the actual upcoming program in the tune-in. The objective of the TV stations is to maximize total advertisement revenues (for simplicity, it is assumed that there is no discounting).

Viewers have the option of switching to the other station or simply turning the TV off after sampling a few minutes of a program. I assume that the amount of time required for

[21] They would never air more than one tune-in because tune-ins are assumed to be fully informative, and viewers do not switch stations in the first period.

learning the true location of a program is fixed and same for both programs and for all individuals. This sampling process entails a cost of c, and is referred to as the "sampling cost". A viewer incurs one unit of the sampling cost if she samples the programs at both stations and ends up watching the one that yields a higher utility. If an individual decides the turn her TV off after sampling one of the programs, then her net utility is $-c$. If she does so after sampling both programs, then her net utility is $-2c$. Since the locations of the first programs are known beforehand, viewers do not engage in sampling in the first period. However, sampling one or both of the stations may be optimal in the second period. An individual's objective is to make a decision at each time that maximizes her total utility.

I maintain the following three assumptions throughout the analysis.

Assumption 3.1 $\frac{1}{4} + c < v < \frac{1}{2} - c$*, where* $c > 0$*.*

Assumption 3.2 $\frac{1}{4} + \frac{1}{2A} < v < \frac{1}{2} - \frac{1}{A}$*, where* $A > 0$*.*

Assumption 3.3 *The first programs at stations* Y *and* Z *are located at* $\frac{1}{4}$ *and* $\frac{3}{4}$*, respectively, and this is common knowledge.*

The first and the second assumptions are made in order to rule out unreasonable equilibria. This shall be more clear as the analysis proceeds. Note that it imposes an upper bound on the value of the sampling cost, and a lower bound on the number of non-program breaks. To be more specific, it is implied that $0 < c < \frac{1}{8}$ and $A > 6$. The third assumption is made in order to simplify the analysis. Combined with the first assumption, it implies that viewers on the lower half of the unit line watch station Y and the ones on the upper half watch station Z.

The timing of the moves is as follows. First, people make their first period viewing decisions. Then the first program starts, and the TV stations make their tune-in decisions during the first program. Having watched the first program, people update their beliefs about the second programs depending on whether or not they were exposed to a tune-in. The second programs start and people make their optimal sampling decisions. After each individual completes sampling one or both (or none) of the stations, they make their final second period viewing choices and the payoffs are realized.

3.3.1 Equilibrium

The equilibrium concept used is perfect Bayesian equilibrium (PBE). That is, the TV stations make optimal tune-in decisions taking the location of their rival's program and the rationality of people into account, and people make optimal sampling and viewing decisions after observing the tune-in decision of the station they have watched. In particular, people's inferences (or posterior beliefs) after the first period about the locations of the second programs must be correct, and the TV stations should not have any incentive to deviate.

As discussed in the introduction section of this chapter, TV stations may choose to behave strategically due to their knowledge of the rival station's program. However, regardless of the location of the rival's program, a station clearly never airs a tune-in for a program that none of its current viewers would like to watch. This case arises for station Y

when $y = 1$, and for station Z when $z = 0$. Given that a station cannot communicate any information with the viewers of the other station, it does not pay off for either station to air a tune-in for such programs. Let $q_i(y, z)$ be a binary variable that assumes a value of 1 if station i airs a tune-in when the two programs are located at (y, z), and 0 otherwise. So, if station Y airs a tune-in for $y = 0$ when $z = 0$, then we have $q_Y(0, 0) = 1$. The following lemma is immediate.

Lemma 3.1 $q_Y(1, z) = 0$ *for all* z*, and* $q_Z(y, 0) = 0$ *for all* y*.*

Next, consider a situation in which neither of the stations air a tune-in for their upcoming programs regardless of their locations. Suppose that these strategies constitute a PBE. In such a "no tune-in" equilibrium, people's priors would be unchanged. This means that all viewers are indifferent between sampling either station. So, if sampling occurs, a random half of viewers initially sample Y and the remaining ones initially sample Z. Suppose a λ-type viewer chooses to sample one of the programs. If λ is such that $0 \leq \lambda \leq \frac{1}{2} - v - c$, then this viewer knows that she would only watch a program located at 0. If the program that she first samples is not at 0, should she also sample the program at the other station? Unless the other program happens to be located at 0, she would turn her TV off, and her net utility would be $-2c$ since she would have sampled both programs and ended up taking the outside option. So, the expected utility of sampling the other station is $\frac{1}{3}(v - c - \lambda) + \frac{2}{3}(-2c)$. On the other hand, if she switches off without sampling the other station, she would enjoy a utility of $-c$. Hence, she should engage in a second sampling if $\frac{1}{3}(v - c - \lambda) + \frac{2}{3}(-2c) \geq -c$, or equivalently if $v - \lambda \geq 2c$. The left-hand side is decreasing in λ, so if this inequality is satisfied at $\lambda = \frac{1}{2} - v - c$, it has to be true for all $\lambda \leq \frac{1}{2} - v - c$. Evaluating at $\lambda = \frac{1}{2} - v - c$, we get $2v - c \geq \frac{1}{2}$ which is always true by Assumption 3.1.

We also need to check if engaging in sampling is optimal at all for this person. Expected utility of doing so is $\frac{1}{3}(v-\lambda)+\frac{2}{3}\left[\frac{1}{3}(v-c-\lambda)+\frac{2}{3}(-2c)\right]$, where the second term is due to the fact that it is also optimal to sample the other station when the first program sampled is not at 0. If this is nonnegative, then it is optimal to engage in sampling for viewers with locations $\lambda \leq \frac{1}{2} - v - c$. Rearranging, expected utility becomes $\frac{1}{3}[v - \lambda - 2c]$, which is the same condition as in the previous paragraph, and therefore is nonnegative.

Now, take a viewer with location $\lambda \in \left[\frac{1}{2} - v - c, \frac{1}{4}\right]$ and suppose that this viewer samples station Y. She stays at Y if y is located at 0. If it turns out that $y = \frac{1}{2}$, she may also want to check out station Z in the hope of finding out $z = 0$. But there is also the chance that z is $\frac{1}{2}$ or 1. If $z = 1$, she would switch back to station Y. If, on the other hand, $z = \frac{1}{2}$, she would be indifferent between the two stations. So, the expected utility of switching to station Z when $y = \frac{1}{2}$ is $\frac{1}{3}(v-c-\lambda)+\frac{2}{3}\left(v-c-\frac{1}{2}+\lambda\right)$. If this expression is greater than the utility of staying at Y, $v-\left(\frac{1}{2}-\lambda\right)$, she should switch to and sample the program at station Z. This is satisfied when $\lambda < \frac{1}{4} - \frac{3c}{2}$. So, when $y = \frac{1}{2}$, it is optimal to also sample Z for the viewers with locations $\frac{1}{2} - v - c \leq \lambda < \frac{1}{4} - \frac{3c}{2}$. Finally, suppose it turns out that $y = 1$. The expected utility of switching to Z is $\frac{1}{3}(v-c-\lambda)+\frac{1}{3}\left(v-c-\frac{1}{2}+\lambda\right)+\frac{1}{3}(-2c)$ which equals $\frac{1}{3}\left(2v-4c-\frac{1}{2}\right)$. This is greater than $-c$ when $2v - c > \frac{1}{2}$, which is again true by Assumption 3.1.

The analysis above equally applies to other possibilities as well. So, as a general rule, stopping sampling is optimal when the location of the program sampled first is at most at a distance $\frac{1}{4}+\frac{3c}{2}$ from a viewer's own location. We can now express the audience shares of stations Y and Z for all possible values of (y,z) under the assumption that there are no tune-ins. Suppose, for instance, that $(y,z)=\left(0,\frac{1}{2}\right)$. A random half of the viewers sample station Y first. Among these viewers, those with $\lambda\leq\frac{1}{4}+\frac{3c}{2}$ stay at Y while the rest switch to Z. Since $z=\frac{1}{2}$, those with $\lambda>\frac{1}{2}+v+c$ turn their TVs off. From among the other half who chose to sample Z first, the ones with $\lambda\in\left[\frac{1}{4}-\frac{3c}{2},\frac{3}{4}+\frac{3c}{2}\right]$ stay at Z while the others switch to Y. Those with $\frac{1}{4}-\frac{3c}{2}<\lambda$ stay at Y. The same is not true for $\lambda>\frac{3}{4}+\frac{3c}{2}$. The program $y=0$ is not favorable for them, so those with locations $\lambda\in\left[\frac{3}{4}+\frac{3c}{2},\frac{1}{2}+v+c\right]$ switch back to station Z while the rest switch off. Arguing along similar lines, we get the following audience shares:

Station Y

	$\mathbf{z=0}$	$\mathbf{z=\frac{1}{2}}$	$\mathbf{z=1}$
$\mathbf{y=0}$	$\frac{v+c}{2}$	$\frac{1}{4}$	$v+c$
$\mathbf{y=\frac{1}{2}}$	$v+c+\frac{1}{4}$	$v+c$	$v+c+\frac{1}{4}$
$\mathbf{y=1}$	$v+c$	$\frac{1}{4}$	$\frac{v+c}{2}$

Station Z

	$\mathbf{z=0}$	$\mathbf{z=\frac{1}{2}}$	$\mathbf{z=1}$
$\mathbf{y=0}$	$\frac{v+c}{2}$	$v+c+\frac{1}{4}$	$v+c$
$\mathbf{y=\frac{1}{2}}$	$\frac{1}{4}$	$v+c$	$\frac{1}{4}$
$\mathbf{y=1}$	$v+c$	$v+c+\frac{1}{4}$	$\frac{v+c}{2}$

Table 3.1 Audience shares of station Y and Z in a hypothetical "no tune-in" PBE.

Now, suppose Y aired a tune-in when $(y,z)=(0,0)$. For the indifferent viewer from the first-period audience of station Y, the expected utility of switching to Z is $\frac{1}{3}(v-\lambda)+$

$\frac{1}{3}\left(v-\frac{1}{2}+\lambda\right)+\frac{1}{3}\left(v-c-\lambda\right)$, where the first term is the utility she would enjoy at Z when $z=0$, the second term is the utility she would enjoy at Z when $z=\frac{1}{2}$, and the third term is the utility she would enjoy at Y when $z=1$. This expression equals the utility of staying at Y, $v-\lambda$, for the viewer located at $\frac{1}{4}+\frac{c}{2}$, so the viewers with $\lambda\leq\frac{1}{4}+\frac{c}{2}$ do not switch to Z. Since this is a unilateral deviation, the behavior of the first-period viewers of station Z remains the same. The ones who switch to Z do not come back to Y since they would incur the sampling cost. So, station Y would gain an extra audience of $\left(\frac{1}{4}+\frac{c}{2}\right)-\left(\frac{v+c}{2}\right)=\frac{1-2v}{4}$ by airing a tune-in, and thus its second-period advertising revenue would go up by $ANp\left(\frac{1-2v}{4}\right)$. The cost of airing a tune-in is the revenue forgone in the first period from a single commercial, which is $Np\left(\frac{1}{2}\right)$. Thus, it is profitable to deviate as long as $\frac{1}{2}-v\geq\frac{1}{A}$.

It will prove useful to analyze in more detail the incentive for airing a tune-in when either station starts from a no tune-in situation. I will henceforth call this situation the "no tune-in" regime. Suppose the program locations are $(y,z)=\left(\frac{1}{2},\frac{1}{2}\right)$. If neither station airs a tune-in, they each receive an expected audience size of $v+c$ in the second period. From the previous analysis, if station Y airs a tune-in, viewers with locations $\lambda\in\left[\frac{1}{4}-\frac{c}{2},\frac{1}{2}\right]$ continue to stay with Y while the others, $\lambda<\frac{1}{4}-\frac{c}{2}$, switch to Z. The ones who switch to Z will stay there once they discover that $z=\frac{1}{2}$, since switching back to Y means missing the first few minutes of the program at Y. When this is a unilateral deviation by Y, behavior of the first-period viewers of Z does not change. A random half of them still switch to Y right after the first period ends. Among these viewers, the ones with locations $\lambda\in\left[\frac{1}{2},\frac{3}{4}+\frac{3c}{2}\right]$

stay at Y. The others go back to Z to check out the program there. Once they discover that $z = \frac{1}{2}$, those with locations at most $v + c$ apart from $\frac{1}{2}$ choose one of the stations at random as their final destination. Similarly, among those who stayed at Z at first, $\lambda \in \left[\frac{3}{4} + \frac{3c}{2}, 1\right]$ also sample Y and a random half of $\lambda \in \left[\frac{3}{4} + \frac{3c}{2}, \frac{1}{2} + v + c\right]$ stay at Y. Hence, station Y collects a total audience size of $\left(\frac{1}{4} + \frac{c}{2}\right) + \frac{1}{2}(v + c)$ as opposed to $v + c$. Since this is a symmetric game, station Z will have the same incentives. When both stations air a tune-in, they each get an audience of $v + c$, the same as before but with different composition. Thus, we get the following audience sizes when $(y, z) = \left(\frac{1}{2}, \frac{1}{2}\right)$,

	$\mathbf{q}_Z = \mathbf{0}$	$\mathbf{q}_Z = \mathbf{1}$
$\mathbf{q}_Y = \mathbf{0}$	$(v + c, v + c)$	$(\frac{3v}{2} - \frac{1}{4} + c, \frac{v}{2} + \frac{1}{4} + c)$
$\mathbf{q}_Y = \mathbf{1}$	$(\frac{v}{2} + \frac{1}{4} + c, \frac{3v}{2} - \frac{1}{4} + c)$	$(v + c, v + c)$

Table 3.2 Audience sizes of (Y, Z) in a "no tune-in" regime when $(y, z) = \left(\frac{1}{2}, \frac{1}{2}\right)$.

The unique Nash equilibrium in this game is $(q_Y, q_Z) = (1, 1)$ provided that $\frac{1}{2} - v \geq \frac{1}{A}$. This is true by Assumption 3.1 and thus the stations face a Prisoners' Dilemma situation. That is why a "no tune-in" regime cannot be maintained in an equilibrium.

