Provision of Informative Advertising: The Case of Broadcasting

Levent Çelik*

November 2005

Abstract

Television (TV) networks forgo almost 20% of their advertising revenues in order to air advertisements for their own programs (tune-ins). This paper explores possible explanations for why they would ‘pass up’ such significant amounts of revenue. The basic model involves a single TV station offering two consecutive programs that are only horizontally differentiated. Viewers are assumed to be uncertain about the degree of the differentiation between the two programs. I show for a wide range of parameter values that the TV station airs a tune-in for its upcoming program unless it is too dissimilar with the program during which the tune-in would be aired. I extend the analysis to include quality uncertainty as well. An important implication is that ‘money burning’ is not always necessary to signal high quality. There are program types that only a TV station whose program is of high quality can afford to advertise. In a further extension, I include a second TV station in the market. The main finding is that the tune-in decision of a station does not depend on the rival station’s upcoming program when the sampling cost is not too high.

Keywords: Informative Advertising, Tune-ins, Uncertainty, Sampling, Information Disclosure.
JEL Classification: D83, M37, L82

*University of Virginia, Department of Economics, 221 Wilson Hall, Charlottesville, VA 22903, email: le2t@virginia.edu. I would like to thank my advisors, Simon Anderson and Maxim Engers, for their valuable comments.
1 Introduction

Many consumer markets exist where the differentiation among products is mainly in their physical attributes. Although the information about the existence of these products may be common knowledge, many people may have limited or even no information about their attributes. This lack of information could be the result of newly introduced products or high costs associated with gaining information about product attributes. The present paper proposes an explanation for the incentives of a firm to provide directly informative advertising about product attributes when the recipients of the advertisements (ads) are rational. The rationality of consumers is crucial, implying that information about a product comes from both the content of the ad and the decision of a firm to advertise.

While it is agreed that the primary role of advertising is to provide information to potential consumers, there has been controversy regarding the nature of this information. One theory, originally developed by Nelson [1974] and later formalized by Kihlstrom and Riordan [1984] and Milgrom and Roberts [1986], is that advertising in and of itself is a signal of high product quality. The signal of high quality is conveyed regardless of the content of the ad. Therefore, high-quality firms engage in ‘money burning’ in order to signal their quality. Another motivation for firms to advertise could be a desire to shift consumers’ tastes (and therefore demand) for their product (see Dixit and Norman [1978]). However, in most real life situations, advertising conveys direct information about product attributes. A good example is the distribution of free samples.

The primary objective of this paper is to characterize the incentives of a firm to advertise the attributes of its product when people are uncertain about them. The model is developed in the context of a television (TV) industry. The TV industry is of particular interest because stations forgo about 20% of their advertising revenues to air preview ads (tune-ins) for their own programs (Shachar and Anand [1998]). These tune-ins provide direct information about program attributes. Although a person can acquire information about

\(^1\) Shachar and Anand [1998] report that in 1995, three major network stations in the U.S. devoted 2 of 12 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%.

\(^2\) It is necessary to distinguish between tune-ins for regular programs, like everyweek sitcoms, and those for special programs, like 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.
the attributes of a program through TV schedules that appear in magazines or through word-of-mouth, some fraction of viewers still remain uninformed due to the costs associated with gaining information. Another advantage of utilizing the TV industry in this context is that network stations do not price their programs in a direct way. Rather they price their programs indirectly by bundling programs with ads. This enables a focus only on the interaction between horizontal differentiation and provision of informative advertising.

I first lay out a benchmark model that has a single TV station airing two consecutive programs. Potential viewers differ in their preferred types of programs. I assume that they know the type of the program to be aired in the first period, but are uncertain about the type 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. He chooses to do so, if the marginal advertising revenue resulting from the increase in the second-period audience size dominates 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 result from tune-ins. As a result of these informational benefits, some viewers watch TV in the first period who would have chosen not to watch if they were not forward-looking.

Next, I relax the assumption that viewers cannot switch off once they choose to watch a program. I do this by introducing a “sampling cost” that is incurred when a person turns the TV off and enjoys the outside option after sampling a few minutes of the program. In this case, the TV station does not have as much market power. The result is that the station has lesser incentive to air a tune-in.

Finally, I discuss the implications of two extensions, adding another dimension of uncertainty, namely quality, and adding one more TV station. In the quality uncertainty case, the basic framework suggests that ‘money burning’ is not always necessary for a TV station to signal high quality. This follows from the presence of program types that only a TV station with a high-quality program can afford to advertise. When there are two television stations in the market, the advertising decision of a station may strategically depend on the type of the other station’s upcoming program. For some, though not all, program types of the rival
station, not airing a tune-in (thus potentially causing some people to switch to the other
station) is optimal.

The paper is organized as follows. The next section discusses the relevant literature. The
benchmark model is introduced in section 3. Section 4 discusses the equilibrium when the
model is extended to allow viewers to turn off their TVs. Section 5 discusses the robustness
of the basic model and introduces two extensions. Finally, section 6 concludes.

