Informative Advertising can be Persuasive[1]

Bharat N. Anand[2] Ron Shachar[3]

Harvard Business School Tel Aviv University

February 23, 2000

[1] We are grateful to Ariel Rubinstein for his beneficial comments.

[2] Department of Competition and Strategy, Soldiers Field Road, Boston, MA 02163. Phone: (617) 495-5082; Fax: (617) 495-0355; email: banand@hbs.edu.

[3] Eitan Berglas School of Economics, Tel Aviv University, Tel Aviv, Israel, 69978. Phone: 972-3-640-9202; Fax: 972-3-640-9908; email: rroonn@post.tau.ac.il.

Abstract

One of the most cited facts about advertising is that exposure to ads increases a consumer's tendency to buy the promoted product. We propose a simple model of *informative advertising* that is consistent with this empirical regularity. In a differentiated-products environment, firms can target their advertising to different sets of consumers, who differ in their preferences over the products. Consumers are uncertain about products' attributes. We show that there exists a separating perfect Bayesian equilibrium where firms target only those consumers who best fit their products, and consumers rationally infer the product attributes of the firm airing these ads. Interestingly, the (noisy) information content in the ads is not used by consumers in equilibrium, but only affects their beliefs off the equilibrium path. There are several other results of interest. First, when ads are extremely noisy, that is ads are non-informative, the signaling equilibrium collapses. Second, the intensity of advertising in equilibrium is a decreasing function of the precision of ads.

JEL Classification: D11, D83, L00, M37.

Keywords*:* Advertising, Information, Product differentiation, Signaling.

"Hardly any busines practice causes economists greater uneasiness than advertising"

(Lester Telser, 1964)

1 Introduction

One of the most well established and well-known facts about advertising is that exposure to ads increases a consumers' tendency to buy the promoted product.[1] This empirical regularity has disturbed economists for a long time. The main concern is that advertising is a form of brainwashing. Advertising is accused of creating wants, distorting tastes and persuading consumers to buy products that they do not need. John Kenneth Galbraith (1971) has gone even further, charging that "the economy for its success requires organized public bamboozlement." Paul Milgrom and John Roberts (1986) and Gary Becker and Kevin Murphy (1993) present models with non-informative advertising which are consistent with the empirical regularity.[2] Recently, however, Avery Abernethy and George Franke (1996) have demonstrated that approximately 84 percent of all ads contain some information, and thus non-informative advertising is an extremely narrow category.[3] Here we propose a simple model of informative advertising that is both consistent with this view of advertising *and* explains the above-cited empirical regularity.

In an era marked by significant product differentiation, the ability of informative advertising to explain the empirical regularity is not obvious. Consider the case of a political campaign. When a candidate

[1] This empirical regularity is well established in the marketing literature (see, for example, Tulin Erdem and Michael Keane [1996]). Recently, a few studies in economics have also revealed the significant effect of exposure to advertising on consumers' choices (Mark Roberts and Larry Samuelson [1988] using data on cigarettes, Scott Stern and Manuel Trajtenberg [1998] using data on physicians' choices of drugs, Aviv Nevo [2000] using data on cereals, and Ron Shachar and Bharat Anand [1998] using data on television shows. It is worth noting that both Dan Ackerberg [1998], using data on yogurt, and Matthew Shum [1999], using data on cereals, find significant effects of advertising among inexperienced consumers only).

[2] Becker and Murphy write: "it is also 'obvious' that many ads provide essentially no information." Milgrom and Roberts write: "a nontrivial amount of advertising (especially on television) has little or no obvious informational content."

[3] Alan Resnik and Bruce Stern (1977) presented a method of measuring advertising information. This method involves content analysis to determine which types of information are presented 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 studies followed this method and Abernethy and Franke (1996) summarize their results in a meta-analysis. They found that 84 percent of all the ads (91,438) have had at least one cue (58 percent have had at least two ads and 33 percent have had at least 3). While newspapers appear to include more informative advertising (98 percent of them have had at least one cue), ads in television are also quite informative (71 percent). The product category that was the most informative was automobile (97 percent) and furniture/home furnishings/appliances (96 percent). Even if these numbers are not exactly right, it is clear that in most cases advertising include information.

informs the voters via ads that she opposes abortion, she is increasing the tendency of some people to vote for her, but decreasing the tendency of others to do so. Thus, a positive correlation between ad exposure and tendency to buy for *all* consumers does not seem consistent with informative advertising.

However, this example also illustrates one of the most important advertising strategies that might be able to explain the empirical regularity: namely, targeting of audiences. The political candidate should have little incentive to inform "pro choice" voters about her views on this matter. Modern media planning tools allow her to send these ads to mainly to "pro life" voters. In the same vein, commercials for beer, for example, appear mostly during sports events, and ads for cleaning products appear mainly during the morning shows on television. Thus, the observation that consumers who are exposed to ads are more likely to buy the promoted product than those who are not might represent an equilibrium relationship, not a causal one, since it might simply indicate that those who were exposed to the ads were more likely to buy the product *a priori*.

This implies that even when using individual level data, consumers' exposure to ads is an endogenous variable from an econometric perspective. In order to correct this bias in estimating the effectiveness of advertising one needs to account for individuals' (observed and unobserved) tastes. In Ron Shachar and Bharat Anand (1998) we find, using individual level data on ads for television shows, that even after correcting for the econometric bias, advertising effectiveness is statistically significant and strong. Indeed, without correcting the endogenity problem, the effectiveness of ads is upward biased. The model presented in Section 2 demonstrates, however, that firms' targeting strategies have another aspect that can explain the powerful effectiveness of advertising that remains after correcting the endogeneity problem.

The logic of the model is that rational consumers understand firms' incentives to target their ads. Thus, consumers who are assumed to face uncertainty about products' attributes, use the intensity of the ads that they were exposed to as a signal on products' attributes. In equilibrium, firms send ads only to the consumers that best fit their product, and consumers base their perception of products attributes only on the intensity of ads. Thus, their uncertainty is completely resolved in equilibrium. We show that there exists such a separating perfect Baysian equilibrium. There are several other results of interest. First, there does not exist any (non-degenerate) pooling equilibrium. Second, the noisy information in the content of the ads is not used by consumers in equilibrium; rather, it only affect their beliefs off the equilibrium path. Third, in equilibrium, ad intensity is a decreasing function of ads precision. Fourth, when ads are extremely noisy (that is, ads are non-informative), the signaling equilibrium collapses. This contrasts with Milgrom and Roberts (1986).

Milgrom and Roberts (1986) demonstrated that in equilibrium ad intensity signals products quality, thus explaining the link between exposure to ads and tendency to buy. Becker and Murphy (1993) viewed advertising as a good that complements the product. Thus, again, advertising intensity increases the tendency to buy. While both theories have a grain of truth, they also bear seeds of discouragement. Milgrom and Roberts view advertising as an activity akin to burning money that can serve as a signal. This has disturbed many commentators for the simple reason that there would appear to be other, more productive, avenues to burn money. For example, high quality firms might contribute large amounts of money for good causes and announce it; this would not only signal the quality of their products, but may gain them "good will" points as well. Furthermore, advertising a la Milgrom and Roberts need not convey any information. However while some ads appear to be consistent with this view, the vast majority of ads do not. Thus, we are left with a theory that might explain only a small fraction of the phenomenon.

While Becker and Murphy (1993) suggest a compelling view of advertising, their theory can be viewed as *de facto* assuming that ad intensity stimulates the tendency to buy the product. Thus, those who are puzzled by the persuasive effect of advertising might view this explanation as a tautology.

2 The model

There are two firms indexed by j, two individuals indexed by i, and two media channels indexed by k. Each firm offers one product. The products are characterized by their prices, p_j, and attributes, X_j. The distribution of product attributes is given by:

$$
\begin{aligned}
X_1 &= 1 \\
X_2 &= -1
\end{aligned}
$$

The individual needs only one product. Thus, she can either buy one product or not buy any. Individual i's utility from consuming the product of firm j is given by:

$$
U_{i,j} = \alpha + \frac{1}{2}\beta_i X_j - p_j \tag{1}
$$

where $\beta_1 = 1$ and $\beta_2 = -1$. Her utility when she does not buy any product is zero, $U_{i,0} = 0$. The individual's decision variable, $d_{i,j}$, is equal to 1 if she chooses alternative j, zero otherwise.

The price of each product is observable and common knowledge to all parties. To begin with, we focus on firms' advertising decisions while taking their choice of prices as given. To simplify the model, we further assume that $p_j = \alpha$ for any j. Later, we relax these assumptions without any effect on our qualitative results.

