The Medium of Advertising

Song Lin*

Hong Kong University of Science and Technology

January 6, 2022 – Version 2.3

Preliminary. Please do not circulate without permission.

Abstract

This article analyzes the role of media in advertising communication. A media platform uses content to attract consumers and must then decide whether and when to expose them to ads. A consumer must decide, given her limited attention, what she will pay attention to at each point in time. Advertising can deliver a product match message stochastically if the consumer is paying attention to the ad. The model allows us to classify media advertising formats into three broad classes: static, sequential, and interactive. For each class, we characterize the optimal consumer attention and media allocation policy as a function of the informativeness of the ad, the values it brings to both the consumer and advertiser, and the value of media content. The analysis identifies the conditions under which the interactive format can be strictly superior to the other formats. It also identifies when advertising can raise or reduce consumer welfare relative to the no-ad benchmark, providing a micro-foundation for understanding the externality of advertising in different media markets. The managerial implications of the findings for media platforms and advertisers are discussed.

Keywords: *advertising, media, information design, two-sided market, limited attention*

*The author thanks Anthony Dukes, Jiwoong Shin, Aniko Öry, and Birger Wernerfelt for helpful discussions and suggestions. This paper has also benefited from the comments of seminar participants at Peking University, Johns Hopkins University, City University of Hong Kong, Xiamen University, Yale University, and University of Science and Technology of China.

1. Introduction

Research on the economics of advertising can be broadly partitioned into two streams. The first and more traditional stream focuses on the persuasive, informative, and complementary effects of advertising, but does not consider how these effects may be endogenously influenced by the type of media in which the ads are embedded. For example, on media platforms that are competing with advertisers for consumers' attention, an ad, though potentially informative, may not even get noticed by the consumers.

The second and more recent stream of research considers media platforms as two-sided markets, but takes the effects of advertising as given. This literature tends to focus on whether market provision of advertising is efficient. Advertising is assumed to be, somewhat arbitrarily, either a distraction (nuisance) or an attraction to consumers. However, either view fits well with some media but poorly with others. The nuisance view of advertising, which implies that consumers prefer to not pay any attention to ads, is incompatible with the traditional literature that treats advertising as weakly enhancing consumer utility.

In this paper, we present a synthesis of the two approaches by analyzing the role of media in the transmission of information between advertisers (sellers) and consumers (buyers). The analysis recognizes two important forces in the information chain: consumers' incentive to pay attention to ads and media platforms' incentive to influence the attention. The basic question is: how should media allocate resources to balance consumers' attention between ads and media content?

To address this question, we formulate the media problem as a dynamic information design problem and illustrate the basic economics using a rudimentary model of one media platform, one consumer, and one advertiser. The media platform provides informational or entertainment content that is valuable to a consumer. It can profit from advertising and thus must decide when to expose the consumer to an ad. At any point, the consumer can decide whether to pay attention to the available media content or to the ad but cannot pay attention to both at the same time. An outside option with constant value is always accessible to the consumer. The ad can inform the consumer of the existence of the advertised product and whether it is a match to her. However, communication takes time. Even when the consumer is paying attention to the ad, the match signal arrives at a random time. It is likely that after a period of time, the consumer learns nothing about the match and thus no transaction takes place. None of the parties possesses superior information about the match value *ex ante*. The media platform, however, can control the exposure time, which ultimately affects the consumer's attention strategy and the signal arrival process. When the match signal arrives, only the consumer obtains it, whereas the media platform and the advertiser remain uninformed.

Due to the differences in communication technologies, the types of media differ substantially in terms of how much control they possess over consumer's attention, and these variations have implications for media policies and consumer welfare. The model naturally permits categorization of media into three general types based on media's control capability. In the first type, a media platform releases its content and ad simultaneously, and thus cannot force a consumer to pay attention to the ad at a particular point in time. The consumer can flexibly determine what to pay attention to. This advertising format is called *static advertising*, and fits well with print media such as newspapers, magazines, and online websites with display ads. In the second type, a media platform can choose when to expose a consumer to an ad or to content, but cannot display both at the same time. This flexibility allows the media platform to show the ad exclusively at specific times, as in media like television and radio. We call this advertising format *sequential advertising*. In the third type, a media platform can release either the ad, the content, or both, at any point in time. The important feature here is that the platform can, after showing the ad for some time, "grant" a consumer an option to skip it. Essentially, such advertising has an interactive feature and thus is termed *interactive advertising*. Many online media platforms such as video streaming platforms that allow viewers to skip ads fall into this category.

In Section 3, we characterize the equilibrium behaviors under each of the three media formats. Under static advertising, a media platform simultaneously introduces its content and an ad to a consumer, who then decides how to allocate her attention. Paying attention to the ad activates the signal arrival process. The consumer then needs to decide whether to pay attention to it and for how long. This is an infinite-horizon optimal stopping problem and the optimal solution to it is a simple threshold strategy. The consumer pays attention to the ad whenever the content value is below the threshold λv_c, where λ is the hazard rate of the exponential process and v_c is the consumer's expected surplus of buying the advertised product conditional on a match. She pays attention to the ad until the signal arrives.

Under sequential advertising, a media platform can expose a consumer to an ad or its content sequentially. Despite greater control over the consumer's ad attention, the platform suffers from the *efficiency lost* due to its inability to observe whether the consumer has obtained the ad signal. This fundamental tension has led the platform to impose a deterministic time window for the ad exposure. The consumer then faces a *finite-horizon* control problem, which in general is intractable. Owing to the exponential arrival assumption, the solution becomes tractable, which is the same as the one for the infinite-horizon counterpart. Anticipating the consumer's behavior, the platform's optimal policy is shown to be a simple function of the expected match values of advertising, the content value to the consumer, and the hazard rate λ that captures the informativeness of the ad. Not surprisingly, the platform prefers to prolong the ad exposure if the match between the consumer and advertiser can bring in more value to both parties, but prefers to limit the exposure time as the consumer values the content more.

The relationship between the ad exposure time and ad informativeness is more subtle. A larger value of λ implies that the signal is more likely to arrive early, indicating a more informative ad. The optimal exposure time for an ad turns out to be non-monotonic in its informativeness: it first (weakly) increases and then decreases with λ. When λ is small, an increase in its value can increase the chance that the consumer will obtain the match signal, motivating the platform to increase the ad exposure time. However, when λ is large, the signal is more likely to arrive very early and thus the platform has a stronger incentive to reduce the ad's exposure time and introduce the content earlier.

Much of the insight obtained from sequential advertising can be readily extended to interactive advertising. However, interactive advertising allows the media platform to simultaneously offer an ad and the media content, giving the consumer the option of choosing between them. As a result, the platform needs to decide whether and when to introduce this option. If the media content is sufficiently attractive, then the consumer has no interest in paying attention to the ad even if it is available. This forces the platform to introduce a *minimum run time* to maintain the consumer's attention to the ad. Accordingly, the platform faces the same problem in sequential advertising that it needs to decide when to terminate the ad's run time. In contrast, if the media content is sufficiently unattractive, then the consumer becomes willing to process the ad. It is then optimal for the platform to introduce the option at the start of the interaction, a situation equivalent to static advertising. This optimal strategy allows the platform to avoid the efficiency lost arising in sequential advertising.

Section 4.1 derives the profitability (welfare) implications of the equilibrium characterization of the three media formats. Given its richer set of options, interactive advertising defines the upper bound for the performance of advertising in any media format. In particular, if the media content is less attractive to consumers *relative* to an ad, then interactive advertising strictly dominates sequential advertising because the former allows consumers to pay as much attention to an ad as they choose, whereas the latter imposes a restriction on the ad exposure time. If the media content is moderately attractive, then interactive advertising is equivalent to sequential advertising because in both cases a media platform should restrict the ad exposure for a finite time period and then present consumers with the media content. This is not feasible in static advertising, under which consumers may avoid the ad altogether. Lastly, if the media content is very attractive to the consumers, then there is no reason for the platform to pursue advertising, and all three formats are equivalent. These findings shed light on when interactive advertising can be a more profitable format, and when static or sequential advertising can perform as well as interactive advertising.

The model provides a simple and intuitive explanation for the phenomenon of ad annoyance. Both practitioners and academic researchers have long noticed two puzzling observations. First, consumers often complain that ads are distracting, a waste of time, or irrelevant, despite their informational values. Second, these accusations vary substantially across different media, with more complaints being raised about ads on television than in newspapers or magazines. The analysis in Section 4.2 resolves both puzzles. Static advertising, such as newspaper ads, does not bother consumers very much because they can flexibly switch their attention to the media content. However, media platforms that adopt a sequential advertising format, like television commercials (or even some cases in interactive advertising), show an ad exclusively for a period of time. Although consumers may not pay any attention to the ad, they cannot immediately resume content consumption, as it is only available only after the ad exposure is terminated. This may be construed as an opportunity cost to the consumers, causing annoyance. However, this negative externality only arises if the media content is moderately attractive. Essentially, the model provides an explanation for the nuisance cost or negative externality of advertising that many analyses have simply taken for granted (e.g., Dukes and Gal-Or 2003, Dukes 2004, Anderson and Coate 2005, Anderson and Gans 2011).