2 Previous Literature

Directly informative advertising has been the topic of several previous studies. Butters [1977]
was the first to model the informative role of advertising. In his paper advertising is the
mechanism through which firms advertise the price of their products. Because consumers
have no knowledge of product existence prior to receiving an ad, the ad indirectly informs
them of this as well. Grossman and Shapiro [1984] study an extended model in which
consumers are heterogeneous in their preferences. Advertising informs them not only about
the existence but also about the characteristics of the products. Common to both of these
papers is that the advertising technology is exogenous, so that people cannot change their
likelihood of receiving an ad. My model is similar to the one used in Grossman and Shapiro.
Consumers are heterogeneous in their preferences over two horizontally differentiated goods.
They use the information advertising provides them about product attributes to choose the
product that yields them the highest expected benefit. My model will depart from theirs
by using a dynamic framework and assuming that product existence is common knowledge.
Another important distinction 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
will receive a tune-in, if and only if she watches the first program.\footnote{To the best of my knowledge, previous work in 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 advertisements. Take the example of low fare alerts that one can receive through an email from Travelocity. Other examples are using a digital video recorder to skip advertisements while watching a TV program or subscribing to “Don’t Call List” to avoid calls by telemarketers. Although this paper does not specifically model how people change their likelihood of receiving an ad, it allows people to receive a tune-in by watching the first program, even when the first-period utility is negative, if doing so has high enough informational benefits.}

3
Three more recent theoretical papers on informative advertising are Meurer and Stahl [1994], Anderson and Renault [2005], and Anand and Shachar [2004]. Meurer and Stahl [1994] analyze the welfare properties of informative advertising in horizontally differentiated markets where a fraction of the consumers are uninformed about the product characteristics. They characterize a unique subgame perfect Nash equilibrium in which the level of advertising provided may be more or less than socially optimal. They assume that all consumers observe prices beforehand and they treat product information as a public good, so that information about one product provides information about the others as well. Anderson and Renault [2005] 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 search costs that consumers face. Furthermore, their results show that the monopolist prefers to convey only limited product information. Anand and Shachar [2004] model advertising as a noisy message that indirectly conveys information to a receiver through the information revealed in the decision to send it. They show that when there are two symmetric senders of different types (high and low types), even highly noisy messages may enable the receiver to fully infer the identities of both senders from their willingness to provide information. Because of the noise in the messages, perfect separation in their model can only happen by advertising at a high frequency.

The focus of the current paper is not on the number of informative ads, but rather on the incentives driving firms to advertise. To analyze the firms’ decision making, I will use models similar to the ones just mentioned, but will extend them in several important dimensions. First, unlike Meurer and Stahl [1994] and Anand and Shachar [2004], I characterize the incentives of firms to reveal information about their products when match information about one product does not necessarily reveal the match value of the other product. Second, I build on the model used by Anderson and Renault [2005] by introducing a second firm. For particular product types, a firm may feel safe enough to not advertise its product when it knows that its current customers will not like the rival firm’s product. Third, unlike Anand and Shachar [2004], advertising in my model is fully informative in terms of content and there are many potential buyers. The buyers all have a common prior about product
type(s). Under monopoly, beliefs are characterized by a continuous distribution. When there are two firms, they are characterized by a discrete distribution comprising three distinct values.

Several empirical studies of the effects of tune-ins on viewing choices of individuals have been done using a micro-level panel dataset on the TV viewing decisions of a large sample of individuals. Those most related to the theoretical model in this paper are two papers written by Anand and Shachar. Their 1998 paper estimates the differential effects of tune-ins on viewing decisions for regular and special shows. In their 2005 paper, advertising content is modeled as a noisy signal of product attributes, which affects the information sets of a priori uncertain viewers. I will improve upon the models used in these empirical papers by assuming agents are forward-looking rather than myopic. This introduces some important dynamic elements. Another contribution will be my attention to the supply side. Not incorporating supply side dynamics creates an endogeneity problem, since the tune-in strategies of TV stations depend on the viewing decisions of people. Furthermore, rational viewers will anticipate the equilibrium strategies of the stations and behave accordingly even in the absence of tune-ins.

3 Benchmark Model

In order to grasp a general idea about what the equilibrium looks like, I present, in this section, a benchmark model with a single TV station and no option of turning the TV off mid show. These assumptions will be relaxed later. There is a single TV station (will be referred to as “he” when convenient) airing two programs, \((x_1, x_2)\), in two consecutive time periods, \(t = 1, 2\), where \(x_t\) represents the type of the program and is located on the unit line. 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 is a number of \(A\) available non-programming breaks during each program, where \(A\) is taken as exogeneous.\(^4\)

---

\(^4\)The assumption that the number of non-programming breaks is fixed is certainly restrictive. However, while U.S. broadcasters are free to choose the number of their commercial breaks, in most European countries, advertising ceilings are imposed on broadcasters. Therefore, in most cases, especially in the prime time, the number of commercial breaks that maximizes a broadcaster’s revenues falls below the imposed ceiling. There are also technical reasons for this assumption. First, if TV stations were allowed to choose the number of non-programming breaks, then people would rationally form priors about them since they are assumed to be
A large number of firms exists, each willing to pay up to $p$ dollars per viewer reached for placing an advertisement during a program. The TV station may choose to devote one of the non-programming 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, 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 advertisement revenues (for simplicity, it is assumed that there is no discounting).

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, 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-programming breaks.\(^5\) 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.\(^6\) I will commonly use the parameter $\lambda$ to represent the location of an individual and refer to a particular individual 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 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 might obtain by seeing a tune-in for the second program. They base their decisions on their prior beliefs about the type of the second program and the equilibrium tune-in strategy of the TV

---

\(^5\)The base utility $v$ also captures the effects of the disutility associated with interruptions during a program. Specifically, the effect of an increase (a decrease) in the nuisance cost of a non-programming break on a viewer’s utility can be captured by lowering (raising) the base utility.

\(^6\)A constant can be included in front of $d$ that measures the disutility associated with one unit of distance from the ideal program type. However, since the value of not watching TV is zero, the utility can easily be expressed as $r - d$, where $r = \frac{v}{t}$. 

6
station, which they rationally anticipate. Their unconditional prior belief about the type of
the second program is summarized by a uniform distribution over $[0, 1]$. To simplify the analysis, it is assumed throughout the paper that the first program is
located at 0.\textsuperscript{7} As a result of this assumption, people whose ideal program types are to the
left of $v$ watch the first program regardless of any associated informational benefits. It is
also assumed that $\frac{1}{4} < v < \bar{v}$, where $\bar{v}$ is less than $\frac{1}{2}$ (its exact value will be derived in the
analysis.) Because of possible informational gains associated with watching the first program,
some people with $\lambda > v$ might also prefer watching TV in the first period in expectations of
a higher second period utility.