2.0.1 Information sets

Each individual knows that there exist two products and knows the distribution of product attributes: $X_j \in \{-1, 1\}$. However, they do not know the true attributes of each product being offered. Each individual's prior probability that a product is of a given type is is $\frac{1}{2}$, that is $\Pr(X_j = 1 | IS_i = \emptyset) = \frac{1}{2}$ for any j, where IS_i is the information set of individual i. The rationale behind this setting is as follows: Because of increasing product diversity and the constant introduction of new alternatives and changes in product attributes, it is difficult for individuals to stay informed about the attributes of all the alternatives. Notice, however, that once the individual knows what firm j is producing, she also knows the attribute of the product of the other firm.

Firms, similarly, may not know the preferences of each individual, but they do know the distribution of preferences. In practise, firms often spend large amounts of money analyzing the markets in which they operate, and the tastes of consumers. Consequently, they have a good assessment of these preferences based on individuals' demographics, life styles and purchase histories. Thus, we assume that the firms know the preferences of each individual.

2.0.2 Information in advertising

Individuals acquire information about the attributes of products (hence, the identity of firms) through advertising. As mentioned earlier, a significant majority of ads include *some* information on the attributes of the products. However, it is also fair to say that the information in ads is noisy. In 1979, the Educational Foundation of the American Association of Advertising Agencies surveyed nearly 2,700 consumers about the content of 60 thirty-second televised communications—including ads, public service communications, and editorial content. They found that about 29 percent of the communications (ads or other content) were miscomprehended by consumers, as measured by a particular series of true-false questions (Jacob Jacoby and Wayne 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 editorial pages. Roughly 20 percent of the material was miscomprehended, with another 15 percent being simply "not understood" (Jacob Jacoby and Wayne Hoyer [1989]). Consequently, it is natural to formulate advertising as a noisy signal:

$$S_{i,j,n} = \begin{cases} X_j \text{ with probability } q \\ -X_j \text{ with probability } 1-q \end{cases} \tag{2}$$

where n indexes the number of the signal and $q \geq \frac{1}{2}$.[4] Notice that in this setting an ad from one firm provides information on the attributes of the other firm's products as well. For example, an individual that

[4] Notice that any $q \neq \frac{1}{2}$ is informative. We prefer to consider the natural case of $q > \frac{1}{2}$ in this paper: in other words, "correct signals are more likely than incorrect signals".

received one signal which was $S_1 = 1$ would, using Bayes' rule, form the following expectations about the product attributes: $E(X_1|S_1 = 1) = (2q-1)$ *and* $E(X_2|S_1 = 1) = (1-2q)$. When $q > \frac{1}{2}$, for individual 1, $E(\beta_1 X_1|S_1 = 1) > 0$ and $E(\beta_1 X_2|S_1 = 1) < 0$.

2.0.3 Media channels

The first individual is exposed only to media channel 1, while the second is only exposed to media channel 2.

In reality there are many classes of media—including 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 men (like Monday Night Football) or shows that appeal especially to women, such as romantic dramas with a feminine cast. Thus, media channels, within-channel categories, and within-category classes, differ in their audience composition. 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, and the audiences have become more fragmented, each of these channels has ended up serving more specific interests. Thus, although our assumption that each individual is exposed to one media channel is a strict one, it captures the basic nature of such channels. We also assume that firms are aware of the media consumption of individuals. In reality, firms have multiple data sources that characterize the audiences of the different media channels, and these characterizations are rather detailed. For example, they include not only the demographics of the audiences, but their lifestyles and consumption habits as well.[5]

2.0.4 Firms' objective function

The cost of an ad in each media channel, c_k, is assumed to be the same, and equal to $c > 0$. Furthermore, we also assume that c is not too big (otherwise, sending ads might never be profitable). Each firm manufactures

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

its product at constant unit cost b_j ($b_j \geq 0$). Thus, the profit of the firm j is:

$$\pi_j(N_{j,1}, N_{j,2}) = (p_j - b_j) \cdot \sum_{i=\{1,2\}} d_{i,j} - c \cdot (N_{j,1} + N_{j,2}) \tag{3}$$

where $N_{j,k}$ is the number of ads that firm j places in media channel k. For simplicity we assume that $b_j = 0$ for any j.

Next, we solve for the optimal ad intensity of firms under the assumption that individuals do not act strategically, by using the ad intensity as a *signal* about the firms' identity. That is, individuals' decisions depend only on the *content* of the ads. Later, we allow individuals to use the information revealed by the intensity of advertising and characterize the rational expectations equilibrium.

2.1 Non Strategic Equilibrium

We start by demonstrating that in the non-strategic equilibrium (NSE) each firm j does not send ads through the media channel that is watched by individual $i \neq j$. Then, we solve for the optimal number of ads sent by each firm j through the media channel that is consumed by individual $i = j$.

2.1.1 The Individual's decision

As mentioned earlier, individual i only watches media channel $k = i$. After being exposed to the ads aired in that channel by both firms, she decides whether to buy one of the products or not. Under the assumptions made above, she gets zero utility when she does not buy a product. Her expected utility from buying product j, for any j, is:

$$E(U_{i,j}|IS_i) = \tfrac{1}{2}\beta_i \left(\frac{1 - (\frac{1-q}{q})^{\bar{s}_{i,j}}}{1 + (\frac{1-q}{q})^{\bar{s}_{i,j}}} \right) \tag{4}$$

where $\bar{s}_{i,j} \equiv \sum_{n=1}^{N_{j,i}} S_{i,j,n} - \sum_{n=1}^{N_{(3-j,i)}} S_{i,j',n}$ and $j' \equiv (3-j)$. Notice that the source of the ad does not matter.

As special cases, if $q = \frac{1}{2}$ or if $\bar{s}_{i,j} = 0$, then $E(U_{i,1}|IS_i) = E(U_{i,2}|IS_i) = 0$. In other words, both alternatives (buy one product or none) yield the same value of the expected utility, zero, if either there is no information in the ads a priori, or if the realizations of the signals cancel each other out. In such cases, we assume that the individual chooses each product with equal probability $\frac{1}{2}$. Similarly, it is easy to show that, if $q > \frac{1}{2}$, then, for any i:

$$E(U_{i,j=i}|IS_i) > E(U_{i,j'\neq i}|IS_i) \Longleftrightarrow \beta_i \bar{s}_{i,j} > 0 \tag{5}$$

In other words, when $q > \frac{1}{2}$ and $\bar{s}_{i,j} \neq 0$, the expected utility from one product is positive, hence the other is negative. In such a case, the individual's decision rule is given by:

$$d_{i,j} = 1 \text{ if } \beta_i \bar{s}_{i,j} > 0 \tag{6}$$

2.1.2 The Firms' strategies

>From the firms' point of view $\bar{s}_{i,j}$ is a random variable. Let $E[\pi_j(N_{j,1}, N_{j,2}); N_{j',1}, N_{j',2})]$ denote the expected profit function for firm j if it airs $N_{j,1}$ ads in channel 1 and $N_{j,2}$ ads in channel 2, and given that the other firm airs $N_{j',1}, N_{j',2}$ ads in the two channels. Thus, her expected profit function is:

$$E\left[\pi_j(N_{j,1}, N_{j,2}; N_{j',1}, N_{j',2})\right] = (p_j - b_j) \cdot \sum_{i=\{1,2\}} \widetilde{R}_{j,i}^{NS}(N_i) - c \cdot [N_{j,1} + N_{j,2}] \tag{7}$$

where the expected revenues to firm j from individual i are given by:

$$\widetilde{R}^{NS}_{j,i}(N_i) \quad = \quad \Pr(\beta_i \bar{s}_{i,j} > 0|N_i) + \tfrac{1}{2} \cdot \Pr(\beta_i \bar{s}_{i,j} = 0|N_i) \tag{8}$$

and N_i, the number of ads that individual i is exposed to, is given by: $N_i \equiv N_{1,i} + N_{2,i}$. Here we already take into account that individual i consumes only the media channel $k = i$. Notice that when firm 2 is sending an ad through media channel 1, and the ad is correct $(S_j = X_j)$ it is leading the first individual to conclude that the other firm is more appropriate for her.

The probability functions in (8) can be rewritten as:

$$\begin{aligned}
\Pr(\beta_i \bar{s}_{i,j} > 0|N_i) &= \sum_{m > \frac{N}{2}} \begin{pmatrix} N_i \\ m \end{pmatrix} (q_{i,j})^m (1 - q_{i,j})^{(N_i - m)} \\
\Pr(\bar{s}_{i,j} = 0|N_i) &= \begin{cases} \begin{pmatrix} N_i \\ \frac{N_i}{2} \end{pmatrix} [q(1-q)]^{\frac{N_i}{2}} & \text{when } N_i \text{ is even} \\ 0 & \text{when } N_i \text{ is odd} \end{cases}
\end{aligned} \tag{9}$$

where $q_{i,j} = q$ if $i = j$ and $q_{i,j} = 1 - q$ if $i \neq j$.

For firm j, the probability that the expected revenues from individual $i = j$ are positive is the sum of the probabilities of all events where the realizations of the ad signals are, in the aggregate, "correct". If, however, N_i is even and the the number of correct signals equal the number of incorrect ones, thus canceling each other out, then the individual obtains no additional information from the ad signals and chooses each product with probability $\frac{1}{2}$. If N_i is odd, however, this event can never occur.

