Quant Mark Econ (2009) 7:237–266
DOI 10.1007/s11129-009-9068-x

Targeted advertising as a signal

Bharat N. Anand · Ron Shachar

Received: 14 July 2008 / Accepted: 23 June 2009 / Published online: 25 July 2009

Abstract This study presents a signaling model of advertising for horizontally differentiated products. The central ingredients of the model are two important characteristics of advertising—targeting, and noisy information content. The theory yields interesting results about the informational role of targeted advertising, and its consequences. First, targeting can itself serve as a signal on product attributes. Second, the effectiveness of targeting depends not only on firms knowing consumer preferences, but on consumers knowing that firms know this. This creates a distinction between strategies of targeting and personalization. Third, the effectiveness of targeting in equilibrium may (far) exceed the information contained directly in the targeted message. Fourth, information content is not, however, superfluous. Specifically, when ads contain *no* information, a targeting equilibrium does not exist. Together, these results reveal how advertising conveys information both through the content of the message and the firm's choice of advertising medium. Furthermore, the model is robust to the various critiques of prior work on ads-as-signals: namely, that ad content is irrelevant, ad exposure is unnecessary, and the choice of ads as signals is inherently arbitrary.

B. N. Anand
Harvard Business School, Soldiers Field Road, Boston, MA 02163, USA
e-mail: banand@hbs.edu

R. Shachar (✉)
Faculty of Management, Tel Aviv University, Tel Aviv, 69978, Israel
e-mail: rroonn@post.tau.ac.il

R. Shachar
Fuqua School of Business, Duke University,
1 Towerview Drive, Durham, NC 27708, USA
e-mail: shachar@duke.edu

Keywords Informative advertising · Signaling game · Targeting · Media selection · Advertising effectiveness

JEL Classification M37 · D82

1 Introduction

One of the most influential theories about advertising is based on the idea that advertisements can serve as signals. Specifically, firms may convey information about their product attributes to consumers indirectly through their expenditures on advertising rather than directly through advertising content. This view goes back at least to Nelson (1974) and subsequently, Kihlstrom and Riordan (1984) and Milgrom and Roberts (1986), followed by a long line of theoretical and empirical analyses. Despite the importance of the signaling perspective, there are certain salient concerns that have been raised about the role of ads as signals. First, the content of ads is irrelevant in that theory. Second, the effectiveness of ads does not depend on exposure. Third, advertising in the signaling theory is merely a nomenclature for any resource dissipative or "money burning" activity that can also serve as a signal.[1] We present a signaling model of advertising that is robust to each of these related critiques. Furthermore, the central ingredients of the model are two features of advertising that are motivated by empirical observation, and that are largely absent in prior signaling models: first, the targeting of advertising, and second, the noisiness of advertising content. In contrast to previous studies of advertising, the model here allows advertising to affect choices both directly via its content and indirectly via its signaling role. The interaction between these different mechanisms for information transmission yields new results on how advertising works, and its effectiveness.

Two facts about advertising are important empirically. First, most advertising is targeted. For example, advertising for handheld devices is targeted towards young professionals, and for rock climbing equipment towards outdoor enthusiasts. Targeting is ubiquitous even in product categories that cater to far broader customer segments, like autos (e.g., BMW, Volvo), beverages (e.g., Pepsi, Snapple), apparel (e.g., Liz Claiborne, Tommy Hilfiger), and

[1] As Bernheim and Redding (2001) note, "[E]xisting theory provides no basis for selecting one (nondiscriminatory) signal over another. This is viewed as an inherent shortcoming of all signaling-cum-money-burning explanations of economic behavior...Indeed, the choice of the means to burn money is inherently arbitrary. Generally, the question of why one method of money burning is chosen above all others has been left open."

entertainment (rap versus country music).[2] The perceived benefit of targeting is that it can reduce wasted advertising by ensuring that advertisements reach the most appropriate consumers for the firm's product.

Second, advertising messages are noisy. By noisiness we mean that ads can be misperceived—i.e., the ad can be interpreted by the consumer as meant by the firm *or not*. Consumer misperception has been documented in a variety of empirical and experimental studies about advertising. There are various reasons for noisiness such as misreading a message, misinterpreting it, not paying enough attention, or vague language.[3]

Both these phenomena have become, if anything, more important over time.[4] Fragmented media and new technologies (e.g., TiVo) make it easier to reach the individuals desired. At the same time, the proliferation of products and messages have made it harder for messages to be clear, thereby increasing noisiness of ads.

Existing theories of advertising have largely ignored these two aspects of advertising, and especially how they interact. This paper fills this gap. We study a model with these two ingredients: firms can target their advertisements to particular groups of consumers, and advertising is noisy. We characterize the resulting equilibrium, and use it to study various questions: Can media selection and targeting enable firms to overcome the noisiness of advertising? Can targeting strategies be augmented in any way to increase their effectiveness? And, how effective is advertising in equilibrium? These questions are not only relevant to firm strategy, but would seem to have important welfare consequences.

[2] Relatedly, one often thinks of some of the largest advertisers, like Proctor & Gamble or General Motors, as directing their ad dollars to a "mass audience." But, even for these firms, their advertising budget invariably comprises spending for many different products that each target a far narrower customer base. For example, General Motors (the world's largest advertiser) spends more than $50 million each on 12 separate brands that include sports cars, family minivans, and heavy trucks. Proctor & Gamble (the second largest advertiser) spends more than $10 million on 45 separate brands, ranging from specialized pet foods like *Iams* to anti-dandruff shampoo like *Head & Shoulders* (*Advertising Age*, June 27, 2005), each with a relatively distinct target customer set.

[3] In other words, the source of the noise can be in the way the ad was created or in the way it was consumed. Indeed, firms face a difficulty in conveying a lot of information in a 30-second advertisement, and consumers may find it hard to pay attention to each advertisement due to factors such as overload and clutter. For example, *The Economist* notes that "The average American is now subjected to 3,000 marketing messages every day and could not possibly take them all in." Even as far back as 1979, Webb and Ray (1979) showed that the more the clutter, the lower are average levels of ad recall.

[4] For example, the correct brand recall of the last commercial seen dropped from 18% in 1965 to 12% in 1974 and 7% in 1981. Advertising textbooks also note that the "major problem facing advertisers today is the difficulty of gaining attention in the face of the increase in advertising clutter" (Batra et al. 1996).

In our model, presented in Section 2, products are differentiated and consumers are heterogeneous in their tastes for product attributes. That is, any particular product has a better fit with the tastes of some consumers than of others, and consumer utility will depend on the resulting "match" between product attributes and their tastes. While firms know the tastes of consumers, consumers are uncertain about product attributes. Firms can send ads through different media channels. However, advertising content is a noisy message on product attributes. Consumer preferences over product attributes are correlated with their choice of media channel, creating a role for targeting.

In this setup, advertising allows a consumer to learn about her match with the characteristics of the product. As Johnson and Myatt (2006) note, the classic taxonomy of advertising is restrictive since it rules out this possibility. Specifically, the model here allows for the simultaneous possibilities that (i) a consumer learns that she likes the product, or (ii) a consumer learns that she dislikes it. In the latter case, exposure to ads ends up decreasing the consumer's tendency to purchase the product.[5]

Before examining the signaling equilibrium, we consider, in Section 3, the benchmark case where consumers only use the information content in the message to update their beliefs about product attributes. In equilibrium, firms target ads—that is, firms send messages only to consumers that are the right match for them (so that media selection choice is correlated with consumer preferences). Since advertising is noisy, consumers do not always choose the product that is the right fit for them.

In this benchmark model, consumers do not exploit the information contained in firms' actions. Since media selection is correlated with product attributes, consumers can use this information in their decision process, and increase their likelihood of choosing products that best fit their needs. Such behavior of consumers will obviously affect the incentives of firms, since firms with a worse match may have an incentive to now mimic the targeting strategies of rivals that have a better match. This raises the question of whether a separating equilibrium exists in such a setting. In Section 4, we show that such an equilibrium exists. Such a perfect Bayesian equilibrium has the following properties: (a) firms target only those consumers who best fit their product, and (b) consumers rationally interpret exposures to ads (in media channels that they watch) as a signal that the promoted product fits their preferences well. Thus, targeting itself (i.e., media selection) can serve as a signal on product attributes. This result reveals how advertising conveys information both through the content of the message and the firm's choice of advertising medium.

An important feature of the model concerns the role of advertising content. We show that media selection can serve as a signal as long as the content of the

[5] Anand and Shachar (2005) demonstrate this "consumption-deterring" effect of advertising, empirically.

ad is not empty. Put it differently, when advertisements are completely noisy, the separating equilibrium breaks down. In this sense, our signaling model is robust to each of the critiques, mentioned above, on the "advertising as a signal" a-la Nelson approach.

These results indicate that the role and effectiveness of targeting are wider than has been realized so far.

A hidden assumption in the signaling equilibrium is that consumers know that firms know their tastes. While one might question how well this assumption corresponds with the real world, it appears that firms are aware of this issue and, as a result, they often adopt strategies of "personalization." Personalization differs from simple targeting in that firms let the recipient know that she is being targeted. Our model suggests that this should be a more effective approach of targeting. We discuss some of these strategies later.

Although the ideas presented here are fairly intuitive, they lack formal foundation. One of the aims of the model is to provide such foundations, and to identify the conditions under which they exist. For example, we show that when ad content is empty (i.e., ads are not informative) the signaling equilibrium does not exist. Furthermore, identifying the conditions under which an equilibrium exists yields additional insights into the signaling role—such as the distinction between personalization and targeting (discussed above).

1.1 Related literature

This paper is related to the existing literature on informative advertising.[6] Previous studies distinguish between the role of directly informative advertising (through ad content) and indirectly informative advertising (through ads as signals). The model here shares common features with the first strand of work since, like Butters (1977), Grossman and Shapiro (1984), Meurer and Stahl (1994), and Anderson and Renault (2006), advertisements contain some information content. Whereas prior work assumes that advertising content resolves all uncertainty about product attributes, we take a different approach and assume that advertisements are noisy messages—there is uncertainty about how ad content will be perceived by consumers. This approach allows us to examine various issues related to noisy ad content, including how noisiness impacts ad effectiveness in equilibrium.