The timing of the game is as follows. First, people make their first period decisions that
maximize their expected two-period utilities. Then the first program starts, and during its
progress, the TV station makes his tune-in decision. When the first program ends, those who
watched the first program update their beliefs for the type of the second program depending
on whether they were exposed to a tune-in or not. Finally people make their second period
optimal decisions and payoffs are realized. As a tie breaking rule, I will assume that the
TV station airs a tune-in when he is indifferent between airing and not airing one, and that
people watch TV when they are indifferent between watching and not watching.

3.1 Equilibrium

The equilibrium concept used is Perfect Bayesian Equilibrium (PBE). That is, we require
the TV station to make an optimal tune-in decision taking the rationality of people into
account, 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 type of
the second program following no tune-in 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 would result following no tune-in. Because a tune-in is assumed to be fully informative,
and people watch a program until the end, the TV station does not air more than one tune-
in. Since the type of the second program is unknown to people prior to their first period

\textsuperscript{7}This assumption does not change any of the main implications of the paper.
decisions, viewers can only form beliefs about the program types for which the TV station will air a tune-in. These beliefs will be described by a set of points $\Omega$ such that people ex-ante anticipate to see a tune-in for the second program during the first one 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 type is larger than $v$. If she watches the first program and sees a tune-in for the second program, she will watch the second program as well only when the advertised program type is at most $v$ units apart from her location. If she watches the first program and does not see a tune-in, then she updates her belief for the type of the second program and keeps watching TV only when her posterior expected utility of doing so is nonnegative. If she does not watch the first program, she will base her decision to watch the second program only on her prior belief about its type. The benefits, $B(\lambda)$, and costs, $C(\lambda)$, of watching the first program for this individual can be written as

$$B(\lambda) = \int_{\lambda-v}^{\lambda+v} u(\lambda, x) 1[x \in \Omega] \, dx + \max \left\{ 0, \int_0^1 u(\lambda, x) 1[x \notin \Omega] \, dx \right\} - \max \left\{ 0, \int_0^1 u(\lambda, x) \, dx \right\}$$

$$C(\lambda) = \lambda - v$$

where $1[\cdot]$ is the indicator function, and $u(\lambda, x) = (v - |\lambda - x|)$. The optimal first period decision of an individual with $\lambda > v$ is to watch TV when $B(\lambda) - C(\lambda) \geq 0$.

$\Omega$ is determined in the equilibrium by people’s anticipations for the TV station’s tune-in decision corresponding to every possible program type. The marginal benefit of airing a tune-in for the TV station is the marginal second period advertisement revenues as a result of higher audience size. The cost is the forgone revenues of an additional advertisement 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 he does not air a tune-in, and $q = 1$ when he does. So, from people’s point of view, the optimal tune-in decision of the TV station as a function of $x_2$ is given by

$$q = \begin{cases} 
1, & A p N \left[ s_2 \left( x_2 \mid q = 1 \right) - s_2 \left( x_2 \mid q = 0 \right) \right] \geq p N s_1 \\
0, & \text{otherwise}
\end{cases}$$

where $s_1$ is the population share of the first period audience size and $s_2 \left( x_2 \mid q \right)$ is the population share of the second period audience size conditional on the realization of $q$. $p N s_1$ is the
amount of advertising revenue that the TV station forgoes. The left-hand side of the same expression denotes the marginal advertising revenues when the TV station airs a tune-in.

**Lemma 1** The only set of consistent beliefs in an equilibrium is in the form of $\Omega = [x_L, x_H]$.

**Proof.** See the Appendix.

As mentioned in the lemma, let $\Omega = [x_L, x_H]$ where $x_L$ is the lower bound and $x_H$ is the upper bound of the set of program types for which people anticipate to see a tune-in. Given $\Omega = [x_L, x_H]$, the benefits of watching the first program can be expressed as

$$B(\lambda) = \min\{x_H, \lambda + v\} \cdot \int_0^{x_L} u(\lambda, x) \, dx + \max\left\{0, \left(\int_0^{x_L} u(\lambda, x) \, dx + \frac{1}{x_H} \int_0^{x_L} u(\lambda, x) \, dx\right)\right\}$$

$$- \max\left\{0, \frac{1}{x_H} \int_0^{x_L} u(\lambda, x) \, dx\right\}$$

Note that all of the three terms in $B(\lambda)$ are continuous functions of $\lambda$. However, since the value of $\lambda$ such that the second term or the third term equals zero could be an interior value (depending on whether these values of $\lambda$ are less than or greater than $v$), $B(\lambda) - C(\lambda)$ could display kinks. Nevertheless, being summation of four continuous functions, $B(\lambda) - C(\lambda)$ must be continuous in $\lambda$. Furthermore, it monotonically decreases as $\lambda$ increases. This is because the decrease in $(v - \lambda)$ as $\lambda$ increases always dominates the aggregate change in the other terms. Therefore, if $B(\lambda) > C(\lambda)$ at $\lambda = v$, then there exists a unique value of $\lambda$ such that $B(\lambda) = C(\lambda)$. Let this value of $\lambda$ be denoted by $\hat{\lambda}$. This critical value of $\lambda$ 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 posterior belief. She will watch the second program if

$$E[u(\lambda, x_2 | x_2 \notin \Omega)] \geq 0,$$

where

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

This function is clearly continuous in $\lambda$. Provided that $x_H$ is high enough, the value of $\lambda$ that equates it to zero is unique. Using the relationship between $B(\lambda) - C(\lambda)$ and $E[u(\lambda, x_2 | x_2 \notin \Omega)]$, as well as the equilibrium existence conditions, the next lemma shows
that for $\Omega$ to be consistent with TV station’s optimal tune-in decision, the equilibrium value of the lower bound of $\Omega$ must be equal to zero. This result proves useful in characterizing the equilibrium.