Next, we present two Lemmas that are used in the subsequent proposition.

Lemma 1 *$\widetilde{R}^{NS}_{j,i}(N_i)$ is a non decreasing function in N_i for $i = j$.*

Proof. See Appendix A. ■

Specifically, the change in $\widetilde{R}_{j,i}^{NS}(N_i)$ when N_i increases by 1 is:

$$\begin{array}{ll} (q_{i,j} - \frac{1}{2}) \Pr(\bar{s}_{i,j} = 0 | N_i) & \text{when } N_i \text{ is even} \\ 0 & \text{when } N_i \text{ is odd} \end{array} \tag{10}$$

The intuition behind the result is that, since $q > 0.5$ (i.e., ads contain *some* information), an increase in the total number of ads that are aired in a channel will (weakly) increase the expected revenues for the firm that is "right" for the consumers watching that channel. However, when N_i is odd, the change in revenues with one more ad is zero.[6] Thus, as the number of ads in channel k increases, the probability that individual $i = k$ chooses firm $j = i$ does not decrease.

Lemma 2 $\Pr(\bar{s}_{i,j} = 0|N_i)$ *is a non increasing function in* N_i*.*

Proof. See Appendix B. ■

This result simply states that as the total number of ads aired in a channel increases, the probability of a tie (i.e., the number of correct and incorrect signals canceling each other out) is likely to fall.

Now, denote as $N_{j,k}^{ns}$ the number of ads aired by firm j in channel k, for any j, k, in any (non-strategic) equilibrium. Also, recall that c is not too big. We now state the main result of this section.

Proposition 3 *In any non-strategic equilibrium,* $N_{1,2}^{ns} = N_{2,1}^{ns} = 0$*. Moreover,* $N_{1,1}^{ns} = N_{2,2}^{ns} = N_{ns}^* > 0$*, where* N_{ns}^* *is finite.*

Proof. Since $\widetilde{R}_{j,i}^{NS}(N_i)$ is non-decreasing in N_i for any $i = j$, and given any number of ads being aired by the rival firm in each channel, each ad sent by firm j through media channel $k \neq j$ weakly decreases

[6] To see this intuitively, note that a change in revenue can occur only when the number of correct signals thus far exceeds or is less than the number of incorrect signals by. Let γ_1 denote the probability of any event where the sum of "correct" signals exceeds the number of incorrect ones by 1; in this case, the addition of one ad will, with probability $(1-q)$, result in the total number of correct and incorrect signals canceling each other out. Conversely, let γ_2 denote the probability of any event where the sum of "incorrect" signals exceeds the number of correct signals by 1; in this case, the addition of one ad will, with probability q, result in the total number of correct and incorrect signals canceling each other out. It is easy to see that $\gamma_1(1-q) = \gamma_2 q$. Therefore, the change in revenues when N_i is odd and one additional ad is aired, is given by: $\frac{1}{2}\gamma_2 q - \frac{1}{2}\gamma_1(1-q)$, which, after substituting for γ_1 in terms of γ_2, equals 0.

its market share. Since ads are costly ($c > 0$) sending *any* ad through channel $k \neq j$ decreases firm j's expected profits. Therefore, $N_{1,2}^{ns} = N_{2,1}^{ns} = 0$ is a dominant strategy.

It is quite clear that $N_{j,j}^{ns}$ must be an odd number. If $N_{j,j}^{ns}$ were even, the firm can decrease the number of ads by 1 lowering the cost of advertising without reducing its expected revenue.

Furthermore, $N_{j,j}^{ns}$ should satisfy the following conditions:

$$\begin{aligned} E\left[\pi_j(N_{j,j}^{ns}, 0)\right] - E\left[\pi_j(N_{j,j}^{ns} - 2, 0)\right] &\geq 0 \\ E\left[\pi_j(N_{j,j}^{ns} + 2, 0)\right] - E\left[\pi_j(N_{j,j}^{ns}, 0)\right] &< 0 \end{aligned} \tag{11}$$

In other words, increasing or decreasing the number of ads by 2 from the equilibrium number should result in negative marginal profits for the firm. (Notice that we add (and substract) 2 from $N_{j,j}^{ns}$ since we are only considering odd numbers). Using (10), this condition (11) can be re-written as:

$$\begin{aligned} \Pr(\bar{s}_{i,j} = 0 | N_{j,j}^{ns} - 1) &\geq \frac{2c}{(q - \frac{1}{2})} \\ \Pr(\bar{s}_{i,j} = 0 | N_{j,j}^{ns} + 1) &< \frac{2c}{(q - \frac{1}{2})} \end{aligned} \tag{12}$$

Since $\Pr(\bar{s}_{i,j} = 0 | N_i)$ is a weakly decreasing function in N_i and c is not too large, there exists an $N_{j,j}^{ns} > 0$ that satisfies the condition (12). ■

This result implies that there is appropriate targeting in equilibrium. Firms will not place ads in the channel whose consumers do not fit its product, since doing so (weakly) reduces share and increases cost. This suggests that it may be natural to consider the case where rational individuals strategically infer the product attributes of each firm simply from the location of ads. We turn to characterizing such strategic equilibria in the next subsection.

$N_{j,j}^{ns}(c, q)$ is obviously a decreasing function of c. The effect of q on $N_{j,j}^{ns}(c, q)$ is not monotonic, however. Notice that for both $q = \frac{1}{2}$ and $q = 1$ the function $(q - \frac{1}{2}) \cdot \Pr(\bar{s}_{i,j} = 0 | N_i) = 0$ (when $q = \frac{1}{2}$, so ads have no information content, a firm is as likely to lose share from airing an ad as to gain share; when

$q = 1$, since all the signals are correct, the probability of a tie is negligible, so the marginal benefit of airing ads is small). When q increases from $\frac{1}{2}$, the probability of a "tie" ($\Pr(\bar{s}_{i,j} = 0|N_i)$) diminishes, but the effectiveness of ads in informing the individuals $(2q-1)$ increases. The first effect reduces the incentive to send ads, while the second increases it. It can be shown that first effect is more significant for q close to $\frac{1}{2}$, while the second dominates for q close to 1.

Finally, notice that for $q < 1$ and any finite $N^{ns}_{j,j}$, the probability that an individual chooses the product that fits her best, is smaller than one ($\Pr(d_{i,i} = 1) < 1$). Thus, with some probability individuals make mistakes in equilibrium—this follows from the noisiness of ad signals.

2.2 Strategic Advertising

Here, we examine the case when individuals incorporate both the statistical information that is revealed in the content of ads, and the signaling information that is revealed by the intensity of advertising, when forming their expectations about products' attributes. We show that in equilibrium, each firm still sends ads only through one media channel. In addition, however, individuals have no uncertainty about products' attributes in equilibrium.

We start by specifying the beliefs of individuals and the strategies of firms, and then show that these beliefs and strategies are part of a perfect Bayesian equilibrium. First, we present a few definitions. Denote the posterior probability function of the firms' types as $\mu(X_j|N_{j,i}, N_{j',i}, \overline{s}_{i,j})$. Also, denote the indicator function as $I\{\cdot\}$. It equals one if the expression in brackets is true, zero otherwise. Finally, denote the probability of $\overline{s}_{i,j}$ (from the point of view of the firm) as $\rho(s_j|N_i)$.

To begin with, we specify the beliefs—denoted by B—of individuals:

$$\mu(X_j = \beta_i|\{N_{j,k}\}, \{\overline{s}_{i,j}\}) = \left\{ \begin{array}{c} 1 \text{ if } N_{j,1} = N^* \text{ and } N_{j',1} \neq N^* \\ 0 \text{ if } N_{j,1} \neq N^* \text{ and } N_{j',1} = N^* \\ Pr(X_j = \beta_i|\overline{s}_{i,j}) = \frac{1}{1+\left(\frac{1-q}{q}\right)^{\beta_i \overline{s}_{i,j}}} \text{ otherwise} \end{array} \right\} \quad (13)$$

where N^* satisfies the conditions:

$$1 - c \cdot N^* \quad \geq \quad \max_{0<N_j<N^*} \widetilde{R}_{j,i=j}(N_j) - c \cdot N_j \tag{IC}$$

$$\widetilde{R}_{j',i=j}(2N^*) - c \cdot N^* \quad < \quad 0 \tag{DC}$$

$\widetilde{R}_{j',i=j}(2N^*) - c \cdot N^*$ denotes the profits that firm $j \neq i$ gets if it mimics the other firm and places N^* ads in channel i; define $\pi^d_{j',j}(N) \equiv \widetilde{R}_{j',i=j}(2N) - c \cdot N$ for notational convenience. Similarly, for firm $j = i$, $1 - cN^*$ denotes its profits if each firm follows their equilibrium strategies; $\max_{0<N_j<N^*} \widetilde{R}_{j,i=j}(N_j) - c \cdot N_j$ denotes its profits if it deviates from its strategy and places $N_j < N^*$ ads in channel $k = i$. Again, for convenience, define $\pi^d_{j,j}(N) \equiv \max_{0<N_j<N} \widetilde{R}_{j,i=j}(N_j) - c \cdot N_j$.

