MOMENT-TO-MOMENT OPTIMAL BRANDING IN TV COMMERCIALS:

PREVENTING AVOIDANCE BY PULSING

Thales Teixeira[1], Michel Wedel[2] and Rik Pieters[3]

September 2008

Conditionally accepted at ***Marketing Science***

1 PhD Student in Marketing, Ross School of Business, University of Michigan, Ann Arbor, MI 48109-1234, USA, Tel. (734) 330-5292, teixeira@bus.umich.edu.

2 Professor of Marketing, Robert H. Smith School of Business, University of Maryland, College Park, MD 20742, USA, Tel. (301) 405-2162, Fax. (301) 405-0146, mwedel@rhsmith.umd.edu.

3 Professor of Marketing, Tilburg University, Faculty of Economics and Business, Department of Marketing, P.O. Box 90153, 5000 LE Tilburg, the Netherlands, Tel: + 31 13 466 32 56, Fax. + 31 13 466 28 75, f.g.m.pieters@uvt.nl.

The authors would like to thank Dominique Claessens for providing the eye tracking data, and Anocha Aribarg and Berk Ataman for providing valuable modeling suggestions. This paper is part of the first author's PhD dissertation in Marketing at the University of Michigan.

MOMENT-TO-MOMENT OPTIMAL BRANDING IN TV-COMMERCIALS:

PREVENTING AVOIDANCE BY PULSING

ABSTRACT

This paper develops a conceptual framework for understanding the impact that branding activity and consumers' concentration of attention have on their moment-to-moment avoidance decisions during television advertising. It formalizes this in a generalized Dynamic Linear Model (gDLM) and estimates it with MCMC methods. Data on commercial avoidance through zapping along with eye tracking on 31 commercials for nearly 2000 participants are used to calibrate the model. New, simple metrics of attention concentration are shown to strongly predict avoidance. Independent of this, central on-screen brand positions but not brand size further raise commercial avoidance. Based on the model estimation, the branding activity under marketing control is optimized for ads in the sample to reduce commercial avoidance. This reveals that pulsing brand presence--while keeping total brand exposure constant--decreases commercial avoidance significantly. Implications for advertising management and theory are addressed.

Key words: Commercial avoidance, branding, attention, gDLM, State Space Model, optimization

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 due to consumers zapping commercials. Commercial avoidance is facilitated by remote controls and by digital video recorders (DVR) that permit consumers to record and replay TV content without having to see all or parts of commercial breaks (Wilbur 2008). 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 US 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 ten years, despite declines in prime-time 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 to optimize brand 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 generalized Dynamic Linear Model (gDLM), which is estimated with MCMC methods. Data on commercial avoidance along with eye-tracking on 31 commercials for nearly 2000 participants are used to calibrate the model. Second, it describes new, simple metrics of consumers' attention concentration based on eye-tracking data and shows that these systematically predict commercial avoidance from moment-to-moment. Third, based on the model estimations, it optimizes 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 pulsing strategies consisting of repeated brief brand insertions.

BRANDING AND ATTENTION EFFECTS

Branding in Commercials

Branding activity is the way in which the brand name and logo 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 (Kosslyn 1975, 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 simultaneously supported by audio (Bryce and Yalch 1993). Prominence endures to the extent that the brand appears more (versus less) frequently and longer (versus shorter) during the commercial.

For consumers, such branding activity provides important information because the brand helps to comprehend ads and learn from them (Curlo and Chamblee 1998, MacInnis and Jaworski 1989, Woltman Elpers et al. 2003). 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 a source of debate in advertising theory, and between marketers and ad agencies trying to balance sales, creativity and others objectives. Some recommend small, unintrusive (Aitchinson 1999) and others 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), late (Fazio et al. 1992) or early-and-late (Rossiter and Percy 1997, Stewart and Koslow 1989).

There is evidence that under conditions of forced exposure--when consumers cannot avoid watching the commercials--early (Baker et al. 2004) and late (Fazio et al. 1992), more frequent and longer branding (Stewart and Furse 1986) can improve comprehension, recall and persuasion. This is in line with basic memory research that frequency and duration contribute to learning (Brown and Craik 2000). Also, Bryce and Yalch (1993) found that, under forced exposure, video-transmitted content is much better learned than the same content in audio, resulting in an 8-to-1 advantage in recall tests after a single exposure. This is consistent with picture superiority effects on memory (Childers and Houston 1984). However, consumers in practice do have increasing control over commercial exposure, which is important. When consumers stop watching commercials before their natural end, later branding activity in the commercial cannot have the beneficial effects reported for forced exposure conditions. We are not aware of research that has examined the influence of the momentary prominence of brands such as due to their size and centrality in television commercials on commercial avoidance. If and how branding activity in commercials impacts on consumers' moment-to-moment avoidance decisions remains as yet largely unknown, and our purpose is to shed further light on this issue.

Television commercials have the characteristics of a narrative (Deighton 1985), a story that communicates how a brand is relevant to solving consumer needs. The narrative has hedonic and informational components to convey the brand message and at the same time entertain and retain consumers. Brands convey information, and high levels of brand prominence are thus liable to increasing the likelihood of commercial avoidance because of information overload (Woltman Elpers et al. 2003). Moreover, increased levels of branding activity decrease the "soft sell" and increase the "hard sell" character of commercials, and people generally resist the forceful persuasion that comes with the hard sell (Aaker and Bruzzone 1985, Greyser 1973). 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 and centrality) and dynamic (frequency and duration) characteristics of branding activity have on this likelihood. 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

Like in the visual arts, advertising tries to focus and direct viewers' attention. It aims to point attention at certain parts of the depicted scene, and direct it across scenes in an orchestrated fashion to let the intended narrative unfold (Solso 1994). We propose that to the extent that commercials are able to concentrate consumers' attention they are better able to retain them behaviorally as well thus preventing commercial avoidance. This is consistent with art theory's (Arnheim 1988) emphasis on "centers of gravity" that 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. Simply stated by 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) distinguished two types of visual attention that an individual viewer can express during perception of artful stimuli, termed specific and diversive exploration, and speculated 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 stimulus characteristics (Itti 2005). 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 due to the overriding effects of consumers' idiosyncratic 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 concentrated his eye-fixations, and that eye-fixations were most dispersed under a free viewing instruction. Working with print advertising, Pieters and Wedel (2007) found 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 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.

Insert figure 1 about here

Figure 1 summarizes our predictions about the influence of aggregate and individual concentration versus dispersion 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, the higher the likelihood of avoidance by this consumer is expected to be. We predict an interaction effect between aggregate and individual attention concentration, 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 but not the single individual, who wanders off and leaves. These measures of attention concentration capture the extent to which the creative content of commercials is successful in focusing and retaining consumers.

In establishing the net contribution of branding activity on commercial avoidance, we therefore account for these consumers' attention concentration patterns. If attention concentration would predict commercial avoidance independent of branding activity, this would indicate the central function that attention guidance by the creative content of commercials plays in ad effectiveness (Tellis et al. 2005 measure important effects of creative cues). To assess the branding and attention effects appropriately, other ad, brand and person characteristics need to be controlled for. We focus on potentially important, objective ad characteristics that may co-vary moment-to-moment with branding and attention.

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) (Donderi 2006), and the pacing of scenes across the commercial (dynamic) (Germeys and d'Ydewalle 2007). Visual complexity refers to all non-representational perceptual information, such as colors, lines, contrasts (Itti 2005). Pacing indicates the speed at which different scenes are presented in dynamic stimuli (Lang 2000). Pacing 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).

Visual complexity 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 at high levels too 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 translate into avoidance being lowest at those levels here). Berlyne (1971) observed a similar pattern in research on the rated appreciation of paintings varying in levels of visual complexity. This effect has been shown on the appreciation of other stationary stimuli as well (Donderi 2006). We extend this by studying avoidance decisions for dynamic visual stimuli.