In a broader sense, this study helps to improve our understanding of the effects of media formats on advertising communication. This has been particularly relevant given the growing sophistication of media formats fueled by the growth of digital technology and the growing dependence of media platforms on advertising revenue. Profit-driven media need to understand how to design advertising formats and how to manage the trade-off between ad exposure and content exposure at a more granular level. These decisions can be heavily influenced by consumers' attention behavior. Advertisers need to carefully select the most effective media formats for their advertising campaigns. One of the challenges is the measurement of advertising effectiveness, which has received abundant attention in the empirical literature. The strategic behaviors of both media and consumers may further complicate the measurement task (Gordon et al. 2021). There is a need for research into the interactions between these forces.

Related Literature

There is a vast literature on whether and how advertising affects consumers and how it interacts with other marketing variables.[1] However, relatively few efforts have been devoted to understanding how the format of media affects advertising communication. More recently, researchers have began to study media advertising in the context of of two-sided markets (e.g., Armstrong 2006, Rochet and Tirole 2006, Weyl 2010) – that is, media platforms seek to attract both advertisers and consumers and benefit from their interactions. Much of the focus has been on the pricing problem, particularly on how pricing policies are shaped by the cross-side externality. In contrast, the focus of this paper is the information design problem of media platforms, a problem that has received relatively little attention. As noted above, the theory developed in this paper can rationalize the cross-side externality that has been assumed in the extant work.

The optimal advertising allocation problem has a long history in marketing, economics, and operations research. A classic approach looks at the optimal control problem of advertising spending over time (e.g., Vidale and Wolfe 1957,

[1] A comprehensive review here is unnecessary and beyond the scope of this paper. There are several excellent reviews with different focuses (e.g., Bagwell 2007, Renault 2015).

Nerlove and Arrow 1962, Little 1979, Mahajan and Muller 1986). Sethi [1977] and Feichtinger et al. [1994] provide a comprehensive review of this line of research. Although our theory similarly builds on a continuous-time control problem, the approach taken is rather different. We explicitly model the interaction process between advertisers and consumers, instead of assuming an aggregate model that relates advertising spending to product sales. Thus, while the literature speaks to the problem of scheduling advertising over long periods (e.g., weeks or months), our model focuses on advertising allocation decisions for a single ad in a short period. More importantly, extant studies are largely motivated by the problem of advertising spending on television, and thus are unable to provide guideline for advertising in the digital age, which features both targetability and interactivity. The theory developed in this study fills this gap and provides insight into how various media formats influence advertising. Methodologically, the proposed theory solves both the media and consumers' control problems simultaneously, in contrast to studies that focuses on the firm's problem without a micro model of consumers.

The information design problem studied here also connects two streams of research in information economics. The first stream investigates how individuals (e.g., consumers, job seekers) actively acquire information over time, but assumes information is exogenously given. The seminal papers by McCall [1970] and Weitzman [1979] have inspired extensive work on optimal information search. The second stream focuses on how an individual (sender) persuades another (receiver) to change her actions within a symmetric information framework but assumes that receivers passively update beliefs given the information provided by senders. The static Bayesian persuasion framework was first introduced by Kamenica and Gentzkow [2011] and has subsequently been extended to dynamic settings in which senders can control the flow of information over time (e.g., Ely et al. 2015, Ely 2017). However, even in these dynamic settings, receivers are often assumed to be myopic with no dynamic incentive to acquire information. This study integrates both research streams in the analysis of a media market: media control the flow of information with the intention of influencing consumers, who in turn have dynamic incentive to acquire information.

2. The Model

We model a simple market with one media platform, one consumer, and one advertiser. The platform can produce media content M continuously over an infinite time horizon $T = \infty$, bringing information or entertainment value to the consumer. In reality, media content can take a variety of forms, such as news articles, television programs, online videos, music streaming, news feeds on social media, etc. In general, the flow utility of the consumer's content consumption may vary over time. Without loss of generality, we assume that the consumer values the media content at a constant rate: $m(t) = m \geq 0$ for any time t.[2]

In addition to media content, the platform can expose the consumer to an ad A sponsored by the advertiser. In theory, at any time t, the full set of options for the platform's decision $d(t)$ is $\mathcal{D} = \{A, M, \{A, M\}\}$, where the subset $\{A, M\}$ implies that the platform allocates both advertising and media content simultaneously to the consumer.[3] In practice, however, there are technological constraints on the platform's actions. Thus, the platform's actions are often confined to a subset of $\mathcal{D}$. This observation suggests an approach that classifies advertising based on the media formats.

Definition 1 (A Taxonomy of Media Advertising Formats):

1. ***Static Advertising***: $\mathcal{D}_1 = \{\{A, M\}\}$. *The media platform can only run the ad and the media content simultaneously. The consumer can choose to process one of the two. Example: print media ads.*

[2] Section 5 discusses an extension of the model with time-varying content.

[3] Here, when both the ad and media content are bundled, it does not mean that the consumer will process both. Instead, she needs to choose which one to process due to limited attention, a point discussed in more detail below.

2. ***Sequential Advertising**: $\mathcal{D}_2 = \{A, M\}$. The media platform can run either the ad or the content in any sequence but cannot run them simultaneously. Example: television and radio commercials, online video ads.*
3. ***Interactive Advertising**: $\mathcal{D}_3 = \{A, M, \{A, M\}\}$. The media platform can run the ad and the content either simultaneously or sequentially at any point in time. Example: online skippable ads like YouTube's "Trueview" ads.*

As will be shown in the following analysis, this taxonomy helps us to understand the fundamental differences in the behaviors of media and consumers under different advertising formats. It is worth pointing out that ads on many online or digital media platforms are not necessarily interactive. For example, some websites display static banner ads that are combinations of text and images. On other websites, the banners ads may show a short video and have sequential or interactive characteristics. The variability of online media calls for a deeper understanding of the nature of ad formats, which is precisely the focus of the main analysis.

Whenever the consumer decides to pay attention to the running ad, the ad information process is activated. This information process must be distinguished from the ad exposure process, as exposure does not guarantee that the consumer will pay attention to the ad. Thus, it is useful to define the ad process time x, which is the time the consumer takes to process the ad. Note that this process time is endogenously related to the universal timeline t and hence the function $x(t)$ will be used in the analysis. The role of advertising is both to inform the consumer of the existence of the advertised product (or the advertised brand more generally) and to provide her with information about whether the product is a match to her. The information, or the signal, is stochastic in two dimensions. First, it arrives at a random time $\sigma \geq 0$, which follows an arrival distribution $F_\sigma(x) \equiv Pr(\sigma \leq x)$. To keep the analysis tractable, the signal arrival process is assumed to follow an exponential distribution with parameter λ. This parameter captures the *informativeness* of the ad: for a fixed time period, an ad with a higher value of λ is more likely to produce the signal. Second, upon arrival, there is a probability $\rho > 0$ the signal tells the consumer that there is a match, and a probability $1 - \rho$ that the signal indicates that there is no match. Conditional on a match, the one-off benefit to the consumer is $\tilde{v}_c \geq 0$ and to the advertiser is $\tilde{v}_a \geq 0$.[4] To simplify notation, let $v_c \equiv \rho\tilde{v}_c$ and $v_a \equiv \rho\tilde{v}_a$ denote the expected ad values to the consumer and the advertiser conditional on the signal arrival. Unlike the signal arrival process that speaks to the efficiency of the ad, the realization of the match value captures the effectiveness of the ad. Further, the arrival of the signal and the realization of the match value are independent. This setup allows us to keep the decision problems of the consumer and the media platform tractable.

It is assumed that at any point in time the consumer can access to an outside option B, which captures any option available other than the media content or the ad. This option may generate reward (possibly stochastically) to the consumer over time. Examples of an outside option include a bathroom break while watching a television program, chatting with a friend while reading a newspaper, or checking emails while watching an online video. This outside option can be viewed as the opportunity cost of paying attention to the ad. As such, it can be interpreted as a search cost.[5] To keep the analysis simple, it is assumed that this option yields a constant reward $b(t) = b$. To avoid triviality, we assume that $b < m$. In later part of the analysis, b will be further normalized to zero.

The consumer has a unit budget of attention to allocate. At any point in time, she chooses an attention action $c(t)$ from a choice set $\mathcal{C}(t)$ that is partially determined by the media platform. Formally, $\mathcal{C}(t) = \{B, d(t)\}$. For example, if the platform runs the ad at time t, $d(t) = A$, then the consumer can choose to

[4]The assumption implies that if the signal does not arrive and thus the consumer remains uninformed about the value of the advertised product, the consumer will not make a purchase. This can occur if the prior belief about the product value is sufficiently low, or if the alternative way of acquiring information is too costly.

[5]As argued in Nelson [1974], the marginal cost to a consumer looking at an ad is primarily a time cost. "This time cost will vary by the alternative use of the time used in watching the advertisements." (p.745, Nelson 1974)

either pay attention to the ad (i.e., $c(t) = A$), or take the outside option (i.e., $c(t) = B$). In both cases, she does not consume the media content. If, however, $d(t) = \{A, M\}$, then she has the additional option of skipping the ad and consuming the content (i.e., $c(t) = M$).

The consumer discounts future rewards. The standard approach is to assume that the present value of a future reward at time t is discounted exponentially by a factor of e^{-rt}. Here we adopt the interpretation that the discounting is due to a random termination of the problem which, if occurs, results in no ad processing and content consumption. This stochastic approach simplifies the analysis because all of the players share the same discount factor, which leads to a simple formula for the optimal ad allocation.