Later, we show that beliefs B satisfy the refinement conditions of Kreps and Wilson (1982). The logic behind these beliefs can be stated as follows: if the actions are consistent with the equilibrium strategies $\{N^*, 0\}$, then the identity of the firms is revealed perfectly. The same holds true if one of the firms follows the equilibrium strategy of the "correct" firm (i.e., $N = N^*$), whereas the other does not. If one of the firms follows the equilibrium path (specifically, N^*) while the other does not, the identity of the firms is revealed. When neither firm follows equilibrium strategies, then individuals base their expectations of firm type only on the statistical information revealed via the ads, and not on the intensity of advertising.

The firms' expected profit function in this case is

$$\begin{aligned} & E\left[\pi^S_j(N_{j,1}, N_{j,2})\right] \\ = & \sum_{i=\{1,2\}} \sum_{s_j=-N_i}^{N_i} I\{\mu(X_j | N_{j,i}, N_{(3-j),i}, s_j) > \frac{1}{2}\} \rho(s_j | N_i) - c \cdot [N_{j,1} + N_{j,2}] \end{aligned} \tag{14}$$

Next we present three Lemmas that are used in the main proposition. In Lemma 3, we show that firm $j \neq i$ will not find it profitable to mimic the other firm if N^* exceeds some number N^L. Then, in Lemma 4, we show that there exists some number N^H such that for all $N^* \leq N^H$, firm $j = i$ will not find it profitable to deviate from its equilibrium strategy and place fewer than N^* ads in channel i. Finally, Lemma 5 asserts that $N^H > N^L$ for $q > \frac{1}{2}$, therefore for all $N^L \leq N \leq N^H$, $N^* = N$ can be supported in a perfect Bayesian equilibrium.

Lemma 4 (N^L) *There exists an $N^L \geq 2$ that satisfies the conditions: $\pi^d_{j',j}(N^L) < 0$ and $\pi^d_{j',j}(N^L - 1) \geq 0$. Furthermore, for any $N > N^L$, $\pi^d_{j',j}(N) < 0$.*

Proof. See Appendix C. ■

Lemma 5 (N^H) *There exists an N^H that satisfies the conditions: $1 - c \cdot N^H \geq \pi^d_{j,j}(N^H)$ and $1 - c \cdot (N^H + 1) < \pi^d_{j,j}(N^H + 1)$. Furthermore, for any $N < N^H$, $1 - c \cdot N > \pi^d_{j,j}(N)$.*

Proof. See Appendix D. ■

Lemma 6 *For $q > \frac{1}{2}$, $N^L \leq N^H$.*

Proof. We first show that $\pi^d_{j',j}(N^L - 1) \leq 1 - c \cdot N^L - \pi^d_{j,j}(N^L)$ and then use this relationship to show that $N^H < N^L$ leads to a contradiction.

Notice that $\widetilde{R}_{j',i=j}(z) = 1 - \widetilde{R}_{j,i=j}(z)$. Thus, $\pi^d_{j',j}(N^L - 1) = 1 - c \cdot N^L - \{\widetilde{R}_{j,i=j}\left(2 \cdot (N^L - 1)\right) - c \cdot 1\}$. Next we show that $\{\widetilde{R}_{j,i=j}(2 \cdot (N^L - 1)) - c \cdot 1\} \geq \pi^d_{j,j}(N^L)$. Since $2 \cdot (N^L - 1) \geq N^L$ for any $N^L \geq 2$, and $\widetilde{R}_{j,i=j}(z)$ is a non decreasing function in z, it follows that $\widetilde{R}_{j,i=j}(2 \cdot (N^L - 1)) \geq \widetilde{R}_{j,i=j}(z)$ for any $z < N^L$. Therefore, $\pi^d_{j,j}(N^L) = \max_{0 < N_j < N^L} \{\widetilde{R}_{j,i=j}(N_j) - c \cdot N_j\} \leq \{\widetilde{R}_{j,i=j}(2 \cdot (N^L - 1)) - c \cdot 1\}$. (Recall that from Proposition 3 we know that $N^*_{ns} \geq 1$ and thus the cost in $\pi^d_{j,j}(N^L)$ is at least $c \cdot 1$). Thus,

$$\pi^d_{j',j}(N^L - 1) \leq 1 - c \cdot N^L - \pi^d_{j,j}(N^L) \tag{15}$$

Now, if $N^H < N^L$ then by the definition of N^H, we have: $1 - c \cdot N^L - \pi^d_{j,j}(N^L) < 0$. However, from the definition of N^L, we know that $\pi^d_{j',j}(N^L - 1) > 0$. This contradicts the inequality in (15). ■

The intuition behind Lemma 6 is illustrated in Figure 1. In order to simplify the intuition the non-continous function $\widetilde{R}_{j',i=j}(z)$ is presented as a continous function there. The straight line represents the cost of advertising to each firm as a function of N. The lower concave line represents the market share of the "wrong" firm when both firms send the same number of ads, and the upper concave lines indicates the market share of the "wrong" firm when only one firm is sending ads. Notice that the two firms are competing on the market share of the "wrong" firm. Off the equilibrium, the "wrong" firm gets a positive market share, while in equilibrium it gets zero. Thus, the two concave lines represent the incentives of the two firms. The lower line represents the incentive of the "wrong" firm to deviate from the equilibrium strategy and the upper line stands for the incentive of the "right" firm *not* to deviate from its equilibrium strategy. Since the incentive of the "right" firm is always higher than the incentive of the "wrong" firm, $N^H \geq N^L$.

We are now ready to state the main result:

Proposition 7 (PBE) *For $q > \frac{1}{2}$, there exists a perfect Bayesian equilibrium where beliefs are given by B and the firms' pure strategies are given by: $N_{j,k=j} = N^* \in \{N^L, N^H\}, N_{j',k=j} = 0$, for all j.*

Proof. Given beliefs and $N_{j,j} = N^* \in \{N^L, N^H\}$ (for any j), firm j' will choose not to deviate from $N_{j',j} = 0$ since: (a) any $0 < N_{j',j} < N^*$ involves the cost $c \cdot N_{j',j}$ without any revenues; (b) $N_{j',j} = N^*$ leads to losses from Lemma 4; and (c) $N_{j',j} > N^*$ leads to larger losses, since $\widetilde{R}_{j',i}(N_i)$ is decreasing in N_i.

Given beliefs and $N_{j',j} = 0$ (for any j'), firm j will optimally choose $N_{j,j} = N^* \in \{N^L, N^H\}$ since: (a) from Lemma 5, any $N < N^*$ results in lower profits; (b) any $N > N^*$ involves additional costs without any increase in revenues.

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

Thus, $N^* \in \{N^L, ...N^H\}$. Each firm sends ads only through the media channel that is watched by those consumers who best "fit" its product. All uncertainty about product attributes is resolved through ad location, and each individual chooses the product that is "right" for her with certainty.

The equilibrium number of ads in each channel, N^*, depends on q and c. We examine here the effect of these parameters on N^*, for $N^* = N^L$.[7] An increase in c decreases the incentive of firm j to send ads in media channel k, for any $j, j \neq k$, since doing so would increase its costs of such deception without affecting the benefits (which depend only on q); thus, $N^* = N^L$ decreases. An increase in q also decreases the incentives to mimic the rival and deceive consumers, since it lowers the benefits (equal to $\widetilde{R}_{j',j}(2N)$ for any N) without affecting the costs; thus, $N^* = N^L$ decreases. As in the non-strategic equilibrium, then, the intensity of ads decreases in q and c.

An interesting feature of the separating equilibria is that, in equilibrium, individuals ignore the content of the ads. Moreover, doing so leads individuals to make the best choices. The following example is illustrative. Consider a case where the content of the ads aired by firm 1 indicates that the second firm's product is the best match for individuals of type 1, i.e., $\overline{s}_{1,1} < 0$. (Although this is not too likely to occur, it is possible as long as $q < 1$). However, since individual 1 is only exposed to ads by firm 1, she ends up, in equilibrium, choosing to buy the product of firm 1, thus making the right choice. Thus, in the perfect Bayesian equilibrium, individuals always choose the product that gives them the highest utility. In contrast, recall that when the individual is not behaving strategically, she occasionally chooses a product that is not the best fit for her. As q increases, however, it the probability of mistakes falls.[8]

Next, compare the number of ads aired by firms in the strategic and non-strategic cases. Although both N^* and N^{ns} depend on the informativeness of ads, the functional dependence on q is different in the two cases. In the non-strategic case, the benefit of airing additional ads is that ads break ties. Conditional on a tie, the marginal benefit increases in q. However, the probability of a tie decreases in q, so the unconditional marginal benefit of airing more ads is ambiguous. In the strategic case, however, the effect of q is only to reduce the benefits to mimicking by the "wrong" firm, since q only affects beliefs off the equilibrium path. Therefore, an increase in q unambiguously decreases the gains to deception, and reduce the need for the

[7] 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 market share of firms is $\frac{1}{2}$, any $N^* > N^L$ results in firms incurring higher costs without affecting firms' revenues or consumer welfare.