In addition, product category, hedonic versus utilitarian, and brand familiarity, low versus high, are controlled for (Pieters and Wedel 2004, Rossiter and Percy 1997). 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 these ad, brand and person characteristics, branding activity in commercials and attention concentration of consumers jointly influence the moment-to-moment commercial avoidance decisions of consumers. Before specifying our 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. These advertised known (Citroen, T-Mobile) and unknown (Radio 538, KWF), national (Albert Hein, Unox) and international (Mastercard, Kodak) brands, from a variety of different product categories (food, durables, public and services, electronics, telecom, clothing), with utilitarian (checking account) and hedonic (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 1998 regular television viewers (age 20 to 62, 48.3% male) and consumers of the advertised products, who were paid for participation. Their demographics matched those of the target population. 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, participants were led to a non-distracting room and seated in a comfortable chair at approximately 55 cm distance of a 21-inch LCD monitor, with a 1280 x 1024 pixel resolution. The instruction on the screen asked people 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. 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). Our 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. Our avoidance rates are higher than other similar reports (Krugman et al 1995, Siddarth and Chattopadhyay 1998), but lower than reported on current DVR usage patterns (Wilbur 2008).

Infrared corneal-reflection eyetracking 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 non-obtrusive to the participant, allowing for head movements within normal boundaries (about 30 x 30 x 30 cm) while facing the television screen. Spatial precision of data collection was 0.5 degrees 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 ms frames, which results in an average of 750 consecutive frames (moments) for every 30 second commercial.

Measures

Commercial avoidance. The dependent variable is 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 for 31 distinct television commercials, each of a maximum of 750 ad frames, for a total of 1998 consumers. In the econometrics literature the term used is "unbalanced panel data".

Branding activity. Branding activity 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 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, for instance because it is competing with other scene objects or occluded by them (Janiszewski 1998). "Mode" indicates whether the brand was present (1) in audio mode or not (0) in a particular frame. "Cardinality" captures how many times a brand appears non-consecutively in video mode during a specific commercial up to that point, from the first (1) to last (n) brand appearance. Finally, "duration" indicates how long in seconds a brand with the same cardinality was present consecutively in video mode up to that point.

Control Variables. The amount 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, JPG, PDF formats, have been developed in computer vision research to enable different hard- and software to use the same data. To the extent that the visual images contain little visual detail, color, contrast, and contain many redundancies, the algorithms cause larger compressions (Sprott et al. 2002). This makes file size a suitable general measures 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 not (0) in each frame of the commercials (Lang 1990) using video editing software. Cuts are due to changing camera positions between scenes, and edits due to changing camera positions within scenes, and both increase complexity, because viewers need to integrate the visual information across discontinuities. Cuts and edits have larger complexity 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 96% for brand and 78% for product, with disagreements resolved by discussion).

Data Aggregation

Data processing and analysis is challenging with 750 frames for each of 31 commercials for which eye-movement data are available for a total of 1998 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 seconds ads. In this duration, an eye-fixation and decision can be made (Calvo and Lang 2004, Rayner 1998), and it 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.

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:

$$\mathrm{U}_{\mathrm{ict}}^{\mathrm{avoid}} = \mathrm{D}_{\mathrm{ict}}^{\mathrm{avoid}} + \varepsilon_{\mathrm{ict}}^{\mathrm{avoid}}, \quad \text{with} \quad \varepsilon_{\mathrm{ict}}^{\mathrm{avoid}} \sim \mathrm{N}(0,1) \tag{1}$$

where *i* is individual, *c* is commercial and *t* is time-frame of the ad. Thus the probability that individual *i* avoids commercial *c* at time-frame *t*, given parameters Θ_t, is:

$$\mathrm{P}(\mathrm{y}_{\mathrm{ict}} = 1 \mid \Theta_{\mathrm{t}}) = \Phi(\mathrm{D}_{\mathrm{ict}}) \\ \text{where:} \quad \mathrm{y}_{\mathrm{t}} = \begin{cases} 1 \rightarrow \text{avoid at frame t} \\ 0 \rightarrow \text{watch at frame t} \end{cases} \tag{2}$$

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

$$\mathrm{D}^{\mathrm{avoid}}_{\mathrm{ict}} = \mu_{\mathrm{i}} + \alpha_{\mathrm{c}} + \mathrm{B}_{\mathrm{ct}} + \left(\gamma^1 \overline{\mathrm{AC}_{\mathrm{ct}}} + \gamma^2 \mathrm{AC}_{\mathrm{ict}} + \gamma^3 \overline{\mathrm{AC}_{\mathrm{ct}}} \times \mathrm{AC}_{\mathrm{ict}}\right) + \mathrm{TVC}_{\mathrm{ct}}\,. \tag{3}$$

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 random effects shrinkage is warranted by the large number of individuals and commercials in a similar fashion to Gustafson and Siddarth (2007). Details are in appendix 1.

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

$$\mathrm{B}_{\mathrm{t}} = \theta^1_{\mathrm{t}}\mathrm{Presence}_{\mathrm{t}} + \theta^2_{\mathrm{t}}\mathrm{Cardinality}_{\mathrm{t}} + \theta^3_{\mathrm{t}}\mathrm{Duration}_{\mathrm{t}} + \theta^4_{\mathrm{t}}\mathrm{Size}_{\mathrm{t}} + \theta^5\mathrm{Mode}_{\mathrm{t}} + \theta^6\mathrm{Position}_{\mathrm{t}} + \theta^7\mathrm{Separation}_{\mathrm{t}}, \\ \text{with } \tilde{\theta}_{\mathrm{t}} = \Xi + \mathrm{G}\tilde{\theta}_{\mathrm{t}-1} + \omega_{\mathrm{t}}\,, \text{ and } \tilde{\theta}_{\mathrm{t}} = (\theta^1_{\mathrm{t}}, \theta^2_{\mathrm{t}}, \theta^3_{\mathrm{t}}, \theta^4_{\mathrm{t}}) \tag{4}$$

Because branding activity may build up irritation over the exposure to the ad 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, $\tilde{\theta}_t$. The ad elements and factors that affect the dynamics of continued attention to TV 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). We believe branding to be one of such factors.

The fourth term (in parenthesis) in equation 3 reflects the attention concentration 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 concentration ($\overline{AC}_{ct}$) 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 ($\overline{AC} = 0$), 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 concentration (AC_{ict}) for each consumer i, commercial c and time-frame t. AC_{ict} ranges from 0 to 2686976 ($1280^2 + 1024^2$). Thus, we have:

$$\begin{aligned}&\text{Aggregate Attention Concentration}: \overline{\mathrm{AC}}_{ct} = \frac{1}{N}\sum_{i=1}^{N}\left(f_{ict} - \tfrac{1}{N}\sum_{i=1}^{N} f_{ict}\right)'\left(f_{ict} - \tfrac{1}{N}\sum_{i=1}^{N} f_{ict}\right)\\ &\text{Individual Attention Concentration}: \mathrm{AC}_{ict} = \left(f_{ict} - \tfrac{1}{N}\sum_{i=1}^{N} f_{ict}\right)'\left(f_{ict} - \tfrac{1}{N}\sum_{i=1}^{N} f_{ict}\right)\end{aligned} \quad , \tag{5}$$

The parameters γ^1, γ^2 and γ^3 (eq. 3) capture the effects of these attention concentration measures and their interaction. The final term TVC_{ct} in equation (3) 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 (IC_{ct} + IC_{ct-1}) in the event of an edit or cut ($Pacing_{ct} = 1$) or the image complexity of the current frame otherwise ($Pacing_{ct} = 0$). The quadratic term of visual complexity in equation (6) allows for a U-shaped effect on avoidance likelihood. Finally, *PaceType* is a dummy variable indicating a cut (=1) or edit (= 0).

