MSI

Marketing Science Institute Working Paper Series 2021

Report No. 21-132

How Viewer Tuning, Presence and Attention Respond to Ad Content and Predict Brand Search Lift

Matthew McGranaghan, Jura Liaukonyte, and Kenneth C. Wilbur

How Viewer Tuning, Presence and Attention Respond to Ad Content and Predict Brand Search Lift[*]

Matthew McGranaghan[†] Jura Liaukonyte[‡] Kenneth C. Wilbur[§]

September 2021

Abstract

New technology measures TV viewer tuning, presence and attention, enabling the first distinctions between TV ad viewability and actual ad viewing. We compare new and traditional viewing metrics, finding that 30% of TV ads play to empty rooms. We then use broadcast networks' verifiably quasi-random ordering of ads within commercial breaks to identify causal effects of ads on viewing behaviors among 4 million advertising exposures. We measure ad metadata and machine-code content features for 6,650 frequent ad videos. We find that recreational product ads preserve audience tuning and presence. Prescription drug advertisements decrease tuning and presence, more so for drugs that treat more prevalent and severe conditions. We also investigate whether new viewing data can inform advertiser objectives, finding that attention helps predict brand search lift after ads.

[*]We thank TVision Insights, especially Tristan Webster, Yan Liu, and Dan Schiffman, for enabling this research and providing valuable feedback. We also thank Thales Teixeira at Decoupling.co for connecting us with TVision, Snehanshu Tiwari and Hayoye Yang for superlative research assistance, and seminar participants at Bocconi, BU, Cornell, Delaware, Georgetown, Indiana, Minnesota, UCSD, UT-Dallas, Virginia Tech, Western Ontario and Wisconsin for helpful suggestions. Authors have no funding or conflicts of interest to report. An earlier version dated July 2019 was called "Watching People Watch TV." Mistakes, omissions and opinions are ours alone.

[†]University of Delaware, Alfred Lerner College of Business and Economics, mmcgran@udel.edu

[‡]Cornell University, S.C. Johnson College of Business, jurate@cornell.edu

[§]University of California at San Diego, Rady School of Management, kcwilbur@ucsd.edu

1 Introduction

This paper asks three questions about TV advertising viewing. How do traditional tuning data compare to new TV viewing metrics? How do new viewing metrics respond to ad content? Can new viewing metrics help predict ad response?

We investigate viewing behavior measured using cameras, microphones and algorithms in a paid sample of 1,155 consenting households. *Tuning* is measured by comparing television audio to a database of known programs. Individual viewer *presence* is measured using person-detection and facial-recognition algorithms. *Attention* is measured as the co-occurrence of eyes-open and eyes-on-screen inferences. All three viewing metrics are measured *in situ*, passively and continuously.

Viewer presence detection distinguishes true ad exposures from ads that air to empty rooms. Viewers are absent from the room during 30% of the ads that play on their TV during active viewing sessions. Viewers are about four times more likely to leave the room during an ad than to tune away.

We use ad metadata and machines to measure three sets of content features in 6,650 frequent ad videos. We verify and exploit broadcast networks' practice of quasi-random ordering of ads within breaks to identify causal effects of ad content on tuning, presence and attention. Viewer tuning and presence during ads fall less during recreational product ads. Prescription drug ads reduce tuning and presence more than average, more so for drugs that treat severe and prevalent conditions. Attention decreases across the first three ad slots in a break and falls with ad duration.

Finally, we investigate whether the new viewing metrics can help predict brand search lift after TV ads. Attention helps predict online search response to ads, as does distinguishing ad exposures from viewable ads. Therefore, the new viewing measures may help inform advertiser objectives.

Next, we discuss the relevant advertising literature. Section 2 describes the new viewing and ad content metrics. Section 3 specifies the model and causal identification strategy. Section 4 presents the ad content findings and then explains the drug results using treated condition attributes. Section 5 relates viewing metrics to online search. Section 6 concludes with limitations and possible extensions.

Relationship to Previous Literature. Advertising studies usually balance viewing behavior measurement quality and sample size. For example, many papers study how ads change TV tuning in large field samples (Danaher, 1995; Shachar and Emerson, 2000; Goettler and Shachar, 2001; Wilbur, 2008; Schweidel and Kent, 2010; Swaminathan and Kent, 2013; Wilbur, 2016). A distinct literature studies how ads change viewer attention and emotion in small laboratory samples (Zhang et al., 2009; Teixeira et al., 2012; Liu et al., 2018). A third literature studies ad viewing in small-scale ethnographic samples (Jayasinghe and Ritson, 2013; Voorveld and Viswanathan, 2015).

The current paper is likely the first to combine lab-like ad viewing metrics with large field samples. We do know of one paper that has studied similar ad viewing data. Liu et al. (2021) quantify suspense and surprise during baseball games and find that in-game suspense decreases consumer attention during commercials, whereas in-game surprise enhances ad attention.

We join a growing number of studies that combine data on TV ad avoidance and ad response. The first we know of was Zufryden et al. (1993), who found that households' TV ad "zapping" decisions correlate with their future packaged good purchases. Siddarth and Chattopadhyay (1998) and Tuchman et al. (2018) used household purchase data to predict TV ad avoidance, finding that consumers are less likely to avoid ads for brands that they have previously bought. Bronnenberg et al. (2010) analyzed a field experiment that treated households with free digital video recorders, finding a tight null treatment effect on packaged good purchases. Deng and

Mela (2018) combined device-level ad avoidance and sales data to study the consequences of microtargeted TV advertising. We contribute to this literature by quantifying how new viewing metrics respond to ad content and predict brand search lift.

2 Data and Descriptive Results

New metrics require careful definition, description and comparison to traditional metrics. We describe the viewing data, introduce the ad data, and then describe their covariation.

2.1 Ad Viewing

Viewer tuning, presence, and attention data are provided by TVision Insights, an analytics firm founded to modernize television audience measurements. The data cover 3,659 viewers in TVision's panel of 1,155 consenting households between July 2016 and June 2017.

Measurement TVision installs cameras and microphones on each household's primary TV. Initial set-up includes training facial recognition algorithms on each household member. Infrared sensors measure depth and aid detection in low light conditions. Image data are processed in real time at the frame level five to six times per second on average. Images are not stored or transmitted outside of the home. TVision combines audio data with industry-standard automated content recognition (ACR) services to measure television *tuning*, i.e., the television network and timestamp of the audio stream.

TVision uses person detection algorithms to identify human bodies – sets of heads, shoulders, and arms – in the cameras' field of view. Person detection technology is similar to real-time face and body recognition algorithms used in smartphone apps, e.g. Instagram filters. For each face, the software either identifies the household member or assigns a unique guest identifier. *Presence* is the detection and recognition of a particular viewer in the room.

TVision software measures when viewers' eyes are open and infers head orientation based on the relative sizes, positions and angles of facial features. *Attention* is the co-occurrence of eyes-open and eyes-on-screen inferences.[1]

At the time the sample data were produced, TVision equipment measured tuning, presence and attention continuously and then sampled one measurement for each viewer in each ad second. The data provided to us report average behaviors across viewer-seconds within each viewer and ad insertion. So, viewing behaviors within 30-second ad exposures are based on 30 underlying measurements per viewer.

Comparisons to Extant Advertising Audience Measurements. Nielsen and TVision both measure tuning continuously and passively. However, two major advantages of TVision data are passive, continuous presence measurement and attention data.

Traditional television audience measurements are based on digital devices—mostly smart TVs and set-top boxes—and Nielsen "People Meters." Digital devices measure tuning passively and continuously in millions of households, but do not measure which household members are watching at which times, or whether anyone is watching at all. People Meters measure tuning passively and continuously in representative samples of tens of thousands of households, and they additionally measure viewer presence in an intrusive and intermittent fashion. People Meters use a red light to prompt Nielsen panelists to log in on a special remote control at the start of each viewing session and once every 15-45 minutes thereafter. Nielsen combines viewer presence data with tuning data to determine audience demographics and infer when viewing sessions may have concluded.

[1] This is only one possible measure of attention. A viewer may actively process ad audio while looking away from the screen; or stare at the screen yet be entirely absorbed in other thoughts; or focus on a program but blinking or head movements may lead to average attention well below 100%.

Media buyers have long known that Nielsen audience estimates overstate advertisement audiences. Ephron (2006) argued that,

> [C]ommercial-minute data... show losses of audience of about 2 to 10 percent during commercials compared to programs... Researchers, who read the fine print, qualify a Nielsen commercial exposure as '*an opportunity to see*' a commercial. And given the opportunity, it's obvious the probability is a lot less than one. So the Nielsen commercial-minute audience is an overstatement of people seeing commercials.

In contrast, TVision's passive and continuous presence measurements avoid disrupting natural viewing behaviors and distinguish opportunities to see from actual ad exposures.

In what follows, we define an *Opportunity to See (OTS)* as a viewer's television tuned to an ad insertion for at least two seconds, for any commercial break in which the viewer is present for at least two seconds in the first ad slot of the break. Selecting viewers present at the start of the break removes inactive viewing sessions from the sample. The two-second threshold is inspired by the Media Rating Council's definition of a "viewable impression" in which 100% of an ad's pixels play on a screen for at least 2 seconds (Knauer, 2019).

The definition of a "viewable impression" does not require a human to be exposed to the ad. Industry reports estimate that 10-30% of digital ad spend is lost to ad fraud, often because ads are served to machines instead of to humans (Gordon et al., 2021). For example, an analysis by the IAB Tech Lab indicated that only 59.8% of ad clicks could be confirmed as human traffic (Swant, 2019). TVision presence data may offer the first passive, continuous human detection data in the history of video advertising.

We define an *ad exposure* as any OTS in which a viewer is detected as present for at least two seconds. An example can illustrate how ad OTS differ from ad exposures. Suppose a viewer watches a program that goes to an ad break. The break starts with a Coca-Cola ad, then a Geico ad, then follows with 5 other ads. The viewer leaves the room halfway through the Geico ad in the second slot and does not return until after the break ends. The viewer has had 7 opportunities to see ads and two ad exposures (Coca-Cola and Geico).

The other major advantage of TVision data is attention measurement. TVision provides the first continuous, passive measurements of television viewers' ad attention in natural viewing environments. Viewer attention is becoming increasingly scarce as consumers increasingly use smartphones or tablets alongside television; attention measurements may help improve advertiser choices.

Ad Viewing Descriptives. Is tuning a reliable proxy for presence or attention? We compare the three behaviors in samples of ad OTS and ad exposures. The following graphics focus on viewers with at least 50 ad exposures and commercial breaks on top-4 broadcast networks between 7:00 A.M. and 1:00 A.M.[2]

Figure 1 compares densities of viewers' average tuning, presence and attention behaviors during OTS and ad exposures. The average viewer's television remains tuned to 96.3% of viewable TV ad seconds. However, the average viewer remains present for only 54.6% of all ad seconds during OTS, with substantial heterogeneity in average viewer presence resulting in a 10-90th percentile range of 28.2% to 76.7%. Further, the average viewer only pays attention to 7.7% of ad seconds during OTS. In fact, 7.2% of viewers disregard 99% or more of all viewable ad seconds on average. Ignoring the distinction between OTS and exposures overestimates ad viewing because 29.8% of the observations occur when ads play to empty rooms.

Comparing OTS densities to exposure densities, the tuning distribution changes little, with an average of 96.2% during ad exposures. However, average viewer presence increases from 54.6% to 85.3% and variation across viewers in average presence falls by about half after filtering out non-exposures. Average viewer attention increases from 7.7% to 11.7%. Only 3.1% of viewers disregard more than 99% of all ad seconds during exposures.

[2]We select 50 to minimize sampling variance from infrequent guests or mistaken person classifications.

Figure 2 depicts covariation among viewers' average behaviors. Each point plots an individual viewer's average of two viewing behaviors among all ad exposures observed. All three panels show diffusion around strong central tendencies, indicating that the three behaviors are correlated yet still quite distinct. For example, within the subset of viewers who average 95% tuning, their average presence ranges from 75% to 94%. Within the subset of viewers who average 85% presence, their average attention ranges from 1% to about 22%. In sum, people engage in different ad viewing behaviors at quite different rates. Thus tuning is an incomplete proxy for presence or attention.

Figure 3 illustrates average tuning, presence and attention within six age/gender groups in the OTS and exposure samples. Tuning varies minimally across demographics and samples. However, the OTS sample shows large differences in viewer presence across demographic groups, with older females showing the highest average presence at 67.3% and younger males showing the lowest average presence at 50.6%. However, the ad exposure sample shows muted variation with mean presence ranging from $85.8\% - 90.8\%$ across groups. Therefore, viewers physically depart during ads (i.e., nonpresence) about 3-4 times more often than they change channels during ads (i.e., nontuning).

Like presence, ad attention increases with viewer age within both genders. However, unlike presence, removing non-exposures from the OTS sample does not change variation across groups much; instead, it mostly induces a level shift in mean attention. The level shifts imply that people leave the room during ads they are unlikely to have watched otherwise. Overall, patterns of tuning, presence, and attention during ad exposures are consistent with a theory that older viewers are more likely to avoid ads by changing channels and younger viewers are more likely to avoid ads by leaving the room or diverting their visual attention.

Appendix Figure A1 graphs mean tuning, presence and attention by ad slot, based on OTS and exposures. Exposure data show that average tuning gradually rises as ad-averse viewers select out of the break. Average presence also rises uniformly after the first slot. Average attention is nearly constant at 13.5% after the first slot. Together, these findings suggest that passive measurements of viewer presence and attention offer richer information than tuning alone.

2.2 Sample Selection, Ad Features and Preliminary Evidence

TVision ad insertion data document the ad environment – network, date, air time, program, genre and episode. Ad metadata provide the ad creative name, product name, brand name, product category, and ad duration.[3]

Ad exposures and attention build throughout the day and peak during the evening prime time hours (see Appendix Figure A2). The estimation sample selects ad insertions between 7:00 AM and 1:00 AM on the four major broadcast networks (ABC, CBS, FOX, NBC) from July 2016 to June 2017. In total, we observe 4,257,112 exposures of 3,659 unique viewers to 6,650 unique ad creatives in 167 product categories. Regular viewers—defined as those with at least 50 active viewing days—view 22.5 sample ads per person per day, and attend to 1.1 out of 8.9 exposed ad minutes.

Ad Features. The three most general sets of ad features are ad creative identifiers, brand identifiers and product category identifiers. An ad creative identifier summarizes all content in a unique ad creative and bounds the behavioral variation ad content could explain. Ad creatives with fewer than 50 exposures are grouped into a composite, covering 2% of all exposures.

Advertised product categories describe things like beer, cancer drugs and pick-up trucks.

[3]We checked the TVision ad insertion data against the official advertising schedule for Super Bowl 51 and against Kantar Media, a reliable commercial source of ad insertion data. The TVision ad data contained all 63 national ads in the correct order. The average insertion time difference was 4.9 seconds, consistent with standard asynchronies in local broadcast affiliate streams. We also found a very high correspondence between TVision ad insertion data and Kantar data in other programs.

Product categories capture stylistic and thematic similarities across ads, such as humor and good times in beer ads or images of toughness in truck ads. They also reflect regulatory requirements about ad content, such as treated condition or potential side effects in drug ads (FDA, 2020).

We follow a long literature on ad content (Resnik and Stern, 1977; Anderson et al., 2013; Liaukonyte et al., 2015; Tucker, 2015; Anderson et al., 2016; Lee et al., 2018; Tsai and Honka, 2021) and supplement ad metadata with three sets of ad features. First, a TV advertising measurement company called iSpot.tv provides an online database of TV ads. We algorithmically downloaded ad videos from iSpot covering 85% of exposures to national ads in the estimation sample ($\sim$65% of uniquely labeled ad creatives). We also scraped ad content features from iSpot webpages. For each ad, we observe a Tagline identifier, a sentiment score ranging from zero to one based on the positivity of the words in the audio transcript, a promotion identifier, a commercial Music identifier, a Movie identifier, an "engagement rating" based on the volume of digital activity related to the ad creative, and a professional actors indicator. iSpot also classifies the "mood" of each ad as active, emotional, funny, informative, or sexy.

Second, we constructed a set of machine-coded ad content features using machine learning algorithms collated by Schwenzow et al. (2021). We retained measures with precision and recall scores of at least 50%, including number of scene transitions; average colorfulness, saturation, and luminosity; and percentages of ad video seconds showing facial expressions of Surprise, Happiness or Neutral emotion.

Finally, we used Google Cloud Vision (GCV) to tag recognized images within ad videos. GCV identifies over 1,000 common image tags in 70 categories, based on a large validation set of human-tagged images and videos.[4] We took two steps to filter out tags likely to be inaccurate

[4]Blanchard et al. (2020) provides a detailed overview of the video coding platform and finds that image tags help to predict new product adoption. Kubany et al. (2020) found GCV performs well compared to competing image recognition services.

or redundant. First, we sought to limit errors in variables by only retaining tags that describe concrete nouns and verbs. Second, we sought to limit collinearity by retaining only those 32 tags for which 50% or more of variance remained unexplained in a regression of the tag on product category, iSpot and machine-coded ad features.

Table A1 summarizes iSpot and machine-coded features, and Table A2 displays GCV image tags and their frequencies in bold. We note three important caveats. First, ad content feature coding is incomplete because no current method can fully characterize video content in interpretable ways. Unobserved features may correlate with coded features and complicate interpretation of feature coding results. Second, classical errors-in-variables issues may bias ad content feature parameter estimates toward zero. Third, ad videos were unavailable for 35% of all creatives covering 15% of sample exposures, so all ad content features implicitly interact video availability with feature measurement.

Linking Ad Viewing to Ad Features. Figure 4 shows how ad viewing changes with ad environment and ad features during ad exposures. Broadcast networks with lower average tuning tend to have higher average attention, a pattern that repeats when comparing prime time to other dayparts. Program genres show some different patterns. For example, ads during Football games have the highest tuning and highest attention whereas Drama ads have the lowest tuning with moderate attention.

Shorter ads retain more viewers and attention than longer ads on a per-second basis. Comparing 15-second ads to 30-second ads, mean tuning per ad second falls from 98% to 94%, mean presence falls from 94.1% to 85.6%, and mean attention falls from 13.5% to 12.5%.

Advertised product category also correlates with ad viewing. Figure 4 provides mean viewing behaviors for the 10 most-tuned and the 10 least-tuned advertised product categories. The Casinos

& gambling category is both the most-tuned and most-attended ad category. Entertainment & games ads are highly tuned but attended much less, perhaps because they are more likely to generate second-screening behaviors. Eight of the ten least-tuned ad categories are for prescription drugs, and they all receive less attention than the remaining two least-tuned product categories.

