MANAGEMENT SCIENCE
Vol. 52, No. 5, May 2006, pp. 713–725
ISSN 0025-1909 | EISSN 1526-5501 | 06 | 5205 | 0713

DOI 10.1287/mnsc.1050.0489

Targeted Advertising Strategies on Television

Esther Gal-Or
Katz School of Business, University of Pittsburgh, 368A Mervis Hall,
Pittsburgh, Pennsylvania 15260, esther@katz.pitt.edu

Mordechai Gal-Or
A. J. Palumbo School of Business Administration, Duquesne University, 470 Rockwell Hall,
Pittsburgh, Pennsylvania 15282, galor@duq.edu

Jerrold H. May
Katz School of Business, University of Pittsburgh, 214 Mervis Hall,
Pittsburgh, Pennsylvania 15260, jerrymay@katz.pitt.edu

William E. Spangler
A. J. Palumbo School of Business Administration, Duquesne University, 924 Rockwell Hall,
Pittsburgh, Pennsylvania 15282, spangler@duq.edu

The personal video recorder (PVR) facilitates the use of targeted advertising by allowing companies to monitor television viewing behavior and to build demographic profiles of viewers from the data that are collected. Our research explores the extent to which an advertiser should allocate resources to increase the quality of its targeting. We present a game-theoretic model that extends the conventional measurement of targeting quality by exploring the trade-off between two measures: accuracy and recognition. Accuracy measures the likelihood that any target segment prediction is correct, while recognition conversely measures the likelihood that any member of the target segment is identified. We find that the relative resources allocated to improving accuracy and recognition depend upon the size of the population of viewers, the propensity of viewers to skip commercials, the overall cost of airing commercials, and the competitive environment. Furthermore, the incentives to improve accuracy are markedly different from those to improve recognition. Although improving accuracy does not affect the extent of price competition, improving recognition leads to intensified price competition and reduced profitability in the product market. Thus, when facing a competitor that pursues a strategy to improve its recognition of potential customers, an advertiser should choose to reduce its investment in recognition and increase its investment in accuracy.

Key words: addressable advertising; targeting measures; media content distributor; skipping commercials; game theory; price competition
History: Accepted by Jagmohan S. Raju, marketing; received November 24, 2004. This paper was with the authors 2 months for 3 revisions.

1. Introduction

The past five years have witnessed the growth of the digital personal video recorder (PVR), which allows viewers to digitally record television programming while also allowing them to eliminate, or "zap," in-stream commercials. Because most PVR subscribers report skipping the majority of ads, continued market penetration of the PVR poses a profound threat to television networks and advertisers (Wooley 2003). Forrester Research predicts that 27% of households in the United States will have PVRs by 2007 (Rose 2003) and expects a corresponding 20% decrease in viewing of television commercials over the same period (Chunovic 2001). In a fundamental way, the PVR jeopardizes the decades-old business model by which advertising has provided revenue for television networks and financed so-called "free" programming. The potential impact has been characterized within the industry as "cataclysmic" (Rose 2003).

Nevertheless, this threat to the traditional business model is an opportunity for companies to market their products more effectively, and to more efficiently allocate marketing resources. In the context of advertising, one approach is for companies to use new technologies to better identify those individuals or households who are most likely to be interested in their products or services, and present those potential customers with commercials that they will choose to watch. Because a PVR is programmable and addressable, it can record the viewing patterns of identifiable households and disclose their choices to others. PVRs from companies such as Tivo are able to collect moment-by-moment information regarding which shows and commercials are watched, which commercials are skipped, and which shows and commercials are rewound and watched again (Rose 2003). Advertisers and content distributors[1] then can use data mining methods to infer characteristics of the household from its viewing choices—i.e., to build a profile of the household—and to send targeted advertising to the PVR that is relevant to that type of household (Spangler et al. 2003).

Given that advertisers are motivated to pursue targeted advertising, the objective of our research is to understand the extent to which an advertiser should allocate resources to increase the quality of its targeting. We present a game-theoretic model that contributes to the literature by extending the conventional measurement of targeting quality and by basing the model on customer viewing behavior. Traditionally, researchers have used a unidimensional descriptor of targeting quality. In Chen and Iyer (2002) the descriptor is referred to as addressability, and it measures the probability of reaching consumers that belong to the target population. Chen et al. (2001) use the concept of targetability to measure the expected proportion of customers who are correctly classified. In contrast to those unidimensional measures, we introduce a bidimensional descriptor to evaluate the quality of the targeting, using what we term *accuracy* and *recognition*. Both accuracy and recognition are conditional probability measures. Accuracy is the conditional probability that a customer is actually in the target segment given that he/she was predicted to be in the target segment. In the definition of recognition, the order of the conditioning is reversed. That is, recognition is the conditional probability that an individual is predicted to be in the target segment given that he/she is actually in the target segment. Hence, the latter measure indicates how many of the actual members of the target segment are identified, or recognized, by that targeting mechanism. Note that the single targetability measure discussed in Chen et al. (2001) is a weighted average of the two conditional probabilities that we term accuracy and recognition. The addressability measure of Chen and Iyer (2002) is equivalent to our concept of recognition. As we later demonstrate, improvements in accuracy may have different implications on profitability than improvements in recognition. As a result, in deciding how to allocate resources to improve the quality of targeting, marketers must take both measures into consideration because of the inherent trade-off between the two of them. In general, improving accuracy tends to require a more conservative selection process, which necessarily decreases the selected pool and may reduce recognition. Conversely, improving recognition requires a more liberal selection process, which in turn may decrease accuracy. The trade-off between those two measures of the quality of targeting is related to the trade-off that exists also between the Type 1 and Type 2 errors in hypothesis testing. Our analysis highlights the relationship among those different measures of the quality of the statistical analysis.

We demonstrate that the relative resources allocated to improving accuracy and to improving recognition depend crucially upon the competitive environment in which advertisers operate. Focusing on advertisers who compete in a given product market, we derive accuracy and recognition reaction functions to assess the incentives of a given advertiser to improve the quality of targeting viewers as a function of choices made by its competitors. We find, in particular, that the incentives to improve accuracy can be markedly different from those to improve recognition because the two types of improvements have different implications with respect to the extent of price competition in the product market. Although improving accuracy does not affect the extent of price competition, improving recognition leads to intensified price competition and to reduced profitability in the product market. As a result, when facing a competitor that chooses to improve the extent of recognition of potential customers, an advertiser should choose, in response, reduced investment in improved recognition and increased investment in improved accuracy.

It is noteworthy that the previous literature makes different assumptions concerning the capabilities of the targeting technology. Chen and Iyer (2002), for instance, assume that the technology allows marketers to identify groups of consumers interested in a certain product class, without being able to distinguish between those who prefer one brand over another. In contrast, Chen et al. (2001), Gal-Or and Gal-Or (2005), and Iyer et al. (2003) consider the implications of a more powerful technology that facilitates also the identification of consumers according to their extent of loyalty to the different brands that compete in the given product category. In this paper, we assume (as did Chen and Iyer 2002) that the technology allows advertisers to only predict whether viewers are interested in the general product class in which two brands compete. The prediction is based on demographic and/or psychographic characteristics that are inferred from the viewing habits of the population. Our assumption implies, for example, that while the targeting technology can predict that the viewer belongs to an age bracket likely to be interested in a sports car, it cannot identify whether the viewer would prefer a Maserati over a Porsche.