$$\begin{aligned}\mathrm{TVC}_{ct} &= \beta^0 \mathrm{PaceType} + \beta^1(\mathrm{VC}_{ct}) + \beta^2(\mathrm{VC}_{ct})^2, \text{with}\\ \mathrm{VC}_{ct} &= \mathrm{IC}_{ct} + \mathrm{Pacing}_{ct}\cdot \mathrm{IC}_{ct-1}\end{aligned} \tag{6}$$

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

We develop a generalized Dynamic Linear Model (gDLM) (Gamerman 1998, West and Harrison 1997), by rewriting equations (1) through (6) in a State Space formulation as in equation (7). For details, see Appendix 1. DLM models have been used in contexts of advertisement scheduling (Bass et al. 2007; Naik et at. 1998), and in dynamics of innovation effects on firm sales (van Heerde et al. 2004).

$$\begin{aligned} &\mathrm{f}\left(\mathrm{E}\left[\mathrm{Y}_{\mathrm{t}}\right]\right)=\mathrm{F}_{\mathrm{t}}\Theta_{\mathrm{t}}+\varepsilon_{\mathrm{t}}, \\ &\Theta_{\mathrm{t}}=\mathrm{G}\Theta_{\mathrm{t}-1}+\omega_{\mathrm{t}} \end{aligned} \tag{7}$$

Y is the commercial avoidance indicator variable; *f* is the probit link function; *E*[] is the expectation operator; $\Theta_{\mathrm{t}}=\{\mu_{\mathrm{i}},\alpha_{\mathrm{c}},\widetilde{\theta}_{\mathrm{t}},\theta^5,\theta^6,\theta^7,\beta^0,\beta^1,\beta^2,\gamma^1,\gamma^2,\gamma^3\}$ is 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; ε,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 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). The MCMC chains are run for 60,000 iterations on 1998 viewers, 31 commercials, and 125 time-frames, totaling 293k observations. The posterior distributions of the parameters of 1750 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 are in Appendix 1. Analysis of synthetic data with the MCMC algorithm shows good recovery of all true parameter values (Appendix 2).

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. 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 (gender and age) and brand (familiarity) and product (utilitarian) variables. Model 2 includes the visual complexity measures in addition to the variables in Model 1, and it outperforms this as shown by the higher LML. Model 3 includes the attention concentration variables in addition to the variables in Model 2, and it outperforms the latter. In Model 4 the branding activity variables are added to Model 2, and it outperforms that model. Finally, the full Model 5 clearly performed best 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 contribute significantly to predicting commercial avoidance, and that branding effects contribute significantly to predicting commercial avoidance even when all other effects are accounted for (Model 5 versus Model 3). Likewise, the attention concentration measures contribute significantly to predicting commercial avoidance, even when all other effects are accounted for (Model 5 versus Model 4).

Insert table 1 and 2 about here

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 5. 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. Yet, when brands were simultaneously present in audio mode, as opposed to just video or no brand, probabilities to avoid the commercial decreased marginally.

Insert table 3 about here

In clear support of our predictions, consumers' attention concentration 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 all consumers' attention simultaneously, increased consumers' probability to stop viewing the commercial. Also, consumers who fail to look where all other consumers concentrate their attention have 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.

The predicted U-shaped effect of visual complexity on avoidance emerged as well, as reflected in the significant effect of complexity-squared and the insignificance of linear term. This is the first time evidence is provided for an "optimum level" of visual complexity for commercials at which avoidance is minimal, while 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 towards the end. This in itself is a reassuring result because it indicates that there is no point in time, apart from start and finish, when viewers systematically tend to stop viewing more, and which is not accounted for by other covariates in the model. Presence of the brand predominantly increases the avoidance probability 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. No strong significant effects emerged for brand cardinality. Higher cardinality of brand presence decreased avoidance towards the second half (marginally significant). Just the opposite effect emerged for duration: longer brand duration increase avoidance in the middle (marginally significant) with the effect dying out towards the end.

Insert figure 2 about here

Since variables were standardized, we can compare their relative importance directly. This shows that the order of importance is (1) attention concentration metrics, with a combined posterior (absolute) mean effect of 1.50, (2) the branding variables, with a total posterior (absolute) mean effect of 0.49, (3) the visual complexity measures with 0.10, (4) the product-brand control variables (brand familiarity and product category) with 0.04 and, (5) the demographic control variables (age and gender) with 0.02.

OPTIMIZATION OF BRANDING ACTIVITY

Marketing managers try to maximize the prominence of their brands in commercials, for instance, by exposing them early, in the middle of the screen, separated from the rest and long, but at the same time try to maximize the likelihood of retaining consumers, which is a difficult trade-off. That is, they aim to maximize the opportunity-to-see the brand, across viewers and time, for a minimum pre-defined level of brand 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 agency's job 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 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 whole screen would always be covered with the brand. In reality, the 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.

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 non-avoidance as zero (0), we have:

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

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:

$$\mathrm{OTS}_c(\Theta) = \frac{1}{R} \sum_{r=1}^{R} \frac{1}{N_i N_t} \sum_{i=1}^{I_c} \sum_{t=1}^{T_c} P\left(\hat{Y}_{ict} = 0 \mid \Theta^{(r)}\right) \qquad \forall c = 1,\ldots,C \tag{9}$$

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 post-production 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 post-production, and the same goes for the position of the brand. Therefore, we optimize brand presence and size, as constant, and cardinality and duration, as dynamic branding features, with all other variables remaining unchanged from their current values.

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, since 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 back to the probability space. Equation 10 describes the optimization problem:

$$\min_{\left\{\substack{\text{size}_t \\ \text{presence}_t}\right\}} \sum_{r=1}^{R}\sum_{i=1}^{I_c}\sum_{t=1}^{T} D_{cit}^{zap} \big/ \Theta^{(r)} \qquad \forall c = 1,\ldots,C$$
$$\begin{aligned} &\text{with} \quad \text{presence}_t \in \{0,1\} \\ &\text{with} \quad \text{size}_t \in [0.5\%, 75\%] \\ &\text{with} \quad \text{BAL} = \text{BAL}_c \pm 5\% \end{aligned} \tag{10}$$

It is noteworthy that the above objective function is linear in presence_t and size_t but has non-linear constraints, and thus may not possess corner solutions. The solution to this optimization problem is a 2 (presence and size) x 125 (total number of time-frames) matrix for each of 31 commercials.

We perform the optimization using a combination of a gradient method and a genetic algorithm (Sekhon and Mebane Jr. 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. While 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 non-linear constraints (BAL_c).

“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, duration of the brand. The combination of the effects of these variables will dictate if and when brands increase or decrease avoidance likelihood. In 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 instants, 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 (75%)

and for smallest brand size possible (0.5%). Since 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 towards 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. Loosely speaking, for each time period, if the duration parameter is larger than the cardinality parameter, then adding a new non-adjacent brand will decrease avoidance. And similarly, if the opposite occurs, then adding a brand in a subsequent frame is more desirable. These parameters are plotted on the rightmost graph of Figure 3 which shows that, for a predefined branding level, non-consecutive 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.

Insert figure 3 about here

Thus, we expect that the optimization results for these commercials should indicate that brands be larger towards the beginning (not in first 1 second) and in the very end, while in the middle portion size will not be critical; and brand appearances be short and frequent in the first 4/5th of the commercial, and then change to longer brands in the last 1/5th.

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 4 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 caused by increases in cardinality, as predicted in the previous section. Total duration (sum of number of frames with brands) is decreased for those ads with comparatively high original total duration and increased for those with comparatively low total duration. Also, if duration is decreased from the original to the improved version, then size is increased, and visa versa. Thus, managers need to make concomitant changes and trade-offs in branding duration and size.

Insert table 4 about here

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

Insert figure 4 about here

The improved solutions have frequent but brief brand-appearances, reminiscent of pulsing. This result is analogous to the finding of pulsing in the advertising effectiveness literature (Feichtinger et al. 1994, Feinberg 2001). It is due to the linearity of the model and the mixed continuous and discrete decision variables, combined with the fact that discrete adjacent brand placements have higher “cost” than non-adjacent ones (Hahn and Hyun 1991).

