David A. Schweidel & Robert J. Kent

# Predictors of the Gap Between Program and Commercial Audiences: An Investigation Using Live Tuning Data

Audience measures serve as the basis on which billions of dollars of television advertising are purchased each year. However, these measures often fall short of providing marketers and media planners with information about the size of the audience that has the opportunity to see advertisements during commercial breaks. Although program audience measures remain relevant for product placement, the size of the gap between the audience potentially exposed to programming and the audience exposed to commercial breaks has implications for advertisers and networks, affecting decisions such as program selection and ad pricing. The authors use a full television season of live tuning data to study variation across programs in the size of the program audience and the size of the gap between the potential program audiences and the potential commercial audiences, which is found to exceed 20% of the program audience in some cases. Across program genres, the authors find that dramas have increased program popularity and reduced ad avoidance, whereas reality television experiences increased ad avoidance. The analysis reveals the importance of incorporating show-specific random effects because their omission can result in spurious attribution of differences in ad avoidance and program popularity to genre. The authors discuss the implications for networks and advertisers.

*Keywords*: advertising, television, ad avoidance, media, marketing metrics



Television advertising has traditionally been purchased on the basis of program audience measures. These metrics are used both to select the programs in which to advertise and to negotiate the rates for billions of dollars in advertising spending. While measures of the program audience are appropriate for marketers considering branded entertainment (i.e., product placement), program ratings have shortcomings in advertising planning. Chief among them is the discrepancy between measures of the audience with the opportunity to see the program and measures of the audience with the opportunity to see the program's commercial breaks. Advertisers are not primarily concerned with program audiences but rather with the audience that particular commercials may reach.

According to Poltrack (2006), "the relationship between program audiences and commercial audiences was found to be very stable over time and by program." If this is the case, established program audience measures may be effective proxies for the potential commercial audience. As such, media planners and their clients can simply use these metrics to choose programs in which to place commercials. However, the appropriateness of program audience measures for such decisions is more tenuous if the relationship between viewers' opportunities to see programs and their opportunities to see commercial breaks varies significantly across programs or within a given episode.

Steinberg and Hampp (2007) report that program audiences may be approximately 5%–10% higher than potential commercial audiences. If the price charged for advertising is based on the average program audience, this discrepancy effectively increases the cost of reaching a given number of viewers. As Ephron (2007) notes, given great programmer competition and the billions of dollars spent annually on television advertising, even small-sounding differences in the audience size can have important effects on ad rates and marketers' satisfaction with ad buys: "the networks would have killed for five percent this year." Gloede (2006) argues that advertisers will angle for lower rates and/or other compensation, such as messages integrated into programming or ad clutter reductions, if some shows are found to lose more of the program audience when commercials appear (e.g., Ephron 2006; Steinberg 2005; Steinberg and Barnes 2006).

With nearly $70 billion in television advertising spending in 2007 (Wilbur, Goeree, and Ridder 2009) and the average price of 30-second spots in network programs ranging from under $50,000 to more than $400,000 (Steinberg

David A. Schweidel is Assistant Professor of Marketing, School of Business, University of Wisconsin–Madison (e-mail: dschweidel@bus.wisc.edu). Robert J. Kent is Associate Professor of Marketing, Department of Business Administration, University of Delaware (e-mail: kentr@lerner.udel.edu). The authors thank TNS Media Research for providing the data used in this research. The first author acknowledges the financial support of the Center for Brand and Product Management at the Wisconsin School of Business.


ISSN: 0022-2429 (print), 1547-7185 (electronic)

*Journal of Marketing*
Vol. 74 (May 2010), 18–33

and Williams 2007), it is not surprising that more than 80% of surveyed members of the Association of National Advertisers expressed interest in obtaining precise measures of viewers' opportunities to see commercials (Hampp 2007). The industry has responded to this call with metrics that focus on the audience during commercial breaks. However, the latest national audience metrics average across all commercial break minutes within a program episode (e.g., Atkinson 2008), which may disappoint marketers that want information specific to particular advertising units (Atkinson 2007).

While these commercial minute ratings are available at a national level, audience measures for local markets are reported as program ratings (McClellan 2008). Media planners have expressed a desire for commercial audience measures in local markets (e.g., Friedman 2009). As Kaul and Wittink (1995) note, national and local advertising each play an important role in a brand's marketing strategy, with local advertising primarily focusing on product availability and price and national advertising primarily focusing on brand positioning. In developing advertising and pricing strategies, marketers must balance national and local messages (e.g., Ephron 2001; Popkowski-Leszczyc and Rao 1990). Thus, local measures of commercial break audiences are of value to national brands, as well as regional brands and retailers, because they provide information on the reach of their local advertisements and allow them to add efficiency to marketing activities (e.g., Abraham and Lodish 1990).

Program audience measures may include consumers who avoid commercials, and thus these measures may overstate the size of the audience with the opportunity to see advertisements (e.g., Briggs and Stuart 2006; Ephron 2007; Wilbur 2008b). Some media planners have also expressed an interest in maintaining "live-only" audience measures (which do not include digital video recorder [DVR] playback) in the interests of clients with time-sensitive advertisements (e.g., Friedman 2009; McClellan 2008). For example, retailers may time advertising buys and slot relevant ad messages to take advantage of customers' increased tendency to shop on weekends (e.g., Fox and Hoch 2005). Similarly, film studios may advertise new movies with timing designed to attract patrons during their critical opening weekend (e.g., Krider and Weinberg 1998; Lehmann and Weinberg 2000). Such advertisers may find diminished value in the audience reached even a few days later.

While there continues to be discussion about the inclusion of time-shifted viewing in audience measures, the strong interest in and importance of commercial audience measures for local markets is evident. Increasingly available set-top box data offer a means by which to begin addressing this concern (Arango 2008; Kang 2008).[1] Marketers and media firms have expressed keen interest in using such data to understand viewers' opportunities to see programs and commercials. The recently formed Coalition for Innovative Media Measurement, a consortium of advertisers (including Procter & Gamble and Unilever), agencies (including GroupM, Omnicom Media Group, and Starcom MediaVest Group), and content providers (including CBS Paramount, Disney Media Networks, NBC Universal, News Corporation, Time Warner, and Viacom), notes that examination of "the current and future potential of television measurement through set-top box data" is one of two key areas of interest (Reuters 2009; see also Li 2009). As we discuss subsequently, using set-top box data to examine live tuning to programs and commercials offers new insights and opportunities for researchers and practitioners.

The need for a better understanding of commercial and program exposure patterns goes beyond the television advertising market. Marketers continue to increase advertising in mobile and online marketing. Cross-media campaigns (e.g., Naik and Raman 2003) require comparable metrics across different media. If the extent to which the program audience declines during commercial breaks differs substantially across programs, program audience measures are more limited than they might appear. Moreover, if there is considerable variation in lost opportunities to see advertisements across the commercial breaks of a single episode, measures of the average commercial audience across a program may be misleading. Variation in the extent of the program audience decline during commercial breaks may also affect measures of advertising effectiveness and wearout (e.g., Bass et al. 2007; Naik 1999) because program audience measures will tend to overstate viewers' opportunities to see advertisements. In turn, this may affect the optimal scheduling of advertising campaigns (e.g., Bruce 2008; Naik, Mantrala, and Sawyer 1998).

The intent of the current research is twofold. First, we use a full season of set-top box data to examine and contrast live tuning to programs and commercial breaks in prime-time network television. While prior research has separately examined program tuning and ad avoidance, these factors have not been considered together. Comparing the live tuning to programs and commercial breaks, our analysis reveals considerable variation across programs and over the course of the season, as well as variation attributable to commercial break characteristics. These results call into question the notion of a stable relationship between program and commercial break audiences.

Second, we examine the impact of show-specific random effects (colloquially, an "X-factor") on program ratings and ad avoidance. While previous research has often used data from a short time frame and has not considered programs' X-factors, the use of a full season of data enables us to identify the programs that are more (or less) popular than their characteristics would suggest, as well the programs that experience more (or less) ad avoidance. The findings reveal that the impact of these effects on both aspects of tuning is substantial and may outweigh the program genre effects that have been documented previously. Moreover, the omission of these X-factors results in spurious inferences about the popularity of certain genres and the levels of ad avoidance they are expected to experience.

Our analysis offers actionable insights for advertisers and networks. The results suggest complementary use of measures of program and commercial break tuning levels for marketers using product placement and traditional

[1]We discuss set-top box data—both its advantages and limitations—in more detail in our "Data" section.

advertising. Moreover, our analysis reveals that the impact of program genre on the extent of ad avoidance and program popularity found in prior studies may be more limited than previously thought. Taken together, the findings offer insight into the extent to which live program audience measures overstate the size of the audience with the opportunity to see commercials, providing marketers with essential information for choosing programs and negotiating advertising rates. This information is also of value to networks in setting advertising costs for different programs, potentially reducing the need to deliver costly "make goods" as a consequence of overestimating commercial break tuning levels or the chances that they will undercharge advertisers as a result of underestimating commercial break tuning levels.

In the following section, we briefly review related literature on television tuning behavior and ad avoidance during commercial breaks. We then describe the data used in the analysis and the modeling framework. We present the findings and then discuss the implications of the results. We conclude with a discussion of the limitations of this research and directions for future work.

## Related Research

Several models have been developed to understand viewers' tuning behavior. Rust and Alpert (1984) use a flow model to capture viewers' program choices in each half-hour interval. Rust, Kamakura, and Alpert (1992) employ peoplemeter data to understand the preferences of different types of viewers. Shachar and Emerson (2000) identify a set of viewer and programming characteristics (e.g., genre) to model television-viewing behavior. Wilbur (2008b) proposes a two-sided model that considers both the viewing and the advertising markets and finds that the attractiveness of programs for viewers and advertisers is related to show characteristics. Wilbur, Goeree, and Ridder (2009) examine the effects of product placement and advertising times on program popularity. Such analyses have formed the basis for purchasing advertising as marketers consider buying commercial time in programs that are popular with their intended markets. As Shachar and Emerson (2000, p. 177) state, such data "provide the standard of ratings for both network executives and advertising agencies." However, these studies focus on television program choices and do not consider the extent of ad avoidance.