The targeting literature has considered two different objectives in targeting different segments of consumers. While Chen et al. (2001) and Chen and Iyer

[1] That is, cable, satellite, and broadcasting companies.

(2002) focus on the use of the targeting technology to facilitate price discrimination, Gal-Or and Gal-Or (2005) and Iyer et al. (2003) focus on the application of the technology to the customization of ads. Similarly, the targeting technology in our model is aimed at facilitating the customization of commercials on television by using the PVR. In contrast to the previous studies, we incorporate in our model the feature that viewers can actively skip commercials, and their propensity to do so depends upon the quality of the targeting technology. Hence, the levels of accuracy and recognition selected by advertisers affect not only the extent of correct classification of different segments of consumers, but also the size of the audience that views the commercials. We point out, in particular, that improved quality of targeting viewers by noncompetitors does not have any adverse affect on the extent of price competition in product markets. Its only effect is to confer a positive externality on all noncompeting advertisers, because viewers are less inclined to zap through commercials when they expect them to be better aligned with their interests. A reduced propensity to zap on the part of viewers increases, in turn, the de facto size of the viewing population, thus strengthening the incentives for all noncompeting advertisers to improve the quality of their targeting technology. Hence, both accuracy and recognition rise when a noncompeting advertiser improves the quality of its targeting technology.

Because the content distributor delivers the targeted commercials of all advertisers through a PVR, we also investigate the preferences of the content distributor with regard to the levels of accuracy and recognition that advertisers select. We show that when the content distributor charges higher rates to air commercials, advertisers find it optimal to cut back on levels of recognition and increase levels of accuracy. However, the profits of the content distributor rise if advertisers choose to increase recognition and reduce accuracy. Higher levels of recognition for a fixed level of accuracy imply that the size of the viewing audience is larger, while lower levels of accuracy for a fixed level of recognition imply that advertisers need to send more commercials. Both effects improve the profits of a content distributor who is paid at a rate tied to the number of commercials that are watched by viewers. Hence, in choosing its price for airing commercials, the content distributor should balance the positive implication of charging higher rates against the adverse effect such higher rates have on the recognition and accuracy levels selected by advertisers. The earlier literature did not consider at all the incentives of the entity that provides the targeting technology to the competing firms in the product market.

Prior research supports our assumption that viewers are less likely to zap through commercials when advertisers improve their targeting technology. For example, Siddarth and Chattopadhyay (1998) found that the likelihood of zapping an ad is lower for viewers who tend to make purchases in the product category. They argue for the segmentation of markets based in part on consumer behavior (Rossiter and Percy 1997), as well as the use of more "precise" media targeting such as direct matching (Assael and Poltrack 1991). Others have suggested that improvements in targeting and customer addressability should improve the feasibility of identifying and selling to small market niches (Blattberg and Deighton 1991), while also increasing the overall effectiveness of advertising and (therefore) advertising expenditures (Iyer et al. 2003). That is also true for television advertising. As PVR technology continues to improve, the ability of advertisers to precisely target and deliver relevant advertising to their customers also is expected to improve, thus reducing the likelihood that customers will choose to skip correctly targeted commercials.

The remainder of this paper is organized as follows. In §2, we describe the primary assumptions of the model. In §3, we derive recognition and accuracy reaction functions for competing advertisers, and characterize the symmetric equilibrium when they choose identical levels of targeting quality. In §4, we derive the optimal price for airing commercials as selected by the content distributor. Section 5 contains conclusions and future research directions. The proofs of all lemmas and propositions are included in a technical appendix that is available on the *Management Science* website (http://mansci.pubs.informs.org/ecompanion.html).

2. The Model

The population is divided into two segments, denoted as Segment 0 and Segment 1. The segments may be divided by age, gender, income bracket, and so on. Segment 1 comprises a fraction a of the population, and Segment 0 includes the remaining $(1-a)$ of the population. There is a group of advertisers who are trying to reach Segment 1. Two of those advertisers, which we designate as Advertisers 1 and 2, compete against each other in a certain industry. All remaining advertisers in the group operate in different industries, and are noncompetitors as a result. Designate the group of noncompeting advertisers whose products appeal to Segment 1 by S_1. Similarly, let S_0 designate all advertisers whose products appeal to Segment 0 of the population. Group S_0 wishes to deliver advertising messages, therefore, to Segment 0.

The PVR, in this context, serves as a communication interface between advertisers and individual households, in that it collects viewing patterns of households and subsequently distributes targeted advertising to them. Between those two events, an advertiser performs an analysis of the viewing patterns to identify or target individual viewers and then selects advertising that is relevant to the targeted viewer. The degree to which relevant advertising is selected is dependent on the quality of the targeting technology that identifies the characteristics of individual households from their viewing behaviors. We introduce two different measures to evaluate the quality of the targeting technology. For the advertisers wishing to target Segment 1 of the population, those measures are defined as follows:

$$\begin{aligned}\text{recognition} &= R\\ &= \text{Prob}(\text{predicted segment}=1 \mid \text{actual segment}=1),\\ \text{accuracy} &= A\\ &= \text{Prob}(\text{actual segment}=1 \mid \text{predicted segment}=1).\end{aligned}$$

Recognition measures the likelihood that an individual in Segment 1 is recognized correctly by the technology. Accuracy measures the rate at which the technology is accurate in predicting that a viewer belongs to Segment 1.

Two group classification problems such as the one we consider are commonly analyzed using receiver operating characteristic (ROC) curves, and their performance is tabulated in confusion matrices such as the one shown in Table 1, which counts the frequency in each cell (Hand 1997).

The ratio (true positives)/(true positives + false negatives) is called the *sensitivity* of the model, and the ratio (true negatives)/(true negatives + false positives) is called its *specificity*. Denote the conditional probabilities of the errors committed by a model as

$$\begin{aligned}\alpha &= 1-\text{sensitivity}\\ &= \text{Prob}(\text{predicted segment}=0 \mid \text{actual segment}=1)\\ &= \frac{\text{false negatives}}{\text{true positives}+\text{false negatives}},\\ \beta &= 1-\text{specificity}\\ &= \text{Prob}(\text{predicted segment}=1 \mid \text{actual segment}=0)\\ &= \frac{\text{false positives}}{\text{true negatives}+\text{false positives}}.\end{aligned}$$

In a hypothesis-testing context, if Segment 1 corresponds to the null hypothesis being true, then α is the probability of committing a Type 1 error and β is the probability of committing a Type 2 error.