This paper is related to the literature on ads as signals as well. Virtually all prior work on signaling applies to vertically differentiated products and experience goods. As far as we know, ours is the first paper to identify a signaling role of advertising for horizontally differentiated products and search goods. Conventional wisdom is that the signaling role is not particularly relevant

[6]This contrasts with other theories of advertising, for example persuasive advertising, or advertising as complements (Becker and Murphy 1993).

in these settings, where advertising content can directly inform consumers. For example, Meurer and Stahl (1994) note that "in vertically differentiated markets, prices or advertising may signal the quality of products. In contrast, signaling opportunities do not arise in symmetric horizontally differentiated markets." As mentioned above, the reason a signaling role arises in our model stems from two features of advertising that have been previously ignored in signaling models: targeting, and noisiness of advertising content.

While prior work based on a vertical setting suggests that ad intensity is a signal, our model, based on horizontal differentiation, demonstrates that media selection is a signal. Beyond this, there are additional substantive economic differences between the two approaches.[7] Specifically, our model is robust to all the traditional critiques that have been raised about signaling theories of advertising—namely, that ad content is irrelevant in that theory, that the effectiveness of ads does not depend on exposure; and that, as a result, the choice of advertising in these models as a means to burn money is inherently arbitrary. Since, in the model here, the effectiveness of advertisements as signals depends on consumers' exposure to the message content, none of these critiques has bite any longer.

This discussion above clarifies another source of departure of our model from prior work – specifically, it integrates the directly and indirectly informative roles of advertising. The interaction between these two roles of advertising yields new results on the advertising mechanism (discussed earlier), and its effectiveness.

The previous discussion focuses on the differences between this study and prior work on the question of how advertising works. In addition, the model presented here offers a different perspective on the question of how much information there is in ads. The literature on advertising has seen a lively debate over this question. Theories of ad content, for example, have been criticized by some observers who question how much directly informative advertising there actually is. Indeed, theories of indirectly informative advertising were motivated by the observation that "a nontrivial amount of advertising (especially on TV) has little or no obvious informational content" (Milgrom and Roberts 1986). Similarly, Becker and Murphy (1993) note that "it is 'obvious' that many ads provide essentially no information." The model here offers a different perspective. Specifically, it suggests that on the question of how much information there is in ads, the only relevant distinction may

[7] (a) There are other, more technical, differences. For example, repeat purchases are replaced with informative advertising as the source of the single crossing property (that gives the high quality or good match firm a greater incentive to incur the advertising cost in equilibrium). (b) Recently Mayzlin and Shin (2008) presented an interesting variation of "advertising as a signal" in a vertical setting. In their model, it is the content of the ad that signals quality. Specifically, they show that in equilibrium the high quality firm engages in vague image-based advertising, while the medium quality firm engages in attribute-based advertising.

be between ads with *any* information—even if arbitrarily small—versus ads with none. The reason is that, as described, even when ads have arbitrarily little information content, ad effectiveness can be quite large. Only when ads are completely noisy do they lose their effectiveness. As a result, even little information can be enough information. Note, though, that (as discussed in Section 4) this result is sensitive to the assumption that when firms send ads, the relevant consumer is exposed to them with certainty.

The two important building blocks used here—the targeting of advertising and the noisiness of marketing communications—have been examined separately in two recent papers. Iyer et al. (2005) study the targeting of advertising. One of their main finding—"…firms advertise more to consumers who have a strong preference for their product…."—is similar to the result of our benchmark model with naïve consumers. The result is similar despite the differences in setting between the two papers: for example, in their model ads provide full information (about the existence and attributes of the product). Furthermore, their setting is based on the three-segments model presented by Narasimhan (1988).

Noisy communications are studied in Anand and Shachar (2007).[8] Our model builds on that but extends it to the realistic setting of advertising in two ways. First, while in Anand and Shachar (2007) the sender is limited to one message (with varying precision), here the number of ads is endogenous. Second, while in that model the sender (practically) uses one media channel, here we account for a critical aspect of advertising—media selection. As a result, while Anand and Shachar (2007) offer general implications of communication, the model here reveals several implications that are specific to advertising (e.g., targeting as a signal) and sheds new light on aspects of the phenomenon that have occupied previous scholars (such as the "burning money" critique).

The rest of the paper proceeds as follows. Section 2 describes the setup of the model, Section 3 characterizes the equilibrium in the benchmark model (that we refer to also as the "non-strategic" case), and Section 4 studies the signaling equilibrium. Section 5 concludes.

2 Model

The central ingredients of our model are as follows. First, products are differentiated, and consumers are heterogeneous and uncertain about product attributes. Second, advertising content is a noisy message on product attributes. Third, firms can send advertisements to consumers about their

[8] While Bhardwaj et al. (2008) study a different marketing communication tool (salespeople) they are also interested in imperfect communications abilities. In their case, the imperfection is not due to the noisiness of the message but rather to the limited "bandwidth" between the firm and customers.

products through different media channels. Fourth, consumer preferences over product attributes are correlated with their choice of media channel, creating a role for targeting. Last, for any particular media channel, firms know the preference distribution of consumers who are exposed to it.[9]

The first and third assumptions are standard. For the others, we offer several stylized facts below that clarify why these assumptions are quite reasonable.

Informative advertising The formulation of advertising content as a noisy message on product attributes is motivated by two sets of stylized facts. One set of facts shows that advertising content is informative, and the other shows that it is noisy.

First, the informative content in advertising has been documented in a variety of marketing studies that employ "content analysis." Resnik and Stern (1977) formulated this method to determine which types of information are present in an ad. They presented 14 information categories or "cues" , such as price, quality, performance etc. The relevant questions about performance, for example, were "What does the product do and how well it do what it is designed to do in comparison to alternative purchases". Many subsequent studies have followed this approach, and Abernethy and Franke (1996) summarize their results in a meta-analysis. They found that 84% of all the ads (91,438) have at least one cue (58% have at least two cues and 33% have had at least 3). While newspapers appear to include more informative advertising (98% of them have at least one cue), ads in television are also quite informative (71%). The product category that was the most informative was automobile (97%) and furniture/home furnishings/appliances (96%).

A second set of studies grounds the assumption that, while informative, it is fair to say that advertising content is noisy. Noisiness may be due to consumer misperception, as described earlier. Indeed, in 1979, the Educational Foundation of the American Association of Advertising Agencies surveyed nearly 2,700 consumers about the content of 60 30-second televised communications—including ads, public service communications, and editorial content. They found that about 29% of the communications (ads or other content) were miscomprehended by consumers, as measured by a particular series of true-false questions (Jacoby and Hoyer 1982). A second study was conducted a few years later and included print (magazine) ads and editorial content. In this study, about 1,250 consumers were asked questions about fifty-four full-page magazine ads and another fifty-four editorial pages. Roughly 20% of the material was miscomprehended, with another 15% being simply "not understood" (Jacoby and Hoyer 1989).

[9] Chen et al. (2001) discuss the role of both targetability (i.e., firms' knowledge about consumers' preferences), and addressability (i.e., the segmentation of media channels) in individual marketing. Their focus, however, is on targeted pricing not on the targeting of ads.

These facts suggest that there is uncertainty about the perception of ad content. As a remark on terminology, we sometimes refer to this as “noisy advertising” or “noisy content” hereafter.

Segmented media markets In practice, there are many classes of media—for example, television, radio, newspapers, magazines, billboards, and direct mail—and many categories within each media class. For example, on television there are morning shows, daytime shows, prime time shows, late night shows, etc. Even within each category there exist various options to advertisers. For example, within prime time, there are shows that appeal mostly to Republicans (like *Law and Order*) or shows that appeal especially to Democrats (like *Grey’s Anatomy*).[10] Thus, it is not surprising that the audience composition differs across and within media channels. (For example, while Fox’s audience is relatively young, CBS’ audience is relatively old). Furthermore, as the number of media channels has grown in the last decade, each of these channels has ended up serving more specific interests.

Firms’ Information Sets Firms have multiple data sources that characterize the audiences of the different media channels through which they can send advertisements. Moreover, these characterizations are rather detailed. For example, such information includes not only the demographics of the audiences, but their lifestyles and consumption habits as well.[11] These observations ground the assumption that firms are aware of the media consumption of individuals.

2.1 Model setup

There are two firms indexed by s (for “sender”), two consumers indexed by i, and two media channels indexed by k. Each firm offers one product. The outside alternative is denoted $s = 0$.

The utility of consumer i for alternative s is denoted $u_{i,s}$ and given by:

$$u_{i,s} = \begin{cases} \frac{1}{2}\beta_i x_s \text{ for } s \in \{1, 2\} \\ 0 \text{ for } s = 0 \end{cases} \tag{1}$$

where x_s is the product attribute of firm s, and β_i is the taste parameter of consumer i. While in previous studies of “advertising as a signal” (a la Nelson),

[10]The popularity of *Law and Order* among Republicans and as a result its attractiveness for the Republican candidate in the 2004 elections are discussed in a *New York Times* article (Jim Rutenberg, July 18, 2004). While the article also mentions programs that were appealing to the Democratic candidate, none of them was a prime time one. We mention *Grey’s Anatomy* because based on the discussion in the article this prime time show seems to be appealing to Democrats.

[11]Such data sets are offered, for example, by research firms like the Nielsen Research company, Information Resources, Inc, and Simmons Market Research Bureau.

the price usually serves as a signal too, here we exclude prices from our model. The reason is straightforward: previous studies focus on vertically differentiated products (i.e., on differences in quality). Since quality usually correlates with both cost and consumers' utility, the price of a product is a natural signal. However, since our research focuses on horizontally differentiated products in a symmetric setting, it is not reasonable to expect that prices will play a similar role.[12] Therefore, in order to simplify and focus the analysis, we exclude prices from our model. Later, in Section 5 we extend the model to include prices and other "vertical" characteristics of products.

