TV Ads and Search Spikes:
Toward a Deeper Understanding

Rex Y. Du[*], Linli Xu[†], Kenneth C. Wilbur[‡]

February 23, 2017[§]

Abstract

Television advertisements lead some multitasking viewers to take immediate, measurable actions online. We analyze minute-by-minute TV ad insertion and online search data for daily fantasy sports and pick-up truck brands. TV ads with small audiences can produce detectable search spikes for the advertised brand, with 75% of incremental searches occurring within two minutes. TV ads also generate post-ad searches for competitor brands. Search spikes vary with ad content: they are larger after brand-focused ads than after price-focused ads, and after less-informative ads than after more-informative ads. Search spikes also vary with contextual media factors (e.g., TV network, daypart, program genre), but their effects differ across brands. Taken as a whole, our findings suggest that marketers should consider post-ad search spikes in conjunction with other metrics when evaluating and purchasing TV ads.

Keywords: Advertising, Analytics, Attribution, Digital, Internet, Marketing, Online Search, Television

[*]University of Houston, rexdu@bauer.uh.edu, http://www.bauer.uh.edu/rexdu.

[†]University of Minnesota, linlixu@umn.edu, https://sites.google.com/a/umn.edu/linlixu.

[‡]University of California, San Diego, kennethcwilbur@gmail.com, http://kennethcwilbur.com.

1 Introduction

ComScore (2016) reports that Americans spend an average of 111 hours per month with TV and 74 hours per month with their smartphone. Nielsen (2016) reports that the average American spends 33 hours weekly with TV and 12 hours per week using a smartphone.[1] Hitwise (2016) reports that 60% of all online searches come from mobile devices. Casual empiricism, consumer surveys and passive device usage measurements all suggest that consumers frequently use television and smartphone simultaneously, especially during the traditional evening "Prime Time" TV viewing hours.

One consequence of television/smartphone multitasking is well established: TV ads lead some viewers to search online for the advertised brands. In fact, a small industry has recently developed to help marketers measure the effectiveness of TV ads through post-ad spikes in online searches. Several marketing research consultancies—including C3 Metrics, Google Analytics 360, Neustar MarketShare, TVadSync, TVSquared Advantage and Wywy SearchSync among others—offer proprietary solutions for attributing search spikes to TV ad spots.

However, limited research is publicly available beyond the main effect of TV ads causing search spikes. In this paper, we seek to deepen our understanding of how television advertisers should (or should not) leverage post-ad search spikes when evaluating and purchasing TV ads. More specifically, we set out to answer the following questions:

1. *Can the relationship between TV ads and post-ad search spikes be estimated using ordinary ad insertions?* Most evidence reported in the existing literature relies on ad spots in "must see" TV programs with audiences numbering in the tens of millions of viewers (e.g., the Olympics or Super Bowl). However, the large majority of TV advertising dollars go to ad spots shown to audiences between 100,000 and 10,000,000 viewers. To the best of our knowledge, this is the first study that links minute-by-minute online searches with ordinary TV ad insertions.

2. *How do TV ads influence post-ad searches for competing brands?* If a TV ad reminds

---

[1]The "TV" statistic counts live and time-shifted viewing of traditional broadcast and cable networks; it excludes internet video and over-the-top streaming services.

a viewer of a product category need, she may search for rival brands in addition to (or in place of) the advertising brand. Lewis and Nguyen (2015) document competitive spillovers in online display advertising. Experiments showed that banner ads on yahoo.com increased search for competing brands by 5-14% as much as the incremental search lift for the advertised brand.[2] Similar phenomena have not been reported in the context of TV ads and online searches.

3. *How do media factors moderate post-ad search spikes?* For a given audience size, almost nothing is known about how post-ad search spikes vary with contextual media factors such as TV network, program genre, daypart, slot in the break, or audience interest in the advertised category.