In addition to the literature on program choices, a separate but related stream of research has addressed the gap between program and commercial break audiences. Danaher (1995) uses second-by-second peoplemeter data to compare program and commercial break audience measures on three networks in New Zealand. In constructing his ratio of the commercial break rating to the program rating, he averages ratings across the seconds of the commercial breaks, comparing this average rating with the average rating of programming seconds. He finds that the audience level declines during commercial breaks by approximately 5%, which is consistent with recent reports (e.g., Ephron 2007; Steinberg and Hampp 2007). He shows that there is a considerable range in estimates of the extent of ad avoidance, which is related to the program type, program rating, and commercial break characteristics, such as the number of advertisements in the break. Because the study uses data from a single week, however, it does not consider the impact of show-specific effects on the extent of ad avoidance. Some programs may systematically maintain more (or less) of their program audience during commercial breaks. Although this may be attributable in part to program characteristics, such as the genre, it may also be driven by a program's X-factor.

Van Meurs (1998) examines a panel of households using peoplemeters in the Netherlands during a four-month period. He decomposes channel switching during commercial breaks into its two underlying components: panelists tuning to a commercial break and those tuning away from the break. He finds that commercial break and program characteristics influence both types of channel-switching behavior. As with Danaher (1995), however, Van Meurs (1998) does not consider show-specific effects in his analysis.

Siddarth and Chattopadhyay (1998) also examine the propensity for viewers to switch channels. They find that previous exposure to commercials increases the propensity for viewers to change the channel and decreases the time until it occurs. They also find that viewers are more prone to change the channel if the commercial occurs within two minutes of the half hour or hour. Although the study reveals the drivers of channel-switching behavior for specific commercials, many of which the marketer can control, it does not consider the impact of the program in which the commercial airs. As such, it does not offer guidance for selecting programs (or genres) in which to advertise. Siddarth and Chattopadhyay also do not consider the presence of commercials on other networks, which may affect changes in the audience level.

In studying the effect of DVRs on the advertising industry, Wilbur (2008a) discusses the reasons that viewers may avoid advertising. Among the reasons are the presence of substitute activities, the lack of engagement generated by commercials, advertising wearout, and a lack of interest in the advertised products. He suggests the use of an adjusted CPM (cost per thousand) to take into account ad skipping that is facilitated by DVRs. More generally, advertising decisions, such as pricing and program selection, should consider the overall extent of ad avoidance, whether due to ad skipping, channel switching, or other activities, as well as the popularity of the program. In each case, there may be variation across programs and commercial breaks.

In this research, we model the average program rating in a half-hour period and the extent of ad avoidance during each commercial break using live set-top box tuning data. These two components play a central role in understanding viewers' opportunities to see commercials at the time they are broadcast. While previous research has separately examined the drivers of program popularity and the relative audience levels during commercial breaks, we explore them together to understand how factors may differentially affect both aspects of tuning behavior, which can affect the selection of programs for both traditional advertising and product placement (e.g., Karrh, McKee, and Pardun 2003). In doing so, we demonstrate the applicability of increasingly

available set-top box tuning data to both media planners and networks (see Arango 2008; Li 2009).

## Data

We employ tuning, programming, and advertising data provided by TNS Media Research. Tuning data were collected from digital cable set-top boxes in a major metropolitan area in the United States. The number of digital cable set-top boxes tuned to each of five broadcast networks (ABC, CBS, CW, FOX, and NBC) was collected for weekdays (Monday–Friday), second-by-second from 8 P.M. to 11 P.M., from September 2007 through May 2008. We also collected the total number of set-top boxes in the area, which enabled us to calculate second-by-second ratings (the fraction of all set-top boxes in the area tuned to a particular channel) for each network.

When we compare the data used in the current research with that used in previous studies, two aspects are particularly noteworthy. First, set-top boxes provide tuning data. That is, they record the channel to which the set-top box is tuned each second, thus providing an indication of opportunity to see the programming and commercials at the time they are broadcast.[2] Because this information is recorded passively, data from a large number of set-top boxes can be collected in local markets (more than 400,000 set-top boxes were installed in the metropolitan area we examine).

Several previous studies have employed data from peoplemeters, in which participants must indicate that they are present in the room once or twice per hour (e.g., Wilbur 2008b). Peoplemeters rely on the compliance of participants, which has been criticized recently (Schneider 2009). Wilbur (2008b) employs data collected by audimeters, devices that passively record the program to which a television is tuned and report tuning behavior in 15-minute intervals. Given the relative coarseness of such data, as Wilbur notes, households may avoid advertisements but still be counted as part of the audience. None of these means of data collection (audimeters, peoplemeters, and set-top boxes) directly measure viewers' attention to or engagement with the programming or commercial breaks. As such, tuning data only enable us to draw conclusions regarding the potential for exposure—that is, the opportunities to see programming and commercial breaks (Ephron 2007). Although this is a limitation, households cannot be influenced by commercial breaks they do not tune in to, and thus the difference between the program and the commercial break tuning audiences is important to advertisers and media buyers.

Second, the data we employ do not include DVR usage. Although new streams of set-top box data may include DVR usage, such data were not available at the time of this study. As such, our conclusions regarding ad avoidance during prime time are restricted to "live tuning." In analyzing TiVo log files, Bronnenberg, Dube, and Mela (2009) report that more than 90% of watched shows were viewed live. Of the programs that were recorded, more than two-thirds of commercials were skipped. Consistent with Bronnenberg, Dube, and Mela's findings, a Nielsen (2008) report found that households watched an average of 142.5 hours of television per month, and the amount of time-shifted television was less than 10% of this total. With DVR penetration of 20%–30% at the time of the study, homes with DVRs dividing television use between live and time-shifted viewing, and a high incidence of skipped commercials in time-shifted viewing (e.g., Bronnenberg, Dube, and Mela 2009; Wilbur 2008a), live tuning data reflect the majority of viewers' opportunities to see programs and commercial breaks. As we noted previously, the live audience is also of particular interest to advertisers with time-sensitive messages (e.g., McClellan 2008).

Although set-top box data do not provide information on people's presence in the room, accounting only for live tuning, such data offer distinct advantages. In particular, we are able to obtain tuning data at a granular level, which enables us to examine observed ad avoidance of individual commercial breaks. Thus, data from set-top boxes have the potential to provide the metrics in local markets that advertisers continue to request (McClellan 2008).

Our season-long data span a longer time frame than the data employed in several previous studies. With multiple airings of the same programs, we can examine variation in program popularity and ad avoidance across programs due to unobserved factors, the importance of which has been documented in the marketing literature (e.g., Rossi and Allenby 2003) but not considered in this context. The results reveal the importance of incorporating show-specific random effects when modeling program and commercial break ratings. In particular, conclusions on program genre effects are affected by the inclusion of these show-specific X-factors, with important implications for advertisers and programmers.

The tuning data were paired with programming schedules to identify the show that aired in each half-hour block. Programs were coded along the following dimensions: Programming genres were coded according to the information available at TV.com (e.g., Wilbur 2008b). When multiple genres were listed, we used the first genre listed. Five genres accounted for more than 90% of the half-hour blocks during our observation period: drama (coded as GENRE = 1), reality (GENRE = 2), sitcom (GENRE = 3), news (GENRE = 4), and game shows (GENRE = 5). The remaining genres were aggregated and treated as an "other" genre. Using the information available at TV.com, we also created a dummy variable to indicate whether an episode was a new airing (REPEAT = 0) or a repeat (REPEAT = 1).

Another aspect of the data we control for was the Writers Guild of America (WGA) strike, which ran from November 5, 2007, to February 12, 2008. For each half-hour block, we created a categorical variable to indicate whether the period occurred before the strike (WGA = 1), during the strike (WGA = 2), or after the strike (WGA = 3). In addition to the respective main effects, we constructed an interaction effect between REPEAT and WGA to allow for the possibility that repeat episodes were more (or less) popular among viewers during and after the strike. We

[2]To account for cases in which a set-top box has not demonstrated activity for a prolonged period, our data provider applies a capping algorithm. This is similar to the practice employed in analyzing dwell times on Web sites (e.g., Montgomery et al. 2004).

denote these four programming characteristics (GENRE, REPEAT, WGA, and REPEAT × WGA), along with a network-specific indicator variable, as $X_{dtc}$.

To provide a summary of programs' relative popularities and explore the tuning data, we next examine ratings. We calculate the half-hour program ratings by averaging the second-by-second ratings across programming seconds in each half-hour period (e.g., Danaher 1995). We present the distribution in Figure 1.

Although the majority of half-hour program ratings are in the range of 1%–2% of set-top boxes, we observe some that exceed 5%. We also observe variation across networks, with ABC and CBS having the highest half-hour program ratings (2.70% and 2.53%, respectively), followed by FOX, NBC, and CW (2.25%, 1.80%, and .97%, respectively).

The programs with episodes tuned to by more than 5% of set-top boxes included ABC's *Dancing with the Stars*, *Grey's Anatomy*, *Samantha Who*? and *The Bachelor*; CBS's *CSI: Crime Scene Investigation* and *Without a Trace*; and FOX's *American Idol*, *The Moment of Truth*, and *House*. From this exploratory analysis, we might conclude that reality television and dramas allow advertisers to reach the largest audiences in this market. However, this would be premature for two reasons that we explore in greater detail. First, the popularity of these programs may reflect the appeal of particular shows rather than all shows of the genre. Second, the half-hour program ratings do not consider the level at which tuning changes during commercial breaks, which may vary across programs.

To probe these issues further, we next present second-by-second plots of the ratings for a selection of programs. We first consider two reality television programs that air on CW: *America's Next Top Model* and *Pussycat Dolls Present: Girlicious*. Figure 2, Panel A, exhibits the second-by-second ratings for individual episodes of these programs.

Although both programs are reality television shows that appear on the same network, *America's Next Top Model* draws a larger audience than *Girlicious*. As might be expected, the observed tuning audience drops considerably during the commercial breaks, evidenced by the declines in the second-by-second ratings. Comparing these episodes, we find a difference in the observed levels of ad avoidance. While this episode of *America's Next Top Model* has an average rating during commercial seconds of 1.51% and an average rating during programming seconds of 1.78%, resulting in a difference of 15% between the commercial and the program ratings, this episode of *Girlicious* has an average rating of .95% during commercial seconds and of 1.15% during programming, yielding a difference of 18%.