Using the notation in Table 1, the recognition R and accuracy A are given by

$$\begin{aligned}R &= \text{Prob}(\text{predicted segment}=1 \mid \text{actual segment}=1)\\ &= (\text{true positives})/(\text{true positives}+\text{false negatives}),\\ A &= \text{Prob}(\text{actual segment}=1 \mid \text{predict segment}=1)\\ &= (\text{true positives})/(\text{true positives}+\text{false positives}).\end{aligned}$$

Let a denote the proportion of the population that belongs to Segment 1, and $(1-a)$ the proportion that belongs to Segment 0. The relationship between the recognition R and the accuracy A and α and β for a given advertiser can be expressed as follows:

$$R = 1-\alpha, \tag{1}$$

$$A = \frac{(1-\alpha)a}{[(1-\alpha)a+\beta(1-a)]}. \tag{2}$$

The probability density function described in Table 2 summarizes the relationship expressed in Equations (1) and (2).

We assume that the quality of the targeting technology is determined by the joint efforts of the content distributor and the advertisers. As pointed out earlier, because PVRs are programmable, they can record and transmit household viewing patterns. The content distributor can utilize data mining techniques to predict the demographic profiles of households based upon their viewing choices. Those techniques may enable the distributor to predict, in particular, whether or not a given viewer belongs to Segment 1 of the population. For example, viewers of the *Survivor* reality show and the sitcom *Friends*, and viewers who tend to watch shows on weekday evenings but not on the weekends, might tend to be unmarried individuals in their early twenties who earn between $35,000–$50,000 per year. The distributor can selectively deliver relevant commercials to the set-top boxes of those individuals dealing, for instance, with online matchmaking sites or sports cars.

Table 1 Confusion Matrix

	Predicted segment	
Actual segment	0	1
0	True negatives	False positives
1	False negatives	True positives

Table 2 Probability Density of Confusion Matrix

	Predicted segment	
Actual segment	0	1
0	$(1-\beta)(1-a)$	$\beta(1-a)$
1	αa	$(1-\alpha)a$

While the content distributor determines the basic underlying quality of its profiling engine, individual advertisers may choose to improve upon the data mining techniques developed by the content distributor by conducting supplementary surveys of the population. They could survey, for instance, a sample of households who are predicted to belong to Segment 1 to verify their actual type. By doing so, an advertiser might be able to increase the level of accuracy of his predictions (see the online appendix). Alternatively, he could survey a sample of those who are predicted to belong to Segment 0, and by doing so improve both the recognition and accuracy levels of his predictions (see the online appendix). By choosing to survey a carefully selected mixed sample consisting of viewers from both predicted segments, an advertiser may be able to gain full control over the extent of improvement along the two different measures of the quality of the targeting technology. However, such improvements are costly to the advertiser, given that they require direct surveying of individuals. We designate by $TC(R, A)$ the cost an advertiser must incur to obtain the recognition and accuracy levels R and A, respectively, in targeting individuals who belong to Segment 1 of the population. The advertiser incurs this cost in addition to the direct payments he must make to the content distributor for delivering his commercials to the target population. We assume that the cost function $TC(\cdot)$ is a strictly increasing and convex function of its arguments.

Note that while the targeting technology permits advertisers to distinguish among different demographic segments, it does not facilitate recognizing the relative preferences of individuals within a segment between two competing brands of a certain product class.[2] Hence, while single individuals in their twenties can be identified, the technology cannot distinguish between those who prefer one matchmaking service over another. In the context of our model, the television distributor can distinguish between Segment 1 and Segment 0 of the population, and try to deliver, therefore, commercials on behalf of only advertisers interested in reaching a particular segment. However, within this segment, the targeting technology cannot predict the individuals who are more inclined to purchase Product 1 over its competitor in the market, Product 2.

We assume that the content distributor charges each advertiser an amount of c dollars for each commercial that is watched by the viewers.[3] Because the PVR technology has the capability to record the number of commercials that viewers choose to watch, such a payment scheme is enforceable. It implies, in particular, that advertisers are not charged for commercials that consumers choose to zap.[4]

To illustrate how the quality of the targeting technology, as measured by the variables R and A, affects the profitability of each advertiser, we consider first the case that the advertiser's product price is fixed at the level p and that viewers are unable to zap through commercials. For simplicity, assume also that if an individual in Segment 1 is correctly recognized and watches a commercial from a product that is relevant to her interests, she purchases this product with probability 1. From Table 2, the fraction of the population predicted to belong to Segment 1 amounts to $[\beta(1-a)+(1-\alpha)a]$. Hence, the distributor sends advertising messages on behalf of Advertiser 1 to that fraction of the population. Because viewers cannot zap through commercials, the advertiser has to pay $c[\beta(1-a)+(1-\alpha)a]$ to the distributor. However, only those viewers who actually belong to Segment 1 have an interest in Advertiser's 1 product. Hence, Table 2 indicates that only the fraction $(1-\alpha)a$ of the population will purchase Product 1, yielding revenue $p(1-\alpha)a$ to Advertiser 1. As a result, for fixed-error rates α and β, the profits of an advertiser attempting to reach viewers in Segment 1 (gross of the cost of implementing the selected levels R and A), can be expressed as follows:

$$\Pi = p(1-\alpha)a - c[\beta(1-a)+(1-\alpha)a]. \tag{3}$$

Using Equations (1) and (2) to substitute for rates α and β in terms of R and A yields

$$\Pi = Ra\left[p - \frac{c}{A}\right]. \tag{4}$$

Hence, the expected profits of the advertiser (gross of the cost $TC(R, A)$) increase with both the levels of recognition and accuracy that it selects. Increased recognition guarantees that it is able to reach the correct segment of viewers more frequently, and increased accuracy guarantees that it has to deliver fewer ads to reach the desired segment of the population (for every ad received by the target population it has to deliver $1/A$ ads).

3. The Behavior of Competing Advertisers 1 and 2

Returning now to an environment in which viewers may choose to zap through commercials, we assume

[2] Gal-Or and Gal-Or (2005) and Iyer et al. (2003) consider a targeting technology that does permit an advertiser to identify those consumers who prefer their product over those of their competitors in the marketplace.

[3] We consider two-part tariff schemes in the online appendix.

[4] This is the common method of payment for ads on the Internet, where advertisers are charged only for "click through" ads.

that the frequency of zapping depends on the extent of misclassification that the distributor experiences in delivering commercials to viewers.[5] Specifically, we assume that the proportion of consumers who watch commercials declines linearly in the average rate of misclassification in delivering advertising messages to them. Essentially, the content distributor establishes a certain reputation regarding its ability to target commercials correctly. Each segment of consumers responds to that reputation when deciding whether or not to skip commercials. The rate at which consumers skip commercials is linearly and inversely related to the average error rate that is relevant to the segment.

The extent of misclassifications committed by Advertisers 1 and 2 are measured in terms of their error rates α_i, β_i, $i = 1, 2$, where α_i affects Segment 1 of the population and β_i affects Segment 0. For the remaining advertisers, let M_1 denote the average rate of misclassification affecting Segment 1 of the population, and M_0 the similar rate affecting Segment 0.