4. *Are the moderating effects of contextual media factors common across competing brands?* If post-ad search spikes vary primarily as a function of context (e.g., TV network, program genre, daypart), then brands can learn from each other's media plans. This would likely increase competitive advertising clutter and imply that media plans designed to optimize post-ad search spikes may increase competition as more consumers become informed about multiple competing brands. On the other hand, post-ad search spikes may vary with brands' target audience characteristics; for example, they may depend on consumers' prior brand knowledge, as in Draganska et al. (2014). In that case, each brand would need to conduct its own analyses to optimize its media plan, leading to potentially less competitive advertising clutter.

5. *How do post-ad search spikes differ between brand- and price-focused ad content?* Brand-oriented advertising targets a broader slice of the audience, at an earlier stage in the purchase funnel, when consumers are less knowledgeable about available products. Price-focused advertising, by contrast, is more relevant to the smaller segment of prospects who are nearing a purchase. Therefore, post-ad search spikes may differ

---

[2]A similar result was found by Sahni (2016), who showed that advertisements for restaurants increased phone referrals to similar restaurants. These results complement a larger literature showing positive spillovers from one competitor's advertising to rivals' sales, including Anderson and Simester (2013) in catalog retailing, Joo et al. (2016) in cruises, and Shapiro (forthcoming) in prescription drugs.

substantially between brand- and price-focused ads.

6. *How do post-ad search spikes vary with the informativeness of ad content?* If more informative ads attract more consumers to the brand, then search spikes might be larger after more informative ads. On the other hand, providing more information in an ad might reduce some consumers' need to search. Further, providing less information may invite more consumers to search for the brand by signaling a high level of product quality, as in the central theoretical prediction of Mayzlin and Shin (2011).

In the remainder of the paper we report two empirical studies of how minute-by-minute TV ad insertions influence minute-by-minute brand searches online. The first is a pilot study analyzing two daily fantasy sports brands over a period of three months; it provides preliminary answers to questions 1–3. Post-ad search spikes turned out to be readily detectable for these brands, even for ads with small audiences. Spillovers to competing brands' search volume were positive and approximately one-fifth as large as the effect on the advertised brand's search volume. After controlling for audience size and brand fixed effects, ad content and media factors explained an incremental 54% of the variation in post-ad search spikes.

The second study is carried out with three pick-up truck brands; it merges brand search data from half a million minutes with audience data for more than 40,000 national TV ad insertions. This main study replicates the results of the pilot study, and further shows that 1) about 75% of incremental search occurs either in the minute the ad begins or the following minute; 2) the moderating effects of media factors are brand-specific; 3) ad audiences who are more likely to buy in the advertised product category produce larger search spikes; 4) brand-focused ads produce larger search spikes than price-focused ads; and 5) search spikes are smaller after more informative ads.

Taken together, our findings show that, for TV advertisers who seek to maximize consumers' online information gathering, post-ad search spikes offer a causal attribution measure that is responsive, reliable and readily available. However, for TV advertisers who have additional goals (e.g., information provision, communicating current pricing, converting prospects to purchasers, etc.), we advise caution in treating post-ad search spikes as the single metric of interest in assessing TV ad effectiveness. In fact, a larger search spike could even be a

negative signal in some cases, as it might convey reduced effectiveness in communicating product information.

2 Prior Literature

Although most advertisers would prefer to estimate ads' influence on sales, recent literature has challenged the idea that such effects are reliably estimable. Effects of ads on sales are typically small (Sethuraman et al. 2011), easily confounded (Lewis et al. 2011) and experiments are often underpowered (Lewis and Rao 2015). Therefore, an advertiser might reasonably consider substituting digital responses like brand searches as an intermediate correlate of sales. For example, a C-level executive at a leading performance marketing agency recently said that "Search behavior is one of the most powerful signals we have because it shows us consumer intent."

In the literature, several studies have established the main effect of TV ads on online search. Zigmond and Stipp (2010, 2011) published the first case studies showing that TV ads during the 2008 and 2010 Olympics generated large spikes in Google searches for the advertised brands. Lewis and Reiley (2013) replicated this analysis for dozens of brands during the 2011 Super Bowl, finding that the heights of search spikes varied across brands by two orders of magnitude. Joo et al. (2014) found that TV ads for mature brands increased both the number of category-related searches and the advertised brand's share of category search. Nearly all evidence to date focuses on estimating the main effect; little is known about potential moderators.