Table A3 presents variance decompositions of ad viewing behaviors on individual sets of viewer, break and ad features. Viewer identifiers are the best predictors of presence and attention, explaining 53 times more variation in attention than the traditional targeting variables of age and gender combined. This finding congrues with prior research quantifying the profitability of individually targeted advertising (Deng and Mela, 2018). Ad environment variables also correlate with viewing behaviors, including slot within the break, network and program genre.

The ad features correlate weakly with presence but explain less variation in tuning and remarkably little variation in attention. One of the strongest correlates is ad duration, explaining 1.6% of tuning and 4.6% of presence but just .1% of attention. Another is ad category, which explains 0.5% of tuning, 1.0% of presence and 0.1% of attention. In summary, the variance decompositions presage difficulty in detecting effects of ad content on advertising attention.

3 Empirical Framework

We describe the model, causal identification, and results interpretation.

3.1 Model Specification

We develop an empirical model in the "causal effects" paradigm described by Chintagunta and Nair (2011). The model explains tuning, presence, and attention behaviors as functions of ad features, slot and time-within-break features, and viewer-break interaction effects. b indexes ad breaks, each of which is a set of consecutive ads inserted into a specific network-program-date-time combination. Each ad slot within a break is indexed with s, so every (b, s) combination identifies a unique insertion of the particular ad creative that was aired in slot s of break b.

Let y^j_{ibs} be viewing behavior j for viewer i exposed to the s^{th} advertisement in ad break b. Ad viewing behavior is modeled as follows:

$$y^j_{ibs} = x'_{bs}\boldsymbol{\beta}^j + g(1_s, l_{bs}, t_{bs}; \Theta^j) + \delta^j_{ib} + \varepsilon^j_{ibs} \tag{1}$$

x_{bs} is a vector of ad features, such as ad creative fixed effects, or product category fixed effects and ad content features. $\boldsymbol{\beta}^j$ represents how ad characteristics change mean viewing behaviors.[5]

The function $g(1_s, l_{bs}, t_{bs}; \Theta^j)$ estimates average changes in viewing behaviors during commercial breaks. The slot-specific indicator variables 1_s capture typical changes in viewing behaviors across ad slots. Ads in the sample range from 15 to 120 seconds, so it is also important for g to accommodate differences in advertisement durations, denoted l_{bs}, as well as the total time elapsed since the beginning of the break, t_{bs}.

δ^j_{ib} captures heterogeneity across viewers, breaks, and viewer-break combinations. δ^j_{ib} is an interaction effect that inherently nests: i) viewer-specific effects including viewer habits or viewing environment idiosyncrasies; ii) break-specific effects including time, program or network shocks, *e.g.* the program in which the break occurs, how much time has passed since the last break, the season of the year or the time of day; and iii) viewer-break interaction effects, such as how engaged viewer i is with the program or whether the viewer is watching the break during time-shifted programming. The flexibility of δ^j_{ib} comes from its high dimensionality given that there are 994,186 (i, b) combinations in the estimation sample. We estimate $\delta_{i,b}$ parameters using the method of

[5]We decided against using a discrete choice model because (i) response behaviors are continuous and (ii) choice sets are unobserved but vary across viewers and viewing sessions, *e.g.* during live or time-shifted viewing.

alternating projections (Guimaraes and Portugal, 2010).[6]

The error term, ε^j_{ibs}, captures any remaining omitted factors such as measurement error.

3.2 Causal Identification: Theory and Evidence

A small but growing literature has recently established that advertising endogeneity problems can be unusually severe. Lewis and Rao (2015) showed that small model specification errors can overwhelm treatment effects of digital banner ads on sales. Gordon et al. (2019) found that observational methods failed to recover experimental treatment estimates of Facebook ads on sales, even in huge samples with numerous covariates. Shapiro et al. (2021) found that careful endogeneity controls estimated smaller effects of TV ads on packaged good sales than correlational approaches.

The ideal experiment to identify ad content effects on viewing behaviors would randomize ads across audiences, brands, ad breaks and slots. However, we know that advertisers and viewers both self-select into commercial breaks (Tuchman et al., 2018). Therefore, we assume x_{bs} correlates with δ^j_{bs} in the *causal* models and estimate the δ^j_{bs} parameters. For comparison, we also report results of *descriptive* models in which the δ^j_{bs} parameters are treated as unobservables.

We then rely on broadcast TV networks' quasi-random ordering of ads within breaks, which implies x_{bs} is uncorrelated with ε^j_{ibs}. The television industry has long known or assumed that viewing behaviors change across ad slots within the commercial break, as confirmed in Figure A1. However, broadcast networks do not sell specific ad slots to advertisers. Advertisers purchase ad insertions based on networks, dates and quarter-hours, typically months in advance and without guarantees of what program the ad will be inserted into. The exclusion of ad slots from standard ad contractual terms can be explained by observing that Nielsen audience estimates do not vary

[6] 3.5% of sample ad exposures occur during viewer-break combinations with a single ad exposure, when viewer-break fixed effects are inseparable from ad effects. Ad effects are identical whether we drop or retain single-slot viewer-break combinations.

meaningfully between consecutive ad slots, likely due to the relative imprecision of People Meter presence measurements. Instead, standard TV ad sales contracts promise to rotate ads across slots on an "equitable" basis across commercial breaks (Mandese, 2004).

Quasi-random ordering of TV ads within commercial breaks is verifiable. If networks assign ads to slots using independent random draws, then the distribution of ad creatives' average slots should be Normal, by the Law of Large Numbers. To check, we define each ad insertion's position within its break as $\frac{s-1}{S_b-1}$, where S_b is the number of slots in break b. Thus, every ad position lies in [0,1] and the measure is comparable across various ad and break durations.

Figure 5, Panel A, plots the empirical distribution of average ad positions for the 1,384 advertised products with at least 50 ad insertions on broadcast networks.[7] The distribution appears approximately Normal. Panel B compares the empirical distribution of average ad positions to quantiles of a Normal distribution with the same mean and variance. There is a remarkably close correspondence. All 8 of the largest positive outliers are ads for sports programs that were most likely house ads run by program producers (e.g., NFL Online, USGA Organization, FedEx Sponsored Event). Overall, ad positions are verifiably consistent with networks' contractual promises of quasi-random ad ordering.

Quasi-random ordering does carry an important caveat. Some cable networks price ads by slot. In fact, average cable network ad slots depart meaningfully from quasi-random placement, as shown in Figure 5 Panel C. Therefore we excluded cable networks from the sample. Quasi-random ordering is also unlikely in addressable TV or other programmatic ad sales contexts.

Numerous papers rely on quasi-random ordering of TV ads within breaks to identify causal TV ad effects.[8] However, to the best of our knowledge, no prior study has empirically

[7]Figure 5 excludes promotions for network programs, which often appear before or after commercial breaks.

[8]Those include studies of TV ad avoidance (Wilbur et al., 2013); brand website visits and sales (Liaukonyte et al., 2015; He and Klein, 2018; Meder et al., 2019); social media chatter (Fossen and Schweidel, 2019); brand search and price search (Du et al., 2019); subsequent TV ads' audience and resulting digital chatter (Fossen et al., 2020); brand awareness, consideration and purchase (Tsai and Honka, 2021).

confirmed quasi-random assignments of ads to slots, so the confirmation may be a contribution to methodologically similar studies.

3.3 Interpretation of Effects

Consumers often use ad blockers or digital video recorders to avoid ad exposures, but avoidance behavior is seldom observable in advertising data. Therefore, most advertising studies estimate intent-to-treat (ITT) effects (Gordon et al., 2021; Tuchman et al., 2018). TVision data are unusual in that they enable direct measurements of ad treatments (i.e., exposures), enabling us to distinguish treatment effects from ITT effects.

We interpret the ad content effects on viewing behaviors as local average treatment effects (LATE) (Imbens and Angrist, 1994). LATE, by definition, conditions on all forms of selection, including advertiser and viewer selection into breaks. LATE estimates quantify how targeted ads changed viewing behaviors within targeted contexts, and therefore can inform advertiser evaluations of ad effects within targeted contexts.

LATE estimates may not extrapolate to untreated contexts, motivating the question of whether Average Treatment Effects (ATE) are estimable. If all potential contexts were treated, then LATE and ATE are equal. However, that is unlikely, because viewers and advertisers each select into breaks, and as a result advertised categories can co-occur within breaks.[9]

We sought to quantify the extent of advertiser targeting in the data. The average viewer is exposed to ads from 39% of the 167 ad categories; and the exposure-weighted average is 70%. The average ad category co-occurs within the same break as 76% of the 166 other ad categories; and the exposure-weighted average is 93%. Figure A3 illustrates the distributions of viewer category exposures and category co-occurrences within breaks while Figure A4 plots coverage

[9]We considered using Inverse Probability Weighting (IPW), as in single-treatment settings (Gordon et al., 2019; Rafieian and Yoganarasimhan, 2020). However, there is no consensus on how best to implement IPW in settings with either multivariate treatments or multiple treatments (Lopez and Gutman, 2017); our setting features both.

of ad categories against each viewer sorted by exposures. Overall, we conclude that broadcast TV ad targeting seems limited, consistent with its perceived role as a mass medium.

4 Findings

We present ad creative and category effects; duration, slot and time effects; ad feature effects; robustness checks; and a deeper case study of drug category ad effects.

4.1 Ad Creative Results

Figure 6 depicts six distributions of ad creative effects: one each for the descriptive and causal models, within each of the tuning, presence, and attention regressions. Each distribution characterizes 6,650 parameter estimates. We demean the distributions to focus on their shapes.

All six distributions are unimodal and nearly symmetric. The causal effect distributions vary less than their descriptive counterparts, with the greatest compression observed in the attention estimates. Specifically, the standard deviation of the causal tuning distribution is 12% smaller than in the descriptive tuning distribution, 20% smaller for presence, and 37% smaller for attention.

The greater variation in the descriptive distributions shows that ad effects on viewing behaviors covary with factors that predict ad assignments to highly-viewed breaks, such as viewer factors, break factors and viewer-break factors. The only difference in the models that generate the different results is whether the δ_{ib}^{j} parameters are estimated jointly with the ad creative effects or treated as unobservables as part of the error term.

Still, despite the compression within the causal effect distributions, the tails of those distributions contain some surprisingly large ad creative point estimates. For example, 5% of the point estimates in the causal tuning distribution exceed .039 in absolute value, more than the difference between tuning's average and its upper bound (.963 and 1.0, respectively). 5% of the presence point estimates exceed .048 in absolute value, and 5% of the attention point estimates exceed 0.030 in absolute value, both of which are surprisingly large compared to sample averages (e.g., .129 for attention), especially when considering that the regression separately accounts for slot effects and time-into-break effects.

We sought to better understand how individual ad creative estimates relate to sample sizes. Figure 7 turns the causal distributions on their side, showing how creative estimates covary with log(exposures). The x axis runs from e^4 to e^9, indicating that sample ad creatives vary 100-fold in total exposures, from 50 to 5,076. The trend lines show that the creative estimates are nearly uncorrelated with the number of exposures. The outlying estimates are all infrequent ads; the most frequently viewed ad creative estimates are more concentrated around zero.

It is possible to bootstrap ad creative standard errors, but we prefer not to risk interpreting noise. We also investigated replacing the ad creative fixed effects with the 1,504 brand fixed effects, but again found a pattern quite similar to Figures 6 and 7: outlier estimates tended to be brands with limited ad exposures and thus indistinguishable from noise. We focus instead on a model which shrinks the ad creative effects toward product category identifiers and ad content features.

4.2 Ad Category Results

Figure 8 depicts the causal and descriptive distributions of product category effects. These distributions are unimodal with the causal distributions again being tighter than their descriptive counterparts. However, the degree of compression is similar across the three viewing behaviors. We observe 22%, 26%, and 21% decreases in standard deviations between the descriptive and causal models for tuning, presence, and attention, respectively ($SD^T_{Desc} = 0.0010$, $SD^T_{Caus} = 0.008$; $SD^P_{Desc} = 0.013$, $SD^P_{Caus} = 0.009$; $SD^A_{Desc} = 0.009$, $SD^A_{Caus} = 0.007$). Ad category effect distributions are tighter than ad creative effect distributions: 5% of the point estimates in the causal tuning distribution exceed 0.015 in absolute value, 5% of the presence point estimates exceed 0.017 in absolute value, and 5 of the attention point estimates exceed 0.011 in absolute value.

Table 1 shows evidence that product category estimates contain reliable signals. A 5% error rate predicts 8.35 false positives among the 167 category effects in any of the six regressions due to random chance alone. The causal model results indicate 32 significant category effects on tuning, 20 significant effects on presence, but only 8 significant effects on attention. Therefore we report but do not really interpret the category effects in the attention regression.

Figure 9 shows that ad category parameter significance is not driven by low-powered outliers. The majority of significant results do not occur among the lowest-powered coefficients. The tuning and presence panels show that many significant positive and significant negative category causal effects are well powered.

Figure 10 unpacks the results presented in Figure 9, highlighting the 40 highest and lowest ad category fixed effects as ranked by tuning estimates. Many of the largest category effects relate to recreation, including Hunting & Fishing; Casinos & Gambling; Wine, Alcohol & E-Cigarettes; Dating; Sports; Movies; and Airlines, whose television ad content promotes leisure travel. Many of the most negative category effects relate to prescription drugs, including drug categories treating Cancer; Depression, Bipolar & Insomnia; Alzheimers & Multiple Sclerosis; Psoriasis, Skin & Nails; Osteoporosis & Arthritis; Varied conditions; Bladder & Gastrointestinal; and Stroke, Cholesterol & Heart Disease.

The presence regression estimates mostly align with the tuning results, but exhibit larger standard errors. The categories that reliably retain viewer presence include Wine, Alcohol & E-Cigarettes; Underwear; Car Rental; Sports; Clothing; Speakers & Headphones; Movies; Legal Services; and Shoes & Socks. On the other end, 6 of the 7 largest significant negative findings again feature prescription drugs.

66 of the 80 category effects on attention have confidence intervals that lie entirely between -2% and 2% of ad seconds.[10] Attention requires both tuning and presence, by definition, so we expected attention results to resemble tuning and presence results. It is possible that ingrained habits drive viewer attention more than on-screen ad content. Another possibility is that ad content is polarizing: if some content reliably attracts attention from viewers interested in the product market, it may simultaneously lead uninterested viewers to divert their attention, change channels or leave the room.

4.3 Duration, Slot and Time Effects

Table A5 presents ad duration, slot and time-elapsed effect estimates. Figure 11 graphs how ad durations change viewing behaviors. 30-second ads reduce tuning by an absolute 2.7%, presence by 5.6%, and attention by 0.7%. 60-second ad duration effects are approximately double the absolute 30-second duration effects, per ad second. 60-second ads reduce tuning by an absolute 5.5%, presence by 10.9%, and attention by 1.5%.

Next we look at slot and time-elapsed effects. The modal ad break contains seven slots and the modal ad duration is 30 seconds. Figure 12 graphs slot and time-elapsed effects on viewing behaviors for a hypothetical break consisting of seven 30-second ads. Standard errors of the combined effects are calculated by bootstrapping out of the joint asymptotic distribution of the parameter estimates, including off-diagonal terms. Confidence intervals widen throughout the break because audience-remaining calculations in later slots depend on more earlier slots' parameter estimates.

Tuning decreases across slots with changes driven primarily by the time-elapsed variables. Presence shows a similar absolute decrease but is significantly impacted by both slot effects

[10] Limited absolute changes could still be appreciable on a relative basis given that average attention is near 13%.

and time-elapsed variables, as shown in Table 2. Attention decreases in the first two slots, but confidence intervals overlap from positions three to seven, and time-elapsed variables are not significant predictors. Therefore, although audience falls throughout the break, attention does not always decrease after the third slot.

4.4 Ad Feature Results

We measured ad content because ad viewing behaviors may respond to stimuli displayed on television screens. However, we interpret the following ad feature results with caution, given the caveats about unobserved features, feature measurement error, and feature availability.

Figure 13 presents iSpot and machine-coded ad feature causal effects on tuning, presence and attention. As before, tuning and presence results are more precisely estimated than attention, but all of the effects are small on an absolute basis. Sales-related content like taglines and promotions reduce tuning and presence, similar to findings in Teixeira et al. (2010). Surprisingly, a higher sentiment score reduces tuning, and professional actors reduce both tuning and presence. It seems plausible that brands may try to make otherwise unattractive ad messages more palatable by using professional actors or more positive scripts.

Another surprise is that the movie dummy reduces tuning and presence, given that the movies product category effect increases ad viewing. We investigated this more carefully by examining the overlap between the movies category dummy and the iSpot movie classifier. The iSpot variable indicates both theatrical movie trailers and also the presence of theatrical movie brands in co-branded advertisements, such as for packaged goods, cars, fast food and retail chains. If we drop the movie category dummy, then the iSpot movie effect becomes positive and significant. Therefore, it appears that movie ads increase viewing, but movie co-branding in non-movie ads reduces tuning and presence.

The engagement variable measures ad traffic across iSpot's video, social and search channels, and increases ad viewing behaviors. Ads classified by iSpot as having a "sexy" mood reduce tuning, but other mood variable effects are estimated imprecisely.

Two of the machine-coded features have significant effects. The number of scenes within an ad reduces tuning and presence, as does the duration of neutral facial expressions shown on screen. Other machine-coded features generally have point estimates near zero.

Figure 14 presents the effects of Google Cloud Vision features on viewing behaviors. The confidence intervals are much smaller than in Figure 13 and the large majority of feature labels do not have significant effects. Further, those few features that do have significant effects resist easy interpretation. For example, one might have predicted that Infant or Party might have increased viewing, but Infant is near zero and Party is negative. We again recall the caveat that labeled features may correlate with important unlabeled features, such as when brands pair less attractive selling messages with more attention-grabbing stimuli.

4.5 Robustness Checks

We view the ad category results as the most interesting set of causal effects, so we investigate how stable they are to alternate model specifications. Figure A5 shows that ad category findings are nearly identical when ad content features are removed from the model. Remarkably, none of the 480 depicted category point estimates falls outside the other model's confidence interval, suggesting that the category findings may be highly reliable.

Tuning is a device-level property, but viewing often occurs among groups of viewers. Figure A6 shows that ad category effects on tuning are generally robust to exclusion of multiple-viewer viewing sessions.

We estimated a version of the model including an indicator of ad awards (Clio, Effie or Emmy).

The award dummy was directionally positive but not statistically distinguishable from zero.

Finally, we investigated the ability of the data to estimate heterogeneous ad effects. Figure A7 shows that ad category confidence intervals overlap when the sample is partitioned between demographic groups.

4.6 Case Study: Prescription Drugs

Pharmaceutical drug advertisements tend to cause viewers to tune away and leave the room. Moreover, the most negative effects relate to serious conditions such as cancer, depression and Alzheimer's. We quantify how drug category attributes relate to ad effects on viewing behaviors.