**Lemma 2** $x_L = 0$ in equilibrium. That is $\Omega = [0, x_H]$.

**Proof.** See the Appendix.

It directly follows from this lemma that $x_H > \hat{\lambda}$. This is because, the two program types, $x_2 = 0$ and $x_2 = \hat{\lambda}$, yield the same second-period audience size if advertised during the first program.

**Corollary 1** The equilibrium value of $x_H$ is greater than $\hat{\lambda}$, since $s_2(0 \mid q = 1) = s_2(\hat{\lambda} \mid q = 1)$.

Now that the form of the PBE has been specified, it is possible to characterize the viewing decisions of people. Let $\lambda^o$ denote the upper bound of the set of people that would watch TV at $t = 2$ only based on their prior for $x_2$. Formally

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

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

**Lemma 3** In equilibrium, $\hat{\lambda} < x_H < \hat{\lambda} + v$.

**Proof.** See the Appendix.

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

**Lemma 4** No one from the first-period audience keeps watching TV unless the TV station airs a tune-in for the second program. That is $E[u(\lambda, x_2 \mid x_2 \notin \Omega)] < 0$ for all $\lambda \leq \hat{\lambda}$.

**Proof.** See the Appendix.
From earlier discussion, the necessary condition for the existence of a PBE is given by
\[ s_2(x_2 \mid q = 1) - s_2(x_2 \mid q = 0) \geq \frac{\hat{\lambda}}{A} \text{ for } \forall x_2 \in [0, x_H] \text{ with strict equality if } x_2 = x_H \tag{1} \]

where \( \hat{\lambda} \) is given by the solution to
\[ \int_0^\mu (v - |\hat{\lambda} - x|) dx - \max \left\{ 0, \frac{1}{\mu} (v - |\hat{\lambda} - x|) dx \right\} = \hat{\lambda} - v \tag{2} \]

Note from expression (2) that if \( v \) is large so that \( x_H = 1 \), there is nothing to gain from watching the first program for an individual with \( \lambda > v \). 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 \( \bar{v} \), will be specified in the following proposition.

**Proposition 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 people with \( \lambda \leq \hat{\lambda} \) watch the first program, where \( \hat{\lambda} > v \) and is given by
\[
\hat{\lambda} = \begin{cases} 
\frac{\left(\sqrt{A^2 + 2v(1 + v)} - A\right) A}{1 - \sqrt{1 - \left(v^2 + \frac{1}{2}\right)\left(1 - \frac{1}{2A^2}\right)}} & , v < \hat{v} \\
\frac{v}{1 - \frac{1}{2A^2}} & , \hat{v} \leq v < \bar{v} \\
\frac{v}{1 - \frac{1}{2A^2}} & , v \geq \bar{v}
\end{cases}
\tag{3}
\]

In this expression, \( \hat{v} \) is the value of \( v \) that solves \( \hat{\lambda} = (1 - \lambda^v) \), and \( \bar{v} \) is the value of \( v \) that solves \( \hat{\lambda} = v \). The TV station airs a tune-in for all \( x_2 \leq x_H^* \), where \( x_H^* = v + \left(1 - \frac{1}{A}\right) \hat{\lambda} \).

**Proof.** See the Appendix.

The unique PBE of the model is described by a binary tune-in decision (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 only if the location of the second program is not too far from the location of the first program (which is taken to be zero). Before deciding to watch TV in the first period, people consider both their first period utilities and the informational benefits associated with doing so. In case there are no tune-ins during the first program, people who watched it correctly infer where the second program could possibly lie in and make their decisions accordingly.
Knowing that people are rational and that they 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 before and for all reasonable $A$. See the proof of Proposition (1) in the Appendix for a more detailed explanation regarding the restrictions on the parameter values.

4 Benchmark Model with Program Sampling

In this section, people can sample the first few minutes of the second program, if they wish, before they make their final second period decision. While this process fully reveals the true type 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 received had she chosen the outside option as the first thing, rather than sampling the program. Therefore, if an individual chooses to turn the TV off after she sampled it, her net second period benefit will be $-c$.

The model is slightly more complicated than the benchmark case because the viewers have more options. Given anticipations $\Omega$ 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 type is on the segment $(v, v + c)$. If she watches the first program and sees a tune-in, then she will watch the second program too, only when the advertised program type is at most $v$ units apart from her location. If she watches the first program and does not see a tune-in, then she updates her beliefs for the type of the second program and chooses to sample it when 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 type turns out to be more than $v + c$ units apart from her location. If she does not watch the first program, she will base her decision to sample the second program on her prior belief about its type. She samples it if the expected benefit of doing so exceeds its cost. In this case, she keeps watching until the end unless the program type turns out to be more than $v + c$ units apart from her location. So, we can express the
benefits of watching the first program for an individual with $\lambda > v$ as

$$B(\lambda) = \int_{\lambda-v}^{\lambda+v} u(\lambda, x) \mathbb{1}[x \in \Omega] \, dx + \max \left\{ 0, \int_0^{\lambda+v+c} u(\lambda, x) \mathbb{1}[x \notin \Omega] \, dx - (1 - (\lambda + v + c)) c \right\}$$

$$- \max \left\{ 0, \int_0^{\lambda+v+c} u(\lambda, x) \, dx - (1 - (\lambda + v + c)) c \right\}$$

Lemma 5 Suppose $c$ is small enough so that $\int_0^{\lambda+v+c} u(\lambda, x) \mathbb{1}[x \notin \Omega] \, dx - (1 - (\lambda + v + c)) c$ is positive. Then no viewer with $\lambda > v$ watches TV in the first period.

Proof. See the Appendix.

