Marketing Science
Vol. 29, No. 5, September–October 2010, pp. 783–804
ISSN 0732-2399 | EISSN 1526-548X | 10 | 2905 | 0783

DOI 10.1287/mksc.1100.0567

Moment-to-Moment Optimal Branding in TV Commercials: Preventing Avoidance by Pulsing

Thales S. Teixeira
Marketing Unit, Harvard Business School, Boston, Massachusetts 02163, tteixeira@hbs.edu

Michel Wedel
Smith School of Business, University of Maryland, College Park, Maryland 20742, mwedel@rhsmith.umd.edu

Rik Pieters
Department of Marketing, Tilburg University, 5000 LE Tilburg, The Netherlands, f.g.m.pieters@uvt.nl

We develop a conceptual framework about the impact that branding activity (the audiovisual representation of brands) and consumers' focused versus dispersed attention have on consumer moment-to-moment avoidance decisions during television advertising. We formalize this framework in a dynamic probit model and estimate it with Markov chain Monte Carlo methods. Data on avoidance through zapping, along with eye tracking on 31 commercials for nearly 2,000 participants, are used to calibrate the model. New, simple metrics of attention dispersion are shown to strongly predict avoidance. Independent of this, central on-screen brand positions, but not brand size, further promote commercial avoidance. Based on the model estimation, we optimize the branding activity that is under marketing control for ads in the sample to reduce commercial avoidance. This reveals that brand pulsing—while keeping total brand exposure constant—decreases commercial avoidance significantly. Both numerical simulations and a controlled experiment using regular and edited commercials, respectively, provide evidence of the benefits of brand pulsing to ward off commercial avoidance. Implications for advertising management and theory are addressed.

Key words: commercial avoidance; branding; attention; dynamic probit model; optimization; pulsing
History: Received: February 12, 2008; accepted: February 7, 2010; processed by Gary Russell. Published online in *Articles in Advance* May 27, 2010.

Introduction

Effective television advertising contributes to sales and long-term brand equity by building and sustaining brand awareness, associations, and attitudes. However, the effectiveness of television advertising may be slipping as a result of consumers' zapping of commercials. Commercial avoidance is facilitated by remote controls and by digital video recorders (DVRs) that permit consumers to record and replay TV content without having to see all or parts of commercial breaks. Early reports already indicated that during television commercials, eyes-on-screen, a metric of commercial contact, declined by 47%, with only 7% of the consumers giving ads total attention and 53% reporting divided attention (Krugman et al. 1995). Currently, about 17% of U.S. households are estimated to have DVRs (Steinberg and Hampp 2007) and around 87% skip past ads frequently (Grover and Fine 2006), and these numbers are growing. In addition, the networks have been imposing hefty price increases for ads by raising their per-viewer rates 110% in the past 10 years, despite declines in primetime audiences of up to 30% (Woolley 2003). Jointly, this leads to inefficiencies in marketing expenditures, increasing costs per viewer, and potential erosions of brand equity. It urges brand and advertising managers to understand the determinants of commercial avoidance and how to best retain consumers' attention from moment to moment during television commercials, in order to optimize brand (the audiovisual representation of it) exposure. This is the focus of the current study.

Specifically, the present research examines the influence that branding in television advertising and consumers' attention have on commercial avoidance. It makes three contributions. First, it provides a conceptual framework for understanding the impact that patterns of branding activity have on their avoidance decisions from moment to moment during television advertising. It formalizes this in a dynamic probit model, which is estimated with Markov chain Monte Carlo (MCMC) methods. Data on commercial avoidance, along with eye tracking on 31 commercials for nearly 2,000 participants, are used to calibrate the model. Second, we propose new, simple metrics of consumers' attention dispersion based on eye-tracking data and show that these systematically predict commercial avoidance from moment to moment. Third, based on the model estimations, we optimize branding activity for the sample of ads in question to reduce commercial avoidance. This demonstrates the significant reductions in commercial avoidance that can be attained by changing the pattern of branding activity by using brand pulsing strategies consisting of repeated brief brand insertions in the ad. Fourth, a controlled lab experiment in which commercials are edited based on the recommendations that follow from the model estimations provides further evidence for the benefits of brand pulsing to ward off commercial avoidance.

Branding and Attention Effects

Branding in Commercials

Branding activity is the way in which brand identity symbols (name, logo, typeface, trademark, or packshot) are present at each moment and across time in the commercial. This activity determines the prominence or conspicuity of the brand in commercials—that is, the extent to which it stands out from other objects and endures in the ad scenes, based on general rules of perception (Palmer 1999). At each moment during the commercial, the brand is more prominent to the extent that it appears larger (versus smaller), more central (versus peripheral), and more separated from its background (versus embedded) visually (Janiszewski 1998, Wedel and Pieters 2000) and is simultaneously supported by audio (Bryce and Yalch 1993). Brand prominence is enduring to the extent that the brand appears more (versus less) frequently and for longer (versus shorter) time periods during the commercial.

For consumers, such activity entails important information because the brand helps to comprehend ads and learn from them. Once the brand is identified, consumers can call upon their own personal experiences and memories to establish a context for the ad and its message. For management, branding in commercials is an important decision variable, because of advertising's intended contribution to sales and brand equity. Branding activity is also a source of debate in advertising theory and between marketers and ad agencies, who are trying to balance sales, creativity, and other objectives. Some recommend small, nonintrusive branding (Aitchinson 1999), and others recommend large, intrusive branding (Book and Schick 1997). Likewise, there are recommendations to place the brand as early as possible in commercials (Baker et al. 2004, Stewart and Furse 1986), later in the commercials (Fazio et al. 1992), or early and late (Stewart and Koslow 1989).

There is some evidence that under conditions of forced exposure (when consumers cannot avoid watching the commercials), both early (Baker et al. 2004) and late (Fazio et al. 1992), more frequent and longer branding (Stewart and Furse 1986) can improve comprehension, memory, and persuasion. Also, under forced exposure, video-transmitted content is better learned than the same content in audio, resulting in an eight-to-one advantage in memory tests after a single exposure (Bryce and Yalch 1993). However, consumers in practice have increasing control over commercial exposure, which is important. When consumers stop watching commercials before they naturally end, later branding activity in the commercial cannot have the beneficial effects that have been reported for forced exposure conditions. Moreover, what if one of the main objectives for advertisers investing heavily in commercials—namely, to expose the brand—is actually related to the consumer's decision to continue or stop watching the commercial? We are not aware of research that has examined the influence of the moment-to-moment prominence of brands, such as a result of their size and centrality, on avoidance of television commercials. If and how branding activity in commercials impacts consumers' moment-to-moment avoidance decisions remains as yet largely unknown; our purpose is to shed further light on this issue.

Television commercials are narratives aimed to convey the brand message and at the same time entertain and retain consumers. However, intense branding activity decreases the "soft-sell" narrative character of commercials and increases the "hard-sell" character, and people generally resist the forceful persuasion that comes with the latter (Aaker and Bruzzone 1985, Greyser 1973). Also, because brands convey information, their prominent presence in television commercials increases the likelihood of commercial avoidance as a result of information overload (Woltman Elpers et al. 2003). Therefore, we predict that higher intensities of branding activity increase the likelihood of avoidance at each moment during the commercial and establish the contribution that the momentary (size, separation, and centrality) and dynamic (frequency and duration) characteristics of branding activity have on this likelihood. Brands carry associations idiosyncratic to each consumer, and these certainly should play a role in their avoidance decisions. This research, however, focuses on systematic, common effects across all consumer–brand dyads and commercial contexts that may contribute to self-controlled termination of exposure. In determining these branding effects, it is important to control for factors that may independently affect moment-to-moment commercial avoidance decisions.

Attention Concentration by Commercials

Similar to the visual arts, advertising tries to focus and direct viewers' attention. It aims to point attention to certain parts of the depicted scene and direct it across scenes in an orchestrated fashion to let the intended narrative unfold. We propose that to the extent that commercials are able to concentrate consumers' attention, they are better able to retain consumers behaviorally as well, thus preventing commercial avoidance. This is consistent with art theory's (Arnheim 1988) emphasis on "centers of gravity," which concentrate the viewer's eyes on the essentials in paintings, statues, or buildings, and with speculations in advertising (Heeter and Greenberg 1985, Perse 1998) that viewers with less focused attention do not actively follow the ad script and may decide to zap away. In the words of Gustafson and Siddarth (2007, p. 587), "...a reasonable hypothesis is that all zaps are associated with looks that have ended, although all completed looks will not end in a zap."

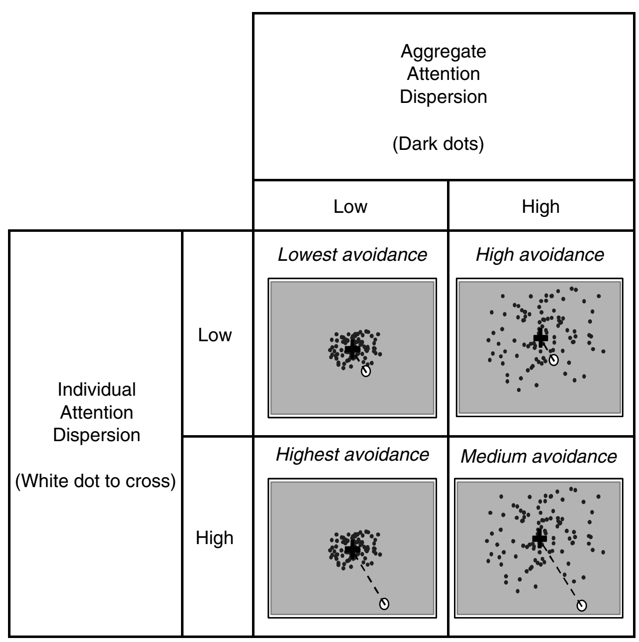
In aesthetic psychology, Berlyne (1971) distinguishes two types of visual attention that an individual viewer can express during perception of artful stimuli, termed specific exploration and diversive exploration, and speculates that each would be reflected in distinct patterns of eye fixations (moments that the eye is relatively still and focused on a specific location in space). Specific exploration would lead to concentrated eye fixations on precise locations of the visual scene to seek out detailed information. Diversive exploration would lead to dispersed eye fixations across larger regions of the scene to search for new stimulation or grasp the gist. Then, to the extent that commercials are successful in focusing and conducting attention, eye fixations of consumers at each moment across the duration of the commercial will be more concentrated at specific locations. Such a dense pattern of eye fixations would reflect desirable bottom-up control of consumers' focal attention by characteristics of the commercial. We predict that under such conditions of concentrated attention—with all consumers held together by the commercial—the likelihood of commercial avoidance will be low.

Conversely, dispersed patterns of eye fixations reflect a lack of bottom-up control as a result of the overriding effects of consumers' goals or tendencies to freely explore the scene. For instance, in an early eye-tracking study with a single participant viewing a painting, Yarbus (1967) observed that specific task instructions led to widely different locations on which the participant focused the eye and that eye fixations were most dispersed under a free viewing instruction. Working with print advertising, Pieters and Wedel (2007) find that goals, as specific instances of top-down factors (residing in the consumer), induced distinct spatial attention patterns. Thus, the more that idiosyncratic personal factors dominate attention, the more dispersed the aggregate eye fixations across commercials will be. We predict that under such conditions of dominant top-down and limited bottom-up control of attention by the commercial, as expressed in dispersed eye-fixation patterns of consumers, the likelihood of commercial avoidance will be high.

Not only should aggregate patterns of attention dispersion across consumers be predictive of commercial avoidance, but patterns of individual consumers should do so as well. That is, when television commercials successfully concentrate focal attention of most consumers as a group but fail to do so for a specific consumer—who wanders off from the virtual flock—the likelihood that this consumer avoids the commercial will be high.

Figure 1 summarizes our predictions about the influence of aggregate and individual dispersion versus concentration of focal attention on commercial avoidance. It indicates that the less concentrated (i.e., more dispersed) the aggregate focal attention of consumers is, the higher the likelihood of commercial avoidance is expected to be. Also, the less concentrated (i.e., more dispersed) the focal attention of an individual consumer relative to the other consumers is, the higher the likelihood of avoidance by this consumer is expected to be. We predict an interaction effect between aggregate and individual attention dispersion, such that avoidance is expected to be highest when a consumer's attention is dispersed from all other consumers, who among themselves have a concentrated pattern of focal attention (lower left cell of Figure 1). Then, the commercial is successful in concentrating the attention of most consumers but not the single individual, who wanders off and leaves. These measures of attention dispersion capture the extent to which the creative content of commercials is successful in focusing and retaining consumers.

Figure 1 **Attention Concentration and Commercial Avoidance**



In establishing the net contribution of branding activity on commercial avoidance, we therefore account for these consumers' attention dispersion patterns. If we were to find that attention dispersion predicts commercial avoidance independent of branding activity, this would be strong evidence for the central function that attention guidance by the creative content of commercials plays in ad effectiveness. To assess these branding and attention effects appropriately, other ad, brand, and person characteristics need to be controlled for. We focus on objective ad characteristics that may covary moment to moment with branding, attention, and baseline zapping levels.

Controlling for Ad, Brand, and Person Effects

Film, television, and advertising producers tailor the visual complexity of commercials and other video stimuli to engage viewers and prevent them from channel switching (Lang et al. 2005). The overall visual complexity of commercials at any point in time is jointly determined by the amount of visual material in separate scenes (momentary) and by the pacing of scenes across the commercial (dynamic) (Germeys and d'Ydewalle 2007). Visual complexity refers to all nonrepresentational perceptual material, such as different colors, lines, and luminance contrasts, in the commercial with more material increasing the visual complexity (Donderi 2006). Pacing is the speed at which different scenes are presented in dynamic stimuli (Lang 2000). It is reflected in discontinuities in the video stream and accomplished by cuts and edits (Bolls et al. 2003, Germeys and d'Ydewalle 2007, Lang 2000), with more cuts and edits increasing the pace.