Background. Previous research has found that pharmaceutical advertising tends to increase drug category consumption (Iizuka and Jin, 2005, 2007; Sinkinson and Starc, 2018; Shapiro, 2018, 2020). If drug advertising enhances general welfare, then we should care about factors that may influence drug ad pricing. Digital ad sellers Facebook and Google typically include some element of consumer acceptance or rejection of ads in their advertising pricing algorithms, as earlier ads affect attention paid to subsequent ads. Broadcast television networks do not publish ad pricing algorithms, but they too may use audience reactions to price ads. Here we seek to quantify what drug category factors correlate with larger or smaller audience losses during drug ads.

Viewers may avoid pharmaceutical ads for two distinct but related reasons. First, drug ads may present viewers with unpleasant reminders of prevalent adverse health conditions.[11] Second, advertised drugs may present viewers with unpleasant reminders of particularly severe health conditions.

[11] Alternatively, drug ads for more prevalent conditions may be more relevant to a wider set of viewers, but we question whether those viewers would prefer to receive such messages during television program consumption.

Data. We collected objective measures of treated condition prevalence and severity for each pharmaceutical category. Prevalence is measured as the case rate, or the percentage of US residents who experience the disease or condition in a year. Severity is measured in Disability-Adjusted Life Years (DALYs), which estimates years of life lost due to premature death and years of healthy life lost due to poor health or disability, reflecting both mortality and morbidity in a single measure. Prevalence and severity measures come from U.S. data compiled by the Institute for Health Metrics and Evaluation's Global Burden of Disease studies.

Table A4 presents the category prevalence and severity data. The two measures correlate at -0.2 at the category level. A few categories have both high prevalence and high severity (e.g., stroke, cholesterol and heart disease; depression, bipolar and insomnia). Most other categories have either high prevalence or high severity.

Model and Results. We run a second stage regression of pharmaceutical category effect estimates from Equation 1 on disease prevalence and severity using Equation 2.

$$\hat{\beta}_k^j = \alpha^j + \gamma_1^j \text{PREV}_k + \gamma_2^j \text{SEV}_k + \varepsilon_k^j \tag{2}$$

$\hat{\beta}_k^j$ is the causal effect estimate of prescription category k ads on viewing behavior j (tuning, presence, or attention); PREV_k and SEV_k are the prevalence and severity of drug category k.

We account for first-stage estimation error using the estimated asymptotic joint distributions of the point estimates. Specifically, we estimate Equation 2 using Generalized Least Squares (GLS) with the estimated variance-covariance matrix, $\hat{\Omega}_{Rx}^j$, where $\hat{\Omega}_{Rx}^j$ is the relevant subset of the full variance-covariance matrix of the parameters estimated in the first stage regression, covering only the pharmaceutical category effects, including the off-diagonal terms. Intuitively, estimating Equation 2 via GLS gives more weight to the more precise drug category estimates.

Table 3 presents the estimation results from Equation 2. The table also shows restricted models that contain each predictor individually. We view the results in Column 3 as the most informative specification. We interpret the results as descriptive as drug firms choose ads strategically and no instruments are available for unobserved category attributes.

Drug category prevalence and severity are both associated with more negative category effects on TV ad tuning and presence. The point estimates show increasing drug category severity by one million DALYs is associated with a 0.07% decrease in tuning and a 0.05% decrease in presence. Increasing drug category prevalence by 1% of the population correlates with a 0.05% decrease in tuning and a 0.5% decrease in presence.

Discussion. The second stage results relating disease prevalence and severity to viewing behavior during drug ads are correlational but we think they are useful for three reasons. First, they point to underlying factors that predict viewer response to advertisements, helping to make some logical sense of the many disparate ad category results. Second, the US regulates direct-to-consumer advertising, but is one of the only two developed economies that allows it. Drug ads remain controversial (Sheehan, 2013), so quantifying the general market's response to drug TV ads may help to inform ad content regulations. Third, prior literature has found that drug ads increase socially desirable outcomes and that drug brands under-advertise relative to the social optimum (Shapiro, 2018, 2020). It is possible that television ad pricing policies may restrict positive public health outcomes by further limiting the reach of drug ads.

5 Can Viewing Metrics Predict Ad Response?

Next, we investigate whether new ad viewing metrics can predict audience response to ads. To do this, we require an ad response measure that is publicly available, reliably measured, temporally disaggregated, relevant and comparable across brands. We focus on online brand search, which has been studied extensively in relation to TV ads (Zigmond and Stipp, 2010; Chandrasekaran et al., 2018; Du et al., 2019; Liaukonyte and Žaldokas, 2021; Lambrecht et al., 2021) and found to predict changes in brand attitudes (Dotson et al., 2017) and sales (He et al., 2014).

The analysis combines data from TVision Insights and Google Trends. We select a sample of all 180 national ads aired during two NFL football playoff games on January 22, 2017, as they had large live audiences.[12]

Specifically, we ask the following questions: Can ad viewing measures predict brand search lift? How do empirical relationships differ between ad OTS and ad exposures? Answers may indicate how new viewing metrics relate to advertiser objectives.

Ad Data. Figure 15 illustrates four total viewing measures—OTS tuning, exposure tuning, presence and attention—across the 180 ad slots, in chronological order with vertical lines between ad breaks. As before, ad OTS and ad exposures differ in the number of potentially exposed viewers who are absent during tuned ads. As expected, viewing typically decreases within breaks and increases between most breaks.

Search Volume Data. We downloaded minute-by-minute Google Trends reports for all brands advertised in the two games, using the main brand keyword for each ad. Google provides relative search volume indices, so every query also included the control keyword "Pizza Hut" to set a standard, comparable scale across Google Trends reports.

Figure 16 depicts the search data for 25 brand keywords from a variety of categories; full data are in Figure A8. Baseline search levels and lifts varied across brands. Search sometimes spiked

[12] 21% of TVision panelists were exposed to the ads. For comparison, Nielsen estimated audience ratings of 24.4 and 25.0 for the two games (*Sports Media Watch*, 2017). The differences likely reflect sample selection and measurement differences between TVision and Nielsen, as well as sampling error. We do not analyze Super Bowl data because many viewers gather specifically to watch Super Bowl ads, a nonmodal behavior.

without an ad (e.g., Apple, Verizon), likely due to TV ads on other channels or in-game sponsor messages.

Model. Equation 3 relates brand search lift to viewing behaviors as follows:

$$LIFT_s = \eta Y_s^k + \phi Z_s + \varepsilon_s \tag{3}$$

where s indexes the 180 ad slots, each of which represents a unique combination of brand b and minute t. $LIFT_s$ measures brand search in minutes t and $t+1$ minus brand search in minutes $t-1$ and $t-2$. Y_s^k is one of the four total viewing measures, indexed by k, in slot s. The smallest correlation among total viewing measures is 0.83 (presence/attention), so including multiple viewing measures induces multicollinearity. Z_s includes an intercept, a first-slot dummy, and, in some specifications, 22 quarter-hour dummies to control for time-varying drivers of advertiser targeting and brand search (e.g., tension within the game).

Results. Table 4 presents eight regressions, one for each of the four total viewing measures, with and without quarter-hour dummies. There are two key results:

1. OTS Tuning does not reliably predict search lift. However, exposure tuning, presence and attention all significantly predict search lift without quarter-hour fixed effects. Attention explains the most variation in search lift, followed by exposure tuning.

2. Attention significantly predicts search lift in the regressions with quarter-hour fixed effects.

Interpretation. The regressions suggest that advertisers can best explain search lift by measuring attention directly, or at least distinguishing between ad OTS and exposures. We conclude that the new viewing measures may help inform advertiser objectives.

We note two important caveats. First, this excercise is a small study of only two particular football games. A larger-scale study could yield more representative, better powered results, which may vary across brands, audiences, creatives, programs or time. Second, we focused on brand search lift because of its convenient measurement across brands and time. Individual advertisers likely should consider proprietary response metrics (e.g., website visits, app usage, sales), as they would better indicate ad profits.[13]

6 Discussion

New TV ad viewing metrics offer the first passive, continuous, *in situ* measures of video ad exposures, showing that 30% of TV ads play to empty rooms. We constructed ad features and used a verifiably quasi-experimental identification strategy to estimate how ad content influences ad viewing. Recreational product categories tend to preserve tuning and presence, whereas prescription drug ads tend to reduce tuning and presence, especially for prevalent and severe conditions. Attention falls during longer ads and early in commercial breaks. A supplementary, spot-level analysis finds that attention predicts brand search lift better than traditional metrics.

Implications for advertisers. Brand advertisers typically do not have granular outcome data to discipline their campaign choices, raising the question of whether attention data could serve as an intermediate success metric. The brand search lift analysis reinforces the idea that attention may predict ad response. However, viewing metrics have to be filtered carefully between targeted and nontargeted ad viewers. For example, the pharmaceutical category results show how the mass audience reacts to drug ads, but they do not imply that drug ads don't work. A cancer patient likely reacts quite differently to a cancer drug ad than the median viewer.

[13]In particular, if a brand advertises on its own keyword, it may have to pay Google a significant toll (Simonov and Hill, 2021) and some resulting searches could be accelerated in time rather than incremental (Lambrecht et al., 2021).

Performance campaigns can directly test the empirical value of the new viewing metrics. The search lift analysis suggests that measuring attention and distinguishing OTS from exposures both may help to predict ad response and improve media buying.

Implications for platforms and viewers. Ads that reduce viewing decrease subsequent ads' audience. In general, video ad attention depends on program environment and earlier ads (Webb, 1979; Burke and Srull, 1988; Rajaram et al., 2021; Joo et al., 2020; Liu et al., 2021). Closely related evidence exists in mobile advertising (Rafieian and Yoganarasimhan, 2020) and search advertising (Gomes et al., 2009). Importantly, highly involving ads also may help retain audience in subsequent ad slots (Fossen et al., 2020).

Media platforms—e.g., Hulu, NBC, New York Times, YouTube—restrict ad content to help preserve viewer attention and brand safety. Many also design ad markets that select, target, order and price ads to internalize short-run negative externalities between ads (Wilbur et al., 2013; Theocharous et al., 2015; Kar et al., 2015; Deng and Mela, 2018; Rajaram et al., 2021; Rafieian, 2019).

Attention preservation also has long-run implications. In particular, it may affect the number and types of viewers attracted to a platform, and the habitual ad responses they develop. Incorporating attention metrics in ad market design could be a rare "win-win-win" for advertisers, platforms, and consumers.

Limitations The current paper has several important limitations, as does all research. We relied exclusively on quasi-random variation to estimate causal effects, but that may have over-controlled for endogeneity. We estimated local average treatment effects rather than average treatment effects or heterogeneous treatment effects. We presume that consumers often divert attention during ads because of second screening behaviors, but we are not yet able to measure such behaviors directly. We have not investigated coviewing, ad frequency or other related topics.

Extensions. Many opportunities remain unexplored. TV ad field experiments are too rare. Randomization could help identify effects of ad insertion, targeting and content on viewing metrics and sales. Further, no one has directly quantified how presence or attention influence conversions, or vice versa.

Conclusion. In summary, new viewing metrics differ meaningfully from traditional metrics, respond to ad content and help to predict ad response. There has been limited competition among TV ad measurement providers in the past, but improved viewer presence and attention measures could improve video ad transactions and overall market efficiency.

References

Anderson, S. P., Ciliberto, F., and Liaukonyte, J. (2013). Information content of advertising: Empirical evidence from the otc analgesic industry. *International Journal of Industrial Organization*, 31(5):355–367.

Anderson, S. P., Ciliberto, F., Liaukonyte, J., and Renault, R. (2016). Push-me pull-you: comparative advertising in the otc analgesics industry. *The RAND Journal of Economics*, 47(4):1029–1056.

Blanchard, S. J., Noseworthy, T. J., Pancer, E., and Poole, M. (2020). Visual scene elements and new product success: An analysis of crowdfunding product images using automated image classification. *Working Paper*.

Bronnenberg, B. J., Dube, J.-P., and Mela, C. F. (2010). Do digital video recorders influence sales? *Journal of Marketing Research*, 47:998–1010.

Burke, R. R. and Srull, T. K. (1988). Competitive interference and consumer memory for advertising. *Journal of Consumer Research*, 15(1):55–68.

Chandrasekaran, D., Srinivasan, R., and Sihi, D. (2018). Effects of offline ad content on online brand search: insights from super bowl advertising. *Journal of the Academy of Marketing Science*, 46:403–430.

Chintagunta, P. K. and Nair, H. S. (2011). Structural workshop paper—discrete-choice models of consumer demand in marketing. *Marketing Science*, 30(6):977–996.

Danaher, P. J. (1995). What happens to television ratings during commercial breaks? *Journal of Advertising Research*, 35(1):37–37.

Deng, Y. and Mela, C. F. (2018). TV viewing and advertising targeting. *Journal of Marketing Research*, 55(1):99–118.

Dotson, J. P., Fan, R. R., McDonnell Feit, E., Oldham, J. D., and Yeh, Y.-H. (2017). Brand attitudes and search engine queries. *Journal of Interactive Marketing*, 37:105–116.

Du, R. Y., Xu, L., and Wilbur, K. C. (2019). Immediate responses of online brand search and price search to tv ads. *Journal of Marketing*, 83(4):81–100.

Ephron, E. (2006). Who watches the watchers. Archive.org link. Accessed: 2020-12-15.

FDA (2020). Prescription drug advertising. Archive.org link. Accessed: 2020-12-15.

Fossen, B. L., Mallapragada, G., and De, A. (2020). Impact of political television advertisements on viewers' response to subsequent advertisements. *Marketing Science*.

Fossen, B. L. and Schweidel, D. A. (2019). Social TV, advertising, and sales: Are social shows good for advertisers? *Marketing Science*, 38(2):274–295.

Goettler, R. L. and Shachar, R. (2001). Spatial competition in the network television industry. *RAND Journal of Economics*, pages 624–656.

Gomes, R., Immorlica, N., and Markakis, E. (2009). Externalities in keyword auctions: An empirical and theoretical assessment. *Proc. of International Workshop on Internet and Network Economics 2009*, pages 172–183.

Gordon, B. R., Jerath, K., Katona, Z., Narayanan, S., Shin, J., and Wilbur, K. C. (2021). Inefficiencies in digital advertising markets. *Journal of Marketing*, 85:7–25.

Gordon, B. R., Zettelmeyer, F., Bhargava, N., and Chapsky, D. (2019). A comparison of approaches to advertising measurement: Evidence from big field experiments at Facebook. *Marketing Science*, 38(2):193–225.

Guimaraes, P. and Portugal, P. (2010). A simple feasible procedure to fit models with high-dimensional fixed effects. *The Stata Journal*, 10(4):628–649.

He, C. and Klein, T. (2018). Advertising as a reminder: Evidence from the Dutch State Lottery. *Available at SSRN 3184756*.

He, Y., Du, R. Y., and Damangir, S. (2014). Decomposing the impact of advertising: Augmenting sales with online search data. *Journal of Marketing Research*, 51:300–319.

Iizuka, T. and Jin, G. Z. (2005). The effect of prescription drug advertising on doctor visits. *Journal of Economics Management Strategy*, 14(3):701–727.

Iizuka, T. and Jin, G. Z. (2007). Direct to consumer advertising and prescription choice. *The Journal of Industrial Economics*, 55(4):771–771.

Imbens, G. W. and Angrist, J. D. (1994). Identification and estimation of local average treatment effects. *Econometrica*, 62(2):467–475.

Jayasinghe, L. and Ritson, M. (2013). Everyday advertising context: An ethnography of advertising response in the family living room. *Journal of Consumer Research*, 40(1):104–121.

Joo, M., Liu, W., and Wilbur, K. C. (2020). Divergent temporal courses for liking versus wanting in response to persuasion. *Emotion*, 20(2):261–270.

Kar, W., Swaminathan, V., and Albuquerque, P. (2015). Selection and ordering of linear online video ads. In *Proceedings of the 9th ACM Conference on Recommender Systems*, pages 203–210.

Knauer, L. (2019). Revised MRC viewability guidelines: Everything you need to know. Archive.org link. Accessed: 2021-03-26.

Kubany, A., Ishay, S. B., Ohayon, R. S., Shmilovici, A., Rokach, L., and Doitshman, T. (2020). Comparison of state-of-the-art deep learning APIs for image multi-label classification using semantic metrics. *Expert Systems with Applications*, 161:113656.

Lambrecht, A., Tucker, C. E., and Zhang, X. (2021). Does TV advertising increase online sales: The role of inter-temporal substitution. *Working Paper. Available at SSRN 3530105*.

Lee, D., Hosanagar, K., and Nair, H. S. (2018). Advertising content and consumer engagement on social media: Evidence from facebook. *Management Science*, 64(11):5105–5131.

Lewis, R. A. and Rao, J. M. (2015). The unfavorable economics of measuring the returns to advertising. *The Quarterly Journal of Economics*, 130(4):1941–1973.

Liaukonyte, J., Teixeira, T., and Wilbur, K. C. (2015). Television advertising and online shopping. *Marketing Science*, 34(3):311–330.

Liaukonyte, J. and Žaldokas, A. (2021). Background noise? TV advertising affects real-time investor behavior. *Management Science*.

Liu, X., Shi, S. W., Teixeira, T., and Wedel, M. (2018). Video content marketing: The making of clips. *Journal of Marketing*, 82(July):86–101.

Liu, X., Shum, M., and Uetake, K. (2021). Passive and active attention to baseball telecasts: Implications for content (re-)design. *Working Paper. Available at SSRN 3717894*.

Lopez, M. J. and Gutman, R. (2017). Estimation of causal effects with multiple treatments: a review and new ideas. *Statistical Science*, pages 432–454.

Mandese, J. (2004). Equitable's no longer questionable: Data reveals nets position some advertisers better than others. Archive.org link. Accessed: 2020-1-07.

Meder, C. J., Kemper, J., and Brettel, M. (2019). How creative cues determine online response to TV commercials: An empirical investigation for digital-native brands.

Rafieian, O. (2019). Optimizing user engagement through adaptive ad sequencing. *Working Paper*.

Rafieian, O. and Yoganarasimhan, H. (2020). How does variety of previous ads influence consumer's ad response? *Working Paper. Available at SSRN 3647373*.

Rajaram, P., Manchanda, P., and Schwartz, E. M. (2021). Finding the sweet spot: Ad targeting on streaming media. *Working Paper. Available at SSRN 3496039*.

Resnik, A. and Stern, B. L. (1977). An analysis of information content in television advertising. *Journal of Marketing*, 41(1):50–53.

Schweidel, D. A. and Kent, R. J. (2010). Predictors of the gap between program and commercial audiences: An investigation using live tuning data. *Journal of Marketing*, 74(3):18–33.

Schwenzow, J., Hartmann, J., Schikowsky, A., and Heitmann, M. (2021). Understanding videos at scale: How to extract insights for business research. *Journal of Business Research*, 123:367 – 379.