Lemma 3.2 *By Assumptions 3.1 and 3.2, it must be true that* $q_Y(0, 0) = q_Y\left(\frac{1}{2}, \frac{1}{2}\right) = q_Z\left(\frac{1}{2}, \frac{1}{2}\right) = q_Z(1, 1) = 1$ *is satisfied in a symmetric PBE.*

Note that when $(y, z) = \left(0, \frac{1}{2}\right)$ or $\left(\frac{1}{2}, 0\right)$, deviating from a "no tune-in" regime is profitable for station Y if $A\left[\left(\frac{1}{4} + \frac{c}{2}\right) - \frac{1}{4}\right] \geq \frac{1}{2}$, or equivalently if $Ac \geq 1$. Although this condition is different from what was shown to be necessary for deviation when $(y,z) \in \left\{(0,0), \left(\frac{1}{2},\frac{1}{2}\right)\right\}$, it is in general more restrictive compared to $\frac{1}{2} - v \geq \frac{1}{A}$. Together with Assumption 3.1, $Ac \geq 1$ implies $\frac{1}{2}-v \geq \frac{1}{A}$. However, $\frac{1}{2}-v \geq \frac{1}{A}$ does not necessarily imply $Ac \geq 1$. An important thing to note is that viewers' optimal sampling behavior depends on their inferences from the observed tune-in decisions of the TV stations. As will be argued shortly, a strategy in which station Y airs a tune-in only when $(y,z) \in \left\{(0,0), \left(\frac{1}{2},\frac{1}{2}\right)\right\}$ cannot actually be part of a PBE. So, it is not nesessary to assume that $Ac \geq 1$ for this outcome to arise.

To see this, consider an equilibrium in which station Y airs a tune-in only when $(y,z) \in \left\{(0,0), \left(\frac{1}{2},\frac{1}{2}\right)\right\}$, and station Z does so only when $(y,z) \in \left\{\left(\frac{1}{2},\frac{1}{2}\right), (1,1)\right\}$. For these strategies to constitute a PBE, viewers' inferences from observed tune-in decisions must be correct. Therefore, when station Y advertises $y = 0$, the first-period viewers of Y infer that $z = 0$ as well. This means that each station ends up with an audience size of $\frac{v}{2}$ as each viewer will watch the first station they choose to sample. Both stations are actually worse off compared to the "no tune-in" regime. However, it is in fact optimal for station Y to switch back to the "no tune-in" regime if Z is going to air a tune when $(y,z) \in \left\{\left(\frac{1}{2},\frac{1}{2}\right), (1,1)\right\}$. If a λ-type viewer who is initially indifferent between the two stations continues to stay at Y after not seeing a tune-in, she will infer (incorrectly) that z is either $\frac{1}{2}$ or 1 upon seeing that $y = 0$. If $\lambda > \frac{1}{4} + c$, it is worth checking out station Z, too. But when she discovers that z is also 0, she will one more time be indifferent between the two stations provided that $\lambda \leq v + c$. Similarly, if she starts at Z and sees that $z = 0$, she will infer that y is either $\frac{1}{2}$ or 1, and will switch to Y if $\lambda > \frac{1}{4} + c$. The second-period audience size of station Y is thus $\frac{v+c}{2}$ as opposed to $\frac{v}{2}$, which means that station Y has an incentive to not air a tune-in when $(y, z) = (0, 0)$. But viewers anticipate this correctly and it was previously shown that this cannot be an equilibrium either.

One way to get around this problem is airing a tune-in more often. That is, when a TV station airs a tune-in for a particular program location, its viewers should not be able to infer the exact location of the other program. For station Y, the strategy of airing a tune-in for $y = 0$ only when $z = 0, 1$ (similarly, the strategy $q_Y\left(\frac{1}{2}, z\right) = 1$ only when $z = \frac{1}{2}, 1$) cannot happen in equilibrium. This is because station Y would also air a tune-in when $(y, z) = \left(0, \frac{1}{2}\right)$ so as to (incorrectly) signal to its viewers that z is either 0 or 1. So, no first-period viewer of Y would switch to Z, and thus Y would get an audience size of v. If instead Y did not air a tune-in – as would be anticipated by viewers in a PBE – all of its current viewers would switch to station Z and would infer that y is either 0 or 1 upon seeing that $z = \frac{1}{2}$. In this case, those viewers with $\lambda < \frac{1}{4} - c$ would switch back to station Y and stay there upon discovering $y = 0$. So, airing a tune-in is profitable when $v - \left(\frac{1}{4} - c\right) \geq \frac{1}{2A}$, which is true by Assumption 3.2. Note that this condition is satisfied even when the sampling cost is infinitesimally small.

This leaves us with two possible strategies. For station Y, these strategies are (i) air a tune-in unless y is 1, (ii) air a tune-in unless y or z is 1. Similarly, airing a tune-in unless $z = 0$ and unless z or $y = 0$ are the only two possible strategies for station Z. In what follows, I will refer to a strategy in which a station's tune-in decision does not depend on the program of the other station as a non-strategic behavior, and to an equilibrium that involves non-strategic behavior as a non-strategic equilibrium. Similarly, a tune-in strategy that depends on the program of the other station will be referred as a strategic behavior, and the corresponding equilibrium as a strategic equilibrium.

Suppose each station behaves strategically. How would viewers behave if this were a PBE? Since the two stations are identical in every aspect except for the locations of the first programs, the viewing behavior of people in the second period will be symmetric with respect to which station they watched in the first period. Therefore, I will only find the optimal sampling and final viewing decisions of the viewers who chose to watch Y in the first period. There are three cases; Y airs a tune-in for $y = 0$, Y airs a tune-in for $y = \frac{1}{2}$, and Y does not air a tune-in.

Case (1): *Y airs a tune-in for $y = 0$.*

In this case, the viewers of Y infer that $z \in \left\{0, \frac{1}{2}\right\}$. Those with locations closer to $\frac{1}{2}$ will have a tendency to switch to Z. Whatever the location of z turns out, none of these people would come back to Y. So, the solution is simple; $\lambda \leq \frac{1}{4}$ stay with Y, the others switch to Z. Those who switch to Z will have the sampling cost sunk, and therefore $\frac{1}{4} < \lambda \leq v + c$ will stay with Z when $z = 0$. The others just switch off in this case. If z turns out $\frac{1}{2}$, then all of them stay with Z.

Case (2): *Y airs a tune-in for $y = \frac{1}{2}$.*

In this case, the viewers of Y infer that $z \in \left\{0, \frac{1}{2}\right\}$. Those with locations closer to 0 will have a tendency to switch to Z. Similar with case (1), $\lambda \geq \frac{1}{4}$ stay with Y, the others switch to Z. Those who switch to Z will have the sampling cost sunk, and therefore $\frac{1}{2} - (v + c) < \lambda \leq \frac{1}{2}$ will stay with Z when $z = \frac{1}{2}$. If z turns out 0, then all of them stay with Z.

Case (3): Y *does not air a tune-in.*

The inference of viewers in this case is that Y did not air a tune-in because either $y = 1$ or $z = 1$ (or both). There are five possibilities:

$$(y, z) \in \left\{ (0, 1), \left(\frac{1}{2}, 1\right), (1, 0), \left(1, \frac{1}{2}\right), (1, 1) \right\}$$

So the posterior probability that $y = 0$ is same with the probability that $z = 0$, which is $\frac{1}{5}$. Similarly, $\Pr\left(y = \frac{1}{2}\right) = \Pr\left(z = \frac{1}{2}\right) = \frac{1}{5}$, and $\Pr(y = 1) = \Pr(z = 1) = \frac{3}{5}$. This means that viewers are indifferent between the two stations, and a random half will choose Z first. For those who stayed with Y, the actual location of y will determine their further behavior.

If $y = 0$, they infer that $z = 1$. So viewers with locations less than $v + c$ stay with Y, and the rest switch off. Note that for $v < \lambda \leq v + c$ switching off yields a disutility of c, so it is better to stay tuned.

If $y = \frac{1}{2}$, they infer that $z = 1$. So viewers with locations $\frac{1}{2} - (v + c) \leq \lambda \leq \frac{1}{2}$ stay with Y, and the rest switch off.

If $y = 1$, they infer that $z \in \left\{0, \frac{1}{2}, 1\right\}$, each with equal probability. If the viewers with locations $0 \leq \lambda < \frac{1}{2} - (v + c)$ choose to sample the program at station Z, they will only stay there when $z = 0$. So, the expected utility of sampling Z for a generic λ-

viewer from this interval, $E\left[U_Z^\lambda\right]$, given that station Y did not air a tune-in (i.e. given that $q_Y^{-1}(0) = (1, z)$) is,

$$E\left[U_Z^\lambda \mid q_Y^{-1}(0) = (1,z)\right] = \frac{1}{3}(v - c - \lambda) - \frac{2}{3}(2c)$$

Note that the highest utility a viewer may get in this case is $v - c$, since she started sampling with station Y and incurred the sampling cost. Viewers would stay tuned if the expected utility of sampling Z is not less than $-c$. Otherwise they turn their TVs off right after the first program ends. For $\lambda = \frac{1}{2} - (v + c)$, the expected utility of sampling is $\frac{1}{3}\left(2v - \frac{1}{2}\right) - \frac{2}{3}(2c)$. This is at least as great as $-c$ if $2v - \frac{1}{2} \geq c$, which is true by Assumption 3.1. Since $E\left[U_Z^\lambda \mid q_Y^{-1}(0) = (1,z)\right]$ is decreasing in λ, all of these viewers would choose to sample Z. Viewers with locations $\frac{1}{2} - (v + c) < \lambda \leq \frac{1}{4}$ would stay with Z unless $z = 1$. So, their expected utility is,

$$\begin{aligned} E\left[U_Z^\lambda \mid q_Y^{-1}(0) = (1,z)\right] &= \frac{1}{3}\left[(v - c - \lambda) + \left(v - c - \left(\frac{1}{2} - \lambda\right)\right) - (2c)\right] \\ &= \frac{1}{3}\left(2v - 4c - \frac{1}{2}\right) \end{aligned}$$

This expression is greater than or equal to $-c$ when $2v - \frac{1}{2} \geq c$, which is the same condition as before. Hence, it is satisfied for all $\lambda \in \left[\frac{1}{4}, \frac{1}{2}\right]$. The choices of viewers with locations on $\left[\frac{1}{4}, \frac{1}{2}\right]$ are just symmetric with those on $\left[0, \frac{1}{4}\right]$, so they all sample Z as well. If the program turns out to be located at 0 or $\frac{1}{2}$, station Z gets an audience size of $N(v + c)$. If $z = 1$, everyone switches off.

For those of $0 \leq \lambda \leq \frac{1}{2}$ who switched to Z initially, the subsequent choices are similar. Now, I need to check if sampling one of the stations is desirable at all, conditional on not seeing a tune-in. For $0 \leq \lambda < \frac{1}{2} - (v + c)$, the expected utility of starting sampling with station Y is,

$$E\left[U_Z^\lambda \mid q_Y = 0\right] = \frac{1}{5}(v - \lambda) + \frac{1}{5}(-c) + \frac{3}{5}E\left[U_Z^\lambda \mid q_Y^{-1}(0) = (1, z)\right]$$

Similarly, for $\frac{1}{2} - (v + c) \leq \lambda < \frac{1}{4}$, it is,

$$E\left[U_Z^\lambda \mid q_Y = 0\right] = \frac{1}{5}(v - \lambda) + \frac{1}{5}\left(v - \frac{1}{2} + \lambda\right) + \frac{3}{5}E\left[U_Z^\lambda \mid q_Y^{-1}(0) = (1, z)\right]$$

We need this value to be nonnegative for a viewer to sample Y. For $0 \leq \lambda < \frac{1}{2} - (v + c)$, $E\left[U_Z^\lambda \mid q_Y = 0\right] = \frac{2}{5}(v - c - \lambda) - \frac{2}{5}(2c)$. This is negative if λ is greater than $3c - v$. If $\frac{1}{2} - (v + c)$ is less than (or equal to) $3c - v$, then all of these people engage in sampling. $\frac{1}{2} - (v + c) \leq 3c - v$ if $c \geq \frac{1}{8}$. By Assumption 3.1, we must have $\frac{1}{4} + c < \frac{1}{2} - c$, which implies $c < \frac{1}{8}$. By monotonicity of $E\left[U_Z^\lambda \mid q_Y = 0\right]$ (increasing up to $\lambda = \frac{1}{4}$, and decreasing thereafter), we can conclude that sampling is desirable conditional on $q_Y = 0$.

We are now ready to calculate the audience share of a station. The table below gives the total fraction of the population choosing station Y to watch (after sampling, if any) in the second period for all possible program locations.

	$\mathbf{z=0}$	$\mathbf{z=\frac{1}{2}}$	$\mathbf{z=1}$
$\mathbf{y=0}$	$\frac{1}{4}$	$\frac{1}{4}$	$v+c$
$\mathbf{y=\frac{1}{2}}$	$v+c+\frac{1}{4}$	$v+c$	$v+c+\frac{1}{4}$
$\mathbf{y=1}$	$v+c$	$\frac{1}{4}$	$v+c-\frac{1}{4}$

Table 3.3 Audience share of Y in the strategic PBE.

Does station Y have any incentive to deviate? Suppose $(y, z) = (0, 0)$. If station Y deviates and does not air a tune-in, then a random half of its viewers stay with it while the other half switch. Those who stayed would think that $z = 1$ upon seeing that $y = 0$, and the ones with locations less than $v + c$ would continue staying. Those who initially switched to Z would think that $y = 1$ upon seeing $z = 0$, and therefore none of them would switch back to Y. So, station Y would end up with an audience share of $\frac{v+c}{2}$. It is profitable to deviate if

$$A\left(\frac{1}{4} - \frac{v+c}{2}\right) < \frac{1}{2}$$

where the left hand side is the marginal per-viewer revenue of a tune-in and the right hand side is the per-viewer cost of a tune-in. So, Y would not deviate if $v + c + \frac{1}{A} \leq \frac{1}{2}$. The same is true for $(y, z) = \left(0, \frac{1}{2}\right), \left(\frac{1}{2}, 0\right)$ and $\left(\frac{1}{2}, \frac{1}{2}\right)$. Note that deviation is not profitable when $y = 1$ since station Y can only communicate with its own viewers, and none of them would watch a program located at 1. It remains to analyze if it is profitable for Y to deviate when $(y, z) = (0, 1)$ or $\left(\frac{1}{2}, 1\right)$. In both cases, station Y is already getting the highest possible audience share from its first period without a tune-in. So, airing a tune-in cannot increase Y's audience size. Therefore, deviation is not profitable in these two cases, either.