Note that the media platform does not observe either the arrival of an ad signal or the realization of the match value. In general, it does not observe the consumer's action.[6] This information structure fits well with the reality — newspapers, magazines, television and radio producers do not know whether a particular consumer has paid attention to an ad. Given this information structure, the platform's action at time t is independent of the history of the consumer's actions prior to time t. Hence, under Model $\mathcal{D}_k$, the platform chooses the action $d(t) \in \mathcal{D}_k$ at time t; the consumer then selects an action $c(t) \in \mathcal{C}(t)$ based on the information she has accumulated as of time t.

An equilibrium solution requires simultaneously solving both the platform's allocation problem and the consumer's attention problem. The equilibrium behaviors, of course, depend on the media format. These behaviors under the different formats are the subject of the analysis in the remainder of the paper.

3. Consumer Attention and Media Policy

[6] In the case of interactive advertising, a media platform can observe a consumer's choice if it offers her the choice of an ad or content. However, in this scenario, the observation does not change the platform's incentive.

3.1 Model $\mathcal{D}_1$: Static Advertising

Under this advertising model, the consumer can decide freely whether to pay attention to the ad or to the media content. Following the assumption that $d(t) \in \mathcal{D}_1 = \{A, M\}, \forall t$, the media platform has limited capability to divert the consumer' attention to the ad. Its action, once determined at the start of the game, is fixed over time. It remains to analyze the optimal strategy of the consumer given this information environment. The analysis also lays the foundation for analysis of more complex advertising models. Lemma 1 below presents the result.

Lemma 1 (Consumer Attention under Static Advertising):

1. *If $\lambda v_c \leq m$, the consumer pays attention to the media content only.*
2. *If $\lambda v_c > m$, the consumer pays attention to the ad first and consumes the content after the ad signal has arrived.*

PROOF: As the consumer has access to both the outside option and the media content at any decision time and $b < m$, she will never choose the outside option. The consumer problem is reduced to an optimal control problem with two options A and M. This is in fact a two-armed bandit problem with a safe arm M and an uncertain arm A.[7] The well-known solution to this problem is the Gittins Index solution (Gittins 1979). Before the arrival of the match signal, we can define the index of the ad process as a function of the state $x(t)$, which is the process time of the ad:

$$G_A(x) \equiv \sup_{t>x} \frac{v_c \int_x^t f(s)e^{-rs}ds}{\int_x^t [1-F(s)]e^{-rs}ds} = \lambda v_c. \tag{1}$$

The equality of the above equation follows from the distributional assumption that F is exponential. Thus, the index value does not depend on the process

[7] The uncertain arm A defined in the model is equivalent to a *job process* in the job scheduling literature in operations research.

time as long as the signal has not arrived, simplifying the solution: the consumer processes the ad if only if the ad index is greater than the reward of the safe arm, $\lambda v_c > m$. Once the signal has arrived, the uncertainty is resolved and thus the consumer consumes the content thereafter. ■

If $\lambda v_c \leq m$, following the optimal strategy, the consumer expects a total value of

$$V^*_{\mathcal{D}_1} = \int_0^\infty m e^{-rt} dt = \frac{m}{r}. \tag{2}$$

As the platform extracts all the consumer surplus, its profit is $\Pi^*_{\mathcal{D}_1} = V^*_{\mathcal{D}_1} = m/r$. In contrast, if $\lambda v_c > m$, the consumer's total expected value following the optimal strategy is

$$V^*_{\mathcal{D}_1} = \int_0^\infty (v_c + \frac{m}{r}) e^{-rt} dF(t) = \frac{\lambda(r v_c + m)}{r(\lambda + r)}. \tag{3}$$

The media platform collects both consumer and advertiser surplus, earning

$$\Pi^*_{\mathcal{D}_1} = \int_0^\infty (v_c + v_a + \frac{m}{r}) e^{-rt} dF(t) = \frac{\lambda(r(v_c + v_a) + m)}{r(\lambda + r)}. \tag{4}$$

To profit from advertising, the platform needs to ensure that the profit expressed in Equation (4) is greater than the ad-free profit m/r. This requires that $\lambda(v_c + v_a) > m$, a condition that can be satisfied as long as $\lambda v_c > m$. The optimal policy for the platform can then be summarized as follows.

Proposition 1 (Media Policy under Static Advertising): *Under Model $\mathcal{D}_1$, the optimal media policy is to introduce the ad only if $\lambda v_c > m$, and offer no-ad content otherwise.*

For the consumer to process the ad and the platform to profit from it, the ad needs to be sufficiently informative and valuable, $\lambda v_c > m$. This is a rather strong restriction under static advertising. For a slightly less informative or valuable ad that satisfies $\lambda v_c < m < \lambda(v_c + v_a)$, it would be in the platform's interest to introduce advertising as long as the consumer is willing to process it. Sequential advertising addresses these limitations.

3.2 Model $\mathcal{D}_2$: Sequential Advertising

Under this class of model, media have more control over consumer attention. Specifically, at any point in time, a media platform can choose to either run an ad or offer the media content, but not both simultaneously. Consumers' attention decisions thus depend on how the platform allocates the two. Next, we first analyze the optimal strategy of a consumer under any arbitrary policy, and then solve for the optimal media policy.

The sequential nature of the media format raises the issue of commitment problem. However, because the media platform does not observe the consumer's action, the commitment problem becomes irrelevant here. Hence, the media strategy has the simple form of history-independent allocation of advertising. Consider an arbitrary allocation policy π of a media platform that runs an ad for a total duration of $S = \int \mathbb{1}\{d(t) = A\}dt$. Note that the class of policies with the same ad duration S is quite large. A policy might specify that the ad is interrupted at some point in time and then resumes at a later point. Lemma 2 summarizes the optimal strategy of the consumer.

Lemma 2 (Consumer Attention under Sequential Advertising):

1. *If $\lambda v_c \leq b$, the consumer always chooses the outside option when the ad is running, and consumes the content whenever it is available;*
2. *If $\lambda v_c > b$, the consumer pays attention to the ad when it is running until the match signal arrives or the ad terminates, whichever comes first; she consumes the content whenever it is available.*

PROOF: Given any arbitrary allocation policy π with ad exposure time S (possibly infinite), let us consider the reduced problem π_0: the ad is run *continuously* from start for a duration of S without interruption, followed by the provision of media content thereafter. That is,

$$\pi_0 : \quad d(t) = \begin{cases} A & \text{if } t < S; \\ M & \text{if } t \geq S. \end{cases}$$

The consumer again faces a bandit problem similar to that in Model $\mathcal{D}_1$ but with two important distinctions. First, there is a fixed time window S for the ad process. Second, the media content is not available when $t < S$ but becomes the only available option when $t \geq S$. These two features together suggest that the problem can be decomposed into two parts: the first part is an optimal control problem with $c(t) \in \mathcal{C}(t) = \{A, B\}$ for $t < S$, and the second part is simply choosing the media content, that is, $c(t) = M$ for all $t \geq S$.

The first part is a *finite-horizon* bandit problem, under which the Gittins Index strategy is not guaranteed to be optimal. However, under the assumption that the ad process follows an exponential process, the flow payoff of paying attention to advertising is non-increasing with probability one, which satisfies the *deteriorating arm* condition (Weber 1992). Then the index strategy remains optimal under a finite horizon.[8] Following the index strategy, the consumer chooses to process the ad if and only if $\lambda v_c > b$.

Note that the original problem under policy π can be viewed as introducing interruptions to the ad process given in the reduced problem π_0, while keeping the total ad exposure time S fixed. During these interruptions, e.g., when the ad is not running, the consumer's only choice is to consume the media content. These interruptions neither affect the state (process time) nor the payoff of the ad process. Thus, following the index strategy for problem π_0 is optimal for the original problem π. ■

Given the optimal strategy of the consumer, the platform can choose an optimal allocation policy to maximize profit. In this simple model, the platform can extract all the value generated by a match between the consumer and the advertiser by charging the consumer a subscription fee and charging the advertiser an advertising fee. However, the platform cannot base its fee on whether the consumer actually pays attention to the ad, because the attention is unob-

[8]Intuitively, in solving the index value as defined in Equation 1, the consumer finds it optimal to stop immediately after the next infinitesimal time interval. This implies that the time horizon appears to be irrelevant in her decision. She acts myopically despite the forward-looking incentive.

servable and thus it is not a feasible basis for a contract (Rochet and Tirole 2006). In essence, the platform is maximizing the total welfare of all parties. The following proposition summarizes the result.

Proposition 2 (Media Policy under Sequential Advertising): *Under Model $\mathcal{D}_2$,*

1. *if $\lambda v_c \leq b$, the media platform does not run the ad; and*
2. *if $\lambda v_c > b$, the optimal media policy is to run the ad continuously from the start for a period of $S^* = \max\{\frac{1}{\lambda} \ln \frac{\lambda(v_c+v_a)}{m}, 0\}$ and to present the media content thereafter.*

PROOF: The first part follows straightforwardly from Lemma 2. To prove the second part, note that if the platform chooses A at time t, then Lemma 2 implies that its flow payoff is $(v_c + v_a)f(x(t))$. If the platform chooses M at any time, the flow payoff is always m. Note that if the platform suspends the ad at some time t and then resumes the ad at a later time $t' > t$, the state does not change. Then the platform's problem is an optimal stopping problem: if it is optimal to choose M at time t, it will continue to do so thereafter. Hence, the optimal policy takes the form of continuously running the ad for a period of S and then running the content thereafter. The optimal S must solve the following problem

$$\max_{S\geq 0} \Pi_{\mathcal{D}_2}(S) = \int_0^S (v_c + v_a)e^{-rs} dF(s) + \int_S^\infty m e^{-rs} ds. \tag{5}$$

The first-order condition is given by

$$\lambda(v_c + v_a)e^{-(\lambda+r)S} - me^{-rS} = 0. \tag{6}$$

It immediately follows that the interior solution is

$$S^* = \frac{1}{\lambda} \ln \frac{\lambda(v_c + v_a)}{m}, \tag{7}$$

which is positive as long as $\lambda(v_c+v_a) > m$. Otherwise, the corner solution $S^* = 0$ applies. ■

Intuitively, the optimal ad exposure time S^* strikes a balance between the values generated by the consumer-advertiser match and the opportunity cost of delaying content provision to the consumer. This optimal ad exposure time is a simple function of the market parameters, allowing us to make sharp predictions about how these parameters influence the time windows for advertising. The following proposition characterizes these relationships.