Shachar, R. and Emerson, J. W. (2000). Cast demographics, unobserved segments, and heterogeneous switching costs in a television viewing choice model. *Journal of Marketing Research*, 37(2):173–186.

Shapiro, B. (2020). Promoting wellness or waste? evidence from antidepressant advertising. *American Economic Journal: Microeconomics*, (Forthcoming).

Shapiro, B., Hitsch, G. J., and Tuchman, A. (2021). TV advertising effectiveness and profitability: Generalizable results from 288 brands. *Econometrica*, 89(4):1855–1879.

Shapiro, B. T. (2018). Positive spillovers and free riding in advertising of prescription pharmaceuticals: The case of antidepressants. *Journal of Political Economy*, 126(1):381–437.

Sheehan, K. B. (2013). *Controversies in Contemporary Advertising*. Sage Publications.

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

Simonov, A. and Hill, S. (2021). Competitive advertising on brand search: Traffic stealing and click quality. *Marketing Science*.

Sinkinson, M. and Starc, A. (2018). Ask your doctor? direct-to-consumer advertising of pharmaceuticals. *The Review of Economic Studies*, 86(2):836–881.

Sports Media Watch (2017). 2016 NFL TV ratings. Archive.org link. Accessed: 2021-08-10.

Swaminathan, S. and Kent, R. (2013). Second-by-second analysis of advertising exposure in TV pods: The dynamics of position, length, and timing. *Journal of Advertising Research*, 53(1):91–100.

Swant, M. (2019). Results of iab tech lab's blockchain pilot shows how many digital ads are seen by humans. Archive.org link. Accessed: 2021-03-26.

Teixeira, T., Wedel, M., and Pieters, R. (2012). Emotion-induced engagement in internet video advertisements. *Journal of Marketing Research*, 49(2):144–159.

Teixeira, T. S., Wedel, M., and Pieters, R. (2010). Moment-to-moment optimal branding in TV commercials: Preventing avoidance by pulsing. *Marketing Science*, 29(5):783–804.

Theocharous, G., Thomas, P. S., and Ghavamzadeh, M. (2015). Personalized ad recommendation systems for life-time value optimization with guarantees. In *Twenty-Fourth International Joint Conference on Artificial Intelligence*.

Tsai, Y.-L. and Honka, E. (2021). Informational and non-informational advertising content. *Marketing Science*, (Forthcoming).

Tuchman, A. E., Nair, H. S., and Gardete, P. M. (2018). Television ad-skipping, consumption complementarities and the consumer demand for advertising. *Quantitative Marketing and Economics*, 16(2):111–174.

Tucker, C. E. (2015). The reach and persuasiveness of viral video ads. *Marketing Science*, 34(2):281–296.

Voorveld, H. A. and Viswanathan, V. (2015). An observational study on how situational factors influence media multitasking with TV: The role of genres, dayparts, and social viewing. *Media Psychology*, 18(4):499–526.

Webb, P. H. (1979). Consumer initial processing in a difficult media environment. *Journal of Consumer Research*, 6(3):225–236.

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

Wilbur, K. C. (2016). Advertising content and television advertising avoidance. *Journal of Media Economics*, 29(2):51–72.

Wilbur, K. C., Xu, L., and Kempe, D. (2013). Correcting audience externalities in television advertising. *Marketing Science*, 32(6):892–912.

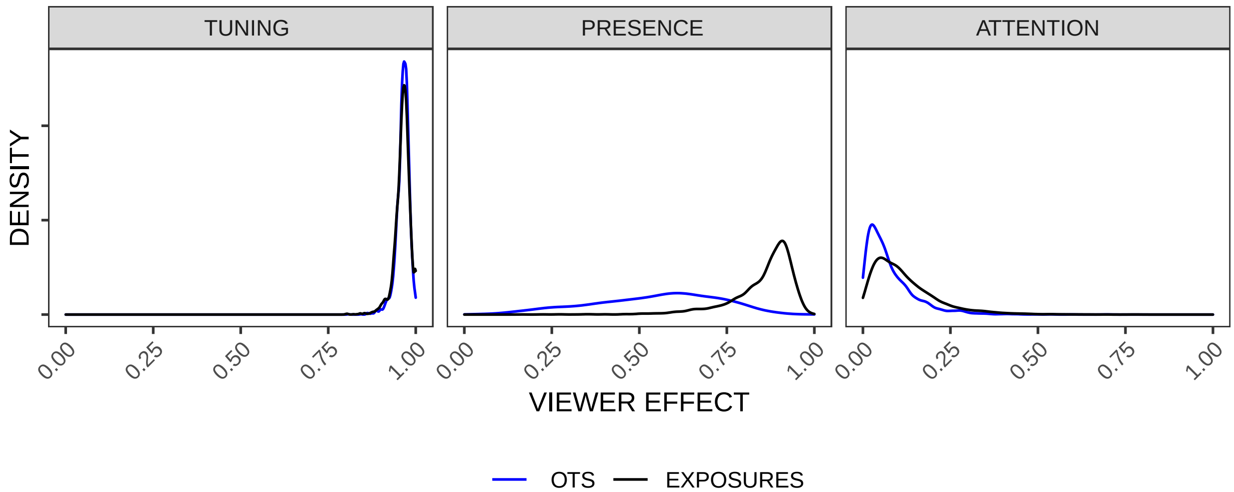
Zhang, J., Wedel, M., and Pieters, R. (2009). Sales effects of attention to feature advertisements: A bayesian mediation analysis. *Journal of Marketing Research*, 46(5):669–681.

Zigmond, D. and Stipp, H. (2010). Assessing a new advertising effect: Measurement of the impact of television commercials on internet search queries. *Journal of Advertising Research*, 50(2):162–168.

Zufryden, F. S., Pedrick, J. H., and Sankaralingam, A. (1993). Zapping and its impact on brand purchase behavior. *Journal of Advertising Research*, 33(1):58–67.

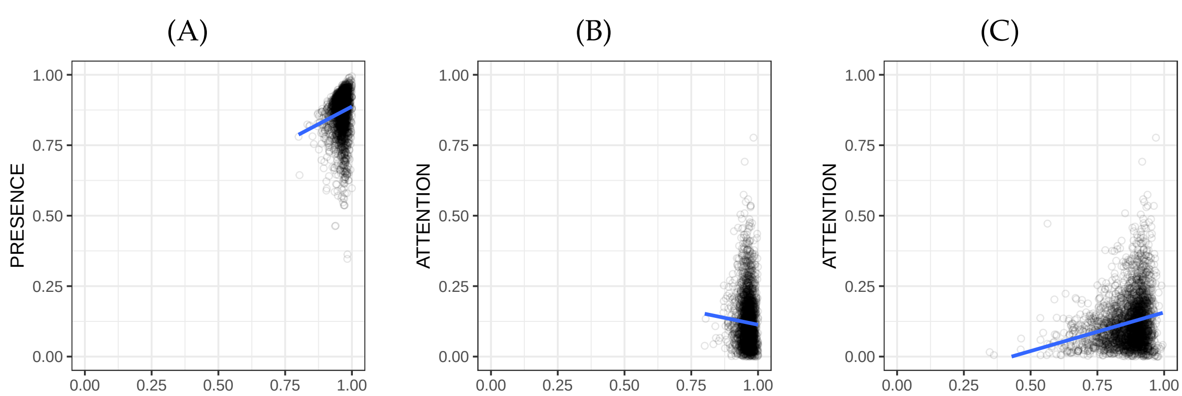
Figures

Figure 1: Densities of Average Viewer Behaviors



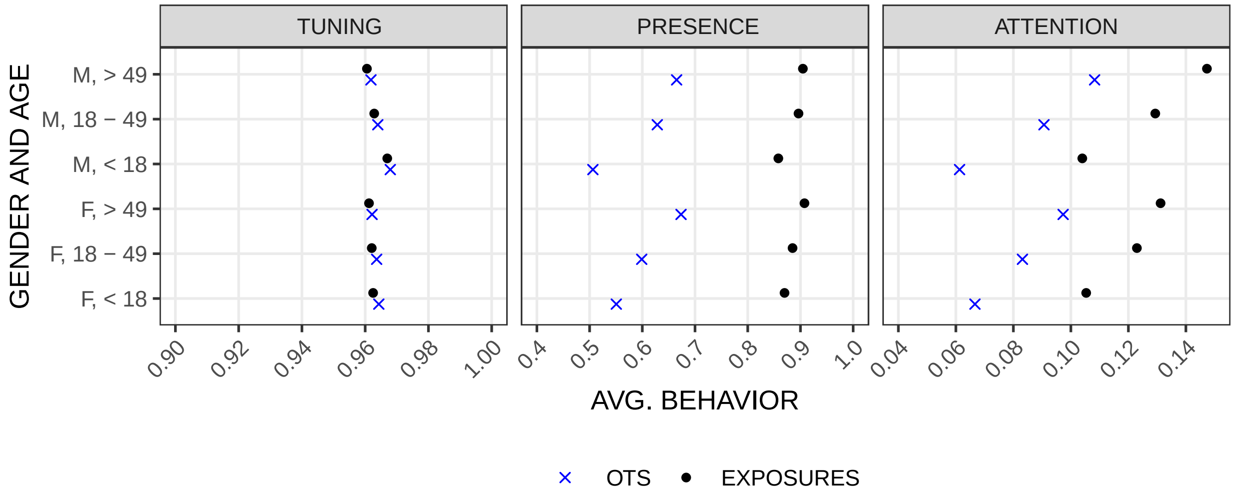
Notes: The three panels show densities of the viewer-level average behaviors among the 3,659 viewers. Opportunities To See (OTS) are defined as ads which are tuned for at least two seconds, during breaks when the viewer was present for at least 2 seconds in the first slot of the break. Exposures (EXP) are defined as any ad during which the viewer is both tuned and present for at least two seconds during the ad's slot.

Figure 2: Covariation in Viewer-Level Average Tuning, Presence, and Attention



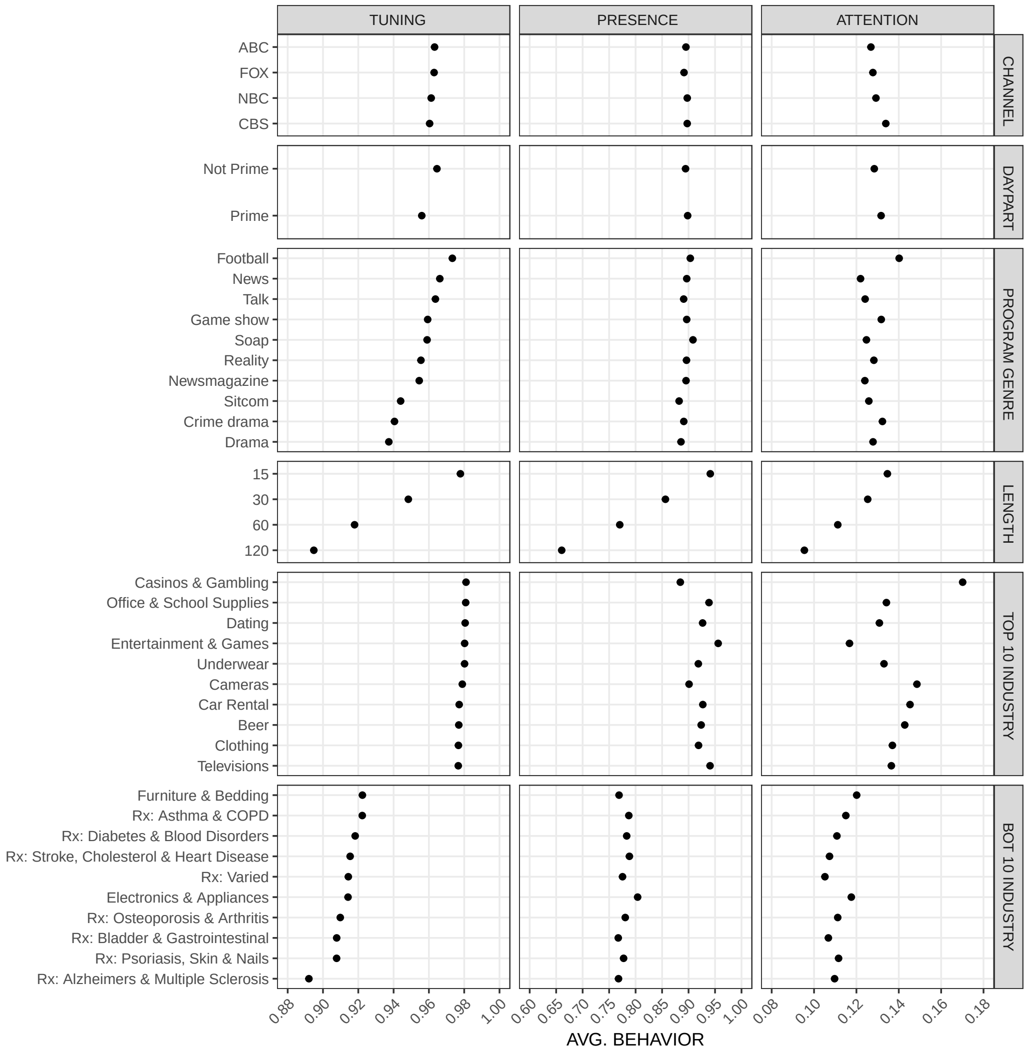
Notes: Each panel presents pairwise comparisons of viewer average (A) tuning and presence, (B) tuning and attention, and (C) presence and attention behaviors during advertising exposures.

Figure 3: Ad Viewing by Viewer Gender and Age



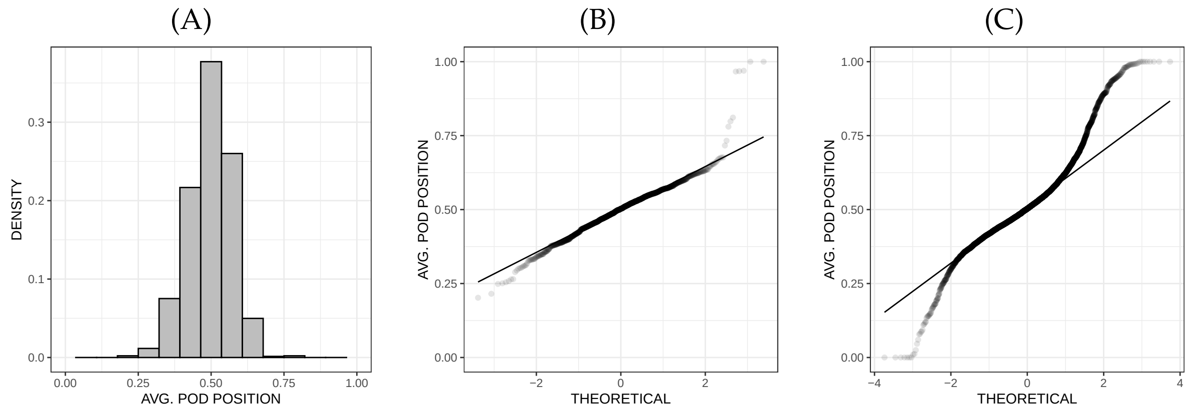
Notes: The three panels show averages of viewing behaviors by viewer gender and age. Opportunities to see (OTS) are defined as ads which are tuned for at least two seconds, during breaks when the viewer was present at the start of the break. Exposures are defined as any ad during which the viewer is both tuned and present for at least two seconds.

Figure 4: Viewing Behaviors during Ad Exposures by Break and Ad Characteristics



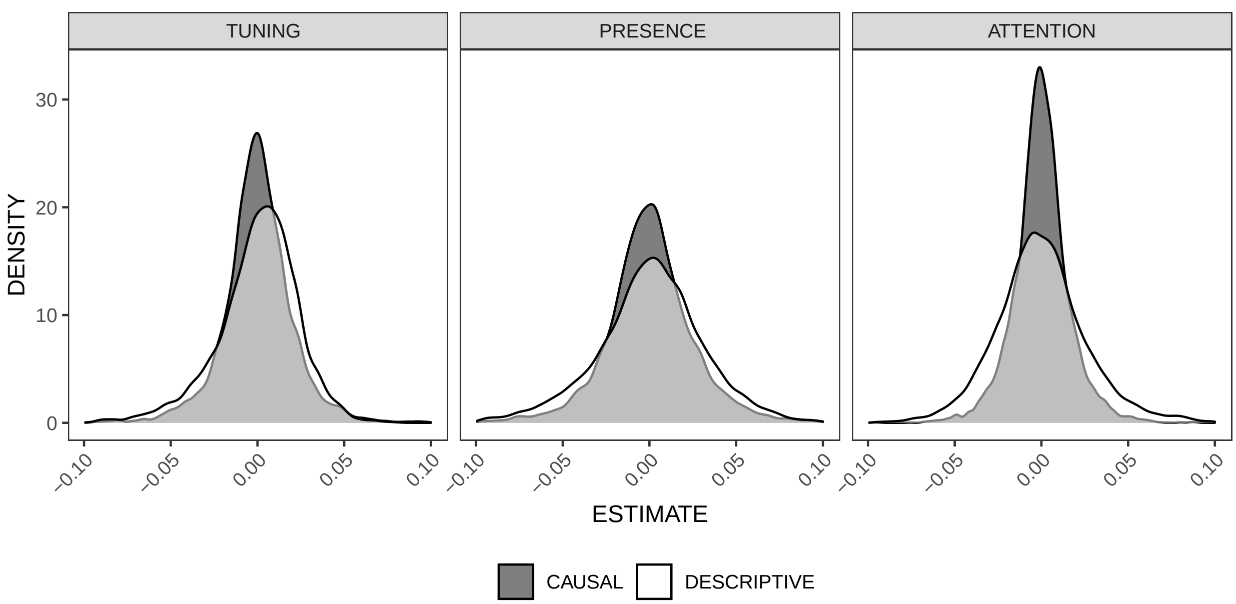
Notes: This figure presents average tuning, presence, and attention behaviors across channels, dayparts, program genres, ad lengths, and top and bottom ad categories (as ranked by average tuning).