There is product differentiation, $x_1 = 1$ and $x_2 = -1$, and heterogenous consumers, $\beta_1 = 1$ and $\beta_2 = -1$. Thus, the utility of consumer 1 from the product of firm 1 is positive, and from firm 2's product is negative. As a note on terminology, hereafter we refer to firm 1's product as having the "better match" with consumer 1 and firm 2's product as having the "worse match" with this consumer. The opposite holds for consumer 2.

The consumer's decision variable, $d_{i,s}$, is equal to 1 if she chooses alternative s, zero otherwise.

Information Set The consumers know that there are two products in the market, as well as the distribution of product attributes: $x_s \in \{-1, 1\}$. However, they do not know which firm offers what product. Their prior probability that either firm offers a product which is the right match for them is 0.5.

Informative Advertising Consumers acquire information about each product's attributes through advertising. The *perception of advertising content* is formulated as a noisy message on product attributes. The perceived content can be represented by a variable that takes on two values, −1 and 1. Specifically, let $m_{i,s,n}$ denote consumer i's perception of the content of the n'th ad sent by firm s (recall that firms can send multiple ads). Then, $m_{i,s,n}$ is given by:

$$m_{i,s,n} = \begin{cases} x_s \text{ with probability } q \\ -x_s \text{ with probability } 1-q \end{cases} \tag{2}$$

where $q \geq \frac{1}{2}$.[13] Thus, there is a probability q that the message is interpreted correctly by the consumer (i.e., that the consumer's perception of the product attribute is equal to the actual attribute). Notice that in this setting an ad from one firm provides information on the attributes of the other firm's product as

[12]Notice, though, that if we were studying an asymmetric setting (e.g., two consumers with $\beta = 1$ and one consumer with $\beta = -1$) prices may have had a signaling role.
[13]Notice that any $q \neq \frac{1}{2}$ is informative. Here, we consider the natural case of $q > \frac{1}{2}$: in other words, "correct messages are more likely than incorrect messages".

well. Specifically, any message that indicates that firm 1 is offering the product whose attribute is equal to 1 ($x_1 = 1$) obviously also indicates that firm 2 offers the product whose attribute is equal to -1 ($x_2 = -1$).[14] Thus, based on this single message, the consumer updates her expected utility from *both* products. Notice that this phenomenon is immediate from product differentiation and consumers' knowledge (i.e., that each consumer knows that there is one firm that is better for her than the other).

Media Channels The first consumer is exposed only to media channel 1, while the second is only exposed to media channel 2. Notice that this is a simplifying assumption that is consistent with the stylized fact that preferences are correlated with media channels, as described earlier.[15] Later, in Section 5 we relax this assumption and demonstrate that while the precise characterization of the equilibrium is different, the qualitative conclusions are not.

To simplify the model's presentation and notation, we focus henceforth on consumer 1 and suppress the consumer and media channel indices. It is easy to show that the results for consumer 2 are mirror images of those for consumer 1.

Furthermore, in order to make the presentation more intuitive, it is convenient to refer to firm 1 as being of type H and firm 2 of type L. (Recall that firm 1 offers a product $x_1 = 1$, and firm 2 offers a product $x_2 = -1$. Since the consumer's taste parameter is $\beta_1 = 1$, firm 1's product fits better with the consumer's tastes than the product offered by firm 2.)

Firms' objective function The cost of sending an ad is $c > 0$. It is assumed that c is not too big (otherwise, sending ads might never be profitable). The profit of firm s is:

$$\pi_s(n_s) = d_s - cn_s \tag{3}$$

where n_s is the number of ads that firm s places in the media channel.[16]

Sections 3 and 4 characterize the equilibrium of this model under different assumptions. In Section 3, consumers' decisions depend only on the content of the ads, whereas in Section 4 they also use the number of ads sent by each firm

[14]For example, a consumer who received one message that was $m_{i,1,1} = 1$ would, using Bayes' rule, form the following expectations about the product attributes: $E(x_1|m_{i,1,1} = 1) = (2q - 1)$ and $E(x_2|m_{i,1,1} = 1) = (1 - 2q)$. Based on this single message, the consumer updates her expected utility from *both* products. Specifically, since $q > \frac{1}{2}$, $E(\beta_1 x_1|m_{i,1,1} = 1) > 0$ and $E(\beta_1 x_2|m_{i,1,1} = 1) < 0$.

[15]The basic setup of our model is similar to Meurer and Stahl (1994) in the sense that they also focus on advertising for horizontally differentiated products. However, they do not allow for targeting of advertisements (from, say, firm 1 to consumer 1), and assume that advertising is perfectly informative ($q = 1$).

[16]The hidden assumption behind this profit function is that the difference between any product's price and marginal cost is 1.

as a signal about the firm's identity. Since the consumers described in Section 3 do not use all the relevant information we refer to them as naïve.

3 Equilibrium with a naïve consumer

Before examining the signaling equilibrium, it is useful to start by characterizing the equilibrium when the consumer is naïve, for two main reasons. First, it clarifies the incentives for consumers to behave strategically. Second, it introduces some formal functions that are useful in the analysis of the signaling equilibrium. Specifically, as will be demonstrated in the next section, when the off-the-equilibrium beliefs (of the signaling game) are being provoked the consumer is using only the information from the content of the ads. In such a case, the profit functions of the firms will be identical to the ones discussed in this section.

We start by demonstrating that in equilibrium only firm 1 (i.e., type H) sends ads. Then, we solve for 1's optimal number of ads, n^{nc} (where we use nc in order to represent "naïve consumer").

Using Bayes rule, it is easy to show that the probability that s is of type H (denoted by μ_s^0) based only on the content of the ads is:

$$\mu_s^0(g_s) = \left[1 + \left[\frac{1-q}{q}\right]^{g_s}\right]^{-1}$$

where g_s represents the number of ads that indicate that firm s is of type H minus the number of ads indicating otherwise. Formally, $g_s = \sum_{j=1}^{n_s} m_{s,j} - \sum_{j=1}^{n_{s'}} m_{s',j}$ where $s' = 3 - s$ (i.e., s' is the competitor of s). This means that g_s is a sufficient statistic in this case. Notice that since an ad from one firm provides information on the attributes of the other firm's product also, g_s depends on the ads of both senders.

The consumer's decision rule is quite simple. She selects the firm s for which g_s is positive. If g_s is equal to zero, she selects one of the firms randomly with probability 0.5.

From the firms' point of view g_s is a random variable, denoted by $\widetilde{g}_s$. The expected payoff function of firm s is:

$$E(d_s|n) - cn_s = \left[\Pr(\widetilde{g}_s > 0|n) + 0.5\Pr(\widetilde{g}_s = 0|n)\right] - cn_s \tag{4}$$

where n represents the number of ads sent by both firms together—i.e., $n = n_1 + n_2$. The probability functions in Eq. 4 are described and discussed in the Appendix A.1. In the appendix we also establish the intuitive result that the more ads to which a consumer is exposed, the more likely she is to choose the product that is the right match for her. This follows from the informative (albeit noisy) nature of advertising. Now, since the probability that the right firm is selected is increasing in n, it should follow that in any equilibrium firm 2 should not want to incur the cost of sending ads, but firm 1 should. The appendix then includes two lemmas that lead to the following proposition.

Proposition 1 (Equilibrium with a with a naïve consumer) *In any equilibrium (with a naïve consumer), firm* 2 *does not send any ads, and firm* 1 *sends* $n_1 = n^{nc} > 0$*, where* n^{nc} *is finite.*

Proof See Appendix A.1. □

This result means that while firm 1 sends ads, its competitor does not. Let π^{nc} denote the expected payoff of firm 1 in this equilibrium (this notation is used in the next section).

Recall that we are focusing on consumer 1 (and, thus, media channel 1) for simplicity. It is immediate to show that when firms compete over two consumers, firm 1 sends ads only in the media channel to which consumer 1 is exposed and firm 2 sends ads only in the media channel to which consumer 2 is exposed. That is, advertising is targeted to the consumer whose tastes best fit the product attributes.

In equilibrium, $\Pr(d_1 = 1|n = n^{nc}) < 1$. This results from the fact that $q < 1$ and n^{nc} is finite. In other words, since ads are noisy, there is a positive probability that the consumer purchases a product that does not fit her preferences. Consumers might minimize the probability of such mistakes by acting strategically and incorporating the information in firms' strategies in their decisions. Specifically, a consumer might be able to infer firms' identities from their actions. The next section examines the implications of such strategic behavior.

It is often thought that as advertisements become more precise the tendency to advertise increases. It turns out that this is not the case for the model here. Specifically, the effect of q on n^{nc} is not monotonic. In Appendix A.1. we show that the marginal expected revenue is represented by $(q - 0.5)\Pr(\widetilde{g}_s = 0|n)$.[17] Notice that for both $q = 0.5$ and $q = 1$ the expected marginal revenue is 0. When $q = 0.5$, the marginal expected revenue is equal to zero because ads are not informative and therefore do not change the consumer's behavior. When $q = 1$, ads are not noisy and the probability of a tie is zero. In this case as well, the marginal expected revenue is zero. When q increases from 0.5, the probability of a "tie" (i.e., $\Pr(\widetilde{g}_s = 0|n)$) diminishes, but the effectiveness of ads in informing the consumers (which is $q - 0.5$) increases. The first effect reduces the incentive to send more ads, while the second increases it. It can be shown that the first effect is more important for q close to 0.5, while the second dominates for q close to 1.

[17] The role of the probability of a "tie" (i.e., $\Pr(\widetilde{g}_s = 0|n)$) is similar to the one played by the "pivotal voter" in elections (see, for example, Shachar and Nalebuff 1999). The logic behind this role is that an additional (i.e., marginal) ad would not affect the consumer if g_s is far enough from zero (say $g_s = 2$). The only way that one ad can change the expected revenue is if it creates a tie or break a tie. For a more formal intuition, please follow the proof of Lemma 8.