To validate our findings, we compare in Table 5 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 5 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--which shows again that brands are avoided by viewers, but that brand pulsing reduces this.

Insert table 5 about here

DISCUSSION

Branding activity in commercials has significant, negative effects on the moment-to-moment likelihood that commercials retain consumers. Frequent appearances of brands for sustained periods, in particular when centrally on the screen, increase the likelihood that consumers stop watching commercials notably. However, our optimization procedure suggests that what consumers find intrusive is not the amount of branding per se. It is long sustained brand appearances that may have bothered most. Thus, we were able to lower commercial avoidance rates by merely changing the pattern of brand activity, while keeping the brand activity level per commercial the same. A pulsing strategy where brands are shown more frequently but briefly instead of infrequently but longer, decreased commercial avoidance rates. Avoidance rates under this optimized branding strategy were even lower than if the brands would not have appeared in the commercials at all. This suggests a parallel between optimal effectiveness of brand pulses over time within commercials and commercial pulses over time in a campaign (Feichtinger et al. 1994, Feinberg 2001). 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 brands completely at the end and long, as reflected in the growing incidence of soft selling, mystery commercials. But his may have detrimental effects on memory and learning (Brown and Craik 2000). Our findings show that intrusiveness can be reduced more, without sacrificing brand activity level, using a brand pulsing strategy. The gDLM and optimization approach on which the present findings are based hold the promise of improving the effectiveness of television advertising through the insights it provides 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. And since adaptations can be made post-production, improvements are virtually costless.

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 concentration), 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 (individual attention concentration), the lower the likelihood of commercial avoidance was. The interactive effect between the two attention concentration metrics suggests that higher aggregate and individual attention concentration 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 or not. The findings support that, indeed, as often speculated upon in advertising, the power to orchestrate attention is a crucial condition to advertising effectiveness. The proposed attention concentration metrics can be readily derived by eye-tracking of commercials and may prove useful in other advertising effectiveness research as well.

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, both under 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. The proposed model allows marketing managers and advertisers objective tests of 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.

More than ever, consumers can easily avoid commercials at any point in time. This provides new opportunities to use avoidance decisions in advertising testing before and during campaigns (Wilbur 2008). The proposed procedure that relates objective characteristics of commercials and attention-metrics obtained through eye-tracking to consumers' moment-to-moment avoidance decisions holds the promise to increase television advertising's effectiveness.

Research Opportunities

One of the limitations of our research design is that viewers watched sequences of back-to-back commercials without programs, and this may have increased the likelihoods of avoidance decisions relative to those obtained under natural viewing conditions at home. The research design was chosen to maximize internal validity and control over commercial exposure conditions, so that these were comparable. Although there is no reason to expect that the research context may have prompted consumers to become sensitive to qualitatively different factors than at home, only future research can provide definitive answers to this question.

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 concentration 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 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 concentration in commercials.

Finally, in our optimized commercials, the 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). While it wasn't directly a decision variable, this proposed number of non-consecutive brand insertions is far from the maximum cardinality observed in our dataset, 6, but it is not uncommon in advertising practice. Examples are the recent commercial "The Happiness Factory" for Coca-Cola (Wieden+Kennedy 2006) with cardinality of 17 (short version), and the 2008 Coco-Cola "Super-bowl commercial", with a cardinality of 13. Although the average avoidance rates of these commercials is unknown, it exemplifies that the prescribed strategy is possible and that high levels of brand pulsing are being used by successful firms. Follow-up research might proceed to construct the proposed optimal commercials, and examine and compare their avoidance likelihoods from moment-to-moment using the procedure proposed here. Such future research may further inform managers of ways to maximize brand exposure while preventing TV-commercial avoidance.

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., 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, CA.

Baker, W., H. Honea, C. Russell. 2004. Do not wait to reveal the brand name. *J. Advertising* **33**(3) 77-85.

Bass, F., N. Bruce, S. Majumdar, B. P. S. Murthi. 2007. Wearout effects of different advertising themes: A dynamic Bayesian model of the ad-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 in dynamic models with latent factors. *Monography of Official Statistics*. Forthcoming.

Bolls, P. D., D. D. Muehling, K. Yoon. 2003. The effect of television commercial pacing on viewers' attention and memory. *J. Marketing Communications* **9** 17-28.

Book, A. C., C. D. Schick. 1997. *Fundamentals of Copy & 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. of Current Issues and Res. in Advertising* **15**(1) 1-20.

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

Carter, C., R. Kohn. 1994. On Gibbs sampling for state space models. *Biometrika* **81** 541-553.

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

Childers, T. L., M. J. Houston. 1984. Conditions for a picture-superiority effect on consumer memory. *J. Consumer Res.* **11**(Sept) 643-654

Curlo, E., R. Chamblee. 1998. Ad processing and persuasion: The role of brand identification. *Psychology and Marketing* **15**(3) 279-299.

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

Deighton, J. 1985. Rhetorical strategies in advertising. *Advances in Consumer Res*. **12** 432-436.

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

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

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, NL 357-367.

Germeys, F., G. d'Ydewalle. 2007. The psychology of film: Perceiving beyond the cut. *Psychological Res.* **71** 458-466.

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. Hartl, S. Sethi. 1994. Dynamic optimal control models in advertising: Recent developments. *Management Sci*. **40**(2) 195-226.

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

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

Gamerman, D. 1998. Markov Chain Monte Carlo for dynamic generalized linear models. *Biometrika* **85** 215-227.

Gelfand, A., 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. *Business Week* (May), 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. Applied Statist.* **34**(5) 585-609.

Hahn, M., J. 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** 15-19.

Itti, L. 2005. Quantifying the contribution of low-level saliency to human eye movmenets in dynamic scenes. *Visual Cognition* **12**(6) 1093-1123.

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

Kosslyn, S. 1975. Information representation in visual images. *Cognitive Psych.* **7** 341-370.

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** 1-12.

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** 57-81.

Lang, A. 1990. Involuntary attention and physiological arousal evoked by structural features and emotional content in TV commercials. *Communication Res.* **17** 275-299.

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

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. of Broadcasting & Electronic Media* **49**(1) 3-22.

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. of Broadcasting & Electronic Media* **44** 94-109.

MacInnis, D., B. Jaworski. 1989. Information processing from advertisements: Toward an integrative framework. *J. Marketing* **53** 1-23.

Martin, A., K. 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.

Naik, P. A., M. K. Mantrala, A. G. Sawyer. 1998. Planning media schedules in the presence of dynamic advertising quality. *Marketing Sci.* **17**(3) 214–35.

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. *Advances in Consumer Res.* **18** 94-99.

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

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

Pieters, R., M. Wedel, J. Zhang. 2007. Optimal feature advertising design under competitive clutter. *Management Sci.* Forthcoming.

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

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

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

Sekhon, J., W. Mebane Jr. 1998. Genetic optimization using derivatives: Theory and applications to nonlinear models. *Political Anal.* **7** 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., J. Bolliger, D. Mladenoff. 2002. Self-organized criticality in forest-landscape evolution. *Physics Lett. A* **297**(3-4) 267-271.

Steinberg, B., A. Hampp. 2007. DVR ad skipping happens, but not always. *Advertising Age* (May 31).

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** 21-32.

Tellis. G. J., R. K. Chandy, D. MacInnis, P. Thaivanich. 2005. Modeling the microeffects of television advertising: Which ad works, when, where, for how long, and why? *Marketing Sci.* **24**(3) 359-366.

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

van Heerde, H., C. F. Mela, P. Manchanda. 2004. The dynamic effect of innovation on market structure. *J. Marketing Res.* **41** 166–183.

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. *Review of Marketing Res*. **4** 123-147.

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

Wieden+Kennedy. 2006. The happiness factory. http://www.coca-cola.com/HF/index.jsp.