Figure 5: Randomization Check



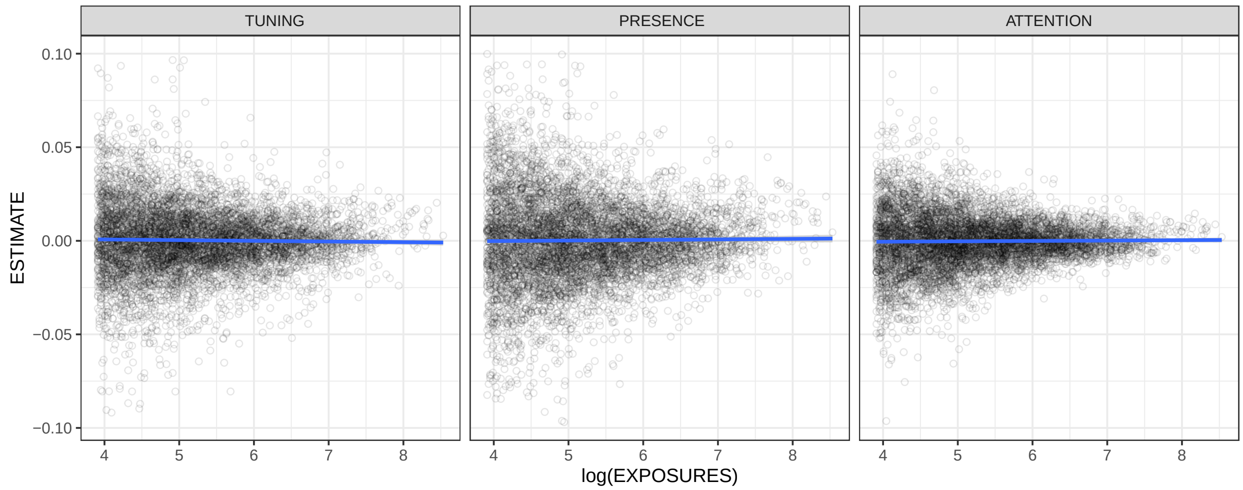
Notes: Panel A shows the empirical distribution of average ad position during broadcast network commercial breaks for the estimation sample of 1,384 advertised brands with at least 50 ad exposures. Panel B compares the empirical distribution of average ad positions on broadcast networks to quantiles of a Normal distribution (QQ-plot) with the same mean and variance. Panel C does the same QQ-plot, but for average ad positions on cable networks.

Figure 6: Distributions of Ad Creative Estimates



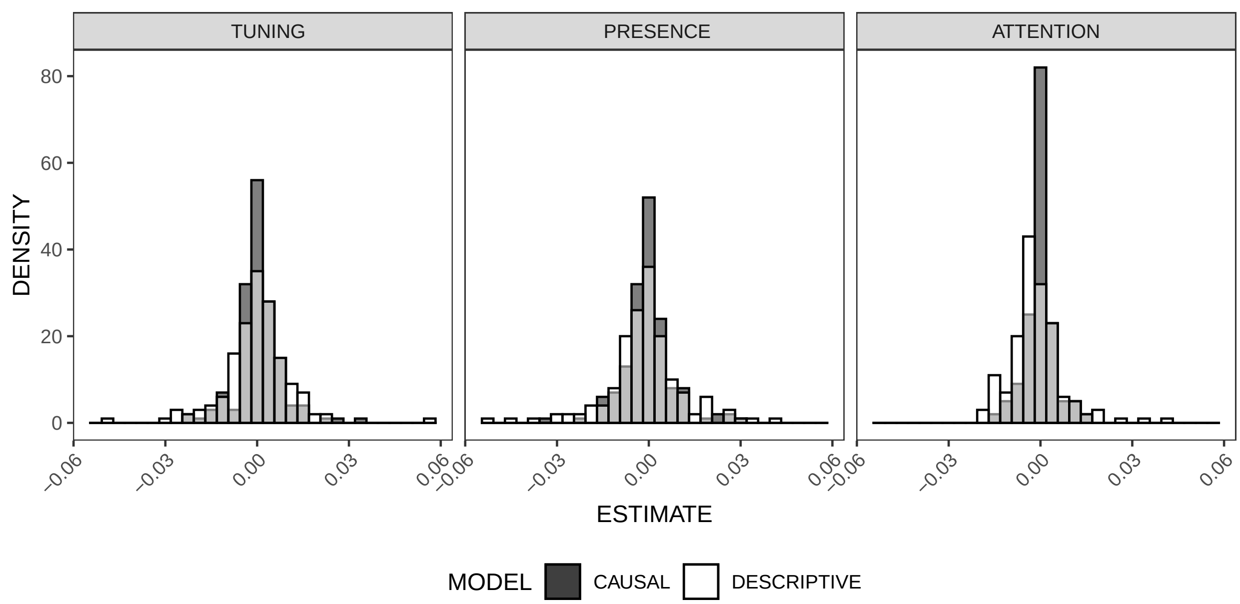
Notes: The three panels show distributions of 6,650 ad creative fixed effect estimates on tuning, presence, and attention from the descriptive and causal models. Distributions are demeaned to aide comparisons across models and outcomes.

Figure 7: Ad Creative Estimates by Sample Size



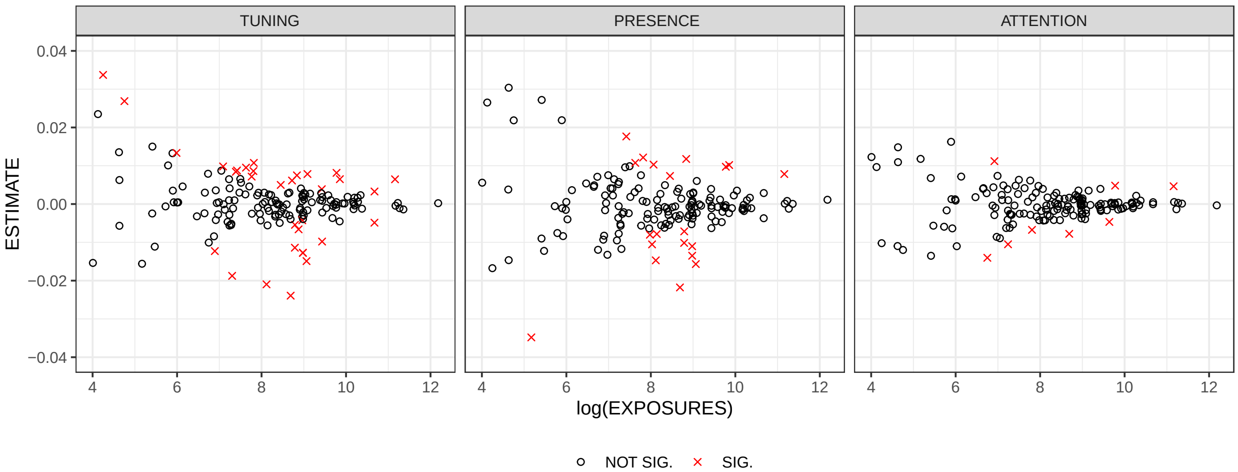
Notes: The three panels show scatterplots of causal ad creative estimates versus the log number of exposures for each ad creative. Distributions are demeaned to aide comparisons across models and outcomes. The blue trend-line shows linear fit.

Figure 8: Distributions of Ad Category Estimates



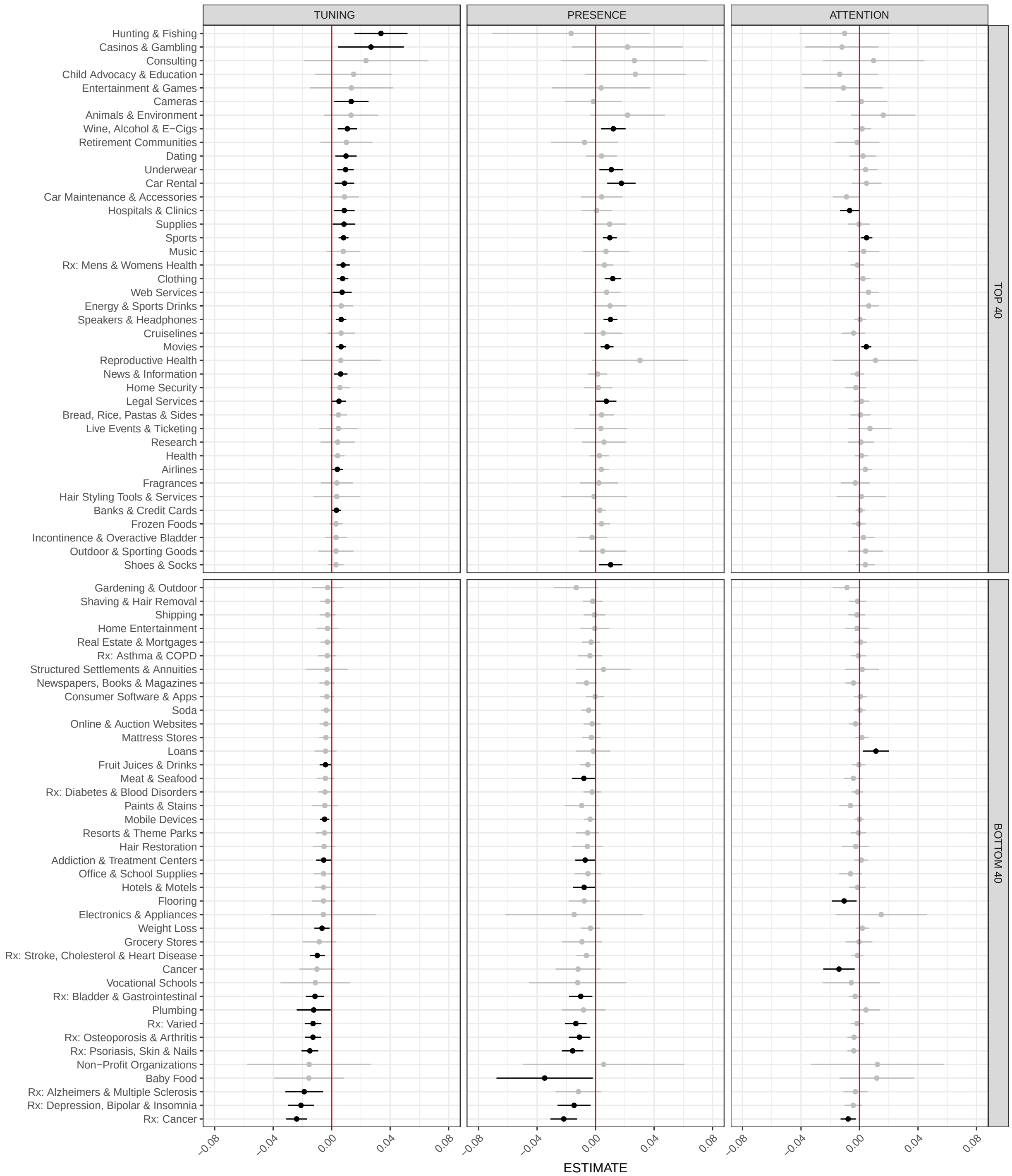
Notes: The three panels show distributions of 167 ad category estimates on tuning, presence, and attention from the descriptive and causal models. Distributions are demeaned to aide comparisons across models and outcomes.

Figure 9: Ad Category Estimates by Sample Size



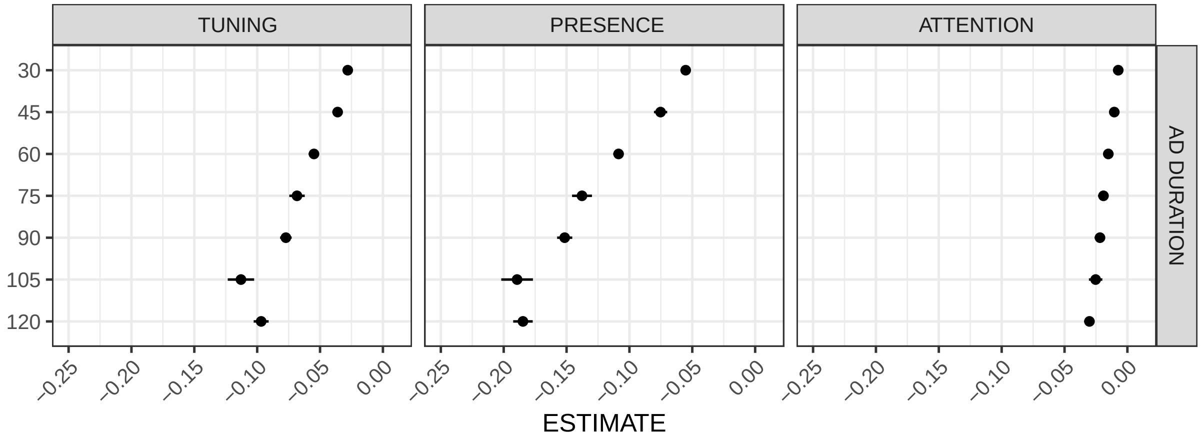
Notes: The three panels show scatterplots of causal ad category estimates versus the logged number of exposures. Distributions are demeaned to aide comparisons across models and outcomes. A red mark indicates that a estimate's 95% confidence interval does not include zero.

Figure 10: Top and Bottom 40 Ad Category Causal Effects



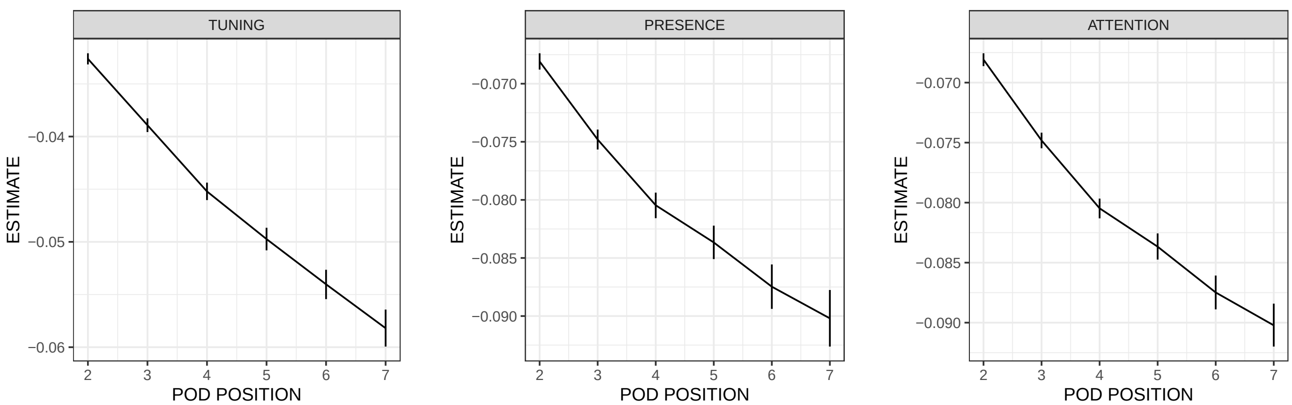
Notes: Each panel presents the top 40 and bottom 40 ad category causal effects, ranked by tuning point estimates. Whiskers represent 95% confidence intervals. Bolded point estimates and whiskers indicate 95% confidence intervals that do not include zero.

Figure 11: Ad Duration Effects



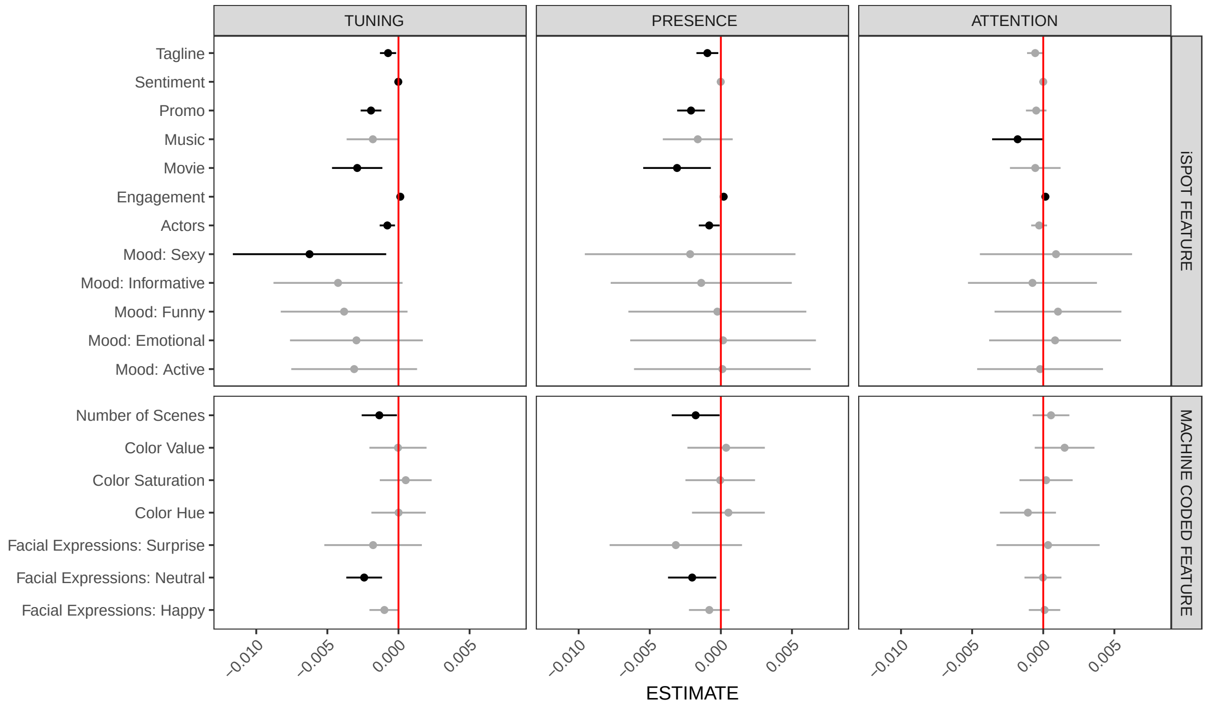
Notes: Ad durations are measured in seconds. Whiskers represent 95% confidence intervals. Every 95% confidence interval in the graph does not include zero.

Figure 12: Position and Time-Elapsed Effects in a Modal Break



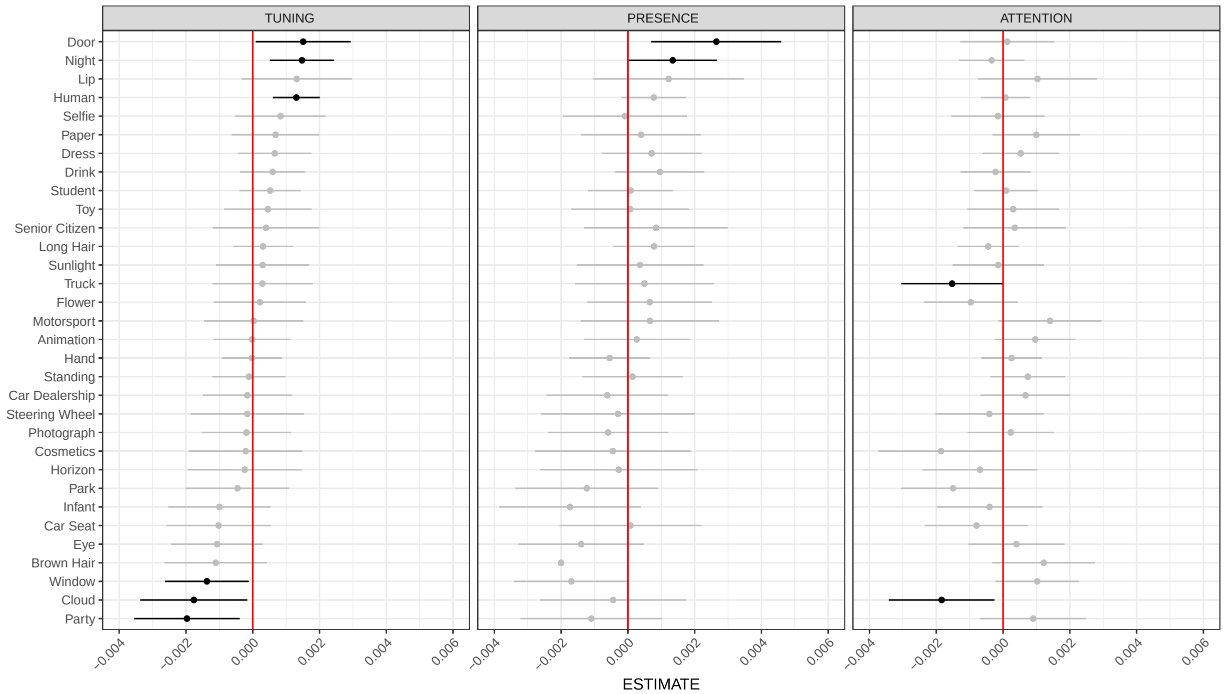
Notes: The figure combines pod position and time-elapsed effects for a hypothetical break consisting of seven 30-second ads. Whiskers represent 95% confidence intervals.