Given the above assumptions, the expected number of viewers in Segment 1 who remain tuned in during any commercial sent to them is equal to

$$\tilde{a} = aP_r^1, \quad \text{where } P_r^1 \equiv \left[\gamma_0 - \gamma_1 \frac{\alpha_1 + \alpha_2}{2} - \gamma_2 M_1\right], \tag{5}$$

and the number of viewers in Segment 0 who remain tuned in during commercials is equal to

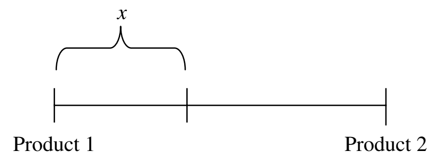
$$\begin{gathered} \widetilde{(1-a)} = (1-a)P_r^0, \\ P_r^0 \equiv \left[\delta_0 - \delta_1 \frac{\beta_1 + \beta_2}{2} - \delta_2 M_0\right], \end{gathered} \tag{6}$$

where $0 \le P_r^1 \le 1$, $0 \le P_r^0 \le 1$, and $\gamma_0, \gamma_1, \gamma_2, \delta_0, \delta_1, \delta_2 \ge 0$.

P_r^1 and P_r^0 in (5) and (6) are the probabilities that viewers in Segment 1 and Segment 0, respectively, view any one commercial. Those probabilities vary inversely with the error rates committed when delivering relevant messages to viewers.

Because our primary objective is to understand the incentives of competing advertisers to invest in the quality of the targeting technology, we now turn to the description of the nature of competition in the product market between Advertisers 1 and 2. We model the competition between those two advertisers using the Hotelling model. Assume that the consumers of Segment 1 wish to buy a single unit of the product of only one of those advertisers. Their preferences between the two products are uniformly distributed on a line of unit length with the products located at the endpoints of this line (see Figure 1). Hence, a consumer located closer to the left endpoint prefers the product of Advertiser 1 to that of Advertiser 2.

Figure 1 **Distribution of Preferences Between the Two Products**



Let x denote the location of a consumer in the distribution of preferences, as measured by distance from the left endpoint of the line. The consumer has the reservation price v for a product that constitutes a perfect match with her preferences. Her willingness to pay declines to $(v - tx)$ when she purchases a product that is located x units of distance away from her location on the line. The transportation parameter t measures the extent of differentiation between the two products. Larger values of t imply that the consumer is more loyal to her preferred product.

If a consumer in Segment 1 is familiar with only one of the two products, say product i, she purchases it as long as her net utility from the purchase is nonnegative. Let p_i be the price of product i and x the distance of the consumer from this product. Then, if $v - tx - p_i \ge 0$, the consumer buys product i. On the other hand, if a consumer in Segment 1 is familiar with the products of both Advertisers 1 and 2, she compares the prices of the two products, and contingent on her location on the line, chooses the product that offers her the higher net utility. If $v - tx - p_1 > v - t(1-x) - p_2$ and $v - tx - p_1 \ge 0$, the consumer buys Product 1.

Assuming that the prices of the products are sufficiently low so that $v - tx - p_i > 0$ for $i = 1, 2$, a consumer who is familiar with both products will be indifferent as to which one she purchases if

$$v - tx^* - p_1 = v - t(1 - x^*) - p_2. \tag{7}$$

Solving for x^* from (7) yields that the consumer buys the product of Advertiser 1 if

$$x \le \frac{1}{2} + \frac{p_2 - p_1}{2t} \equiv x^*, \tag{8}$$

otherwise, she buys the product of Advertiser 2.

For given levels of errors of (1 − sensitivity), α_1 and α_2, selected by the two advertisers, the probability that a viewer of a commercial who belongs to Segment 1 is familiar with both advertisers amounts

[5] As noted, Siddarth and Chattopdahyay (1998) found that the likelihood of zapping declines for viewers who tend to make purchases in the product category.

to $(1-\alpha_1)(1-\alpha_2)$. The probability that the same viewer is familiar only with Advertiser i amounts to $(1-\alpha_i)\alpha_j$.

From the above derivation, it follows, therefore, that the revenue Advertiser i expects from a viewer who belongs to Segment 1 is

$$(1-\alpha_1)(1-\alpha_2)\left[\frac{1}{2}+\frac{p_j-p_i}{2t}\right]p_i+(1-\alpha_i)\alpha_j p_i, \quad i,j=1,2;\ i\neq j. \tag{9}$$

The first term measures revenue from a viewer who is familiar with both advertisers, and the second term measures revenue from a viewer who is only familiar with Advertiser i.

Assuming that only individuals in Segment 1 have an interest in buying Products 1 or 2, we can derive the expected revenue of Advertisers 1 and 2 as follows:

expected revenues of Advertiser i

$$=\tilde{a}\left[(1-\alpha_1)(1-\alpha_2)\left(\frac{1}{2}+\frac{p_j-p_i}{2t}\right)p_i+(1-\alpha_i)\alpha_j p_i\right], \quad i=1,2;\ i\neq j,$$

where we incorporate the possibility that viewers of Segment 1 may choose to skip through commercials, implying that the fraction of the audience included in Segment 1 declines from its original size a to $\tilde{a}$. As previously explained, such a decline depends on the average rate of misclassification that the entire population of advertisers, trying to reach Segment 1, experiences.

The payment that Advertiser i makes to the content distributor depends on the total number of commercials that are watched either correctly by Segment 1 viewers or incorrectly by Segment 0 viewers. The average number of viewers in Segment 1 who are expected to watch the commercial of Advertiser i is $\tilde{a}(1-\alpha_i)$, and the average number of viewers in Segment 0 expected to watch its commercial is $(\widetilde{1-a})\beta_i$. As a result, the expected payment to the content distributor is given by $(\tilde{a}(1-\alpha_i)+(\widetilde{1-a})\beta_i)c$. Note that the advertiser incurs costs both when delivering messages correctly to Segment 1 customers (at the rate $1-\alpha_i$) and when delivering messages incorrectly to Segment 0 customers (at the rate β_i). The cost is incurred, however, only for the population of viewers who actually watch the commercials, which is measured by $\tilde{a}$ and $(\widetilde{1-a})$, not by a and $(1-a)$.

Based on the above derivations, the expected profits of Advertiser i gross of the cost it incurs in maintaining a certain quality of the targeting technology (measured either by α_i and β_i or A_i and R_i) are

$$\Pi_i=\tilde{a}\left[(1-\alpha_1)(1-\alpha_2)\left(\frac{1}{2}+\frac{p_j-p_i}{2t}\right)p_i+(1-\alpha_i)\alpha_j p_i-(1-\alpha_i)c\right]-(\widetilde{1-a})\beta_i c, \quad i,j=1,2;\ i\neq j. \tag{10}$$

Using (1) and (2) to rewrite (10) in terms of R_i and A_i yields

$$\Pi_i=\tilde{a}\left[R_1R_2\left(\frac{1}{2}+\frac{p_j-p_i}{2t}\right)p_i+R_i(1-R_j)p_i-\frac{R_i c}{A_i}\right], \quad i,j=1,2;\ i\neq j. \tag{11}$$

We model the game between Advertisers 1 and 2 as consisting of two stages. In the first stage, advertisers independently choose their levels of recognition and accuracy. In the second stage, advertisers act as Bertrand oligopolists to independently choose the prices of their products. Restricting attention to subgame perfect equilibria, we start by considering the second-stage game. At this stage, each advertiser chooses the price of its product, p_i, as a function of recognition and accuracy levels selected by the parties in the first stage of the game. Solving for the pure strategy[6] Nash equilibria of this stage yields the results reported in Lemma 1.