Further research has linked TV ads to other dimensions of online response.[3] Kitts et al. (2014) found large increases in new visitors to a brand's website within minutes of TV ad insertions; website traffic from returning visitors showed no change. Hu, Du and Damangir (2014) incorporated online search data into a marketing-mix modeling framework, showing that brand search volume helps to predict weekly automotive sales better both in and out

---

[3]The reverse path has also been well established: online promotions significantly alter television viewing (Godes and Mayzlin 2004, Gong et al. forthcoming). Lamberton and Stephen (2016) and Kannan and Li (2016) review the literature in greater depth.

of sample. Liaukonyte et al. (2015) showed that brand website traffic and sales increased within two minutes of TV ad insertions, with patterns of effects that depended on advertising content. He and Klein (2016) estimate the effect of TV ads on online sales of lottery tickets, finding strong effects lasting up to four hours. Tirunillai and Tellis (forthcoming) and Fossen and Schweidel (forthcoming) found that TV advertising improves multiple aspects of online word-of-mouth, including quantity, diffusion and valence.

Beyond the main effect of TV advertising on online searches, we seek to contribute to the prior literature by investigating competitive spillovers and highlighting the moderating effects of ad content and media factors. More broadly, we hope our paper can enhance our understanding of how and when marketers should (or should not) leverage post-ad digital responses when optimizing their advertising campaigns.

3   Pilot Study: Daily Fantasy Sports

In 2014-2015, Daily Fantasy Sports (DFS) brands DraftKings (DK) and FanDuel (FD) blanketed the sports television landscape with an estimated $375 million in TV advertising across thousands of national spots. DFS services are offered exclusively online, so it was natural to expect that the TV ads would generate online search, offering an opportunity to test the initial research questions. Another advantage of using this category is that brand names are unique so that brand searches are easily distinguished from unrelated search terms. However, it was not clear *a priori* whether spikes in brand searches could be reliably detected after "ordinary" TV ads, i.e. those with relatively small audiences.

3.1   Data

We merged minute-by-minute data from three sources.

TV ad insertion data were taken from Kantar Media's "Stradegy" database, a comprehensive source of competitive advertising intelligence that includes all ads played on major national networks and local broadcast stations.[4] For each ad insertion, we observe the date, start time, duration, advertised brand, TV network, program genre, ad creative identifier,

---

[4]95% of DFS expenditures went to national ads, so we do not analyze local ads in this section.

and a cost estimate.[5]

Ad audience data come from comScore's "TV Essentials" database. ComScore collects viewing data passively from 52 million digital set-top boxes in 22 million households.[6] In recent years, comScore has emerged as a formidable rival to Nielsen due to its nearly 1,000-fold advantage in sample size.[7] We observe national audience estimates for every ad insertion in every minute on each national TV network.

Finally, brand search volume was constructed by combining information from Google AdWords Keyword Planner and Google Trends.[8] Keyword Planner provides total brand search volume estimates for the sample period whereas Google Trends provide brand search indices by week, hour-within-each-week and minute-within-each-hour. We combined Keyword Planner total brand search volume estimates with Google Trends brand search indices to obtain, sequentially, brand search volume estimates for each week, each hour within each week, and finally, each minute within each hour.

We merged the three datasets at the minute level for the period August 1—October 31, 2015. Ad audience data were assigned to the minute in which the ad began.

3.2   Model-Free Evidence

Figure 1 plots minute-by-minute search volume for DFS brands between 3:00 P.M. and midnight Eastern Time on Thursday, September 10, 2015. The date was chosen to coincide with the first professional football game of the season, which was watched by 31 million viewers.

Both brands' search levels were fairly stable from minute to minute, with small fluctuations through the course of the day, and some spectacular but short-lived outliers during Prime Time. DraftKings was typically searched more than FanDuel, except when FD

---

[5]TV networks report ad costs to the Standard Rate and Data Service as telecast-level averages.