Proposition 3.1 *The following constitutes a symmetric PBE if* $v + c + \frac{1}{A} \leq \frac{1}{2}$: Y *airs a tune-in unless* y *or* z *is* 1, Z *airs a tune-in unless* y *or* z *is* 0.

When $v + c + \frac{1}{A} > \frac{1}{2}$, people have no reason to expect the strategies in Proposition 3.1 to be played by the TV stations. This condition is satisfied when $v + c$ is large and/or the number of non-program breaks is small. Intuitively, a larger value of v is associated with a higher audience size since more viewers end up watching TV. A higher sampling cost means that if sampling occurs in the absence of a tune-in, a higher fraction of those who sample stay tuned. When the number of non-program breaks is small, the marginal benefit of promoting the upcoming program is lower. So, in all three cases, the incentive for passing up on airing a tune-in is higher.

From viewers' point of view, the ex-ante expected value of station Y's per-viewer profits is the weighted average of the profits in each of the possible nine cases. Per-viewer revenue in the first period is $\frac{pA}{2}$ in each case. Per-viewer revenue in the second period is the average of the audience shares given in the table for all of the nine cases, multiplied with pA. Since Y is expected to air a tune-in in four of the nine cases, its expected per-viewer costs are $\frac{4p}{9}$ times the audience share in the first period (which is $\frac{1}{2}$). So, the ex-ante expected per-viewer profits of station Y can be expressed as (the superscript S stands for strategic),

$$E\left[\Pi_j^S\right] = \left[\frac{A}{2} + \frac{(6(v+c)+1)A}{9} - \frac{2}{9}\right]p, \quad j = Y, Z$$

What happens when $v+c+\frac{1}{A} > \frac{1}{2}$? Based on the analysis so far, one possibility is when each station airs an additional tune-in relative to the PBE in Proposition 3.1. For station Y, this is when $q_Y(0,z) = 1$ for all z and $q_Y\left(\frac{1}{2}, z\right) = 1$ unless $z = 1$, or $q_Y\left(\frac{1}{2}, z\right) = 1$ for all z and $q_Y(0,z) = 1$ unless $z = 1$ (symmetric for Z). These are summarized in Table 3.4 below (for station Y only).

Strategy 1

	$\mathbf{z=0}$	$\mathbf{z=\frac{1}{2}}$	$\mathbf{z=1}$
$\mathbf{y=0}$	1	1	1
$\mathbf{y=\frac{1}{2}}$	1	1	0
$\mathbf{y=1}$	0	0	0

Strategy 2

	$\mathbf{z=0}$	$\mathbf{z=\frac{1}{2}}$	$\mathbf{z=1}$
$\mathbf{y=0}$	1	1	0
$\mathbf{y=\frac{1}{2}}$	1	1	1
$\mathbf{y=1}$	0	0	0

Table 3.4 The value of $q_Y(y,z)$ in two alternative regimes.

When Y does not air a tune-in, all of its viewers will switch to Z since it is highly likely that $y = 1$. If z turns out 0 or $\frac{1}{2}$, then they are certain that $y = 1$, and none of them come back to Y. When it turns out that $z = 1$, however, viewers will get confused. Station Y might have played the first or the second strategy. If it played the first strategy, y could be $\frac{1}{2}$ or 1 with equal chances. If it played the second one, on the other hand, y is 0 or 1. Without any further information, viewers just assume that two strategies are equally likely to be played, and therefore their inference will be $\Pr(y=0) = \Pr\left(y=\frac{1}{2}\right) = \frac{1}{4}$, $\Pr(y=1) = \frac{1}{2}$. But sampling Y would be optimal for all $\lambda \in \left[0, \frac{1}{2}\right]$ with these posteriors. This means that station Y could have done better by deviating, and reverting back to the strategy in Proposition (1). Since viewers anticipate this beforehand, we cannot have either of these strategies being played in a symmetric PBE.

A second, and the only other, possibility when $v + c + \frac{1}{A} > \frac{1}{2}$ is the non-strategic equilibrium in which the TV stations air a tune-in unless their second program is located at the farther end of the unit line compared to their first-period program. In this case, the priors of viewers who watch Y in the first period about z are unchanged regardless of the tune-in decision of Y.

Case (1): *Y airs a tune-in for $y = 0$.*

Since there is also the chance that $z = 1$, only the viewers with $\lambda > \frac{1}{4} + \frac{c}{2}$ sample Z. If z turns out to be located at 1, the ones with $\lambda \leq v$ come back to Y. If $z = 0$ or $\frac{1}{2}$, none of them come back.

Case (2): *Y airs a tune-in for $y = \frac{1}{2}$.*

This case is symmetric with Case (1).

Case (3): *Y does not air a tune-in.*

In this case, it is inferred that $y = 1$, and therefore none of the current viewers of Y will watch it.

The audience share of station Y for all possible program locations is summarized in Table 3.5.

	$\mathbf{z=0}$	$\mathbf{z=\frac{1}{2}}$	$\mathbf{z=1}$
$\mathbf{y=0}$	$\frac{1}{4}+\frac{c}{2}$	$\frac{1}{4}+\frac{c}{2}$	v
$\mathbf{y=\frac{1}{2}}$	$v+\frac{1}{4}+\frac{3c}{2}$	$v+c$	$v+\frac{1}{4}-\frac{c}{2}$
$\mathbf{y=1}$	$v+c$	$\frac{1}{4}-\frac{c}{2}$	$v-\frac{1}{4}+\frac{c}{2}$

Table 3.5 Audience share of station Y in the non-strategic PBE.

Note that deviation is not possible in this case, since none of Y's current viewers would keep watching or would come back later when $q_Y = 0$. So, the unique symmetric PBE when $v + c + \frac{1}{A} > \frac{1}{2}$ is the one in which the two stations play non-strategically.

Proposition 3.2 *When $v + c + \frac{1}{A} > \frac{1}{2}$, the unique symmetric PBE is the one in which Y airs a tune-in unless $y = 1$, and Z airs a tune-in unless $z = 0$.*

Arguing along the same lines as before, the ex-ante expected per-viewer profits of a station can be expressed as (the superscript NS stands for non-strategic),

$$E\left[\Pi_j^{NS}\right] = \left[\frac{A}{2} + \frac{(6v+4c+1)\,A}{9} - \frac{1}{3}\right] p, \quad j = Y, Z$$

Simple comparison yields that $E\left[\Pi_j^{S}\right]$ is always greater than $E\left[\Pi_j^{NS}\right]$. Even though it is on average more profitable to behave strategically, the existence of profitable deviations induces the TV stations to behave non-strategically. When $v + \frac{1}{A} > \frac{1}{2}$, even an infinitesimally small value of the sampling cost gives rise to the non-strategic equilibrium.

When $v + c + \frac{1}{A} \leq \frac{1}{2}$, both equilibria can be supported as PBEs. However, as long as viewers rationally expect the TV stations to play the less costly strategies, the non-strategic equilibrium can be ruled out. To be more precise, provided that $v + c + \frac{1}{A} \leq \frac{1}{2}$, it is always optimal to play strategically for the TV stations when the viewers expect them to do so. However, if the viewers are pessimistic in the sense that they only expect the worse when they do not see a tune-in, the unique PBE is the non-strategic one.

3.3.2 Social Value of Tune-ins

In this section, I analyze the effects of a possible ban on the use of tune-ins. I compare the expected social welfare under a hypothetical "no tune-in" regime with that under no restrictions. In the appendix to this chapter, I find the expected utility of a random viewer in all of the possible three situations: the strategic equilibrium (S), the non-strategic equilibrium (NS), and a "no tune-in" equilibrium (NT).

In a regime of no tune-ins, ex-ante expected per-viewer profit of a station in the second period is just the average of the audience shares given in Table 3.1, multiplied with the number of commercials and the per-viewer price. So, the total ex-ante expected per-viewer profits of a station are given by,

$$E\left[\Pi_j^{NT}\right] = A\left[\frac{1}{2} + \frac{6\left(v + c\right) + 1}{9}\right]p, \quad j = Y, Z$$

Let W denote the social welfare which is defined as the summation of station profits and viewer well-being. Then, the change in expected social welfare when the non-strategic PBE arises is expressed as,

$$\begin{aligned} E\left[W^{NS}-W^{NT}\right] &= N\left(E_{\lambda}\left[U_{\lambda}^{NS}-U_{\lambda}^{NT}\right]+2E\left[\Pi_{j}^{NS}-\Pi_{j}^{NT}\right]\right) \\ &= N\left(\left(\frac{15}{2}-c-6v\right)\frac{c}{9}-2\left(\frac{p}{3}+\frac{2cA}{9}\right)\right) \end{aligned}$$

Note that $\left(\frac{15}{2}-c-6v-4A\right)<0$ by Assumption 3.1 (even a much smaller value of A would imply the same result). So, $E\left[W^{NS}-W^{NT}\right]=N\left(\left(\frac{15}{2}-c-6v-4A\right)\frac{c}{9}-\frac{2p}{3}\right)<0$ for all parameter values. This means that it is welfare improving to ban the use of tune-ins when $v+c+\frac{1}{A}>\frac{1}{2}$, since non-strategic equilibrium is the unique symmetric PBE for these parameter values. Although viewers are obviously better off when there are tune-ins, it may be the case that lost revenues are too high, and therefore it is better to ban tune-ins. The primary reason for why the stations lose that much revenue is that fewer people watch TV when there are more tune-ins in general. In the absence of a ban, the “no tune-in” regime is not sustainable as an equilibrium because of unilateral deviations.

The same result does not carry over to the strategic equilibrium. The reason is that the expected audience size in the strategic equilibrium is equal with that in the “no tune-in” regime. Recalling that $E\left[\Pi_{j}^{S}\right]=\left[\frac{A}{2}+\frac{(6(v+c)+1)A}{9}-\frac{2}{9}\right]p$, the change in expected social welfare when the strategic PBE is the outcome is,

$$
\begin{aligned}
E\left[W^S - W^{NT}\right] &= N\left(E_\lambda\left[U_\lambda^S - U_\lambda^{NT}\right] + 2E\left[\Pi_j^S - \Pi_j^{NT}\right]\right) \\
&= N\left(\frac{c}{3} - \frac{4p}{9}\right)
\end{aligned}
$$

which is negative when $p > \frac{3c}{4}$.

Proposition 3.3 *If the outcome is the strategic PBE, it is welfare improving to ban tune-ins only when $p > \frac{3c}{4}$. If it is the non-strategic PBE, on the other hand, it is always welfare improving to ban tune-ins.*

It immediately follows from Proposition 3.3 that it may be welfare improving if the two stations collude and maximize total advertisement revenues. It is optimal to air no tune-ins in such a case, and as long as the conditions of Proposition 3.3 hold, this is better for the society as a whole.

Tune-ins are clearly beneficial for viewers. Without tune-ins, viewers would engage in too much sampling and some of them would end up watching TV although this yields a negative utility. If viewers had complete information about program attributes, TV stations would serve to a smaller audience size. However, information is incomplete and it is not feasible to inform everyone about TV programs. In a non-strategic equilibrium, TV stations are forced by market forces to air too many tune-ins. This is due to two factors. First, an equilibrium with no tune-ins is not feasible because of the oligopoly structure; without tune-ins, more people would switch away. Second, strategic equilibrium is not credible when c and/or $\frac{1}{A}$ is large. Had viewers believed that the strategic PBE would arise, neither station would have incentive to air any tune-ins at all. When c is large, a station can still capture an audience size that is high enough to make it worthwhile to deviate. When A is small, second-period revenue is not that large anyway, so deviation is again profitable. But, viewers are rational and perfectly anticipate these incentives beforehand. As a result, the stations are forced to air more tune-ins. The higher the number of tune-ins, the better choices people make, which implies a smaller audience size in the second period. So, the stations are double jeopardized when the strategic equilibrium cannot be attained; a higher number of tune-ins and fewer viewers on average. The revenue they lose in such a situation is larger than the increase in the well-being of viewers. Therefore, banning tune-ins is welfare enhancing.

The former one of the two factors above is also present in the strategic equilibrium. However, since c or $\frac{1}{A}$ is small enough, strategic equilibrium is credible. Therefore, the second factor does not arise. When the strategic equilibrium is attainable, stations do not end up airing too many tune-ins. Consumer surplus is now lower since viewers more often get stuck watching a program that is a bad match. When the per-viewer commercial price is low relative to the sampling cost, the decrease in the well-being of viewers is smaller than the increase in the revenues of the TV station due to fewer tune-ins, and therefore no intervention is necessary.[22]

[22] 88.5 million U.S. viewers watched the 2000 Super Bowl. The average price for a 30-second commercial was $2.1 million. So, the per-viewer price was approximately 2.4 cents. Although it is impossible to make an ordinal comparison of the per-viewer price and the sampling cost, common sense suggests that an average viewer loses more following an unsuccessful sampling.

3.4 Discussion and Conclusion

In this chapter, I have introduced a framework in which advertising decision of a firm provides information to people about the attributes of the product of the other firm in a horizontally differentiated duopoly market. I have chosen the TV industry for this analysis for several reasons. Most importantly, tune-ins are directly informative advertisements, they are special to the TV industry and they are exclusive (exclusive in the sense that a TV station cannot advertise to other stations' audiences).