Wilbur, K.C. 2008. How the digital video recorder changes traditional television advertising. *J. Advertising*, Forthcoming.

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.

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

APPENDIX 1

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) through (6):

$$
\begin{aligned}
&\mathrm{D}_{\mathrm{ict}} = \mu_{\mathrm{i}} + \alpha_{\mathrm{c}} + \mathrm{B}_{\mathrm{ct}} + \left(\gamma^1 \overline{\mathrm{AC}}_{\mathrm{ct}} + \gamma^2 \mathrm{AC}_{\mathrm{ict}} + \gamma^3 \overline{\mathrm{AC}}_{\mathrm{ct}} \times \mathrm{AC}_{\mathrm{ict}}\right) + \mathrm{TVC}_{\mathrm{ct}} \\
&\mathrm{B}_{\mathrm{ct}} = \theta_{\mathrm{t}} \cdot \mathrm{Branding}_{\mathrm{ct}} \\
&\qquad \theta_{\mathrm{t}} = \left(\theta_{\mathrm{t}}^1, \theta_{\mathrm{t}}^2, \theta_{\mathrm{t}}^3, \theta_{\mathrm{t}}^4, \theta^5, \theta^6, \theta^7\right)^{\mathrm{T}} \sim \mathrm{N}(\theta^*, \Sigma^*) \\
&\mathrm{TVC}_{\mathrm{ct}} = \beta^0 \cdot \mathrm{PaceType} + \beta^1 \cdot \mathrm{Visual\,Complexity}_{\mathrm{ct}} + \beta^2 \cdot \mathrm{Visual\,Complexity}_{\mathrm{ct}}^2 \\
&\qquad \beta = \left(\beta^0, \beta^1, \beta^2\right)^{\mathrm{T}} \\
&\left(\gamma^1 \overline{\mathrm{AC}}_{\mathrm{ct}} + \gamma^2 \mathrm{AC}_{\mathrm{ict}} + \gamma^3 \overline{\mathrm{AC}}_{\mathrm{ct}} \times \mathrm{AC}_{\mathrm{ict}}\right) = \gamma^1 \mathrm{Aggreg.Conc.}_{\mathrm{ict}} + \gamma^2 \mathrm{Indiv.Conc.}_{\mathrm{ct}} + \gamma^3 \mathrm{Aggreg.Indiv.Conc.}_{\mathrm{ct}} \\
&\qquad \gamma = \left(\gamma^1, \gamma^2, \gamma^3\right)^{\mathrm{T}} \sim \mathrm{N}(\gamma^*, \Gamma^*) \\
&\mu_{\mathrm{i}} = \Lambda^1 \cdot \mathrm{Age}_{\mathrm{i}} + \Lambda^2 \cdot \mathrm{Gender}_{\mathrm{i}} + \mathrm{V}_\lambda \\
&\qquad \Lambda \sim \mathrm{N}(\Lambda_0, \Sigma_\Lambda) \\
&\alpha_{\mathrm{c}} = K^1 \cdot \mathrm{Pr\,oductCategory}_{\mathrm{c}} + K^2 \cdot \mathrm{BrandFamiliarity}_{\mathrm{c}} + \mathrm{V}_\kappa \\
&\qquad K \sim \mathrm{N}(K_0, \Sigma_K)
\end{aligned}
$$

Let, time t = 1,…,T, commercial c = 1,…,C and individual i = 1,…,I. The basic relationship of equations (1) and (2) form the complete Utility Model specification and are expressed as:

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

The State Space (gDLM) 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):

$$
\begin{aligned}
&U_{ict} = \mu_i + \alpha_c + F_{ct}\Theta_t + X_{ict}^3 \Psi + \varepsilon_{ict} \\
&\Theta_t = \Xi + G\Theta_{t-1} + \omega_t \\
&\mu_i \sim N(X_i^2 \Lambda, V_\lambda) \\
&\alpha_c \sim N(X_c^1 \mathrm{K}, V_\kappa) \\
&\omega_t \sim N(0, V_\omega) \qquad \varepsilon_{ict} \sim N(0,1)
\end{aligned}
$$

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

$$\mathrm{U}_{\mathrm{ict}} = \mathrm{D}_{\mathrm{ict}} + \varepsilon_{\mathrm{ict}}$$

$$\underbrace{\mathrm{U}_{\mathrm{ict}}}_{\text{Probit}} = \underbrace{(\mathrm{X}_{\mathrm{i}}^2 \Lambda + \lambda_{\mathrm{i}}) + (\mathrm{X}_{\mathrm{c}}^1 K + \kappa_{\mathrm{c}})}_{\text{HB: Variance component model}} + \underbrace{\mathrm{F}_{\mathrm{ct}}(\mathrm{G}^{\mathrm{t}}\Theta_0 + \sum_0^{\mathrm{t}-1} \mathrm{G}^{\mathrm{j}}\omega_{\mathrm{t-j}})}_{\text{Forward Filter Backward Sample}} + \underbrace{\mathrm{X}_{\mathrm{ict}}^3 \Psi}_{\text{Bayes Regresssion}} + \varepsilon_{\mathrm{ict}}, \quad \begin{array}{l} \varepsilon_{ict} \sim N(0,1) \\ \lambda_i \sim N(0, V_\lambda) \\ \kappa_c \sim N(0, V_\kappa) \\ \omega_t \sim N(0, V_\omega) \end{array}$$

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

$$\mathrm{X} = \{\mathrm{F}_{\mathrm{ct}} \quad \mathrm{X}_{\mathrm{c}}^1 \quad \mathrm{X}_{\mathrm{i}}^2 \quad \mathrm{X}_{\mathrm{ict}}^3\} = \left(\begin{bmatrix} 1 \\ \mathrm{Presence}_{\mathrm{ct}} \\ \mathrm{Cardinality}_{\mathrm{ct}} \\ \mathrm{Duration}_{\mathrm{ct}} \\ \mathrm{Size}_{\mathrm{ct}} \end{bmatrix} \begin{bmatrix} \mathrm{ProductCategory}_{\mathrm{c}} \\ \mathrm{Brand\ Familiarity}_{\mathrm{c}} \end{bmatrix} \begin{bmatrix} \mathrm{Age}_{\mathrm{i}} \\ \mathrm{Gender}_{\mathrm{i}} \end{bmatrix} \begin{bmatrix} \mathrm{Mode}_{\mathrm{ct}} \\ \mathrm{Position}_{\mathrm{ct}} \\ \mathrm{Separation}_{\mathrm{ct}} \\ \mathrm{PaceType}_{\mathrm{ct}} \\ \mathrm{VisualComplexity}_{\mathrm{ct}} \\ \mathrm{VisualComplexity}_{\mathrm{ct}}^2 \\ \mathrm{Indv.Concen.}_{\mathrm{ict}} \\ \mathrm{Aggreg.Concen.}_{\mathrm{ct}} \\ \mathrm{Indv.*Aggreg.Concen.}_{\mathrm{ict}} \end{bmatrix} \right)$$

The prior distribution of parameters are diffuse conjugate distributions:

$$\begin{array}{l}
\mathrm{U}_{\mathrm{ict}}, \mu_{\mathrm{i}}, \alpha_{\mathrm{c}} \rightarrow \text{specified from model} \\
\Theta_0 \sim \mathrm{N}_5\left(\mathrm{m}_0 = 0, \mathrm{C}_0 = 10^5 \mathrm{I}\right) \\
\Psi \sim \mathrm{N}_9\left(\mathrm{n}_0 = 0, \mathrm{S}_0 = 10^5 \mathrm{I}\right) \\
\Lambda \sim \mathrm{N}_2\left(\Lambda_0 = 0, \Sigma_\Lambda = 10^5 \mathrm{I}\right) \\
K \sim \mathrm{N}_2\left(K_0 = 0, \Sigma_K = 10^5 \mathrm{I}\right) \\
\mathrm{V}_\lambda^{-1} \sim \mathrm{G}\left(\rho_\lambda = 2+1, (\rho_\lambda \mathrm{R}_\lambda)^{-1} = (3 \cdot 0.0001)^{-1}\right) \\
\mathrm{V}_\kappa^{-1} \sim \mathrm{G}\left(\rho_\kappa = 2+1, (\rho_\kappa \mathrm{R}_\kappa)^{-1} = (3 \cdot 0.0001)^{-1}\right) \\
\mathrm{V}_\omega^{-1} \sim \mathrm{W}_5\left(\rho_\omega = 5+1, (\rho_\omega \mathrm{R}_\omega)^{-1} = (6 \cdot 0.0001 \cdot \mathrm{I}_5)^{-1}\right)
\end{array}$$