[8] The probability of mistakes depends both on the noise in ads and the number of ads aired. As q increases, ads are less noisy, hence the likelihood 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). The first effect, however, dominates the second as q increases (it is easy to see this diagrammatically).

"correct" firm to air more ads. Using simulations, we find that if q is small (close to $\frac{1}{2}$) or large (close to 1), $N^* > N^{ns}$; for q not too small and not too large, however, $N^* \leq N^{ns}$. This result is intuitive, since, in the non-strategic case, recall that the incentive to send ads is low when q is small or large.

Furthermore, while in the NSE the individual choose the "wrong" firm with a positive probability. In the PBE the probability of choosing the "wrong" firm is zero.

The previous Proposition studied the PBE when $q > \frac{1}{2}$, the following Lemma and Proposition examines the case of $q = \frac{1}{2}$. The following Lemma is quite clear from Figure 2.

Proposition 8 ***Lemma 9*** *For $q = \frac{1}{2}$, $N^H < N^L$.*

Proof. When $q = \frac{1}{2}$, $\widetilde{R}_{j',j}(N) = \widetilde{R}_{j,j}(N) = \frac{1}{2}$ for any N. In particular, $\widetilde{R}_{j',j}(2N) = \frac{1}{2}$, for any N and $\max\limits_{0<N'<N} \widetilde{R}_{j,j}(N') - c \cdot N' = \frac{1}{2}$ for any N. Consequently, from the definition of N^H, it follows that $c \cdot N^H \leq \frac{1}{2}$ and $c \cdot (N^H + 1) > \frac{1}{2}$. Similarly, from the definition of N^L, it follows that $c \cdot (N^L - 1) \leq \frac{1}{2}$ and $c \cdot (N^L) > \frac{1}{2}$. Rewriting these expressions gives: $\frac{1}{2c} < N^L \leq (\frac{1}{2c} + 1)$ and $\frac{1}{2c} \geq N^H > (\frac{1}{2c} - 1)$. ■

Proposition 10 *When $q = \frac{1}{2}$ a perfect Bayesian equilibrium where beliefs are given by B does not exist.*

Proof. Since $N^H < N^L$ there is not any N^* that can satisfy the conditions of beliefs B. ■

This means that when $q = \frac{1}{2} + \varepsilon$ the separting equilibrium exists, but when $q = \frac{1}{2}$ it does not hold any more. In other words, as long as advertising is informative, the empirical regularity can be explained with the signaling game. Furthermore, the information in the ads can be very noisy. However, when ads are not informative, the signaling model cannot explain the empirical regularity.

So far we have shown that under beliefs B there exists a set of equilibria that can explain the empirical regularity as long as advertising is informative ($q > \frac{1}{2}$). Next, we show that a pooling equilibium does not exist when beliefs are B.

Proposition 11 *When $q > \frac{1}{2}$ a pooling equilibrium where the beliefs are given by B does not exist.*

Proof. When $N_{j,k=j} = N_{j',k=j} < N^L$, firm j will deviate. Since $\widetilde{R}_{j',i=j}(2 \cdot (N^L - 1)) > c \cdot (N^L - 1)$, we know that $1 - \widetilde{R}_{j,i=j}(2 \cdot (N^L - 1)) > c \cdot (N^L - 1)$ and thus $1 - c \cdot N^L > \widetilde{R}_{j,i=j}(2 \cdot (N^L - 1)) - c \cdot 1 >$

$\widetilde{R}_{j,i=j}(N_{j,k=j}+N_{j',k=j})-c\cdot N_{j,k=j}$. This means that firm j would benefit by increasing its ad intensity to N^L.

When $N_{j,k=j}=N_{j',k=j}\geq N^L$, firm j' will deviate, because its profits are zero.[9] ■

It turns out that the main Proposition, 7, can be stated in a more general way. Next we show that B is the only set of beliefs that leads to a *separating* equilibrium within a larger set of beliefs, B'.

$$\mu(X_j=\beta_i|N_{j,k=j},N_{j',k=j},\overline{s}_{i,j})=\begin{cases}1 & \text{if} \quad (N_{j,k=j},N_{j',k=j})\in n'_{12}\\ \Pr(X_j=\beta_i|\overline{s}_{i,j}) & \text{if} \quad (N_{j,k=j},N_{j',k=j})\in n''_{12}\\ 0 & \text{if} \quad (N_{j,k=j},N_{j',k=j})\in n'''_{12}\end{cases} \qquad (B')$$

where n_{12} is a subset of all pairs of non-negative numbers $N_{j,k}$ and $N_{j',k}$, n'_{12},n''_{12}, and n'''_{12} are mutually exclusive, and $B\epsilon B'$. The only (intuitive) restriction on B' is that $\mu(\bullet)$ is equal to one of three values: 0, $\Pr(X_j=\beta_i|\overline{s}_{i,j})$ and 1. Notice that since the individual does not know a-priori whose the "right" firm for her the sets n'_{12} and n'''_{12} are symmetric.[10] Furthermore, the pairs $N_{j,k=j}=N_{j',k=j}$ are in n''_{12}.

Proposition 12 *The only beliefs in B' that lead to a* ***separating*** *equilibrium are B.*

Proof. Since we are only interested in characterizing separating equilibria, equilibria in the set n''_{12} are not of interest. Further, because of symmetry, we only need consider separating equilibria in n'_{12}; the same logic applies to equilibria in n'''_{12}. We start by ruling out pairs that cannot be an equilibrium, and show that the only set of strategies and beliefs that can form a separating equilibrium are the ones stated in the proposition.

All the pairs in n'_{12} with $N_{j',k=j}>0$ cannot lead to a separating equilibrium, because firm j' has no incentive to send ads to get zero market share. This establishes that $N_{j',k=j}=0$ in any such equilibrium.

Among all pairs $(N_{j,k=j},0)\in n'_{12}$, any pair where $N_{j,k=j}<N^L$ cannot be an equilibrium since

[9] Notice that for some $N^*\epsilon\{N^L,N^H\}$ this proof will not hold.

[10] Let the actions of the two firms j and j' be $(N_{j,k},N_{j',k})$ for any k. Then, if individuals' beliefs of the firm types are: (p_1,p_2), then symmetry implies that if the actions were $(N_{j',k},N_{j,k})$, then beliefs on types must be (p_2,p_1).

firm j' can profitably deviate and mimic firm j. However, even pairs $(N^L \leq N_{j,k=j}, 0) \in n'_{12}$ can not be an equilibrium, if the pairs $(N_{j,k=j} < N^L, 0)$ are in n'_{12}. Thus, n'_{12} is restricted to include only those pairs where $N_{j,k=j} \geq N^L$. Furthermore, the pairs $(N^L \leq N_{j,k=j}, 0) \in n'_{12}$ can not be an equilibrium, if the pairs $(N_{j,k=j}, N_{j',k=j} < N^L)$ are *not* in n'_{12}, because firm j' can deviate and get profits. Thus, n'_{12} should include all pairs where $(N_{j,k=j} \geq N^L, N_{j',k=j} < N^L)$ and only these pairs.[11] ■

We are now ready to relax some of the assumptions made above.

3 Extensions

3.1 Endogenous Price

Once gain we study the non strategic equilibirum first.

3.1.1 Non Strategic Equilibrium

The individual decision When the price is not equal to α, the individual might decide not to buy any product. This would happen when the price is high and the posterior probability of each type is close to 0.5. In such a case the expected utility from any product would be negative. For example, when the price of both product is the same and high $p_j = \alpha + 0.49$ for any j, and the posterior probability is $\Pr(X_j = \beta_i | \bullet) = 0.51$, then the expected utility from product of firm j is $\alpha + \frac{1}{2}(0.51 - 0.49) - (\alpha + 0.49) = (-0.47) < 0$ and the expected utility from the product of the other firm is even smaller. It can be shown that the decision rule of the individual is:

$$\begin{aligned} d_{i,j} &= 1 \Longleftrightarrow \beta_i \bar{s}_{i,j} \geq \max\left(s^*_{i,j}(p_j, \alpha, q), s^{**}_j(p_j, p_{(3-j)}, q)\right) \qquad (16) \\ \text{where } s^*_{i,j}(p_j, \alpha, q) &\equiv \frac{\ln\left(\frac{1+2(p_j-\alpha)}{1-2(p_j-\alpha)}\right)}{\ln\left(\frac{q}{1-q}\right)} \text{ and } s^{**}_j(p_j, p_{(3-j)}, q) \equiv \frac{\ln\left(\frac{1+(p_j-p_{(3-j)})}{1-(p_j-p_{(3-j)})}\right)}{\ln\left(\frac{q}{1-q}\right)} \end{aligned}$$

[11] Notice that it does not matter in which set are the pairs $(N_{j,k=j} \geq N^L, N_{j',k=j} \geq N^L)$. Thus, the beliefs B'' where these pairs are in n'_{12} also lead to a separating equilibrium.