Visual complexity of images can influence ease of perception, memory, attitudes (Bolls et al. 2003, Germeys and d'Ydewalle 2007, Lang et al. 2005, Pavelchak et al. 1991, d'Ydewalle et al. 1998), and perhaps avoidance decisions. That is, at low levels of visual complexity, commercials may be insufficiently engaging and challenging, whereas at high levels they may be too arousing and demanding. Therefore, we expect a Yerkes and Dodson (1908) type of U-shaped relationship between the amount of visual complexity in scenes and the likelihood of commercial avoidance at each moment during the commercial, with the lowest avoidance likelihood at intermediate complexity levels and the highest levels at the low and high ends of the complexity spectrum (in fact, the original curve is an inverted U with performance being highest at intermediate levels, which translates into avoidance being lowest at those levels here). Berlyne (1971) observed a similar pattern in research on the appreciation of paintings varying in levels of visual complexity, which has been replicated for other stationary stimuli as well (Donderi 2006). We extend this by studying avoidance decisions for dynamic visual stimuli.

In addition, in our empirical study, product category (hedonic versus utilitarian) and brand familiarity (low versus high) are controlled for (Pieters and Wedel 2004). Finally, two demographic factors, gender and age, are controlled for, based on findings that males compared to females and younger compared to older consumers generally zap more (Cronin 1995, Heeter and Greenberg 1985).

In sum, we predict that, while controlling for ad, brand, and person characteristics, branding activity in commercials and attention dispersion of consumers jointly influence the moment-to-moment commercial avoidance decisions of consumers. Before specifying the analytic model that allows us to examine specific branding effects in detail, the data on which it is calibrated are described.

Data

Stimuli and Participants

The data for this research were collected by the marketing research company Verify International (Rotterdam, The Netherlands). A sample of 31 regular, newly aired commercials of 25, 30, and 35 seconds were selected. They featured known (e.g., Citroën, T-Mobile) and unknown (e.g., Radio 538, KWF) companies and national (e.g., Albert Heijn, UNOX) and international (e.g., MasterCard, Kodak) brands from a variety of different product categories (e.g., food, durables, public services, electronics, telecom, clothing), with utilitarian (e.g., checking account) and hedonic (e.g., chocolate) purchase motivations. By selecting newly aired commercials, the chances that participants had been exposed to the commercials before are minimized.

Participants were a random sample of 1,998 regular television viewers (aged 20 to 62, 48% male) and consumers of the advertised products, who were paid for participation. Their demographics matched those of the target population. Although all participants watched a long reel of commercials, the data available to us had a maximum of four television commercials per person. On average, each commercial was watched by 111 participants.

Data Collection

Data collection took place at the facilities of the company. Upon entering, each participant was led to a nondistracting room and seated in a comfortable chair at a distance of approximately 55 cm from a 21-inch LCD monitor with a 1280×1024 pixel resolution. The instruction on the screen asked the participant to watch the commercials and to stop watching any commercial at any time by zapping. Immediately after zapping a commercial or after it ended without the participant zapping, the next commercial in the sequence appeared. The order of the commercials was randomized across participants to control for serial-position effects. Filler ads were shown between the target ads but no program content was shown because the study focuses on commercial avoidance, not on channel switching, surfing, or grazing (Cronin 1995, Tse and Lee 2001). This experimental setup mimics the common situation of "roadblocking," in which blocks of commercials are aired at the same time on different channels so that consumers zapping away from one commercial zap into another one. These avoidance rates are higher than in similar reports (Krugman et al. 1995, Siddarth and Chattopadhyay 1998) but lower than other more recent ones (Tse and Lee 2001) and are reported on current DVR usage patterns (Wilbur 2008).

Infrared corneal-reflection eye-tracking methodology was used to record the focal positions of the viewer's right eye, in an X and Y coordinate system (Duchowski 2003). The method is nonobtrusive to the participant, allowing for head movements within normal boundaries (about $30 \times 30 \times 30$ cm) while facing the television screen. Spatial precision of data collection was 0.5° of visual angle at a sampling rate of 20 ms (50 Hz). To match them to the frequency of standard video frame presentation, the data were combined into 40-millisecond frames, which results in an average of 750 consecutive frames (moments) for every 30-second commercial.

Measures

Commercial Avoidance. The dependent variable consists of every recorded avoidance decision, when a participant chooses to stop watching a particular commercial by pushing the button ($1 =$ avoid, $0 =$ else). The dependent variable is a binary cross-sectional (consumers), repeated measures (ads) time series, because we have decisions to zap or not to zap for 31 distinct television commercials, each of a maximum of 750 ad frames, for a total of 1,998 consumers. That is, we have unbalanced panel data, truncated at each zapping incidence.

Branding Activity. Branding activity (name, logo, typeface, trademark, or packshot) was recorded semi-automatically by means of specialized video manipulation/editing software for each frame of a commercial. We identified the brand's (a) presence, (b) size, (c) position, (d) separation, and (e) mode per frame as stationary characteristics, and we identified its (f) cardinality and (g) duration across frames as dynamic characteristics, as defined next.

Presence indicates whether the brand is on screen (1) or not (0) during a particular frame. *Size* is the proportion of the screen, in square pixels, occupied by the smallest rectangle enveloping the brand at each frame and is zero when the brand is absent (Pieters and Wedel 2004). *Position* indicates whether the brand takes a central (1) or peripheral (0) position on the screen. For this, an imaginary rectangle with the same 4:5 aspect ratio as the 21-inch LCD monitor was defined such that the length of the longest dimension is equal to the viewing angle of the parafoveal field of the eye: 5° from a central axis, to the left and to the right (Duchowski 2003, Rayner 1998). The brand is central if the rectangle boundary to define brand size intersects with the parafoveal rectangle (above) in the center of the LCD screen, and it is peripheral otherwise. *Separation* indicates whether the brand is well separated from its background (1) or not (0)—for instance, because it is competing with other scene objects or occluded by them (Janiszewski 1998). *Mode* indicates whether the brand was additionally present (1) in audio mode or not (0) in a particular frame. *Cardinality* is a count of how many times the brand has appeared in video mode, in nonconsecutive blocks of frames, up to the time-point in question, from the first (1) to last (n) brand appearance. Finally, *Duration* indicates, in seconds, the total time that a brand has appeared in the current block (*Cardinality*) up to the time-point in question. Thus, each time a new brand appears in any visual form on screen, cardinality increases by one and the "duration counter" starts from zero again, increasing frame by frame until this brand exits the screen.

Control Variables. The level of visual complexity in consecutive frames of the commercial was assessed by the file size in kilobytes of the GIF-compressed image, as in recent, similar applications (Calvo and Lang 2004, Sprott et al. 2002). Compression algorithms, such as for the GIF, JPEG, and PDF formats, have been developed in computer vision research to enable different hardware and software to use the same data. To the extent that the visual images contain little visual detail, color, and contrast, and that they contain many redundancies, the algorithms cause larger compressions (Sprott et al. 2002). This makes file size a suitable general measure of the visual complexity of images. In support, research has found the file size of images such as charts, Web images, and photos to correlate highly and significantly (0.82) with human judgments of visual complexity (Calvo and Lang 2004, Donderi 2006). *Pacing* was measured by the presence of cuts and edits (1) versus absence (0) in each frame of the commercials using video editing software. Cuts are a result of changing camera positions between scenes and edits a result of changing camera positions within scenes, and both increase complexity because viewers need to integrate the visual information across discontinuities.[1] Cuts and edits have larger complexity

[1] Brand appearances (i.e., changes in cardinality) do not necessarily reflect a discontinuity or pace change. They will if the brand appears simultaneously with a change of a scene, but this need not be the case.

effects than more subtle production choices such as zooms and camera moves (Lang et al. 2000).

Information about the *Gender* (1 = male, 0 = female) and *Age* (years) was available from company records. *Brand familiarity* (familiar = 1, unfamiliar = 0) and *Product category* (utilitarian = 1, hedonic = 0) were coded by two independent judges (initial agreement was 96% for brand and 78% for product, with disagreements resolved by discussion).

Data Aggregation

Data processing and analysis is challenging; there are 750 frames for each of the 31 commercials for which eye-movement data are available for a total of 1,998 consumers. To strike a balance between keeping the analysis task manageable and retaining sufficient detail, we averaged the eye movement data to intervals of approximately 240 ms (4.17 Hz) for the 30-second ads, which is the shortest observed consecutive time a brand is on screen in the data. Participants may see the brand and react by zapping within one or two intervals (Calvo and Lang 2004, Mihaylova et al. 1999, Rayner 1998).[2] The interval is shorter than the typical interval between pacing events (Germeys and d'Ydewalle 2007). The aggregation led to a total of 125 frames. To equate 25- and 35-second ads with 30-second ads, we use similar procedures by lowering sampling rates to 5 and 3.57 Hz, respectively, with the differences being perceptually undistinguishable. Frame lengths are chosen to be uniform across all commercials; not doing so would make it hard to link the frame images to exact fixation points of the eye-tracking data.

Model

We assume that an individual's decision to continue watching a specific commercial at time point t or to avoid it is based on the (negative) utility derived up to that time point from the commercial:

$$U_{ict}^{\text{avoid}} = D_{ict}^{\text{avoid}} + \varepsilon_{ict}^{\text{avoid}}, \quad \text{with } \varepsilon_{ict}^{\text{avoid}} \sim N(0, 1), \tag{1}$$

where i is the individual, c is the commercial, and t is the time frame of the ad. The variance of the error term is fixed to one for identification because utility is defined up to a scale factor. Thus the probability that individual i avoids commercial c at time frame t, given parameters Θ_t, and $\Phi(\,)$, the cumulative Normal density function, is

$$P(y_{ict} = 1 \mid \Theta_t) = \Phi(D_{ict}),$$
$$\text{where } y_t = \begin{cases} 1 \rightarrow \text{avoid at frame } t, \\ 0 \rightarrow \text{watch at frame } t. \end{cases} \tag{2}$$

[2] In the section on robustness checks, we provide the results of an experiment used to support this claim and investigate the robustness of our results to this data aggregation.

Five terms make up the deterministic component of the utility (D_{ict}):

$$D_{ict}^{\text{avoid}} = \mu_i + \alpha_c + B_{ct} + \left(\gamma^1 \,\text{AAD}_{ct} + \gamma^2 \,\text{IAD}_{ict} + \gamma^3 \,\text{AAD}_{ct} \times \text{IAD}_{ict}\right) + \text{TVC}_{ct}. \tag{3}$$

The time-constant intercepts μ_i and α_c are estimated for each individual and commercial, respectively, and are a linear function of individual-specific demographics (age and gender) and brand familiarity and product category (utilitarian or hedonic), respectively. This specification is parsimonious, given the large number of individuals and commercials, and is similar to Gustafson and Siddarth (2007). Details are in the appendix.

The branding effects B_{ct} are commercial and time specific (to simplify notation, we suppress subscripts c in Equation (4)) and are specified as

$$\begin{aligned} B_t &= \theta_t^1\, \textit{Presence}_t + \theta_t^2\, \textit{Cardinality}_t + \theta_t^3\, \textit{Duration}_t \\ &\quad + \theta_t^4\, \textit{Size}_t + \theta^5\, \textit{Mode}_t + \theta^6\, \textit{Position}_t \\ &\quad + \theta^7\, \textit{Separation}_t, \quad \text{with} \\ \tilde{\theta}_t &= \Xi + G\tilde{\theta}_{t-1} + \omega_t \quad \text{and} \quad \tilde{\theta}_t = (\theta_t^1, \theta_t^2, \theta_t^3, \theta_t^4). \end{aligned} \tag{4}$$

Because branding activity may build up irritation over the exposure to the commercial if it becomes too intrusive (presence and size of brand) and enduring (cardinality and duration) (Aaker and Bruzzone 1985, Greyser 1973), the parameters capturing the effects of these branding variables are specified to be time dependent, $\tilde{\theta}_t$. We specify only the branding effects to vary over time to keep the model parsimonious and because we have no theory predicting that the effects of the other variables should be time varying. Factors that affect the dynamics of attention to commercials "…have generally been ignored by previous research on advertising, even though recent research has established that consumers' real-time response to a commercial vary significantly over the time of its airing" (Gustafson and Siddarth 2007, p. 605). We believe brands to be one of such factors. Consequently, time-varying parameters of brand presence allow the effect of a one-second of brand exposure in the beginning of an ad to be different from when the viewer has potentially seen more of the brands toward the end of the ad.