Lemma 1. (i) *The product prices selected by Advertisers* 1 *and* 2 *at the pure strategy equilibrium of the second-stage game satisfy the following equation*:

$$p_i=t\left[1+\frac{2}{3}\left(\frac{1-R_i}{R_i}+\frac{2(1-R_j)}{R_j}\right)\right], \quad i,j=1,2;\ i\neq j. \tag{12}$$

(ii) *The expected profits of the advertisers as a function of the recognition and accuracy levels selected in the first stage are expressed as follows*:

$$v_i=\tilde{a}\left\{\frac{tR_1R_2}{2}\left[1+\frac{2}{3}\left(\frac{1-R_i}{R_i}+\frac{2(1-R_j)}{R_j}\right)\right]^2-\frac{R_i c}{A_i}\right\}-TC(R_i,A_i), \quad i,j=1,2;\ i\neq j, \tag{13}$$

where

$$\tilde{a}=a\left[\gamma_0-\gamma_1\frac{2-R_1-R_2}{2}-\gamma_2 M_1\right]. \tag{14}$$

[6] Note that in Chen and Iyer (2002) only mixed strategy equilibria exist in the pricing subgame. In their model, the firms can price discriminate and choose their prices contingent upon the location x of the consumer on the line.

Note that increased recognition selected by either Advertiser 1 or 2 results in intensified price competition between their products because viewers are more likely to be familiar with both products. When consumers are better informed, producers are forced to compete more aggressively. Moreover, according to (12), the price chosen by a given advertiser is more responsive to the level of recognition selected by the competitor than the level selected by the advertiser itself. As well, note that prices are independent of the levels of accuracy selected by the two competing advertisers.

Improved recognition is not necessarily beneficial to the advertisers even if it can be obtained at no extra cost. From (13), the expected profits of Advertiser i may actually decline with R_i, even when $\partial TC(R_i, A_i)/\partial R_i = 0$. Because higher levels of recognition intensify the extent of price competition between products, advertisers may be reluctant to improve the recognition level of their statistical predictions. In contrast, an advertiser who does not face competition on price will try to improve its recognition to the maximum level possible if such an improvement can be obtained at no extra cost. From (4), an advertiser whose product price p is fixed strictly benefits from higher levels of recognition.

In Lemma 2, we derive conditions to guarantee that the local equilibrium described in Lemma 1 is indeed global. Specifically, no advertiser has an incentive to consider large deviations from the asserted equilibrium. Two types of large deviations should be considered. The first is aimed at forcing the competitor out of the market altogether. For advertiser j, such a deviation implies that $p_j^D = p_i - t$. The second type of deviation is aimed at exploiting the group of consumers who are only informed about one of the products. The deviating advertiser can entertain the possibility of acting as a monopolist vis-à-vis this group of consumers by charging them the price $p_j^{DD} = v - wt$, where $0 < w \leq 1$. Such a price guarantees to the deviating advertiser a segment of size w of the consumers who are only informed about product j. The condition derived in Lemma 2 guarantees that either type of deviation is unprofitable to the advertisers. As a result, we are guaranteed that a pure strategy equilibrium in the pricing subgame exists.[7]

Lemma 2. (i) *When*

$$(R_j - R_i) < \frac{3R_iR_j[7R_j + 2R_i - 6R_iR_j]}{2(R_j + 2R_i)}$$

and

$$\max\left\{2, \left(\frac{1}{R_i} + \frac{1}{R_j} - \frac{1}{2}\right)\right\} \leq \frac{v}{t} \leq 1 + \frac{[2R_i + 4R_j - 3R_iR_j]^2}{18R_iR_j^2(1-R_1)}, \quad i, j = 1, 2;\ i \neq j,$$

the local equilibrium described in Lemma 1 *is global.*

(ii) *If $R_i = R_j = R$, the above condition reduces to*

$$\max\left\{2, \left(\frac{2}{R} - \frac{1}{2}\right)\right\} \leq \frac{v}{t} \leq 1 + \frac{[2-R]^2}{2R(1-R)}. \tag{15}$$

The condition that guarantees the existence of a pure strategy equilibrium in the pricing subgame consists of two parts. The first part requires that the extent of recognition of the two advertisers is not significantly different. When one advertiser is significantly more recognized than the other, there is an increased probability that consumers are only familiar with the product of that advertiser, making its deviation to monopoly pricing profitable. The second part of the condition imposes lower and upper bounds on the value of the ratio (v/t). While the lower bound is necessary to guarantee that the market is fully covered, the upper bound is necessary to guarantee that no firm has an incentive to take advantage of uninformed consumers by charging them the monopoly price.[8]

We now turn to the first stage of the game, when the levels of recognition and accuracy are selected by the advertisers. We use the first-stage payoff function (13) to derive the equilibrium levels of recognition and accuracy. Given that those two different measures of quality of the targeting technology are derived from the confusion matrix of Table 1, their values are not independent. In Lemma 3, we derive a condition to ensure that simple first-order conditions obtained by differentiating the payoff function (13) are still the necessary conditions for an interior Nash equilibrium of the first-stage game.

Lemma 3. *If, at the interior equilibrium, the incidence of "true negatives" in Table* 1 *is strictly positive, then the necessary conditions that characterize the interior equilibrium of the first-stage game satisfy the equations $\partial v_i/\partial A_i = \partial v_i/\partial R_i = 0$. Moreover, if the objective v_i is a strictly concave function of R_i and A_i, those conditions are both necessary and sufficient.*

[7] The nonexistence of pure strategy equilibria in the Hotelling model has been discussed in d'Aspremont et al. (1979) and in Gal-Or (1982). In the case of nonexistence, mixed strategy equilibria should be considered, instead.

[8] With imperfect information, there is a positive probability that consumers are only informed about a single product. If the ratio (v/t) is too big, it is impossible to eliminate the incentive of firms to deviate to monopoly pricing to take advantage of the consumers who are poorly informed. This explains the upper bound on (v/t) that is derived in Lemma 2. We thank an anonymous referee for this observation.

We can now use the payoff function in (13) to characterize the choice of strategies selected by the advertisers in the first stage of the game. We wish to focus, in particular, on the slope of the recognition and accuracy reaction functions. In Proposition 1, we evaluate the signs of those slopes, assuming the existence of an interior equilibrium.