Proposition 3 (Comparative Statics under Sequential Advertising): *Under Model $\mathcal{D}_2$, the optimal exposure time of advertising*

1. *weakly decreases with the content quality m,*
2. *weakly increases with the expected match values for the consumer and advertiser (v_c, v_a), and*
3. *weakly increases as the ad informativeness λ increases if $\lambda < em/(v_c + v_a)$, but weakly decreases in λ if $\lambda > em/(v_c + v_a)$.*

PROOF: The first two parts are trivial given the solution in Equation 7. The third result is obtained by examining the derivative of the interior solution S^* with respect to λ,

$$\frac{\partial S^*}{\partial \lambda} = \frac{1}{\lambda^2}\big[1 - \ln \frac{\lambda(v_c + v_a)}{m}\big], \tag{8}$$

which is positive if $\lambda < em/(v_c + v_a)$, and negative if $\lambda > em/(v_c + v_a)$. ■

The first two parts of Proposition 3 are fairly intuitive. If the media content becomes more attractive to the consumer, then the platform should reduce the exposure time of the ad. However, if the total value brought about by successfully matching the consumer and advertiser is higher, then the consumer should be exposed to the ad for a longer period to increase her chance of obtaining the match signal.

The third part of the proposition illustrates the nonmonotonic relationship between the optimal exposure time for advertising and the informativeness of the ad captured by λ. Figure 1 provides an illustration. Because λ is the hazard rate of the exponential distribution, the probability of observing the signal, given no signal before time x, is $\lambda = f(x)/(1 - F(x))$, $\forall x$. If λ is very small, an increase in λ can effectively increase the chance of observing the signal. This motivates the platform to run the ad for a longer period. Conversely, if λ is sufficiently large, the ad signal arrives much sooner. In that case, it is more important for the platform to reduce the ad exposure time so that it can introduce the content sooner.

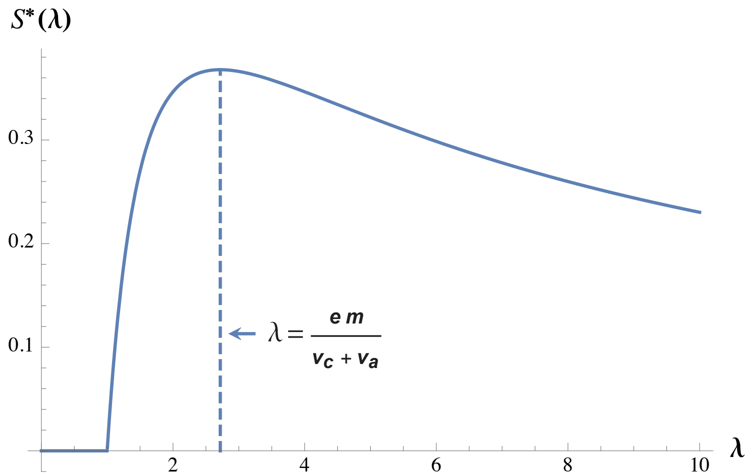


Figure 1: Optimal Ad Exposure as a Function of Ad Efficiency ($v_c + v_a = 1, m = 1$)

Clearly the media platform benefits from the ability to focus the consumer's attention exclusively on the ad. However, in this scenario, the platform needs to decide when to terminate the ad and switch to the media content, without being able to observe whether the consumer has obtained the match signal. The interactive media format analyzed in the next subsection addresses the problem.

3.3 Model $\mathcal{D}_3$: Interactive Advertising

Under the interactive advertising model, the media platform has one additional control option $d(t) = \{A, M\}$, which allows the consumer to choose between processing the ad and consuming the media content. Essentially, interactive advertising integrates both static and sequential advertising by enabling one of them at any time t. Whenever the platform is focusing on sequential advertising, either advertising and content is running exclusively. Again, during the running time, the platform does not directly observe the consumer's action and thus the commitment issue is irrelevant here. However, when the platform switches to static adverting, allowing the consumer to make a choice option between the ad and the content, it can observe the decision outcome. For example, an online video streaming platform can allow users to skip an ad to watch a video. By clicking on the skip option, a user reveals to the platform that she prefers to watch the video content instead of attending to the ad. The observation of the consumer's action implies that the platform can condition its policy based on the consumer's action whenever $d(t) = \{A, M\}$ is introduced.

The commitment problem introduces complexity to the analysis. We first draw the connection to the earlier results on the consumer's problem by assuming that the platform can commit to a policy that is independent of the consumer's decision. That is, whenever $d(t) = \{A, M\}$ is introduced, it specifies for how long this option will last before expiration and commits to it.

Lemma 3 (Consumer Attention under Interactive Advertising with Commitment):

1. *When the platform exclusively runs A or M, the consumer follows the strategy under sequential advertising in Lemma 2.*
2. *When the platform introduces the option $\{A, M\}$, the consumer follows the strategy follows the strategy under static advertising in Lemma 1.*

PROOF: The first part immediately follows by noticing the equivalence of the consumer problem to that under sequential advertising analyzed in subsection 3.2. The second part can be obtained by extending the result in subsection 3.1 to the finite-horizon case and following the same logic adopted in the proof of Lemma 2. ■

Note that the above result does not fully solve the consumer problem because it is restricted to the committed case. However, as we show next, this indeed is the problem that the consumer faces in an equilibrium. The following lemma suggests that the optimal media policy has a simple form that involves commitment.

Lemma 4 (Format of Media Policy under Interactive Advertising): *Under Model $\mathcal{D}_3$, the optimal media policy takes the following form:*

1. *the media platform runs the ad A exclusively from the start for a period of S_m, followed by granting the consumer the optional control $\{A, M\}$ thereafter, and*
2. *the media platform never overturns the consumer's decision.*

PROOF: See the appendix. ■

Lemma 4 suggests that the optimal strategy is featured by a *minimum run time* S_m that forces the consumer to attend to the ad and a *skip* control that allows the consumer to skip through the ad to consume the media content immediately. Indeed, this format is widely used in practice. For example, YouTube's skippable video ads entail a period of five seconds that the an ad must be watched and an option to skip after the five seconds.

To understand the intuition behind Lemma 4, we need to distinguish two cases. First, if $\lambda v_c \leq m$, then the consumer prefers to consume the media content over the ad when both are available. This holds true regardless of whether or not the consumer has already obtained the ad signal prior to the introduction of the skip option. Hence, even after observing that the consumer has chosen the media content, the platform has no information about the signal arrival. The problem to the media platform reduces to an optimal stopping one: when to introduce the skip control, or equivalently, the choice of S_m.

Second, if $\lambda v_c > m$, then the consumer prefers to process the ad over the media content when both are available, as long as the ad signal has not arrived. In this case, the consumer's action can reveal the signal arrival. At the point of S_m when the platform observes that the consumer does not choose A, it infers that she has already obtained the ad signal and thus it is optimal for the platform to let her continue to consume the content. There is no incentive to force the consumer to pay attention to the ad. In contrast, when the platform observes that the consumer continues to choose A after the skip option is introduced, it learns that the signal has not arrived. It is then in the interest of both the consumer and the platform to continue processing the ad until the signal arrival. Any strategy that interrupting this process is suboptimal.

Based on the simplifying results in Lemma 4, we can readily obtain the optimal media policy under interactive advertising, as summarized in the following proposition.

Proposition 4 (Media Policy under Interactive Advertising): *Under Model $\mathcal{D}_3$,*

1. *if $\lambda v_c \leq b$, the media platform does not run the ad, and*
2. *if $\lambda v_c > b$, the optimal media policy is to continuously run the ad from the start for a minimum period of $S_m^* = \frac{1}{\lambda} \ln \frac{\lambda(v_c+v_a)}{m}$ if $\lambda v_c < m < \lambda(v_c + v_a)$, and $S_m^* = 0$ otherwise. After S_m^*, the skip option is introduced.*

PROOF: By Lemma 3, the consumer will pay attention to the ad until the signal arrival or the termination of the ad S_m, whichever comes first. For $t > S_m$, the consumer chooses to skip the ad and consume the media content if $\lambda v_c \leq m$. Hence, the optimal allocation policy is the same as that under sequential advertising: the media platform runs the ad for a fixed time window S_m^* as long as $m < \lambda(v_c + v_a)$.