The fourth term (in parentheses) in Equation (3) reflects the attention dispersion of consumers in each time frame. We have the eye fixation (f_{ict}) for individual i and commercial c at time frame t in $x - y$ pixel coordinates. Extending ideas of Germeys and d'Ydewalle (2007), we propose the variance of f_{ict} as a measure of aggregate attention dispersion (AAD_{ct}) across consumers i for each commercial c at time frame t. Attention concentration is at a maximum when all eye fixations are on exactly the same screen pixel (AAD = 0) and decreases when eye fixations become more spatially dispersed. In addition, we propose the squared Euclidian distance between an individual's eye fixation and the centroid of eye fixations for all other consumers as a measure of individual attention dispersion (IAD_{ict}) for each consumer i, commercial c, and time frame t. The centroid indicates the average point of focus of all viewers and may be indicative of the desired location of attention. IAD_{ict} ranges from 0 to 2,686,976 ($1{,}280^2 + 1{,}024^2$, based on our screen resolution). Thus, we have

(Individual Attention Dispersion):

$$\mathrm{IAD}_{ict} = \left(f_{ict} - \frac{1}{N}\sum_{i=1}^{N} f_{ict}\right)'\left(f_{ict} - \frac{1}{N}\sum_{i=1}^{N} f_{ict}\right); \tag{5}$$

(Aggregate Attention Dispersion):

$$\mathrm{AAD}_{ct} = \frac{1}{N}\sum_{i=1}^{N} \mathrm{IAD}_{ict}.$$

The parameters γ^1, γ^2, and γ^3 (Equation (3)) capture the effects of these attention dispersion measures and their interaction. The final term in Equation (3), TVC_{ct}, captures the effect of the total visual complexity of commercial c at time frame t. The visual complexity effects are specified in Equation (6). For every time frame, we define visual complexity to be the sum of the consecutive image complexities ($\mathrm{IC}_{ct} + \mathrm{IC}_{ct-1}$) in the event of an edit or cut ($Pacing_{ct} = 1$) or the image complexity of the current frame otherwise ($Pacing_{ct} = 0$). This is in line with the viewer's effort when integrating images that are completely different, or very similar. The quadratic term of visual complexity allows for a U-shaped effect on avoidance likelihood. Finally, *PaceType* is a dummy variable indicating a cut (= 1) or edit (= 0).

$$\mathrm{TVC}_{ct} = \beta^0 PaceType + \beta^1(\mathrm{VC}_{ct}) + \beta^2(\mathrm{VC}_{ct})^2, \quad \text{with}$$
$$\mathrm{VC}_{ct} = \mathrm{IC}_{ct} + Pacing_{ct} \cdot \mathrm{IC}_{ct-1}. \tag{6}$$

To summarize, the model describes commercial avoidance as a utility-based decision that is made on a moment-to-moment basis. It specifies specific branding parameters to be time varying to allow for the evolution of their effects. It accounts for observable individual and commercial heterogeneity, partially by the eye-tracking data and by covariates, and for other unobserved sources of heterogeneity by assuming normal distributions of all parameters.

Estimation Procedure and Inferences

Dynamic linear models have been used in advertisement contexts that have similar dynamics (Bass et al. 2007, Naik et al. 1998). Here, we develop a dynamic probit model (Gamerman 1998, West and Harrison 1997) by rewriting Equations (1)–(6) in a state space formulation as in Equation (1):

$$f(Y_t) = F_t\Theta_t + \varepsilon_t,$$
$$\Theta_t = G\Theta_{t-1} + \omega_t, \tag{7}$$

where Y is the commercial avoidance indicator variable; f is the probit link function; $\Theta_t = \{\mu_i, \alpha_c, \tilde{\theta}_t, \theta^5, \theta^6, \theta^7, \beta^0, \beta^1, \beta^2, \gamma^1, \gamma^2, \gamma^3\}$, the vector of parameters previously defined; F_t is the vector of covariates, blocked by time-varying and invariant ones; G is the evolution matrix of the time-varying parameters; and ε and ω are independently distributed with contemporaneously independent time-varying error terms. We specify the evolution matrix, $G = I$, so that Θ_t follows a random walk, which strikes a balance between sequential independence and time invariance (Martin and Quinn 2002).

We use a MCMC Gibbs sampling in blocks given the hierarchical Bayes (HB) structure of the model (Billio et al. 2007, Gamerman 1998), using the forward filtering backward sampling algorithm (Carter and Kohn 1994, Frühwirth-Schnatter 1994). In essence, the estimation is done by drawing the latent values for utilities for all i, c, and t, drawing from a truncated Normal distribution, and then proceeding with sampling the remainder of the parameters using the draws of the latent utilities. The MCMC chains are run for 60,000 iterations on 1,998 viewers, 31 commercials, and a maximum of 125 time frames, totaling 293,000 observations. The posterior distributions of the parameters of 1,750 draws were extracted, thinning 1 in 5 draws after a burn-in period of 51,250. Starting values were obtained from the maximum likelihood parameter estimates from an ordinary probit model. Details of the estimation are provided in the appendix. Analysis of synthetic data with the MCMC algorithm shows good recovery of all true parameter values. Convergence of a Gibbs sampler was checked through visual inspection of likelihood and diagnostic plots for key model parameters.

Results

Sample Statistics and Model Comparisons

Table 1 provides sample statistics for the 17 independent variables. All independent variables were standardized before analysis to facilitate comparison of parameter estimates.[3] The condition number of the X-matrix was 3.02. With the exception of the visual complexity (VC and VC^2 were orthogonalized

[3] For each covariate, the mean was subtracted and was divided by the covariate's standard deviation. The interaction and squared terms were first computed and then standardized to facilitate comparison of effect sizes.

Table 1 **Summary of the Independent Variable**

Variable	Variation across units	Mean	Std. dev.	Minimum	Maximum
Branding activity					
Presence (present = 1)	Ad, time	22%	41.2%	0	1
Size (percentage of screen)	Ad, time	2.9%	8.8%	0.1%	61.5%
Position (central = 1)	Ad, time	13.9%	34.5%	0	1
Separation (separated = 1)	Ad, time	89.1%	31.2%	0	1
Mode (audio = 1)	Ad, time	3.2%	17.5%	0	1
Cardinality (1, 2, ...)	Ad, time	0.79	1.33	0	6
Duration (seconds)	Ad, time	1.89	3.62	0	30
Attention dispersion					
Aggregate dispersion (pixels^2)	Ad, time	104,212	486,780	2,434	7,311,820
Individual dispersion (pixels)[b]	Ad, time, indiv.	147	289	0	27,478
Aggregate × *Individual dispersion*	Ad, time, indiv.	32,256,972	1.2E+09	0	1.0E+11
Control variables					
Age (years)	Individual	38.3	10.9	20	62
Gender (male = 1)	Individual	48.3%	50.0%	0	1
Brand familiarity (familiar = 1)	Ad	89.8%	30.3%	0	1
Product category (utilitarian = 1)	Ad	60.0%	49.0%	0	1
PacingType[a] (cut = 1)	Ad, time	44.4%	49.7%	0	1
Visual complexity (kilobytes)	Ad, time	180	69	2	662
Visual complexity^2	Ad, time	37,156	32,628	4	438,244

[a]Conditional on a camera shot change.
[b]The square root of IAD was used in the model to reduce the skewness of this distance measure.

via Gramm–Schmidt), all other independent variables have a variance inflation factor (VIF) $\ll 10$, which indicates that collinearity is not a significant problem (Kutner et al. 2004). Note that the Euclidian, as opposed to the squared Euclidian (Equation (5)), distance was used for IAD to reduce skewness.

To examine the contribution of sets of explanatory variables, we first compared the full model to four nested models, using the log-marginal likelihood (LML). We estimated the LML using Chib's (1995) method, which requires running several additional chains after the original MCMC chain but is more appropriate than using the harmonic mean estimator. The results are in Table 2.

Model 1 is the benchmark containing only the demographics, brand familiarity, and product category type. Model 2 adds the visual complexity measures to model 1, and model 2 outperforms model 1 as shown by the higher LML. Model 3 includes the attention dispersion variables in addition to the variables in model 2 and outperforms the latter. In model 4 the branding activity variables are added to model 2, and model 4 outperforms model 2. Model 5 includes all variables, but no unobserved heterogeneity and no dynamic effects. Its LML is worse than that of all other models except for model 1. Model 6 includes all variables, as well as unobserved commercial and individual heterogeneity, but no dynamic effects. It performs better than models 1 and 5 but worse than all other models. Finally, the full model 7 clearly performed best among all models in terms of the LML. It predicts commercial avoidance with an average absolute error of only 6.5% across the 31 commercials.

The model comparisons reveal that all sets of variables as well as heterogeneity and dynamics contribute significantly to predicting commercial avoidance and that branding effects contribute significantly to

Table 2 **Model Comparisons**

	Model components included						
Model	Demographics, product–brand	Visual complexity	Attention dispersion	Branding activity	Heterogeneity	Dynamics	LML
1	Yes	No	No	No	Yes	No	−11,231
2	Yes	Yes	No	No	Yes	No	−10,021
3	Yes	Yes	Yes	No	Yes	No	−9,966
4	Yes	Yes	No	Yes	Yes	Yes	−9,690
5	Yes	Yes	Yes	Yes	No	No	−10,909
6	Yes	Yes	Yes	Yes	Yes	No	−10,237
7	Yes	Yes	Yes	Yes	Yes	Yes	−9,639

Table 3 Determinants of Commercial Avoidance

			Percentiles of the posterior distribution				
Parameter	Mean	SE	5%	10%	50%	90%	95%
Intercept ($t = 0$)	−3.641**	0.219	−4.004	−3.925	−3.622	−3.376	−3.301
Branding activity							
Presence ($t = 0$)	0.335**	0.099	0.174	0.212	0.332	0.465	0.507
Size ($t = 0$)	−0.001	0.115	−0.200	−0.147	0.001	0.143	0.189
Position (central = 1)	0.033**	0.009	0.016	0.019	0.033	0.046	0.050
Separation (separated = 1)	0.014*	0.011	−0.005	0.000	0.014	0.026	0.029
Mode (audio = 1)	−0.011*	0.010	−0.027	−0.023	−0.011	−0.000	0.003
Cardinality ($t = 0$)	0.014+	0.097	−0.144	−0.109	0.011	0.139	0.186
Duration ($t = 0$)	0.085+	0.096	−0.070	−0.035	0.082	0.207	0.249
Attention dispersion							
Aggregate dispersion	0.055**	0.021	0.013	0.027	0.057	0.078	0.085
Individual dispersion	0.199**	0.011	0.181	0.185	0.200	0.214	0.218
Aggregate × *Individual dispersion*	−1.249**	0.108	−1.377	−1.355	−1.284	−1.054	−1.038
Control variables							
Age (years)	−0.003	0.012	−0.023	−0.019	−0.004	0.013	0.017
Gender (male = 1)	0.020**	0.011	0.001	0.005	0.021	0.035	0.039
Brand familiarity ($f = 1$)	0.001	0.030	−0.047	−0.035	0.000	0.039	0.053
Product category ($u = 1$)	0.037	0.030	−0.012	−0.001	0.037	0.075	0.087
PacingType (cut = 1)	0.000	0.010	−0.018	−0.014	0.000	0.014	0.017
Visual complexity	−0.008	0.011	−0.027	−0.023	−0.008	0.005	0.009
Visual complexity^2	0.088**	0.033	0.033	0.046	0.089	0.128	0.140

*Indicates that 90% confidence interval does not contain zero; **indicates that 95% posterior confidence interval does not contain zero; +indicates that 90% posterior confidence interval does not contain zero for some time periods.

predicting commercial avoidance even when all other effects are accounted for (model 7 versus model 3). Likewise, the attention dispersion measures contribute significantly to predicting commercial avoidance, even when all other effects are accounted for (model 7 versus model 4). Finally, even when all explanatory variables are included, including dynamic effects contributes significantly to predicting commercial avoidance (model 7 versus model 6).[4]

Determinants of Commercial Avoidance

Table 3 provides the mean, standard error, and main percentiles of the posterior distributions from the MCMC draws for the full model 7. As benchmarks, the estimates of models 6 (the static HB probit) and 5 (the static probit) are also provided in Table 4. We will not discuss the parameter estimates of those models here, but it suffices to note that the estimated effects of several of the branding activity variables are different from the full model 7 (implications of the main differences will be pointed out below).

In support of our hypotheses, branding activity had significant effects on the moment-to-moment decision to continue or stop watching the commercial. Specifically, the presence of a brand, independent of the other branding variables, significantly increased the probability to stop watching the commercial (posterior mean estimate = 0.335). Also, when the brand appeared more central and well separated from the rest of the scene, and later and longer in the commercial (for some periods), the probability to stop watching the commercial increased as well. The size of the brand did not have an independent effect once the other branding and all other effects were accounted for. However, when brands were simultaneously present in audio mode, as opposed to just video mode or no brand apperance, probabilities to avoid the commercial decreased marginally.

In support of our predictions, attention dispersion strongly predicted the probability to stop viewing the commercials, over and above the effects of all other variables. Specifically, at each moment, a commercial's failure to concentrate consumers' attention simultaneously increased their probability to stop viewing the commercial. Also, consumers who failed to look where all other consumers concentrate attention had a higher probability to stop viewing the commercial. The probability to stop viewing was lowest when consumers on the aggregate, and each of them individually, concentrated their attention on the same locations in the commercial. This reveals the importance that the attention concentration power of commercials has frame for frame in retaining consumers. As predicted for the interaction of IAD and AAD, in particular when IAD is high and AAD is low, commercial avoidance is most likely to occur for a particular consumer, as depicted in Figure 1.

[4] We also included lag 1 and 2 measures of branding variables, but this did not improve model fit, as we discuss in the section on robustness checks.

Table 4 Model Comparisons

	Dynamic HB probit (7)		Static HB probit (6)		Static probit (5)	
LML	−9,639		−10,237		−10,909	
Parameter	Mean	SE	Mean	SE	Mean	SE
Intercept[a]	−2.602**	0.082	−2.577**	0.022	−2.548**	0.010
Presence[a]	0.092**	0.045	0.096**	0.011	0.091**	0.011
Cardinality[a]	−0.013+	0.041	0.024**	0.012	−0.003	0.010
Duration[a]	0.040+	0.039	0.024**	0.013	0.015	0.009
Size[a]	−0.030	0.051	0.005	0.012	−0.007	0.010
Mode (plus audio = 1)	−0.011*	0.009	−0.012	0.090	−0.010	0.008
Position (central = 1)	0.033**	0.011	0.023**	0.010	0.020*	0.009
Competing (nested = 1)	−0.014*	0.010	−0.014*	0.010	−0.023**	0.008
PaceType (cut = 1)	0.000	0.010	−0.019	0.009	−0.016*	0.009
Scene complexity	−0.008	0.011	−0.012	0.011	0.000	0.009
Scene complexity^2	0.088**	0.033	0.068**	0.031	0.018	0.026
Individual dispersion	0.199**	0.011	0.203**	0.011	0.202**	0.014
Aggregate dispersion	0.055**	0.021	0.071**	0.018	0.109**	0.024
Individual ∗ aggregate dispersion	−1.249**	0.108	−1.414**	0.096	−1.910**	0.298
Age	−0.003	0.012	−0.002	0.011	−0.003	0.008
Gender (male = 1)	0.020**	0.011	0.018**	0.011	0.017**	0.008
Brand familiarity (familiar = 1)	0.001	0.030	−0.010	0.028	−0.017*	0.009
Product category (utilitarian = 1)	0.037	0.030	0.042*	0.029	0.067**	0.010

[a]Indicates that the parameter was averaged across time in model 7 to compare values with static models 5 and 6.