[6]ComScore receives a census of usage data provided by large cable, satellite and fiber-optic network operator partners. It stratifies the data to reflect unrepresented subpopulations (e.g., antenna-only households).

[7]Nielsen estimates national audience sizes using a more invasive technology, PeopleMeters, in a permission sample of 26,000 households (VAB 2016).

[8]Data were collected for all queries containing "draftkings," "draft kings", "draftking," "draft king," "fanduel" and "fan duel".

Figure 1: DFS Search Volume: Sept. 10, 2015, 3:00-11:59 P.M. E.S.T.
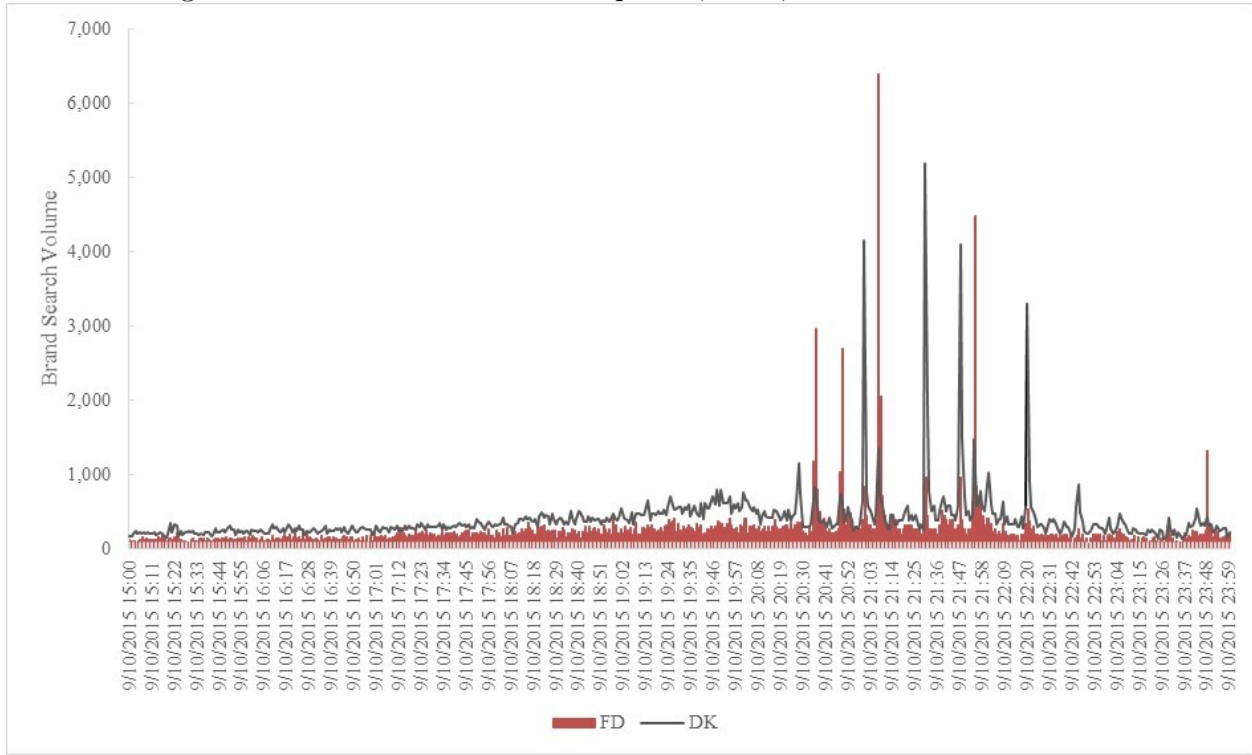
searches spiked.

Figure 2 zooms in on the period from 9:01 to 9:59 P.M., part of the professional football season-opening broadcast. This period contained two DraftKings ads and two FanDuel ads. The gray bars represent minutes of the telecast that contained ads.