In Figure 2, Panel B, we present a similar comparison between episodes of the FOX dramas *Prison Break* and *Bones*. In this episode of *Prison Break*, the average rating during programming seconds is 2.06%, compared with its average rating of 1.85% during commercial seconds—a difference of approximately 10%. This episode of *Bones* has an average rating of 1.83% during programming, which drops only 5% to 1.74% during commercials. While *Prison Break* has a tuning audience during programming that is 12% larger than *Bones*' audience, *Prison Break*'s commercial audience is only 6% larger, demonstrating that the program audience measure may not serve as an adequate proxy for opportunity to see commercials. Note also that these programs both experience an audience build over time, and thus the audience levels during commercial breaks within a single episode may vary.

As the two comparisons in Figure 2 reveal, the network and genre may not fully explain the popularity of programs and the extent to which the programs experience ad avoidance. Although variation in the gap between the live program and the commercial break tuning audiences that we observe in these four episodes may simply be due to randomness, after we control for program genre, some shows (in these examples, *Pussycat Dolls Present: Girlicious* and *Prison Break*) may systematically experience more ad avoidance. If this is the case, when pricing advertising time, networks should account not only for differences in the popularity of programs but also for differences in the amount of ad avoidance the programs experience. Advertisers should also take these issues into consideration when planning and buying television ad time.



**FIGURE 1**
**Distribution of Average Half-Hour Program Ratings**

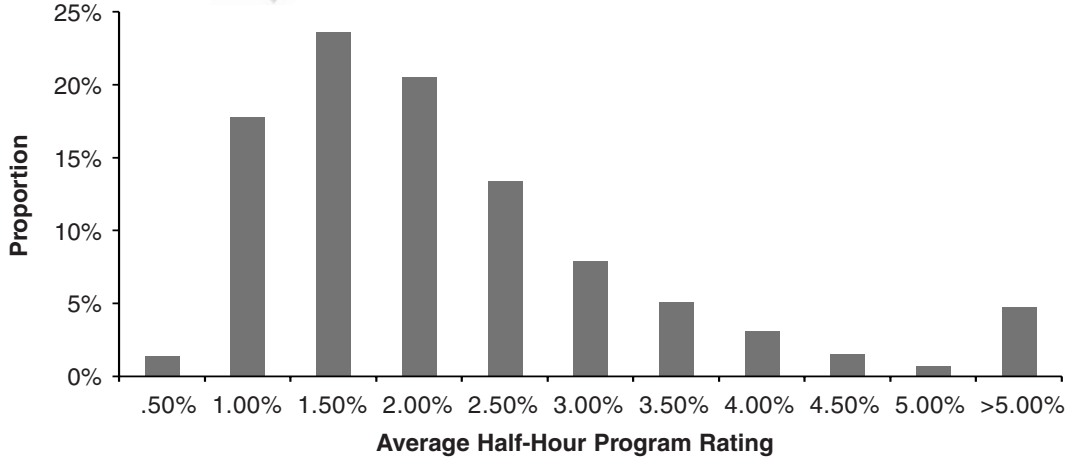
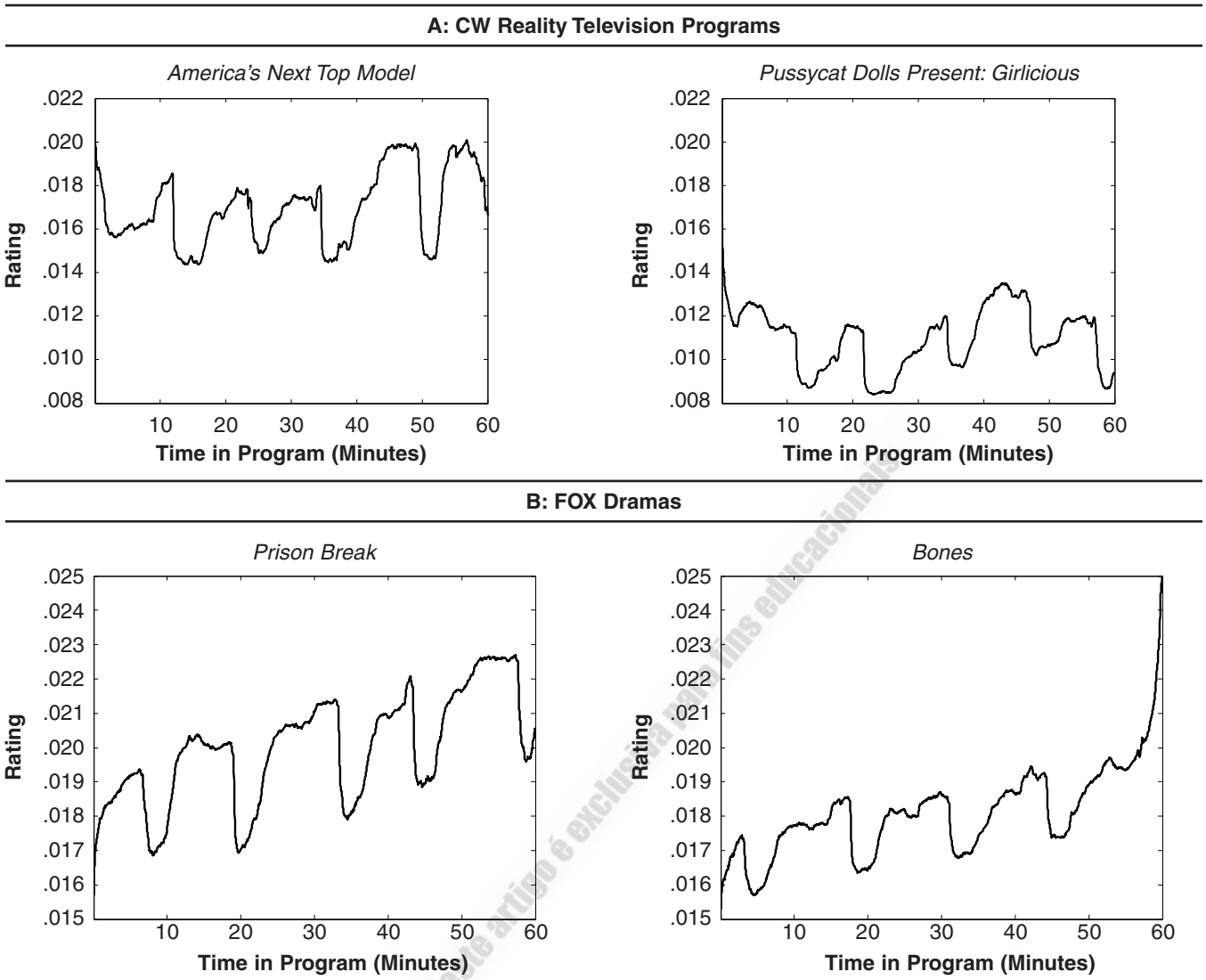


**FIGURE 2**
**Rating Comparisons**



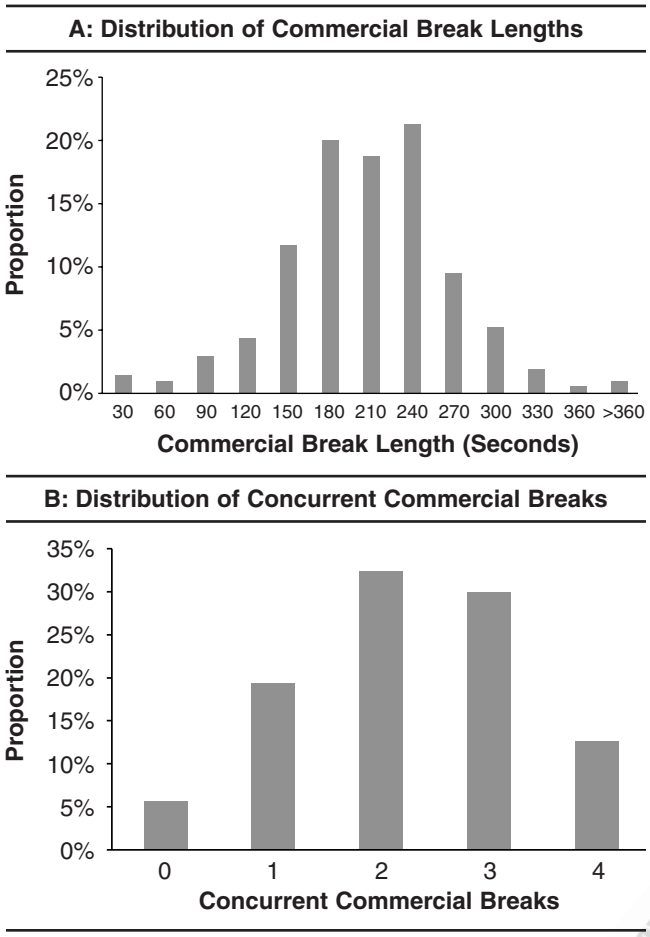

To complement the tuning data, we collected advertising schedules on the five networks. This data consist of the time at which each commercial break started and ended. From this, we constructed variables for the length of each break in minutes (LENGTH), the number of breaks that aired on the remaining four networks concurrently (CONCUR), and an indicator variable for whether the break occurred within two minutes of the half-hour blocks changing (HALFHOUR).[3] We denote these three variables, in conjunction with $X_{dtc}$, as $Q_{jdtc}$, which are the descriptors we use to model the change in audience levels during commercial breaks.

In Figure 3, we provide the distribution of commercial break lengths and the distribution of the number of concurrent commercial breaks. Figure 3, Panel A, shows that the majority of commercial breaks last between three and four minutes. Longer commercial breaks provide viewers with increased opportunity to switch channels without missing the programming, while shorter commercial breaks limit such activity. Figure 3, Panel B, shows that commercial breaks tend to occur at the same times across networks. This overlap may diminish the appeal of switching channels during the commercial breaks. Of all commercial breaks during the observation period, 27% aired around changes in the half-hour blocks. We might contend that breaks near the half hour are subject to more switching because programs are concluding. However, they may exhibit a smaller decline because viewers remain tuned to see the final scenes or previews for next week's episode.

Next, we examine the size of the observed commercial break tuning audience relative to that of the live program audience. We refer to the ratio of the commercial break rating (calculated each second and averaged across the seconds of the commercial break) to the half-hour program rat-

[3]To construct the CONCUR variable for commercial breaks on network c, we examine the advertising schedules on the remaining four networks and count the number of networks with a commercial break that had any overlap with time at which the break aired on network c.