If $\lambda v_c > m$, the consumer will continue to pay attention to the ad for $t > S_m$ as long as the signal has not arrived. In this case, any platform's choice of $S_m > 0$ will be strictly dominated by $S_m = 0$. To see that, suppose $S_m > 0$. If the ad signal arrives at some time $t < S_m$, then the consumer has no access to the media content and thus can only choose the outside option within the time period $[\sigma, S_m]$. This implies that the platform could have improved the nprofit by letting the consumer choose M within $[\sigma, S_m]$, a contradiction. Hence, it must be that $S_m^* = 0$ in this case. ■

4. Implications for Media Profitability and Advertising Externality

4.1 Media Profitability/Welfare

The preceding analysis reveals that different media formats can influence the communication outcomes between advertisers and consumers. In this section, we explore the implications for the profitability of media platforms, which in our model is equivalent to social welfare.[9] Because interactive advertising (Model $\mathcal{D}_3$) enables the richest set of actions for the media platform, this media advertising format clearly dominates, at least weakly, the other formats. The more interesting question, however, is under what circumstances can static advertising (Model $\mathcal{D}_1$) and sequential advertising (Model $\mathcal{D}_2$) achieve the same profits as the interactive advertising format. Given the proof of Proposition 4, the answer can be readily obtained as follows.

Proposition 5 (Profitability of Advertising Models):

1. $\Pi_{\mathcal{D}_3} = \Pi_{\mathcal{D}_1} > \Pi_{\mathcal{D}_2}$, *if* $m < \lambda v_c$;
2. $\Pi_{\mathcal{D}_3} = \Pi_{\mathcal{D}_2} > \Pi_{\mathcal{D}_1}$, *if* $\lambda v_c < m < \lambda(v_c + v_a)$;
3. $\Pi_{\mathcal{D}_3} = \Pi_{\mathcal{D}_2} = \Pi_{\mathcal{D}_1}$, *if* $m > \lambda(v_c + v_a)$.

If the media content is less attractive, $m < \lambda v_c$, static advertising is equivalent to interactive advertising. This is because matching the advertiser and consumer delivers more value than profiting solely from the content, and thus it is optimal to have the consumer to pay attention to the ad until a signal arrives. Both static advertising and interactive advertising achieve exactly that. In contrast, sequential advertising is strictly dominated by both of these models because of its restriction on the time window for running the ad.

[9] As noted above, in this simple model, the media platform extracts all surplus from both the consumer and advertiser. Hence, the social welfare is equivalent to the platform's profit.

If the media content is moderately attractive, $\lambda v_c < m < \lambda(v_c + v_a)$, sequential advertising becomes equivalent to interactive advertising. This result immediately follows from Proposition 4: it is sufficient for the platform to let the consumer choose the media content after showing her the ad for a period of time, because she is motivated to do so anyway. However, under static advertising, granting the consumer the complete control of whether to process the ad or to consume the content encourages her to skip the ad entirely. This approach leaves money on the table because the platform could have extracted the value from matching the advertiser and the consumer.

If the media content is very attractive, $m > \lambda(v_c + v_a)$, all three models produce the same profit, because they all admit the strategy of offering no ads and focusing on the media content.

Proposition 5 provides an alternative, but complementary, perspective on why (and when) online or digital advertising is more attractive than advertising on traditional media such as newspapers and television. Conventional wisdom suggests that it is the ability to target specific consumer segment or individuals that makes online advertising attractive. Proposition 5 highlights the importance of the interactive nature of online media communication, a feature often overlooked in studies of online advertising. Print media and broadcast media lack the technological capability to engage in instantaneous two-way communication. Note that our model is simply based on a representative consumer, and it implicitly assumes that all media formats can target individuals. Without targetability, print media and broadcast media could perform even worse compared with interactive online media. Hence, the targeting and interactive features together are what make online advertising appealing than traditional advertising.

Proposition 5 also explains why online platforms may not always choose to adopt interactive advertising. YouTube, for example, adopts a variety of advertising formats including banner ads and non-skippable ads, even though skippable ads are technologically feasible and have been shown to be effective. In situations where static and sequential ads can do just as well as interactive ads, it makes more sense to choose them over interactive ads, which generally involves higher operational costs.

4.2 Advertising Externality

There are mixed views on how consumers perceive advertising. It is widely held that ads typically contain information about products or brands, thus facilitating consumers' decision making (Nelson 1974). This implies that consumers generally expect a positive net benefit from viewing an ad.[10] However, in media environments, advertising may cause negative utility. Common complaints from consumers about advertising include, for example, "the ad is too distracting", "it is just wasting my time", and "it is not relevant/helpful at all". Hence, many studies on two-sided media market have assumed that advertising is a nuisance to consumers (e.g., Dukes and Gal-Or 2003, Dukes 2004, Anderson and Coate 2005, Anderson and Gans 2011). A few cautiously point out that this assumption does not hold in all media formats (e.g., Kaiser and Song 2009). The nuisance view is particularly pronounced in studies of radio and television advertisements. In print media such as newspaper and magazines, consumers often view advertising more positively.

This subsection provides a unified analysis to resolve these disparate views of advertising. To that end, we must be precise about what it is meant by the utility or externality of advertising to a consumer. Note that a frequent complaint in the ad nuisance narratives is that ads distract consumers' attention from media content. Hence, the time meant for content consumption that is lost to advertising can form the basis for defining advertising externality. Let $V^0 = m/r$ denote the consumer's expected value under a counterfactual scenario, in which only the media content is available whereas no ad is shown. Let V^* denote the total value the consumer expects to gain on the platform by following the opti-

[10] At a minimum, consumers do not lose much (e.g., not purchase a product) if an ad turns out to not deliver much information.

mal attention strategy in equilibrium. The externality of advertising, ΔV, is thus defined as follows:

Definition 2 (Advertising Externality): *The externality of advertising to the consumer is the difference between the expected value of the optimal attention strategy under advertising and the expected value when no advertising is present:* $\Delta V \equiv V^* - V^0$.

This definition of advertising externality is aligned with the literature on media markets and advertising (e.g., Becker and Murphy 1993, Anderson and Gabszewicz 2006). Under this definition, negative externality may arise when a consumer finds that the benefit of advertising falls short of the benefit of alternative use of their time. Without loss of generality, in this subsection the outside option is normalized to zero, i.e., $b = 0$. It is obvious that negative externality can hardly arise under static advertising, as stated in the following proposition.

Proposition 6 (Externality under Static Advertising): *Under Model $\mathcal{D}_1$, the externality of advertising to the consumer is always nonnegative:* $\Delta V_{\mathcal{D}_1} \geq 0$.

PROOF: If $m > \lambda v_c$, then $V^*_{\mathcal{D}_1} = V^0 = \frac{m}{r}$. If $m < \lambda v_c$, then $V^*_{\mathcal{D}_1} = \frac{\lambda(rv_c+m)}{r(\lambda+r)} > \frac{m}{r}$. ■

Proposition 6 states that there is no nuisance cost of advertising so long as the consumer follows the optimal attention strategy under static advertising. This is fairly intuitive. A consumer with full control of the decision to either consume the media content or to process an ad should not be annoyed by the availability of advertising as she can always choose not to pay attention to it.

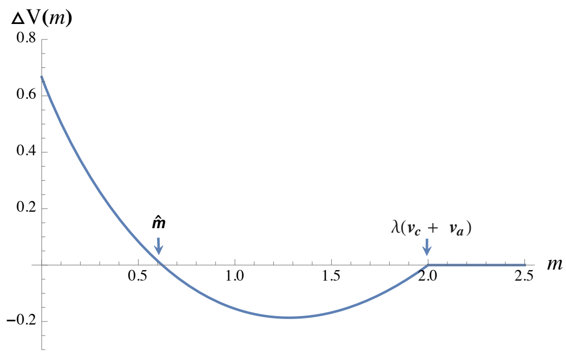
Turning to sequential advertising, we first note that the total value the consumer expects in equilibrium is

$$V^*_{\mathcal{D}_2} = \int_0^{S^*} v_c e^{-rt} dF(t) + \int_{S^*}^{\infty} m e^{-rt} dt. \tag{9}$$

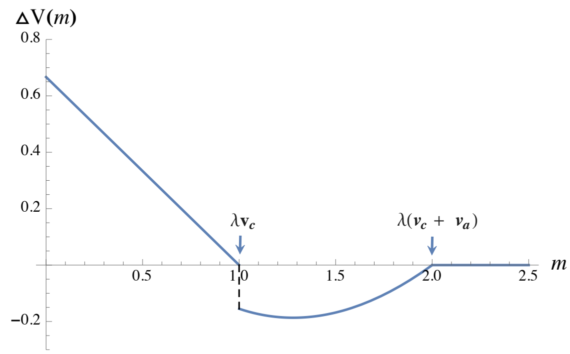
Following Definition 2, the externality of advertising is measured by comparing $V^*_{\mathcal{D}_2}$ with the no-advertising benchmark $V^0 = m/r$. The following proposition presents the result of this comparison and Figure 2a provides an illustration of the result.

Proposition 7 (Externality under Sequential Advertising): *Under Model $\mathcal{D}_2$, there exists a threshold $\hat{m}$ such that the externality of advertising to the consumer, $\Delta V_{\mathcal{D}_2}$, is positive if $m < \hat{m}$, negative if $\hat{m} < m < \lambda(v_c + v_a)$, and zero if $m > \lambda(v_c + v_a)$.*

PROOF: See the appendix. ■



(a) Sequential Advertising



(b) Interactive Advertising

Figure 2: Externality of Advertising as a Function of Content Quality ($v_c = v_a = 1$, $\lambda = 1$, $r = 0.5$)

The key difference between Proposition 7 and Proposition 6 is that under sequential advertising, as the media platform has more control over what the consumer can attend to, the consumer may be hurt by having advertising. While this makes intuitive sense, it is not at all obvious *a priori*, because the consumer can always choose not to pay attention to an ad. Indeed, under static advertising, the consumer is not hurt precisely because she can withhold her attention at any time. The situation of sequential advertising is different. The ad is run in a fixed time window and because the ad signal arrives stochastically, there is always a positive probability that some time is wasted— that is, the consumer obtains the ad signal but has to wait for the media content. When the media content is not very attractive (i.e., m is small), the consumer does not loss much if she pays more attention to the ad. When the media content is very attractive (i.e., m is large), it is optimal for the platform to abandon advertising. It is only when m is within an intermediate range that the consumer finds the ad annoying because she has to wait until S^* to enjoy the media content.