In order to estimate the unique observation equation via Gibbs sampling, let $\Phi = \{\Theta_0, ..., \Theta_{\mathrm{T}}, \Psi, \mu_{\mathrm{i}=1}, ..., \mu_{\mathrm{i}=1}, \Lambda, \mathrm{V}_\lambda, \alpha_{\mathrm{c}=1}, ..., \alpha_{\mathrm{c}=\mathrm{C}}, K, \mathrm{V}_\kappa, \mathrm{V}_\omega\}$ 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.

1. Probit (Albert and Chib 1993)

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

2. Forward Filtering-Backward Sampling (Fruhwirth-Schnatter 1994, Lachaab et al. 2006)

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

$V_{\varepsilon,t} = V_\varepsilon \otimes I_{C_t I_t}$.

Forward Filter: Loop forward in time and sample Normal distributions

$$\begin{aligned}
&\mathit{\Theta}_t \mid \mathit{\Omega}_{t-1}, \mathit{\Phi}_{-\mathit{\Theta}_t} \sim \mathrm{N}_5(\mathrm{m}_t, \mathrm{C}_t), \qquad \forall t = 1,\ldots,T \\
&\gamma_t = \mathit{\Xi} + \mathrm{G}\mathrm{m}_{t-1}, \\
&\mathit{\Gamma}_t = \mathrm{G}\mathrm{C}_{t-1}\mathrm{G}^{\mathrm{T}} + \mathrm{V}_\omega \\
&\mathrm{C}_t^{-1} = \mathit{\Gamma}_t^{-1} + \widetilde{\mathrm{F}}_t^{\mathrm{T}}\mathrm{V}_{\varepsilon,t}^{-1}\widetilde{\mathrm{F}}_t \\
&\mathrm{m}_t = \mathrm{C}_t(\mathit{\Gamma}_t^{-1}\gamma_t + \widetilde{\mathrm{F}}_t^{\mathrm{T}}\mathrm{V}_{\varepsilon,t}\widetilde{\mathrm{U}}_t^*)
\end{aligned}$$

With dimensions G = 5 × 5, Ξ = 5 × 1 , Θ = (5×125) × 1, $F_t = (C_t \times I_t) \times 5$, $U_t^* = (C_t \times I_t) \times 1$,

$\gamma_t = 5 \times 1$, $\Gamma_t = 5 \times 5$, $C_t = 5 \times 5$, $m_t = 5 \times 1$.

Backward Sampler: Loop backwards in time and sample Normal distributions

$$\Theta_T \mid \Omega_T, \Phi_{-\Theta_T} \sim N_5(m_T, C_T)$$

$$\begin{aligned}
&\mathit{\Theta}_t \mid \mathit{\Theta}_{t+1}, \mathit{\Omega}_{t-1}, \mathit{\Phi}_{-\mathit{\Theta}_t} \sim \mathrm{N}_5(\mathrm{q}_t, \mathrm{Q}_t), \qquad \forall t = T-1,\ldots,0 \\
&\mathrm{Q}_t^{-1} = \mathrm{C}_t^{-1} + \mathrm{G}^{\mathrm{T}}\mathrm{V}_\omega^{-1}\mathrm{G} \\
&\mathrm{q}_t = \mathrm{Q}_t[\mathrm{C}_t^{-1}\mathrm{m}_t + \mathrm{G}^{\mathrm{T}}\mathrm{V}_\omega^{-1}(\mathit{\Theta}_{t+1} - \mathit{\Xi})]
\end{aligned}$$

With dimensions: $Q_t = 5 \times 5$, $q_t = 5 \times 1$

3. Conjugate sampling (Lachaab et al. 2006)

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

4. Bayesian Regression

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

$$\Psi \mid \Omega_T, \Phi_{-\Psi} \sim N_9\left(\mathrm{M}_\Psi, V_\Psi\right)$$
$$V_\Psi = (\widetilde{X}^{3}{}'\widetilde{X}^{3} + S_0^{-1})^{-1}$$
$$\mathrm{M}_\Psi = V_\Psi(\widetilde{X}^{3}{}'\widetilde{U}^{**} + S_0^{-1}n_0)$$

5. HB: Variance Component Model (Gelfand et al. 1990)

Individual-specific baseline intercepts

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

$$\mu_i \mid \Omega_T, \Phi_{-\mu_i} \sim \mathrm{N}\left(\frac{\mathrm{V}_\lambda \sum_{ct} \mathrm{U}^{***}_{ict} + \Lambda' \mathrm{X}^2_i}{\mathrm{C}_i \mathrm{T}_i \mathrm{V}_\lambda + 1}, \frac{\mathrm{V}_\lambda}{\mathrm{C}_i \mathrm{T}_i \mathrm{V}_\lambda + 1}\right), \forall \mathrm{i} = 1,\ldots,\mathrm{I}$$
$$\Lambda \mid \Omega_T, \Phi_{-\Lambda} \sim \mathrm{N}_2\left(\mathrm{V_n}[(\mathrm{X}^2{}' \cdot \frac{1}{\mathrm{V}_\lambda})\mu + \Sigma^{-1}_\lambda \cdot \Lambda_0],\ \ \mathrm{Var_n}\right)$$
$$\mathrm{Var_n} = \left((\mathrm{X}^2{}'\mathrm{X}^2 \cdot \frac{1}{\mathrm{V}_\lambda}) + \Sigma^{-1}_\Lambda\right)^{-1}$$
$$\mathrm{X}^2 = \mathrm{stack}(\mathrm{X}^2_i{}'),\ \ \mu = \mathrm{stack}(\mu_i)$$
$$\mathrm{V}_\lambda \mid \Omega_T, \Phi_{-\mathrm{V}_\lambda} \sim \mathrm{IG}\left(\rho_\lambda + \tfrac{1}{2}, \mathrm{R}_\lambda + \tfrac{1}{2}(\mu - \mathrm{X}^2\Lambda)^{\mathrm{T}}(\mu - \mathrm{X}^2\Lambda)\right)$$

Commercial-specific baseline intercepts

Let $U_{ict} - \hat{\mu}_i - F_{ct}\hat{\Theta}_t - X^3_{ict}\hat{\Psi} = U^{****}_{ict} = \alpha_c + \varepsilon_{ict}$

$$\alpha_c \mid \Omega_T, \Phi_{-\alpha_c} \sim \mathrm{N}\left(\frac{\mathrm{V}_\kappa \sum_{it} \mathrm{U}^{****}_{ict} + K' \mathrm{X}^1_c}{\mathrm{I}_c \mathrm{T}_c \mathrm{V}_\kappa + 1}, \frac{\mathrm{V}_\kappa}{\mathrm{I}_c \mathrm{T}_c \mathrm{V}_\kappa + 1}\right), \forall \mathrm{c} = 1,\ldots,\mathrm{C}$$
$$K \mid \Omega_T, \Phi_{-K} \sim \mathrm{N}_2\left(\mathrm{V_C}[(\mathrm{X}^1{}' \cdot \frac{1}{\mathrm{V}_\kappa})\alpha + \Sigma^{-1}_\kappa \cdot K_0],\ \ \mathrm{Var_C}\right)$$
$$\mathrm{Var_C} = \left((\mathrm{X}^1{}'\mathrm{X}^1 \cdot \frac{1}{\mathrm{V}_\kappa}) + \Sigma^{-1}_\kappa\right)^{-1}$$
$$\mathrm{X}^1 = \mathrm{stack}(\mathrm{X}^1_c{}'),\ \ \alpha = \mathrm{stack}(\alpha_c)$$
$$\mathrm{V}_\kappa \mid \Omega_T, \Phi_{-\mathrm{V}_\kappa} \sim \mathrm{IG}\left(\rho_\kappa + \tfrac{\mathrm{C}}{2}, \mathrm{R}_\kappa + \tfrac{1}{2}(\alpha - \mathrm{X}^1 K)^{\mathrm{T}}(\alpha - \mathrm{X}^1 K)\right)$$