Figure 13: iSpot and Machine-coded Ad Feature Causal Effects



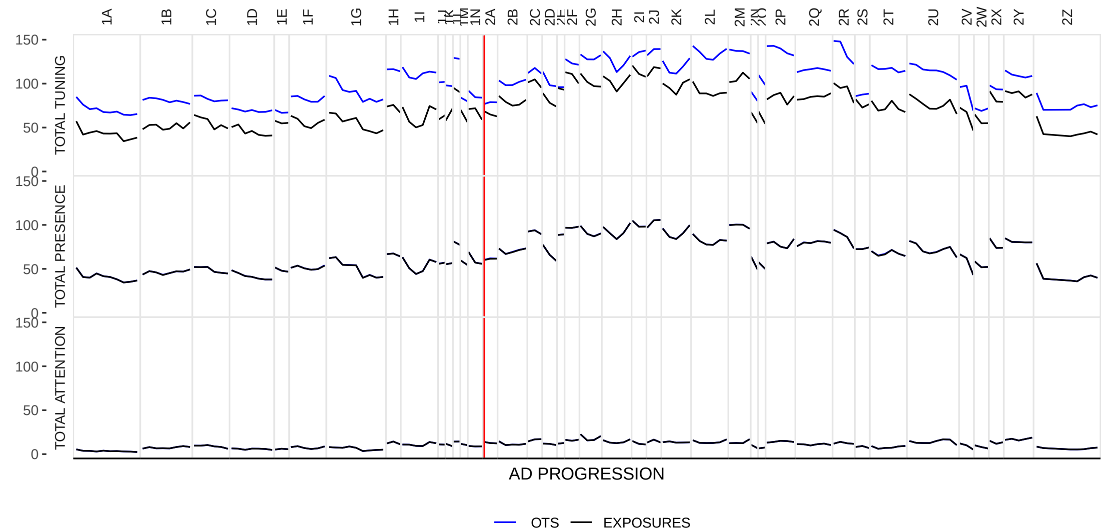
Notes: Whiskers represent 95% confidence intervals. Bolded point estimates and whiskers indicate 95% confidence intervals that do not include zero.

Figure 14: Google Cloud Vision Ad Feature Causal Effects



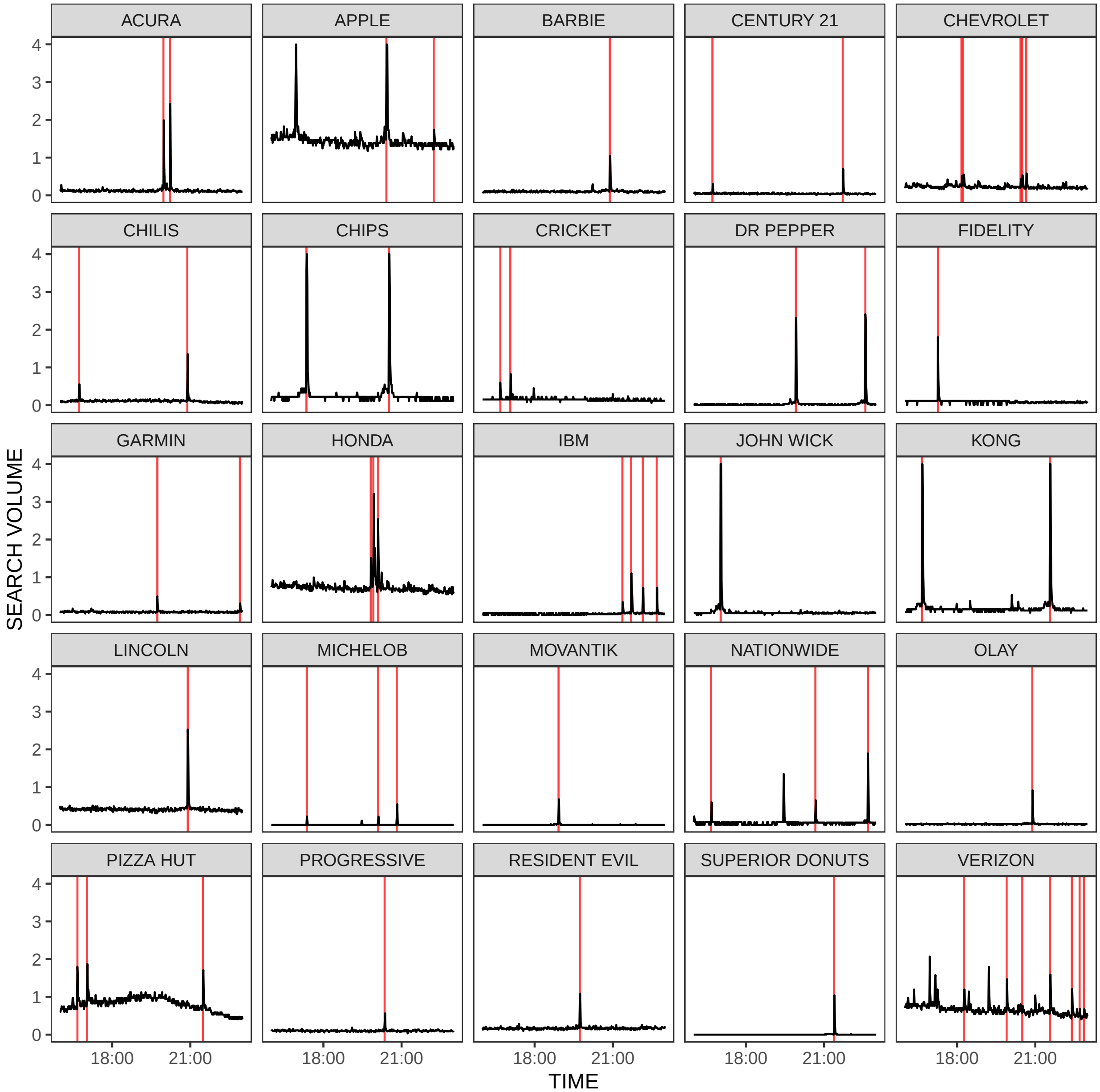
Notes: This figure presents Google Cloud Vision (GCV) feature estimates, ranked by tuning estimates. Features are tags created by GCV based on ad creative videos. Whiskers represent 95% confidence intervals. Bolded point estimates and whiskers indicate 95% confidence intervals that do not include zero.

Figure 15: Total Viewing Measures across two NFL Playoff Games in 2017



Notes: Panels show numbers of viewers tuned (in OTS and Exposures), present, and attentive across consecutive ad slots in ad breaks with at least two national ads. Ad breaks are separated by gray vertical bars. The red line marks when the first game ends.

Figure 16: Examples of Brand Search Lift



Notes: Panels show Google search volume for 25 brands from a variety of categories, in units normalized to the average per-minute search for the keyword "Pizza Hut" in a reference hour, truncated at 4. Red lines denote minutes brand ads began during the two NFL playoff games. The Y axis is truncated so as to highlight baseline variation. Appendix Figure A8 shows untruncated data for all brands.

Table 1: Counts of Significant Ad Category Estimates

Model	Tuning	Presence	Attention
Descriptive	68	45	33
Causal	32	20	8

Notes: Table entries count how many of the 167 ad category estimates are statistically significantly different from the average ad at the 95% confidence level. A 5% Type I error rate predicts 8.35 false positive results in expectation.

Table 2: Slot, Duration, and Time-Elapsed Parameter Estimates

Variable	Tuning	Presence	Attention
Pod Position == 2	0.00049	-0.00816 ***	-0.00352 ***
	(0.00036)	(0.00048)	(0.00040)
3	-0.00068	-0.01024 ***	-0.00509 ***
	(0.00056)	(0.00076)	(0.00067)
4	-0.00185 *	-0.01123 ***	-0.00603 ***
	(0.00075)	(0.00101)	(0.00090)
5	-0.00124	-0.00976 ***	-0.00602 ***
	(0.00093)	(0.00125)	(0.00110)
6	-0.00042	-0.00891 ***	-0.00643 ***
	(0.00108)	(0.00145)	(0.00129)
7	0.00053	-0.00700 ***	-0.00617 ***
	(0.00122)	(0.00164)	(0.00146)
8	0.00194	-0.00632 ***	-0.00658 ***
	(0.00135)	(0.00182)	(0.00162)
9	0.00128	-0.00675 ***	-0.00763 ***
	(0.00148)	(0.00199)	(0.00178)
10	0.00163	-0.00586 **	-0.00689 ***
	(0.00161)	(0.00217)	(0.00195)
11	0.00125	-0.00609 *	-0.00827 ***
	(0.00175)	(0.00236)	(0.00214)
12	0.00407 *	-0.00356	-0.00783 ***
	(0.00190)	(0.00258)	(0.00235)
13	0.00363	-0.00574 *	-0.00845 **
	(0.00208)	(0.00283)	(0.00261)
14	0.00223	-0.00916 **	-0.00974 ***
	(0.00231)	(0.00315)	(0.00291)
15	-0.00025	-0.00844 *	-0.00908 **
	(0.00262)	(0.00356)	(0.00336)
16	-0.00640 *	-0.01139 **	-0.01059 **
	(0.00313)	(0.00418)	(0.00383)
17	-0.01824 ***	-0.02730 ***	-0.00937 *
	(0.00406)	(0.00530)	(0.00457)
18	-0.04495 ***	-0.05422 ***	-0.01302 *
	(0.00622)	(0.00774)	(0.00622)
Ad Duration == 30	-0.02798 ***	-0.05526 ***	-0.00730 ***
	(0.00018)	(0.00025)	(0.00019)
45	-0.03603 ***	-0.07523 ***	-0.01044 ***
	(0.00197)	(0.00268)	(0.00163)
60	-0.05485 ***	-0.10864 ***	-0.01517 ***
	(0.00070)	(0.00094)	(0.00056)
75	-0.06833 ***	-0.13772 ***	-0.01904 ***
	(0.00313)	(0.00403)	(0.00198)
90	-0.07718 ***	-0.15149 ***	-0.02182 ***
	(0.00235)	(0.00307)	(0.00155)
105	-0.11291 ***	-0.18936 ***	-0.02527 ***
	(0.00535)	(0.00643)	(0.00272)
120	-0.09683 ***	-0.18470 ***	-0.03020 ***
	(0.00303)	(0.00396)	(0.00197)
Time Elapsed	-1.749e-04 ***	-1.624e-04 ***	-7.167e-06
	(1.216e-05)	(1.647e-05)	(1.424e-05)
Time Elapsed^2	1.330e-07 ***	2.339e-07 ***	1.601e-08
	(3.267e-08)	(4.625e-08)	(4.339e-08)
R^2	0.366	0.571	0.711
N	4,257,112	4,257,112	4,257,112

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The table presents pod-position, ad duration, and time-elapsed estimates for the causal model. SE in parentheses.

Table 3: Descriptive Results of Pharmaceutical Category Causal Ad Effects on Drug Category Prevalence and Severity

	DV = Tuning Rx Effects		
	(1)	(2)	(3)
Constant	-0.00454 ***	-0.00731 ***	-0.00389 ***
	(0.00001)	(0.00001)	(0.00001)
Severity	-0.00071 ***		-0.00072 ***
	(0.00000)		(0.00000)
Prevalence		-0.00037 ***	-0.00459 ***
		(0.00005)	(0.00003)

	DV = Presence Rx Effects		
	(1)	(2)	(3)
Constant	-0.00177 ***	-0.00333 ***	-0.00098 ***
	(0.00001)	(0.00001)	(0.00001)
Severity	-0.00046 ***		-0.00048 ***
	(0.00000)		(0.00000)
Prevalence		-0.00262 ***	-0.00564 ***
		(0.00005)	(0.00004)

	DV = Attention Rx Effects		
	(1)	(2)	(3)
Constant	-0.00075***	-0.00127 ***	-0.0001 ***
	(0.00000)	(0.00000)	(0.00000)
Severity	-0.00012		-0.0027 ***
	(0.00000)		(0.00001)
Prevalence		0.00018 ***	-0.0037 ***
		(0.00001)	(0.00003)

Notes: This table reports the results of the regression specification in Equation 2 where dependent variable is each of the three viewing measures. Columns (1) and (2) report results for severity and prevalence separately, while Column (3) includes both. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 4: Search Lift Parameter Estimates

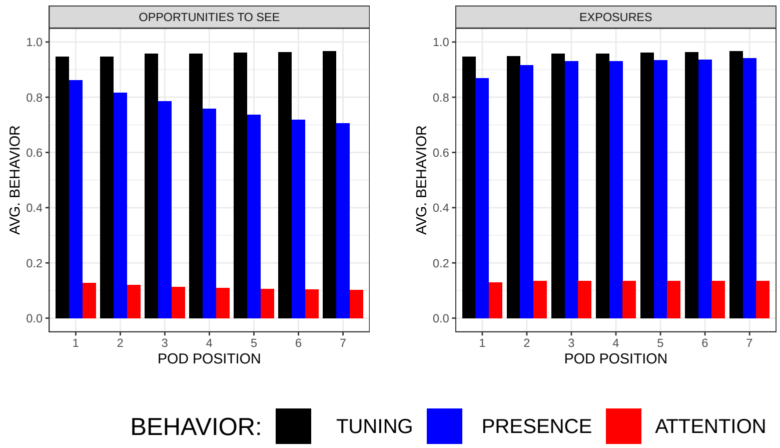
	Dependent Variable: Search Lift							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OTS Tuning	0.006				0.006			
	(0.005)				(0.01)			
Exposure Tuning		0.015*				0.024		
		(0.006)				(0.013)		
Presence			0.013 *				0.017	
			(0.006)				(0.014)	
Attention				0.064 *				0.106 *
				(0.029)				(0.052)
First Ad FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
1/4 Hour FE	No	No	No	No	Yes	Yes	Yes	Yes
R^2	0.035	0.082	0.07	0.116	0.158	0.176	0.165	0.180
N	180	180	180	180	180	180	180	180

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Specifications (1-4) do not include quarter-hour fixed effects. (5-8) do include quarter-hour fixed effects. SE in parentheses.

Appendix

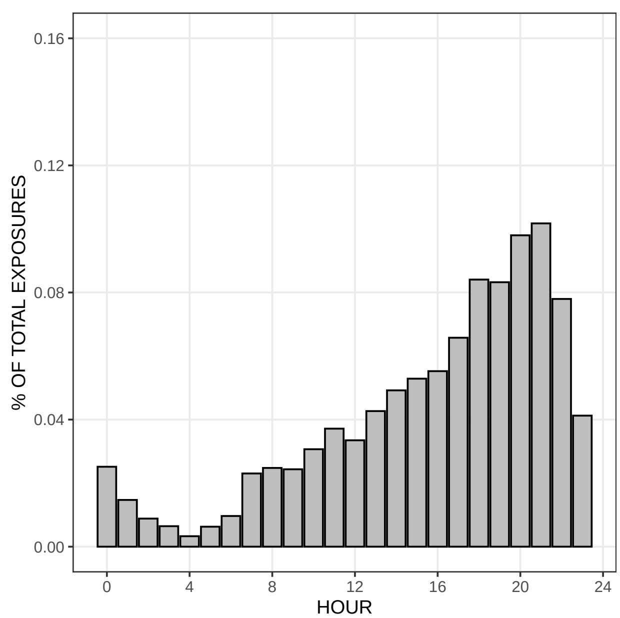
Figures

Figure A1: Average Viewing Behaviors in all 7-Slot Breaks



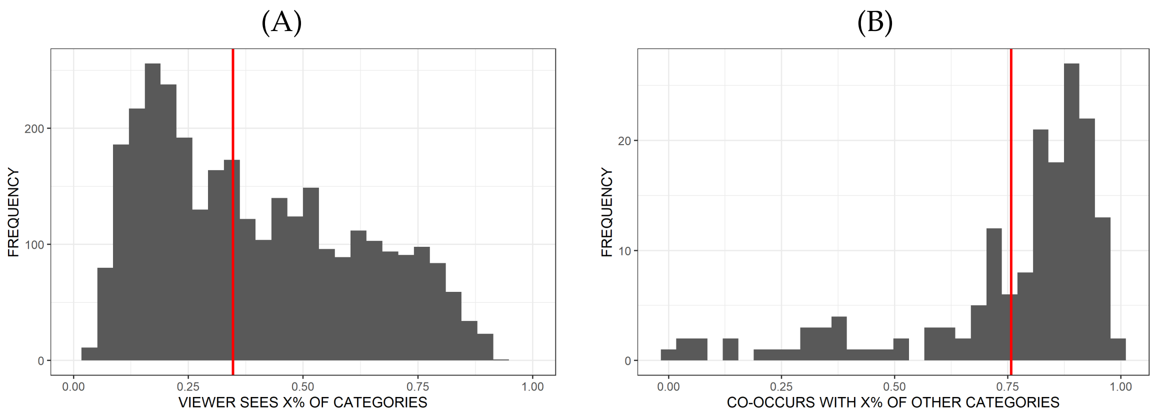
Notes: The figure focuses on audience retention by excluding viewers who join breaks mid-stream. Changes across slots would be smaller in all three panels if they incorporated viewers who join commercial breaks after the first slot.