4 Signaling equilibrium

Here, we examine the case where consumers incorporate both the statistical information that is revealed in the content of the ads, and the signaling information that is revealed by the number of advertisements, when forming their expectations about product attributes. We show that there exists a separating equilibrium in which firm 1 sends ads and firm 2 does not, although (unlike the case discussed in the previous section) by sending ads firm 2 can increase its expected revenue. In addition, consumers have no uncertainty about product attributes in equilibrium.

We start by specifying the beliefs of consumers and the strategies of firms, and then show that these beliefs and strategies are part of a separating equilibrium. First, we present a few definitions. Denote the posterior probability function that firm s is of type H as $\mu_s^1(n_s, n_{s'}, g_s)$ (recall that $s' = 3 - s$). These beliefs are given by:

$$\mu_s^1(n_s, n_{s'}, g_s) = \left\{ \begin{array}{c} 1 \text{ if } n_s \geq n^* \text{ and } n_{s'} < n^* \\ 0 \text{ if } n_s < n^* \text{ and } n_{s'} \geq n^* \\ \mu_s^0(g_s) \text{ otherwise} \end{array} \right\} \tag{5}$$

where n^* satisfies the conditions:

$$1 - cn^* \geq \max_{0 \leq n} E(d_1|n) - cn$$

$$\text{and } 0 \geq E\left(d_2|2n^*\right) - cn^*$$

The logic behind these beliefs can be stated as follows. In equilibrium only one firm, H, should find it profitable to send n^* ads. Thus, if one firm sends n^* (or more) ads while its competitor sends fewer ads than that, the identity of the firms is revealed perfectly. Otherwise, the consumer bases her expectation regarding firm type only on the statistical information revealed via the ads, and not on the intensity of advertising.

An interesting feature of these beliefs is the role of competition. The existence of a competitor provides additional information to the receiver regarding whether or not a sender has deviated—namely, deviation is known to occur when the actions of both senders are the same (say, both send at least n^* ads). In the standard one-sender, two-type model, this channel is not present and, as a result (as demonstrated below) there is no separating equilibrium. Another way to see this is that in a standard one-sender two-type model either the cost or the benefit would need to be type-specific in order to get a separating equilibrium; in our setting this assumption is not needed. Thus, competition between the senders plays an important role in the model and resulting equilibrium. It is worth noting that although competition is a common feature of most IO models, it is typically absent from signaling games.

The logic behind the two conditions for n^* is the following. The first condition requires that at n^* firm 1 does not want to deviate. Notice that $1 - cn^*$ is the profit of firm 1 in equilibrium while $\max_{0 \leq n} E(d_1|n) - cn = \pi^{nc}$, which is simply the expected payoff of firm 1 in the equilibrium with a naïve consumer. In other words, π^{nc} is the expected payoff of firm 1 if it deviates and provokes the off-the-equilibrium beliefs, leading consumers then to only consider the statistical information that is revealed in the content of the ads. The second condition requires that at n^* firm 2 does not want to deviate. In equilibrium its profit is zero. However, if it were to imitate firm 1 and send n^* ads then (i) the number of ads received by the consumer would be $2n^*$ and (ii) the off-the-equilibrium beliefs will be provoked leading the consumer to consider only the information in the content of the ads and leaving firm 2 with the expected market share of $E(d_2|2n^*)$.

Next, we present three Lemmas that are used in the main proposition. In Lemma 2, we show that firm 2 will not find it profitable to mimic the other firm if n^* exceeds some number $\underline{n}$. Then, in Lemma 3, we show that there exists some number $\overline{n}$ such that for all $n^* \leq \overline{n}$, firm 1 will not find it profitable to deviate from its equilibrium strategy and send fewer than n^* ads. Finally, Lemma 4 asserts that $\overline{n} \geq \underline{n}$ for $q > 0.5$. It follows that for all $\underline{n} \leq n \leq \overline{n}, n = n^*$ can be supported in a perfect Bayesian equilibrium.

Lemma 2 ($\underline{n}$) *There exists an* $\underline{n} > 0$ *that satisfies the conditions:* $\left[E(d_2|2\underline{n}) - c\underline{n}\right] < 0$ *and* $\left[E(d_2|2\left(\underline{n} - 1\right)) - c\left(\underline{n} - 1\right)\right] \geq 0$. *Furthermore, for any* $n > \underline{n}$, $\left[E(d_2|2n) - cn\right] < 0$.

Proof $E(d_2|2n) - cn$ is the expected payoff function of firm 2 from imitating firm 1. This function is (a) equal to 0.5 at $n = 0$, (b) approaches $-\infty$ as $n \to \infty$, and (c) is monotonically decreasing in n (since $E(d_2|2n)$ is a non increasing function in n—see Lemma 8 in Appendix A.1.—and cn is increasing in n.) This establishes the existence of an $\underline{n}$ that satisfies the conditions in the first part of the lemma. The second part of the lemma follows from the fact that $E(d_2|2n) - cn$ is monotonically decreasing in n. □

Lemma 3 ($\overline{n}$) *There exists an* $\overline{n}$ *that satisfies the conditions:* $1 - c\overline{n} \geq \pi^{nc}$ *and* $1 - c(\overline{n} + 1) < \pi^{nc}$. *Furthermore, for any* $n < \overline{n}$, $1 - cn > \pi^{nc}$.

Proof $1 - cn - \pi^{nc}$ is (a) positive at $n = 0$ (since π^{nc} must be smaller than 1), (b) approaches $-\infty$ as $n \to \infty$, and (c) monotonically decreasing in n. □

Lemma 4 *For* $q > 0.5$, $\underline{n} \leq \overline{n}$.

Proof Since $1 - cn - \pi^{nc}$ is monotonically decreasing in n, we only need to show that $1 - c\underline{n} - \pi^{nc} \geq 0$.

Recall that $\pi^{nc} = E(d_1|n^{nc}) - cn^{nc}$, and thus $1 - c\underline{n} - \pi^{nc} = \left[1 - E(d_1|n^{nc})\right] + c\left[n^{nc} - \underline{n}\right]$.

If $n^{nc} \geq \underline{n}$, it is immediate that $\left[1 - E(d_1|n^{nc})\right] + c\left[n^{nc} - \underline{n}\right] > 0$.

Next, consider the case where $n^{nc} < \underline{n}$. $\left[1 - E(d_1|n^{nc})\right] + c\left[n^{nc} - \underline{n}\right] = E(d_2|n^{nc}) - c\left[\underline{n} - n^{nc}\right] \geq E\left[d_2|(\underline{n}-1)\right] - c\left[\underline{n}-1\right] \geq E\left[d_2|2(\underline{n}-1)\right] - c\left[\underline{n}-1\right] \geq 0$ by Lemma 2. □

We are now ready to state the main result:

Proposition 5 (PBE) *For $q > 0.5$, there exists a perfect Bayesian equilibrium where beliefs are given by $B1$, firm 2 sends no ads, and firm 1 sends n_1 ads, where $n_1 = n^* \in \left[\underline{n}, \overline{n}\right]$.*

Proof Given beliefs and $n_1 = n^* \in \left[\underline{n}, \overline{n}\right]$, firm 2 will choose not to deviate from its pure strategy of sending zero ads since: (a) any $n_2 < n^*$ involves a cost without any revenues, (b) $n_2 = n^*$ leads to losses from Lemma 2, and (c) it is immediate that $n_2 > n^*$ leads to even larger losses.

Given beliefs and the fact that firm 2 sends no ads, firm 1 will optimally choose $n_1 = n^* \in \left[\underline{n}, \overline{n}\right]$ since, (a) from Lemma 3, any $n_1 < n^*$ leads to lower expected payoff, and (b) any $n > n^*$ involves additional costs without any increase in revenues.

It is trivial to show that beliefs $B1$ are consistent, given these strategies of firms. □

Thus, in equilibrium firm 1 sends ads and its competitor does not. All uncertainty about product attributes is resolved and the consumer chooses the product that is best for her with certainty.

A separating equilibrium obtains despite the assumption that the benefits and the costs are not type specific (i.e., benefits and costs are the same for both firms). Separation is enabled by the informative nature of advertising: specifically, although the content of the ads is noisy, in expectation ads "favor" firm 1 over 2 (for the relevant consumer). While the content of the ads is not used on the equilibrium path, it is used by the consumer off-the-equilibrium path. Thus, if the two firms were to send n^* ads, the profit of firm 1 would be higher than that of firm 2. The proof demonstrates that at n^* (i) firm 2 makes losses from imitating firm 1 and (ii) firm 1's profit is higher than its profit in the equilibrium with naïve consumers (which represents its profit from deviation).

Notice an interesting difference from the case where consumers are naïve. There, firm 2 did not send any ads since by doing so it increased its costs and decreased its expected revenues (due to the informative nature of ads). In the case studied here, by contrast, the expected revenue of firm 2 in equilibrium is zero and by sending ads (i.e., mimicking firm 1) it can increase its expected revenue. However, this increase is not large enough to cover the associated costs.

Once again, recall that we are focusing on consumer 1 (and, thus, media channel 1) for simplicity. It is immediate to show that when firms compete over two consumers, firm 1 sends ads only in the media channel to which consumer 1 is exposed and firm 2 sends ads only in the media channel to which consumer

2 is exposed. That is, advertising is targeted to the consumer whose tastes best fit the product attributes.

The equilibrium number of ads in each channel depends on q and c. We examine here the effect of these parameters on n^*, for $n = \underline{n}$.[18] Recall that $\underline{n}$ is the threshold of the number of ads that deter firm 2 from mimicking firm 1, i.e., $\left[E(d_2|2\underline{n}) - c\underline{n}\right] \cong 0$. It is trivial to show that an increase in either c or q decreases $n^* = \underline{n}$. The effect of costs on n^* is standard. The effect of q is also quite intuitive. Specifically, by mimicking firm 1, firm 2 provokes the off-the-equilibrium beliefs. Its expected revenue in such a case depends on the precision of ads—as ads become more precise, the consumer is more likely to prefer firm 1 over firm 2. Thus, as q increases, the incentive of firm 2 to mimic firm 1 decreases.