APPENDIX 2

PARAMETER RECOVERY

To assess accuracy of parameter recovery we simulate a dataset with the same variables as in the model, generated randomly with identical mean and variance as in our dataset, for eight individuals, six ads and twelve time-frames, totaling 368 observations. The gDLM model was run with this simulated dataset for an additional 50,000 iterations, after a burn-in of 50,000, taking samples of the parameter posteriors at every 50 iterations. All parameter priors were conjugate, diffuse, as specified in Appendix I. Table A1 provides the most important estimates (aggregate). All hyper parameters estimated were recovered within a 95% confidence interval. The Gibbs sampler converged well before the burn-in, generating identifiable parameters with fast decaying autocorrelation and good mixing properties, recovering the true parameter values well.

Table A1: Key parameters with true values and posterior percentiles recovered

		Posterior percentile				
Parameter	True value	2.5%	5.0%	50.0%	95.0%	97.5%
Presence	-1.5	-1.745	-1.712	-1.593	-1.480	-1.458
Mode	-1.1	-1.291	-1.271	-1.156	-1.053	-1.035
Position	-0.7	-0.814	-0.796	-0.702	-0.607	-0.589
Separation	-0.3	-0.421	-0.408	-0.318	-0.234	-0.212
VisualComplexity	0.5	0.163	0.197	0.487	0.743	0.798
VisualComplexity²	0.9	0.846	0.862	0.963	1.075	1.098
Indv. Concentration	1.3	1.255	1.269	1.369	1.480	1.510
Aggr Concentration	1.7	1.612	1.641	1.757	1.897	1.938
Aggr.x Indiv. Concen.	2.1	2.009	2.034	2.167	2.317	2.340
Intercept	-0.5	-0.704	-0.684	-0.539	-0.388	-0.355
Cardinality	1.0	0.837	0.883	1.050	1.242	1.288
Duration	2.0	1.822	1.854	2.008	2.182	2.204
Size	2.0	1.853	1.873	2.034	2.207	2.240
Brand familiarity	-1.0	-1.306	-1.257	-1.077	-0.896	-0.864
Product category	-2.0	-2.443	-2.391	-2.170	-2.007	-1.986
Age	-1.0	-1.418	-1.367	-1.099	-0.833	-0.778
Gender	-2.0	-2.328	-2.268	-1.963	-1.643	-1.581
V_Intercept	1E-3	5.0E-4	5.9E-4	2.0E-3	8.3E-3	1.1E-2
V_Cardinality	1E-3	6.3E-4	7.2E-4	2.7E-3	1.3E-2	1.9E-2
V_Duration	1E-3	6.0E-5	7.0E-5	2.4E-4	1.5E-3	2.0E-3
V_Size	1E-3	6.0E-5	7.0E-5	2.7E-4	1.6E-3	2.3E-3
V_alpha	0.1	0.085	0.093	0.169	0.336	0.382
V_mu	0.1	0.049	0.055	0.099	0.192	0.217

TABLE 1
SUMMARY OF THE INDEPENDENT VARIABLES

Variable	Variation across units	Mean	Std. Dev.	Minimum	Maximum
Branding activity:					
Presence (present = 1)	ad, time	22%	41.2%	0	1
Size (% 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 concentration:					
Aggregate concentration (pixels^2)	ad, time	104212	486780	2434	7311820
Individual concentration (pixels^2)	ad, time, indv.	147	289	0	27478
Aggregate × Indiv. concentration	ad, time, indv.	32256972	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
Pacing type* (cut = 1)	ad, time	44.4%	49.7%	0	1
Visual complexity (Kbytes)	ad, time	180	69	2	662
Visual complexity^2	ad, time	37156	32628	4	438244

*Conditional on a camera shot change.

TABLE 2
MODEL COMPARISONS

	Independent Variables Included				
Model	Demographics, product-brand	Visual Complexity	Attention Concentration	Branding Activity	LML
1	Yes	No	No	No	-11231
2	Yes	Yes	No	No	-10021
3	Yes	Yes	Yes	No	-9966
4	Yes	Yes	No	Yes	-9690
5	Yes	Yes	Yes	Yes	-9639

TABLE 3

DETERMINANTS OF COMMERCIAL AVOIDANCE

Parameter	Mean	SE	Percentiles of the Posterior Distribution 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 concentration:							
Aggregate concentration	0.055^{**}	0.021	0.013	0.027	0.057	0.078	0.085
Individual concentration	0.199^{**}	0.011	0.181	0.185	0.200	0.214	0.218
Aggregate × Indiv. concentration	-1.249^{**}	0.108	-1.377	-1.355	-1.284	-1.054	-1.038
Control variables:							
Age	-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
Pacing type (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

Note: ** indicates 95% posterior confidence interval doesn't contain zero; * indicates 90% confidence interval doesn't contain zero; $^{+}$ indicates 90% posterior confidence interval doesn't contain zero for some time periods.

TABLE 4

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*
5. Mona	46.2	1.02	3	29	3.5	43.4	17	34	3.2	6.0
6. Citroen	49.2	1.67	4	44	3.8	46.0	16	43	3.7	6.5*
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*
10. Mona	48.9	2.31	2	21	11.0	46.7	13	23	10.5	4.5*
11. Unox	47.6	5.54	2	29	19.1	42.5	24	39	14.3	10.7*
12. Telefoon-gids	49.8	0.61	2	21	2.9	48.8	5	20	3.2	2.0*
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*
18. Electro World	49.3	10.9	4	43	25.4	44.7	29	48	22.9	9.4
19. Master-card	47.1	7.02	1	21	33.4	41.4	27	37	19.1	12.1*
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*
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*
28. Min. of Justice	59.4	0.38	1	11	3.5	57.4	11	21	1.9	3.5*
29. Wadden	41.1	10.5	1	32	32.8	33.8	30	41	26.4	17.7
30. PostBank	55.5	12.1	1	77	15.8	47.5	22	67	18.5	14.4*
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. = Duration of the brand appearance; * indicates that the overall post-optimization avoidance reductions for these ads are larger than their estimation error.