The first FanDuel commercial began at 9:06 P.M., immediately followed by a 25-fold spike in FD search, from 255 in the minute before the ad to 6,397 in the minute after. DraftKings' first ad aired in minute 29, followed by a 20-fold increase in DK search. The final two commercials produced 11- and 18-fold search spikes. In total, these four advertisements produced 32,861 incremental DFS searches, or about 0.42 incremental searches per 1,000 ad viewers.[9]

Four features of the figure are notable.

1. In all four cases, brand search volume reverted to its pre-ad baseline within about five

_____

[9]This calculation takes the search volume in the minute prior to each ad insertion as the baseline for the post-insertion period.

Figure 2: DFS TV Ads and Search Spikes: Sept. 10, 2015, 9:01-9:59 P.M. E.S.T.
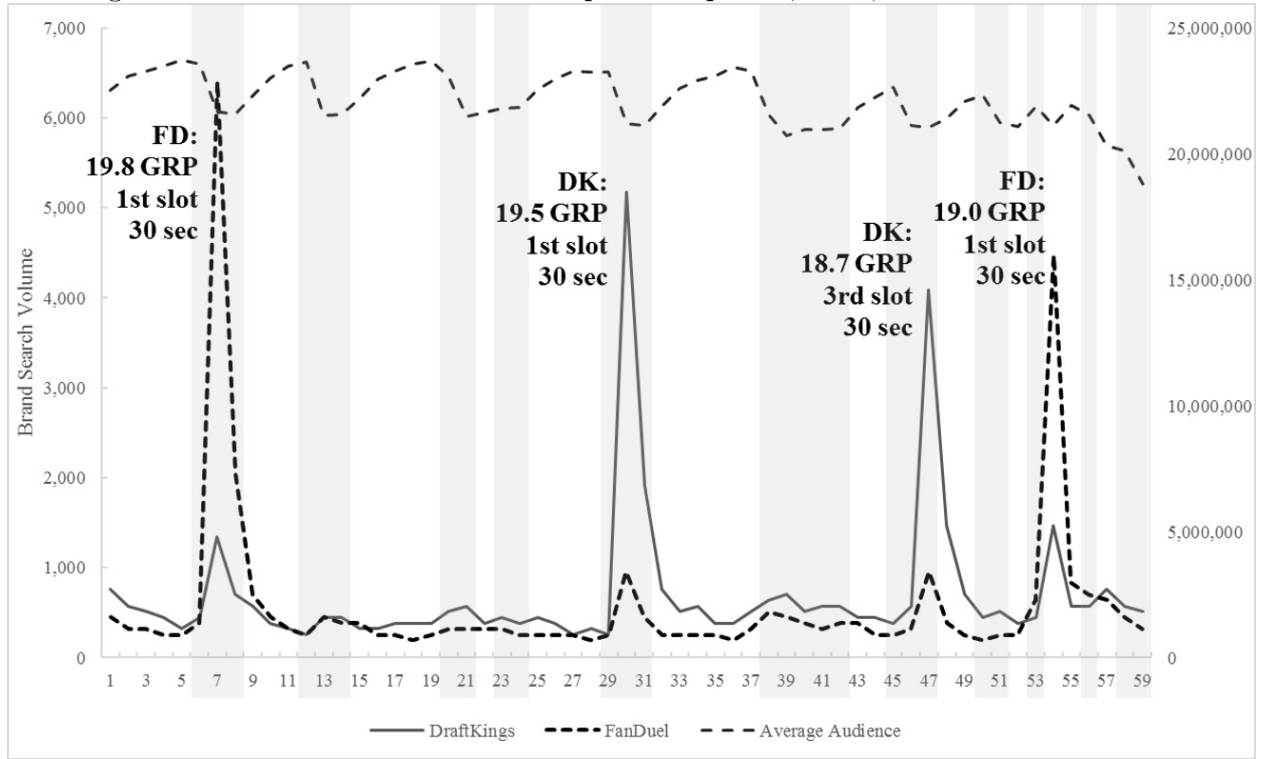
minutes or less.

2. All ads produced search spikes for both brands, with competitive spillovers one-fifth the size of own effects, on average.