An interesting feature of the separating equilibria is that, in equilibrium, consumers ignore the content of the ads. Moreover, doing so leads them to make the best choices. This is due to the fact that the content of the ad is noisy while (in equilibrium) the actions of the firms are not. As a result even when the content of ads aired by firm 1 indicates that its competitor offers the best product for the consumer (although this is not too likely to occur, it is possible as long as $q < 1$), she will, correctly, select the product of firm 1. In other words, in the perfect Bayesian equilibrium, consumers always choose the product that gives them the highest utility. In contrast, recall that when the consumer is naïve, she occasionally chooses a product that is not the best fit for her. As q increases, however, the probability of mistakes falls.[19]

Next, compare the number of ads aired by firms in the signaling equilibrium with the "naïve equilibrium". One can show that $n^* \leq n^{nc}$ when q is not too low or too high, otherwise $n^* > n^{nc}$. The reason is that although both n^* and n^{nc} depend on the informativeness of ads, their dependence on q is due to different sources. While the effect of q on n^* is monotonic, its effect on n^{nc} is not. Recall that in the naïve case, the incentive to send ads is low when q is either small or large.

The previous proposition studied the PBE when $q > 0.5$. The following Lemma and Proposition examine the case of $q = 0.5$.

Lemma 6 *For $q = 0.5$, $\underline{n} > \overline{n}$.*

[18]This may be argued to be the most reasonable equilibrium, both positively and normatively. Since, in any equilibrium, individuals choose firms whose products best fit their tastes, and the firm's revenue is the same for any $n^* > \underline{n}$, it follows that any $n^* > \underline{n}$ results in firms incurring higher costs without affecting firms' revenues or consumer welfare.

[19]The probability of mistakes depends both on the noise in ads and on the number of ads aired. As q increases, ads are less noisy so that the probability of mistakes falls. On the other hand, the number of ads aired by each firm may fall with q (recall that n^{ns} is not monotonic in q). It is easy to show, however, that the first effect dominates the second as q increases.

Proof When $q = 0.5$, for any n, we get that $E(d_1|n) = 0.5$ and $E(d_2|n) = 0.5$. Thus, it is easy to show that $\underline{n} > \frac{1}{2c} > \overline{n}$. □

Proposition 7 *When* $q = 0.5$ *there does not exist any perfect Bayesian equilibrium in which beliefs are given by* $B1$.

Proof Since $\underline{n} > \overline{n}$ there is no n^* that can satisfy the conditions of beliefs $B1$. □

This means that when $q = 0.5 + \varepsilon$ a separating equilibrium exists, but when $q = 0.5$ it does not hold any more. In other words, the signaling role of advertising obtains even when ads are extremely, though not completely, noisy. The reason is that if the information content in a message were arbitrarily close to zero, firms can still reveal their type by endogenously increasing the number of ads they send. However, if there were *no* information content in ads, then more ads convey no additional information, so the separating equilibrium breaks down.

An important assumption behind this discontinuity result is that when firms send ads, the relevant consumer is exposed to them with certainty. If, instead, the exposure to the ads sent is not deterministic, then the consumer is uncertain about the exact number of ads sent by firms through the media channels to which she is exposed.[20] In such a case, one would expect that the consumer does not rely exclusively on ad intensity in forming beliefs about product attributes, but rather on both ad intensity and ad content.

While the discontinuity result is sensitive to the assumption of perfect observability, it still provides interesting insight into the question of how informative ads *need* to be. Previous theories of content-free advertising were motivated by the fact that "a nontrivial amount of advertising (especially on television) has little or no obvious informational content" (Milgrom and Roberts 1986). Becker and Murphy (1993) similarly write that "it is obvious that many ads provide essentially no information." The discontinuity result here offers a different perspective, by creating a sharp distinction between ads with very little information and ads with no information. Specifically, the *effectiveness* of ads with arbitrarily little information content can be quite different than for ads with no information—implying that "little information" can be "enough information."

[20] Hertzendorf (1993) revised the model of "advertising as a signal" a-la Nelson to account for the possibility that consumers may see fewer ads than were originally purchased. While he found that ads can still serve as a signal, he also demonstrated that "…under reasonable condition…price and advertising expenditures would never be simultaneously employed to signal quality."

Finally, in a working paper version,[21] we show that a pooling equilibrium does not exist when beliefs are $B1$, and that the separating equilibrium holds for a more general class of beliefs.

5 Extensions

This section relaxes certain assumptions of the simple model presented above, in order to both enrich the model and to shed light on the robustness of the results. We consider two extensions. First, we relax the assumption of perfect correlation between consumer preferences and media choices. Second, we include price in the analysis.

5.1 Imperfect correlation between preferences and media choices

So far we have assumed that the firms can expect to find consumer 1 only in media channel 1 and consumer 2 only in the other media channel. Here we relax this "perfect correlation" assumption. There are various ways to formulate the model so that imperfect correlation obtains. Some formulations, however, imply that additional assumptions are being relaxed as well; in order to have a clean examination of the "perfect correlation" assumption we adopt an extension in which only this assumption is being changed.

Let the probability that consumer i is exposed to media channel k be γ for $k = i$ and $(1-\gamma)$ for $k \neq i$. As long as $\gamma \neq \frac{1}{2}$ media choices are type specific (i.e., there is a correlation between preferences and media choices). For an easier interpretation we assume that $\gamma \in \left(\frac{1}{2}, 1\right)$. The sequence of events is as follows: (1) first, firms decide on their advertising spending without knowing the composition of consumers in each media channel, (2) next, the consumers choose a media channel (based on the probabilities above) and are exposed to the relevant ads, and (3) last, consumers choose a product. Under this formulation, while firms do not know the audience composition of each media channel, they know the media choice probabilities. Note that consumers' choice of media is for the duration of the game.[22]

The solution of this revised model is described in Appendix A.2. Since the media choices are still type specific, the central qualitative result of the

[21] This version is available from the authors upon request.

[22] (a) Assuming otherwise (e.g. there are multiple time periods within the duration of the game and the consumer can switch between the media channels) would lead to additional significant changes in the model. (b) We get similar results to those obtained below for the following setting: there is a unit mass of consumers of each type; a fraction γ of segment 1 watches media channel 1 and the rest watches media 2; and, the viewing habits of the other segment are the same. Thus, γ of the audience of media 1 is of segment 1 and $(1-\gamma)$ is of segment 2.

model remains—i.e., there is a separating equilibrium in which each firm sends ads only through the media channel that is more likely to be chosen by "its" consumer. However, since the media choice is probabilistic (rather than deterministic) the precise characterization of the equilibrium is somewhat different.

To examine the impact of γ on the number of ads in equilibrium we focus (as earlier) on the case of $n = \underline{n}$. We find that as the media channels become less precise (i.e., γ decreases toward $\frac{1}{2}$) the number of ads decreases. The reason for this is interesting. In the separating equilibrium the expected revenue of firm s among consumers who are exposed to media channel $k = s$ is γ. This is due, of course, to the possibility that the consumer who prefers its product will not be exposed to this media channel. As the media channels become less precise this revenue decreases and as a result the incentive of firm s' to imitate firm s is diminishing. Thus, the number of ads required to deter the "wrong" firm from imitating (and thus the number of ads needed to separate between the players) is smaller.

5.2 Endogenous prices

As discussed above, since our research focuses on horizontally differentiated products in a symmetric setting, it is not reasonable to expect that prices will play a signaling role. Nevertheless, it is interesting to endogenize prices. Accordingly, we study a case in which consumers' utility from product s is $\alpha + \frac{1}{2}\beta_i x_s - p_s$ (where p stands for price) and firms set prices that maximize their profits, but prices are not included in consumer beliefs. In Appendix A.3. we characterize this model and its solution and show that (as long as α is not too high) in equilibrium $p_s = \alpha + \frac{1}{2}$; and, $n_s = n^* > 0$ in media channel $k = s$ and 0 in media channel $k \neq s$. Thus, the profit of each firm is $\alpha + \frac{1}{2} - cn^*$. This means that the signaling role of media selection does not disappear even when prices are accounted for (as long as α is not too high).

However, when α is high, it is impossible to find an n^* that can deter the "wrong" firm from imitating and at the same time is not too high for the "right" firm (in the sense that the "right" firm is not satisfied with the off-the-equilibrium profit). The reason is simple and can be described as follows. Consumers' willingness to pay for the product is high, and thus the difference between the two products (due to their attributes) is relatively small. When prices are endogenous the "wrong" firm can "compensate" the consumer via lower prices, and still make a nice profit. As a result it is more difficult to deter such firm from imitating. For more details see Appendix A.3.

6 Conclusion

This paper studies a model where firms target advertisements to particular consumers, and advertising is noisy. We characterize the resulting equilibrium and use it to shed light on various questions. We show that targeting can serve as a signal on product attributes. Furthermore, the effectiveness of targeting in equilibrium may (far) exceed the information contained directly in the targeted message. At the same time, information content is not superfluous since it affects consumer beliefs off the equilibrium path. These results reveal how advertising conveys information both through the content of the message and the firm's choice of advertising medium. We conclude here by discussing some implications for firm strategy, welfare, and future research.

The analysis sheds new light on the effectiveness of targeting strategies. Conventional wisdom is that targeting is effective due to the fact that it avoids waste by not sending ads to consumers who are not likely to buy the product. The model here suggests a new role for targeting—firms' media selection can serve as a signal on product attributes. The reason is that consumers know that in equilibrium it is optimal only for the firm that offers products that best fit her tastes to advertise in the media channels to which she is exposed.

This reasoning builds on a central, albeit hidden, assumption in the model: the consumer knows that firms know her tastes (and target their ads accordingly). Without this assumption, targeting cannot serve as a signal. This also suggests an interesting distinction between targeting and personalization of ads. Personalization means that the firm informs the recipient she is being targeted. Worth noting is a recent trend among firms towards personalization of ads. For example, Google and Amazon make an effort to ensure that the recipient of an ad knows that the ad was designed for people like her.[23] The prediction that strategies of targeting might have different effectiveness than strategies of personalization also lends itself easily to experimental testing.