It is possible that no tune-ins prevail in equilibrium if the sampling cost is low. As discussed earlier, if the beliefs are such that $\Omega$ is nonempty, then it must be true that $v \in \Omega$. This directly follows from the specification of the model; $x_2 = v$ is one of the program types that, if advertised, provides the TV station with the highest possible audience size. By Lemma (5), if some viewers with $\lambda > v$ watch the first program, then it must be true that

$$\int_0^{\lambda+v+c} u(\lambda, x) \mathbb{1}[x \notin \Omega] \, dx - (1 - (\lambda + v + c)) c < 0$$

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

We can also be certain that no one 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 the first period audience would watch the second program.

To characterize the no tune-in equilibrium, suppose beliefs of people are given by $\Omega = [\epsilon, v - \epsilon]$, where $\epsilon$ is very small. The second period decisions of the first period audience conditional on being exposed to no tune-in is determined by the sign of

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

Those people for whom the left-hand side is nonnegative watch the second program. It is easy to show that this condition is satisfied for $\lambda + c \geq \sqrt{2(1-v)} c$ when $\epsilon = 0$. So,
conditional on $q = 0$, the total mass of people 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 these people would watch the second program. Letting $\varepsilon \to 0$, if $s_2(x_2 \mid q = 1) - s_2(x_2 \mid q = 0) < \frac{s_1}{A}$ for every $x_2 \in [0,v]$, then there is no reason for people to believe that $\Omega$ is nonempty. This inequality is satisfied for every $x_2 \in [0,v]$ when $c$ is such that $\sqrt{2(1-v)c} - c < \frac{v}{2}$. 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 $\Omega = \emptyset$, when $c < c_1$. Intuitively, when the sampling cost is sufficiently low, the TV station has no incentive to advertise his program, since everyone will find out about it anyhow.

Now suppose that $c \geq c_1$, and $\Omega = \emptyset$. The TV station will confirm this belief if $s_2(x_2 \mid q = 1) - s_2(x_2 \mid q = 0) < \frac{s_1}{A}$ for every $x_2 \in [0,v]$. When $\Omega = \emptyset$, only $\lambda \leq v$ watch the first program, since there are no possible informational gains associated with watching it. If these people receive a tune-in while watching, then all of them will watch the second program as well, although they will think that the TV station made a mistake. If they don’t receive a tune-in, on the other hand, then only those with ideal program types satisfying the following condition will 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 $\Omega = \emptyset$. This condition is satisfied for $\lambda \geq v + c - \sqrt{2((v+c)^2 - c)}$. So, the beliefs will be confirmed when $v + c - \sqrt{2((v+c)^2 - c)} < \frac{v}{\frac{s_1}{A}}$, 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 $\frac{s_1}{A}$. People believe that the TV station will not air a tune-in for any program type unless this inequality is satisfied. For values of $c \geq c_1$ that satisfy this inequality, $q = 0$ for all $x_2 \in [0,1]$, and $\Omega = \emptyset$ 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, when $c_1 < c < c_2$, where $c_2 = 1 - \sqrt{(1-2v)^2 - \left( \frac{v}{A} \right)^2} - 2v$, the cost of watching the first program to obtain information is more costly than just sampling the second program, and therefore no one
with $\lambda > v$ watch it. However, since the sampling cost is relatively low, in the absence of a tune-in, some people with $\lambda \leq v$ also choose to sample the second program. If $c \geq c_2$, on the other hand, the cost of sampling the second program conditional on not watching the first program dominates the cost of watching the first program. All of these equilibria are summarized in the following proposition. The derivation of them is very similar to the ones that were presented in the previous section, and therefore has been omitted.

Proposition 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 people with $\lambda > v$ watch the first program, and no tune-in for the second program takes place.

(ii) When $c_1 \leq c < c_2$, no people with $\lambda > v$ watch the first program, and TV station airs a tune-in for all $x_2 \leq x^*_H$, where $x^*_H = 1 - \frac{(c + v)^2}{2c}$.

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

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

and the TV station airs a tune-in for all $x_2 \leq x^*_H$, where $x^*_H = v + \left(1 - \frac{1}{A}\right) \hat{\lambda}$.

(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 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 (1). The picture below shows how the location of the marginal viewer in the first period, and in turn the equilibrium value of $x^*_H$ change as the sampling cost increases function.

---

8$c_2$ can be found by characterizing the location of 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 $v$. 

---
FIGURE 1

To summarize, in this extended model, two different equilibria may arise depending on the size of the sampling cost. If the sampling cost is low, then the unique PBE exhibits no tune-ins. If it is high, then the unique PBE involves a tune-in for the upcoming program unless the two programs are too dissimilar. For moderate levels of sampling cost, there are beliefs that also support no tune-ins in equilibrium. An important implication of the model is that if the sampling cost not too low, it is rationally anticipated that the TV station will air a tune-in for a very wide range of program types. Therefore no one from the first period audience watches the second program unless they were exposed to a tune-in for it. Another important result is that the TV station airs a tune-in for the program types that make at least \( (1/A) \)% of the first period audience watch the second program.

The structure of the sampling cost used in this paper has the same interpretation as search cost in the literature on search where there is no free recalling. This interpretation...
is suitable for the television industry, since a person misses the first part of the program that she ends up watching if she first opted for sampling the programs in other stations. Comparing the results with the benchmark model, we see that the fear of the chance that an individual will end up watching a bad program till the end causes some people to gather information by watching the first program. This comes at a cost for the TV station. No one from the first period audience keeps watching TV unless they were exposed to a tune-in. When it is possible for people to sample a program for a while, the fear of undergoing large utility losses is not a strong motivating factor. Therefore, people do not bear the direct cost of watching the first program just to alleviate their informational constraints.