Notice that if the prices are the same, $s_j^{**}(p_j, p_{(3-j)}, q) = 0$. Furthermore, if the prices are the same and $p_j > \alpha$ then for small enough $\bar{s}_{i,j}$, the individual would not buy any product (since her expected utility from both products is negative).

The Firms' Strategies The firms choose the price and the ad intensity that maximize their profit function. Given the prices, the optimal number of ads for each firm in each media channel satisfies the following conditions:

$$q(2q-1)\Pr\left((\beta_i \bar{s}_{i,j}+1) > \beta_i \cdot \max\left(s_{i,j}^*(p_j,\alpha,q), s_j^{**}(p_j,p_{(3-j)},q)\right) > \beta_i \bar{s}_{i,j} | (N_{j,i}+N_{(3-j),i})\right) \geq 2c$$

and

$$q(2q-1)\Pr\left((\beta_i \bar{s}_{i,j}+1) > \beta_i \cdot \max\left(s_{i,j}^*(p_j,\alpha,q), s_j^{**}(p_j,p_{(3-j)},q)\right) > \beta_i \bar{s}_{i,j} | (N_{j,i}+2+N_{(3-j),i})\right) < 2c$$

Notice that the firms consider adding two ads at the margin, since the optimal number would be either odd or even depending on $\max\left(s_{i,j}^*(p_j,\alpha,q), s_j^{**}(p_j,p_{(3-j)},q)\right)$. For example, when $\max\left(s_{i,j}^*(p_j,\alpha,q), s_j^{**}(p_j,p_{(3-j)},q)\right) = 0.1$ the optimal number of ads (of both firms) in a media channel should be even. Notice that with zero ads the individual would not buy any product, with one ad (and $q = \frac{3}{4}$) she would buy the product of the "right" firm with probability $\frac{3}{4}$ and the wrong firm with probability $\frac{1}{4}$. With two ads the probabilities are ($\frac{9}{16}$ and $\frac{1}{16}$) etc. Thus, the optimal number of ads in a media channel would not be even, since any firm can lower the number of ads by one, increase its market share and decrease its cost. Thus, in equilibrium for any pair of prices $N_{j,j} > 0$ $N_{j,(3-j)} = 0$.

Furthermore, using simulations we find that in equilibrium $p_j = p_{(3-j)}$. That is the two firms choose the same price. This is not surprising given the symmetric set-up. The prices are always between α and $\alpha + 0.5$. This make sense since for any price below α the expected number of consumers of each firm is one, and thus its revenues is p_j. Thus, the for any $p_j < \alpha$ there is a higher price $p_j < p_j' \leq \alpha$ that yields a higher profit. For any $p_j > (\alpha + \frac{1}{2})$ the expected number of consumers is zero. We also find that when q increases the prices increases and the number of ads decreases. This is also quite intuitive. If there was not uncertainty in this model, then the optimal price would have been $p_j = (\alpha + \frac{1}{2})$ (we assume that when the expected utility is zero the individual buys the product). On the other hand if there is uncertainty and $q = \frac{1}{2}$ the firms would not send any ads and cannot charge any price above α. Thus, when q increases the price increases, and so are the profits of the firms. On the other hand, when c increases the firms send less ads and thus charge lower prices. Notice that when the firms send less ads the posterior probability of each firm is closer to 0.5 and thus the likelihood of the individual not buying any product increases.

To summarize, with endogenous prices, the prices of both firms are the same, they are a positive function of q and a negative function of c. Furthermore, each firm sends ads only through the media channel consumed by the individual who best fit its product. This is the motivation for the strategic equilibrium.

3.1.2 Strategic Advertising

In such a setting a seperating PBE with targeting exists and firms charge the highest price $p_j = \alpha + \frac{1}{2}$. They charge the highest price, since in equilibrium individuals make their choices without uncertainty. The proof is in Appendix E. Here we present the beliefs and discuss the equilibrium briefly.

The beliefs (denoted with B1) are:

$$\mu(X_j = \beta_i | N_{j,i}, N_{(3-j),i}, \bar{s}_{i,j}) = \left\{ \begin{array}{c} 1 \text{ if } N_{j,i} = N^{PBE},\ N_{(3-j),i} = 0 \\ 0 \text{ if } N_{j,i} = 0,\ N_{(3-j),i} = N^{PBE} \\ Pr(X_j = \beta_i | \bar{s}_{i,j}) = \frac{1}{1+(\frac{1-q}{q})^{\beta_i \bar{s}_{i,j}}} \text{ otherwise} \end{array} \right\} \tag{B1}$$

where N^{PBE} satisfies the conditions:

$$\alpha + \frac{1}{2} - c \cdot N^{PBE} > \max_p \{p \cdot \left[1 + \Pr\left(\left[\beta_i \bar{s}_{i,(3-j)} > \beta_i \cdot s^*_{i,(3-j)}(p, \alpha, q)\right] | 2 \cdot N^{PBE}\right)\right] - 2c \cdot N^{PBE}\} \tag{DC1}$$

$$\alpha + \frac{1}{2} - c \cdot N^{PBE} > \max_{p, N < N^{PBE}} \{p \cdot \Pr\left(\left[\beta_i \bar{s}_{i,j} > \beta_i \cdot s^*_{i,j}(p, \alpha, q)\right] | N\right) - c \cdot N\} \tag{IC1}$$

We consider beliefs that do not inlcude prices. That is prices do not serve as a signal. The deterance condition means that when firm j send N^{PBE} signals in media channel $k = j$, the other firm cannot find a price such that its profit in both markets is higher than $\alpha + \frac{1}{2} - c \cdot N^{PBE}$. Since prices are not part of the beliefs, firm $(3-j)$ would still get the individual $i = (3-j)$ with certainty as long as it send N^{PBE} ads in media channel $k = (3-j)$. Thus, its profit in its market would be $(p - c \cdot N^{PBE})$ and the profit in the other market would be $\left[p \cdot \Pr\left(\left[\beta_i \bar{s}_{i,(3-j)} > \beta_i \cdot s^*_{i,(3-j)}(p,\alpha,q)\right] |2 \cdot N^{PBE}\right) - c \cdot N^{PBE}\right]$. On the other hand, if firm j do not send N^{PBE} ads, the individual is unsure about its identity and firm j cannot charge $p = (\alpha + \frac{1}{2})$ any more. Thus, its profit would be $\max_{p,N<N^{PBE}} \{p \cdot \Pr\left(\left[\beta_i \bar{s}_{i,j} > \beta_i \cdot s^*_{i,j}(p,\alpha,q)\right] |N\right) - c \cdot N\}$.

Therefore, allowing prices to be determine endogenously does not charge the main result–each firm send ads only to the consumers that best fit its product, and consumers are using the ads intensity as signals.

3.2 Individuals do not watch all the time

In reality indivduals do not watch all the time. Thus, they cannot be sure that the firm sent N^{PBE} ads. It can be shown that when individuals consume only part of the media channel both firms would send ads through both media channels with the "right" firm sending more ads to the consumers who best fit its product. In such a case the individuals are using both the content of the ads and the ads intensity when making their choices. Thus, ad intensity still serve as a signal.

4 Other ways informative advertising may look presuasive

There are other ways that informative advertising might look like presuasive advertising. In those ways the focus is not on the targeting strategies of the firms. Thus, we assume at this point that there is only one media channel.

4.1 Risk Aversion

So far we have assumed that the individuals are risk neutral. When they are risk averse their expected utility from each firm depends not only on the expected attribute of each product but also on other moments of its distribution. For example, when the utility is quadratic, the expected utility depends also on the variance of the attributes. The smaller the variance, the more likely the individual chooses the product. It might even be the case, that a product that might look less suitable for the individual will be chosen by her, since it has a smaller variance. In our model the variance depends on the number of the ads received by the individual. The higher the number of ads consumed by the individual, the smaller the variance of the attributes and the more likely the consumer chooses the product. Thus, there is a positive correlation between the number of ads and the sales of the product. In this way risk aversion can lead to the empirical regularity that this study is trying to explain.