3. Ad minutes with non-DFS commercials were not associated with noticeable DFS search spikes. This suggests the DFS search spikes are caused by the presence of DFS ads, rather than by the absence of the game.

4. The figure shows no dips following search spikes. This suggests that DFS ads produced incremental searches rather than "stealing" searches that otherwise would have occurred a few minutes later.

These four features were common to all of the data visualizations we inspected.

9

3.3 Analysis

Our goal in the pilot study is to estimate the "Cumulative Abnormal Search" (CAS) for a subset of ads, and then to relate CAS to ad characteristics.

We identified all DFS commercials that (i) *did not follow* any other DFS advertisement by less than 21 minutes and (ii) *were not followed by* any other DFS advertisement by less than 5 minutes.[10] This procedure resulted in 332 ad insertions for which we can be fairly certain that any search spikes following a TV ad are caused solely by the ad insertion in question. However, a drawback of this identification strategy is that these ad insertions are not representative of the population: they were disproportionately seen by smaller audiences, aired in August and October, and outside of prime time (daytime or late night).

Figure 3 depicts the research design. The minute-by-minute baseline search prior to each ad insertion minute is modeled as a linear function of lagged search in each of the previous five minutes. For each ad, we use the minutes immediately prior to the insertion minute to estimate a linear regression specific to that advertisement.[11] We then use that regression to project what the baseline search volume would have been in the five minutes following the ad, under the counterfactual that the ad did not take place. Finally, we define the CAS as the cumulative difference between observed and predicted brand search volume in the five minutes after the ad.

This pre/post design seeks to rule out confounds. DraftKings and FanDuel advertised extensively during the sample period, including numerous non-television media. By partialing out minute-by-minute changes in local search trends, we are able to isolate the effect of TV ads on search based on the knowledge that DFS brands were unable to coordinate timing at the minute level between TV and non-TV advertising efforts.[12]

---

[10]

[11]A longer "pre" period results in more stable estimates of post-advertisement baseline search, but fewer ad insertions available for analysis. We dropped minutes 21–19 prior to the ad insertion minute because search during these minutes could be inflated by another ad that was aired 22 minutes prior to the ad insertion in question. Minutes 1-13 prior are used to estimate the five-lag regression model in Equation 1, as minute 13 prior is the first minute that does not include minute 19 as a predictor.

[12]Standard TV advertising contracts specify timing in a particular quarter-hour on a particular date, because TV networks do not precommit to starting commercial breaks at specific times (Liaukonyte et

Figure 3: Depiction of DFS "Isolationist" Approach
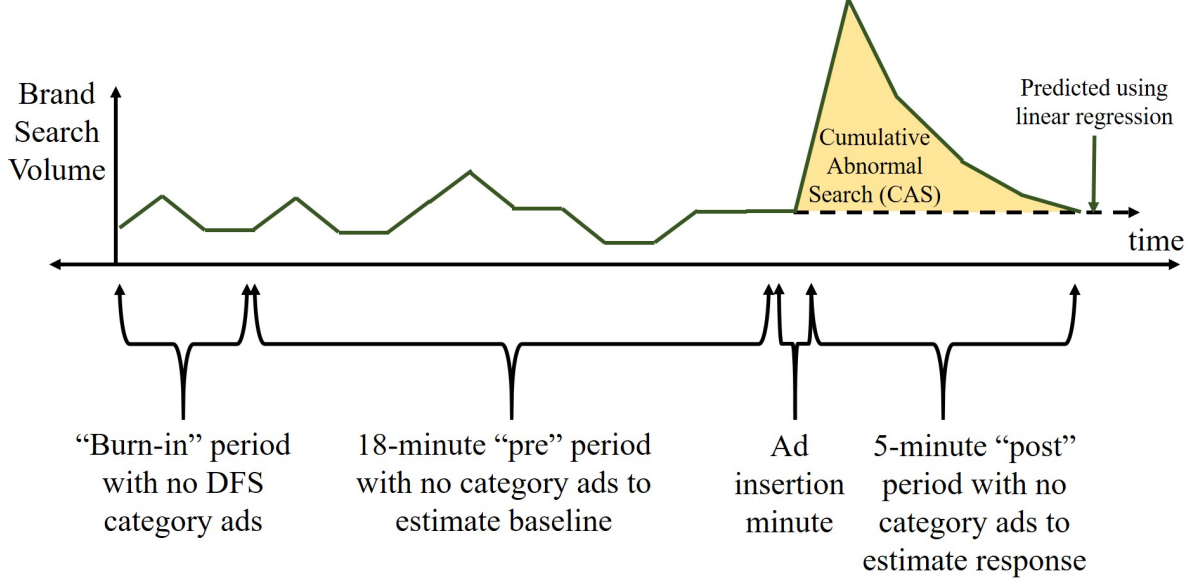
More specifically, we proceed in three steps to estimate CAS. First, for each TV ad, we run a linear regression using pre-ad search data (332 regressions total). For each insertion i of brand b_i which occurred in minute m_i, we use search data from the 13 minutes prior to the insertion minute($i.e.$ for $m = m_i - 13, ..., m_i - 1$) to estimate the following equation:

$$(1) \qquad y_m^i = \sum_{t=1}^{5} y_{m-t}^i \beta_t^i + \epsilon_m^i$$

where y_m^i is search for brand b_i in minute m, β_t^i is the tendency for y_m^i to increase with its tth lag y_{m-t}^i, and ϵ_m^i is a white-noise error.

Second, we use the fitted values from Equation 1 to predict what search for brand b_i would have been in minutes $m = m_i, ..., m_i + 4$, had the advertisement not occurred. The counterfactual search level in minute m_i is defined as $\hat{y}_{m_i}^i = \sum_{t=1}^{5} y_{m_i-t}^i \hat{\beta}_t^i$. For each of the four minutes $m_i + \tau$ after the ad insertion (for $\tau = 1, ..., 4$), the counterfactual search level is defined as $\hat{y}_{m_i+\tau}^i = \sum_{t=1}^{\tau} \hat{y}_{m_i+\tau-t}^i \hat{\beta}_t^i + \sum_{t=\tau+1}^{5} y_{m_i+\tau-t}^i \hat{\beta}_t^i$.

Third, we calculate CAS as the difference between the predicted and observed search

al. 2015, Wilbur et al. 2013). Because advertisers could not select particular ad minutes, time-varying unobservables at the minute level are unlikely to contaminate the causal estimates.

levels:

$$(2) \qquad CAS_i = \sum_{m=m_i}^{m_i+4} (\hat{y}_m^i - y_m^i).$$

3.4   Results

All 332 ads produced positive estimates of CAS. Figure 4 shows that CAS and ad audience size are strongly related, with a correlation of 0.5. On average, an ad in this subsample was seen by 221,000 viewers and produced 116 incremental brand searches. The average of 0.52 incremental searches per 1,000 ad viewers was 24% higher than the ads analyzed in subsection 3.2. It might be that must-see TV program audiences are not ideal for brands seeking to maximize post-ad brand search, as high program engagement or co-viewing behavior could reduce the incidence of multitasking with smartphones.

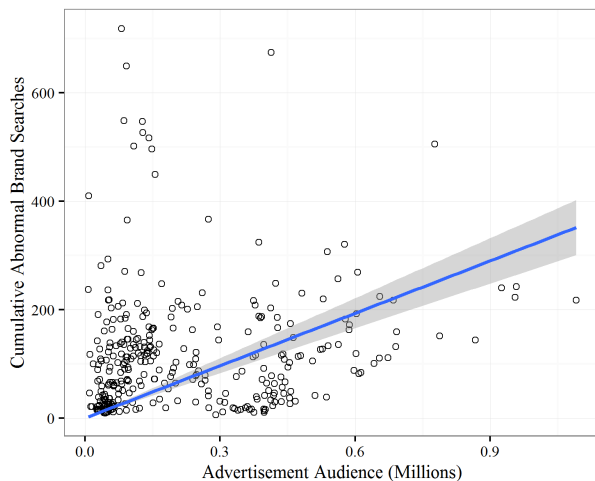Figure 4: Ad Audience Size and Cumulative Abnormal Search

Figure 5 splits the relationship between CAS and ad audience size by brand and by month.[13] FanDuel ads produced larger search spikes per viewer than DraftKings ads. Calendar month correlated perfectly with advertised products: baseball and golf were advertised in August, with football in October. The October ads show a much stronger response per viewer than the August ads, motivating further investigation of how ad content may moderate search spikes.