It is worth highlighting that under sequential advertising, the content value m *per se* has no direct impact on the consumer's attention strategy. It only affects consumer utility indirectly through the endogenous change of the exposure time S^*, which is determined by the media platform. A similar logic holds for the match value v_a that is expected by the advertiser. Although v_a does not directly affect consumer's attention, it endogenously changes the ad's exposure time. That is, the consumer becomes more annoyed by the ad if the advertiser reaps more profit from the match, because in this situation the platform is motivated to extend the ad exposure time and thus increase the chance of signal arrival. In contrast, if the consumer expects a higher value from the ad captured by v_c, she will be less annoyed by prolonged ad exposure because she has a stronger incentive to keep paying attention to the ad. These observations are summarized by the following corollary.

Corollary 1 *The advertising externality, $\Delta V_{\mathcal{D}_2}$, increases with consumer's expected match value v_c, but decreases with advertiser's expected value v_a.*

PROOF: See the appendix. ■

Last, let us turn to the case of interactive advertising. Given the equilibrium policy derived in Proposition 4, the consumer's expected value of following the optimal strategy is

$$V^*_{\mathcal{D}_3} = \begin{cases} V^*_{\mathcal{D}_1} & \text{if } m < \lambda v_c; \\ V^*_{\mathcal{D}_2} & \text{if } m > \lambda v_c. \end{cases} \tag{10}$$

It is then straightforward to characterize the advertising externality under this format of advertising. It is presented in the following proposition:

Proposition 8 (Externality under Interactive Advertising): *Under Model $\mathcal{D}_3$, the externality of advertising to the consumer, $\Delta V_{\mathcal{D}_3}$, is positive if $m < \lambda v_c$, negative if $\lambda v_c < m < \lambda(v_c + v_a)$, and zero if $m > \lambda(v_c + v_a)$. In addition, $\Delta V_{\mathcal{D}_3} \geq \Delta V_{\mathcal{D}_2}$.*

PROOF: See the appendix. ■

Figure 2b illustrates the pattern of advertising externality under the interactive format. Similar to sequential advertising, interactive advertising can also lead to a negative externality. Although in this format the media platform provides the skip option to the consumer, it introduces the option only after a minimum exposure time for the ad, which distracts the consumer from the media content. Nevertheless, the consumer expects an overall higher value under interactive advertising than under sequential advertising. First, the range of parameters for which negative externality occurs is reduced under interactive advertising because $\hat{m} < \lambda v_c$ (recall that $\hat{m}$ is the threshold that defines the negative externality under sequential advertising). Second, when the content is less attractive, i.e., $m < \lambda v_c$, the consumer can exercise the skip control from the start, achieving a higher value than that under sequential advertising.

5. Beyond the Simple Model

The framework developed thus far allows us to succinctly analyze how media formats influence advertising communication. Some simplifying assumptions are made for the purpose of expositional clarity. In this section, we briefly discuss some directions for stretching the model to accommodate other aspects of media advertising.

5.1 Time-Varying Content Value

One direction for enriching the model is to allow time-varying media content. To illustrate this, let us focus on the case of sequential advertising.[11] A few additional assumptions and notations are necessary. Assume that the content process $m(t)$ follows a Markov process that is independent of the advertising process. Let $x(t), y(t)$ denote the process time of the ad and the content, respectively, at time t. Whenever the consumer processes the content at time t, her flow payoff is $m(y(t))$. If the consumer freezes the content process and processes the ad instead, then the state $y(t)$ remains unchanged. If the consumer always processes the content from the start $(t = 0)$, then $y(t) = t$, and the maximum total payoff from the media content in this case is assumed to be bounded: $\int_0^\infty m(t)dt < \infty$. The evolutions of states x and y are independent of each other. It is also assumed that $m(y)$ is known to both the consumer and the platform for every possible y.

Proposition 9 *Under time-varying media content, a media platform shows an ad whenever $\lambda(v_a + v_c) > M^*(y)$, where $M^*(y)$ is defined as the solution to the following equation:*

$$M^*(y) \equiv \sup_\tau \frac{\mathbb{E}[\int_0^\tau m(y(s))e^{-rs}ds|y(0) = y]}{\mathbb{E}[\int_0^\tau e^{-rs}ds|y(0) = y]}. \tag{11}$$

This result follows straightforwardly from the index policy. The ad allocation policy in the main analysis can be understood as a special case of the above result. In the more general media environment in which content value fluctuates, the optimal policy in Proposition 9 implies that a media platform may insert an ad a number of times throughout a content exposure session. This is consistent with the observation that television programs place commercials throughout the programming period.

[11] The model here remains silent on many institutional details that may further influence media behavior. Zhou [2004] provides an early analysis on how a monopoly television network can structure commercial breaks based on the appeal of the program.

5.2 Multiple Ads

Media platforms often show multiple ads from different advertisers. The model can be readily extended to capture this practice. We again focus on the case of sequential advertising with constant media content. Suppose that there are K advertisers available in the advertising market. The k-th advertiser can gain v_{ak} and deliver an expected value of v_{ck} if its ad signal reaches the consumer. The ad process of the k-th advertiser follows an exponential process with hazard rate λ_k and is independent across ads.

Proposition 10 *Under sequential advertising with multiple ads, a media platform shows the ad from advertiser k as long as $\lambda_k(v_{ak}+v_{ck}) > m$, and sequentially in the order of decreasing magnitude of $\lambda_k(v_{ak}+v_{ck})$. Furthermore, the exposure time of each ad is $S_k^* = \frac{1}{\lambda_k}\ln\frac{\lambda_k(v_{ck}+v_{ak})}{m}$.*

Proposition 10 illustrates that the decision to show an ad depends on the threshold rule that is independent across ads. This is quite natural given the independence assumption. More importantly, the proposition sheds light on how media platforms should prioritize different ads. Ads with higher index value $\lambda_k(v_{ak}+v_{ck})$ should be prioritized and shown first.

5.3 Ads with Entertainment Value

Ads often carry more than just information. To draw attention, advertisers deliberately make ads interesting, humorous, and entertaining, and they may involve storytelling (Dukes and Liu 2020). These ads typically deliver benefits to consumers, in addition to product information. The simplest way to capture this idea is to assume that there is a flow utility w attached to an ad, independent of the ad's signal arrival process. Consumers are more likely to attend to an ad that offers a flow benefit, motivating media platforms to show such ads and to increase their exposure time.

Proposition 11 *Under sequential advertising, a media platform shows an ad with constant entertainment value as long as* $\lambda(v_a + v_c) > m - w$ *for a duration of* $S^* = \frac{1}{\lambda} \ln \frac{\lambda(v_c+v_a)}{m-w}$.

Proposition 11 illustrates that as an ad contains more entertainment value, consumers become more willing to process the ad. Furthermore, the media platform becomes more willing to extend the ad exposure time, which in turn increases the chance that consumers will receive the ad signal. The complementarity between entertainment value and informational value to some extent explains why a considerable fraction of many television commercials is devoted to entertaining consumers.

6. Concluding Remarks

Despite the interdependence of advertising and media, the influence of media formats on advertising communication has received relatively little attention in the advertising literature. We therefore take a preliminary step towards a better understanding of the fundamental differences and connections between various media formats. These questions are explored in a tractable framework of an information design problem in which media platforms determine the allocation of ads and media content, and consumers determine what to pay attention to at different times. For media platforms, the core of the problem is the tradeoff between drawing consumers' attention to their media content and ensuring the delivery of advertising information that may match the consumers and advertisers. This study characterizes the equilibrium behaviors of both consumers and media platforms under each media format, allowing us to compare the profitability of different formats. An important contribution of the model is that it provides a simple and intuitive explanation for the phenomenon of ad annoyance, which has been a critical parameter for many analyses in the media and advertising literature.

Clearly, this preliminary study raises more questions than it answers. Nevertheless, as shown in Section 5, the model can be flexibly extended to study a host of problems. The simple model presented here does not take into account consumer heterogeneity. An obvious area of inquiry is to look at the implications of consumer heterogeneity for advertising allocation policy, particularly when media platforms do not have the capability of targeting individuals. As consumer heterogeneity is assumed away, pricing decisions are also ignored in this study, allowing us to focus on the information design problem. Future research could examine the interaction between pricing and information design. Another interesting direction for future studies is how multiple media may jointly affect advertising communication. This is particularly relevant when media platforms offer multiple channels or when multiple media platforms are competing for consumer attention. Finally, although the consumer in our model is fully rational with forward-looking capability, it turns out that the optimal attention strategy is a myopic (or one-step look-ahead) policy. Future studies that involve complex media environments may consider assuming myopic consumers to simplify analysis without loss of generality.

Appendix

A. Proof of Lemma 4

To prove this lemma, we first prove a set of useful intermediate results.