Figure A2: Ad Exposures by Hour of the Day



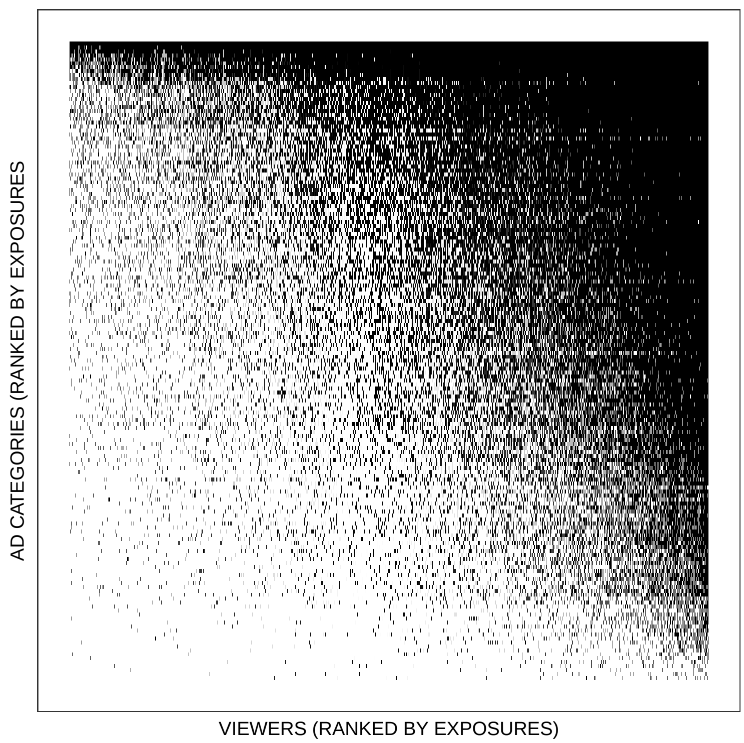
Notes: This figure plots the histogram of the number of ad exposures by hour of the day. Hourly attention and hourly exposures correlate at .99.

Figure A3: Ad-Category and Viewer-Category Co-occurrences



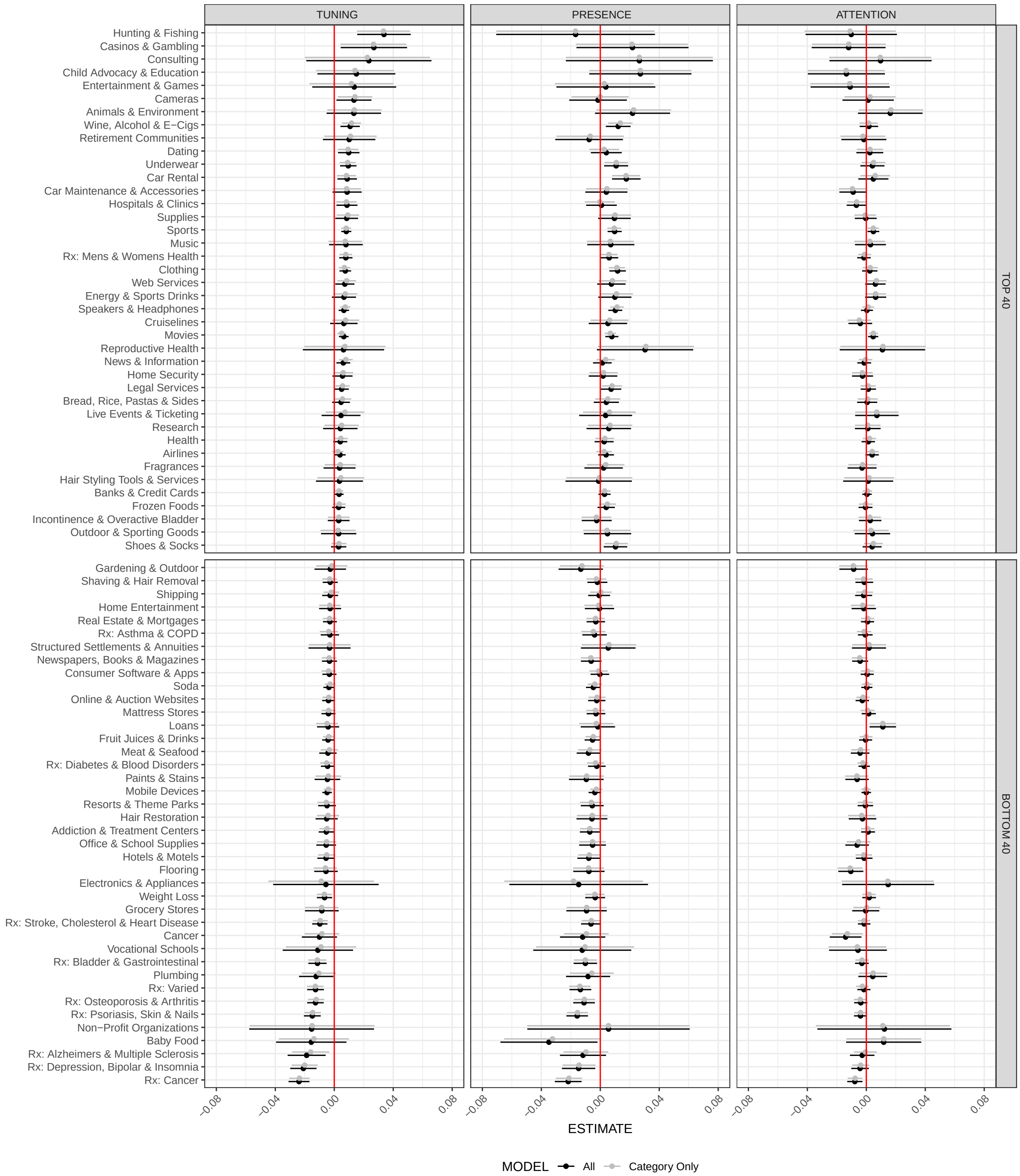
Notes: Panel (A) summarizes the co-occurrence of ad categories. A co-occurrence value of 0.75 implies that the ad category co-occurs in breaks with 75% of other ad categories. The red vertical line denotes the average across all ad categories. Panel (B) summarizes the average number of distinct ad categories seen by viewers. The red vertical line denotes the median number of categories seen across all viewers.

Figure A4: Viewer and Ad Category Coverage



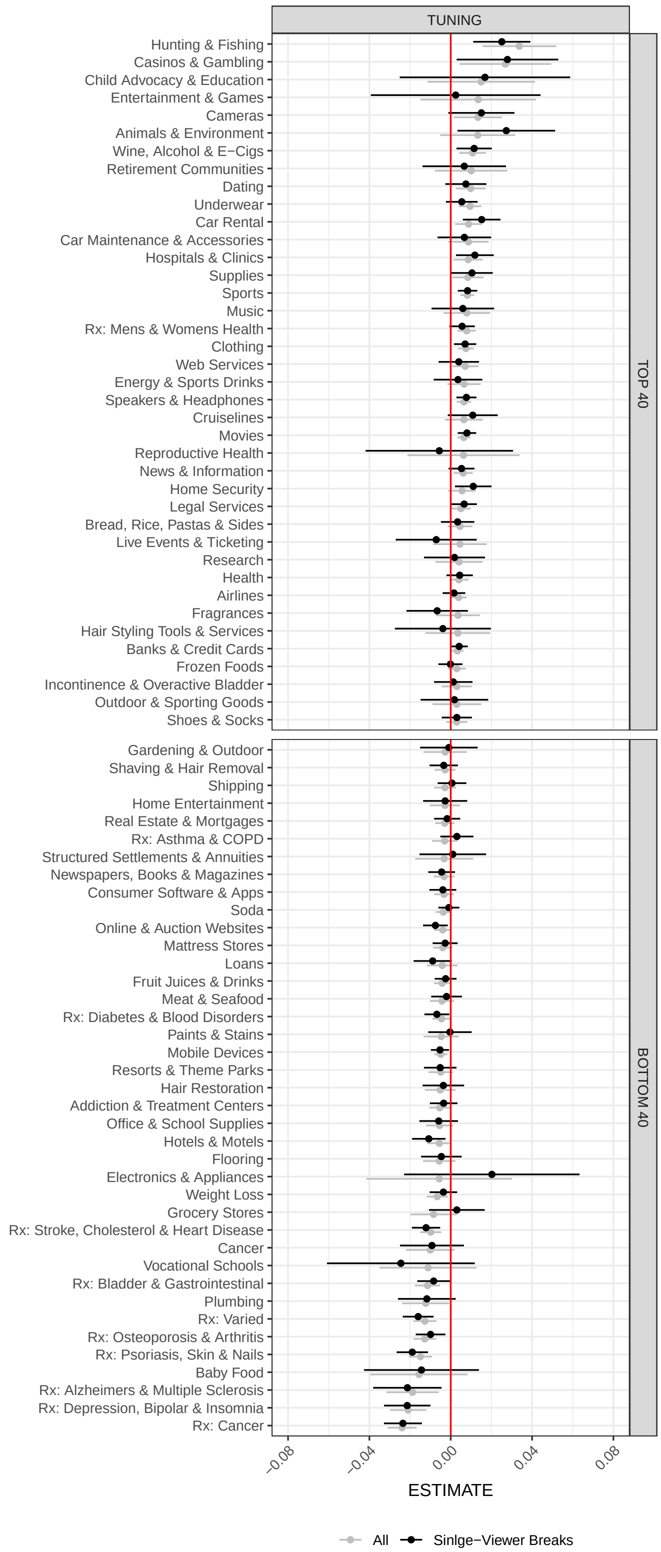
Notes: This figure shows the coverage of ad categories across viewers. The x-axis denotes individual viewers, ranked by the number of ad exposures; the y-axis expresses ad categories, ranked by number of ad exposures. Each cell in this matrix denotes whether a viewer saw a particular ad category – black indicating "yes", white, "no". Intuitively, coverage improves as we move right (viewers who have seen more ads) and up (ad categories that advertise more).

Figure A5: Category Estimates with and without Ad Content Features



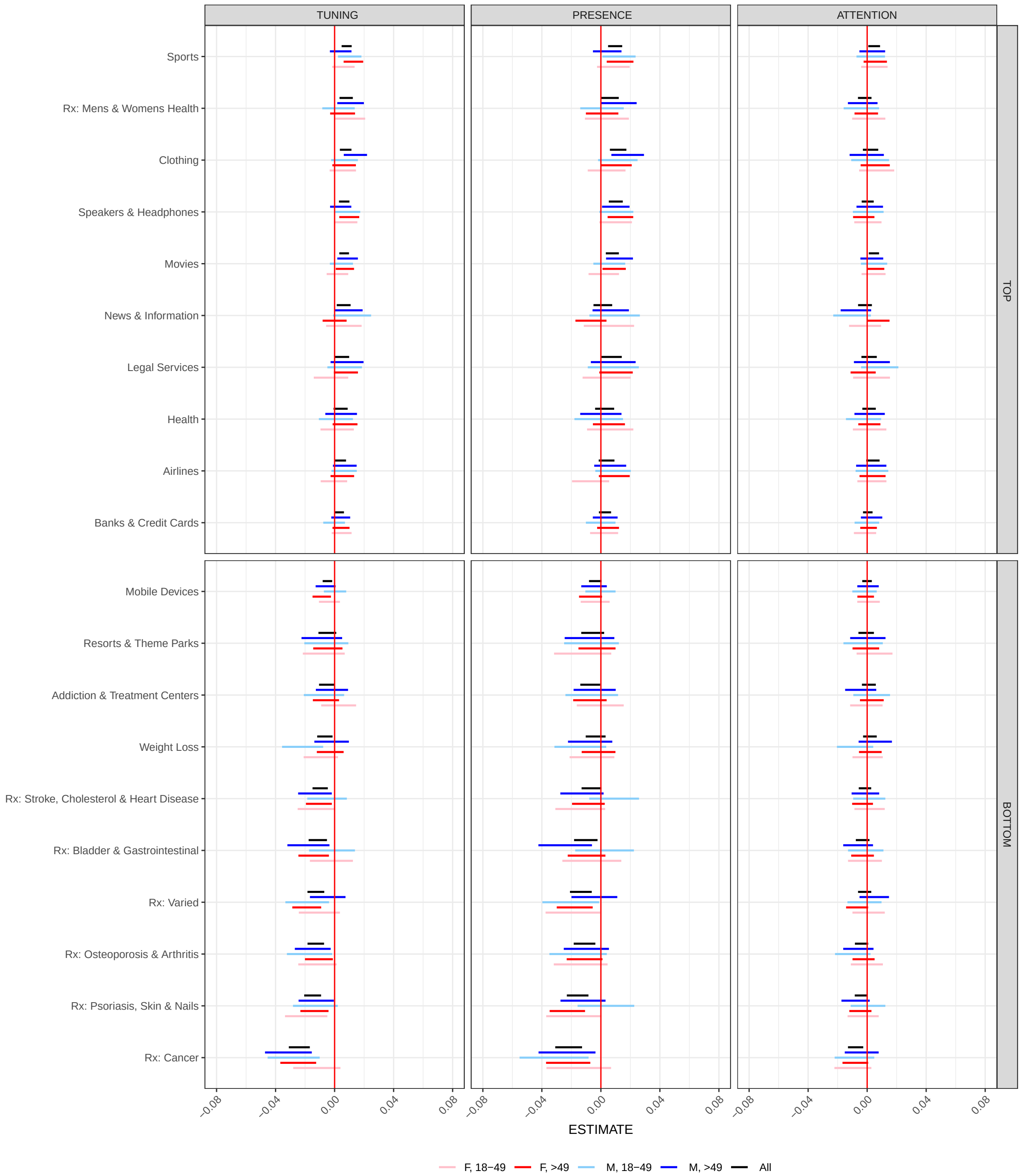
Notes: Black indicates category 95% confidence intervals from the full causal model. Gray indicates 95% confidence intervals from a model that excludes ad content feature data (iSpot features, machine-coded features and GCV features).

Figure A6: Category Estimates on Tuning with and without Multiple-viewer Sessions



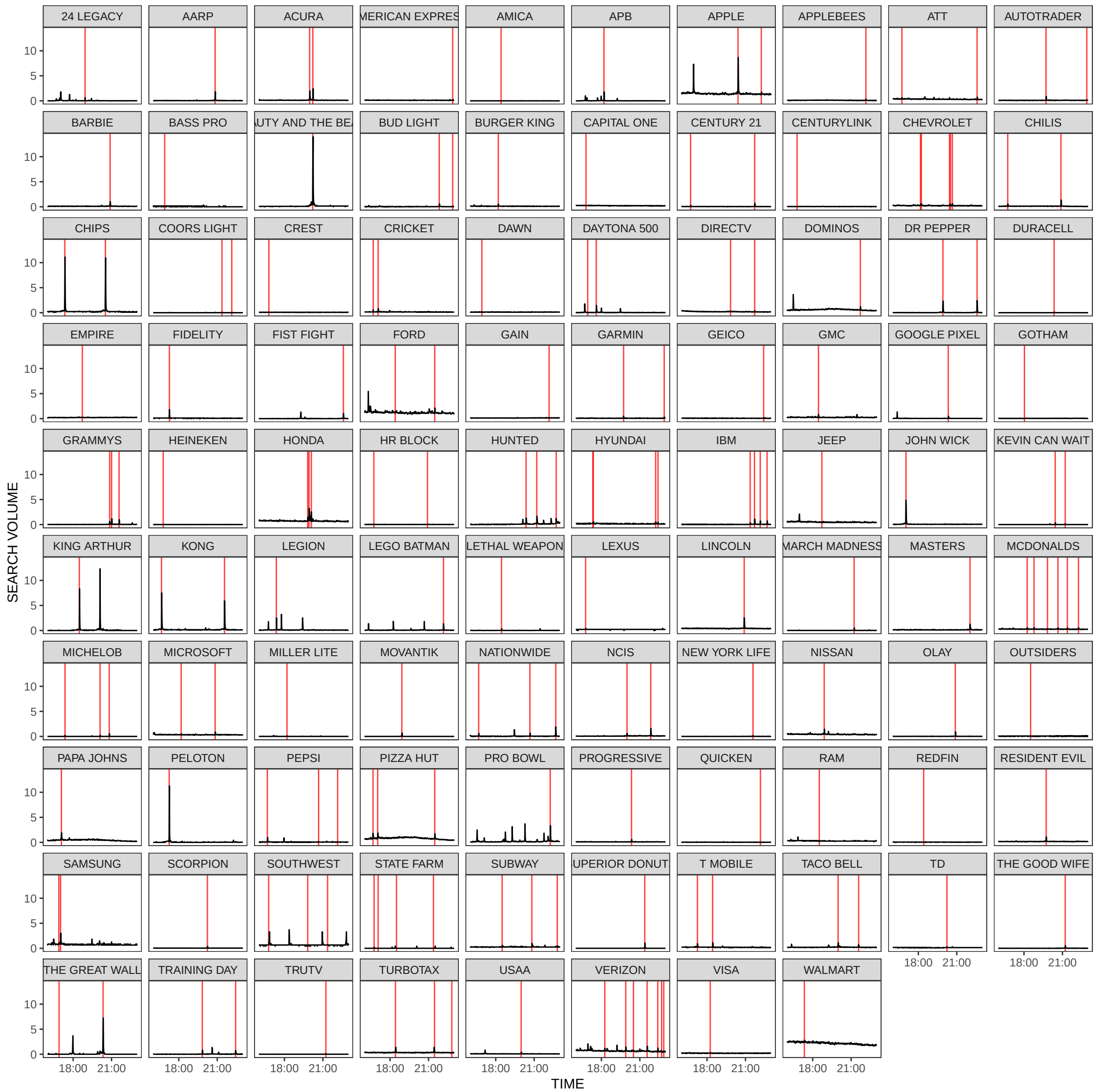
Notes: Each panel presents the top 40 and bottom 40 ad category causal effects separately for all ad breaks and single-viewer ad breaks.

Figure A7: Category Estimates overall and within Demographic Partitions



Notes: Each panel presents top 10 and bottom 10 ad category causal effects on tuning partitioned by demographic groups. Estimates for viewers under 18 years old confidence intervals also overlap, but they are substantially wider due to lower statistical power, so we exclude them for compactness.

Figure A8: Search Lift for All Brands



Notes: Panels show Google search volume for all 180 national ads, in units normalized to the average per-minute search for the keyword “Pizza Hut.” Red lines denote minutes that brand ads began during the two NFL playoff games.