The model can shed light on empirical research in advertising as well, and in turn its welfare analysis. Specifically, empirical work finds that advertising exposures affect purchase probabilities even after accounting for an informative effect.[24] This "direct effect" of advertising has typically been interpreted as 'persuasion', or sometimes even 'brainwashing' and 'manipulation of tastes'. The study here suggests a different behavioral source of such a direct effect—namely, strategic Bayesian consumers. Specifically, even with noisy informative content, ad exposures serve as a signal about the positive match between the consumer's tastes and attributes of the promoted product. This implies that (a) consumers should respond positively to advertising exposures and (b) the informative effect of ads in equilibrium is greater than that simply due to consumer learning from ad content. Therefore, the observed correlation between ad intensity and purchase probability in empirical research may result from informative advertising rather than a direct effect on utility.

[23]Consider Gmail, the mail service offered by Google. This service includes a very large memory and no spam. Instead, Google guarantees that commercial mail will be directly linked to the interest of the recipient. To achieve this, the firm scans all the personal mail received by the client and determines her area of interest. The reason this strategy goes beyond simple targeting is that Google ensures that the recipient is aware that the ads were designed for people like her. Targeting, on the other hand, does not guarantee that the recipient knows she is being targeted.

[24]See, for example, Anand and Shachar (2005).

Appendix

A.1 Characterizations and proofs of the "naïve equilibrium"

The probability functions in Eq. 4 can be rewritten as

$$\begin{aligned} \Pr(\widetilde{g}_s > 0|n) &= \sum_{k>\frac{n}{2}} \binom{n}{k}(q_s)^k(1-q_s)^{(n-k)} \\ \Pr(\widetilde{g}_s = 0|n) &= \begin{cases} \binom{n}{\frac{n}{2}}\left[q(1-q)\right]^{\frac{n}{2}} & \text{when } n \text{ is even} \\ 0 & \text{when } n \text{ is odd} \end{cases} \end{aligned} \tag{6}$$

where $q_1 = q$, and $q_2 = 1 - q$. For firm 1, $\Pr(\widetilde{g}_s > 0|n)$ depends on all the events in which the number of "correct" realizations of the messages is larger than the number of "incorrect" ones. For its competitor, obviously, the opposite holds. The probability of a "tie" (i.e., the number of correct and incorrect messages canceling each other out), $\Pr(\widetilde{g}_s = 0|n)$, is the same for both senders. It is a positive value when n is even, and zero otherwise.[25]

We start by showing that the probability that the consumer selects firm 1 increases (weakly) in the total number of ads received, n.

Lemma 8 (Ads increase probability of selecting firm 1) *$E(d_1|n)$ is a non decreasing function in n.*

Proof $E(d_1|n)$ is equal to:

$$\Pr(\widetilde{g}_1 > 1|n) + \Pr(\widetilde{g}_1 = 1|n) + 0.5\Pr(\widetilde{g}_1 = 0|n)$$

Notice that:

$$\Pr(\widetilde{g}_1 > 0|n+1) = \Pr(\widetilde{g}_1 > 1|n) + q\Pr(\widetilde{g}_1 = 1|n) + q\Pr(\widetilde{g}_1 = 0|n)$$

and

$$\Pr(\widetilde{g}_1 = 0|n+1) = (1-q)\Pr(\widetilde{g}_1 = 1|n) + q\Pr(\widetilde{g}_1 = -1|n)$$

Thus,

$$\begin{aligned} &E(d_1|n+1) - E(d_1|n) \\ &\quad = (q-0.5)\Pr(\widetilde{g}_1 = 0|n) + 0.5\left[q\Pr(\widetilde{g}_1 = -1|n) - (1-q)\Pr(\widetilde{g}_1 = 1|n)\right] \end{aligned}$$

Notice that the first term is relevant only when n is even, and that the second term is relevant only when n is odd.

[25] Notice that when n is even, there is a positive probability that the number of "correct" messages equals the number of "incorrect" ones. If n is odd, of course, this event can never occur.

Furthermore, the second term is equal to zero, since:

$$\Pr(\widetilde{g}_1 = 1|n) = \binom{n}{\frac{n+1}{2}} q^{\frac{n+1}{2}} (1-q)^{\frac{n-1}{2}},$$

$$\Pr(\widetilde{g}_1 = -1|n) = \binom{n}{\frac{n-1}{2}} q^{\frac{n-1}{2}} (1-q)^{\frac{n+1}{2}}, \text{ and } \binom{n}{\frac{n+1}{2}} = \binom{n}{\frac{n-1}{2}}.$$

Thus, when n is odd $\left[E(d_1|n+1) - E(d_1|n)\right] = 0$ and when n is even it is equal to $(q - 0.5) \Pr(\widetilde{g}_1 = 0|n)$.

Since $q > 0.5$, $E(d_1|n)$ is a non decreasing function in n. □

Since the consumer prefers type H over type L, it is not surprising that getting additional informative ads increases the probability that she will choose firm 1. This also means that by sending ads to the consumer, firm 2 (weakly) decreases its expected payoff. Thus, one would expect that firm 2 would not send ads and firm 1 would (this is established in Proposition 1 in the text). We proceed to prove this result.

To begin with, it is easy to show (and this is also clear from the proof of Lemma 8) that the change in $E(d_1|n)$ when n increases by 1 is:

$$\begin{array}{ll} (q_s - 0.5) \Pr(\widetilde{g}_s = 0|n) & \text{when } n \text{ is even} \\ 0 & \text{when } n \text{ is odd} \end{array} \tag{7}$$

This means that the increase in the expected payoff (as a result of an additional ad) depends on the probability that the realizations of the previous ads lead to a "tie". This can only happen when the number of ads is even. When n is odd, the change in the firm's expected payoff from sending one more ad is zero.[26]

The following Lemma shows that the probability of a tie, and therefore the marginal expected revenue of firm 1, decreases in n.

Lemma 9 (Probability of a tie falls with more ads) $\Pr(\widetilde{g}_1 = 0|n)$ *is a non increasing function in n for even n.*

Proof For even n, $\Pr(\widetilde{g}_1 = 0|n) = \binom{n}{\frac{n}{2}} \left[q(1-q)\right]^{\frac{n}{2}}$.

---

[26]When n is odd, the expected revenue of 1 can increase only if $g_1 = -1$ and the perceived content of the additional ad is 1 (in which case it creates a tie). The probability of the first event is $\Pr(g_1 = -1|n)$ and the probability of the second event is q. Since the increase in the expected revenue in such a case is $\frac{1}{2}$ (going from 0 to $\frac{1}{2}$), the unconditional increase in the expected revenue is $\frac{1}{2} q \Pr(g_1 = -1|n)$. Accordingly, it is easy to show that the unconditional decrease in the expected revenue is $\frac{1}{2}(1-q) \Pr(g_1 = 1|n)$. Thus, in order to show that for an odd n an additional ad does not change the expected revenue, we need to show that $q \Pr(g_1 = -1|n) = (1-q) \Pr(g_1 = 1|n)$. This is immediate from the fact that for any set of realizations that leads to $g_s = -1$ there is a "mirror" set that leads to $g_s = -1$ (for example, $\{-1, 1, -1\}$ and $\{1, -1, 1\}$). It is obvious that the probability of the first set is equal to $\frac{1-q}{q}$ times the probability of the second set.

It is easy to show that

$$\binom{n+2}{\frac{n+2}{2}} = 4\left(\frac{n+1}{n+2}\right)\binom{n}{\frac{n}{2}}$$

Thus,

$$\frac{\Pr(\widetilde{g_1}=0|n+2)}{\Pr(\widetilde{g_1}=0|n)} = 4\left(\frac{n+1}{n+2}\right)\left[q(1-q)\right]$$

Since $q > 0.5$, we get that $4\left[q(1-q)\right] < 1$. Obviously $\left(\frac{n+1}{n+2}\right) < 1$, and thus $\Pr(\widetilde{g_1}=0|n+2) < \Pr(\widetilde{g_1}=0|n)$. □

The rationale behind this Lemma is the following. Notice that the expected value of g_s is $n(2q-1)$. Since $q > 0.5$ this is bigger than 0. Therefore, the probability that the number of incorrect realizations equals the number of correct realizations is diminishing in n.

Lemma 9, coupled with the fact that the marginal cost of an ad is constant, suggests that as long as c is not too large, firm 1 should send a positive (and finite) number of ads, and firm 2 should not send any ad. This is established in the Proposition 1, which is presented in the text and its proof appears next.

Proof [Proposition 1] We first show that firm 2 does not send any ads in equilibrium. Lemma 8 states that $E(d_1|n)$ is non-decreasing in n. Thus, each ad sent by 2 weakly decreases its expected payoff (which is $1 - E(d_1|n)$). Since ads are costly ($c > 0$) sending *any* ad decreases the payoff of 2. Therefore, sending no ads is a dominant strategy for firm 2.

Next, we show that $n^{nc} > 0$.

Notice that n^{nc} must be an odd number. Recall, from Lemma 8, that, when n is odd, the change in $E(d_1|n)$ when n increases by 1, is equal to zero. Thus, for an even n, if firm 1 were to decrease the number of messages that it sends by 1, it does not change its expected revenues but lowers its cost.

Furthermore, n^{nc} should satisfy the following conditions:

$$\begin{aligned}\left[E(d_1|n^{nc}) - cn^{nc}\right] - \left[E(d_1|n^{nc}-2) - c\,(n^{nc}-2)\right] \geq 0\\ \left[E(d_1|n^{nc}) - cn^{nc}\right] - \left[E(d_1|n^{nc}+2) - c\,(n^{nc}+2)\right] > 0\end{aligned} \tag{8}$$

In other words, increasing or decreasing n^{nc} by 2 results in negative marginal payoff for firm 1. (Notice that we add (and subtract) 2 from n^{nc} since we are only considering odd numbers).