Proposition 1. *Assuming the existence of an interior equilibrium in the first stage*:

(i) *The slopes of the reaction functions of the two advertisers who compete in the same product market satisfy the following*:

$$\frac{\partial R_i}{\partial R_j} < 0, \qquad \frac{\partial A_i}{\partial R_j} > 0, \qquad \frac{\partial R_i}{\partial A_j} = \frac{\partial A_i}{\partial A_j} = 0 \qquad \text{for } i, j = 1, 2;\ i \neq j.$$

(ii) *The recognition and accuracy levels of Advertisers* 1 *and* 2 *are inversely related to the average error rates of advertisers established in other product markets.*

According to Proposition 1, the recognition level of a given firm is inversely related to the recognition choice of its competitor. In contrast, its level of accuracy is directly related to the recognition level chosen by the competitor. When the competitor increases its recognition level, two things happen. First, there is a larger fraction of Segment 1 individuals who are less inclined to zap through commercials (i.e., $\tilde{a}$ is larger) because they expect to receive relevant ads, consistent with their interests, at a higher frequency. Given that the de facto size of Segment 1 viewers becomes larger, the marginal return to investing in improved targeting is higher (increasing either R_i or A_i is more beneficial because of this effect). The second effect of the competitor's decision to increase recognition is that price competition in the second stage intensifies, because consumers are likely to be better informed about the competitor's product. To alleviate that intensified competition effect, the advertiser reduces its own recognition, which, by Lemma 1, yields higher prices (reducing R_i is beneficial due to this second effect). Proposition 1 asserts that the second intensified price competition effect on R_i more than outweighs the first, increased viewership effect, thus yielding downward sloping recognition reaction functions. On the other hand, because the choice of accuracy level is only affected by the first, increased viewership effect, the proposition argues that the level of accuracy is increasing when the competitor improves its recognition.

Note that the level of accuracy chosen by the competitor affects neither the recognition nor the accuracy levels of a given advertiser. Such increased accuracy by the competitor does not influence the de facto number of Segment 1 viewers because the error rates that such viewers are exposed to depend on the accuracy levels selected by advertisers whose target population is Segment 0 and not Segment 1. Moreover, the accuracy level chosen by the competing advertiser does not influence the extent of price competition in the second stage, as implied by Equation (12). Proposition 1 also states that the recognition and accuracy levels of the two competing advertisers depend on the behavior of the advertisers in other product markets because such advertisers affect the de facto size of Segment 1 viewers who watch commercials. As the error rates of those advertisers increase, viewers in Segment 1 are more likely to skip commercials because they expect to be sent more irrelevant commercials and fewer relevant ones. As a result, the incentive of Advertisers 1 and 2 to improve targeting quality declines.

Before we characterize the symmetric equilibrium of the two-stage game, in Lemma 4 we derive a sufficient condition to guarantee that no firm has an incentive to deviate from an interior symmetric equilibrium. Small deviations from an interior, local equilibrium are unprofitable if the condition for reaction function stability holds (see Vives 1999, pp. 47–52). To guarantee that the local equilibrium is also global, large deviations of the kind entertained in Lemma 2 should also be considered. In Lemma 4, we derive a condition to ensure that such large deviations are unprofitable.

Lemma 4. *If at the symmetric equilibrium* (*i.e.*, $R_1 = R_2 = R$ *and* $A_1 = A_2 = A$) *the level of recognition satisfies the condition that*

$$\max\left\{2, \frac{2}{R} - \frac{1}{2}\right\} \leq \frac{v}{t} \leq 1 + \frac{(4-R)^2}{18R(2-R)},$$

no firm has an incentive to unilaterally change its recognition level so that it serves only poorly informed consumers (*i.e., acts as a monopoly vis-à-vis this group of consumers*). *The above region for* (v/t) *values is nonempty provided that* $R > 0.4665$.

With the aid of Lemma 4, we now consider an example that satisfies the conditions of the lemma. Using numerical calculations, we illustrate that no firm wishes to deviate from the symmetric local equilibrium. In the example, $TC(R, A) = 2R^2 + 2A^2$, $\gamma_0 = \gamma_1 = t = 1$, $M_1 = 0$, and $c = 0$. Because $c = 0$, firms do not invest resources to improve accuracy, thus $A_1 = A_2 = 0$. The recognition levels, product prices, and the profits of the firms are reported in Table 3 for three different values of a (i.e., $a = 4$, 4.1, and 4.2). The last column of the table reports the feasible values of the parameter v, that according to Lemma 4, eliminate incentives to deviate from the proposed local equilibrium.

Table 3 Characterization of the Local Symmetric Equilibrium

a	R^*	p^*	Π^*	Feasible v values
4	0.574	2.48324	1.6752	[2.9832, 3.6675]
4.1	0.5799	2.4488	1.72484	[2.94881, 3.66777]
4.2	0.585	2.4159	1.77448	[2.91, 3.66889]

Based on the calculations in Table 3, we choose $v = 3.5$ to evaluate whether large deviations from the proposed equilibria are profitable.[9]

The entries of Table 4 illustrate that any deviation either to higher or lower values of recognition yield lower profits to the deviating firm. Deviating to a higher value $R_D > R^*$ sustains price competition between the firms (from (12) when $R_D > R^*$ price competition intensifies). In contrast, for the example considered, lower values of $R_D < R^*$ result in an increase of prices that transforms the market of each firm to a local monopoly,[10] even vis-à-vis the segment of consumers who are fully informed (the firms charge the price $v - t = 2.5$ in this region). Because unilateral deviation to either higher or lower recognition levels is unprofitable, the local symmetric equilibrium is, indeed, global.

It is noteworthy that in the example we consider a symmetric equilibrium exists. In fact, this symmetric equilibrium may even yield the maximum level of recognition of one (i.e., $R = 1$) if we assume also that $TC(R, A) \equiv 0$. That result contradicts the one obtained in Chen and Iyer (2002), who found that when the cost of addressability (recognition in our model) is zero, only asymmetric equilibria can arise.[11] The primary reason for the difference in results stems from the assumption in Chen and Iyer that firms can practice price discrimination. Specifically, each firm can recognize the exact location of each consumer on the line and charge him a different price contingent on this location. The ability to price discriminate yields fierce price competition and zero profits when each firm obtains perfect addressability. To alleviate this competition, firms end up choosing different levels of addressability at the equilibrium when the cost of addressability is zero. In contrast, in our model, where targeting is for the purpose of delivering TV commercials to consumers interested in the general product class, firms cannot practice price discrimination, because all consumers end up facing the same prices when shopping for the products. As a result, even when firms choose the maximum level of recognition (i.e., $R = 1$), their profits are not necessarily driven down to zero.

Table 4 Expected Profits from Unilateral Deviation ($R_j = R_D$ and $R_i = R^*$)

	R_D	$a = 4.0$	$a = 4.1$	$a = 4.2$
Deviation leads to local monopoly	0.15	0.342206	0.3534	0.364642
	0.25	0.609457	0.63666	0.650564
	0.35	0.907999	0.939373	0.970747
	0.45	1.2378	1.28152	1.32519
	0.55	1.59896	1.65645	1.7139
Deviation sustains price competition	0.60	1.67504	1.72474	1.77444
	0.65	1.67341	1.72339	1.77332
	0.75	1.66273	1.71358	1.76434
	0.85	1.6392	1.69123	1.74314
	0.95	1.60057	1.65401	1.70728

In Proposition 2, we describe how changes in the values of the parameters affect the levels of recognition and accuracy selected at the symmetric equilibrium.