5 Extensions

The following is a discussion of two extentions to the model. First, I will discuss the implications of developing a model with both quality and attribute uncertainty. Afterwards, I will describe an adaptation of the benchmark model to include two TV stations. For the sake of exposition, and because of space constraints, I have omitted the derivations of the propositions.

5.1 Quality Uncertainty

Suppose that people are not only uncertain about the attributes of a program, but also about its quality. Any direct information the TV station provides in a tune-in is not reliable, since the low-quality station can mimic the high-quality one. The benchmark model introduced in this paper actually enables me to characterize a separating equilibrium in which people learn the true quality of the program. As an example, suppose that there are only two possible quality levels, high or low, and that people’s utility function is given by $u(\lambda, x) = v_j - |\lambda - x|$, $j = H, L$, where $v_H > v_L$. Suppose the first program is of high quality. A TV station offering a high-quality program in the second period wants to correctly identify himself, since a low-quality TV station has an incentive to mimic him. Because the sampling cost is sunk, some viewers will stay tuned to a low-quality program even after they realize it is of low quality. Based on the results obtained in the benchmark model, a candidate equilibrium is the one in which people anticipate the high-quality station to air a tune-in for a wider range of
program types compared to the low-quality station. Therefore, if \( v_H - v_L \) is not too small, a separating equilibrium for some program types can be found in which the high-quality TV station distinguishes himself by only revealing horizontal information about his upcoming program.

For values of \( c \) that are not very small, there exists a single viewer who is just indifferent between watching and not watching the first program. As before, let the location of this individual be denoted by \( \hat{\lambda} \). The main result is presented in Proposition (3).

**Proposition 3** Suppose that \( v_H - v_L > \frac{\hat{\lambda}}{A} > c \). Then the following constitutes a PBE:

(i) For \( x_2 \leq v_L + \frac{\hat{\lambda}}{A} \), there is no separating equilibrium in quality. Each type airs ‘one’ tune-in.

(ii) For \( v_L + \frac{\hat{\lambda}}{A} < x_2 \leq \left(1 - \frac{1}{A}\right) \hat{\lambda} + (v_L + c) \), the high-quality station airs ‘two’ tune-ins, while the low-quality station airs ‘one’ tune-in for \( v_L + \frac{\hat{\lambda}}{A} < x_2 \leq \left(1 - \frac{1}{A}\right) \hat{\lambda} + v_L \) and airs none for \( x_2 > \left(1 - \frac{1}{A}\right) \hat{\lambda} + v_L \).

(iii) For \( \left(1 - \frac{1}{A}\right) \hat{\lambda} + (v_L + c) < x_2 \leq \left(1 - \frac{1}{A}\right) \hat{\lambda} + v_H \), only the high-quality station airs ‘one’ tune-in.

(iv) For \( x_2 > \left(1 - \frac{1}{A}\right) \hat{\lambda} + v_H \), no tune-in by either type takes place.

These strategies satisfy individual rationality and incentive compatibility constraints for both types. In case (i), a high-quality station cannot distinguish himself. This is because both types end up with the same amount of audience size for the specified values of \( x_2 \). However, when \( x_2 \) takes on the values given in case (ii), the high-quality station can signal his quality by burning money. The reason for why the high-quality station airs ‘two’ tune-ins comes from the specification of the benchmark model, and from the assumption given in Proposition (3). More specifically, if \( \frac{\hat{\lambda}}{A} < c \), then the same result holds for a higher number of tune-ins. In general, money burning is expected to happen whenever the low-quality station can mimic the high-quality one. This offers an explanation as to why there are relatively more tune-ins for some program types.

Finally, and most strikingly, for some program types, the high-quality station is able
to signal his quality by airing only one tune-in. This result is similar to the one in the benchmark model. Intuitively, when the program type advertised is very unfavorable to the first-period audience, it has to be the high-quality station who advertised it. So, for some program types, a high-quality station does not need to engage in “money burning” in order to signal his type. These results are based on a model which assumes that people can perfectly detect the quality level of a program when they watch it and can turn their TV off if they desire.

5.2 Duopoly

When there are multiple TV stations, each station may act strategically depending on the type of the program he is offering and the types of others’ programs. Suppose there are two TV stations, Y and Z, whose upcoming programs may be located at either of the three points, 0, $\frac{1}{2}$ and 1, with equal probabilities. Suppose also that their first period programs are such that people with $\lambda < \frac{1}{2}$ watch Y while the remaining people watch Z. All the other specifications remain the same. I propose two candidate equilibria in this setting, one strategic and one non-strategic. Which one we observe in equilibrium will depend on the size of the sampling cost. In the first one, station Y’s (station Z’s) strategy is given as follows: air a tune-in for $x_Y = 0$ (for $x_Z = 1$) regardless of the other station’s program, air a tune-in for $x_Y = \frac{1}{2}$ (for $x_Z = \frac{1}{2}$) unless the other station’s program is located at $x_Z = 1$ (at $x_Y = 0$), and never air a tune-in for $x_Y = 1$ (for $x_Z = 0$). Intuitively, the strategy of not advertising the upcoming program may arise in two situations: (1) when advertising a program does not make enough people stay tuned, and (2) when a station knows that those who switch to the other station as a result of not seeing a tune-in will come back. The former happens for station Y when $x_Y = 1$, and for station Z when $x_Z = 0$. The latter happens when the sampling cost is low. Low sampling cost ensures that some viewers will come back after sampling the program at the other station. When the sampling cost is high, a station can no longer expect the people who initially switched to the other station to come back.

**Proposition 4** The following constitutes a PBE when there are two TV stations, Y and Z: (i) If $Ac < \frac{1}{2}$, then station Y (station Z) always airs a tune-in for $x_Y = 0$ (for $x_Z = 1$)
regardless of the other station’s program, airs a tune-in for \( x_Y = \frac{1}{2} \) (for \( x_Z = \frac{1}{2} \)) unless the other station’s program is located at \( x_Z = 1 \) (at \( x_Y = 0 \)), and never airs a tune-in for \( x_Y = 1 \) (for \( x_Z = 0 \)).