**FIGURE 3**
**Summary of Commercial Break Data**



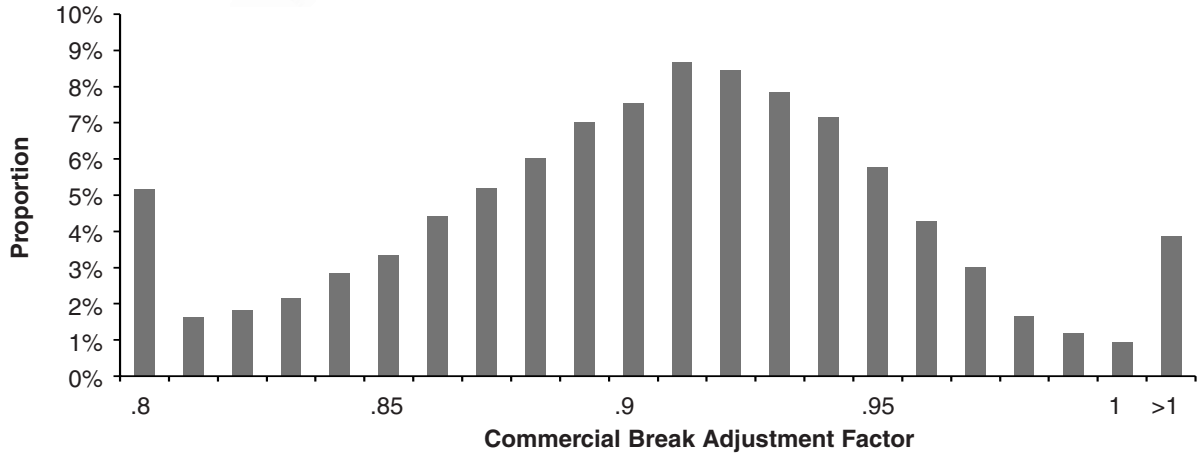

ing (calculated as the average rating of noncommercial seconds of the program) as an "adjustment factor" and present the distribution in Figure 4.

Although we expect that the potential commercial break audience will be smaller than the audience of the program in which it airs (indicated by adjustment factors less than 1), we observe that approximately 4% of commercial breaks have ratings that exceed the program ratings (adjustment factors greater than 1). Such commercial breaks may occur early in programs that display strong audience build, and some advertisers benefit from this effect. We also observe considerable variation in the adjustment factors, with values less than .80 reflecting more than a 20% difference from the program rating. The average adjustment factor across all commercial breaks was .90, reflecting approximately twice as much ad avoidance as Danaher (1995) reports. We also observe slight variation across networks (average for ABC = .91, CBS = .92, CW = .88, FOX = .89, NBC = .90). However, these differences may be influenced by variation in the adjustment factors across programs. Revisiting the programs with high ratings reported previously, we observe that some had higher-than-average adjustment factors (*Grey's Anatomy* = .93, and *CSI: Crime Scene Investigation* = .93), while others did not maintain as much of their tuning audience during commercial breaks (*Samantha Who?* = .87, and *The Bachelor* = .87).

To highlight the importance of considering both program popularity and the change in audience levels during commercial breaks, we provide a comparison of average half-hour program ratings and commercial break ratings for selected programs. Although *Hell's Kitchen* (3.92% program rating, 3.10% commercial break rating) draws a larger program audience than *Survivor: China* (3.11% program rating, 2.78% commercial break rating), the gap between measures of the set-top boxes tuned to their commercial breaks is considerably smaller. While *Hell's Kitchen* draws a larger average program audience than *CSI: Crime Scene Investigation* (3.65% program rating, 3.35% commercial break rating), more set-top boxes are tuned to the commercial breaks during *CSI*. As we discuss subsequently, such reversals (when program appeal is based on commercial versus program ratings) offer opportunities to advertisers and are relevant to networks. To systematically examine the relationship between program and potential commercial break audiences, as well as the drivers of each, we next detail our modeling framework.



**FIGURE 4**
**Distribution of Commercial Break Adjustment Factors**



## Model Development

We first describe a model for the program rating in a half-hour period. We then describe a model for the commercial break adjustment factor.

### *Program Rating Model*

We begin with a model of the program audience rating in a half-hour period on each of the five networks. On each day d, we observe six half-hour measures of ratings from 8 P.M. to 11 P.M. (which we denote by the subscript t) for channel c, which we denote as $R_{dtc}$. Because $R_{dtc}$ is a fraction of set-top boxes, we model the rating $R_{dtc}$ as follows:

$$R_{dtc} = \frac{e^{A_{dtc}}}{1 + \sum_{c'=1}^{5} e^{A_{dtc'}}}, \tag{1}$$

where $A_{dtc}$ indicates the attractiveness of the program airing on channel c during half-hour t on day d. The fraction of set-top boxes not tuned to one of the five networks that we consider (i.e., "the outside good") is then given by the following:

$$1 - \sum_{c'=1}^{5} R_{dtc'} = \frac{1}{1 + \sum_{c'=1}^{5} e^{A_{dtc'}}}. \tag{2}$$

From Equations 1 and 2, we can express $A_{dtc}$ as follows:

$$A_{dtc} = \ln\left(\frac{R_{dtc}}{1 - \sum_{c'=1}^{5} R_{dtc'}}\right) = \ln\left(R_{dtc}\right) - \ln\left(1 - \sum_{c'=1}^{5} R_{dtc'}\right). \tag{3}$$

To understand the factors that affect program attractiveness, we incorporate a set of descriptive variables ($X_{dtc}$) to explain the observed variation in the program rating and to permit correlation among observations with common traits. In addition, programs may have an unobservable X-factor that makes them more attractive. For example, while *American Idol* and *The Next Great American Band* are both reality programs that aired on FOX, *American Idol* is a stronger show that dominates the nights on which it airs. To link both observed and unobserved factors to $A_{dtc}$ (and, thus, $R_{dtc}$), we model $A_{dtc}$ as follows:

$$\overline{A_{dtc}} = \alpha_{h(d,t)} + X'_{dtc} \times \beta + \delta_{p(d,t,c)}, \text{ and} \tag{4}$$

$$A_{dtc} = \overline{A_{dtc}} + \varepsilon_{dtc}, \tag{5}$$

where $\alpha_h$ is a weekday-by-half-hour intercept to reflect different levels in the appeal of network television programming, $\delta_p$ is a program-specific random effect, and $\varepsilon_{dt}$ is an error term that follows a multivariate normal distribution with mean 0 and covariance matrix $\Sigma$.[4] Thus, at its core, our program rating model is a multivariate regression with random effects, which can be replicated with ease and used with any set of program characteristics X (e.g., a program's target demographic).

### *Observed Ad Avoidance During Commercial Breaks*

Given the objective of reaching the audience during commercial breaks, media planners must understand the amount by which the commercial break audience differs from the program audience. While some researchers have suggested that a fairly constant fraction of the audience tunes away during commercial breaks (e.g., Steinberg and Hampp 2007), others find greater variation (e.g., Danaher 1995). To examine the effects of program and commercial break characteristics on the extent of observed ad avoidance, as well as a program's X-factor with respect to ad avoidance, we model an adjustment factor that allows for variation across programs and within a single episode.

Let $B_{jdtc}$ denote the average commercial break rating during the jth commercial break on channel c in half-hour t on day d. If the premise of a constant level of ad avoidance holds, the ratio $B_{jdtc}/R_{dtc}$ would be constant, modulo random error. Allowing for variation from commercial break to commercial break would indicate the following:

$$\log\left(\frac{B_{jdtc}}{R_{dtc}}\right) = -\gamma + \varphi_{jdtc}, \tag{6}$$

where $\varphi_{jdtc}$ is the zero-centered normal error and $\exp(-\gamma + \varphi_{jdtc})$ reflects the extent of ad avoidance. Larger values of $\gamma$ correspond to a larger difference between the program and the commercial break ratings (i.e., increased ad avoidance).

To examine the effects of observable and unobservable characteristics on ad avoidance, as an alternative to Equation 6, we consider an adjustment factor given by the following:[5]

$$\log\left(\frac{B_{jdtc}}{R_{dtc}}\right) = -\gamma_{jdtc} + \varphi_{jdtc}, \text{ and} \tag{7}$$

$$\gamma_{jdtc} = Q'_{jdtc} \times \kappa + \psi_{p(d,t,c)}, \tag{8}$$

where $\kappa$ captures the effects of program and commercial break characteristics and $\psi_p$ is a show-specific random effect.

Live program and commercial break tuning audience metrics are relevant to marketers and networks. For example, program ratings can inform product placement decisions, but they may offer limited insight into the size of the audience that has the opportunity to see commercial breaks. Some shows do not attract as large a program audience as others, which may make them less appealing for product placement. However, if they maintain a larger fraction of the program audience during commercial breaks, these programs may be bargains for advertisers. Conversely, more

[4] $A_{dtc}$ is the log of the ratio of two fractions. Because this lies in the interval $(-\infty, \infty)$, we assume a multivariate normal distribution for the error term.

[5] We considered adding a weekday-by-half-hour intercept in Equation 8, analogous to $\alpha_h$ in Equation 4. We did not find substantial variation in the values of this intercept or any differences in our substantive findings. Therefore, we present the more parsimonious model that omits this term.

popular programs may be potential "lemons" if the audience size more sharply declines during commercials. By examining the effects of observed program characteristics (i.e., genre and network) and show-specific X-factors on both program popularity (through $\beta$) and the extent of ad avoidance (through $\kappa$), we can identify the characteristics related to such patterns.[6]

These results offer practical implications for advertisers and networks. Advertisers relying on program ratings, as is the case in local markets, may discover that they are not reaching their gross-rating-point targets because of the extent to which certain shows experience ad avoidance. They may also attempt to identify programs that represent potential bargains. On the other side of the advertising market, networks require an understanding of the popularity of programs and the ad avoidance they experience. If relatively popular programs experience considerable ad avoidance, the use of commercial ratings may compel networks to offer "make goods." Networks may also find that other programs are underpriced if they experience lower levels of ad avoidance than would be anticipated from the show's characteristics.