Using Eq. 7, these conditions can be re-written as

$$\begin{aligned}\Pr(\widetilde{g_1}=0|n^{nc}-1) \geq \frac{2c}{(q-0.5)}\\ \Pr(\widetilde{g_1}=0|n^{nc}+1) < \frac{2c}{(q-0.5)}\end{aligned} \tag{9}$$

Since $\Pr(\widetilde{g_1}=0|n)$ is a weakly decreasing function in n and c is not too large, there exists an $n^{nc} > 0$ that satisfies these conditions. □

A.2 Imperfect correlation between preferences and media choices

Here we describe the *main steps* in reproducing the model results under the assumption that each consumer might be exposed to any of the media channels.

To simplify the model's presentation and notation, we focus henceforth on media channel 1 and suppress the media channel indices.

A.2.1 Naïve equilibrium

The firms' objective function is:

$$E(\pi_s|n) = \gamma E\left(d_{1,s}|n\right) + (1-\gamma)\, E\left(d_{2,s}|n\right) - cn_s \tag{10}$$

where

$$E\left(d_{1,s}|n\right) = \sum_{k>\frac{n}{2}} \binom{n}{k} (q_s)^k (1-q_s)^{(n-k)} + \frac{1}{2} I\left\{n \text{ is even}\right\} \binom{n}{\frac{n}{2}} \left[q(1-q)\right]^{\frac{n}{2}}$$

and

$$E\left(d_{2,s}|n\right) = \sum_{k<\frac{n}{2}} \binom{n}{k} (q_s)^k (1-q_s)^{(n-k)} + \frac{1}{2} I\left\{n \text{ is even}\right\} \binom{n}{\frac{n}{2}} \left[q(1-q)\right]^{\frac{n}{2}}$$

Notice that this formulation accounts for the *possibility* that the perceptions of the two consumers of the same ads is not the same.

Thus

$$E(\pi_s|n) = (1-\gamma) + (2\gamma - 1)\Omega_s(n) - cn_s$$

where

$$\Omega_s(n) \equiv \left[\sum_{k>\frac{n}{2}} \binom{n}{k} (q_s)^k (1-q_s)^{(n-k)} + \frac{1}{2} I\left\{n \text{ is even}\right\} \binom{n}{\frac{n}{2}} \left[q(1-q)\right]^{\frac{n}{2}}\right]$$

Notice that $\Omega_s(n)$ is equal to (i) the value of $E(d_s|n)$ in the "perfect correlation model" (i.e., the one described in the text), and also to (ii) $1 - \Omega_{s'}(n)$. Furthermore, note that if $\gamma = 1$ we get that $E(\pi_s|n) = \Omega_s(n) - cn_s$. In other words, the "perfect correlation model" is nested in this one.

It is straightforward to reproduce the rest of the results for the naïve equilibrium. From Lemma 8 we know that $\Omega_1(n)$ is a non decreasing function in n, and thus it is immediate that (given $\gamma > \frac{1}{2}$) while firm 1 will send ads through media channel 1, firm 2 will not. To solve for n^{nc} we note that the change in $(1-\gamma) + (2\gamma - 1)\Omega_1(n)$ when n increases by 1 is $(2\gamma - 1)\sigma(n)$ where $\sigma(n) \equiv I\left\{n \text{ is even}\right\} (q - 0.5)\binom{n}{\frac{n}{2}} \left[q(1-q)\right]^{\frac{n}{2}}$ and, of course, $\sigma(n)$ is exactly the probability of a "tie" in the "perfect correlation model". Thus, as γ increases (i.e., media channel is more "precise") the effectivity of ads increases. Indeed, at $\gamma = 1$ the effectiveness is the same as in the "perfect correlation model". As a result n^{nc} (which is an odd number) should satisfy the following conditions:

$$\begin{aligned}(2\gamma - 1)\sigma(n^{nc} - 1) &\geq 2c \\ (2\gamma - 1)\sigma(n^{nc} + 1) &> 2c\end{aligned} \tag{11}$$

where the marginal cost is $2c$ since the firm is always choosing an odd number.

This result (in comparison to the "perfect correlation model") is very intuitive. It suggests that as γ decreases (toward $\frac{1}{2}$) firm 1 is decreasing the number of ads that it is sending, because it become less effective.

A.2.2 Signaling equilibrium

The only change in the beliefs compared to the "perfect correlation model" relates to the conditions that n^* satisfies. They are:

$$\gamma - cn^* \geq \max_{0\leq n}(1 - \gamma) + (2\gamma - 1)\Omega_1(n) - cn$$

$$\text{and } \Omega_2(2n^*) \leq \frac{cn^*}{(2\gamma - 1)}$$

Notice that $\max_{0\leq n}(1 - \gamma) + (2\gamma - 1)\Omega_1(n) - cn = \pi^{ns}$ which is the expected payoff of firm 1 in the naïve equilibrium. Thus, the first condition implies that firm 1 would not like to deviate. The second condition can be written as $(2\gamma - 1)\Omega_2(2n^*) - cn^* \leq 0$ and since the profit of firm 2 in equilibrium is $(1 - \gamma)$ and if it imitates firm 1 its profit is $(1 - \gamma) + (2\gamma - 1)\Omega_2(2n^*) - cn^*$, the second condition implies that firm 2 would not like to deviate.

Next, we reproduce the three lemmas and the main proposition of the basic model.

Lemma 10 ($\underline{n}$) *There exists an $\underline{n} > 0$ that satisfies the conditions: $\Omega_2(2\underline{n}) < \frac{c\underline{n}}{(2\gamma-1)}$ and $\Omega_2(2(\underline{n} - 1)) \geq \frac{c(\underline{n}-1)}{(2\gamma-1)}$. Furthermore, for any $n > \underline{n}$, $\Omega_2(2n) < \frac{cn}{(2\gamma-1)}$.*

Proof The function $\Omega_2(2n) - \frac{cn}{(2\gamma-1)}$ is (a) equal to 0.5 at $n = 0$, (b) approaches $-\infty$ as $n \to \infty$, and (c) is monotonically decreasing in n (since$1-\ \Omega_2(2n)$ is non-decreasing in n—from Lemma 8—and $\frac{cn}{(2\gamma-1)}$ is increasing in n). This establishes the existence of an $\underline{n}$ that satisfies the conditions in the first part of the Lemma. The second part of the lemma follows from the fact that $\Omega_2(2n) - \frac{cn}{(2\gamma-1)}$ is monotonically decreasing in n. □

Lemma 11 ($\overline{n}$) *There exists an $\overline{n}$ that satisfies the conditions: $\gamma - c\overline{n} \geq \pi^{nc}$ and $\gamma - c(\overline{n} + 1) < \pi^{nc}$. Furthermore, for any $n < \overline{n}$, $\gamma - cn > \pi^{nc}$.*

Proof

$$\gamma - cn - \pi^{nc} = (2\gamma - 1) - cn - (2\gamma - 1)\left[\max_{0\leq n}\Omega_1(n) - \frac{cn}{(2\gamma - 1)}\right]$$

is (a) positive at $n = 0$ (since $1 > \max_{0 \le n} \Omega_1(n)$), (b) approaches $-\infty$ as $n \to \infty$, and (c) monotonically decreasing in n. □

Lemma 12 *For $q > 0.5$, $\underline{n} \le \overline{n}$.*

Proof Since $\gamma - cn - \pi^{nc}$ is monotonically decreasing in n, we only need to show that $\gamma - c\underline{n} - \pi^{nc} \ge 0$.

Recall that $\pi^{ns} = (1-\gamma) + (2\gamma-1)\Omega_1(n^{nc}) - cn^{nc}$, and thus $\gamma - c\underline{n} - \pi^{ns} = (2\gamma - 1)\left[1 - \Omega_1(n^{nc})\right] + c\left[n^{nc} - \underline{n}\right]$.

If $n^{nc} \ge \underline{n}$, it is immediate that $(2\gamma - 1)\left[1 - \Omega_1(n^{nc})\right] + c\left[n^{nc} - \underline{n}\right] > 0$.

Next, consider the case where $n^{nc} < \underline{n}$. It is easy to show that since $(2\gamma-1)(1-\Omega_1(2\left(\underline{n}-1\right))) - c\left(\underline{n} - 1\right) = (2\gamma - 1)\Omega_2(2\left(\underline{n}-1\right)) - c\left(\underline{n} - 1\right) \ge 0$ we can get $(2\gamma-1)(1-\Omega_1(2\left(\underline{n}-1\right))) - c\left(\underline{n} - n^{nc}\right) \ge 0$ and thus $(2\gamma - 1)(1 - \Omega_1(2n^{nc})) - c\left(\underline{n} - n^{nc}\right) \ge 0$. □

Given these lemmas reproducing the PBE is immediate.

To examine the impact of γ on the number of ads in equilibrium we focus here (as explained in the text) on the case of $n = \underline{n}$. We find that as the media channels becoming less precise (i.e., γ decreases toward $\frac{1}{2}$) the number of ads decreases. The reason for this is quite interesting. In the separating equilibrium the expected revenue of firm s among the consumers who are exposed to media channel $k = s$ is γ. This is due, of course, to the possibility that the consumer who prefers its product will not be exposed to this media channel. As the media channels becoming less precise this revenue decreases and as a result the incentive of firm s' to imitate firm s is diminishing. Thus, the number of ads required to deter the "wrong" firm from imitating (and thus the number of ads needed to separate between the players) is smaller.