*Indicates that 90% confidence interval does not contain zero; **indicates that 95% posterior confidence interval does not contain that zero; +indicates that 90% posterior confidence interval does not contain zero for several time periods.

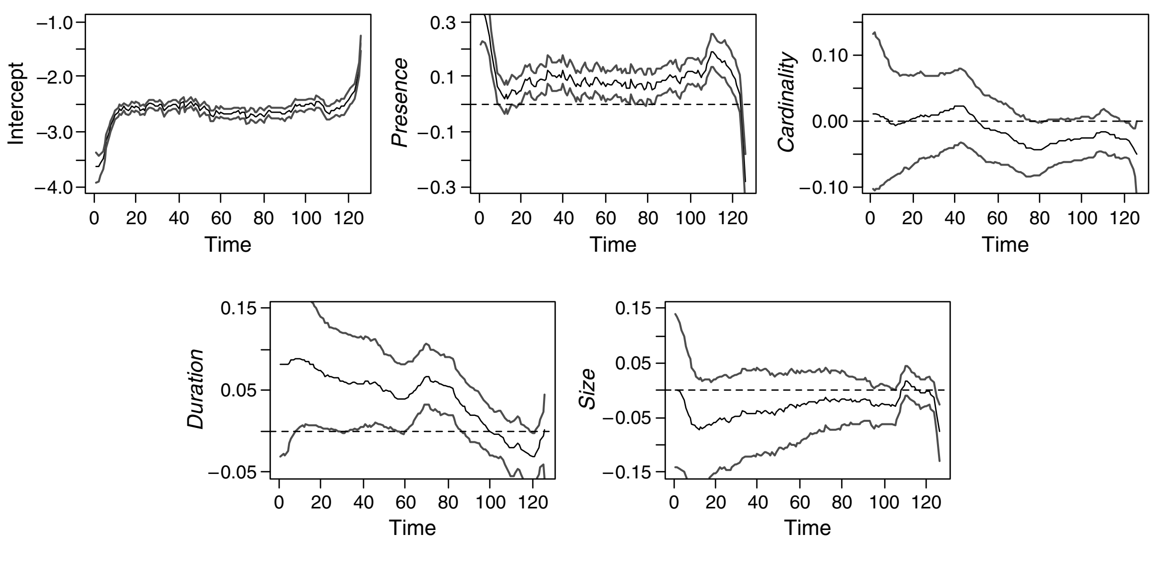
The predicted U-shaped effect of visual complexity on avoidance emerged as well, as reflected in the significant effect of complexity-squared and the nonsignificant linear term. This is the first evidence for an "optimum level" of visual complexity for commercials at which avoidance is minimal, whereas both lower and higher levels of visual complexity increase avoidance probabilities. Finally, as expected, males are significantly more likely to avoid commercials than females are. None of the other control variables was significant.

Parameter Evolution

Figure 2 plots the stochastic paths (and 90% confidence intervals) of the posterior parameter values of the intercept and dynamic effects of the brand's presence, cardinality, duration, and size, which are time varying. Baseline avoidance levels (intercept, top left in Figure 2) are fairly constant throughout the commercials, with less avoidance in the beginning, a stable and long period in the middle, and an increase toward the end. This in itself is a reassuring result because it indicates that there is no point in time, apart from the start and end, when viewers systematically tend to zap more, and this behavior is not accounted for by other covariates in the model. Brand presence drives the avoidance probability up throughout the commercial, except in the last few time frames, where brands are generally expected to appear, and consumers expect the commercial to naturally end soon. Apart from the start and end, the effect of brand presence slightly increases over time at frames 10 to 40 and 80 to 110. No strong significant effects emerged for brand cardinality. Higher cardinality of brand presence decreased avoidance toward the second half (marginally significant). Just the opposite effect emerged for duration: prolonged brand presence increased avoidance in the middle (significant) with the effect dying out toward the end.

Because variables were standardized, we can compare their relative importance directly. This shows

Figure 2 Time-Varying Parameters of Branding Activity: Posterior Median and 90% Confidence Bands



that the order of importance from highest to lowest is (1) attention dispersion metrics, with a combined posterior (absolute) mean effect of 1.50; (2) branding variables, with a total effect of 0.49; (3) visual complexity measures with 0.10; (4) product–brand control variables (brand familiarity and product category) with 0.04; and (5) demographic control variables (age and gender) with 0.02 of combined mean effect.

Optimization of Branding Activity

Marketing managers try to maximize the prominence of their brands in commercials, for instance, by exposing them early, long, in the middle of the screen, and separated from the rest of the commercial. At the same time, managers try to maximize the likelihood of retaining consumers, which is a difficult trade-off, given the results of the previous section. High levels of zapping are detrimental to the entire TV advertising industry. Networks lose ratings, individual advertisers lose viewers, subsequent commercials lose potential viewership, and the medium itself fails to engage consumers in the brand communication of firms. Therefore managers aim to maximize the opportunity to see the brand, across viewers and time, for a minimum predefined level of branding activity. We will next do this optimally, based on our model. We assume that brand owner and ad agency have established the minimum branding level in the commercial as a precondition. It is then the ad agency's responsibility to maximize opportunities to see the brand and simultaneously minimize the likelihood of avoiding the commercial from moment to moment. This decision will be respected in our optimization (within a ±5% tolerance). Formally, we define the brand activity level of commercial c (BAL_c) as the sum across time frames of the relative size of the brand, conditional on a brand presence. According to this definition, the brand activity level varies from 0, when there would be no brand appearances in the commercial, to 125 (125 frames × 100%), in which case the image would always be completely covered with the brand. In practice, observed brand activity levels are much smaller and do not show too much variability across ads, with an average of 4.65 frames, or equivalent to 1,116 ms of BAL (not to be confused with the total duration of the brand on screen, which will always be greater than BAL).

The goal of improving patterns of branding can be translated into minimizing the avoidance likelihood for a commercial c subject to a certain minimum brand activity level BAL_c. Formally, the maximization criterion opportunity-to-see (OTS) for a particular commercial c, evaluated over all I participants for the duration T, and coding avoidance as one (1) and nonavoidance as zero (0), is

$$\text{OTS}_c = \int_{\Theta} \frac{1}{N_i N_t} \sum_{i=1}^{I_c} \sum_{t=1}^{T_c} P(\widehat{Y}_{\text{ict}} = 0 \mid \Theta) P(\Theta \mid \text{data})\, d\Theta \quad \forall\, c = 1, \ldots, C. \tag{8}$$

Both uncertainties in the decision space as well as in the parameter space are taken into account in the optimization routine (Rossi and Allenby 2003). This objective function is integrated over the posterior distribution of Θ, which is approximated by averaging across the R draws of the MCMC chain:

$$\text{OTS}_c = \frac{1}{R} \sum_{r=1}^{R} \frac{1}{N_i N_t} \sum_{i=1}^{I_c} \sum_{t=1}^{T_c} P(\widehat{Y}_{ict} = 0 \mid \Theta^{(r)}) \quad \forall\, c = 1, \ldots, C. \tag{9}$$

We focus on branding decisions that can be made both before and after the actual production of the commercial and even while running the campaign to allow marketing managers and agencies optimal flexibility. Some postproduction changes in branding cannot be easily made without making large aesthetic compromises. For example, whether the brand is embedded within the scene or not cannot be easily manipulated postproduction, and the same goes for the position of the brand. Therefore, we optimize brand presence and size as instantaneous, and cardinality and duration as dynamic branding features, with all other variables remaining unchanged from their current values.[5]

To ensure a realistic solution for the optimum branding patterns, constraints are placed on the variables to be in the range of the observed values in our data (see Table 1). Size and presence of brand are the two decision arguments, because presence at $t = 1, \ldots, T$ determines cardinality_t and duration_t. Brand presence is a dichotomous variable, assuming values one or zero for presence or absence, respectively, and size is taken to vary from 0.5% to 75%, subject to the constraint that total brand activity stays the same (±5% tolerance). The probability of commercial avoidance is a monotonic increasing function of utilities in our model with convex constraints, so that we solve the following set of C decision problems, one for each commercial, in the utility space and map the solution back to the probability space. Equation (10) describes the optimization problem:

$$\min_{\left\{ \substack{Size_t \\ Presence_t} \right\}} \sum_{r=1}^{R} \sum_{i=1}^{I_c} \sum_{t=1}^{T} D_{cit}^{\text{zap}} \Big/ \Theta^{(r)} \quad \forall\, c = 1, \ldots, C$$
$$\text{with } Presence_t \in \{0, 1\},$$
$$\text{with } Size_t \in [0.5\%, 75\%],$$
$$\text{with BAL} = \text{BAL}_c \pm 5\%. \tag{10}$$

[5] Although challenging, in the next section we show that these brand features can even be manipulated postproduction.

It is noteworthy that the above objective function is linear in $Presence_t$ and $Size_t$ (see Equations (3) and (4)) but has nonlinear constraints (per the definition of BAL) and thus may not yield corner solutions. The solution to this optimization problem is a 2 (presence and size) × 125 (total number of time frames) matrix for each of the 31 commercials.

We perform the optimization using a combination of a gradient method and a genetic algorithm (Sekhon and Mebane 1998). This combines the benefit of a deterministic fast steepest descent, when the gradient of this multidimensional function can be calculated, with the benefit of stochastic search, to avoid local optima solutions. Because of the computational burden we use $R = 10$ in the optimization. Although this approach substantially reduces the likelihood that a solution is only a local as opposed to a global optimum, it does not guarantee global optimality because of the high dimensionality of the problem (250 decision variables), mixed continuous and discrete decision variables, and nonlinear constraints.

"Ceteris Paribus" Analysis

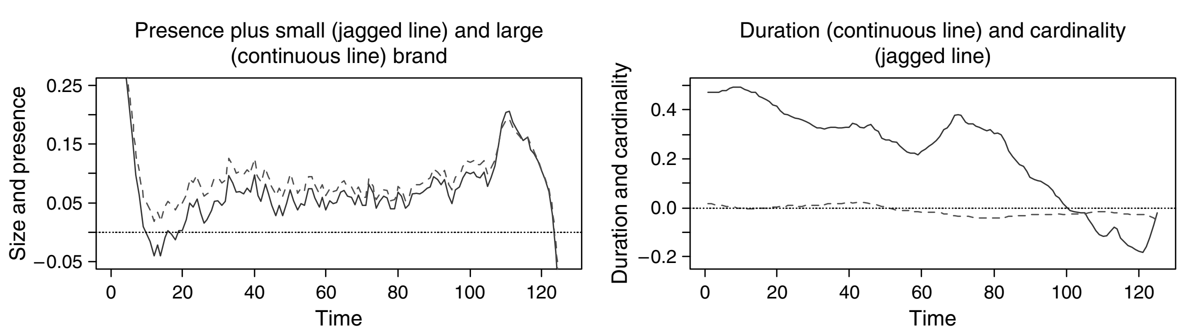
The optimal effect of branding on (minimal) avoidance likelihood depends predominantly on four variables and their estimated time-varying effects: presence, size, cardinality, and duration of the brand. The combination of the effects of these variables will dictate if and when brands increase or decrease avoidance likelihood. By itself, ceteris paribus, a brand presence will increase avoidance, but taking cardinality, duration, and size into account, it may in fact decrease avoidance at certain moments, as is the case, for example, for a large brand shown in the beginning. This is illustrated in the leftmost graph of Figure 3, where the parameter for presence is added to the parameter for size for the largest brand size (75%) and for smallest brand size possible (0.5%). Because the Y-axis shows the contribution to avoidance, larger positive values increase and larger negative values decrease avoidance. Notice how the line for the largest brand size is almost always below that for the smallest size, indicating less avoidance. Only toward the end of the ad ($t > 110$) do smaller brands cause less avoidance. Also, note that for a brief period in the beginning ($10 < t < 20$), large brand appearances systematically decrease avoidance from moment to moment.

For dynamic branding effects, one needs to compare the parameter estimates of cardinality and duration. Generally speaking, for each time period, if the duration parameter is larger than the cardinality parameter, both measured in units of time frames, then adding a new nonadjacent frame with a brand will decrease avoidance in comparison to adding an adjacent brand. Similarly, if the opposite occurs, then adding a brand in a subsequent frame is more desirable. These parameters are plotted on the right-hand graph of Figure 3, which shows that for a predefined branding level, nonconsecutive brand placements will decrease overall avoidance more than consecutive brand placements from the start of the ad up until the 105th frame. After that, and until almost the end of the commercial, the opposite is true: clumping brands together in time is preferable.

If the proposed model did not have time-varying parameters on cardinality and duration, these differential impacts of consecutive versus nonconsecutive branded frame insertions could not be effectively assessed. In particular, the static HB probit model's parameters in Table 4 show that effects of cardinality and duration are both significant and equal in magnitude up to the third decimal place, rendering the zapping effects caused by either spreading out branded frames in time or clumping them together indistinguishable.