This argument can be described formally in the following simple model. The utility has an ideal point structure which can also be written as:

$$U_{i,j} = \alpha_0 + \alpha_1 X_j + \alpha_2 X_j^2 - p_j \tag{17}$$

where $\alpha_0 > 0$, $\alpha_1 > 0$ and $\alpha_2 < 0$. While prices are known to the individuals, the attribute X_j is unknown. The attributes of the two firms are drawn independently from a normal distribution with mean μ_x and variance $\frac{1}{\theta_x}$. Advertising is a noisy signal of this attribute:

$$S_{j,n} \tilde{} N(X_j, \frac{1}{\theta_s}) \tag{18}$$

where n index the number of the signal.

For simplicity we assume that α_0 is large enough so that the individual always prefers to buy a product, and that she needs only one product. Under these conditions the decision rule of the individual is:

$$\begin{aligned} d_{i,j} \;\; = \;\; 1 \;\; & \Longleftrightarrow \alpha_1 \left[E(X_j|\{S_j\}) - E(X_{j'}|\{S_{j'}\})\right] + \alpha_2 \left[E(X_j|\{S_j\})^2 - E(X_{j'}|\{S_{j'}\})^2\right] \qquad (19) \\ & -\alpha_2 \left[V(X_{j'}|\{S_{j'}\}) - V(X_j|\{S_j\})\right] + \left[p_{j'} - p_j\right] \end{aligned}$$

Using Bayes rule we know that:

$$E(X_j|\{S_j\}) = \frac{\theta_x\mu_x + \theta_s\sum_{n=1}^{N_j} S_{j,n}}{\theta_x + N_j\theta_s} \tag{20}$$
$$V(X_j|\{S_j\}) = \frac{1}{\theta_x + N_j\theta_s}$$

where N_j is the number of ads received by the individual about alternative j and $j' = (3-j)$. Thus, as N_j increases the element $-\alpha_2 [V(X_{j'}|\{S_{j'}\}) - V(X_j|\{S_j\})]$ increases and the tendency of the individual to choose alternative j increases. Notice that when N_j increases there is another effect: the estimate of the attribute X_j improves. Thus, an increase of N_j might actually hurt a firm which received a low X_j. This is the "informative effect". However, if α_2 is large enough, the risk aversion dominates the result. Furthermore, when the two firms are close enough to μ_x the "informative effect" is small and N_j has a positive effect on sales.

This simple model should be considered only as an example for the way that risk aversion can lead to an observed behavior that looks like persuasive advertising.

4.2 Existence

Another way to explain the observed persuasive effect with "informative advertising" is by acknowledging that most consumers are not aware about all the alternatives that they are facing. The standard assumption that the individuals are aware of all her alternatives seems to us as quite extreme in the modern economy. It is hard to believe that most people are familiar with all the automobiles that exists, for example. Advertising firms know that a major role of advertising is to inform consumers on the existence of alternatives. This was exactly the role of advertising in Gerard Butters (1977). Advertising that inform the individuals about the existence of an alternative might look persuasive. Consider the following example, when making her choice the individual remembers only a sub-set of the alternatives. Increase in the advertising intensity increases the probability that the individual remembers this alternative when she decides. Since this alternative is the best choice for some of the individuals, an increase in the probability of remembering it leads to an increase in its market share of this product. Such an effect looks like persuasive advertising, since it creates a positive correlation between ad intensity and sales.

5 Conclusion

A significant characteristic of today's economy is the heterogeneity of products. In almost every aspect of life, new products and services are being introduced to satisfy some specific segment of the population. Even products that appear homogeneous—like mobile phones—are designed to attract different segments of consumers. For example, they come in different colors, some subtle, others jumpy. Some phones can only receive calls, others get special discounts for specific calls, etc. At the same time, media channels have also become increasingly segment-specific. The number of magazines, television channels, and websites is growing incredibly rapidly. Most of these channels address the interests of a specific segment. For example, there are not just specific journals on furniture, but even within this category, there are magazines on high quality furniture, "well-priced" furniture, modern furniture, classical furniture, etc. Finally, ads for products are very different in their style. Some ads contain shiny or bright colors, other appear with softer colors. Some feature characters with loud, screaming voices, others with low, calm voices. Some feature families, others feature singles. These differences suggest the informativeness of advertising even though they may not appear to contain any explicit, tangible information.[12]

These three characteristics—product differentiation, heterogeneous media channels (targeting different segments), and informative advertising—are the criticial elements of our model. In this setting, we show that informative advertising can lead to behavior to does not seem rational at first—namely, exposure to ads increases the tendency to buy the promoted product. We show that such behavior is consistent with strategic, rational behavior of consumers. The logic behind this claim is quite intuitive. Firms direct their informative ads to those consumers who their product fits best. Consumers, aware of this strategy, use the ad intensity as a signal of product attributes. That is, the consumer thinks: "If I have seen many ads for

[12] For example, somoe of the ads that target teenagers use language that is so unclear to us (we are both older than thirty) that we do not always know what they mean.

this product, it is probably because it is the best product for me." It turns out that ads do not have to be very informative for this logic to hold. Even if the information in ads is very noisy, firms separate themselves successfully in equilibrium. We also show that: (a) a pooling equilibrium does not exist in this setting; (b) a separating equilibrium does not exist when ads are completely uninformative; and (c) these results are robust to the inclusion of prices, and non-exclusive channel-watching behavior.

This model can be extended in various ways to make it more realistic, and applied to data. For example, as mentioned above, we assume that each type of individual watches only one media channel. In reality, some individuals occasionally consume media channels that do not necessarily fit their characteristics exactly. However, such adaptations to reality would not change the main implications of the study.

6 Appendixes

6.1 Appendix A: Proof of Lemma 1

We can rewrite $ER_{j,i}^{NS}(N_i)$ as:

$$ER_{j,i}^{NS}(N_i) = \Pr(\beta_i \bar{s}_{i,j} > 1|N_i) + \Pr(\beta_i \bar{s}_{i,j} = 1|N_i) + \frac{1}{2} \cdot \Pr(\bar{s}_{i,j} = 0|N_i)$$

Notice that (for $i = j$):

$$\begin{aligned}
&\Pr(\beta_i \bar{s}_{i,j} > 0|N_i + 1) = \\
&\qquad (q + (1-q)) \cdot \Pr(\beta_i \bar{s}_{i,j} > 1|N_i) \\
&\qquad + q \cdot \Pr(\beta_i \bar{s}_{i,j} = 1|N_i) + q \cdot \Pr(\beta_i \bar{s}_{i,j} = 0|N_i)
\end{aligned}$$

and

$$\Pr(\beta_i \bar{s}_{i,j} = 0 | N_i + 1) =$$
$$(1-q) \cdot \Pr(\beta_i \bar{s}_{i,j} = 1 | N_i)$$
$$+ q \cdot \Pr(\beta_i \bar{s}_{i,j} = -1 | N_i)$$

Thus,

$$ER_{j,i}^{NS}(N_i + 1) - ER_{j,i}^{NS}(N_i) =$$
$$(q - \frac{1}{2}) \cdot \Pr(\beta_i \bar{s}_{i,j} = 0 | N_i)$$
$$+ \frac{1}{2} \cdot [q \cdot \Pr(\beta_i \bar{s}_{i,j} = -1 | N_i) - (1-q) \cdot \Pr(\beta_i \bar{s}_{i,j} = 1 | N_i)]$$

Notice that the second line is relevant only when N_i is even and the third line is relevant only when N_i is odd. Furthermore, notice that

$$[q \cdot \Pr(\beta_i \bar{s}_{i,j} = -1 | N_i) - (1-q) \cdot \Pr(\beta_i \bar{s}_{i,j} = 1 | N_i)] = 0$$

since

$$\Pr(\beta_i \bar{s}_{i,j} = 1 | N_i) = \begin{pmatrix} N_i \\ \frac{N_1+1}{2} \end{pmatrix} q^{\frac{N_1+1}{2}} (1-q)^{\frac{N_1-1}{2}}$$

and

$$\Pr(\beta_i \bar{s}_{i,j} = -1 | N_i) = \begin{pmatrix} N_i \\ \frac{N_1-1}{2} \end{pmatrix} q^{\frac{N_1-1}{2}} (1-q)^{\frac{N_1+1}{2}}$$

(Notice that $\begin{pmatrix} N_i \\ \frac{N_1+1}{2} \end{pmatrix} = \begin{pmatrix} N_i \\ \frac{N_1-1}{2} \end{pmatrix}$).

Thus, when N_i is odd $\left[ER_{j,i}^{NS}(N_i+1) - ER_{j,i}^{NS}(N_i)\right] = 0$ and when N_i is even $\left[ER_{j,i}^{NS}(N_i+1) - ER_{j,i}^{NS}(N_i)\right] = (q - \frac{1}{2}) \cdot \Pr(\beta_i \bar{s}_{i,j} = 0 | N_i)$. Since $q > \frac{1}{2}$, $ER_{j,i}^{NS}(N_i)$ is a non decreasing function in N_i for $i = j$.