(ii) If \( Ac \geq \frac{1}{2} \), then station \( Y \) (station \( Z \)) always airs a tune-in for \( x_Y = 0 \) and \( x_Y = \frac{1}{2} \) (for \( x_Z = \frac{1}{2} \) and \( x_Z = 1 \)) regardless of the other station’s program, and never airs a tune-in for \( x_Y = 1 \) (for \( x_Z = 0 \)).

If the sampling cost of viewers is sufficiently low, then the tune-in decision of a station strategically depends on the type of the other station’s program. A TV station chooses to act strategically only if it is sure that a big majority of people who initially switched to the other station will come back. Otherwise, both stations act non-strategically in equilibrium, in which case the monopoly outcome arises. This finding suggests that, the higher the number of TV stations, the less likely their preview advertising decisions are to depend on each others’ program types. With a high number of TV stations, the chances that a switching viewer will ever come back is much low. As a result, the equilibrium tune-in decision of each TV station is expected to be very similar to the one derived in the single TV station case when there are many TV stations.

6 Conclusion

In this paper, I have presented a model that explains the incentives of a firm to provide information about its product when the potential buyers are rational decision makers. Rationality of people plays a crucial role in the derivation of the equilibrium. It implies that the decision of a firm to not provide information actually reveals useful information to people. This point has largely been ignored in the previous literature. By incorporating this, the current paper is a first 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 the program characteristics. In addition, although the TV industry was used as an example, the results presented are generalizable.

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 size of the sampling cost viewers incur. If the sampling cost is sufficiently low, the unique PBE exhibits no tune-ins. For higher levels of sampling cost, the TV station airs a tune-in for its upcoming program unless it is too dissimilar with the program during which it would air the tune-in. This equilibrium is unique if the sampling cost is sufficiently high.

The model also enables me to analyze quality signaling when viewers are uncertain about the quality of the upcoming program. ‘Money burning’ is not always necessary for a TV station to signal high quality. This is because there are program types that only a TV station with a high-quality program can afford to advertise. Lastly, when there are two TV stations in the market, the advertising decision of a station may strategically depend on the type of the rival station’s upcoming program. Not airing a tune-in (thus potentially causing some people to switch to the other station) is optimal when a station knows that those who may switch to the other station will not like the program there and therefore will come back. This strategy constitutes a PBE only if the sampling cost is not too high. When the sampling cost is high, tune-in decision of a station is independent from the type of the rival station’s upcoming program. This suggests that strategic preview advertising is not likely to arise in equilibrium if the number of stations is higher.

Some restrictive assumptions have been made in the analysis. Particularly, it was assumed that horizontal attributes of programs can be described by a one-dimensional variable. In reality, it is more probably that TV programs are differentiated along several 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.
Appendix

Proof of Lemma 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 some people who did not watch the first program. These people base their decisions on their beliefs and therefore their magnitude is independent from \( x_2 \). Let \( \Delta = \{ \lambda > v \mid B(\lambda) \geq C(\lambda) \} \), i.e. 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. \( \Delta \) is determined by \( \Omega \) 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 \( \Delta \) 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 from \( x_2 = 0 \) to \( x_2 = (\max(\Delta) - v) \), attains its maximum at \( x_2 \in [(\max(\Delta) - v), v] \), and starts monotonically decreasing at \( x_2 = v \). Hence, conditional on existence of a PBE, \( s_2(x_2 \mid q = 1) \) could 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. Assume on the contrary that \( \Omega = \{ [x_L, x_H] \mid x_L > 0 \} \). Since \( x_2 = 0 \) does not lie in \( \Omega \), 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 the following condition:

\[
s_2(x_2 \mid q = 1) - s_2(x_2 \mid q = 0) \geq \frac{\hat{\lambda}}{A} \text{ for all } x_2 \in [x_L, x_H] \text{ with strict equality if } x_2 = x_L, x_H
\]

Since \( s_2(x_2 \mid q = 0) \) and \( \frac{\hat{\lambda}}{K} \) 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 value at the values of \( x_2 \) over \([\hat{\lambda} - v, v]\). We are now ready to present the formal proof.

Step 1:

\[
E [u(\lambda, x_2 | x_2 \notin \Omega)] \geq 0 \iff \int_0^1 u(\lambda, x) \, dx - \int_{x_L}^{x_H} u(\lambda, x) \, dx \geq 0
\]

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

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

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

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, \( E [u(\hat{\lambda}, x_2 | x_2 \notin \Omega)] \geq 0 \), then \( \int_0^1 u(\hat{\lambda}, x) \, dx \) must be positive, too.

Suppose \( E [u(\hat{\lambda}, x_2 | x_2 \notin \Omega)] \geq 0 \). Remember that \( \hat{\lambda} \) is the solution to \( B(\lambda) - C(\lambda) = 0 \) Then

\[
B(\hat{\lambda}) - C(\hat{\lambda}) = \int_{\hat{\lambda} - v}^{x_H} (v - (\hat{\lambda} - x)) \, dx - \int_{x_L}^{x_H} (v - (\hat{\lambda} - x)) \, dx - (\hat{\lambda} - v) = 0
\]

\((\hat{\lambda} - v)\) is always at least as big 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 \( \Omega = \{[x_L, x_H] | x_L > 0\} \) constitutes an equilibrium that is consistent with TV station’s optimal tune-in decision. Therefore, it must be true that \( E [u(\hat{\lambda}, x_2 | x_2 \notin \Omega)] < 0 \).