Thus, we expect that the optimization results for these commercials should indicate that brands be larger toward the beginning (not in the first second) and at the very end, with size not being critical in the middle portion of the commercial. Also, brand appearances should be short and frequent in the first four-fifths of the commercial and be longer and less frequent in the last one-fifth. Little additional insights can be foreseen. The need for this optimization exercise lies therein that cardinality, duration, and size may all trade-off, even while brand activity level

Figure 3 **Time-Varying Parameters**



remains the same before and after. This constraint that total screen time occupied by brands is constant, in addition to the fact that the upper bounds of cardinality and duration are a function of BAL for each commercial and the time-varying effects, warrants a formal numerical optimization procedure as described by Equation (10).

Optimization Results

The brand activity level in commercials ranged from 0.38 to as much as 15.25 (mean = 4.65). All 31 commercials were individually optimized, subject to their BAL remaining unchanged. The optimization procedure was carried out in a Linux Grid server based on processors with 3.0 GHz of speed and 15 GB of memory, taking from 49 to 172 hours of CPU clock time to arrive at the solution depending on the specific commercial. Table 5 presents the results.

On average, avoidance dropped by 7.9% in the optimized compared to the original commercials, with a range from 2.0% to 19.1%. All improved ads were predicted to be avoided less than their original counterparts, and for 12 out of the 31 ads the magnitude of the reduction was larger than the estimation error. The reduction in avoidance is mainly caused by increases in brand cardinality, as predicted in the previous section. Also, apart from the extremes (start and finish of ads), brands that appear later cause more zapping than ones that appear earlier, with larger brands only causing marginally less zapping in the first half of the commercial, as the left graph of Figure 3 shows. Total brand duration (the sum of number of frames with a brand appearance) is decreased for those ads with comparatively high original total duration and increased for those with comparatively low total duration, trading it off with size. In other words, if total duration is decreased from the original to the improved version, then size is increased, and vice versa. Thus, managers need to make trade-offs in branding duration and size to strike a balance. The extent to which each of the above issues (increase in cardinality, earlier brand appearance, total duration,

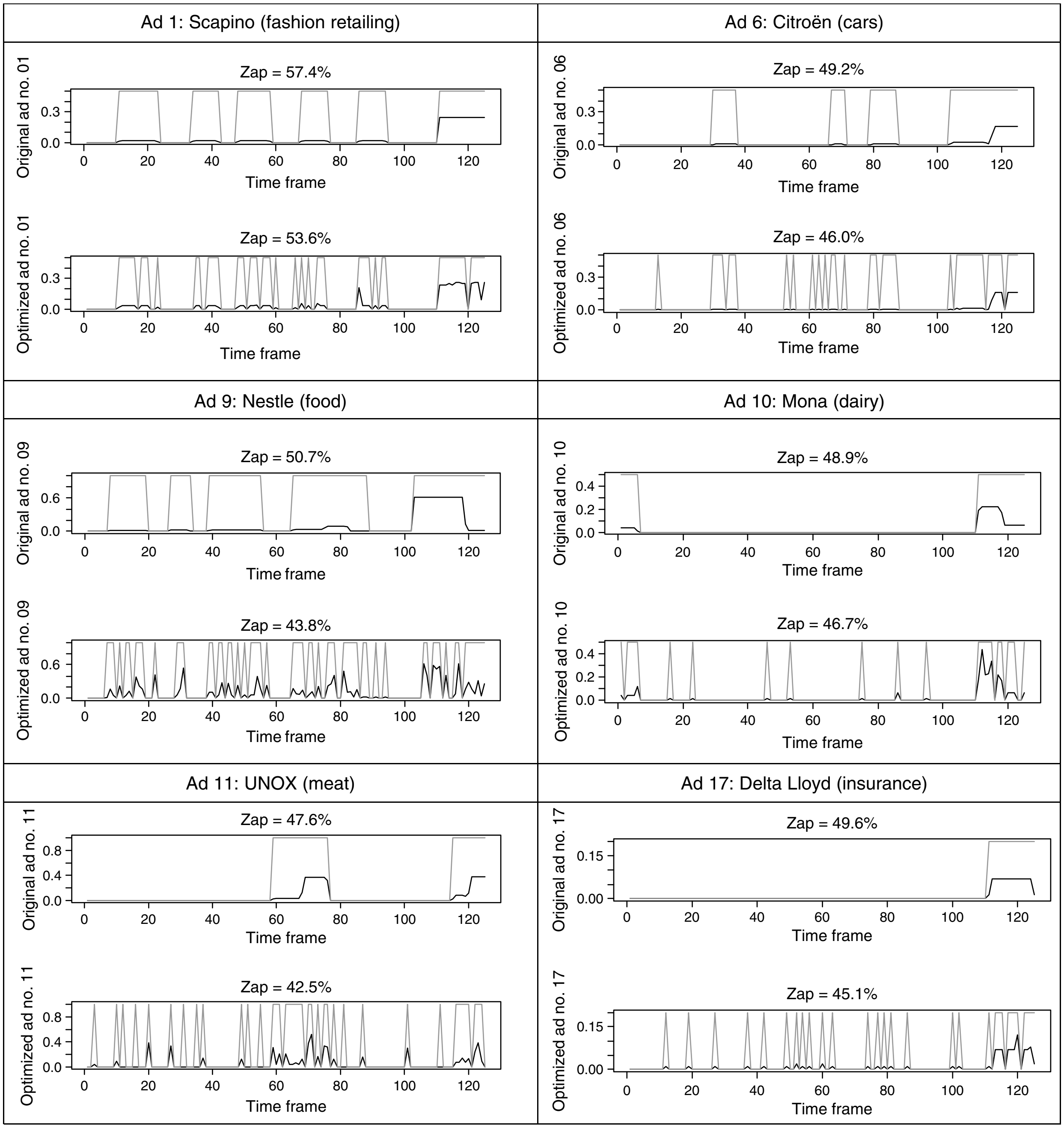
Table 5 **Brand Activity in Original and Optimized Ads**

Advertised brand	Original ad					Optimized ad				Reduction in commercial avoidance (%)
	Est. CA (%)	BAL	Card.	Dur.	Mean size (%)	Est. CA (%)	Card.	Dur.	Mean size (%)	
1. Scapino	57.4	4.40	6	66	6.7	53.6	18	53	8.7	6.7
2. Schimmelnagels	55.9	2.72	2	44	6.2	53.4	13	45	6.2	4.5
3. Staatsloterij	49.4	2.53	1	34	7.4	46.4	14	41	6.3	6.1
4. Hertog	46.5	6.79	1	25	27.2	40.4	29	41	16.6	13.2[a]
5. Mona	46.2	1.02	3	29	3.5	43.4	17	34	3.2	6.0
6. Citroën	49.2	1.67	4	44	3.8	46.0	16	43	3.7	6.5[a]
7. Sportlife	43.5	2.77	2	14	19.8	39.8	20	25	11.5	8.5
8. Post Bank	53.0	2.75	2	24	11.5	50.5	13	32	8.6	4.7
9. Nestle	51.3	12.0	5	83	14.5	41.5	32	48	26.3	19.1[a]
10. Mona	48.9	2.31	2	21	11.0	46.7	13	23	10.5	4.5[a]
11. UNOX	47.6	5.54	2	29	19.1	42.5	24	39	14.3	10.7[a]
12. Telefoongids	49.8	0.61	2	21	2.9	48.8	5	20	3.2	2.0[a]
13. NN-1	57.0	4.15	2	20	20.7	53.4	19	32	13.4	6.4
14. Albert Heijn	53.6	4.92	3	42	11.7	50.1	17	43	12.0	6.5
15. Achmea	53.5	0.73	1	9	8.1	51.2	13	21	3.5	4.3
16. Red Band	53.1	2.98	3	53	5.6	50.3	10	48	6.4	5.3
17. Delta Lloyd	49.6	0.94	1	15	6.3	45.1	22	30	3.3	9.1[a]
18. Electro World	49.3	10.9	4	43	25.4	44.7	29	48	22.9	9.4
19. MasterCard	47.1	7.02	1	21	33.4	41.4	27	37	19.1	12.1[a]
20. UNOX	56.4	9.16	2	51	18.0	50.8	24	51	18.0	10.0
21. NN-2	53.8	5.28	1	11	48.0	49.4	25	35	15.3	8.1
22. Essent	55.1	2.88	1	38	7.6	53.6	7	39	7.8	2.7
23. SNS Bank	55.0	1.29	1	9	14.4	52.8	16	23	5.6	4.0
24. T-Mobile	46.8	3.29	1	10	32.9	43.8	15	19	17.9	6.3[a]
25. Radio 538	42.9	1.12	1	8	13.9	39.8	13	18	6.5	7.2
26. KWF	51.5	3.74	1	37	10.1	49.8	11	41	9.6	3.3
27. Kodak	50.6	2.35	2	29	8.1	47.7	14	36	6.9	5.8[a]
28. Min. of Justice	59.4	0.38	1	11	3.5	57.4	11	21	1.9	3.5[a]
29. Wadden	41.1	10.5	1	32	32.8	33.8	30	41	26.4	17.7
30. Post Bank	55.5	12.1	1	77	15.8	47.5	22	67	18.5	14.4[a]
31. SMS Land	54.7	15.3	1	125	12.2	46.1	11	115	13.9	15.6

Note. Est. CA = estimated commercial avoidance; BAL = brand activity level in the commercial; Card. = cardinality of the brand appearance; Dur. = total sum duration of the brand appearance.

[a]Indicates that the overall postoptimization avoidance reductions for these ads are larger than their estimation error.

Figure 4 Illustrating Branding Optimization Frame by Frame



Notes. Plots of brand presence and size for six optimized ads are shown. The upper graph is for the original ad and lower graph is for the optimized ad; the light thick line is brand presence and dark thin line is size.

and size) is mostly responsible for the optimal solution depends on the specific time frame, because of the specific way the parameter estimates of these variables vary over time.

Figure 4 shows the brand presence (thick line) and the size of the brand (thin line) for 6 out of the 31 ads optimized for the original ad (upper graph) and the improved ad (lower graph). Notice how, consistent with the predictions in the ceteris paribus analyses, most of the improved ads have more/shorter brand appearances up to around the 100th frame mark and less/longer ones thereafter.

The improved solutions have frequent but brief brand appearances. This result is analogous to the finding of pulsing benefits *across exposures* in the advertising effectiveness literature (Feichtinger et al. 1994, Feinberg 2001) but is shown here *within exposures*. It is due to the linearity of the model and the mixed continuous and discrete decision variables (Hahn and Hyun 1991), combined with the fact that discrete adjacent brand placements have a higher "cost" than nonadjacent ones. The optimization exercise is an attempt to uncover, directionally, the pattern of a generalizeable brand allocation strategy. Practical feasibility of such a strategy is further investigated and discussed in the following sections.

To validate our findings, we compare in Table 6 the avoidance rates obtained from our procedure (strategy 1) with eight alternative branding strategies, including no branding (strategy 2), current branding practice (strategy 3), and branding strategies that systematically vary part of the commercial in which the brand is placed (first half, second half, all) and its size (largest brand size is 75% and smallest is 0.5%). The avoidance rates of the above strategies are averaged across all 31 ads. Table 6 shows that our optimization solution is better than all these alternative strategies. Note how strategy 2, in which there is even no brand placement, outperforms the other strategies—except the proposed strategy (although not significantly so)—which shows again that brands are avoided by viewers but that brand pulsing reduces this. Because of the potentially important implications for advertising strategy, the question is pertinent to what extent the main finding of optimal brand-pulsing strategies are not due to idiosyncratic aspects of our data, commercials, or model. Therefore, we next check the robustness of our findings to several aspects of the data and model and, in the sequel, validate our findings in a controlled lab experiment.

Robustness Checks

The finding in the numerical optimization that brand pulsing minimizes zapping relative to other brand placement strategies is contingent upon (a) the model being adequately specified and (b) the viewer being able to react fast following a brand exposure. In this section we further investigate the robustness of our findings to violations of these assumptions.

First, a brand identification, key-press reaction time (RT) experiment was conducted, using the DirectRT software, on 60 participants across six commercials. Participants were undergraduate students of a university in The Netherlands and received a small financial compensation for their participation. They were seated individually in dimly lit cubicles and were shown each of six commercials on a computer screen. Prior to the commercials, they first saw an image of the brand in each commercial and were instructed to zap by pressing the space bar as soon as the brand appeared in the commercial pod that followed. The average modal reaction time was 460 ms (std. dev. = 115 ms). Per ad, between 0% and 24% of the reaction times were below 240 ms. Across ads, nearly half of the participants (equivalent to 1,000 viewers in our estimation data) reacted within 500 ms of brand exposure, showing that viewers may zap brands within one or two frames of brand exposure as defined in our data aggregation.

To assess the impact of such a delay in zapping reaction on the results, we tested different models with lagged effects of brand features on zapping. Fit, as measured by the log marginal density (LMD), is highest for the dynamic probit model with contemporaneous effects of the brand features presented above (LMD $= -9{,}533$) as opposed to models with lagged effects of these features (lag 1: LMD $= -9{,}597$; lag 2: LMD $= -9{,}572$.[6]) An explanation for this finding is that the time-varying parameters in the model already capture a significant portion of these lagged effects. We also estimated our models after aggregating the data to time windows of 960 ms (four frames), capturing 95% of RTs in the above experiment. We observe no qualitative differences among most critical parameters in the two models, the only exception being that cardinality became significantly negatively associated with zapping.[7] These findings do not undermine the assumptions of our model specification that lead to the optimality of brand pulsing. On the contrary, the stronger effect of cardinality strengthens our finding as seen in the right-hand graph of Figure 3.