I have analyzed the provision of tune-ins in a two-period model of TV broadcasting where programs are provided by two TV stations and are only horizontally differentiated. For simplicity, I assumed that the first programs at each TV station are known to viewers beforehand, and that the market is completely covered in the first period. The locations of the second programs are ex-ante unknown to people, although they know that the TV stations have this information. Therefore, viewers know that the tune-in decision of the TV station they watch in the first period may provide them with information about the location of the other station's program. In this context, I have characterized the symmetric perfect Bayesian equilibria in which tune-in decisions of the stations and the following inferences of viewers are in accordance.

The main aim of the analysis is to characterize the nature of strategic behavior when TV stations are privately informed about both programs. Therefore, I have restricted the parameter values so as to ensure that an equilibrium in which no tune-ins are aired does not exist. These restrictions simply require that the number of non-program breaks is not very small, and that the value of the sampling cost is small relative to the baseline utility people derive when they watch their ideal programs. The latter one implies that viewers always choose to sample a station if their priors for the program at that station are unchanged. However, it is also implied that the baseline utility is not very large so that all of the viewers do not watch the second program until the end. I believe that these assumptions are in line with empirical regularities and do not impose restrictions on the findings.

Existence of two symmetric PBEs has been shown. In the first one – referred to as the strategic PBE in the text – each station's tune-in decision depends on the location of its own as well as the location of its rival's program. A station chooses not to air a tune-in whenever at least one of the programs is such that no first-period viewers of that station would watch it. Not airing a tune-in when a station's own program is not appealing to its first-period viewers is optimal because a tune-in would not bring in any viewers. Similarly, it is not optimal to air a tune-in when a station knows that its first-period viewers will not like the program at the other station, and therefore anyone who may initially switch to the other station will come back.

The second possible PBE – referred to as the non-strategic PBE in the text – is the one in which each station's tune-in decision only depends on the location of its own program. A station chooses not to air a tune-in only when its program is such that none of its first-period viewers would watch it. This PBE is shown to exist as long as Assumption 3.1 is satisfied, and to be the unique equilibrium when either the sampling cost is relatively high or the TV stations have a small number of non-program breaks (or both). In the opposite case, both equilibria are valid. It is at first ambigous as to which one of these equilibria would be selected. If viewers are extremely pessimistic, they would expect the worst scenario upon not seeing a tune-in. This belief structure would eliminate the strategic PBE. However, there is no reason for viewers to be pessimistic in the current model. Since they perfectly know the incentives of the TV stations, they would most likely anticipate that the TV stations would choose the strategies resulting in higher revenues. From their point of view, the expected revenues are higher in the strategic PBE since the stations air fewer tune-ins on average compared to the non-strategic PBE. So, there is not much reason to believe that the stations would behave non-strategically. If, however, viewers were exogenously assumed to be pessimistic, we would have had the non-strategic PBE as the unique PBE.

In the strategic PBE, viewers update their priors about the location of the other station's program upon observing the tune-in decision of the station they have watched in the first period. So, the tune-in decision of the station they have watched in the first period serves as a signal for the program of the other station. There does not exist a fully separating equilibrium in which viewers are able to locate the other station's program with certainty only based on the tune-in decision of a TV station. If a station does not air a tune-in, it is inferred by that station's first-period audience that the upcoming program is probably not a good match for them. This deters them away from staying at the same station. However, they also infer that it could actually be the other station's program that is a bad match. In the strategic PBE, this additional inference exactly offsets the deterrence effect. As a result, viewers are indifferent between choosing either one of the two stations to start watching. After learning the location of the program they chose to sample first, they update their beliefs about the other program and then make their final decisions. At some instances, perfect revelation may occur after sampling one of the stations. Suppose a viewer watches station Y in the first period and station Y does not air a tune-in for its upcoming program. If this viewer continues to watch station Y and discovers after sampling that the program at Y is appealing to her, then she perfectly infers the location of the program at the other station. Had the TV stations behaved non-strategically, this inference would be impossible to reach.

The stations are better off on average in the strategic PBE compared to the non-strategic one. The same is not true for the viewers; unsuccessful sampling occurs less frequently in the strategic PBE since TV stations air fewer tune-ins. From a welfare perspective, it may be desirable to ban tune-ins. In the non-strategic PBE, total revenues of the stations are significantly reduced compared to the strategic PBE and this reduction overweights the increase in the aggregate viewer surplus. The real life is clearly not as simple. It is almost impossible to represent TV programs in a one-dimensional space. It is also not likely that viewers' preferences stay the same over time, and that viewers have nonrandom utilities.Nevertheless, this paper raises a question that has been untouched before, and constitutes a starting point for a more thorough analysis of the welfare properties of tune-ins in a duopoly TV industry.

3.5 Appendix

In this section, I find the utility of a random viewer under three specifications; the strategic PBE (S), the non-strategic PBE (NS), and the "no tune-in" regime (NT). I assume that the viewer located at $\frac{1}{2}$ watches station Y in the first period, and that whenever a viewer is indifferent between staying at a station or sampling the other one (or switching off), she chooses to stay (these assumptions do not change the results since there is a continuum of viewers). The findings in this appendix help me calculate the social welfare in Section 3.4. In each one of the nine possible program locations, average viewer derives a different level of utility since the TV stations, and in turn the viewers, behave differently in each one. Only the first five cases are analyzed in detail. The remaining four cases are symmetric with the first four ones.

Case (1): $(y, z) = (0, 0)$.

S: Station Y does, Z does not air a tune-in. Among those who watched Y in the first period, $\lambda \leq \frac{1}{4}$ stay with Y while the others switch to Z before the second period starts. After seeing that $z = 0$, $\lambda > v + c$ switch off. The ones who watched Z in the first period end up sampling both stations and eventually turn their TVs off. So,

$$U_\lambda^S = \begin{cases} v - \lambda & \text{,if } 0 \leq \lambda \leq v + c \\ -c & \text{,if } v + c < \lambda \leq \frac{1}{2} \\ -2c & \text{,if } \frac{1}{2} < \lambda \leq 1 \end{cases}$$

NS: Station Y does, Z does not air a tune-in. Among those who watched Y in the first period, $\lambda \leq \frac{1}{4} + \frac{c}{2}$ stay with Y after seeing a tune-in for $y = 0$ while the others switch to Z. The ones who watched Z in the first period only sample Y since they infer that $z = 0$. But they eventually turn their TVs off. So,

$$U_{\lambda}^{NS} = \begin{cases} v - \lambda & \text{,if } 0 \leq \lambda \leq v + c \\ -c & \text{,if } v + c < \lambda \leq 1 \end{cases}$$

NT: A random half of viewers start with Y and the other half with Z. Viewers with locations $\lambda \leq \frac{1}{4} + \frac{3c}{2}$ settle in the first station they sampled, thus incurring no sampling cost, while those with $\frac{1}{4} + \frac{3c}{2} \leq \lambda < v + c$ sample both stations and choose one at random. All others switch off after sampling both stations. So,

$$U_{\lambda}^{NT} = \begin{cases} v - \lambda & \text{,if } 0 \leq \lambda \leq \frac{1}{4} + \frac{3c}{2} \\ v - c - \lambda & \text{,if } \frac{1}{4} + \frac{3c}{2} < \lambda \leq v + c \\ -2c & \text{,if } v + c < \lambda \leq 1 \end{cases}$$

So, we have the following:

$$U_{\lambda}^{S} - U_{\lambda}^{NT} = \begin{cases} 0 & \text{,if } 0 \leq \lambda \leq \frac{1}{4} + \frac{3c}{2} \\ c & \text{,if } \frac{1}{4} + \frac{3c}{2} < \lambda \leq \frac{1}{2} \\ 0 & \text{,if } \frac{1}{2} < \lambda \leq 1 \end{cases}$$

$$U_\lambda^{NS} - U_\lambda^{NT} = \begin{cases} 0 & \text{,if } 0 \leq \lambda \leq \frac{1}{4} + \frac{3c}{2} \\ c & \text{,if } \frac{1}{4} + \frac{3c}{2} < \lambda \leq 1 \end{cases}$$

Integrating over λ, we get:

$$\begin{aligned} E_\lambda \left[U_\lambda^S - U_\lambda^{NT} \mid (y,z) = (0,0) \right] &= \left(\frac{1}{4} - \frac{3c}{2} \right) c \\ E_\lambda \left[U_\lambda^{NS} - U_\lambda^{NT} \mid (y,z) = (0,0) \right] &= \left(\frac{3}{4} - \frac{3c}{2} \right) c \end{aligned}$$

Case (2): $(y,z) = \left(0, \frac{1}{2}\right)$.

S: Station Y does, Z does not air a tune-in. Among those who watched Y in the first period, $\lambda \leq \frac{1}{4}$ stay with Y while the others switch to Z and stay there. Among those who watched Z in the first period, a random half stay with Z. After seeing that $z = \frac{1}{2}$, they infer that $y = 0$, so $\frac{1}{2} \leq \lambda \leq \frac{1}{2} + v + c$ stay and the others switch off. The other half start sampling with Y. After seeing that $y = 0$, they infer $z \in \left\{0, \frac{1}{2}, 1\right\}$, so all switch to Z. Those with $\frac{1}{2} \leq \lambda \leq \frac{1}{2} + v + c$ stay, the others switch off. So,

$$U_\lambda^S = \begin{cases} v - \lambda & \text{,if } 0 \leq \lambda \leq \frac{1}{4} \\ v - \left(\frac{1}{2} - \lambda\right) & \text{,if } \frac{1}{4} < \lambda \leq \frac{1}{2} \\ \frac{1}{2}\left(v - \lambda + \frac{1}{2}\right) + \frac{1}{2}\left(v - c - \lambda + \frac{1}{2}\right) & \text{,if } \frac{1}{2} < \lambda \leq \frac{1}{2} + v + c \\ \frac{1}{2}(-c) + \frac{1}{2}(-2c) & \text{,if } \frac{1}{2} + v + c < \lambda \leq 1 \end{cases}$$

NS: Both stations air a tune-in. $\lambda \leq \frac{1}{4} + \frac{c}{2}$ among those who watched Y in the first period stay with Y after seeing a tune-in for $y = 0$ while the others switch to Z and stay there.

Behavior of the ones who watched Z in the first period is similar. Those with $\lambda > \frac{3}{4} + \frac{c}{2}$ initially switch to Y in the hope of finding out $y = 1$. After discovering that $y = 0$, $\frac{3}{4} + \frac{c}{2} < \lambda \leq \frac{1}{2} + v$ come back to Z while the others turn their TVs off. So,

$$U_{\lambda}^{NS} = \begin{cases} v - \lambda & ,\text{if } 0 \leq \lambda \leq \frac{1}{4} + \frac{c}{2} \\ v - \left(\frac{1}{2} - \lambda\right) & ,\text{if } \frac{1}{4} + \frac{c}{2} < \lambda \leq \frac{1}{2} \\ v - \left(\lambda - \frac{1}{2}\right) & ,\text{if } \frac{1}{2} < \lambda \leq \frac{3}{4} + \frac{c}{2} \\ v - c - \left(\lambda - \frac{1}{2}\right) & ,\text{if } \frac{3}{4} + \frac{c}{2} < \lambda \leq \frac{1}{2} + v \\ -c & ,\text{if } \frac{1}{2} + v < \lambda \leq 1 \end{cases}$$

NT: A random half of viewers start with Y and the other half with Z. Viewers with locations $\frac{1}{4} - \frac{3c}{2} \leq \lambda \leq \frac{1}{4} + \frac{3c}{2}$ settle in the first station they sampled, thus incurring no sampling cost, while the others may end up sampling both stations. For $\lambda \leq \frac{1}{4} - \frac{3c}{2}$, if the viewer is lucky and started with Y, she stays there. If she started with Z, she also samples Y. Similarly, $\frac{1}{4} + \frac{3c}{2} \leq \lambda \leq \frac{3}{4} + \frac{3c}{2}$ end up at Z either immediately or after initially sampling Y. All others sample both stations and those with $\frac{3}{4}+\frac{3c}{2}\leq\lambda\leq\frac{1}{2}+v+c$ stay tuned. So,

$$U_{\lambda}^{NT}=\begin{cases}\frac{1}{2}\left(v-\lambda\right)+\frac{1}{2}\left(v-c-\lambda\right) & \text{,if } 0\leq\lambda<\frac{1}{4}-\frac{3c}{2}\\ \frac{1}{2}\left(v-\lambda\right)+\frac{1}{2}\left(v-\frac{1}{2}+\lambda\right) & \text{,if } \frac{1}{4}-\frac{3c}{2}\leq\lambda\leq\frac{1}{4}+\frac{3c}{2}\\ \frac{1}{2}\left(v-\frac{1}{2}+\lambda\right)+\frac{1}{2}\left(v-c-\frac{1}{2}+\lambda\right) & \text{,if } \frac{1}{4}+\frac{3c}{2}<\lambda\leq\frac{1}{2}\\ \frac{1}{2}\left(v-\lambda+\frac{1}{2}\right)+\frac{1}{2}\left(v-c-\lambda+\frac{1}{2}\right) & \text{,if } \frac{1}{2}<\lambda\leq\frac{3}{4}+\frac{3c}{2}\\ v-c-\left(\lambda-\frac{1}{2}\right) & \text{,if } \frac{3}{4}+\frac{3c}{2}<\lambda\leq\frac{1}{2}+v+c\\ -2c & \text{,if } \frac{1}{2}+v+c<\lambda\leq 1\end{cases}$$