We employ a hierarchical Bayesian approach to estimate the model presented in Equations 1–8. Uninformative normal priors are assumed for the vectors of fixed effects ($\alpha$, $\beta$, and $\kappa$) and an inverse-Wishart prior is assumed for $\Sigma$. We assume that $\delta_p \sim N(0, \sigma_{show}^2)$ and that $\psi_p \sim N(0, \sigma_{break}^2)$, with uninformative inverse-gamma prior distributions for $\sigma_{break}$ and $\sigma_{show}$.[7] We estimate a series of models, detailed in the next section, using WinBUGS (http://www.mrc-bsu.cam.ac.uk/bugs/), freely available software that draws from the marginal posterior distributions of the parameters of interest using Markov chain Monte Carlo (MCMC). We ran three independent chains for 25,000 iterations, discarding the first 10,000 iterations of each chain as a burn-in. We assessed convergence both visually and using Gelman and Rubin's (1992) F-test.

[6]Our intention in discussing potential "bargains" and "lemons," based on measures of the tuning audiences during programs and commercial breaks, is to caution against relying solely on program audience measures. Advertisers may select programs according to the target demographic, which may lead them to prefer programs that we term "lemons." We discuss how our approach can be extended to incorporate demographics in the "Conclusion" section.

[7]We considered a model in which the impact of show-specific effects on program attractiveness ($\delta_p$) and on the commercial break adjustment factor ($\psi_p$) were correlated and drawn from a multivariate normal distribution. We did not find support for this model and therefore assume that $\delta_p$ and $\psi_p$ are drawn from independent normal distributions.

## Empirical Analysis

We estimated a series of models to determine the specification for which we would present detailed results and demonstrate the applicability of our framework. We varied the inclusion of show-specific random effects ($\delta_p$ and $\psi_p$) and the inclusion of commercial break characteristics (length, concurrent commercial breaks, and proximity to the half hour or hour), which resulted in four model specifications. We compare these models using the deviance information criterion (DIC), a likelihood-based measure that penalizes more complex models (Spiegelhalter et al. 2002) and in which smaller values of DIC indicate a better-performing model. In the absence of show-specific random effects, the model that includes commercial break characteristics (Model 1; DIC = –33,814) provides a superior fit to the model that lacks them (Model 2; DIC = –33,188). We observe a similar pattern comparing the models with show-specific random effects, with the model including commercial break characteristics (Model 3; DIC = –41,481) performing better than the model without commercial break characteristics (Model 4; DIC = –40,884). Next, we discuss the results of Model 3. To understand the implications of ignoring the show-specific random effects, we provide a comparison of these results with those that would be reached under Model 1.

### *Model Results*

In Table 1, we present the weekday-by-half-hour intercepts ($\alpha_h$ from Equation 4) for the program rating model that incorporates show-specific random effects. We observe a fair amount of variation across weekdays and half-hour periods. Of note, we tend to observe lower intercepts from

**TABLE 1**
**Half-Hour Program Rating Intercepts for Model 3**

| | Monday | Tuesday | Wednesday | Thursday | Friday |
|---|---|---|---|---|---|
| 8:00–8:30 | –3.64 | –3.67 | –3.76 | –3.63 | –3.93 |
| | (–3.76, –3.50) | (–3.79, –3.53) | (–3.89, –3.63) | (–3.76, –3.50) | (–4.05, –3.80) |
| 8:30–9:00 | –3.57 | –3.60 | –3.67 | –3.56 | –3.89 |
| | (–3.69, –3.43) | (–3.72, –3.46) | (–3.80, –3.54) | (–3.69, –3.42) | (–4.01, –3.75) |
| 9:00–9:30 | –3.63 | –3.62 | –3.63 | –3.43 | –3.91 |
| | (–3.76, –3.50) | (–3.74, –3.48) | (–3.76, –3.50) | (–3.56, –3.29) | (–4.03, –3.77) |
| 9:30–10:00 | –3.54 | –3.61 | –3.60 | –3.41 | –3.90 |
| | (–3.66, –3.40) | (–3.73, –3.48) | (–3.72, –3.46) | (–3.54, –3.27) | (–4.02, –3.76) |
| 10:00–10:30 | –3.74 | –3.62 | –3.65 | –3.65 | –3.92 |
| | (–3.86, –3.6) | (–3.75, –3.48) | (–3.77, –3.51) | (–3.78, –3.51) | (–4.04, –3.78) |
| 10:30–11:00 | –3.80 | –3.70 | –3.71 | –3.74 | –3.94 |
| | (–3.92, –3.66) | (–3.82, –3.56) | (–3.84, –3.57) | (–3.87, –3.60) | (–4.06, –3.80) |

Notes: The table presents the posterior means and 95% highest posterior density intervals from the 15,000 draws of three independent MCMC chains.

10 P.M. to 11 P.M., as well as on Fridays, reflecting a diminished appeal of network television at these times. This seems logical because there may be less tuning to television later in the evening and on Friday nights.

Next, we consider how programming characteristics affect program ratings and the commercial break adjustment factors in Table 2. For comparative purposes, we also present estimates for the effects of programming characteristics that would be reached when show-specific random effects are ignored (Model 1).

We observe variation across networks with regard to program ratings and the adjustment factors. In interpreting the effect on program ratings, these parameters reflect the increased popularity of programs on the network compared with ABC. The baseline popularity of ABC programming is reflected in the intercepts in Table 1. We observe that CW, FOX, and NBC tend to have lower-rated programs than ABC, mirroring our exploratory analysis. Notably, when we incorporate show-specific random effects, comparing Model 1 with Model 3, we do not observe a difference between the baseline popularity of ABC and CBS programming.

Turning our attention to the coefficients for the adjustment factors, we again observe variation across the networks. In particular, CW tends to experience more ad avoidance than the other four networks, in addition to drawing a smaller program tuning audience. Advertisers should consider this "double whammy" when purchasing advertising time. Although measures of the program audience (still used in local markets) will reflect the lower average rating of the programs, they will not provide information about the larger audience declines during commercial breaks.

As we expected, programming during the WGA strike and repeat episodes are less attractive and thus should result in lower ratings. Note that the effect of the period after the WGA strike has a larger coefficient (in magnitude) than the period during the strike, resulting in a larger decline in ratings. As the positive interaction terms in the rating models reflect, the adverse effect of repeat episodes diminished

**TABLE 2**
**Model Parameters**

| | Program Rating Parameters (β) | | Adjustment Factor Parameters (κ) | |
|---|---|---|---|---|
| **Variable** | **Model 1** | **Model 3** | **Model 1** | **Model 3** |
| ABC | — | — | .07* | .07* |
| | | | (.07, .08) | (.06, .08) |
| CBS | .10* | .13 | .07* | .07* |
| | (.07, .14) | (–.02, .27) | (.07, .08) | (.05, .07) |
| CW | –.87* | –.93* | .11 | .11* |
| | (–.91, –.83) | (–1.12, –.71)* | (.10, .12) | (.10, .13) |
| FOX | –.25* | –.37* | .09* | .09* |
| | (–.30, –.20) | (–.53, –.20) | (.09, .10) | (.08, .10) |
| NBC | –.28* | –.28* | .09* | .08* |
| | (–.31, –.24) | (–.44, –.14) | (.08, .09) | (.07, .09) |
| WWGA = 2 (during strike) | –.09* | –.11* | –.01* | –.01* |
| | (–.11, –.07) | (–.13, –.09) | (–.01, –.01) | (–.02, –.01) |
| WGA = 3 (after strike) | –.15* | –.20* | –.00 | –.00 |
| | (–.17, –.13) | (–.22, –.18) | (–.00, .00) | (–.00, .00) |
| REPEAT | –.52* | –.51* | –.01* | –.01* |
| | (–.57, –.47) | (–.54, –.48) | (–.01, –.00) | (–.02, –.00) |
| REPEAT × (WGA = 2) | .16* | .14* | –.00 | .00 |
| | (.10, .21) | (.10, .18) | (–.01, .00) | (–.01, .01) |
| REPEAT × (WGA = 3) | .21* | .13* | .00 | .00 |
| | (.15, .27) | (.09, .16) | (–.01, .01) | (–.01, .01) |
| GENRE = 1 (drama) | .11* | .17* | –.02* | –.02* |
| | (.06, .15) | (.03, .30) | (–.03, –.02) | (–.03, –.01) |
| GENRE = 2 (reality television) | .10* | .17 | .01* | .02* |
| | (.05, .15) | (–.00, .34) | (.01, .02) | (.01, .04) |
| GENRE = 3 (sitcom) | –.19* | –.07 | –.01* | –.00 |
| | (–.25, –.14) | (–.24, .09) | (–.02, –.01) | (–.02, .01) |
| GENRE = 4 (news) | –.09* | .01 | –.01* | .00 |
| | (–.14, –.04) | (–.14, .15) | (–.02, –.01) | (–.01, .02) |
| GENRE = 5 (game show) | –.01 | .01 | .00 | .01 |
| | (–.07, .04) | (–.27, .27) | (–.01, .00) | (–.01, .03) |
| LENGTH (minute) | N.A. | N.A. | .01* | .01* |
| | | | (.01, .01) | (.01, .01) |
| CONCUR | N.A. | N.A. | –.003* | –.003* |
| | | | (–.004, –.002) | (–.004, –.002) |
| HALFHOUR | N.A. | N.A. | .005* | .005* |
| | | | (.002, .007) | (.003, .007) |

Notes: The table presents the posterior means and 95% highest posterior density (HPD) intervals from the 15,000 draws of three independent MCMC chains. We denote parameters for which the 95% HPD interval does not contain 0 with an asterisk (*).The shaded cells indicate parameters for which 0 is contained in the interval under Model 3 but not under Model 1. N.A. = not applicable.

during the WGA strike. Although the WGA strike and repeat episodes yielded lower program ratings, we observe that these factors are also associated with reduced ad avoidance. We hesitate to draw conclusions about the enduring effects of the WGA strike because we cannot disentangle this result from seasonality. However, further research could examine this issue using multiple years of data.