Step 2: Given that \( x_H + x_L = \hat{\lambda} \), it can be shown that \( \frac{dE [u(\lambda, x_2 | x_2 \notin \Omega)]}{d\lambda} > 0 \) for \( \lambda < \hat{\lambda} \). We have found that for \( \Omega = \{[x_L, x_H] | x_L > 0\} \) to be consistent with TV station’s optimal tune-in decision, \( E [u(\hat{\lambda}, x_2 | x_2 \notin \Omega)] < 0 \). This implies that \( E [u(\lambda, x_2 | x_2 \notin \Omega)] < 0 \) 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 much bigger \( \hat{\lambda} \) is relative to \( v \), \( s_2(x_2 | q = 1) - s_2(x_2 | q = 0) \) can range between \( v \) and \( \hat{\lambda} \). So, from the equilibrium condition \( s_2(x_2 | q = 1) - s_2(x_2 | q = 0) \geq \frac{\hat{\lambda}}{A} \); 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 TV station’s optimal tune-in decision unless \( x_L = 0 \). \( \blacksquare \)
Proof of Lemma 3. It will be enough to show that \( x_2 = \hat{\lambda} + v \) does not belong to \( \Omega \) in equilibrium. Note that,

\[
s_2(\hat{\lambda} + v | q = 1) = \lambda^o - \max \left\{ \hat{\lambda}, 1 - \lambda^o \right\}
\]
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 \( s_2(\hat{\lambda} + v | q = 1) - s_2(\hat{\lambda} + v | q = 0) \leq 0 \), so that \( x_H \geq \hat{\lambda} + v \) can never happen in equilibrium. ■

Proof of Lemma 4. Suppose, on the contrary, that some people keep watching. Then \( \lambda = \hat{\lambda} \) has to be one of them. If \( \int_0^1 u(\hat{\lambda}, x)dx > 0 \), then the condition \( B(\hat{\lambda}) = C(\hat{\lambda}) \) is expressed as

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

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 left-hand side is \( \frac{(\hat{\lambda} - v)^2}{2} \). But \( (\hat{\lambda} - v) \) is always greater than \( \frac{(\hat{\lambda} - v)^2}{2} \). So, if \( \int_0^1 u(\hat{\lambda}, x)dx > 0 \), \( E[u(\hat{\lambda}, x_2 | x_2 \notin \Omega)] \) cannot be positive.

Now, if \( \int_0^1 u(\hat{\lambda}, x)dx \leq 0 \), then the condition \( B(\hat{\lambda}) = C(\hat{\lambda}) \) becomes

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

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

Proof of Proposition 1. When \( x_2 = x_H \) equation (1) becomes \( \hat{\lambda} - (x_H - v) = \frac{\hat{\lambda}}{A} \) since \( s_2(x_H | q = 1) = \hat{\lambda} - (x_H - v) \) and \( s_2(x_H | q = 0) = 0 \) by Lemma (4). Rearranging, we have \( x_H = v + (1 - \frac{1}{A}) \hat{\lambda} \). The value of \( \hat{\lambda} \) is obtained by simultaneously solving equation (1) when \( x_2 = x_H \), and equation (2). Suppose the equilibrium value of \( \hat{\lambda} \) turns out to be less than \( (1 - \lambda^0) \), in which case the second term in equation (1) equals zero. Then, \( \hat{\lambda} \) is given by the solution to

\[
\int_{\lambda - v}^{v + (1 - \frac{1}{A}) \hat{\lambda}} (v - |\hat{\lambda} - x|)dx = \hat{\lambda} - v
\]

24
The integral on the lefthand side equals \((v^2 - \frac{\hat{\lambda}^2}{2A^2})\). Solving the equation for \(\hat{\lambda}\), we obtain

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

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

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

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

\[
\hat{\lambda} = \frac{1 - \sqrt{1 - (v^2 + \frac{1}{2}) (1 - \frac{1}{2A^2})}}{(1 - \frac{1}{2A^2})}
\]

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 \(x^*_H = v + (1 - \frac{1}{A}) \hat{\lambda}\). Solving equation (1) 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 \(v + (1 - \frac{1}{A}) h(x)\) is less than \(x\) at \(x = 1\), and is greater than \(x\) at \(x = 4v - 1\) if \(v < \frac{1}{2 - \frac{1}{A}}\). This is true for almost all values of \(v\) and \(A\). Hence there must be an interior solution to \(x = v + (1 - \frac{1}{A}) h(x)\), and this has to be unique.

Also note that \(v\) is less than \(\frac{1}{2}\), so this equilibrium does not require any off-the-equilibrium path beliefs. ■
Proof of Lemma 5. First observe that 
\[ \int_{0}^{\lambda+v+c} (v - |\lambda - x|) \, dx + (1 - (\lambda + v + c)) (-c) \]
increasing in \( \lambda \) for \( v < \lambda \leq v + c \). When \( \lambda = v \), it equals \( v^2 - \frac{c^2}{2} - (1 - (2v + c)) c \), which can be rearranged as \( \frac{1}{2} (c^2 - 2(1 - 2v) c + 2v^2) \). 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, \( B(\lambda) \geq C(\lambda) \) when

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

Suppose there exists \( v < \hat{\lambda} \leq v + c \) such that \( B(\hat{\lambda}) = C(\hat{\lambda}) \). Then, \( 1 [x \notin \Omega] = 1 \) 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 \( B(\hat{\lambda}) = C(\hat{\lambda}) \) becomes \( \int_{0}^{\lambda-v} u(\lambda, x) (1 [x \notin \Omega] - 1) \, dx = \lambda - v \). However, the left hand side is at most \( \frac{(\lambda - v)^2}{2} \) for any \( \Omega \). This is equal to \( \lambda - v \) only when \( \lambda = v \). So there is no \( v < \lambda \leq v + c \) such that \( B(\lambda) = C(\lambda) \). People with \( \lambda > v + c \) never watch the first program since their first period disutility exceeds the sampling cost. ■
References