Proposition 2. (i) *An increase in the cost parameter c reduces the level of recognition and increases the level of accuracy that is selected by the advertisers at the symmetric equilibrium. Specifically,*

$$\frac{\partial R}{\partial c} < 0 \quad \textit{and} \quad \frac{\partial A}{\partial c} > 0.$$

(ii) *An increase in the degree of differentiation parameter t or the size of Segment* 1 *population increases both recognition and accuracy. Hence,*

$$\frac{\partial R}{\partial t} > 0, \quad \frac{\partial A}{\partial t} > 0, \quad \frac{\partial R}{\partial a} > 0, \quad \textit{and} \quad \frac{\partial A}{\partial a} > 0.$$

According to Proposition 2, when the content distributor charges a higher price for viewed commercials, competing advertisers choose to increase accuracy and reduce recognition. In essence, the marginal return for improved accuracy is higher when c goes up because the cost of mistakenly delivering ads to the wrong population (viewers in Segment 0) is higher. However, the marginal return to improved recognition declines because the profitability of each advertiser diminishes when its payments to the content distributor are higher. Part (ii) of Proposition 2 asserts that when products are more highly differentiated or when the target population is larger, each advertiser finds it optimal to improve the quality of the targeting technology along both dimensions. Because the product market is more profitable when it is less competitive (larger t) or bigger (larger a), the marginal return to any kind of investment, including investment in improved targeting, is higher.

[9] Note that $p^* < v - t = 2.5$, thus implying an interior solution to the pricing subgame.

[10] Note that the price chosen at the symmetric equilibrium is very close to the monopoly price of 2.5. For $a = 4$, for instance, $p^* = 2.48324$. The deviation to $R_D < R^*$ yields, therefore, the local monopoly outcome.

[11] In the example considered in the appendix, we demonstrate the nonexistence of asymmetric equilibria of the kind derived in this earlier paper (i.e., $R_j < R_i = 1$). We are unable, however, to prove the nonexistence of other types of asymmetric equilibria.

It is noteworthy that the comparative statics with respect to the remaining parameters of the model may be ambiguous. In particular, a change in the coefficient γ_1 that determines the sensitivity of viewers in Segment 1 to improved targeting by Producers 1 and 2, has ambiguous implications on the level of recognition selected at the symmetric equilibrium. This ambiguity is implied, once again, by the two counteracting forces discussed earlier. Higher levels of recognition yield both larger de facto viewership and intensified price competition. Changes in γ_1 affect the magnitude of the two forces in opposite directions, thus making it impossible to determine which of the two dominates.

4. The Pricing Decision of the Content Distributor

In the previous sections, we assumed that the amount the advertisers pay the content distributor is exogenously determined at the level c. In this section, we relax that assumption, and allow the content distributor to choose the payment optimally. We introduce an initial stage, prior to the two in our earlier formulation. At the initial stage, the content distributor chooses the fee charged to advertisers for each commercial that is watched by a viewer.[12] We assume that the choice may be industry specific, so that the content distributor can charge different rates to advertisers in different industries. In deriving the optimal fee to be charged to Advertisers 1 and 2, the content distributor can ignore proceeds from advertisers in other industries. To simplify matters, we assume that the content distributor incurs only fixed costs that are independent of the eventual accuracy and recognition levels selected by the advertisers. In essence, the company delivers a certain quality of targeting viewers by using data mining techniques to analyze historical viewing patterns of individuals in the population. It is only the advertisers themselves, and not the content distributor, that can further improve on the quality of the targeting engine by conducting, for instance, direct surveys of the population. As a result, the development costs incurred by the content distributor are fixed in nature,[13] and are determined independently of the improvement efforts exerted by the advertisers to achieve certain levels of accuracy and recognition A and R.

[12] If the content distributor signs long-term contracts with the advertisers, two-part tariff schemes may also be feasible. We consider those schemes in the appendix.

[13] This assumption can be easily extended to allow for the possibility that the costs of the content distributor depend on the levels of A and R. This extension does not change any of our qualitative results.

After the content distributor chooses its charge c in the first stage, the game proceeds as in our initial formulation. Each advertiser chooses the nature of its efforts to improve accuracy and recognition and, subsequently, the market price of its product. The symmetric characterization of the last two stages for Advertisers 1 and 2 is provided in Proposition 2. That characterization depends, in particular, on the payment level c selected by the content distributor in the first stage.

In view of the above simplifying assumptions, the unit charge set by the content distributor for advertisers in a certain industry maximizes its proceeds from those advertisers. At the symmetric equilibrium, the expected proceeds from Advertisers 1 and 2 amount to

$$2c[\tilde{a}(1-\alpha)+(\widetilde{1-a})\beta],$$

where the error rates α and β translate to the accuracy and recognition levels expressed by Equations (1) and (2). Using those expressions, we can formulate the maximization problem of the content distributor as follows:

$$\max_c TR = 2c\tilde{a}\frac{R(c)}{A(c)}, \tag{16}$$

where $\tilde{a}$ is given in (14) and the dependence of R and A upon c is characterized in Proposition 2.

Note that the expected proceeds of the content distributor increase when advertisers choose to improve recognition and reduce accuracy at the symmetric equilibrium. However, from Proposition 2, obtaining higher levels of recognition and lower levels of accuracy require that the content distributor reduce the fee it charges to advertisers (because $\partial R/\partial c < 0$ and $\partial A/\partial c > 0$). In choosing c, the company should trade off the two counteracting effects of a change in this variable on the expected proceeds from Advertisers 1 and 2. Optimizing (16) with respect to c yields the following first-order condition:

$$\frac{\partial TR}{\partial c} = \frac{2\tilde{a}R(c)}{A(c)}[1+\eta_c^R-\eta_c^A] + 4c\frac{R(c)}{A(c)}\frac{a\gamma_1}{2}R'(c) = 0, \tag{17}$$

where η_c^R and η_c^A are the elasticities of the recognition and accuracy levels with respect to c. To guarantee the existence of an interior equilibrium, we must have that $\partial^2 TR/\partial c^2 < 0$ (see the online appendix for the derivation of this condition).