So, we have the following:

$$U_{\lambda}^{S}-U_{\lambda}^{NT}=\begin{cases}\frac{c}{2} & \text{,if } 0\leq\lambda<\frac{1}{4}-\frac{3c}{2}\\ \frac{1}{4}-\lambda & \text{,if } \frac{1}{4}-\frac{3c}{2}\leq\lambda\leq\frac{1}{4}\\ \lambda-\frac{1}{4} & \text{,if } \frac{1}{4}<\lambda\leq\frac{1}{4}+\frac{3c}{2}\\ \frac{c}{2} & \text{,if } \frac{1}{4}+\frac{3c}{2}<\lambda\leq\frac{1}{2}\\ 0 & \text{,if } \frac{1}{2}<\lambda\leq\frac{3}{4}+\frac{3c}{2}\\ \frac{c}{2} & \text{,if } \frac{3}{4}+\frac{3c}{2}<\lambda\leq 1\end{cases}$$

$$U_\lambda^{NS} - U_\lambda^{NT} = \begin{cases} \frac{c}{2} & \text{,if } 0 \le \lambda < \frac{1}{4} - \frac{3c}{2} \\ \frac{1}{4} - \lambda & \text{,if } \frac{1}{4} - \frac{3c}{2} \le \lambda \le \frac{1}{4} + \frac{c}{2} \\ \lambda - \frac{1}{4} & \text{,if } \frac{1}{4} + \frac{c}{2} < \lambda \le \frac{1}{4} + \frac{3c}{2} \\ \frac{c}{2} & \text{,if } \frac{1}{4} + \frac{3c}{2} < \lambda \le \frac{3}{4} + \frac{c}{2} \\ -\frac{c}{2} & \text{,if } \frac{3}{4} + \frac{c}{2} < \lambda \le \frac{3}{4} + \frac{3c}{2} \\ 0 & \text{,if } \frac{3}{4} + \frac{3c}{2} < \lambda \le \frac{1}{2} + v \\ \left(\lambda - \frac{1}{2}\right) - v & \text{,if } \frac{1}{2} + v < \lambda \le \frac{1}{2} + v + c \\ c & \text{,if } \frac{1}{2} + v + c < \lambda \le 1 \end{cases}$$

Integrating over λ, we get:

$$\begin{aligned} E_\lambda\left[U_\lambda^{S} - U_\lambda^{NT} \mid (y,z) = \left(0, \frac{1}{2}\right)\right] &= \left(\frac{3}{4} - \frac{9c}{2}\right)\frac{c}{2} + \frac{9c^2}{2} \\ &= \left(\frac{3}{8} + \frac{9c}{4}\right)c \\ E_\lambda\left[U_\lambda^{NS} - U_\lambda^{NT} \mid (y,z) = \left(0, \frac{1}{2}\right)\right] &= \left(\frac{3}{4} - \frac{5c}{2}\right)\frac{c}{2} + 4c^2 + \left(\frac{1}{2} - v\right)c \\ &= \left(\frac{7}{8} + \frac{9c}{4} - v\right)c \end{aligned}$$

Case (3): $(y,z) = (0,1)$.

S: Neither station airs a tune-in. Among those who watched Y in the first period, a random half stay with Y and infer that $z = 1$ after seeing $y = 0$. So, $0 \le \lambda \le v + c$ stay and the others switch off. The other half initially switch to Z. All of these viewers also sample Y after discovering that $z = 1$ and $0 \le \lambda \le v + c$ stay. Behavior of the viewers who watched

Z in the first period is just symmetric. So,

$$U_\lambda^S = \begin{cases} \frac{1}{2}(v-\lambda) + \frac{1}{2}(v-c-\lambda) & \text{,if } 0 \le \lambda \le v+c \\ \frac{1}{2}(-c) + \frac{1}{2}(-2c) & \text{,if } v+c < \lambda < 1-(v+c) \\ \frac{1}{2}(v-1+\lambda) + \frac{1}{2}(v-c-1+\lambda) & \text{,if } 1-(v+c) \le \lambda \le 1 \end{cases}$$

NS: Both stations air a tune-in. So, $\lambda \le \frac{1}{4} + \frac{c}{2}$ continue to stay with Y while $\frac{1}{4} + \frac{c}{2} < \lambda \le v$ come back to Y after initially sampling Z. Behavior of the viewers who watched Z in the first period is just symmetric. So,

$$U_\lambda^{NS} = \begin{cases} v-\lambda & \text{,if } 0 \le \lambda \le \frac{1}{4} + \frac{c}{2} \\ v-c-\lambda & \text{,if } \frac{1}{4} + \frac{c}{2} < \lambda \le v \\ -c & \text{,if } v < \lambda < 1-v \\ v-c-(1-\lambda) & \text{,if } 1-v \le \lambda < \frac{3}{4} - \frac{c}{2} \\ v-(1-\lambda) & \text{,if } \frac{3}{4} - \frac{c}{2} \le \lambda \le 1 \end{cases}$$

NT: The viewing choices here are similar with the previous case.

$$U_\lambda^{NT} = \begin{cases} \frac{1}{2}(v-\lambda) + \frac{1}{2}(v-c-\lambda) & \text{,if } 0 \le \lambda \le \frac{1}{4} + \frac{3c}{2} \\ v-c-\lambda & \text{,if } \frac{1}{4} + \frac{3c}{2} < \lambda \le v+c \\ -2c & \text{,if } v+c < \lambda < 1-(v+c) \\ v-c-(1-\lambda) & \text{,if } 1-(v+c) \le \lambda < \frac{3}{4} - \frac{3c}{2} \\ \frac{1}{2}(v-1+\lambda) + \frac{1}{2}(v-c-1+\lambda) & \text{,if } \frac{3}{4} - \frac{3c}{2} \le \lambda \le 1 \end{cases}$$

So, we have the following:

$$U_\lambda^S - U_\lambda^{NT} = \begin{cases} 0 & \text{,if } 0 \le \lambda \le \frac{1}{4} + \frac{3c}{2} \\ \frac{c}{2} & \text{,if } \frac{1}{4} + \frac{3c}{2} \le \lambda < \frac{3}{4} - \frac{3c}{2} \\ 0 & \text{,if } \frac{3}{4} - \frac{3c}{2} \le \lambda \le 1 \end{cases}$$

$$U_\lambda^{NS} - U_\lambda^{NT} = \begin{cases} \frac{c}{2} & \text{,if } 0 \le \lambda \le \frac{1}{4} + \frac{c}{2} \\ -\frac{c}{2} & \text{,if } \frac{1}{4} + \frac{c}{2} < \lambda \le \frac{1}{4} + \frac{3c}{2} \\ 0 & \text{,if } \frac{1}{4} + \frac{3c}{2} < \lambda \le v \\ \lambda - v & \text{,if } v < \lambda \le v + c \\ c & \text{,if } v + c < \lambda < 1 - (v + c) \\ 1 - \lambda - v & \text{,if } 1 - (v + c) < \lambda < 1 - v \\ 0 & \text{,if } 1 - v < \lambda < \frac{3}{4} - \frac{3c}{2} \\ -\frac{c}{2} & \text{,if } \frac{3}{4} - \frac{3c}{2} \le \lambda < \frac{3}{4} - \frac{c}{2} \\ \frac{c}{2} & \text{,if } \frac{3}{4} - \frac{c}{2} \le \lambda \le 1 \end{cases}$$

Integrating over λ, we get:

$$\begin{aligned} E_\lambda \left[U_\lambda^S - U_\lambda^{NT} \mid (y, z) = (0, 1) \right] &= \left(\frac{1}{4} - \frac{3c}{2} \right) c \\ E_\lambda \left[U_\lambda^{NS} - U_\lambda^{NT} \mid (y, z) = (0, 1) \right] &= \left(\frac{5}{4} - \frac{3c}{2} - 2v \right) c \end{aligned}$$

Case (4): $(y, z) = \left(\frac{1}{2}, 0 \right)$.

S: Station Y does, Z does not air a tune-in. Among those who watched Y in the first period, $\lambda \geq \frac{1}{4}$ stay with Y while the others initially switch to Z and stay there after seeing that $z = 0$. Among those who watched Z in the first period, the random half that started sampling with Y are lucky as they infer that $z = 0$. So, those with $\frac{1}{2} \leq \lambda \leq \frac{1}{2} + v + c$ stay, the others turn their TVs off. The other half sample both stations and those with $\frac{1}{2} \leq \lambda \leq \frac{1}{2} + v + c$ end up watching Y. So,

$$U_\lambda^S = \begin{cases} v - \lambda & \text{,if } 0 \leq \lambda < \frac{1}{4} \\ v - \left(\frac{1}{2} - \lambda\right) & \text{,if } \frac{1}{4} \leq \lambda \leq \frac{1}{2} \\ \frac{1}{2}\left(v - \lambda + \frac{1}{2}\right) + \frac{1}{2}\left(v - c - \lambda + \frac{1}{2}\right) & \text{,if } \frac{1}{2} < \lambda \leq \frac{1}{2} + v + c \\ \frac{1}{2}\left(-c\right) + \frac{1}{2}\left(-2c\right) & \text{,if } \frac{1}{2} + v + c < \lambda \leq 1 \end{cases}$$

NS: Station Y does, Z does not air a tune-in. Among those who watched Y in the first period, $\lambda \geq \frac{1}{4} - \frac{c}{2}$ stay with Y while the others initially switch to Z and stay there after seeing that $z = 0$. The viewers who watched Z in the first period switch to Y and those with $\frac{1}{2} \leq \lambda \leq \frac{1}{2} + v + c$ stay there. So,

$$U_\lambda^{NS} = \begin{cases} v - \lambda & \text{,if } 0 \leq \lambda < \frac{1}{4} - \frac{c}{2} \\ v - \left(\frac{1}{2} - \lambda\right) & \text{,if } \frac{1}{4} - \frac{c}{2} \leq \lambda \leq \frac{1}{2} \\ v - \left(\lambda - \frac{1}{2}\right) & \text{,if } \frac{1}{2} < \lambda \leq \frac{1}{2} + v + c \\ -c & \text{,if } \frac{1}{2} + v + c < \lambda \leq 1 \end{cases}$$

NT: Same with Case (2). So,

$$U_\lambda^{NT} = \begin{cases} \frac{1}{2}(v-\lambda) + \frac{1}{2}(v-c-\lambda) & \text{,if } 0 \le \lambda < \frac{1}{4} - \frac{3c}{2} \\ \frac{1}{2}(v-\lambda) + \frac{1}{2}\left(v - \frac{1}{2} + \lambda\right) & \text{,if } \frac{1}{4} - \frac{3c}{2} \le \lambda \le \frac{1}{4} + \frac{3c}{2} \\ \frac{1}{2}\left(v - \frac{1}{2} + \lambda\right) + \frac{1}{2}\left(v - c - \frac{1}{2} + \lambda\right) & \text{,if } \frac{1}{4} + \frac{3c}{2} < \lambda \le \frac{1}{2} \\ \frac{1}{2}\left(v - \lambda + \frac{1}{2}\right) + \frac{1}{2}\left(v - c - \lambda + \frac{1}{2}\right) & \text{,if } \frac{1}{2} < \lambda \le \frac{3}{4} + \frac{3c}{2} \\ v - c - \left(\lambda - \frac{1}{2}\right) & \text{,if } \frac{3}{4} + \frac{3c}{2} < \lambda \le \frac{1}{2} + v + c \\ -2c & \text{,if } \frac{1}{2} + v + c < \lambda \le 1 \end{cases}$$

So, we have the following:

$$U_\lambda^{S} - U_\lambda^{NT} = \begin{cases} \frac{c}{2} & \text{,if } 0 \le \lambda < \frac{1}{4} - \frac{3c}{2} \\ \frac{1}{4} - \lambda & \text{,if } \frac{1}{4} - \frac{3c}{2} \le \lambda < \frac{1}{4} \\ \lambda - \frac{1}{4} & \text{,if } \frac{1}{4} \le \lambda \le \frac{1}{4} + \frac{3c}{2} \\ \frac{c}{2} & \text{,if } \frac{1}{4} + \frac{3c}{2} < \lambda \le \frac{1}{2} \\ 0 & \text{,if } \frac{1}{2} < \lambda \le \frac{3}{4} + \frac{3c}{2} \\ \frac{c}{2} & \text{,if } \frac{3}{4} + \frac{3c}{2} < \lambda \le 1 \end{cases}$$

$$U_\lambda^{NS} - U_\lambda^{NT} = \begin{cases} \frac{c}{2} & \text{,if } 0 \le \lambda < \frac{1}{4} - \frac{3c}{2} \\ \frac{1}{4} - \lambda & \text{,if } \frac{1}{4} - \frac{3c}{2} \le \lambda < \frac{1}{4} - \frac{c}{2} \\ \lambda - \frac{1}{4} & \text{,if } \frac{1}{4} - \frac{c}{2} \le \lambda < \frac{1}{4} + \frac{3c}{2} \\ \frac{c}{2} & \text{,if } \frac{1}{4} + \frac{3c}{2} < \lambda \le \frac{3}{4} + \frac{3c}{2} \\ c & \text{,if } \frac{3}{4} + \frac{3c}{2} < \lambda \le 1 \end{cases}$$

Integrating over λ, we get:

$$\begin{aligned} E_\lambda\left[U_\lambda^S - U_\lambda^{NT} \mid (y,z) = \left(\frac{1}{2},0\right)\right] &= \left(\frac{3}{4} - \frac{9c}{2}\right)\frac{c}{2} + \frac{9c^2}{2} \\ &= \left(\frac{3}{8} + \frac{9c}{4}\right)c \\ E_\lambda\left[U_\lambda^{NS} - U_\lambda^{NT} \mid (y,z) = \left(\frac{1}{2},0\right)\right] &= \left(\frac{5}{4} - \frac{9c}{2}\right)\frac{c}{2} + 4c^2 \\ &= \left(\frac{5}{8} + \frac{7c}{4}\right)c \end{aligned}$$

Case (5): $(y,z) = \left(\frac{1}{2}, \frac{1}{2}\right)$.

S: Both stations air a tune-in. $\frac{1}{4} \leq \lambda \leq \frac{3}{4}$ stay with the stations they watched in the in the first period. The others initially switch to the other station; those with $\lambda < \frac{1}{2} - v - c$ and $\lambda > \frac{1}{2} + v + c$ switch off while the others stay. So,

$$U_\lambda^S = \begin{cases} -c & \text{,if } 0 \leq \lambda < \frac{1}{2} - v - c \\ v - \left(\frac{1}{2} - \lambda\right) & \text{,if } \frac{1}{2} - v - c \leq \lambda \leq \frac{1}{2} \\ v - \left(\lambda - \frac{1}{2}\right) & \text{,if } \frac{1}{2} < \lambda \leq \frac{1}{2} + v + c \\ -c & \text{,if } \frac{1}{2} + v + c < \lambda \leq 1 \end{cases}$$

NS: Both stations air a tune-in. $\frac{1}{4} - \frac{c}{2} \leq \lambda \leq \frac{3}{4} + \frac{c}{2}$ stay with the stations they watched in the in the first period. The others initially switch to the other station and stay there except for $\lambda < \frac{1}{2} - v - c$ and $\lambda > \frac{1}{2} + v + c$. So, U_λ^{NS} remains the same as above.