Next, we turn our attention to the effects of program genre. We find that only dramas are expected to be more popular than the "other" genre. This is somewhat surprising in light of our exploratory analysis, which revealed several drama and reality television programs with high program ratings, but this is explained by examining the genre effects when show-specific random effects are ignored (under Model 1). In the absence of show-specific random effects, we observe that reality television is more popular than the "other" genre, while sitcoms and news programs are less popular. Thus, the omission of show-specific random effects affects our inferences regarding differences that exist in the popularity of different programming genres. While dramas tend to draw larger live tuning audiences, the remaining genres do not differ in this regard. Rather, as we explore in greater detail, differences in programs' X-factors explain the variation in program ratings we observe.

Next, we examine the impact of program genre on the adjustment factors to determine whether more popular genres also maintain more of the tuning audience during commercial breaks. In addition to being expected to have a higher program rating, dramas also experience less ad avoidance than the "other" genre. This is consistent with previous research (e.g., Danaher 1995). We also find that the reality television genre differs from the "other" genre and experiences more ad avoidance. Comparing the results from Models 1 and 3, we again observe the effects of ignoring unobserved differences across programs. Although Model 1 would lead us to believe that sitcoms and news programming experience less ad avoidance, these differences are not significant after we incorporate programs' X-factors.

To provide a more intuitive interpretation of the observed differences in ad avoidance between dramas and reality television programs, we calculated the ratio of the adjustment factor for an "average" drama (assuming $\psi_p = 0$) to the adjustment factor for an "average" reality television program on the same network. We calculate this ratio at each iteration of the MCMC sampler and then compute the average across iterations (e.g., Rossi and Allenby 2003). A comparison of a drama and a reality television program with the same program rating shows that the drama is expected to have a commercial break audience that is 4.2% larger than the reality program.

Examining the impact of characteristics that vary across commercial breaks within the same episode, we observe that the extent of the audience decline diminishes as the number of concurrent commercial breaks increases. Though significant, this difference is fairly small: When we compare a commercial break with concurrent advertisements on two other networks with a commercial break with concurrent advertisements on four other networks, the difference in the commercial break rating is expected to be just .55%. An explanation for this finding is that viewers will not see programming on other channels, making channel switching less appealing. We also find that longer commercial breaks are associated with a larger decline in the commercial break audience. This is expected because longer breaks provide more opportunities to change the channel without missing programming. Although long breaks may allow networks to air more commercials, they may reach a smaller audience; a four-minute commercial break is expected to lose 1.3% more of the program rating than a three-minute commercial break. Commercial breaks that occur near the half-hour block tend to exhibit slightly more ad avoidance, which may be attributable to viewers tuning away to other programs or turning off the set-top boxes.

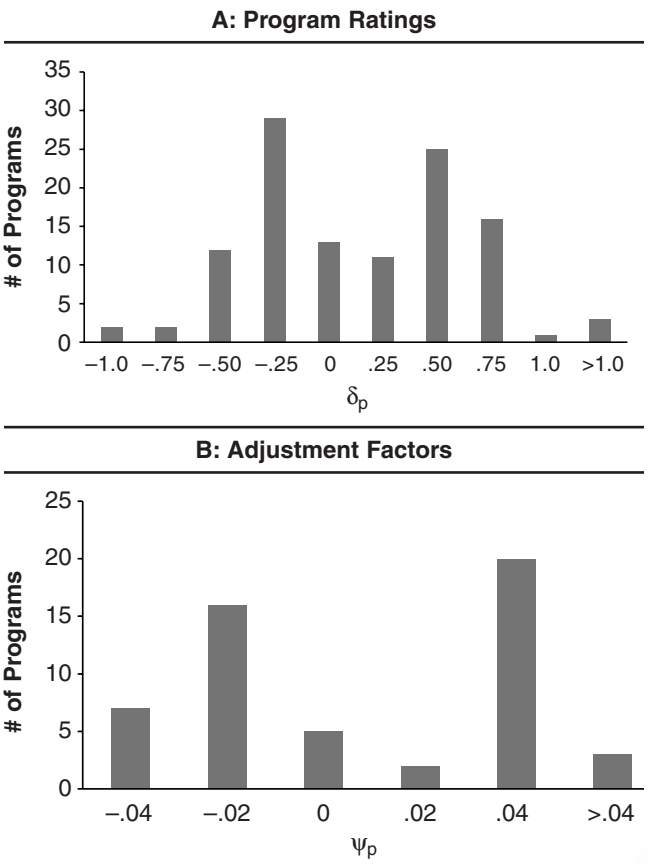
To understand the role of show-specific random effects, we examine their significance in the program ratings model ($\delta_p$) and the adjustment factor model ($\psi_p$). Of the 227 programs that aired during our observation period, 114 had a significant effect for program rating model, and 53 had a significant effect for the adjustment factor model. Furthermore, 31 programs had significant show-specific effects for both model components. The results reveal that *American Idol* has a higher program rating than is expected from its program characteristics and experiences less ad avoidance, both of which contribute to increased commercial break ratings. In contrast, *Big Brother 8* attracts a smaller program audience than is expected and experiences more ad avoidance, leading to lower commercial break ratings.

While these two programs illustrate the show-specific effects $\delta_p$ and $\psi_p$ complementing each other, we also observe programs that have opposing effects, resulting in potential bargains and lemons. For example, *NASCAR in Primetime* and *Aliens in America* have negative posterior mean coefficients in the program rating and the adjustment factor models. Although these programs draw smaller audiences than others with shared characteristics, they are expected to experience less ad avoidance and may be appealing to some media buyers. Conversely, programs such as *Hell's Kitchen*, *Pussycat Dolls Present: Girlicious*, and *Malcolm in the Middle* have higher-than-expected program attractiveness but experience more ad avoidance. Advertisers should give pause when considering such programs because the larger program audiences may not signal similarly high levels of opportunities to see commercial breaks.

In Figure 5, Panel A, we illustrate the distribution of the posterior means of $\delta_p$ for programs for which the 95% highest posterior density interval does not contain 0. Among the programs with the lowest significant values of $\delta_p$ are *NASCAR in Primetime* and *Nashville*, while the programs *The Moment of Truth*, *Dancing with the Stars*, and *American Idol* have the largest values. Compared with the largest genre effect (a posterior mean of .17 for dramas; see Table 2), all the programs with estimates of $\delta_p$ that significantly differ from 0 have posterior means that are greater in magnitude.

In Figure 5, Panel B, we illustrate the distribution of values of $\psi_p$, the show-specific random effect in the adjustment factor model. Comparing the values of $\psi_p$ that are significantly different from 0 with the posterior means of the genre effects in the adjustment factor model, we observe

**FIGURE 5**
**Distribution of Significant Show-Specific Random Effects**



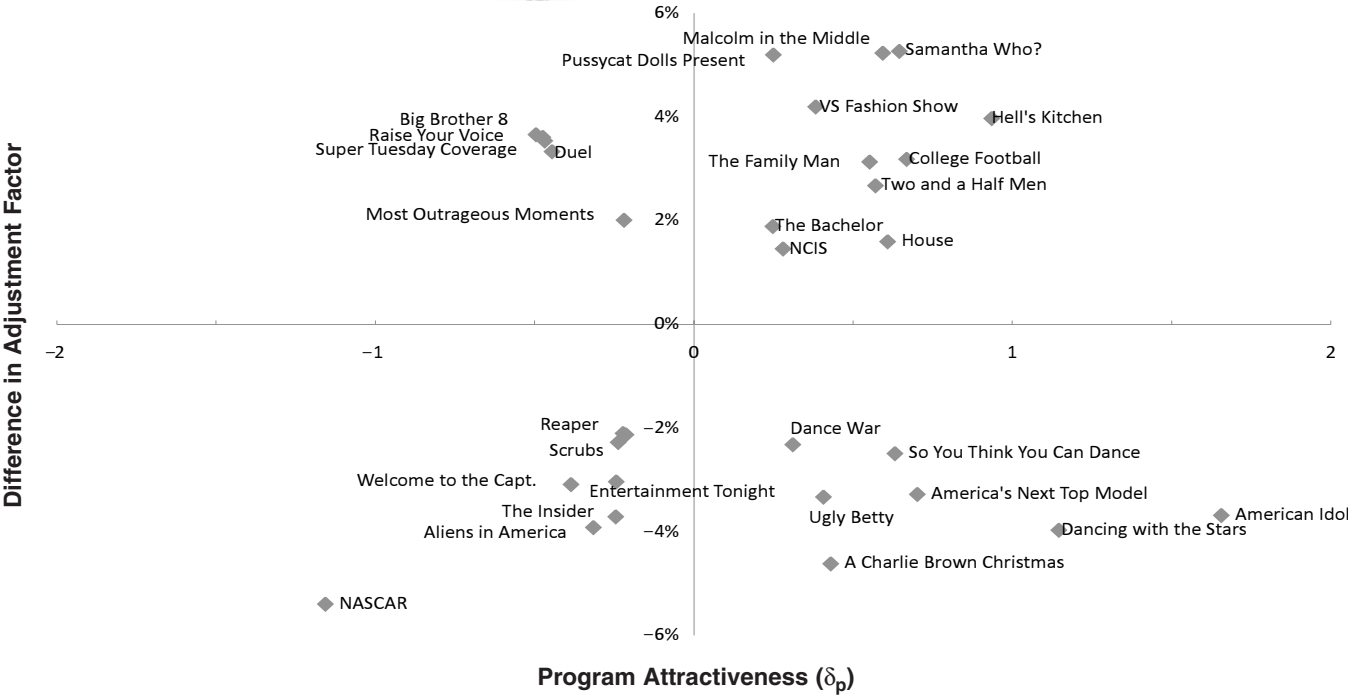

that 46 of the 53 shows have posterior means of $\psi_p$ that are greater in magnitude. Thus, in both the program ratings and the adjustment factor models, the show-specific random effects can outweigh the effects of genre.

To examine programs' X-factors for program popularity and ad avoidance together, we illustrate the show-specific random effects for programs for which $\delta_p$ and $\psi_p$ significantly differ from 0 in Figure 6. For ease of interpretation, we present the posterior mean of $\exp(\psi_p)$, which can be interpreted as the percentage increase or reduction in ad avoidance relative to an "average" program of the same genre and that airs on the same network.

The programs depicted in the upper-left quadrant of Figure 6 are expected to have lower program ratings than their observable characteristics would suggest and experience more ad avoidance. Conversely, programs in the lower-right quadrant will have higher program ratings than expected and experience less ad avoidance.