Lemma 5 (Advantage of Optional Control (Part I)): *Suppose the ad signal has not arrived at time t_0, and consider an exogenous deadline T (possibly infinite). Then within the period $[t_0, T]$, running the media content exclusively is weakly dominated by offering and committing to the optional control $\{A, M\}$ throughout.*

PROOF: First, consider the case $\lambda v_c \leq m$. By Lemma 3, the consumer will always choose to consume the media content M when the optional control $\{A, M\}$ is available. Hence, the total profit the media can extract under the optional control is the same as that under running M exclusively.

Second, consider the case $\lambda v_c > m$. By Lemma 3, when the optional control $\{A, M\}$ is available within $[t_0, T]$, the consumer will first pay attention to the ad A until the signal arrival, or the deadline T, whichever comes first. There are two possible outcomes, depending on the arrival of the ad signal.

Case (a): The ad signal does not arrive before T (i.e., $\sigma > T$). In this case, the consumer pays attention to the ad throughout $[t_0, T]$. The profit accrued is then higher than the profit under running M exclusively because $\lambda v_c > m$.

Case (b): The ad signal eventually arrives before T (i.e., $\sigma < T$). In this case, the profit yielded after σ is the same across the two policies because both lead the consumer to consume M only. However, within $[t_0, \sigma]$, the optional-control policy allows the consumer to process the ad. Thus, by the same argument in Case (a), the profit is higher under the optional-control policy.

Summarizing, by committing to the optional control within $[t_0, T]$, the media platform can achieve a higher profit than running the media content exclusively throughout. ■

Lemma 6 (Advantage of Optional Control (Part II)): *Suppose the ad signal has not arrived at time t_0, and consider an exogenous deadline T (possibly infinite). Suppose further that $\lambda v_c > m$. Then within the period $[t_0, T]$, exposing the ad exclusively for a period of S followed by running the media content exclusively is weakly dominated by offering and committing to the optional control $\{A, M\}$ throughout.*

PROOF: When $\lambda v_c > m$, by Lemma 3, under the optional control $\{A, M\}$ the consumer will first pay attention to the ad A until the signal arrival, or the deadline T, whichever comes first. Let us examine the two possible cases of signal arrival.

Case (a): The ad signal does not arrive before S (i.e., $\sigma > S$). In this case, prior to $t = S$, the consumer pays attention to the ad under both policies. After $t = S$, however, the consumer continues to process the ad until the signal arrival under the optional-control policy, whereas she will be diverted to media content under the sequential policy. Hence, starting from $t = S$, the situation is the same as that in Lemma 5. Then, the optional-control policy dominates the sequential policy.

Case (b): The ad signal arrives before S (i.e., $\sigma < S$). In this case, prior to $t = \sigma$, the consumer pays attention to the ad under both policies. After $t = S$, however, the consumer switches to the media content under the optional-control policy, whereas she has to turn to the outside option for the period of $[\sigma, S]$ under the sequential policy. Thus, there is profit gain under the optional-control policy.

Summing up, by committing to the optional control within $[t_0, T]$, the media platform can achieve a higher profit than running the ad and media content sequentially throughout. ■

We are now prepared to prove Lemma 4.

PROOF: By Lemma 5 and Lemma 6, the optimal media policy entails sequencing of two components: running the ad A exclusively and running the optional control $\{A, M\}$. This is because any M-exclusive run will be weakly dominated by the optional control $\{A, M\}$. Under the optional control $\{A, M\}$, the media platform may or may not commit the course of action throughout. Next, we show that the media platform has no incentive to overturn the consumer's decision whenever the optional control $\{A, M\}$ is introduced. Take the first sequence such that the platform starts with running A exclusively for a period of S_1, followed by a period S_2 of optional control $\{A, M\}$.

Case (1): $\lambda v_c \leq m$. By Lemma 3, the consumer will keep paying attention to the ad during the period of $[0, S_1]$ unless the ad signal has arrived. If the ad signal does arrive before S_1, that is, $\sigma < S_1$, then it will be always optimal for the consumer to choose M after $t = S_1$, provided it is available. If at any point later than S_1, the platform disables the access to M, then it will incur a loss of profit. If the ad signal does not arrive before S_1, that is, $\sigma > S_1$, then the consumer does not have any incentive to process A after $t = S_1$. Given the availability of M after $t = S_1$, the consumer will consume it. Therefore, regardless of whether the signal has arrived or not, the consumer will choose M as long as it is available. Hence, by observing the decision of the consumer at the point of introducing $\{A, M\}$, the platform does not update its information thereafter. Its optimal policy is to let the consumer continue to consume the media content M. This can be achieved by either continuing to offer the optional control $\{A, M\}$ or restricting to the media content M.

Case (2): $\lambda v_c > m$. Again, by Lemma 3, the consumer will keep paying attention to the ad during the period of $[0, S_1]$ unless the ad signal has arrived. If the ad signal does arrive before S_1, that is, $\sigma < S_1$, then it will be always optimal for the consumer to choose M after $t = S_1$, provided it is available. However, if the ad signal does not arrive before S_1, that is, $\sigma > S_1$, then the consumer will continue to process the ad after $t = S_1$. If the platform wants to overturn the consumer's decision, then it can instead introduce the media content exclusively, implying the failure to commit to the optional control $\{A, M\}$ throughout $[S_1, S_2]$. However, according to Lemma 6, this is weakly dominated by the committed optional control. Therefore, it is never optimal for the media platform to violet the commitment for any point after $t = S_1$.

The two cases taken together imply that once the platform has introduced the optional control $\{A, M\}$, it is committed to providing this option thereafter.

■

B. Proof of Proposition 7

PROOF: The total discounted value a consumer expects following the optimal strategy is:

$$
\begin{aligned}
V^*_{\mathcal{D}_2} &= \int_0^{S^*} v_c e^{-rt} dF(t) + \int_{S^*}^{\infty} m e^{-rt} dt && \text{(A-1)} \\
&= \frac{\lambda v_c}{r+\lambda}(1 - e^{-(r+\lambda)S^*}) + \frac{m}{r} e^{-rS^*} && \text{(A-2)} \\
&= \frac{\lambda v_c}{r+\lambda}(1 - e^{-(r+\lambda)S^*}) + \frac{\lambda(v_c+v_a)}{r} e^{-(r+\lambda)S^*} && \text{(A-3)} \\
&= \frac{\lambda v_c}{r+\lambda} + \big(\frac{\lambda(v_c+v_a)}{r} - \frac{\lambda v_c}{r+\lambda}\big) e^{-(r+\lambda)S^*} && \text{(A-4)} \\
&= \frac{\lambda v_c}{r+\lambda} + \big(\frac{\lambda(v_c+v_a)}{r} - \frac{\lambda v_c}{r+\lambda}\big)\big(\frac{m}{\lambda(v_c+v_a)}\big)^{1+\frac{r}{\lambda}}. && \text{(A-5)}
\end{aligned}
$$

where the third equality (A-5) follows from the first-order condition in Equation (6). Let $\Delta V = V^*_{\mathcal{D}_2} - V^0$, where $V^0 = m/r$. Note that

$$
\frac{\partial \Delta V}{\partial m} = \underbrace{\big(\frac{\lambda(v_c+v_a)}{r} - \frac{\lambda v_c}{r+\lambda}\big)}_{>0} \big(\lambda(v_c+v_a)\big)^{-1-\frac{r}{\lambda}} \big(1+\frac{r}{\lambda}\big) m^{\frac{r}{\lambda}} - \frac{1}{r} < 0
$$

for m sufficiently small, and

$$
\frac{\partial^2 \Delta V}{\partial m^2} = \big(\frac{\lambda(v_c+v_a)}{r} - \frac{\lambda v_c}{r+\lambda}\big) \big(\lambda(v_c+v_a)\big)^{-1-\frac{r}{\lambda}} \big(1+\frac{r}{\lambda}\big) \frac{r}{\lambda} m^{\frac{r}{\lambda}-1} > 0.
$$

Together with the facts that $\Delta V(m=0) = \lambda v_c/(r+\lambda)$ and that $\Delta V(m = \lambda(v_c + v_a)) = 0$, it follows that, there exists a threshold $\hat{m} \in (0, \lambda(v_c+v_a))$ such that

$\Delta V > 0$ if $m < \hat{m}$, and $\Delta V < 0$ if $m > \hat{m}$.