NT: Those with $\frac{1}{4} - \frac{3c}{2} \leq \lambda \leq \frac{3}{4} + \frac{3c}{2}$ stay with the stations they sample first. Others sample both stations and those with $\frac{1}{2} - v - c \leq \lambda \leq \frac{1}{4} - \frac{3c}{2}$ and $\frac{3}{4} + \frac{3c}{2} \leq \lambda \leq \frac{1}{2} + v + c$ choose to watch one of them at random. The others turn their TVs off. So,

$$U_\lambda^{NT} = \begin{cases} -2c & \text{,if } 0 \le \lambda < \frac{1}{2} - v - c \\ v - c - \left(\frac{1}{2} - \lambda\right) & \text{,if } \frac{1}{2} - v - c \le \lambda < \frac{1}{4} - \frac{3c}{2} \\ v - \left(\frac{1}{2} - \lambda\right) & \text{,if } \frac{1}{4} - \frac{3c}{2} \le \lambda \le \frac{1}{2} \\ v - \left(\lambda - \frac{1}{2}\right) & \text{,if } \frac{1}{2} < \lambda \le \frac{3}{4} + \frac{3c}{2} \\ v - c - \left(\lambda - \frac{1}{2}\right) & \text{,if } \frac{3}{4} + \frac{3c}{2} < \lambda \le \frac{1}{2} + v + c \\ -2c & \text{,if } \frac{1}{2} + v + c < \lambda \le 1 \end{cases}$$

So, we have the following:

$$U_\lambda^{S} - U_\lambda^{NT} = U_\lambda^{NS} - U_\lambda^{NT} = \begin{cases} c & \text{,if } 0 \le \lambda < \frac{1}{4} - \frac{3c}{2} \\ 0 & \text{,if } \frac{1}{4} - \frac{3c}{2} \le \lambda \le \frac{3}{4} + \frac{3c}{2} \\ c & \text{,if } \frac{3}{4} + \frac{3c}{2} < \lambda \le 1 \end{cases}$$

Integrating over λ, we get:

$$\begin{aligned} E_\lambda\left[U_\lambda^{S} - U_\lambda^{NT} \mid (y,z) = \left(\frac{1}{2}, \frac{1}{2}\right)\right] &= E_\lambda\left[U_\lambda^{NS} - U_\lambda^{NT} \mid (y,z) = \left(\frac{1}{2}, \frac{1}{2}\right)\right] \\ &= \left(\frac{1}{2} - 3c\right) c \end{aligned}$$

The remaining four cases are symmetric with the first four cases, and therefore are omitted.

Finally, integrating over (y, z), we get the following:

$$\begin{aligned} E_\lambda \left[U_\lambda^S - U_\lambda^{NT}\right] &= \frac{1}{9}(3c) \\ E_\lambda \left[U_\lambda^{NS} - U_\lambda^{NT}\right] &= \frac{1}{9}\left(\frac{15}{2} - 6v - c\right)c \end{aligned}$$

These are expressions I use in welfare comparison in Section 3.3.2.

Chapter 4

Ignorant Consumers and Informative Advertising

4.1 Introduction

As discussed in Chapter 1, Grossman and Shapiro (1984) (henceforth, GS) find overprovision of advertising relative to the social optimum when the number of firms is sufficiently large and the products are heterogeneous. This finding rests on two crucial assumptions. First, GS use the "large-group" assumption of Dixit and Stiglitz (1977) which implies that the industry produces an exogenous amount of aggregate output. This assumption is particularly strong since a main role of informative advertising is expanding the market demand by informing consumers of product existence and matching them with products that better suit their tastes. Second, they treat consumers as passive in the sense that they do not engage in search or any other information acquisition activities. They argue that this may be the case when consumers remain unaware about the existence of a product unless they are reached by its ad, or when the search costs are sufficiently high. Thus, the only brands consumers are informed about are the ones whose ads reach them.

In this chapter, I analyze the implications of incorporating active consumer search into GS's model. When consumers are ex-ante unaware of product existence, they naturally do not consider search until they receive an ad. The first ad they receive raises awareness about the existence of a market for the advertised product and puts them in a position in which they may want to gather more information by searching. At this point, however, a technical problem arises: what beliefs are they going to base their optimal search behavior on? They do not know if there are any other firms in the market or if this is a differentiated good. So, one needs to define some sort of a belief formation mechanism to address these issues.

I take a different approach in this chapter and assume that search is costless. The above complications do not arise under this assumption. I continue to assume that there is a single good which can be sold in different varieties (i.e. brands) by different firms. Consumers are initially unaware of the existence of such a good, and hence of the market structure. Firms randomly send their ads in order to inform consumers about the existence, the characteristics and the price of the brand they sell. If the consumers are fully rational, exposure to a single ad enables them to obtain information about the whole market since it raises awareness about the existence of such a good. I then introduce "ignorant" consumers who ignores the possibility that other firms exist when contacted with only one brand's ad containing a price that only a monopolist would choose. Thus, by advertising the monopoly price, a firm may be able to convince a subset of consumers that there are no other varieties of this good. If the ad contains a price that would a monopolist would never choose, on the other hand, these consumers infer that the market is not served by a single firm. Therefore, they conduct a search and find out all the relevant market information.

I next characterize the price-advertising equilibria under both consumer specifications and investigate how market level of advertising compares to the socially optimal amount.

For simplicity, I assume that there are only two brands in the market offered by two firms. When consumers are rational, firms charge a lower price in a symmetric equilibrium compared to its level in GS, and advertising is always undersupplied relative to the socially optimal amount. This is mainly because of an apparent free-riding problem: if my rival is going to choose a price different from the monopoly one, then I do not need to advertise to those consumers who my rival advertises to! When consumers are ignorant, on the other hand, I find for reasonable parameter values that there exists a symmetric Nash equilibrium in which firms set their prices at the monopoly level and possibly undersupply advertising relative to the social optimum. These findings are in sharp contrast with those of GS.

As described in Chapter 1, there is a vast literature on informative advertising. One of the central questions addressed in this literature is whether the market produces too much or too little advertising compared to the socially optimal amount. The market outcome may differ from the socially optimal allocation because of three distinct effects advertising creates. First, advertising can increase the total market size by drawing in new consumers. Second, informative advertising improves the match between consumers and brands. In both cases, the additional surplus is typically divided between sellers and buyers. However, firms do not consider the benefits accruing to consumers when choosing their advertising levels. Therefore, advertising tends to be undersupplied in the market equilibrium. And third, firms advertise in order to increase their own profits. Therefore, they do not take into account the reduction in the profits of the rival firms because of their ads. This causes advertising to be oversupplied in the market equilibrium.

When products are homogeneous, Butters (1977) finds that the market produces the socially optimal amount of advertising. Stegeman (1991) extends Butters' analysis to include heterogeneous valuations and finds that advertising is undersupplied. Stahl (1994) also finds that it is undersupplied when consumers have downward sloping demands. When products are differentiated, GS find that firms oversupply informative advertising irrespective of the degree of differentiation. Contrary to this result, two recent papers, Christou and Vettas (2003) and Hamilton (2004), find that it actually depends on the degree of product differentiation. Christou and Vettas analyze informative advertising in a random-utility model of product differentiation with possible non-localized competition.[23] Hamilton extends GS by allowing for incomplete coverage and considering quantity competition as well.[24] Both papers find that advertising is oversupplied only when products are sufficiently differentiated.[25]

4.2 The Model

The market consists of two firms who sell different varieties (i.e., brands) of the same good to a group of consumers. The product space is a circle of unit circumference in which firms are located $\frac{1}{2}$ unit apart from each other. I assume, for simplicity, that production is costless. Consumers are initially unaware of the existence of the good, and hence of the

[23]GS focus on local deviations around a symmetric equilibrium, thus ignoring the possibility that a deviation to a high price is not profitable.

[24]GS consider only the case of full market coverage in which all partially and fully informed consumers make a purchase.

[25]Hamilton (2004) also finds that advertising is always undersupplied when the two products are homogeneous.

firms. They stay unaware about existence until they are reached by advertising. The role of advertising is to convey information about a firm's product to consumers. Following Butters (1977), firms send independent advertising messages and have no ability to target ads towards particular consumers.

Ads are truthful and contain information about the price and true location of individual varieties. Exposure to an ad makes a consumer learn the existence of such a good and realize her ideal variety. Each consumer is identified by a point on the circle that corresponds to her ideal variety. I assume that consumers are distributed uniformly around the circle with a density of 1. Each consumer purchases at most one unit and receives a gross value of v from consuming her ideal variety. A consumer whose ideal variety differs from a firm's variety by a distance of d units obtains a net benefit of $v - td - p$, where p is the price the consumer pays to purchase the product. Consumers rely on the information received from ads to locate specific brands in the product space.

Consumer search is costless. Since consumers are initially assumed to be completely unaware of the existence of the good, they do not engage in search in the absence of an ad. In subsection 4.2.1, I consider the case in which any consumer who is reached by at least one ad conducts a search, and hence obtains full information about the market. In the next subsection, I consider the case in which a consumer does not engage in search if she is reached by only one firm's ad and the advertised price is that a monopolist would choose. So, she remains unaware of the existence of the other firm when the only advertised price is at the monopoly level. A different price, on the other hand, prompts a consumer to speculate the existence of other firms. Since search is costless, she conducts a search and finds out the location and price information of the other firm. Consumers only purchase the good if they are aware of a brand that offers positive net surplus, and a consumer who is aware of both brands selects the one that offers the highest net surplus.

Let ϕ_i denote the advertising intensity of firm i, which measures the reach of the advertising campaign. So, ϕ_i is interpreted as the fraction of consumers exposed to the ad of firm i at least for once. This divides consumers into four types. With probability $\phi_i \phi_j$ a consumer simultaneously receives ads from both firm i and firm j ($j \neq i$); with probability $\phi_i \left(1 - \phi_j\right)$ a consumer only receives an ad from firm i; with probability $(1 - \phi_i)\, \phi_j$ a consumer only receives an ad from firm j; and with probability $(1 - \phi_i) \left(1 - \phi_j\right)$ a consumer does not receive any ads from either firm. I assume that the cost of achieving a reach of ϕ by an advertising campaign takes a simple form. It is given by $A(\phi) = a\phi^2$, where a is a positive constant.

It is constructive to find the monopoly equilibrium first. An informed consumer finds it worthwhile to purchase the advertised brand if consuming it yields a nonnegative consumer surplus. This requires that $p \leq v - td$ for a consumer located d units away from the location of the brand. Note that $d = \frac{1}{2}$ is the highest possible distance from an advertised brand. When $v < t$, the monopolist would not find it optimal to set a price that is low enough so as to make each recipient of an ad purchase the good. Given that the monopolist chooses an ad reach of ϕ, its total demand is given by

$$D = \begin{cases} \phi \left[2\dfrac{v-p}{t} \right] & \text{, if } p > v - \dfrac{t}{2} \\ \phi & \text{, if } p \leq v - \dfrac{t}{2} \end{cases}$$

The price that maximizes $\pi = \phi \left(\frac{2(v-p)}{t} \right) p - A(\phi)$ is easily found to be $\frac{v}{2}$. Note that the profit function is concave in price when $p > v - \frac{t}{2}$, so the second-order condition is satisfied. For all informed consumers to be served in the monopoly equilibrium, we need that $v \geq t$. When $v \geq t$, the monopolist can increase its price to $v - \dfrac{t}{2}$ without losing any (informed) consumers. Given that $p = v - \frac{t}{2}$, the profits of the monopolist can be expressed as

$$\pi = \phi \left(v - \frac{t}{2} \right) - a\phi^2$$

The profit-maximizing advertising intensity is then $\frac{1}{2a} \left(v - \frac{t}{2} \right)$.

I examine only the case in which the monopoly price is sufficiently low so that an informed consumer located at the most distant location from the advertised brand finds it worthwhile to purchase it. In this sense, the market is completely covered.

Assumption 4.1 $v > t$.

Strict inequality also ensures that consumers can distinguish the monopoly price. This will be necessary for the analysis in the following subsection. I also assume that it is too costly for a monopolist to reach everyone by advertising.

Assumption 4.2 $v - \dfrac{t}{2} < 2a$.

This assumption also ensures that the optimal advertising intensities in the following subsections are less than one.