6.2 Appendix B: Proof of Lemma 2

Recall that for even N_i

$$\Pr(\beta_i \bar{s}_{i,j} = 0 | N_i) = \begin{pmatrix} N_i \\ \frac{N_i}{2} \end{pmatrix} [q(1-q)]^{\frac{N_i}{2}}$$

It is easy to show that

$$\frac{\begin{pmatrix} N_i + 2 \\ \frac{N_i+2}{2} \end{pmatrix}}{\begin{pmatrix} N_i \\ \frac{N_i}{2} \end{pmatrix}} = 4 \cdot \left(\frac{N_i+1}{N_i+2}\right)$$

Thus,

$$\frac{\Pr(\beta_i \bar{s}_{i,j} = 0 | N_i + 2)}{\Pr(\beta_i \bar{s}_{i,j} = 0 | N_i)} = 4 \cdot \left(\frac{N_i+1}{N_i+2}\right) \cdot [q \cdot (1-q)]$$

Since $q > \frac{1}{2}$, $4 \cdot [q \cdot (1-q)] < 1$. Obviously $\left(\frac{N_i+1}{N_i+2}\right) < 1$ and thus $\Pr(\beta_i \bar{s}_{i,j} = 0 | N_i + 2) < \Pr(\beta_i \bar{s}_{i,j} = 0 | N_i)$.

6.3 Appendix C: Proof of Lemma 3

Since $\widetilde{R}_{j',i=j}(N)$ is weakly decreasing in N and $c \cdot N$ is increasing in N, $\pi^d_{j',j}(N)$ is decreasing in N.

Next we show that $\pi^d_{j',j}(1) > 0$ and $\pi^d_{j',j}(\infty) < 0$ to complete the first part of the proof.

$\widetilde{R}_{j',i=j}(1) > c$ by assumption (otherwise, firm $j \neq i$ would *never* have an incentive to mimic firm $j = i$ for any N, and it is trivial to show that there exists a seperating equilibrium). Since adding 1 to an odd number N does not change $\widetilde{R}_{j',i=j}(N)$, we know that $\widetilde{R}_{j',i=j}(2) = \widetilde{R}_{j',i=j}(1)$. Thus, $\widetilde{R}_{j',i=j}(2) - c = \pi^d_{j',j}(1) > 0$.

As $N \to \infty$, $\pi^d_{j',j}(N) \to -\infty$.

This establishes the existence of an N^L that satisfies the conditions in the first part of the Lemma.

The second part of the Lemma follows from the fact that $\pi^d_{j',j}(N)$ is monotonically decreasing in N.

6.4 Appendix D: Proof of Lemma 4

$\pi^d_{j,j}(z)$ is non-decreasing in z, and $c \cdot z$ is strictly increasing in z, therefore $1 - c \cdot z - \pi^d_{j,j}(z)$ is a monotonically decreasing function in z.

Next we show that $1 - c \cdot 1 - \pi^d_{j,j}(1) > 0$ and $1 - c \cdot \infty - \pi^d_{j,j}(\infty) < 0$ to complete the first part of the proof.

Notice that $\pi^d_{j,j}(1) = \frac{1}{2}$. Thus $1 - c \cdot 1 - \pi^d_{j,j}(1) = \frac{1}{2} - c > 0$ by assumption (otherwise, firm j will never place any ads in channel $k = i = j$ in equilibrium).

Obviously, it is easy to show that as $z \to \infty$, $1 - c \cdot z - \pi^d_{j,j}(z) \to -\infty$. This establishes the existence of an N^H which satisfies the conditions in the first part of the Lemma.

The second part of the Lemma follows from the fact that $1 - c \cdot z - \pi^d_{j,j}(z)$ is monotonically decreasing in z.

6.5 Appendix E: Proof of PBE with prices

Definition 13 N^L *satisfies the conditions:*

$$c \cdot N^L > A(N^L) \text{ and } c \cdot (N^L - 1) \leq A(N^L - 1) \tag{21}$$
$$\text{where } A(x) \equiv \max_p \{p \cdot \left[1 + \Pr\left(\left[\beta_i \bar{s}_{i,(3-j)} > \beta_i \cdot s^*_{i,(3-j)}(p, \alpha, q)\right] |2 \cdot x\right)\right]\} - (\alpha + \frac{1}{2})$$

Definition 14 N^H *satisfies the conditions:*

$$B(N^H) > c \cdot N^H \text{ and } B(N^H + 1) \leq c \cdot (N^H + 1) \tag{22}$$
$$\text{where } B(x) \equiv \left(\alpha + \frac{1}{2}\right) - \max_{p, N<x} \{p \cdot \Pr\left(\left[\beta_i \bar{s}_{i,j} > \beta_i \cdot s^*_{i,j}(p, \alpha, q)\right] |N\right) - c \cdot N\}$$

Lemma 15 (N^L) *There exists an* N^L

Proof. *$A(x)$ is a decreasing function in x. Notice that when x increases, the probability of i choosing firm $j \neq i$ decreases unless its price decreased. However if the price decreased as a result of an increase in x, then $A(x)$ must decreased. Otherwise, the price with the lower x was not optimal.*

Furthermore, $A(0) > c \cdot 0$ and $A(\infty) < c \cdot \infty$. ■

Lemma 16 (N^H) *There exists an* N^H

Proof. *$B(x)$ (by definition) is a non increasing function in x.*

Furthermore, $B(0) > c \cdot 0$ and $B(\infty) < c \cdot \infty$

Proposition 17 *For $q > \frac{1}{2}$, $N^H > N^L$*

Proof. [13] *If $q = 1$ then for any $x \geq 1$, $B(x) - A(x) > 0$. If $q = \frac{1}{2} + \varepsilon$ then there exist an x for which $B(x) - A(x) \cong 0$ and $B - A > 0$ for any number above x. We show, using simulation that for any set of parameters (α, c, q) such an $x < N^L$. Thus, for $x = N^L$, $B > A$ and thus $N^H > N^L$.* ■

■

[13] $A(0) = \max(2\alpha, (\alpha + \frac{1}{2})) = 2\alpha$, $B(0) = (\alpha + \frac{1}{2})$, $A(\infty) = 0$, $B(\infty) > 0$. *Furthermore, it can be shown that B is decreasing slower than A.*

Proposition 18 *For $q > \frac{1}{2}$, there exist at least one perfect Bayesian equilibria with beliefs B1 and pure strategies where for each j, $N^*_{j,j} = N^{PBE} > 0$, $N^*_{j,(3-j)} = 0$.*

Proof. *The logic is the same as the proof for PBE without prices.* ■

References

[1] Ackerberg, Daniel (1998) "Empirically distinguishing informative and prestige effects of advertising," Mimeo.

[2] Abernethy, Avery and George Franke (1996) "The Information Content of Advertising: A Meta-Analysis," *Journal of Advertising,* 25 (Summer), 1-17.

[3] Becker, Gary and Kevin Murphy (1993) "A Simple Theory of Advertising as a Good or Bad," *Quarterly Journal of Economics*, 108 (4), 941-964.

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

[5] Erdem, Tulin and Michael Keane (1996) "Decision-Making Under Uncertainty: Capturing Dynamic Brand Choice Processes in Turbulent Consumer Goods Markets," *Marketing Science*, 15 (1), 1-20.

[6] Galbraith, John Kenneth (1971) *The New Industrial State*, Second Edition. New York, The New American Library Inc.

[7] Jacoby, Jacob and Wayne Hoyer (1982) "Viewer Miscomprehension of Televised Communication: Selected Findings," *Journal of Marketing*, 46 (4), 12-26.

[8] Jacoby, Jacob and Wayne Hoyer (1989) "The Comprehension/Miscomprehension of Print Communcation: Selected Findings," *Journal of Consumer Research*, 15(4), 434-443.

[9] Kreps, David and Robert Wilson (1982) "Sequantial equilibrium," *Econometrica*, 50 (4), 863-894.

[10] Milgrom, Paul and John Roberts (1986) "Price and Advertising Signals of product Quality," *Journal of Political Economy*, 94 (4), 796-821.

[11] Nevo, Aviv (2000) "Measuring Market Power in the Ready-to-Eat Cereal Industry," forthcoming *Econometrica.*

[12] Roberts, Mark and Larry Samuelson (1988) "An empirical Analysis of Dynamic, Nonprice Competition in an Oligopolistic Industry," *RAND Journal of Economics*, 19 (2), 200-220.

[13] Shachar, Ron and Bharat Anand (1998) "The Effectiveness and Targeting of Television Advertising," *Journal of Economics and Management Strategy*, 7 (3), 363-396.

[14] Shum, Matthew (1999) "Advertising and swithcing behavior in the breakfast cereals market," Mimeo.

[15] Stern, Scott and Manuel Trajtenberg (1998) "Empirical implications of physician authority in pharmaceutical decision making," The Foerder institute for economic research working paper no. 24-98.

[16] Telser, Lester G. "Advertising and Competition," *Journal of Political Economy*, 72 (6), 537-562.