TABLE 5

ESTIMATED COMMERCIAL AVOIDANCE FOR CURRENT, OPTIMIZED AND BENCHMARK BRANDING STRATEGIES

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

FIGURE 1

ATTENTION CONCENTRATION AND COMMERCIAL AVOIDANCE

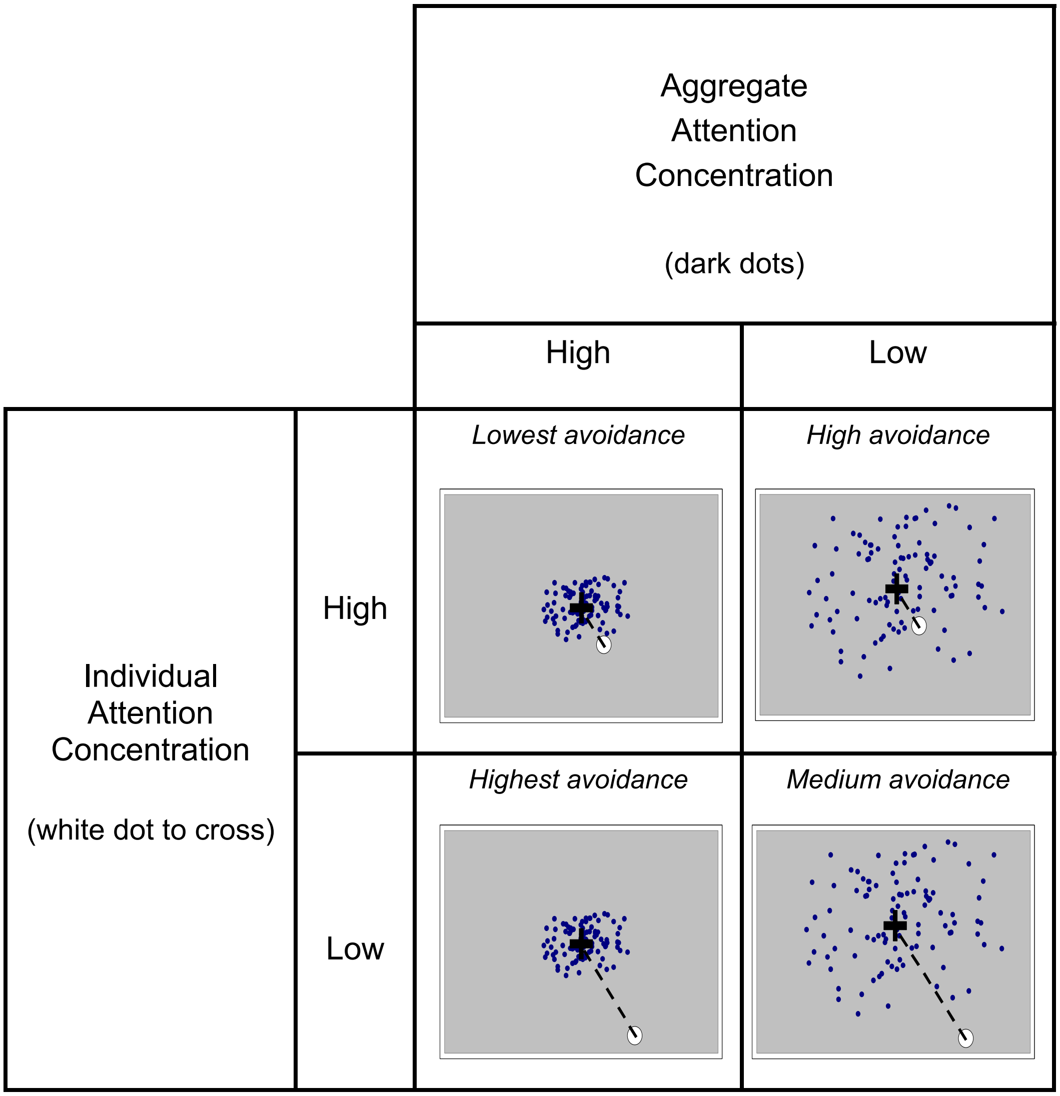


FIGURE 2

FOR TIME-VARYING PARAMETERS OF BRANDING ACTIVITY:
POSTERIOR MEDIAN AND 90% CONFIDENCE BANDS

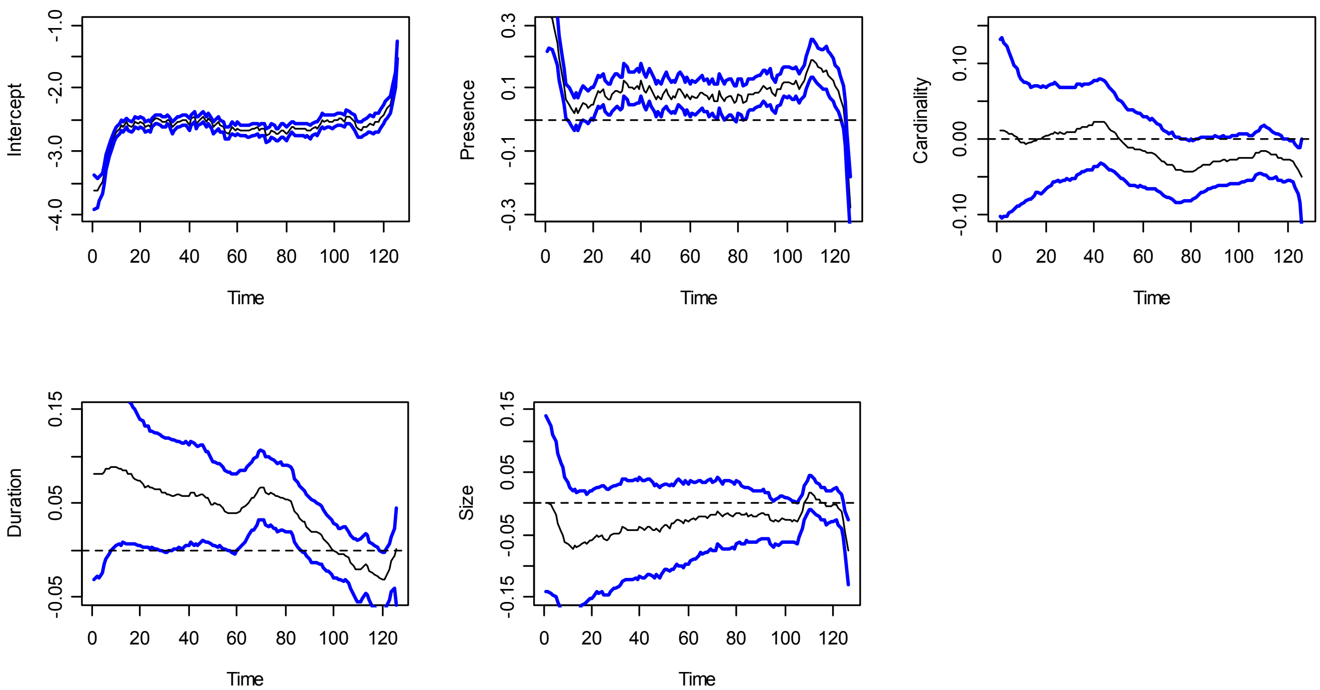
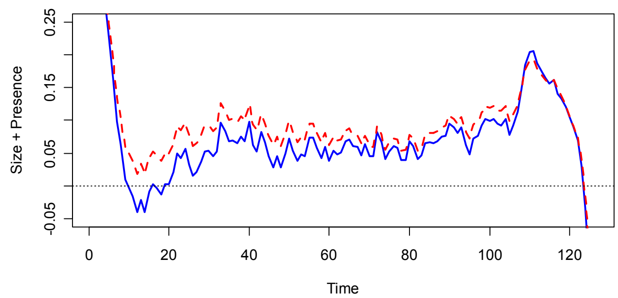


FIGURE 3

TIME VARYING PARAMETERS

Presence plus small (jagged line)
and large (continuous line) brand



Duration (continuous line)
and cardinality (jagged line)

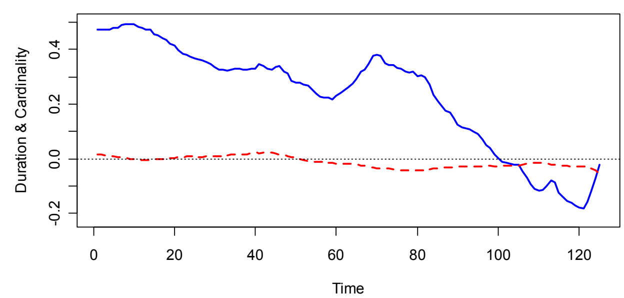


FIGURE 4

ILLUSTRATING BRANDING OPTIMIZATION FRAME-BY-FRAME

(plots of brand presence and size for six optimized ads. Upper graph is original and lower graph is optimized ad; light thick line is brand presence and dark thin line is size)