Consider the remaining two quadrants of Figure 6. The lower-left quadrant shows programs that draw smaller audiences than would be expected but also experience less ad avoidance. As such, these programs may represent potential bargains for advertisers, particularly if pricing is based on program ratings, because it enables advertisers to "spend a little less money and get a lot more delivery" (Birchall and Edgecliffe-Johnson 2009, p. 20). The programs depicted in the upper-right quadrant are the mirror image—though more popular than the network and genre would lead us to expect, they experience more ad avoidance. For such programs, advertisers need to exercise caution in relying on program ratings because more of the program audience will be lost during commercial breaks. In some cases, this will

**FIGURE 6**
**Show-Specific Random Effects Significant in Both Program Rating and Adjustment Factor Models**



result in a difference in the tuning audience size of more than 5%. These results are also of interest to networks, which must consider both program popularity and ad avoidance in pricing advertising.

To further highlight the importance of incorporating show-specific random effects in our program ratings and ad avoidance models, we calculated the mean absolute error (MAE) in program and commercial break ratings under Models 1 and 3. Although they incorporate the same observed program characteristics, the inclusion of show-specific random effects reduced the MAE of program ratings from .70% to .32%. The inclusion of show-specific random effects also reduced the MAE of commercial ratings from .62% to .29%.

### *Scenario Analysis*

To illustrate the differences between program and commercial break measures, we use our model results to compare the expected program and the commercial break ratings for a hypothetical scenario. As our focal program, we use a program on FOX that airs Tuesday at 9:00 P.M. For this illustration, we assume that a reality television program aired on CBS, a drama aired on NBC, and sitcoms aired on ABC and CW. We consider three program genres (drama, reality television, and sitcom), whether the commercial break appears within two minutes of a half hour, two commercial break lengths (two minutes versus four minutes), and two levels of concurrent commercial breaks (commercial breaks on two versus four other networks). This results in 24 profiles, for which we calculate program and commercial break ratings.

At each iteration of the MCMC sampler, we draw values for $\delta_p$ and $\psi_p$. Using these values and the parameter estimates for $\alpha_h$, $\beta$, and $\kappa$ at each iteration, we calculate the program ratings and the adjustment factor. The program and commercial break ratings, averaged across the iterations, appear in Table 3.

Consistent with Table 2, we observe reduced ad avoidance in commercial breaks that are shorter and aired at the same time as more commercial breaks on other networks. Comparing the expected commercial break ratings with the expected program ratings, we observe differences in the range of 8.8%–11.2% for dramas, 12.5%–15.2% for reality television, and 10.5%–12.8% for sitcoms, based on the characteristics of the commercial breaks. Although dramas and reality television programs are expected to have approximately the same program ratings, because of differences in the extent of ad avoidance, dramas are expected to have higher commercial break ratings. In addition to the differences that exist across genres, media planners should be aware of the variation in commercial break ratings attributable to differences in commercial break characteristics.

As this scenario analysis highlights, the difference between program and commercial break ratings is not constant but rather depends on program and commercial break characteristics, in addition to show-specific X-factors. As a result, the extent to which program ratings—the audience measures still used for advertising in local markets—overstate the audience potentially reached during commercial breaks will differ from program to program and, to a lesser extent, may even vary within an episode. Similarly, this reveals the limitations of inferring the size of the program



**TABLE 3**
**Comparison of Programs Versus Commercial Break Audiences**

| Genre | Program Rating | Break Near Half Hour? | Break Length | Concurrent Breaks | Commercial Break Rating |
|---|---|---|---|---|---|
| Drama | 2.13% | Yes | 2 minutes | 2 | 1.93% |
| | | Yes | 2 minutes | 4 | 1.95% |
| | | Yes | 4 minutes | 2 | 1.88% |
| | | Yes | 4 minutes | 4 | 1.89% |
| | | No | 2 minutes | 2 | 1.94% |
| | | No | 2 minutes | 4 | 1.96% |
| | | No | 4 minutes | 2 | 1.89% |
| | | No | 4 minutes | 4 | 1.90% |
| Reality television | 2.12% | Yes | 2 minutes | 2 | 1.84% |
| | | Yes | 2 minutes | 4 | 1.85% |
| | | Yes | 4 minutes | 2 | 1.79% |
| | | Yes | 4 minutes | 4 | 1.80% |
| | | No | 2 minutes | 2 | 1.85% |
| | | No | 2 minutes | 4 | 1.86% |
| | | No | 4 minutes | 2 | 1.80% |
| | | No | 4 minutes | 4 | 1.81% |
| Sitcom | 1.68% | Yes | 2 minutes | 2 | 1.50% |
| | | Yes | 2 minutes | 4 | 1.51% |
| | | Yes | 4 minutes | 2 | 1.46% |
| | | Yes | 4 minutes | 4 | 1.47% |
| | | No | 2 minutes | 2 | 1.51% |
| | | No | 2 minutes | 4 | 1.52% |
| | | No | 4 minutes | 2 | 1.47% |
| | | No | 4 minutes | 4 | 1.48% |

audience, useful for selecting programs for product placement, from measures of the commercial break audience.

## Conclusions and Future Directions

This research makes important advances to the growing body of research on program appeal and viewers' opportunity to see commercials. The research extends previous work in two key ways. First, rather than study television program appeal or ad avoidance separately, we simultaneously examine program and commercial tuning. Second, we incorporate longitudinal data from a full season of prime-time network television, which is a longer time frame than has typically been employed in prior research. With the long time frame and the inclusion of multiple airings of each program, we model show-specific effects and thus extend previous analyses that have examined program genre effects. Doing so reveals that many programs display significantly different program appeal and ad avoidance than their observable characteristics would suggest. As such, marketers engaging in product placements and traditional advertising may be able to find bargains (i.e., to get a greater delivery per ad dollar spent) through careful analysis and scrutiny of tuning activity for particular programs. The results highlight the importance of show-specific effects relative to program genre effects, and these findings may be of use to networks when making programming decisions. Overall, the results demonstrate the value of increasingly available set-top box tuning for researchers and practitioners.

Although measures based on program ratings have been used to make advertising decisions, we observe a wide range in the extent of a program's live tuning audience that is lost during its commercial breaks. Given the variation that exists across programs in terms of ad avoidance, program and commercial break ratings each offer value to marketers. Although summary commercial ratings are being employed at the national level, because of differences in ad avoidance across programs, these measures are less appropriate for gauging the size of audience that can be reached by branded entertainment and product placements, which program ratings can inform.

For local markets, in which program ratings continue to be used as the basis for advertising, our analysis demonstrates how granular tuning data can provide advertisers with show-specific measures of ad avoidance. Such metrics can give advertisers insight into the size of the audience with an opportunity to see their commercials rather than the opportunity to see the program in which the advertisements are placed, thus improving advertisers' media planning precision and enabling them to better negotiate advertising rates. Some marketers have products that are not sold nationwide (e.g., large regional banks, supermarket chains). Other brands vary in sales potential across local markets because of differences in the appeal of the product (e.g., four-wheel drive Subarus have traditionally sold better in smaller northern and western markets) or variance in retailer support (e.g., consumer product brands gain and lose sales in particular markets because of pricing, display, and relationship issues with dominant local supermarket chains). These marketers require information on program appeal and ad avoidance levels for specific local markets. Such information may be used to shift ad spending from national umbrella budgets to the local "spot" advertising markets with the greatest sales potential.

In addition to the value the findings provide to advertisers and media planners, networks may also benefit from a joint understanding of program ratings and ad avoidance. Programs that fail to draw sufficient audiences and/or experience high levels of ad avoidance may disappoint ad buyers and may compel networks to provide costly "make goods" to marketers. Our analysis shows that programs that are more popular than expected for their genre may experience higher or lower levels of ad avoidance than anticipated for the genre, thus limiting the use of program ratings as a proxy for commercial break ratings, and vice versa. Networks can use show-specific measures of expected ad avoidance to price advertising and product placements appropriately. Doing so would enable networks to reduce the number of times that opportunities to see advertisements or programs are underestimated, thus decreasing the number of times they undercharge marketers and reducing the number of times they would need to deliver "make goods" because advertisers were overcharged relative to the tuning audience delivered.

With program-specific measures of ad avoidance, networks can also identify commercial breaks that perform worse than would be expected. By identifying advertisements that are associated with more ad avoidance than anticipated, networks can consider different pod-placement or pricing strategies for advertisements. When pod-placement policies allow, it may be in a network's interest to slot such commercials later in breaks to minimize audience loss or to consider differential pricing (e.g., Wilbur, Goeree, and Ridder 2009). Marketers can also use such information in developing ad creatives.

The addition of show-specific random effects to models of program ratings and ad avoidance affects conclusions regarding the differences that exist across program genres. Previous studies have found that program popularity and the extent of ad avoidance are both related to program genre. However, incorporating show-specific random effects into our analysis of these two aspects of tuning reveals that several of these differences are not significant. That is, it is not so much program genres that vary in popularity and ad avoidance but rather specific programs. As such, the usefulness of genre as a proxy for the extent of show appeal and ad avoidance may be more limited than prior research has suggested. While the model omitting show-specific random effects would lead us to believe that sitcoms draw smaller audiences and reality television programs draw larger audiences, these generalizations do not hold after we allow for unobserved differences across shows. This finding has implications for networks because it may affect their programming decisions as a result of differences in the cost of producing programs of different genres. Knowledge of the impact of genre effects on program appeal and ad avoidance can also inform advertisers' decisions to place advertising in new programs from different genres during the critical up-front market (e.g., Steinberg 2009).

Several areas remain open for further research. In addition to providing audience measures for local markets, set-top box tuning data may allow for the detailed study of program appeal and ad avoidance for small, niche cable channels. As marketers face the challenge of dividing their budget among national advertising and advertising in local markets, our approach could be generalized to allow for differences that are observed across multiple metropolitan areas. Such a hierarchical model could also be used to gain insight into how program appeal and ad avoidance vary across demographic groups. With tuning data that incorporate DVR usage, a similar modeling approach could be employed to study how ad skipping affects the size of the potential audience for advertisements in various programs over message-relevant periods (Neff 2009; Wilbur 2008a). By using both program and commercial break ratings, networks and advertisers can identify program elements and commercial characteristics that maximize viewers' opportunities to see product placement and commercials. Networks can also examine the effects of the order and assortment of commercials (e.g., Schweidel, Bradlow, and Williams 2005) in a break on ad avoidance to minimize audience loss.