Table 6 **Estimated Commercial Avoidance for Current, Optimized, and Benchmark Branding Strategies**

Comparison of branding strategies	Mean estimated commercial avoidance (%)	Std. dev. (%)
1. Optimized brand placement (our model)	47.2	5.2
2. No brands present	49.3	4.3
3. First half of ad with largest brands	51.0	4.4
4. Current branding practice	51.1	4.5
5. Second half of ad with largest brands	51.1	4.2
6. Second half of ad with smallest brands	52.5	4.2
7. All ad with largest brands	55.2	4.1
8. First half of ad with smallest brands	57.0	4.5
9. All ad with smallest brands	62.6	4.2

Note. Smallest brand = 0.5% of screen size, largest brand = 75% of screen size.

[6] We deleted the first second observations for each commercial, needed to initialize the lags, to make the LMDs comparable.

[7] Modality and separation become insignificant and pace type and visual complexity become significant; these results can speculatively be attributable to temporal aggregation bias, because shorter brand and scene features are averaged out within an interval.

Validation Experiment

We now present the results of a lab experiment in which we take six 30-second commercials with varying amounts of total visual branding and alter the number of short nonconsecutive brand pulses to compare the effects on average zapping rates. Because the objective is to validate the optimality of pulsing, ideally one would reconfigure the patterns of brand pulses to mimic those found in the previous optimization section, i.e., inserting an average of about 17 pulses. However, because not all ads are amenable to post hoc reengineering of brand placement without compromising their creative execution and cohesiveness, and because of the authors' inability to do so in all cases without the altered version appearing to be tampered with, we proceed to moderately increase (+3 to +5) or decrease (−2 to −4) the number of brand pulses as seamlessly as possible. Strictly speaking, this experiment is a test of the optimality of pulsing via moderate, not drastic, changes in cardinality of the brand. Consequently, we do not claim to test the optimal frequency of pulsing. Care was also taken to nest the brand while not "hiding" it in the scene and maintaining a constant average brand size.

The stimuli were six[8] commercials, altered to have either a higher or lower number of brand pulses than the original ones, as well as seven other filler ads that were the same across conditions. The ads were chosen randomly from a pool of ads conveniently available to the researchers and provided by Verify International.[9] For each randomly selected commercial, a subjective evaluation was made to decide if modifying the commercial was feasible in the sense of it being technically possible to (1) isolate the brand and replicate it in other time frames without occluding any images, (2) eliminate some of the frames with the original brand or eliminate the nested brand from the scene, and (3) identify every appearance of the brand upon first exposure to the commercial. If the chosen commercial obeyed these conditions, it was used in the experiment, or else another ad was chosen from the available pool. For the chosen commercials, the video-editing software Adobe Premiere Pro™ was used to insert new brand appearances as uniformly as possible in time to mimic the patterns in Figure 4, albeit at moderate pulses.

The experiment used a 2 (original versus altered commercial) × 2 (commercial sequence 1 versus 2) between-subjects design, with six commercials as replicates. The commercial sequence factor was included to account for serial position effects. Thus, each participant had the opportunity to see and zap three (or four) original ads and four (or three) altered ads all interspersed by the seven filler ads. A total of 130 participants (undergraduate students, mean age of 20 years, 61% male) were randomly assigned to the conditions and first individually watched a four-minute TV show on the computer to ease them, after which a commercial pod with 14 commercials was shown in sequence. They could watch each ad or press the space bar (on which the index finger rested at all times) to skip to the next ad in the sequence. After the last commercial, they answered questions regarding the experiment, engaged in other unrelated tasks, were thanked, debriefed, and paid (the equivalent of US$8 for the complete experimental session, which took about one hour).

The results of the experiment are summarized in Table 7. The percentage of commercials zapped across participants ranged from 7% (five participants zapped only one commercial) to 100% (eight participants zapped all commercials), with a mean zapping rate of 60% (std. dev. = 25%). This is very similar to the original data set, with a zapping rate of 55% (std. dev. = 12%) and to zapping rates reported in other studies (Wilbur 2008). Table 7 shows that out of the six commercials, four showed appreciable differences in zapping, ranging from 9% to 25%, between the versions with high and low numbers of brand pulses. In particular, three commercials altered to have *higher number* of pulses showed major decreases in zapping and one commercial altered to have a *lower number* of original brand pulses showed a major increase in zapping. This is an indication that our findings work both ways, at least in the interval between 1 and 10 brand pulses that we studied: as predicted by the model, more pulses decrease zapping and less pulses increase it, holding total BAL constant. For the two commercials for which we did not find an effect an exact reason is unknown, but care should be taken to propose an explanation since all altered commercials are suboptimal; i.e., they are not professionally executed modifications from the original, professionally developed TV ads. Also, note that the number of pulses that we could realistically implement is somewhat lower than those suggested by the optimization results. We could not implement an assessment of whether a much higher pulsing frequency will continue to decrease zapping; we leave that issue for future research.

The zapping rates for versions of the commercials with high pulsing (mean of 67%) were lower than the zapping rates for those with low pulsing (mean of 74%) as shown by a test of proportions at 95% confidence ($p = 0.02$). Two points are worth noting. First, the average relative reduction in zapping of 9.5%

[8] Originally, seven commercials were selected, but one of the commercials used was dropped because the execution was problematic and evaluated by some participants as being tampered with.

[9] Three of the altered commercials were in the original data set and three were not.

Table 7 Comparison Between Zapping Rates of Original and Altered Commercials

			Brand pulses		Zapping rate		
	Brand	Category	Original	Altered	Original (%)	Altered (%)	Change (%)
Increase in brand pulses	Scapino	Clothe	5	10	92	84	−9
	Mona	Desert	3	6	76	65	−14
	Mastercard	Financial	1	5	69	69	0
	Pastrelli	Food	5	8	63	56	−11
Decrease in brand pulses	Dommelsch	Beer	8	6	60	75	25
	Nike	Sports	7	3	66	66	0

Note. Total duration of brand exposure is the same in original and altered commercials.

(compared to the 7.9% reduction in the optimization simulation) is attained by increasing the brand cardinality but not the total duration of brand exposure on screen, which remained the same. Second, it is important to note that the averaged viewing times for the viewers that did zap are not very different from each other: 17.40 seconds for altered commercials and 17.86 seconds for original commercials. Thus, the brand-pulsing strategy within the ad exposures did not erode the viewing time (and possibly, the overall quality of the experience) for the viewers that eventually decided to stop watching before the end.

Apart from the evidence from the parameter estimates in the model, and the evidence from the optimization, the results of this experiment provide strong additional verification that pulsing short brands across time, even if done in a frequency that is lower than potentially optimal, provides benefits in terms of significantly reducing commercial avoidance levels. Moreover, we conjecture that if brand pulsing is to be simultaneously integrated with ad creation and execution as opposed to postexecution, as was done in this experiment, the additional executional coherence will further contribute to reaping the benefits of this strategy.

Discussion

Branding activity in television commercials affects the moment-to-moment likelihood that these commercials retain their viewers. Specifically, inserting brands for sustained periods—in particular, centrally on the screen—increases the likelihood that consumers stop watching a particular commercial notably. However, our model estimation and optimization procedure suggests that what is "intrusive" is not the total branding activity level per se, but long, sustained brand appearances. Thus, we were able to lower avoidance rates of commercials by merely changing the pattern of brand exposure while keeping the brand activity level per commercial the same. A pulsing strategy, in which brands are shown more frequently and for a shorter time instead of infrequently and for a longer time, decreased commercial avoidance rates both in a simulated optimization and in a lab experiment by about 8% and 10%, respectively. For the former, avoidance rates under this optimized branding strategy were even lower than if the brands would not have appeared in the commercials at all, reflecting the value of brands in ads. This suggests a parallel between optimal effectiveness of brand pulses within commercials and commercial pulses in media placement in a campaign (Feichtinger et al. 1994, Feinberg 2001). Because our objective was to provide a solution to the advertiser's problem of how to insert the brand in commercials while retaining viewers' attention, a behavioral foundation for the optimality of pulsing was not provided. Nonetheless, a conjecture is that such brand pulses leave the narrative in the commercial more intact and thereby interfere less with the entertainment goals that consumers generally have when watching television. One common strategy to cope with increased commercial avoidance is to reduce the overall brand activity levels in commercials and place the brand only once, and completely at the end and for a longer time, as reflected in the growing incidence of soft-selling mystery commercials. However, our model parameters show that this strategy can only retain attention for one or two seconds of brand exposure at most, and the single exposure only may adversely affect memory and learning in addition (Brown and Craik 2000). Our findings show that intrusiveness can be reduced more, without sacrificing brand exposure levels, using a brand pulsing strategy. The dynamic probit model, optimization approach, and experimental validation on which the present findings are based hold the promise of improving the effectiveness of television advertising through the insights they provide into the moment-to-moment determinants of commercial avoidance. Because of its focus on branding activities that are largely under managerial control, independent of the creative content of commercials, our procedure can be used both before and after final production and even while the campaigns are in the media. Moreover, because adaptations can be made postproduction, as was done in our follow-up experiment, improvements are virtually costless. Nonetheless, there may be ads where altering the brand's location, without drastically changing its visual complexity, pacing, or other visual elements may be more difficult to do as a consequence of their creative design.

Independent of branding activity and other factors, the ability of commercials to concentrate consumers' visual attention reduced commercial avoidance significantly. Specifically, the smaller the variance in the location of consumers' eye fixations (aggregate attention dispersion), the lower the likelihood of commercial avoidance was. Also, the closer an individual consumer's eye fixations were to the center of other consumers' eye fixations (lower individual attention dispersion), the lower the likelihood of commercial avoidance was. The interactive effect between the two attention dispersion metrics suggests that lower aggregate and individual attention dispersion led to the lowest commercial avoidance likelihoods. As far as we know, these results are the first to show that a commercial's power to concentrate, hold, and direct visual attention directly predict consumers' decisions to stay with the commercial. The findings support that, indeed, as often speculated upon in advertising, the power to orchestrate attention is crucial to advertising effectiveness. The proposed attention dispersion metrics can be readily derived through eye tracking of commercials and may prove useful in advertising effectiveness research in general.

Consumers' moment-to-moment decisions to continue or stop watching commercials also depended on the optimal amount of visual complexity in the commercials, independent of all other factors. That is, under both low and high levels of visual complexity, the likelihood to stop watching commercials was higher than under intermediate levels. We believe that this finding is particularly interesting because it was obtained using objective, novel measures of complexity, based on the pacing of commercials and the density of visual information in the GIF-compressed file size of each frame in the commercials. To our knowledge, the present findings are the first to show that objective measures of visual complexity directly influence consequential consumer decisions. These measures allow marketing managers and advertisers to assess the frame-by-frame visual complexity of their commercials to supplement other quality indicators and opportunities to fine-tune visual complexity levels to reduce commercial avoidance.

Limitations and Research Opportunities

The issue of how brand pulsing impacts brand attitude measures, purchase intention, or many other metrics used to evaluate advertising effectiveness, although important, was not addressed in the current study, as it was outside of the scope. However, to shed some light on this issue, after the validation experiment, we measured on five-point response scales anchored by "Not at all" and "Very strongly" the extent to which participants felt that (1) the commercial made them feel good about the brand, (2) the commercial aroused their interest in the brand, and (3) the commercial made them evaluate the brand more positively. No significant difference was found between commercials altered to exhibit degrees of brand pulsing and their original counterparts for any of the three measures across the six tested commercials. However, we recognize the need to assess whether the optimal brand placement strategy will affect other important ad metrics and future research should attempt to tackle this issue, where our model and methodology may be used as a starting point.

Furthermore, given the discrete (frame-by-frame) nature of our data, both a binary choice model and a (discrete time) hazard model could be applied in this context and should provide similar results (Sueyoshi 1995). We choose the frame-by-frame probit model because it ties in directly with our frame-by-frame optimization of the TV commercials and because frameworks for dealing with time-varying parameters have been well established for this model (Lachaab et al. 2006). However, a hazard model could be a viable alternative.

The limitations of our research design is that viewers watched sequences of commercials back-to-back without programs, and this may have increased the likelihoods of avoidance decisions relative to those obtained under natural viewing conditions at home. Nevertheless, because we employ a frame-by-frame analysis, as opposed to a commercial-by-commercial analysis, we expect that the zapping instants shown to be systematically affected by brand presence would not be qualitatively different. In addition, in the validation experiment, commercials were embedded in a TV show, and the overall zapping rates obtained were very similar. Thus, although there is no reason to expect that the research context may have prompted consumers to become sensitive to qualitatively different factors, only future research can provide definitive answers to this question.

The present study could not assess the impact of program because commercials were shown in sequence. Such sequential presentation of the commercials is somewhat realistic and reflective of several conditions occurring in practice—in particular, those where networks coordinate so-called "roadblocks," which are time-synchronized commercial breaks on different channels. In those situations, a consumer who zaps out of a commercial zaps into another one at the other channel. Our experimental design is reflective of these situations. It is not unlikely, however, that programs or shows have an impact on attention to commercials (Burns and Anderson 1993), yet both zapping rates and the main pulsing finding were replicated in the validation experiment, which did involve such programming content. A systematic assessment of this effect would need to vary the type of program, content, entertainment, and information value, among other elements, to understand the influence on ad avoidance. We suggest this as another important avenue for future research.

Another research opportunity concerns improvements of the elementary metrics of visual attention that we could use here. Eye tracking of dynamic stimuli such as television commercials is challenging because of the doubly dynamic character of the data. That is, the eyes move across scenes that move themselves or have objects that do so. The resulting large streams of such doubly dynamic data are a main reason for the lack of prior eye-tracking research of commercials and other dynamic stimuli (Wedel and Pieters 2008). The present findings demonstrated that our aggregate- and individual-level attention dispersion measures were strongly predictive of commercial avoidance decisions, even though they are independent of where in the scene consumers' attention was actually located. It seems likely that refining the metrics to include the concentration location (or the advertiser's desired location of focus) will increase their predictive validity, and future research may address this. More generally, in view of their predictive validity for avoidance decisions, future research may examine the factors that influence consumers' attention dispersion in commercials.