■

C. Proof of Corollary 1

PROOF: Note that V^0 is independent of v_a and v_c, the impacts of these two parameters on $V^*_{\mathcal{D}_2}$ are the same as their impacts on ΔV. Using the expression of $V^*_{\mathcal{D}_2}$ in Equation A-4 and taking its derivative over v_c, we have

$$
\begin{aligned}
\frac{\partial \Delta V}{\partial v_c} &= \frac{\lambda}{r+\lambda} + \big(\frac{\lambda}{r} - \frac{\lambda}{r+\lambda}\big)e^{-(r+\lambda)S^*} + \big(\frac{\lambda(v_c+v_a)}{r} - \frac{\lambda v_c}{r+\lambda}\big)\frac{\partial e^{-(r+\lambda)S^*}}{\partial v_c} \\
&= \frac{\lambda}{r+\lambda} + \big(\frac{\lambda}{r} - \frac{\lambda}{r+\lambda}\big)e^{-(r+\lambda)S^*} - \big(\frac{\lambda(v_c+v_a)}{r} - \frac{\lambda v_c}{r+\lambda}\big)\frac{r+\lambda}{\lambda(v_c+v_a)}e^{-(r+\lambda)S^*} \\
&= \frac{\lambda}{r+\lambda}(1 - e^{-(r+\lambda)S^*}) + \frac{v_a}{v_c+v_a}e^{-(r+\lambda)S^*} \\
&> 0.
\end{aligned}
$$

With respect to v_a, we have

$$
\begin{aligned}
\frac{\partial \Delta V}{\partial v_a} &= \frac{\lambda}{r}e^{-(r+\lambda)S^*} + \big(\frac{\lambda(v_c+v_a)}{r} - \frac{\lambda v_c}{r+\lambda}\big)\frac{\partial e^{-(r+\lambda)S^*}}{\partial v_a} \\
&= \frac{\lambda}{r}e^{-(r+\lambda)S^*} - \big(\frac{\lambda(v_c+v_a)}{r} - \frac{\lambda v_c}{r+\lambda}\big)\frac{r+\lambda}{\lambda(v_c+v_a)}e^{-(r+\lambda)S^*} \\
&= -\frac{v_a}{v_c+v_a}e^{-(r+\lambda)S^*} \\
&< 0.
\end{aligned}
$$

■

D. Proof of Proposition 8

The proof evokes a variant of Bernoulli's inequality, which is stated below without proof:

Lemma 7 (A Variant of Bernoulli's Inequality): $(1+y)^{\theta} < \frac{1+y}{1+y-\theta y}$, *for any real numbers* $\theta > 1$ *and* $y \in (-1, 0)$.

PROOF: Part(i): Based on Equation 10, we have

$$\Delta V_{\mathcal{D}_3} \equiv V^*_{\mathcal{D}_3} - V^0 = \begin{cases} V^*_{\mathcal{D}_1} - V^0 & \text{if } m < \lambda v_c; \\ V^*_{\mathcal{D}_2} - V^0 & \text{if } m > \lambda v_c. \end{cases}$$

Recall from Proposition 7 that $V^*_{\mathcal{D}_2} - V^0 < 0$ if $\hat{m} < m < \lambda v_c$. According to the proof of Proposition 7, $\partial \Delta V_{\mathcal{D}_2} / \partial m (m = \hat{m}) < 0$. It suffices to determine the relationship between $\hat{m}$ and λv_c. In what follows, we show that $\hat{m} < \lambda v_c$.

When $m = \lambda v_c$, following Equation A-5 we have

$$\begin{aligned} \Delta V_{\mathcal{D}_2} \equiv V^*_{\mathcal{D}_2} - V^0 &= \frac{\lambda v_c}{r+\lambda} + \Big(\frac{\lambda(v_c+v_a)}{r} - \frac{\lambda v_c}{r+\lambda}\Big)\Big(\frac{v_c}{v_c+v_a}\Big)^{1+\frac{r}{\lambda}} - \frac{\lambda v_c}{r} \\ &= -\frac{\lambda^2 v_c}{r(r+\lambda)} + \frac{\lambda\big(\lambda(v_c+v_a)+rv_a\big)}{r(r+\lambda)}\Big(\frac{v_c}{v_c+v_a}\Big)^{1+\frac{r}{\lambda}}, \\ &< -\frac{\lambda^2 v_c}{r(r+\lambda)} + \frac{\lambda\big(\lambda(v_c+v_a)+rv_a\big)}{r(r+\lambda)} \cdot \frac{\lambda v_c}{\lambda(v_c+v_a)+rv_a} \\ &= 0, \end{aligned}$$

where the second equality is obtained by rearranging terms and the inequality follows from Lemma 7 by noting that $y = -v_a/(v_a + v_c)$ and $\theta = 1 + r/\lambda$. Then according to Proposition 7, we have $\hat{m} < \lambda v_c$.

Part(ii): To prove the second result that $\Delta V_{\mathcal{D}_3} \geq \Delta V_{\mathcal{D}_2}$, it suffices to show $V^*_{\mathcal{D}_1} > V^*_{\mathcal{D}_2}$ if $m < \lambda v_c$. Indeed,

$$\begin{aligned} V^*_{\mathcal{D}_1} - V^*_{\mathcal{D}_2} &= \frac{\lambda(rv_c+m)}{r(\lambda+r)} - \frac{\lambda v_c}{r+\lambda} + \Big(\frac{\lambda(v_c+v_a)}{r} - \frac{\lambda v_c}{r+\lambda}\Big)\Big(\frac{m}{\lambda(v_c+v_a)}\Big)^{1+\frac{r}{\lambda}} \\ &= \frac{\lambda}{r+\lambda}\Big[m - \big(\lambda(v_c+v_a)+rv_a\big)\cdot\Big(\frac{m}{\lambda(v_c+v_a)}\Big)^{1+\frac{r}{\lambda}}\Big] \\ &> \frac{\lambda}{r+\lambda}\Big[m - \big(\lambda(v_c+v_a)+rv_a\big)\cdot\frac{m}{(r+\lambda)(v_c+v_a)-\frac{rm}{\lambda}}\Big] \\ &= \frac{m\lambda}{r+\lambda}\Big[1 - \frac{\lambda(v_c+v_a)+rv_a}{(r+\lambda)(v_c+v_a)-\frac{rm}{\lambda}}\Big] \end{aligned}$$

$$> \frac{m\lambda}{r+\lambda}\Big[1 - \frac{\lambda(v_c+v_a)+rv_a}{(r+\lambda)(v_c+v_a)-rv_c}\Big]$$
$$= 0,$$

where the first inequality makes use of Lemma 7 by treating $y = m/\lambda(v_a + v_c) - 1$ and $\theta = 1 + r/\lambda$, and the second inequality follows from $m < \lambda v_c$. ■

References

Simon P Anderson and Stephen Coate. Market provision of broadcasting: A welfare analysis. *The Review of Economic Studies*, 72(4):947–972, 2005.

Simon P Anderson and Jean J Gabszewicz. The media and advertising: a tale of two-sided markets. *Handbook of the Economics of Art and Culture*, 1:567–614, 2006.

Simon P Anderson and Joshua S Gans. Platform siphoning: Ad-avoidance and media content. *American Economic Journal: Microeconomics*, 3(4):1–34, 2011.

Mark Armstrong. Competition in two-sided markets. *The RAND Journal of Economics*, 37(3):668–691, 2006.

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

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

Anthony Dukes. The adverstising market in a product oligopoly. *The Journal of Industrial Economics*, 52(3):327–348, 2004.

Anthony Dukes and Esther Gal-Or. Negotiations and exclusivity contracts for advertising. *Marketing Science*, 22(2):222–245, 2003.

Anthony J Dukes and Qihong Liu. Advertising content and viewer attention: The role of ad formats. *Available at SSRN*, 2020.

Jeffrey Ely, Alexander Frankel, and Emir Kamenica. Suspense and surprise. *Journal of Political Economy*, 123(1):215–260, 2015.

Jeffrey C Ely. Beeps. *American Economic Review*, 107(1):31–53, 2017.

Gustav Feichtinger, Richard F Hartl, and Suresh P Sethi. Dynamic optimal control models in advertising: recent developments. *Management Science*, 40(2): 195–226, 1994.

John C. Gittins. Bandit processes and dynamic allocation indices. *Journal of the Royal Statistical Society B*, 41:148–164, 1979.

Brett R. Gordon, Kinshuk Jerath, Zsolt Katona, Sridhar Narayanan, Jiwoong Shin, and Kenneth C. Wilbur. Inefficiencies in digital advertising markets. *Journal of Marketing*, 85(1):7–25, 2021.

Ulrich Kaiser and Minjae Song. Do media consumers really dislike advertising? an empirical assessment of the role of advertising in print media markets. *International Journal of Industrial Organization*, 27(2):292–301, 2009.

Emir Kamenica and Matthew Gentzkow. Bayesian persuasion. *American Economic Review*, 101(6):2590–2615, 2011.

John DC Little. Aggregate advertising models: The state of the art. *Operations Research*, 27(4):629–667, 1979.

Vijay Mahajan and Eitan Muller. Advertising pulsing policies for generating awareness for new products. *Marketing Science*, 5(2):89–106, 1986.

John Joseph McCall. Economics of information and job search. *The Quarterly Journal of Economics*, pages 113–126, 1970.

Phillip Nelson. Advertising as information. *Journal of Political Economy*, 82(4): 729–754, 1974.

Marc Nerlove and Kenneth J Arrow. Optimal advertising policy under dynamic conditions. *Economica*, pages 129–142, 1962.

Régis Renault. Advertising in markets. In *Handbook of Media Economics*, volume 1, pages 121–204. Elsevier, 2015.

Jean-Charles Rochet and Jean Tirole. Two-sided markets: a progress report. *The RAND Journal of Economics*, 37(3):645–667, 2006.

Suresh P Sethi. Dynamic optimal control models in advertising: a survey. *SIAM Review*, 19(4):685–725, 1977.

ML Vidale and HB Wolfe. An operations-research study of sales response to advertising. *Operations Research*, 5(3):370–381, 1957.

Richard Weber. On the gittins index for multiarmed bandits. *The Annals of Applied Probability*, 2(4):1024–1033, 1992.

Martin L. Weitzman. Optimal search for the best alternative. *Econometrica*, 47 (3):641–654, 1979.

E Glen Weyl. A price theory of multi-sided platforms. *American Economic Review*, 100(4):1642–72, 2010.

Wen Zhou. The choice of commercial breaks in television programs: the number, length and timing. *The Journal of Industrial Economics*, 52(3):315–326, 2004.