## REFERENCES


Abraham, Magid M. and Leonard M. Lodish (1990), "Getting the Most Out of Advertising and Promotion," *Harvard Business Review*, 68 (3), 50–60.

Arango, Tim (2008), "Cable Firms Join Forces to Attract Focused Ads," *The New York Times*, (accessed February 1, 2009), [available at http://www.nytimes.com/2008/03/10/business/media/10cable.html].

Atkinson, Claire (2007), "Soon, You Could Know How Many See Your Spots (Really!)," *Advertising Age*, (accessed March 20, 2007), [available at http://adage.com/abstract.php?article_id=115633].

——— (2008), "How Commercial Ratings Changed the $70B TV Market," *Advertising Age* Commercial-Ratings White Paper, (accessed October 30, 2008), [available at http://adage.com/images/random/0908/Commercial-Ratings%20White%20Paper.pdf].

Bass, Frank M., Norris Bruce, Sumit Manumdar, and B.P.S. Murthi (2007), "Wearout Effects of Different Advertising Themes: A Dynamic Bayesian Model of the Advertising-Sales Relationship," *Marketing Science*, 26 (2), 179–95.

Birchall, Jonathan and Andrew Edgecliffe-Johnson (2009), "Procter & Gamble Says Advertisers Have Upper Hand," *Financial Times*, (May 1), 20.

Briggs, Rex and Greg Stuart (2006), *What Sticks*. Chicago: Kaplan Publishing.

Bronnenberg, Bart J., Jean-Pierre Dube, and Carl F. Mela (2009), "Do DVRs Influence Consumers' Brand Purchases?" Marketing Science Institute Special Report No. 09-208.

Bruce, Norris I. (2008), "Pooling and Dynamic Forgetting Effects in Multitheme Advertising: Tracking the Advertising Sales Relationship with Particle Filters," *Marketing Science*, 27 (4), 659–73.

Danaher, Peter J. (1995), "What Happens to Television Ratings During Commercial Breaks?" *Journal of Advertising Research*, 35 (1), 37–47.

Ephron, Erwin (2001), "Media-Mix Needs to Think Local," *MediaWeek*, (accessed October 1, 2009), [available at http://www.mediaweek.com/mw/esearch/article_display.jsp?vnu_content_id=1143075].

——— (2006), "Who Watches the Watchers?" *MediaWeek*, (accessed February 1, 2009), [available at http://www.mediaweek.com/mw/esearch/article_display.jsp?vnu_content_id=1003380936].

——— (2007), "Blunt Pencil: The Minute That Took Nearly a Year," *MediaWeek*, (accessed February 1, 2009), [available at http://www.mediaweek.com/mw/esearch/article_display.jsp?vnu_content_id=1003550027].

Fox, Edward J. and Stephen J. Hoch (2005), "Cherry-Picking," *Journal of Marketing*, 69 (January), 46–62.

Friedman, Wayne (2009), "Group M's Scanzoni Slams Nielsen Live-Plus," *MediaPost*, (accessed October 1, 2009), [available at http://www.mediapost.com/publications/?fa=Articles.showArticle& art_aid=113370].

Gelman, Andrew and Donald B. Rubin (1992), "Inferences from Iterative Simulation Using Multiple Sequences," *Statistical Science*, 7 (4), 457–72.

Gloede, Bill (2006), "Ticking Clock? Or Time Bomb?" *MediaWeek*, (July 24), 10.

Hampp, Andrew (2007), "Marketers Push for Brand-Specific Commercial Ratings," *Advertising Age*, (accessed March 13, 2007), [available at http://adage.com/abstract.php?article_id=115535].

Kang, Stephanie (2009), "Nielsen Will Get Data from New Channel," *The Wall Street Journal*, (March 12), B4.

Karrh, James A., Kathy Brittain McKee, and Carol J. Pardun (2003), "Practitioners' Evolving Views on Product Placement Effectiveness," *Journal of Advertising Research*, 43 (2), 138–49.

Kaul, Anil and Dick R. Wittink (1995), "Empirical Generalizations About the Impact of Advertising on Price Sensitivity and Price," *Marketing Science*, 14 (3), G151–60.

Krider, Robert E. and Charles B. Weinberg (1998), "Competitive Dynamics and the Introduction of New Products: The Motion Picture Timing Game," *Journal of Marketing Research*, 35 (February), 1–15.

Lehmann, Donald R. and Charles B. Weinberg (2000), "Sales Through Sequential Distribution Channels: An Application to Movies and Videos," *Journal of Marketing*, 64 (July), 18–33.

Li, Kenneth (2009), "Coalition Has Its Sights Set on Monitoring All Viewing Habits," *Financial Times*, (September 11), 14.

McClellan, Steve (2008), "Nielsen Readies New Local Ratings," *AdWeek*, (accessed October 1, 2009), [available at http://www.adweek.com/aw/content_display/news/media/e3ib0d0b65921d8c6ff693b6f3fb570643b].

Montgomery, Alan L., Shibo Li, Kannan Srinivasan, and John C. Liechty (2004), "Modeling Online Browsing and Path Analysis Using Clickstream Data," *Marketing Science*, 23 (4), 579–95.

Naik, Prasad A. (1999), "Estimating the Half-Life of Advertisements," *Marketing Letters*, 10 (3), 351–62.

———, Murali K. Mantrala, and Alan G. Sawyer (1998), "Planning Media Schedules in the Presence of Dynamic Advertising Quality," *Marketing Science*, 17 (3), 214–35.

——— and Kalyan Raman (2003), "Understanding the Impact of Synergy in Multimedia Communications," *Journal of Marketing Research*, 40 (November), 375–88.

Neff, Jack (2009), "Future of Advertising? Print, TV, Online Ads," *Advertising Age*, (accessed June 19, 2009), [available at http://adage.com/abstract.php?article_id=136993].

Nielsen (2008), "A2/M2 Three Screen Report, 3rd Quarter 2008," (accessed October 1, 2009), [available at http://blog.nielsen.com/nielsenwire/wp-content/uploads/2008/11/nielsen_three_screen_report_3q08.pdf].

Poltrack, David F. (2006), "Why TV Needs Commercial Ratings—Now," *Advertising Age*, (accessed Novemeber 17, 2006), [available at http://adage.com/abstract.php?article_id=113089].

Popkowski-Leszczyc, Peter T.L. and Ram C. Rao (1990), "An Empirical Analysis of National and Local Advertising Effect on Price Elasticity," *Marketing Letters*, 1 (2), 149–60.

Reuters (2009), "Media Industry Leaders to Launch the Coalition for Innovative Media Measurement," (accessed October 1, 2009), [available at http://www.reuters.com/article/pressRelease/idUS123082+10-Sep-2009+PRN20090910].

Rossi, Peter E. and Greg M. Allenby (2003), "Bayesian Statistics in Marketing," *Marketing Science*, 22 (3), 304–328.

Rust, Roland T. and Mark I. Alpert (1984), "An Audience Flow Model of Television Viewing Choice," *Marketing Science*, 3 (2), 113–24.

———, Wagner A. Kamakura, and Mark I. Alpert (1992), "Viewer Preference Segmentation and Viewing Choice Models for Network Television," *Journal of Advertising*, 21 (1), 1–17.

Schneider, Michael (2009), "Fox Wants Answers from Nielsen," *Variety*, (accessed October 1, 2009), [available at http://www.variety.com/article/VR1118003924.html?categoryId=1275&cs=1].

Schweidel, David A., Eric T. Bradlow, and Patti Williams (2006), "A Feature-Based Approach to Assessing Advertisement Similarity," *Journal of Marketing Research*, 43 (May), 237–43.

Shachar, Ron and John W. Emerson (2000), "Cast Demographics, Unobserved Segments, and Heterogeneous Switching Costs in a Television Viewing Choice Model," *Journal of Marketing Research*, 37 (May), 173–86.

Siddarth, S. and Amitava Chattopadhyay (1998), "To Zap or Not to Zap: A Study of the Determinants of Channel Switching During Commercials," *Marketing Science*, 17 (2), 124–38.

Spiegelhalter, D.J., N.G. Best, B.P. Carlin, and A. Van der Linde (2002), "Bayesian Measures of Model Complexity and Fit (with Discussion)," *Journal of the Royal Statistics Society, Series B*, 64 (4), 583–640.

Steinberg, Brian (2005), "Building a Metric to Determine Viewer Attention to Ad Spots," *The Wall Street Journal*, (August 10), B3.

——— (2009), "There's Real Business in FUNNY Business as Nets Pile on Sitcoms," *Advertising Age*, (accessed October 1, 2009), [available at http://adage.com/abstract.php?article_id=138600].

——— and Brooks Barnes (2006), "Nielsen Plans to Track Viewership of TV Commercials for First Time," *The Wall Street Journal*, (July 10), A1.

——— and Andrew Hampp (2007), "Commercial Ratings? Nets Talk TiVo Instead," *Advertising Age*, (accessed June 5, 2007), [available at http://adage.com/abstract.php?article_id=117076].

——— and Kimberly D. Williams (2007), "What Fall TV Spots Cost," *Advertising Age*, (accessed October 4, 2007), [available at http://adage.com/images/random/0907/2007_Ad_Age_TV_Price_Survey.pdf].

Van Meurs, Lex (1998), "Zapp! A Study on Switching Behavior During Commercial Breaks," *Journal of Advertising Research*, 38 (1), 43–53.

Wilbur, Kenneth C. (2008a), "How the Digital Video Recorder Changes Traditional Television Advertising," *Journal of Advertising*, 37 (1), 143–49.

——— (2008b), "A Two-Sided, Empirical Model of Television Advertising and Viewing Markets," *Marketing Science*, 27 (3), 356–78.

———, Michelle Sovinsky Goeree, and Geert Ridder (2009), "Effects of Advertising and Product Placement on Television Audiences," University of Southern California Marshall School of Business Working Paper No. MKT 09-09, (accessed June 19, 2009), [available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1151507].

**Fonte: Journal of Marketing, v. 74, n. 3, p. 18-33, May 2010. [Base de Dados]. Disponível em: <http//:web.ebscohost.com>. Acesso em: 19 maio 2010.**