Finally, in our optimized commercials, brand size and duration remained in the range of the current values, but the cardinality did not. That is, the average cardinality went from a low 2.0 (original ads) to a mean 17.7 (improved ads). Although it is not directly a decision variable, this proposed number of nonconsecutive brand insertions is far from the maximum cardinality observed in our data set (6), but it is not uncommon in advertising practice. Examples are the recent commercial "The Happiness Factory" for Coca-Cola[10] with a cardinality of 17 (short version) and the 2008 Coca-Cola "Super Bowl commercial" with a cardinality of 13. Automobile commercials often portray cars being driven around an urban scene. The frequent changes of camera angle shots as well as the driver turning the car "induce" a natural pulsing strategy in which the brand logo (in front and back of the car) is shown in multiple but brief instances during the ad. Examples of this are the 2009 Mercedes Benz ("Narrow" by Merkley + Partners), Audi ("Chase" by Venables, Bell & Partners), and BMW ("Z4 Roadster" by GSD&M Idea City), all of which show 12 to 13 brand pulses in 25 to 30 seconds. Although the average avoidance rates of these commercials are unknown, these commercials exemplify that the prescribed pulsing strategy is possible and that high levels of brand pulsing are being used by successful firms. It is important to note that our validation experiment revealed that the main pulsing result holds up even for more moderate cardinalities. We do not claim to provide conclusive evidence for the optimal frequency of brand pulsing required to minimize commercial zapping, only that brand pulsing works in this direction across a variety of brands, known and unknown, liked or not.

More than ever, consumers can easily avoid commercials at any point in time. The proposed procedure that relates objective characteristics of commercials and attention metrics obtained through eye tracking to consumers' moment-to-moment avoidance decisions can be used in advertising testing before and during campaigns and holds the promise to increase television advertising's effectiveness.

Appendix. Model Specification and Estimation

The deterministic component of the model, D_{ict}, is expressed in a hierarchical Bayes structure as follows, in what is a summary of Equations (3)–(6):

$$
\begin{aligned}
D_{ict} &= \mu_i + \alpha_c + B_{ct} + (\gamma^1 \mathrm{AAD}_{ct} + \gamma^2 \mathrm{IAD}_{ict} \\
&\qquad + \gamma^3 \mathrm{AAD}_{ct} \times \mathrm{IAD}_{ict}) + \mathrm{TVC}_{ct},\\
\mu_i &= \Lambda^1 \cdot \mathit{Age}_i + \Lambda^2 \cdot \mathit{Gender}_i + V_\lambda \quad \Lambda \sim N(\Lambda_0, \Sigma_\Lambda),\\
\alpha_c &= K^1 \cdot \mathit{Product\ category}_c + K^2 \cdot \mathit{Brand\ familiarity}_c + V_\kappa \\
&\qquad K \sim N(K_0, \Sigma_K),\\
B_{ct} &= \theta_t \cdot \mathit{Branding}_{ct}\\
&\qquad \theta_t = (\theta_t^1, \theta_t^2, \theta_t^3, \theta_t^4, \theta^5, \theta^6, \theta^7)^T \sim N(\theta^*, \Sigma^*),\\
&(\gamma^1 \mathrm{AAD}_{ct} + \gamma^2 \mathrm{IAD}_{ict} + \gamma^3 \mathrm{AAD}_{ct} \times \mathrm{IAD}_{ict})\\
&\quad = \gamma^1 \mathit{Aggregate\ dispersion}_{ict} + \gamma^2 \mathit{Individual\ dispersion}_{ct}\\
&\qquad + \gamma^3 \mathit{Aggregate\ dispersion} \times \mathit{Individual\ dispersion}_{ct}\\
&\qquad \gamma = (\gamma^1, \gamma^2, \gamma^3)^T \sim N(\gamma^*, \Gamma^*),\\
\mathrm{TVC}_{ct} &= \beta^0 \cdot \mathit{PaceType} + \beta^1 \cdot \mathit{Visual\ complexity}_{ct}\\
&\qquad + \beta^2 \cdot \mathit{Visual\ complexity}_{ct}^2 \quad \beta = (\beta^0, \beta^1, \beta^2)^T.
\end{aligned}
$$

Let time $t = 1, \ldots, T$, commercial $c = 1, \ldots, C$, and individual $i = 1, \ldots, I$. The basic relationship of Equations (1) and (2) form the complete utility model specification and are expressed as

$$
Y_{ict} = \begin{cases} 1\text{: avoid} & \text{if } U_{ict} \geq 0,\\ 0\text{: watch} & \text{if } U_{ict} < 0, \end{cases}
$$
$$
U_{ict} = D_{ict} + \varepsilon_{ict}.
$$

The state space (dynamic probit model) formulation of the model is (with Ψ_{ic} incorporating θ^5, θ^6, θ^7, β^0, β^1, β^2, γ^1, γ^2, γ^3, and Θ_t incorporating θ_t^1, θ_t^2, θ_t^3, θ_t^4 and an intercept θ_t^0)

$$
U_{ict} = \mu_i + \alpha_c + F_{ct}\Theta_t + X_{ict}^3 \Psi_+ \varepsilon_{ict},
$$

[10] See http://www.coca-cola.com/HF/index.jsp.

$$\Theta_t = \Xi + G\Theta_{t-1} + \omega_t,$$
$$\mu_i \sim N(X_i^2\Lambda, V_\lambda),$$
$$\alpha_c \sim N(X_c^1\mathrm{K}, V_\kappa),$$
$$\omega_t \sim N(0, V_\omega) \quad \varepsilon_{ict} \sim N(0, V_\varepsilon = 1).$$

We let $G = I_5$ and $\Xi = 0$, and thus, specifying the MCMC inference procedures, we rewrite the equations as

$$U_{ict} = D_{ict} + \varepsilon_{ict} \quad \varepsilon_{ict} \sim N(0, V_\varepsilon = 1),$$
$$\underset{\text{Probit}}{U_{ict}} = \underset{\text{HB: Variance component model}}{(X_i^2\Lambda + \lambda_i) + (X_c^1 K + \kappa_c)}$$
$$+ \underset{\text{Forward filter backward sample}}{F_{ct}\left(G^t\Theta_0 + \sum_0^{t-1} G^j\omega_{t-j}\right)}$$
$$+ \underset{\text{Bayes regresssion}}{X_{ict}^3\Psi} + \varepsilon_{ict}, \quad \lambda_i \sim N(0, V_\lambda),$$
$$\kappa_c \sim N(0, V_\kappa),$$
$$\omega_t \sim N(0, V_\omega).$$

The design matrix, composed of the independent variables X and dependent variable Y, is structured in the following way:

$$X = \{F_{ct}X_c^1X_i^2X_{ict}^3\} = \left[\begin{bmatrix} 1 \\ Presence_{ct} \\ Cardinality_{ct} \\ Duration_{ct} \\ Size_{ct} \end{bmatrix} \begin{bmatrix} Product\ category_c \\ Brand\ familiarity_c \end{bmatrix} \begin{bmatrix} Age_i \\ Gender_i \end{bmatrix} \cdot \begin{bmatrix} Mode_{ct} \\ Position_{ct} \\ Separation_{ct} \\ PaceType_{ct} \\ Visual\ complexity_{ct} \\ Visual\ complexity_{ct}^2 \\ Individual\ dispersion_{ict} \\ Aggregate\ dispersion_{ct} \\ Individual * Aggregregate\ dispersion_{ict} \end{bmatrix}\right].$$

The prior distribution of parameters are diffuse conjugate distributions:

$$U_{ict}, \mu_i, \alpha_c \rightarrow \text{specified from model},$$
$$\Theta_0 \sim N_5(m_0 = 0, C_0 = 10^5 I),$$
$$\Psi \sim N_9(n_0 = 0, S_0 = 10^5 I),$$
$$\Lambda \sim N_2(\Lambda_0 = 0, \Sigma_\Lambda = 10^5 I),$$
$$K \sim N_2(K_0 = 0, \Sigma_K = 10^5 I),$$
$$V_\lambda^{-1} \sim G(\rho_\lambda = 2+1, (\rho_\lambda R_\lambda)^{-1} = (3 \cdot 0.0001)^{-1}),$$
$$V_\kappa^{-1} \sim G(\rho_\kappa = 2+1, (\rho_\kappa R_\kappa)^{-1} = (3 \cdot 0.0001)^{-1}),$$
$$V_\omega^{-1} \sim W_5(\rho_\omega = 5+1, (\rho_\omega R_\omega)^{-1} = (6 \cdot 0.0001 \cdot I_5)^{-1}).$$

To estimate the unique observation equation via Gibbs sampling, let $\Phi = \{\Theta_0, \ldots, \Theta_T, \ \Psi, \ \mu_{i=1}, \ldots, \mu_{i=I}, \ \Lambda, \ V_\lambda, \ \alpha_{c=1}, \ldots, \alpha_{c=C}, K, V_\kappa, V_\omega\}$ be the full parameter set and $\Omega_t = \{Y_{i,c,1:t}, X_{i,c,1:t}\}$ the complete data up to time t. The following algorithm describes the estimation steps along with full conditionals for each "sweep" (iteration) of the Gibbs sampler. All model parameters are estimated simultaneously by recursively sampling from their conditional posterior distributions, which are given below.

1. Probit (Albert and Chib 1993):

$$U_{ict} \mid \Omega_T, \Phi \sim \text{Truncated} - N_{\{a,b\}}(D_{ict}, V_\varepsilon = 1),$$
$$Y_{ict} = \begin{cases} 0 \rightarrow a = -\infty, & b = 0; \\ 1 \rightarrow a = 0, & b = +\infty. \end{cases}$$

2. Forward filtering backward sampling (Frühwirth-Schnatter 1994, Lachaab et al. 2006).

Let $U_{ict} - \hat{\mu}_i + \hat{\alpha}_c - X_{ict}^3\hat{\Psi} = U_{ict}^* = F_{ct}\Theta_t + \varepsilon_{ict}$, $\widetilde{U}_t^* = \text{stack}(U_{ict}^*)\ \forall\, i, c \in t$, $\tilde{F}_t = \text{stack}(F_{ct})\ \forall\, c \in t$ and $V_{\varepsilon,t} = V_\varepsilon \otimes I_{C_tI_t}$.

Forward filter: Loop forward in time and sample Normal distributions

$$\Theta_t \mid \Omega_{t-1}, \Phi_{-\Theta_t} \sim N_5(m_t, C_t) \quad \forall\, t = 1, \ldots, T,$$
$$\gamma_t = \Xi + Gm_{t-1},$$
$$\Gamma_t = GC_{t-1}G^T + V_\omega,$$
$$C_t^{-1} = \Gamma_t^{-1} + \tilde{F}_t^T V_{\varepsilon,t}^{-1}\tilde{F}_t,$$
$$m_t = C_t(\Gamma_t^{-1}\gamma_t + \tilde{F}_t^T V_{\varepsilon,t}^{-1}\widetilde{U}_t^*),$$

with dimensions $G = 5 \times 5$, $\Xi = 5 \times 1$, $\Theta = (5 \times 125) \times 1$, $F_t = (C_t \times I_t) \times 5$, $U_t^* = (C_t \times I_t) \times 1$, $\gamma_t = 5 \times 1$, $\Gamma_t = 5 \times 5$, $C_t = 5 \times 5$, and $m_t = 5 \times 1$.

Backward sampler: Loop backward in time and sample Normal distributions

$$\Theta_T \mid \Omega_T, \Phi_{-\Theta_T} \sim N_5(m_T, C_T),$$
$$\Theta_t \mid \Theta_{t+1}, \Omega_{t-1}, \Phi_{-\Theta_t} \sim N_5(q_t, Q_t) \quad \forall\, t = T-1, \ldots, 0,$$
$$Q_t^{-1} = C_t^{-1} + G^T V_\omega^{-1} G,$$
$$q_t = Q_t[C_t^{-1}m_t + G^T V_\omega^{-1}(\Theta_{t+1} - \Xi)],$$

with dimensions: $Q_t = 5 \times 5$ and $q_t = 5 \times 1$.