A.3 Endogenous prices

Here we discuss the differences between the solution of the model with and without endogenous prices. We start, though, by presenting the relevant change in the setting. The utility of consumer i from alternative s is:

$$u_{i,s} = \begin{cases} \alpha + \frac{1}{2}\beta_i x_s - p_s \text{ for } s \in \{1, 2\} \\ 0 \text{ for } s = 0 \end{cases} \tag{12}$$

As discussed in the text, since our research focuses on horizontally differentiated products in a symmetric setting, it is not reasonable to expect that prices will play a similar role. Thus, while we endogenize prices in the sense that firms are setting prices that maximize their profit, we do not include prices in the beliefs of the consumers. Accordingly, the beliefs are the same as in the text with the following change. The conditions that n^* is required to satisfy are:

$$\alpha + \tfrac{1}{2} - cn^* \geq \max_{p,0\leq n} p \Pr\left(\widetilde{g}_s \geq g^*(p,\alpha,q)\,|n\right) - cn \text{ and}$$

$$\alpha + \tfrac{1}{2} - cn^* \geq \max_{p} p\left\{1 + \Pr\left(\widetilde{g}_{s'} \geq g^*(p,\alpha,q)\,|2n^*\right)\right\} - 2cn^* \qquad (13)$$

where $g^*(p,\alpha,q) \equiv \left[\ln\left(\frac{0.5+p-\alpha}{0.5-p+\alpha}\right)\right]\left[\ln\left(\frac{q}{1-q}\right)\right]^{-1}$.

The rationale behind these beliefs is as follows. In equilibrium $p_s = \alpha + \frac{1}{2}$ and $n_s = n^*$ in media channel $k = s$ and 0 in media channel $k \neq s$. Thus, the profit of each of the firms is $\alpha + \frac{1}{2} - cn^*$. The first condition above requires that it is not optimal for s to deviate and send less than n^* in media channel $k = s$. The left hand side of the inequality is the profit of this firm in equilibrium while the right hand side is its profit if it decides to deviate. Notice that if it would deviate it would also charge a different price. It is easy to show that in such a case, the consumer will select firm s if $g_s \geq g^*(p,\alpha,q)$. Otherwise, this consumer will not buy any product because the expected utility from either firm would be negative. The second condition requires that firm s' would not find it optimal to imitate firm s and send n^* in media channel $k = s$. The right hand side represents its profit in the case of imitation. Notice that this firm will still send n^* in media channel $k = s'$ and thus it will get consumer $j = s'$ for sure. On the other hand, it will get consumer $j = s$ only if $g_{s'} \geq g^*(p,\alpha,q)$.[27]

Next we demonstrate the existence of a PBE. The following two definitions would become useful: (i) $A(x) \equiv \max_p p\left\{1 + \Pr\left(\widetilde{g}_{s'} \geq g^*(p,\alpha,q)\,|2x\right)\right\} - \left(\alpha + \frac{1}{2}\right)$ and (ii) $B(x) \equiv \left(\alpha + \frac{1}{2}\right) - \left[\max_{p,n<x} p \Pr\left(\widetilde{g}_s \geq g^*(p,\alpha,q)\,|n\right) - cn\right]$. Notice that $A(x)$ is a decreasing function in x. It is easy to show this by contradiction.[28]

Lemma 13 ($\underline{n}$) *There exists an* $\underline{n} > 0$ *that satisfies the conditions:* $c\underline{n} \geq A(\underline{n})$ *and* $A(\underline{n}-1) > c\left(\underline{n}-1\right)$. *Furthermore, for any* $n > \underline{n}$, $cn \geq A(n)$.

Proof (i) $A(n)$ is decreasing in n, (ii) $A(0) \geq c0$, and (iii) $A(\infty) < c\infty$. □

[27]Notice that if $g_{s'} = g^*$ the consumer is actually indifferent between s and s'. However, it is immediate that firm s' will never set a price so that $g_{s'} = g^*$ since by lowering its price by ε the consumer will prefer it.

[28]Let $p^*(x)$ be the optimal price given x, and accordingly, one can write $A(x)$ as $A(x, p^*(x))$. Let's assume that $A(x)$ is a decreasing in x. Thus, $p^*(x+1)\{1+\Pr(\widetilde{g}_{s'} \geq g^*(p^*(x+1),\alpha,q)|2(x+1))\} > p^*(x)\{1+\Pr(\widetilde{g}_{s'} \geq g^*(p^*(x),\alpha,q)\,|2(x))\}$. However, we know that $p^*(x+1)\{1+\Pr(\widetilde{g}_{s'} \geq g^*(p^*(x+1),\alpha,q)\,|2(x+1))\} < p^*(x+1)\{1+\Pr(\widetilde{g}_{s'} \geq g^*(p^*(x+1),\alpha,q)\,|2x)\}$ since $\Pr(\widetilde{g}_{s'} \geq g^*(p,\alpha,q)\,|2x)$ decreases in x (since this is the "wrong" firm). Thus, we get that $p^*(x+1)\{1+\Pr(\widetilde{g}_{s'} \geq g^*(p^*(x+1),\alpha,q)\,|2x)\} > p^*(x)\{1+\Pr(\widetilde{g}_{s'} \geq g^*(p^*(x),\alpha,q)\,|2x)\}$ which means that $p^*(x)$ was not optimal.

Lemma 14 ($\overline{n}$) *There exists an* $\overline{n}$ *that satisfies the conditions:* $B(\overline{n}) \geq c\overline{n}$ *and* $B(\overline{n}+1) < c(\overline{n}+1)$. *Furthermore, for any* $n < \overline{n}$, $B(n) \geq cn$.

Proof (i) $B(n)$ is non-increasing in n (since as n increases the maximization is less constraint) (ii) $B(0) > c0$, and (iii) $B(\infty) < c\infty$. □

Finally, we numerically identify the subset of the parameter space for which $B(\underline{n}) > 0$ and thus $\overline{n} > \underline{n}$ (i.e., the parameters' values for which a separating equilibrium exists).

We conduct the numerical analysis for $c = 0.05$, $\alpha \geq 0.6$ and $q \in (\frac{1}{2}, 1)$. We find that when $\alpha \in [0.6, 1.2]$ there is a separating equilibrium (i.e., $B(\underline{n}) > 0$ and thus $\overline{n} > \underline{n}$) for any value of q. However, when $\alpha > 1.2$ there are values of q for which $B(\underline{n}) < 0$. Specifically, when α is not too high $B(\underline{n}) < 0$ only for small values of q (i.e., qs which are close to 0.5), but as α increases the range of q for which $B(\underline{n}) < 0$ increases.

The reason that $B(\underline{n}) < 0$ for high values of α can be described as follows. When α is high, consumers' willingness to pay for the product is high. Furthermore, in such a case, the difference between the two products (due to their attributes) is relatively small. When prices are endogenous a firm can "compensate" the consumer who does not match well with its product (via lower prices), and still make a nice profit. As a result it is more difficult to deter such firm from imitating. Of course, this difficulty is more severe when q is small.

References

Abernethy, A., & Franke, G. (1996). The information content of advertising: A meta-analysis. *Journal of Advertising, 25*(Summer), 1–17.

Anand, B., & Shachar, R. (2005). *Advertising, the matchmaker.* http://www.tau.ac.il/˜rroonn/Papers/match05.pdf.

Anand, B., & Shachar, R. (2007). (Noisy) Communication. *Quantitative Marketing and Economics, 5*(3), 211–232.

Anderson, S., & Renault, R. (2006). Advertising content. *American Economic Review, 96*(1), 93–113.

Bagwell, K., & Ramey, G. (1994). Coordination economies, advertising, and search behavior in retail markets. *American Economic Review, 84*(3), 498–517.

Batra, R., Myers J. G., & Aaker, D. (1996). *Advertising management*. Prentice Hall.

Becker, G., & Murphy, K. (1993). A simple theory of advertising as a good or bad. *Quarterly Journal of Economics, 108*(4), 941–964.

Bernheim, D., & Redding, L. (2001). Optimal money burning: Theory and application to corporate dividends. *Journal of Economics & Management Strategy, 10*(4), 463–507.

Bhardwaj, P., Chen, Y., & Godes, D. (2008). Buyer-initiated vs. seller initiated information revelation. *Management Science, 54*(6), 1104–1114.

Butters, G. (1977). Equilibrium distribution of sales and advertising prices. *Review of Economic Studies, 44,* 465–91.

Chen, Y., Narasimhan, C., & Zhang, J. (2001). Individual marketing with imperfect targetability. *Marketing Science, 20*(1), 23–41.

Grossman, G., & Shapiro, C. (1984). Informative advertising with differentiated products. *Review of Economic Studies, L1*(1), 63–81.

Hertzendorf, M. N. (1993). I'm not a high-quality firm—but I play one on TV. *Rand Journal of Economics, 24*(2), 236–247.
Iyer, G., Soberman, D., & Villas-Boas, M. (2005). The targeting of advertising. *Marketing Science, 24*(3), 461–476.
Jacoby, J., & Hoyer, W. (1982). Viewer miscomprehension of televised communication: Selected findings. *Journal of Marketing, 46*(4), 12–26.
Jacoby, J., & Hoyer, W. (1989). The comprehension/miscomprehension of print communcation: Selected findings. *Journal of Consumer Research, 15*(4), 434–443.
Johnson, J., & Myatt, D. (2006). On the simple economics of advertising, marketing, and product design. *American Economic Review, 96*(3), 756–784.
Kihlstrom, R., & Riordan, M. (1984). Advertising as a signal. *Journal of Political Economy, 92*(3), 427–450.
Mayzlin, D., & Shin, J. (2008). *A model of image advertising*. Mimeo.
Meurer, M., & Stahl, D. (1994). Informative advertising and product match. *International Journal of Industrial Organization, 12*(1), 1–19.
Milgrom, P., & Roberts, J. (1986). Price and advertising signals of product quality. *Journal of Political Economy, 94*(4), 796–821.
Narasimhan, C. (1988). Competitive promotional strategies. *Journal of Business, 61*(4), 427–449.
Nelson, P. (1974). Advertising as information. *Journal of Political Economy, 82*(4), 729–754.
Resnik, A., & Stern, B. (1977). An analysis of information content in television advertising. *Journal of Marketing, 41*(1), 50–53.
Shachar, R., & Nalebuff, B. (1999). Follow the leader: Theory and evidence on political participation. *American Economic Review, 89*(3), 525–549.
Webb, P. H., & Ray, M. (1979). Effects of TV clutter. *Journal of Advertising Research, 9*(3), 7–12.