Tables

Table A1: Summary of iSpot and Machine-Coded Ad Content Features

Ad Content	Mean	SD
iSpot Features		
Actors	0.483	0.500
Engagement	4.047	2.986
Promo	0.220	0.415
Tagline	0.583	0.493
Movie	0.032	0.177
Music	0.018	0.132
Sentiment	0.422	0.398
Mood == Active	0.583	0.493
Mood == Emo	0.024	0.154
Mood == Funny	0.203	0.402
Mood == Info.	0.046	0.211
Mood == Sexy	0.008	0.089
ML Coded Features		
Colorfulness	0.280	0.169
Saturation	0.420	0.235
Luminosity	0.357	0.188
Happy	0.187	0.236
Surprise	0.024	0.063
Neutral	0.182	0.197
Scene Frequency	0.425	0.271

Table A2: Google Cloud Vision Feature Tag Frequencies

GCV Feature	% Exposures	GCV Feature	% Exposures	GCV Feature	% Exposures
Logo *	0.143	Mid Size Car	0.025	Downtown	0.014
Graphics	0.135	Transport	0.025	Summer	0.014
Text *	0.116	Ball Game	0.025	Residential Area	0.014
Conversation	0.106	Stage	0.024	**Cloud** *	0.014
People	0.104	Performance Art	0.024	Forest *	0.014
Smile *	0.101	Games	0.024	Town	0.014
Interaction	0.101	Pet *	0.024	Humour	0.014
Human *	0.090	Plant *	0.024	Television *	0.013
Vehicle	0.087	Bottle *	0.024	Formal Wear *	0.013
Facial Expression	0.087	Compact Car	0.024	**Steering Wheel** *	0.013
Song	0.087	Online Advertising	0.024	**Flower** *	0.013
Individual Sports	0.086	Telephone	0.024	**Truck** *	0.013
Sitting	0.084	**Standing** *	0.024	Classroom *	0.013
Car *	0.075	Windshield *	0.024	Junk Food	0.013
Fun	0.073	Lighting	0.024	Mountain *	0.013
Social Group	0.070	Symbol	0.023	Neighbourhood	0.013
Motor Vehicle	0.068	Professional	0.023	**Motorsport** *	0.013
Graphic Design	0.067	House	0.023	Structure	0.013
Food *	0.062	Website *	0.023	Architecture	0.012
Presentation	0.059	Street	0.023	Drinking	0.012
Brand	0.057	**Dress** *	0.023	Snack	0.012
Land Vehicle	0.056	Sport Utility Vehicle *	0.023	Romance	0.012
Emotion	0.055	Team Sport	0.023	**Senior Citizen** *	0.012
Long Hair *	0.054	Urban Area *	0.023	Soccer *	0.012
Mode of Transport	0.053	Dish *	0.022	**Lip** *	0.012
Nature	0.049	Meal	0.022	Music Venue	0.012
Animal *	0.049	Light	0.021	Retail	0.012
Sports	0.045	Cuisine	0.021	Foot *	0.012
Advertising	0.044	Media	0.021	Label	0.011
Driving *	0.044	Pedestrian *	0.021	Glass	0.011
Crowd *	0.044	Automotive Exterior	0.021	Education	0.011
Community	0.042	Furniture *	0.021	Machine	0.011
Tree *	0.041	Dog *	0.021	**Park** *	0.011
Performing Arts	0.040	**Animation** *	0.020	Property	0.011
Performance	0.040	Product	0.020	Metropolis	0.011
Hand *	0.039	Backyard	0.020	**Infant** *	0.011
Happiness	0.038	Lawn *	0.020	Ball *	0.011
Night *	0.038	**Car Dealership** *	0.020	**Horizon** *	0.011
Mobile Device	0.037	**Window** *	0.020	Choreography	0.011
Speech	0.037	Physical Fitness	0.020	Sunglasses	0.011
Font	0.037	Physical Exercise	0.019	**Cosmetics** *	0.011
Singing *	0.037	Portable Communications Device	0.019	**Party** *	0.011
Gadget	0.036	Sport Venue *	0.019	Sandwich *	0.011
Technology	0.035	Television Advertisement	0.019	Hair	0.011
Visual Effects	0.035	Electronics	0.019	Finger	0.011
Mobile Phone	0.035	Learning	0.019	Orator	0.011
Smartphone	0.035	**Photograph** *	0.018	Personal Computer	0.011
Black and White *	0.034	Luxury Vehicle	0.018	Off Road Vehicle	0.011
Sky *	0.034	Leisure	0.018	Concert	0.011
Play *	0.033	City	0.017	Woodland	0.011
Display Device	0.033	Ceremony	0.017	Musical Instrument	0.011
Television Program	0.033	Outdoor Recreation	0.017	Metropolitan Area	0.011
Glasses *	0.032	**Selfie** *	0.017	Terrain	0.011
Dance *	0.032	**Sunlight** *	0.017	Action Game	0.011
Student *	0.032	**Toy** *	0.017	Living Room	0.011
Audience *	0.031	Document *	0.017	Company	0.010
Fashion	0.031	Writing *	0.016	Sports Car	0.010
Grass *	0.031	Wilderness *	0.016	Body of Water	0.010
Automotive Design	0.030	**Eye** *	0.016	Footwear	0.010
Public Space *	0.030	Emblem	0.016	Signage	0.010
Home *	0.030	Yard	0.016	Chair	0.010
Electronic Device	0.030	Special Effects	0.016	Flora	0.010
Liquid *	0.029	Sedan	0.016	Facial Hair	0.010
Drink *	0.028	Player	0.016	Banner	0.010
Eyewear	0.028	Communication	0.015	Rural Area	0.010
Black *	0.028	Suit *	0.015	Kitchen	0.010
Public Speaking *	0.028	Plastic Bottle	0.015	Sign	0.010
Film	0.027	**Paper** *	0.015	Monochrome	0.010
Consumer Electronics	0.027	Landscape *	0.015	Glass Bottle	0.010
Road *	0.027	Compact Sport Utility Vehicle	0.015	Sea	0.010
Recreation	0.026	**Car Seat** *	0.015	Cheering	0.010
Cooking *	0.026	**Brown Hair** *	0.015	Uniform	0.010
Communication Device	0.026	**Door** *	0.015	Shoe	0.010
Eating *	0.026	Web Page	0.014	Nightclub	0.010
Building *	0.026	Fast Food	0.014	Hill	0.010

Notes: GCV features are ranked by the % of exposures in which they occur. Concrete nouns and verbs are indicated by an asterisk. Bolded features correspond to the 32 tags for which 50% or more of variance remained unexplained in a regression of the tag on product category, iSpot and machine-coded ad features.

Table A3: Variance Decompositions of Viewing Behaviors on Viewer, Break, Ad Features

Variable	df	Tuning R^2	Tuning F	Presence R^2	Presence F	Attention R^2	Attention F
Viewer Characteristics							
Viewer ID	3659	0.00848	9.9	0.04635	**56.5**	0.15125	**207.2**
Viewer Age and Gender	6	0.00012	**105.8**	0.00387	**3,306.1**	0.00287	**2,446.8**
Ad Environment							
Pod Position	18	0.00153	**384.5**	0.00228	**572.4**	0.00004	9.0
Channel	4	0.00007	**95.4**	0.00010	**136.0**	0.00015	**211.1**
Program Genre	104	0.00539	**224.0**	0.00156	**64.6**	0.00201	**83.1**
Ad Characteristics							
Ad Len	8	0.01550	**9,574.5**	0.04658	**29,710.6**	0.00080	**486.6**
Ad Industry	167	0.00481	**124.0**	0.01068	**276.9**	0.00051	13.1
Ad Title	6650	0.01381	9.0	0.02076	13.6	0.00332	2.1
iSpot Features							
Actors	2	0.00008	**150.2**	0.00032	**565.9**	0.00001	9.0
Eng	2	0.00001	11.5	0.00001	14.1	0.00005	**88.8**
Promo	2	0.00015	**257.9**	0.00010	**184.6**	0.00001	20.2
Tagline	2	0.00022	**389.3**	0.00065	**1,162.0**	0.00001	9.5
Movie	2	0.00000	3.4	0.00001	14.4	0.00002	30.6
Music	2	0.00000	6.9	0.00001	14.1	0.00000	2.9
Actors	2	0.00008	**150.2**	0.00032	**565.9**	0.00001	9.0
Sentiment	2	0.00028	**505.5**	0.00053	**947.7**	0.00000	4.1
Mood	6	0.00092	**328.2**	0.00251	**893.3**	0.00018	**64.8**
All	13	0.00191	**283.7**	0.00464	**690.4**	0.00023	34.3
ML Ad Coded Features							
Colorfulness	2	0.00007	**115.4**	0.00032	**502.3**	0.00000	1.4
Saturation	2	0.00001	9.8	0.00007	**104.9**	0.00005	**70.8**
Value	2	0.00007	**103.9**	0.00017	**269.6**	0.00001	12.6
Emotion	4	0.00003	13.2	0.00011	**56.6**	0.00001	2.7
Scenes	2	0.00042	**655.9**	0.00127	**1,961.4**	0.00009	**142.7**
All	9	0.00071	**136.8**	0.00238	**461.8**	0.00012	23.1
Google Cloud Vision							
GCV Selected Features	33	0.00069	38.1	0.00182	**101.3**	0.00010	5.6

Notes: Each entry reports a separate variance decomposition of an ad viewing behavior on the set of viewer, break or ad features described in the row header. Bold indicates statistical significance at 95% confidence.

Table A4: Prescription Drug Category Characteristics

Category Name	Prevalence (%)	Severity (MM DALY/year)
Allergies	18.5	0.0
Alzheimer's & Multiple Sclerosis	1.4	1.9
Asthma & COPD	8.0	5.4
Bladder & Gastrointestinal	25.9	0.4
Cancer	3.7	14.2
Depression, Bipolar & Insomnia	16.4	7.2
Diabetes & Blood Disorders	6.3	1.3
Mens & Women's Health	10.0	0.0
Osteoporosis & Arthritis	13.4	1.7
Psoriasis, Skin & Nails	28.1	2.4
Stroke, Cholesterol & Heart Disease	10.9	15.0
Varied	0.5	0.3

Table A5: Slot, Duration and Time-Elapsed Parameter Estimates for Various Specifications

Variable		Tuning			Presence			Attention	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Pod Position == 2	-0.00216 ***	-0.00012	0.00049	-0.00202 ***	-0.00105	-0.00816 ***	-0.00081	-0.00188 ***	-0.00352 ***
	(0.00029)	(0.00036)	(0.00036)	(0.00046)	(0.00055)	(0.00048)	(0.00045)	(0.00042)	(0.0004)
3	0.00557 ***	0.00695 ***	-0.00068	0.00532 ***	0.00498 ***	-0.01024 ***	-0.00056	-0.00204 **	-0.00509 ***
	(0.00038)	(0.0006)	(0.00056)	(0.00061)	(0.00086)	(0.00076)	(0.0006)	(0.00069)	(0.00067)
4	0.00974 ***	0.01197 ***	-0.00185 *	0.00941 ***	0.00919 ***	-0.01123 ***	-0.00086	-0.00231 *	-0.00603 ***
	(0.00047)	(0.00081)	(0.00075)	(0.00076)	(0.00114)	(0.00101)	(0.00074)	(0.00096)	(0.0009)
5	0.01295 ***	0.0168 ***	-0.00124	0.01294 ***	0.01368 ***	-0.00976 ***	-0.00136	-0.00199	-0.00602 ***
	(0.00055)	(0.001)	(0.00093)	(0.0009)	(0.0014)	(0.00125)	(0.00088)	(0.00117)	(0.0011)
6	0.01466 ***	0.02115 ***	-0.00042	0.0141 ***	0.01695 ***	-0.00891 ***	-0.00253 *	-0.0022	-0.00643 ***
	(0.00062)	(0.00117)	(0.00108)	(0.00102)	(0.00163)	(0.00145)	(0.00099)	(0.00135)	(0.00129)
7	0.01604 ***	0.0252 ***	0.00053	0.01665 ***	0.02133 ***	-0.007 ***	-0.0031 **	-0.00176	-0.00617 ***
	(0.00069)	(0.00129)	(0.00122)	(0.00112)	(0.00184)	(0.00164)	(0.0011)	(0.00155)	(0.00146)
8	0.01765 ***	0.02942 ***	0.00194	0.01728 ***	0.02373 ***	-0.00632 ***	-0.00398 ***	-0.00223	-0.00658 ***
	(0.00075)	(0.00147)	(0.00135)	(0.00123)	(0.00203)	(0.00182)	(0.0012)	(0.00175)	(0.00162)
9	0.01738 ***	0.03119 ***	0.00128	0.01641 ***	0.02471 ***	-0.00675 ***	-0.00665 ***	-0.0035	-0.00763 ***
	(0.00082)	(0.00157)	(0.00148)	(0.00133)	(0.00217)	(0.00199)	(0.00131)	(0.00192)	(0.00178)
10	0.01784 ***	0.03352 ***	0.00163	0.01748 ***	0.02703 ***	-0.00586 **	-0.00732 ***	-0.00303	-0.00689 ***
	(0.00089)	(0.00171)	(0.00161)	(0.00145)	(0.00236)	(0.00217)	(0.00142)	(0.00211)	(0.00195)
11	0.01672 ***	0.03406 ***	0.00125	0.01602 ***	0.02744 ***	-0.00609 *	-0.0105 ***	-0.00434	-0.00827 ***
	(0.00098)	(0.00182)	(0.00175)	(0.0016)	(0.00254)	(0.00236)	(0.00156)	(0.0023)	(0.00214)
12	0.01892 ***	0.03801 ***	0.00407 *	0.01798 ***	0.03074 ***	-0.00356	-0.01252 ***	-0.00422	-0.00783 ***
	(0.0011)	(0.00195)	(0.0019)	(0.00179)	(0.00275)	(0.00258)	(0.00175)	(0.00249)	(0.00235)
13	0.01795 ***	0.03792 ***	0.00363	0.01579 ***	0.02922 ***	-0.00574 *	-0.01471 ***	-0.00495	-0.00845 **
	(0.00125)	(0.00212)	(0.00208)	(0.00204)	(0.00298)	(0.00283)	(0.002)	(0.00275)	(0.00261)
14	0.01672 ***	0.03746 ***	0.00223	0.01205 ***	0.0253 ***	-0.00916 **	-0.01776 ***	-0.00694 *	-0.00974 ***
	(0.00146)	(0.00236)	(0.00231)	(0.00238)	(0.00341)	(0.00315)	(0.00233)	(0.00313)	(0.00291)
15	0.01538 ***	0.03548 ***	-0.00025	0.01383 ***	0.02705 ***	-0.00844 *	-0.01799 ***	-0.00658	-0.00908 **
	(0.00176)	(0.00252)	(0.00262)	(0.00286)	(0.00356)	(0.00356)	(0.0028)	(0.00344)	(0.00336)
16	0.00895 ***	0.03065 ***	-0.0064 *	0.00863 *	0.02439 ***	-0.01139 **	-0.02112 ***	-0.00809	-0.01059 **
	(0.00216)	(0.00305)	(0.00313)	(0.00351)	(0.0043)	(0.00418)	(0.00343)	(0.00415)	(0.00383)
17	-0.00152	0.02003 ***	-0.01824 ***	-0.00378	0.01071 *	-0.0273 ***	-0.02239 ***	-0.00669	-0.00937 *
	(0.00283)	(0.0038)	(0.00406)	(0.00461)	(0.00524)	(0.0053)	(0.00451)	(0.00483)	(0.00457)
18	-0.02523 ***	-0.00303	-0.04495 ***	-0.03018 ***	-0.01395	-0.05422 ***	-0.02564 ***	-0.01299	-0.01302 *
	(0.00421)	(0.00606)	(0.00622)	(0.00686)	(0.00751)	(0.00774)	(0.00671)	(0.00678)	(0.00622)
Ad Duration == 30	-0.0291 ***	-0.03003 ***	-0.02798 ***	-0.08444 ***	-0.07883 ***	-0.05526 ***	-0.01008 ***	-0.0101 ***	-0.0073 ***
	(0.00015)	(0.00039)	(0.00018)	(0.00024)	(0.00115)	(0.00025)	(0.00024)	(0.00035)	(0.00019)
45	-0.02962 ***	-0.03727 ***	-0.03603 ***	-0.11095 ***	-0.10974 ***	-0.07523 ***	-0.01513 ***	-0.01577 ***	-0.01044 ***
	(0.00147)	(0.00198)	(0.00197)	(0.00239)	(0.00345)	(0.00268)	(0.00234)	(0.00178)	(0.00163)
60	-0.05361 ***	-0.05757 ***	-0.05485 ***	-0.16616 ***	-0.1571 ***	-0.10864 ***	-0.0223 ***	-0.02118 ***	-0.01517 ***
	(0.00047)	(0.00095)	(0.0007)	(0.00076)	(0.00247)	(0.00094)	(0.00075)	(0.00085)	(0.00056)
75	-0.06089 ***	-0.07033 ***	-0.06833 ***	-0.21306 ***	-0.2034 ***	-0.13772 ***	-0.03014 ***	-0.02773 ***	-0.01904 ***
	(0.00188)	(0.00333)	(0.00313)	(0.00305)	(0.00516)	(0.00403)	(0.00299)	(0.00206)	(0.00198)
90	-0.0742 ***	-0.08126 ***	-0.07718 ***	-0.23377 ***	-0.22145 ***	-0.15149 ***	-0.02935 ***	-0.02836 ***	-0.02182 ***
	(0.00139)	(0.00249)	(0.00235)	(0.00226)	(0.00447)	(0.00307)	(0.00221)	(0.00189)	(0.00155)
105	-0.11626 ***	-0.11315 ***	-0.11291 ***	-0.2893 ***	-0.26518 ***	-0.18936 ***	-0.04638 ***	-0.0339 ***	-0.02527 ***
	(0.00271)	(0.0052)	(0.00535)	(0.00442)	(0.00785)	(0.00643)	(0.00432)	(0.00312)	(0.00272)
120	-0.0813 ***	-0.09591 ***	-0.09683 ***	-0.27804 ***	-0.26746 ***	-0.1847 ***	-0.04053 ***	-0.03885 ***	-0.0302 ***
	(0.00164)	(0.00316)	(0.00303)	(0.00268)	(0.00631)	(0.00396)	(0.00262)	(0.0024)	(0.00197)
Time Elapsed	-9.16E-05 ***	-1.81E-04 ***	-1.75E-04 ***	-6.62E-05 ***	-1.44E-04 ***	-1.62E-04 ***	-2.18E-05 *	-1.68E-06	-7.17E-06
	(6.81E-06)	(1.27E-05)	(1.22E-05)	(1.11E-05)	(1.83E-05)	(1.65E-05)	(1.08E-05)	(1.48E-05)	(1.42E-05)
Time Elapsed^2	9.29E-08 ***	2.08E-07 ***	1.33E-07 ***	6.88E-08 *	1.88E-07 ***	2.34E-07 ***	1.47E-07 ***	2.84E-08	1.60E-08
	(2.13E-08)	(3.17E-08)	(3.27E-08)	(3.46E-08)	(4.82E-08)	(4.63E-08)	(3.39E-08)	(4.55E-08)	(4.34E-08)
Viewer FE		YES			YES	.		YES.	
Break FE.		YES			YES			YES	
Viewer-Break FE			YES			YES			YES
R^2	0.017	0.135	0.366	0.048	0.187	0.571	0.001	0.272	0.710
N	4,257,112	4,257,112	4,257,112	4,257,112	4,257,112	4,257,112	4,257,112	4,257,112	4,257,112

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.