3. Conjugate sampling (Lachaab et al. 2006):

$$V_\omega^{-1} \mid \Omega_T, \Phi_{-V_\omega} \sim W_5\left(\rho_\omega + T, \left(\rho_\omega R_\omega + \sum_{t=1}^{T}(\Theta_t - G\Theta_{t-1})^2\right)^{-1}\right).$$

4. Bayesian regression:

Let $U_{ict} - \hat{\mu}_i - \hat{\alpha}_c - F_{ct}\hat{\Theta}_t = U_{ict}^{**} = X_{ict}^3\Psi + \varepsilon_{ict}$, $\widetilde{U}^{**} = \text{stack}(U_{ict}^{**})_{\forall i,c,t}$, $\widetilde{X}^3 = \text{stack}(X_{ct}^3)_{\forall i,c,t}$:

$$\Psi \mid \Omega_T, \Phi_{-\Psi} \sim N_9(M_\Psi, V_\Psi),$$
$$V_\Psi = (\widetilde{X}^{3'}\widetilde{X}^3 + S_0^{-1})^{-1},$$
$$M_\Psi = V_\Psi(\widetilde{X}^{3'}\widetilde{U}^{**} + S_0^{-1}n_0).$$

5. HB: Variance component model (Gelfand et al. 1990). Individual-specific baseline intercepts:

Let $U_{ict} - \hat{\alpha}_c - F_{ct}\hat{\Theta}_t - X_{ict}^3\hat{\Psi} = U_{ict}^{***} = \mu_i + \varepsilon_{ict}$,

$$\mu_i \mid \Omega_T, \Phi_{-\mu_i} \sim N\left(\frac{V_\lambda \sum_{ct} U_{ict}^{***} + \Lambda' X_i^2}{C_iT_iV_\lambda + 1}, \frac{V_\lambda}{C_iT_iV_\lambda + 1}\right) \quad \forall\, i = 1, \ldots, I,$$

$$\Lambda \mid \Omega_T, \Phi_{-\Lambda} \sim N_2\left(\mathrm{Var}_n\left[\left(X^{2'} \cdot \frac{1}{V_\lambda}\right)\mu + \Sigma_\lambda^{-1} \cdot \Lambda_0\right], \mathrm{Var}_n\right),$$

$$\mathrm{Var}_n = \left(\left(X^{2'}X^2 \cdot \frac{1}{V_\lambda}\right) + \Sigma_\Lambda^{-1}\right)^{-1},$$

$$X^2 = \mathrm{stack}(X_i^{2'}), \quad \mu = \mathrm{stack}(\mu_i),$$

$$V_\lambda \mid \Omega_T, \Phi_{-V_\lambda} \sim \mathrm{IG}\left(\rho_\lambda + \frac{I}{2}, R_\lambda + \frac{1}{2}(\mu - X^2\Lambda)^T(\mu - X^2\Lambda)\right).$$

Commercial-specific baseline intercepts:
Let $U_{ict} - \hat{\mu}_i - F_{ct}\hat{\Theta}_t - X_{ict}^3\hat{\Psi} = U_{ict}^{****} = \alpha_c + \varepsilon_{ict}$,

$$\alpha_c \mid \Omega_T, \Phi_{-\alpha_c} \sim N\left(\frac{V_\kappa \sum_{it} U_{ict}^{****} + K'X_c^1}{I_c T_c V_\kappa + 1}, \frac{V_\kappa}{I_c T_c V_\kappa + 1}\right) \quad \forall c = 1, \ldots, C,$$

$$K \mid \Omega_T, \Phi_{-K} \sim N_2\left(\mathrm{Var}_C\left[\left(X^{1'} \cdot \frac{1}{V_\kappa}\right)\alpha + \Sigma\kappa^{-1} \cdot K_0\right], \mathrm{Var}_C\right),$$

$$\mathrm{Var}_C = \left(\left(X^{1'}X^1 \cdot \frac{1}{V_\kappa}\right) + \Sigma_\kappa^{-1}\right)^{-1},$$

$$X^1 = \mathrm{stack}(X_c^{1'}), \quad \alpha = \mathrm{stack}(\alpha_c),$$

$$V_\kappa \mid \Omega_T, \Phi_{-V_\kappa} \sim \mathrm{IG}\left(\rho_\kappa + \frac{C}{2}, R_\kappa + \frac{1}{2}(\alpha - X^1K)^T(\alpha - X^1K)\right).$$

References

Aaker, D. A., D. E. Bruzzone. 1985. Causes of irritation in advertising. *J. Marketing* **49**(2) 47–57.

Aitchinson, J. 1999. *Cutting Edge Advertising: How to Create the World's Best Print for Brands in the 2lst Century.* Prentice Hall, New York.

Albert, J. H., S. Chib. 1993. Bayesian analysis of binary and polychotomous response data. *J. Amer. Statist. Assoc.* **88**(422) 669–679.

Arnheim, R. 1988. *The Power of the Center: A Study of Composition in the Visual Arts*. University of California Press, Berkeley.

Baker, W. E., H. Honea, C. A. Russell. 2004. Do not wait to reveal the brand name: The effect of brand-name placement on television advertising effectiveness. *J. Advertising* **33**(3) 77–85.

Bass, F. M., N. Bruce, S. Majumdar, B. P. S. Murthi. 2007. Wearout effects of different advertising themes: A dynamic Bayesian model of the advertising-sales relationship. *Marketing Sci.* **26**(2) 179–195.

Berlyne, D. E. 1971. *Aesthetics and Psychobiology*. Appleton-Century-Crofts, New York.

Billio, M., R. Casarin, D. Sartore. 2007. Bayesian inference on dynamic models with latent factors. G. L. Mazzi, G. Savio, eds. *Growth and Cycle in the Eurozone.* Palgrave Macmillan, Basingstoke, Hampshire, UK, 25–44.

Bolls, P. D., D. D. Muehling, K. Yoon. 2003. The effect of television commercial pacing on viewers' attention and memory. *J. Marketing Comm.* **9**(1) 17–28.

Book, A. C., C. D. Schick. 1997. *Fundamentals of Copy and Layout*. NTC/Contemporary Publishing Group, Lincolnwood, IL.

Brown, S. C., F. I. M. Craik. 2000. Encoding and retrieval of information. E. Tulving, F. I. M. Craik, eds. *The Oxford Handbook of Memory*. Oxford University Press, Oxford, UK, 93–107.

Bryce, W. J., R. F. Yalch. 1993. Hearing versus seeing: A comparison of learning of spoken and pictorial information in television advertising. *J. Current Issues Res. Advertising* **15**(1) 1–20.

Burns, J. J., D. R. Anderson. 1993. Attentional inertia and recognition memory in adult television viewing. *Comm. Res.* **20**(6) 777–799.

Calvo, M. G., P. J. Lang. 2004. Gaze patterns when looking at emotional pictures: Motivationally biased attention. *Motivation Emotion* **28**(3) 221–243.

Carter, C. K., R. Kohn. 1994. On Gibbs sampling for state space models. *Biometrika* **81**(3) 541–553.

Chib, S. 1995. Marginal likelihood from the Gibbs output. *J. Amer. Statist. Assoc.* **90**(432) 1313–1321.

Cronin, J. J. 1995. In-home observations of commercial avoidance behavior. *J. Current Issues Res. Advertising* **17**(2) 69–75.

Donderi, D. C. 2006. Visual complexity: A review. *Psych. Bull.* **132**(1) 73–97.

Duchowski, A. T. 2003. *Eye Tracking Methodology: Theory and Practice*. Springer-Verlag, London.

d'Ydewalle, G., G. Desmet, J. Van Rensbergen. 1998. Film perception: The processing of film cuts. G. Underwood, ed. *Eye Guidance in Reading and Scene Perception*. Elsevier Science, Amsterdam, 357–367.

Fazio, R. H., P. M. Herr, M. C. Powell. 1992. On the development and strength of category–brand associations in memory: The case of mystery ads. *J. Consumer Psych.* **1**(1) 1–13.

Feichtinger, G., R. F. Hartl, S. P. Sethi. 1994. Dynamic optimal control models in advertising: Recent developments. *Management Sci.* **40**(2) 195–226.

Feinberg, F. M. 2001. On continuous-time optimal advertising under S-shaped response. *Management Sci.* **47**(11) 1476–1487.

Frühwirth-Schnatter, S. 1994. Data augmentation and dynamic linear models. *J. Time Ser. Anal.* **15**(2) 183–202.

Gamerman, D. 1998. Markov chain Monte Carlo for dynamic generalized linear models. *Biometrika* **85**(1) 215–227.

Germeys, F., G. d'Ydewalle. 2007. The psychology of film: Perceiving beyond the cut. *Psych. Res.* **71**(4) 458–466.

Gelfand, A. E., S. E. Hills, A. Racine-Poon, A. F. M. Smith. 1990. Illustration of Bayesian inference in normal data models using Gibbs sampling. *J. Amer. Statist. Assoc.* **85**(412) 972–985.

Greyser, S. A. 1973. Irritation in advertising. *J. Advertising Res.* **13**(1) 3–10.

Grover, R., J. Fine. 2006. The sound of many hands zapping. *BusinessWeek* (May 22), http://www.businessweek.com/magazine/content/06_21/b3985063.htm.

Gustafson, P., S. Siddarth. 2007. Describing the dynamics of attention to TV commercials: A hierarchical Bayes analysis of the time to zap an ad. *J. Appl. Statist.* **34**(5) 585–609.

Hahn, M., J.–S. Hyun. 1991. Advertising cost interactions and the optimality of pulsing. *Management Sci.* **37**(2) 157–169.

Heeter, C., B. S. Greenberg. 1985. Profiling the zappers. *J. Advertising Res.* **25**(2) 15–19.

Janiszewski, C. 1998. The influence of display characteristics on visual exploratory search behavior. *J. Consumer Res.* **25**(3) 290–301.

Krugman, D. M., G. T. Cameron, C. M. White. 1995. Visual attention to programming and commercials: The use of in-home observations. *J. Advertising* **24**(1) 1–12.

Kutner, M. H., C. J. Nachtsheim, J. Neter. 2004. *Applied Linear Regression Models*. McGraw-Hill, New York.

Lachaab, M., A. Ansari, K. Jedidi, A. Trabelsi. 2006. Modeling preference evolution in discrete choice models: A Bayesian state-space approach. *Quant. Marketing Econom.* **4**(1) 57–81.

Lang, A. 2000. The limited capacity model of mediated message processing. *J. Comm.* **50**(1) 46–70.

Lang, A., S. Zhou, N. Schwartz, P. D. Bolls, R. F. Potter. 2000. The effects of edits on arousal, attention, and memory for television messages: When an edit is an edit can an edit be too much? *J. Broadcasting Electron. Media* **44**(1) 94–109.

Lang, A., M. Shin, S. D. Bradley, Z. Wang, S. Lee, D. Potter. 2005. Wait! Don't turn that dial! More excitement to come! The effects of story length and production pacing in local television news on channel changing behavior and information processing in a free choice environment. *J. Broadcasting Electron. Media* **49**(1) 3–22.

Martin, A. D., K. M. Quinn. 2002. Dynamic ideal point estimation via Markov chain Monte Carlo for the U.S. Supreme Court, 1953–1999. *Political Anal.* **10**(2) 134–153.

Mihaylova, M., V. Stomonyakov, A. Vassilev. 1999. Peripheral and central delay in processing high spatial frequencies: Reaction time and VEP latency studies. *Vision Res.* **39**(4) 699–705.

Naik, P. A., M. K. Mantrala, A. G. Sawyer. 1998. Planning media schedules in the presence of dynamic advertising quality. *Marketing Sci.* **17**(3) 214–235.

Palmer, S. E. 1999. *Vision Science: Photons to Phenomenology*. A Bradford Book, Cambridge, MA.

Pavelchak, M. A., M. P. Gardner, V. C. Broach. 1991. Effect of ad pacing and optimal level of arousal on attitude toward the ad. *Adv. Consumer Res.* **18** 94–99.

Perse, E. M. 1998. Implications of cognitive and affective involvement for channel changing. *J. Comm.* **48**(3) 49–68.

Pieters, R., M. Wedel. 2004. Attention capture and transfer in advertising: Brand, pictorial, and text-size effects. *J. Marketing* **68**(2) 36–50.

Pieters, R., M. Wedel. 2007. Goal control of attention to advertising: The Yarbus implication. *J. Consumer Res.* **34**(2) 224–233.

Pieters, R., M. Wedel, J. Zhang. 2007. Optimal feature advertising design under competitive clutter. *Management Sci.* **53**(11) 1815–1828.

Rayner, K. 1998. Eye movements in reading and information processing: 20 years of research. *Psych. Bull.* **124**(3) 372–422.

Rossi, P. E., G. M. Allenby. 2003. Bayesian statistics and marketing. *Marketing Sci.* **22**(3) 304–328.

Sekhon, J., W. Mebane Jr. 1998. Genetic optimization using derivatives: Theory and applications to nonlinear models. *Political Anal.* **7**(1) 187–210.

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.

Sprott, J. C., J. Bolliger, D. J. Mladenoff. 2002. Self-organized criticality in forest-landscape evolution. *Phys. Lett. A* **297**(3–4) 267–271.

Steinberg, B., A. Hampp. 2007. DVR ad skipping happens, but not always. *Advertising Age* (May 31) http://adage.com/mediaworks/article?article_id=117023.

Stewart, D. W., D. H. Furse. 1986. *Effective Television Advertising*. Lexington Books, Lexington, MA.

Stewart, D. W., S. Koslow. 1989. Executional factors and advertising effectiveness. *J. Advertising* **18**(3) 21–32.

Sueyoshi, G. T. 1995. A class of binary response models for grouped duration data. *J. Appl. Econometrics* **10**(4) 411–431.

Tse, A. C. B., R. P. W. Lee. 2001. Zapping behavior during commercial breaks. *J. Advertising Res.* **41**(3) 25–29.

Wedel, M., R. Pieters. 2000. Eye fixations on advertisements and memory for brands: A model and findings. *Marketing Sci.* **19**(4) 297–312.

Wedel, M., R. Pieters. 2008. A review of eye-tracking research in marketing. *Rev. Marketing Res.* **4** 123–147.

West, M., J. Harrison. 1997. *Bayesian Forecasting and Dynamic Linear Models*. Springer-Verlag, New York.

Wilbur, K. C. 2008. How the digital video recorder changes traditional television advertising. *J. Advertising* **37**(1) 143–149.

Woltman Elpers, J. L. C. M., M. Wedel, F. G. M. Pieters. 2003. Why do consumers stop watching TV commercials?: Two experiments on the influence of moment-to-moment entertainment and information value. *J. Marketing Res.* **40**(4) 437–453.

Woolley, S. 2003. Zap! *Forbes* (September), http://www.forbes.com/free_forbes/2003/0929/076.html.

Yarbus, A. L. 1967. *Eye Movements and Vision*. Plenum, New York.

Yerkes, R. M., J. D. Dodson. 1908. The relation of strength of stimulus to rapidity of habit-formation. *J. Comparative Neurology Psych.* **18**(5) 459–482.