4.2.1 Rational Consumers

When consumers are fully rational, a direct implication of costless search is that all partially informed consumers conduct a search. As a result, all consumers reached by at least one ad obtain full information regarding the market. Regardless of the exact locations of the two brands, the total demand that accrues to firm i in this case is $\left(1-\left(1-\phi_i\right)\left(1-\phi_j\right)\right)\left(\frac{1}{2}+\frac{p_j-p_i}{t}\right)$, where $\left(1-\phi_i\right)\left(1-\phi_j\right)$ is the fraction of consumers who are not reached by the ads of either firm. Each firm maximizes its profits with respect to price and the advertising intensity. The profit function of firm i is

$$\pi_i=\left(1-\left(1-\phi_i\right)\left(1-\phi_j\right)\right)\left(\frac{1}{2}+\frac{p_j-p_i}{t}\right)p_i-a\phi_i^2$$

The first-order conditions with respect to p_i and ϕ_i are

$$\frac{\partial\pi_i}{\partial p_i}=\left(1-\left(1-\phi_i\right)\left(1-\phi_j\right)\right)\left(\frac{1}{2}+\frac{p_j-2p_i}{t}\right)=0$$

$$\frac{\partial\pi_i}{\partial\phi_i}=\left(1-\phi_j\right)\left(\frac{1}{2}+\frac{p_j-p_i}{t}\right)p_i-2a\phi_i=0$$

The second-order conditions are easily seen to be satisfied. At a symmetric Nash equilibrium, both firms choose the same price and the same advertising intensity. Let these be denoted by p^R and ϕ^R, respectively, where the superscript R refers to rational. Inserting the symmetry condition into the first-order conditions, we get

$$\left(1-\left(1-\phi^R\right)^2\right)\left(\frac{1}{2}-\frac{p^R}{t}\right)=0 \tag{4.1}$$

$$\left(1-\phi^R\right)\frac{p^R}{2}-2a\phi^R=0 \tag{4.2}$$

Equation (4.1) implies that $\phi^R>0$, so $\left(1-\left(1-\phi^R\right)^2\right)\neq 0$. Hence, the equilibrium price from equation (4.1) is

$$p^R=\frac{t}{2} \tag{4.3}$$

Inserting this into equation (4.2), we get

$$\phi^R=\frac{t}{8a+t} \tag{(4.4)}$$

The profits of each firm are π^R=$\left(1-\left(1-\phi^R\right)^2\right)\frac{p^R}{2}-a\left(\phi^R\right)^2$. After rearranging, we get

$$\pi^R=\left(3a+\frac{p^R}{2}\right)\left(\phi^R\right)^2 \tag{4.5}$$

4.2.2 Ignorant Consumers

In this subsection, I find the duopoly equilibrium when partially informed consumers do not engage in search if the advertised price is at the monopoly level. A firm has two options in this specification. The first option is to choose the monopoly price, in which case the consumers who only receive that firm's ads would think there are no other firms, and thus would purchase the brand offered by that firm. The second option is to choose a different price, in which case the ad-recipients would understand that there is at least one more firm offering the same good, possibly in a different variety.

Suppose firm 2 sets $p_2 = v - \frac{t}{2}$, which is the price level that a monopolist would choose. Let this price be denoted by p^M. If firm 1 also chooses to set its price equal to p^M, two firms equally split the consumers who receive both firms' ads. A fraction $\phi_1 (1 - \phi_2)$ of consumers only receive firm 1's ads, so they purchase the brand offered by firm 1. If firm 1 chooses a different price, the recipients of its ads infer that firm 1 is not the only firm. Therefore, they conduct a search and find out that there is another brand offering the same good at a price of p^M. Regardless of the exact locations of the two brands, the demand that accrues to firm1 from these fully informed consumers is $\phi_1 \left(\frac{1}{2} + \frac{p_2 - p_1}{t}\right)$ as long as they are located $\frac{1}{2}$ unit apart from each other.

In a symmetric non-cooperative Nash equilibrium in which both firms act as if they offer the only brand of the good, both firms set the price at the monopoly level and then maximize the resulting profits with respect to the intensity of advertising. With a little abuse of notation, I will denote the equilibrium advertising reach and the resulting profits in this situation with the superscript M. Suppose that $p_2 = p^M$. If firm 1 also chooses $p_1 = p^M$, its profits are

$$\pi_1 = p^M \left[\phi_1 (1 - \phi_2) + \frac{\phi_1 \phi_2}{2}\right] - a\phi_1^2$$

The first-order condition with respect to ϕ_1 is

$$\frac{\partial \pi_1}{\partial \phi_1} = p^M \left(1 - \frac{\phi_2}{2}\right) - 2a\phi_1 = 0$$

At a symmetric equilibrium $\phi_1 = \phi_2$. Letting ϕ^M denote this common value of advertising intensity, we get

$$\phi^M = \frac{p^M}{2a + \frac{1}{2}p^M}$$

The resulting profits of each firm in this equilibrium can be found as

$$\pi^M = \frac{a\left(p^M\right)^2}{\left(2a + \frac{1}{2}p^M\right)^2} = a\left(\phi^M\right)^2 \tag{4.6}$$

Firm 1 would choose a different price if it is able to achieve higher profits. Given that $p_2 = p^M$ and $\phi_2 = \phi^M$, if firm 1 chooses a price $p_1 \neq p^M$, its profits are

$$\pi_1 = p_1 \left[\phi_1 \left(\frac{1}{2} + \frac{p^M - p_1}{t}\right)\right] - a\phi_1^2$$

Evaluated at $p^M = v - \frac{t}{2}$, the first term above becomes $p_1\phi_1\left(\frac{v-p_1}{t}\right)$. The first-order conditions with respect to $p_1$and ϕ_1 are

$$\frac{\partial \pi_1}{\partial p_1} = \phi_1 \left(\frac{v - 2p_1}{t}\right) = 0 \tag{4.7}$$

$$\frac{\partial \pi_1}{\partial \phi_1} = p_1 \left(\frac{v - p_1}{t}\right) - 2a\phi_1 = 0 \tag{4.8}$$

Note that π_1 is strictly concave in p_1 and ϕ_1, so the first-order conditions give the profit-maximizing price and advertising intensity. Let these be denoted by p^D and ϕ^D, and the resulting profits be denoted by π^D. Solving equation (4.7) for p_1 and assuming that $t < v \leq 2t$, we get

$$p^D = \frac{v}{2}$$

Substituting this into (4.8) and solving for ϕ_1 gives

$$\phi^D = \frac{v^2/4t}{2a}$$

So, the profits of firm 1 in this non-symmetric situation are

$$\pi^D = \frac{v^4}{64at^2} = a\left(\phi^D\right)^2 \tag{4.9}$$

Firm 1 would choose this strategy if $\pi^D > \pi^M$, which is true when $\phi^D > \phi^M$. However, this is impossible to occur when $t < v < 2t$.

Proposition 4.1 *Under Assumption 4.2, $\pi^M > \pi^D$ for all $t < v < 2t$. That is, both firms choose the monopoly price and the corresponding advertising intensity in a symmetric Nash equilibrium provided that the following three conditions are satisfied: (i) The market is completely covered in equilibrium when there is a single firm, i.e. $v > t$; (ii) A monopolist would find it too costly to send advertising messages to everyone, i.e. $v - \frac{t}{2} < 2a$; (iii) When the market is served by two firms, no firm is able to sell to all of its ad recipients by* *undercutting its rival, i.e.* $v < 2t$.

Proof. Let $\frac{v}{t} = x$, $\frac{a}{t} = y$. Then, ϕ^D and ϕ^M can be expressed as

$$\phi^D = \frac{(x/2)^2}{2y}, \quad \phi^M = \frac{1}{\frac{4y}{2x-1} + \frac{1}{2}}$$

$\phi^M > \phi^D$ if and only if

$$\frac{1}{(x/2)^2} > \frac{\frac{4y}{2x-1} + \frac{1}{2}}{2y} \Leftrightarrow \frac{4}{x^2} - \frac{2}{2x-1} > \frac{1}{4y}$$

By Assumption 2, we have $4y > 2x - 1$, so $\frac{1}{2x-1} > \dfrac{1}{4y}$. It now suffices to show

$$\frac{4}{x^2} - \frac{2}{2x-1} > \frac{1}{2x-1}$$

which would then imply $\phi^M > \phi^D$. Rearranging the above inequality, we get

$$3x^2 - 4(2x-1) < 0$$

The left-hand side equals $(3x-2)(x-2)$, which is negative for all $\frac{2}{3} < x < 2$. So, $\phi^M > \phi^D$ for all $t < v < 2t$ as claimed in the proposition. ■

When $v > 2t$, each consumer who receives firm 1's advertising message in the non-symmetric situation purchases firm 1's brand. Therefore, there is more incentive for undercutting the rival if it is charging the monopoly price. Whether this yields higher profits depends on the size of advertising costs. If it is too costly to advertise, then it is optimal to follow suit and charge the monopoly price.

Given that the two brands are located $\frac{1}{2}$ unit apart from each other, the consumer for whom firm 2's brand is the ideal one derives a utility of $v - p^M = \frac{t}{2}$ from consuming it. If she chooses to consume firm 1's brand, her net utility is $v - \frac{t}{2} - p_1$. This is greater than $\frac{t}{2}$ when $p_1 > v - t$. So, when $v > 2t$, firm 1 optimally sets $p^D = v - t$, thus capturing all consumers who receive at least one ad from firm 1. Substituting p^D into equation (4.8) implies $\phi^D = \frac{p^D}{2a} = \frac{v-t}{2a}$, and the resulting profits are $\pi^D = \frac{(v-t)^2}{4a} = a\left(\phi^D\right)^2$. Note that $\phi^D = \frac{v-t}{2a} < 1$ by Assumption 4.2.

Firm 1 earns a higher level of profits by charging a lower price if $\phi^D > \phi^M$.

$$\phi^D > \phi^M \;\Leftrightarrow\; \frac{v-t}{2a} > \frac{1}{\frac{4a}{2v-t} + \frac{1}{2}}$$

Rearranging the terms above, we get $(v-t)(2v-t) > 4at$ for $\phi^D > \phi^M$.

Proposition 4.2 *When $v > 2t$, both firms choose the monopoly price and the corresponding advertising intensity in a symmetric Nash equilibrium only if $(v-t)(2v-t) < 4at$.*

Comparing this result with the one presented in Proposition 4.1, we see that the level of advertising costs comes into consideration only when consumers' willingness to pay for the good is high enough. This is because a firm does not need to lower its price too much in order to capture from its rival all the consumers it can reach by advertising. However, this leads to a reduced incentive for achieving a high advertising reach, especially when advertising is more costly relative to consumers' willingness to pay. Therefore, when it is too costly to advertise, each firm chooses to charge the monopoly price, and achieve a higher advertising reach.

A comparison with the equilibrium advertising intensity found in the previous subsection yields that firms generally advertise less when consumers are rational. The underlying reason is that advertising in that case is treated as a public good. Since a firm's advertising also informs its recipients about the other brand, their incentive for free-riding on their rival's ads increases. Therefore, they advertise less.

4.3 Welfare Analysis

In this section, I find the socially optimal allocation of advertising for given price levels. Then, I compare the market level of advertising to the level that maximizes social welfare. The welfare standard used is the conventional one of consumer surplus plus profits. Aggregate welfare is given by

$$W = v\left[1-(1-\phi)^2\right] - 2a\phi^2 - T$$

The first term represents consumer benefits gross of disutility associated with consuming a brand that is not the ideal one. The term $1-(1-\phi)^2$ is the fraction of consumers that purchase the good The second term is the total advertising costs that firms incur. The final term, T, represents the aggregate consumer disutility associated with consuming a brand that is not the ideal one.

Consumers that purchase the good can be partitioned into two groups. The first group comprises those consumers who purchase the brand that is closer their ideal brands. The average distance between the ideal brand of a random consumer in this group and the brand that is consumed is $1/8$. The second group comprises those consumers who purchase the brand that is farther from their ideal brands. The average distance between the ideal brand of a random consumer in this group and the brand that is consumed is $3/8$.

4.3.1 Rational Consumers

When consumers are rational, all consumers reached by advertising purchase their most preferred brand. So, the aggregate transportation costs are

$$T = \frac{1-(1-\phi)^2}{8}t$$

The social welfare function becomes

$$W = \left(1-(1-\phi)^2\right)\left(v - \frac{t}{8}\right) - 2a\phi^2$$

which is maximized at $\phi^S = \frac{2v-\frac{t}{4}}{4a+2v-\frac{t}{4}}$. This can be rearranged as

$$\phi^S = \frac{1}{\frac{4a}{2v-(t/4)}+1} \tag{4.10}$$

The market equilibrium, given by equation (4.4), can be rearranged as $\phi^R = \dfrac{1}{\frac{4a}{(t/2)}+1}$.

Since $2v - \frac{t}{4} > \frac{t}{2}$, it follows that $\phi^R < \phi^S$.

Proposition 4.3 *When consumers are rational, the market always underprovides informative advertising.*

The intuition in this case is quite simple. As discussed before, advertising is a public good, and therefore free-riding arises.

4.3.2 Ignorant Consumers

When consumers are ignorant, the probability that a consumer will purchase the closer brand is simply the advertising intensity chosen by the seller of that brand. Hence, the size of the first group is ϕ. Similarly, $\phi(1-\phi)$ is the probability that she will purchase the farther brand. So, when consumers are ignorant, the total transportation costs are

$$T = t\left[\frac{\phi}{8} + \frac{3\phi(1-\phi)}{8}\right]$$

Returning to the expression for aggregate welfare, we have

$$W = v\left[1 - (1-\phi)^2\right] - 2a\phi^2 - t\left[\frac{\phi}{8} + \frac{3\phi(1-\phi)}{8}\right]$$

The first-oder condition with respect to ϕ yields

$$(1-\phi)v - 2a\phi - \frac{(2-3\phi)t}{8} = 0$$

After rearranging, we get

$$\phi^S = \frac{v - \frac{t}{4}}{2a + v - \frac{3t}{8}} \tag{4.11}$$

where ϕ^S is the socially optimal advertising intensity of one firm. The second derivative of W is $\left(-v - 2a + \frac{3t}{8}\right)$, which is negative since $v > t$.

Proposition 4.4 *When consumers are ignorant, the market underprovides informative advertising if* $v - \dfrac{t}{2} < \sqrt{at}$.

Proof. $\phi^M < \phi^S$ iff

$$\left(v - \frac{t}{2}\right)\left(2a + v - \frac{3t}{8}\right) < \left(v - \frac{t}{4}\right)\left(2a + \frac{v}{2} - \frac{t}{4}\right)$$

The right-hand side can be written as $\left(v - \frac{t}{2}\right)\left(2a + \frac{v}{2} - \frac{t}{4}\right) + \frac{t}{4}\left(2a + \frac{v}{2} - \frac{t}{4}\right)$. Taking the first one of these expressions to the left-hand side, we get

$$\left(v - \frac{t}{2}\right)\left(\frac{v}{2} - \frac{t}{8}\right) < \frac{t}{4}\left(2a + \frac{v}{2} - \frac{t}{4}\right)$$

Reorganizing yields

$$2\left(\frac{v^2}{2} - \frac{vt}{2} + \frac{t^2}{8}\right) < at$$

$v^2 - vt + \frac{t^2}{4} = \left(v - \frac{t}{2}\right)^2$. So, we get $\phi^M < \phi^S$ iff $v - \frac{t}{2} < \sqrt{at}$. ■

By Assumption 4.1, $v - \frac{t}{2} > \frac{t}{2}$ So, if $\dfrac{t}{2} \geq \sqrt{at}$, we always get overprovision of advertising in the market equilibrium. This occurs when $4a \leq t$. However, this condition violates

Assumption 4.2. So, under the two assumptions, there is no value of a that causes firms to overprovide advertising regardless of the degree of differentiation. As the two brands become more similar (i.e. as $t \to 0$), firms always overprovide informative advertising regardless of its cost. Otherwise, it depends on the values of a and t. Confining attention to the case when $t < v \leq 2t$, if advertising costs are high, firms generally underprovide advertising.