It is interesting to compare condition (17) with the condition usually derived when a decision maker maximizes revenues. For instance, a monopolistic producer that wishes to maximize revenues sets its product price at the point on the demand function where price elasticity is unitary. In our context, the demand facing the monopolistic content distributor is expressed by the levels of accuracy and recognition selected by the advertisers as a function of the payment c that the content distributor charges. Hence, to obtain the unitary elasticity rule in our context would imply that $(\eta_c^R - \eta_c^A) = -1$. However, if $1 + \eta_c^R - \eta_c^A = 0$ in (17), then $\partial TR/\partial c < 0$ because $R'(c) < 0$ from Proposition 2. Thus, to reach the optimum, the content distributor should reduce its fee below the point at which the combined elasticities are unitary. This deviation leads to higher recognition and lower accuracy levels than those implied by following the unitary elasticity rule. Such a deviation leads to a higher de facto size of Segment 1 viewers who choose to watch commercials. When Advertisers 1 and 2 increase recognition, viewers in Segment 1 are less likely to zap through commercials, implying that $\tilde{a}$ in (14) is larger. The second term in (17) measures the extra effect on $\tilde{a}$, and its implication on the rate the content distributor charges advertisers for viewed commercials.

In the online appendix, we also consider the possibility that the content distributor can utilize nonlinear pricing schemes, such as two-part tariff schedules. We show that such pricing results in higher levels of recognition and lower levels of accuracy than those implemented with linear pricing. Given that the distributor trades off a higher fixed charge against a lower variable fee per viewed commercial, advertisers respond to the lower variable fee by improving recognition and reducing accuracy. Moreover, circumstances may arise under which the distributor eliminates completely the variable charges, and relies only on the fixed payments as its source of revenues. When variable charges are eliminated, advertisers choose zero levels of accuracy at the equilibrium. In this latter case, the distributor can actually implement the "first-best" outcome that maximizes channel profits. It is noteworthy that two-part tariff pricing schemes are not always feasible. Such pricing can be implemented only when the advertisers sign long-term agreements with the content distributor.

5. Concluding Remarks and Possible Extensions

We developed a simple model, aimed at understanding how digital personal video recording technology is likely to affect advertising decisions of product producers. We demonstrated that the extent of resources an advertiser allocates to improving the quality of the targeting technology depends on the size of the viewing audience, the propensity of viewers to zap through commercials, and the overall cost of airing commercials. We introduced two different measures for the quality of the targeting technology, recognition and accuracy, and suggested that by conducting direct surveys of the population of viewers an advertiser can improve on the quality of the targeting engine offered by the content distributor. Moreover, the advertiser can gain full control over the extent of improvement along the two dimensions of quality by choosing an appropriate sample of the population to be surveyed. The initial predictions of the content distributor, as obtained from its profiling engine, can guide the advertiser in determining the identity of the viewers to be surveyed. For instance, surveying viewers who are predicted to belong to the target population increases accuracy, but may reduce the level of recognition in targeting customers.

Our results have important managerial implications both for advertisers and the content distributor who utilizes the PVR to deliver targeted ads to viewers. For advertisers, we found that the relative resources allocated to improving recognition versus accuracy should depend on the competitive environment in which they operate. An advertiser confronted by a competitor that chooses to allocate more resources to improving recognition should respond by cutting its own recognition level and improving accuracy. We also show that advertisers may have different preferences about the appropriate type of investment in improving the quality of the targeting technology than the cable content distributor. Those conflicting preferences are especially pronounced if the content distributor can only utilize linear pricing schemes. With linear pricing, the content distributor always benefits when competing advertisers choose to improve the level of recognition of their targeting quality. In contrast, advertisers may be reluctant to improve the extent of recognition in order to prevent intensified price competition in product markets.

Our main focus in this paper has been on the derivation of symmetric equilibria. While we address (in the online appendix) the possible existence of asymmetric equilibria, we do it only in the context of a specific example. Moreover, even in this case, we consider only the type of asymmetry that was also discussed in the earlier literature ($R_j < R_i = 1$). A most welcome extension of our analysis would explore whether the nonexistence of asymmetric equilibria that we obtain is valid when other types of asymmetric equilibria are considered (i.e., $R_j < R_i < 1$).

Our model can be extended in several ways. We assume that only the advertisers can choose the extent of investment in improving the quality of the targeting technology. An interesting extension is to consider the possibility that the content distributor can also endogenously decide on the extent of resources to allocate to improving its profiling engine. If both the content distributor and the advertisers can improve quality, it is reasonable that they will jointly make the quality decision via negotiations. The outcome of such negotiations is likely to yield improved channel coordination in the choice of investment level to enhance the quality of targeting.

Another interesting extension is to consider the possibility that the targeting technology can provide some information about the relative preferences of viewers among competing brands of a certain product class. As demonstrated in Gal-Or and Gal-Or (2005) and in Iyer et al. (2003), such improved targeting is unambiguously beneficial in alleviating the extent of price competition among competing advertisers.

An online companion to this paper is available on the *Management Science* website (http://mansci.pubs.informs.org/ecompanion.html).

Acknowledgment

The authors thank two anonymous referees and an associate editor for their helpful suggestions.

References

Assael, H., D. F. Poltrack. 1991. Using single-source data to select television programs based on purchasing behavior. *J. Advertising Res.* **31**(6) 9–17.

Blattberg, R. C., J. Deighton. 1991. Interactive marketing: Exploiting the age of addressability. *Sloan Management Rev.* (Fall) 5–14.

Bult, J. R., T. Wansbeek. 1995. Optimal selection for direct mail. *Marketing Sci.* **14**(4) 378–394.

Chen, Y., G. Iyer. 2002. Research note: Consumer addressability and customized pricing. *Marketing Sci.* **21**(2) 197–208.

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

Chunovic, L. 2001. Change ahead for channels. *Electronic Media* **20**(31) 1.

d'Aspremont, C., J. J. Gabszewicz, J. F. Thisse. 1979. On Hotelling's "Stability in competition." *Econometrica* **47**(5) 1145–1150.

Gal-Or, E. 1982. Hotelling's spatial competition as a model of sales. *Econom. Lett.* **9**(1) 1–6.

Gal-Or, E., M. Gal-Or. 2005. Customized advertising via a common media distributor. *Marketing Sci.* **24**(2) 241–253.

Hand, D. J. 1997. *Construction and Assessment of Classification Rules*. John Wiley & Sons, Chichester, UK.

Iyer, G., D. Soberman, J. M. Villas-Boas. 2003. The targeting of advertising. Working paper, University of California, Berkeley, CA.

Rose, F. 2003. The fast-forward, on-demand, network-smashing future of television. *Wired* (11).

Rossiter, J., L. Percy. 1997. *Advertising and Promotion Management*. McGraw Hill, New York.

Siddarth, S., A. Chattopadhyay. 1998. To zap or not to zap: A study of the determinants of channel switching during commercials. *Marketing Sci.* **17**(2) 124–138.

Spangler, W. E., M. Gal-Or, J. H. May. 2003. Using data mining to profile television viewers in the digital TV era. *Comm. ACM* **46**(4) 66–72.

Vives, Xavier. 1999. *Oligopoly Pricing—Old Ideas and New Tools*. MIT Press, Cambridge, MA, and London, England.

Wooley, S. 2003. Zapped to death. *Forbes* **172**(6) 76–82.

Zeisser, M. P. 2002. Living in a post-TiVo world. *McKinsey Quart.* Special to *CNET News.com*, http://news.com.com/2009-1069_3-976102.html.