These findings are in sharp contrast with the existing literature. In the current setting, a firm has an incentive to advertise more since doing so results in a higher number of consumers who are only informed about that firm's brand. This incentive is even higher when the two brands are less differentiated. In such a case, a firm considers the effects of a possible undercutting by its rival. If its rival advertises a lower price, then it only sells to consumers who are only informed about its brand. Therefore, it advertises more. However, this effect vanishes as the degree of differentiation increases. In this case, a firm can also sell to some of the fully informed consumers.

4.4 Conclusion

This chapter has presented a modified duopoly version of the seminal paper of Grossman and Shapiro on informative advertising. I have added costless consumer search into their analysis. In the presence of "ignorant" consumers, the findings are strikingly different. Both firms may advertise the monopoly price, and we may get too little advertising compared to the socially optimal amount. When consumers are fully rational, advertising is always undersupplied. When consumers are not fully rational, advertising is undersupplied if it is sufficiently costly.

The result that monopoly pricing may arise when consumers are ignorant may seem to have been driven by the assumption of ignorant consumers. However, I have made this assumption just to get a nontrivial equilibrium in the presence of costless search. In fact, the same pricing strategy should carry over to a few other, economically more reasonable, specifications. Suppose that search is costly and consumers are initially unaware of product existence. When they get an ad, they immediately come up with a good estimate of the fixed cost of producing such a good. Consider the case when they think the fixed is either low in which case only one firm can survive in this market, is moderate in which case there can only be two firms, or is high in which case the market can accomodate three firms. This also depends on their beliefs for possible product variations which can, for simplicity, be described by a uniform distribution. Then, we can determine the optimal search behavior of partially informed consumers as a function of their search costs. For the monopoly price advertising result to arise, the search cost must exceed a certain minimum threshold value.

References

[1] Anand, Bharat N. and Ron Shachar (1998), "The Effectiveness and Targeting of Television Advertising," *Journal of Economics and Management Strategy*, 7, 363-396.

[2] Anand, Bharat N. and Ron Shachar (2005), "Advertising the Matchmaker," mimeo, Harvard Business School and Tel Aviv University.

[3] Anand, Bharat N. and Ron Shachar (2006), "Targeted Advertising as a Signal," mimeo, Harvard Business School and Tel Aviv University.

[4] Anderson, Simon P. and André de Palma (2006), "Information Congestion," mimeo, Université de Cergy-Pontoise and University of Virginia.

[5] Anderson, Simon P., André de Palma and Jacques-F. Thisse (1992), *Discrete Choice Theory of Product Differentiation*, MIT Press, Cambridge.

[6] Anderson, Simon P. and Jean J. Gabszewicz (2005), "The Media and Advertising: A Tale of Two-Sided Markets," working paper, University of Virginia and CORE, forthcoming in *Handbook of Cultural Economics*, eds. Victor Ginsburgh and David Throsby, Elsevier Science.

[7] Anderson, Simon P. and Régis Renault (1999), " Pricing, Product Diversity, and Search Costs: A Bertrand-Chamberlin-Diamond Model," *Rand Journal of Economics*, 30:4, pp. 719-735.

[8] Anderson, Simon P. and Régis Renault (2006), "Advertising Content," *American Economic Review*, 96:1, pp. 93-113.

[9] Anderson, Simon P. and Stephen Coate (2005), "Market Provision of Broadcasting: A Welfare Analysis," *Review of Economic Studies*, 72:4, pp. 947-972.

[10] Armstrong, Mark (2005), "Competition in Two-Sided Markets," mimeo, University College London. Forthcoming in *Rand Journal of Economics*.

[11] Bagwell, Kyle. (2005), "The Economic Analysis of Advertising," mimeo, Columbia University. Forthcoming in *Handbook of Industrial Organizatio*n *Vol. 3*, eds. Mark Armstrong and Robert Porter, North-Holland, Amsterdam.

[12] Bagwell, Kyle and Garey Ramey (1994), "Advertising and Coordination," *Review of Economic Studies*, 61:1, pp. 153-171.

[13] Bain, Joe S. (1949), "A Note on Pricing in Monopoly and Oligopoly," *American Economic Review*, 39, pp. 448-464.

[14] Becker, Gary and Kevin Murphy (1993), "A Simple Theory of Advertising as a Good or Bad," *Quarterly Journal of Economics*, pp. 942-964.

[15] Bontems, Philippe and Valérie Meunier (2006), "Advertising and Price Signaling of Quality in a Duopoly with Endogenous Locations," mimeo, Université des Sciences Sociales de Toulouse and University of Aarhus. Forthcoming in *Journal of Economics and Management Strategy*.

[16] Butters, Gerard R. (1977), "Equilibrium Distributions of Sales and Advertising Prices," *Review of Economic Studies*, 44:3, pp. 465-491.

[17] Caillaud, Bernard and Bruno Jullien (2001), "Competing Cybermediaries," *European Economic Revie*w, 45:4-6, pp. 797-808.

[18] Caillaud, Bernard and Bruno Jullien (2003), "Chicken and Egg: Competition Among Intermediation Service Providers," *Rand Journal of Economics*, 34:2, pp. 309-328.

[19] Christou, Charalambos and Nicolaos Vettas (2003) "Informative Advertising and Product Differentiation," Center for Economic Policy Research Discussion Paper No. 3953. Forthcoming in *International Journal of Industrial Organization*.

[20] Coate, Stephen (2004), "Political Competition with Campaign Contributions and Informative Advertising," *Journal of the European Economic Association*, 2:5, pp. 772-804.

[21] Crampes, Claude, Carole Haritchabalet and Bruno Jullien (2005), "Advertising, Competition and Entry in Media Industries," mimeo, IDEI and GREMAQ, University of Toulouse.

[22] Crawford, Vincent P. and Joel Sobel (1982), "Strategic Information Transmission," *Econometrica*, 50:6, pp. 1431-1451.

[23] d'Aspremont, Claude, Jean J. Gabszewicz and Jacques-F. Thisse (1979), **"**On Hotelling's Stability in Competition," *Econometrica*, 47:5, pp. 1145-1150.

[24] d'Aspremont, Claude, Jean J. Gabszewicz and Jacques-F. Thisse (1983), "Product Differences and Prices," *Economics Letters*, 11:1-2, pp. 19-23.

[25] de Bijl, Paul W. J. (1997) "Entry Deterrence and Signaling in Markets for Search Goods," *International Journal of Industrial Organization*, 16, pp. 1-19.

[26] Diamond, Peter A. (1971), "A Model of Price Adjustment," *Journal of Economic Theory*, 3, pp. 156-168.

[27] Dixit, Avinash K. and Victor D. Norman (1978), "Advertising and Welfare," *Bell Journal of Economics*, 9:1, pp. 1-17.

[28] Dixit, Avinash K. and Victor D. Norman (1981), "Advertising and Welfare: Reply," *Bell Journal of Economics*, 12:2, pp. 753-754.

[29] Dixit, Avinash K. and Joseph E. Stiglitz (1977), "Monopolistic Competition and Optimum Product Diversity," *American Economic Review*, 67:3, pp. 297-308.

[30] Dukes, Anthony (2004), "The Advertising Market in a Product Oligopoly," *Journal of Industrial Economics*, 52, pp. 327-348.

[31] Dukes, Anthony and Esther Gal-Or (2003a), "Minimum Differentiation in Commercial Media Markets," *Journal of Economics and Management Strategy*, 12, pp. 291-325.

[32] Dukes, Anthony and Esther Gal-Or (2003b), "Negotiations and Exclusivity Contracts for Advertising," *Marketing Science*, 22:2, pp. 222-247.

[33] Emerson, John W. and Ron Shachar (2000), "Cast Demographics, Unobserved Segments, and Heterogeneous Switching Costs in a Television Viewing Choice Model," *Journal of Marketing Research*, 37, pp. 173-186.

[34] Fluet, Claude and Paolo G. Garella (2002), "Advertising and Prices as Signals of Quality in a Regime of Price Rivalry," *International Journal of Industrial Organization*, 20, pp. 907-930.

[35] Gabszewicz, Jean, Didier Laussel and Nathalie Sonnac (2004), "Programming and Advertising Competition in the Broadcasting Industry," *Journal of Economics & Management Strategy*, 13:4, pp. 657-669.

[36] Goettler, Ronald and Ron Shachar (2001), "Spatial Competition in the Network Television Industry," *Rand Journal of Economics*, 32(4).

[37] Grossman, Gene and Carl Shapiro (1984), "Informative Advertising with Differentiated Products," *Review of Economic Studies*, 51, pp. 63-81.

[38] Hamilton, Stephen F. (2004), "Informative Advertising in Concentrated, Differentiated Markets," mimeo, University of Central Florida.

[39] Hertzendorf, Mark N. and Per Baltzer Overgaard (2001a), "Price Competition and Advertising Signals: Signaling by Competing Senders," *Journal of Economics and Management Strategy*, 10:4, pp. 621-663.

[40] Hertzendorf, Mark N. and Per Baltzer Overgaard (2001b), “Prices as Signals of Quality in Duopoly,” mimeo, Federal Trade Commission and University of Aarhus.

[41] Horstman, Ignatius J. and Glenn MacDonald (1994), “When is Advertising a Signal of Product Quality?,” *Journal of Economics and Management Strategy*, 3, pp. 561-584.

[42] Hotelling, Harold (1929), “Stability in Competition,” *Economic Journal*, 39, pp. 41-57.

[43] Kaldor, Nicholas (1949), “The Economic Aspects of Advertising,” *Review of Economic Studies*, 18, pp. 1-27.

[44] Kihlstrom, Richard E. and Michael H. Riordan (1984), “Advertising as a Signal,” *Journal of Political Economy*, 92, pp. 427-450.

[45] Matthews, Steven A. and Doron Fertig (1990), “Advertising Signals of Product Quality,” Discussion Paper No. 881, Northwestern University.

[46] Meurer, Michael and Dale O. Stahl, II (1994), “Informative Advertising and Product Match,” *International Journal of Industrial Organization*, 12, pp. 1-19.

[47] Milgrom, Paul R. (1981), “Good News and Bad News: Representation Theorems and Applications,” *Bell Journal of Economics*, 12, pp. 380-391.

[48] Milgrom, Paul R. and John M. Roberts (1986a), “Prices and Advertising as Signals of Product Quality,” *Journal of Political Economy*, 94, pp. 796-821.

[49] Milgrom, Paul R. and John M. Roberts (1986b), "Relying on the Information of Interested Parties," *Rand Journal of Economics*, 17:1, pp. 18-32.

[50] Moraga-González, José Luis (2000), "Quality Uncertainty and Informative Advertising," *International Journal of Industrial Organization*, 18, pp. 615–640.

[51] Moshkin, Nickolay V. and Ron Shachar (2002), "The Asymmetric Information Model of State Dependence," *Marketing Science*, 21:4, pp. 435-454.

[52] Nelson, Phillip J. (1970), "Information and Consumer Behavior," *Journal of Political Economy*, 78, pp. 311-329.

[53] Nelson, Phillip J. (1974), "Advertising as Information," *Journal of Political Economy*, 82, pp. 729-754.

[54] Nelson, Phillip J. (1975), "The Economic Consequences of Advertising," *Journal of Business*, 48, pp. 213-41.

[55] Roberts, Jacques and Dale O. Stahl, II (1993), "Informative Price Advertising in a Sequential Search Model," *Econometrica*, 61:3, pp. 657-686.

[56] Rochet, Jean-Charles and Jean Tirole (2003), "Platform Competition in Two-Sided Markets," *Journal of the European Economic Association*, 1:4, pp. 990.

[57] Rochet, Jean-Charles and Jean Tirole (2005), "Two-Sided Markets: A Progress Report," IDEI Working Paper No. 275, University of Toulouse. Forthcoming in *Rand Journal of Economics*.

[58] Salop, Steven C. (1979) "Monopolistic Competition with Outside Goods," *Bell Journal of Economics*, 10, pp. 141-156.

[59] Shapiro, Carl (1980), "Advertising and Welfare: Comment," *Bell Journal of Economics*, 11, pp. 749-752.

[60] Soberman, David (2002), "Informative Advertising: An Alternate Viewpoint and Implications," *Review of Marketing Science Working Papers*, 1:3, Working Paper No. 3.

[61] Stahl, Dale O. (1994) "Oligopolistic Pricing and Advertising," *Journal of Economic Theory*, 64, pp. 162-177.

[62] Spence, Michael and Bruce Owen (1977), "Television Programming, Monopolistic Competition, and Welfare," *Quarterly Journal of Economics*, 91:1, pp. 103-126.

[63] Stegeman, Mark (1991) "Advertising in Competitive Markets," *American Economic Review*, 81, pp. 210-223.

[64] Tirole, Jean (1988), *The Theory of Industrial Organization*, MIT Press, Cambridge.

[65] Vickrey, William S. (1964), *Microstatics*, Harcourt, Brace and World, New York. Republished as: Vickrey, William S., Simon P. Anderson and Ralph M. Braid (1999), "Spatial Competition, Monopolistic Competition, and Optimum Product Diversity," *International Journal of Industrial Organization*, 17, pp. 953-963.

[66] Wilbur, Kenneth C. (2005), "Not All Eyeballs are Created Equal: A Structural Equilibrium Model of Television Advertisers, Networks, and Viewers," unpublished dissertation, University of Virginia.

[67] Wolinsky, Asher (1984), "Product Differentiation with Imperfect Information," *Review of Economic Studies*, 51, pp. 53-61.