---

[13]No trend line appears for month 9 because only one ad in the subsample aired in September.

Figure 5: Ad Audience Size and Cumulative Abnormal Search, by Brand and by Month
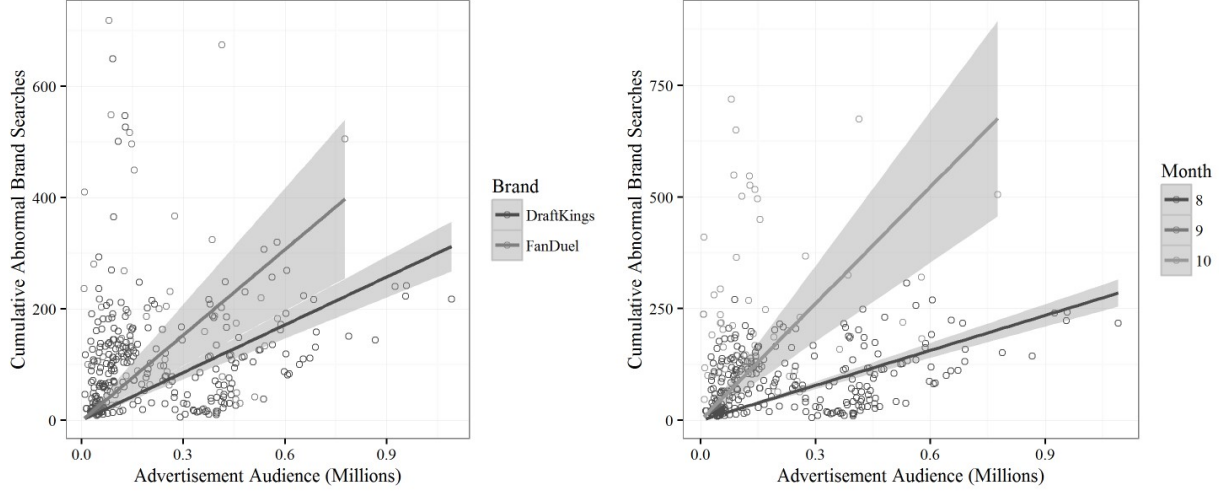
Finally, we ran a series of linear regressions to investigate how much variation in CAS could be explained by advertisement characteristics. All regressions included brand fixed effects and ad audience size. Separate regressions included fixed effects for one of the following: program genre, ad duration, first slot in the break, TV network, ad creative, daypart, or calendar month. One regression included all ad characteristics.

Figure 6 provides the percentage of CAS variation explained by each regression. Advertisement audience size and brand fixed effects explain about 9% of the variation in CAS. Several ad characteristics (calendar month, daypart, ad creative and TV network) have a greater partial R-square statistic. Other variables (duration, program genre) had less explanatory power. The regression that included all variables had far greater explanatory power (R-square = 63%) than any smaller set of advertisement characteristics, suggesting that multiple contextual factors are important predictors of post-ad search spikes.

To summarize what we learned from this pilot study, it seems that media multitasking is so widespread that some brands can now use minute-by-minute data to estimate post-ad search spikes, even for ad insertions with small audiences. Competitive spillovers tend to be positive, and many advertisement characteristics are important predictors of post-ad search spikes. Although we may not extend the quantitative conclusions of our DFS analysis to other product categories, it seems reasonable to suspect that some of the qualitative

Figure 6: % of CAS Variation Explained using Ad Audience Size, Brand Fixed Effects, and Contextual Media Factors
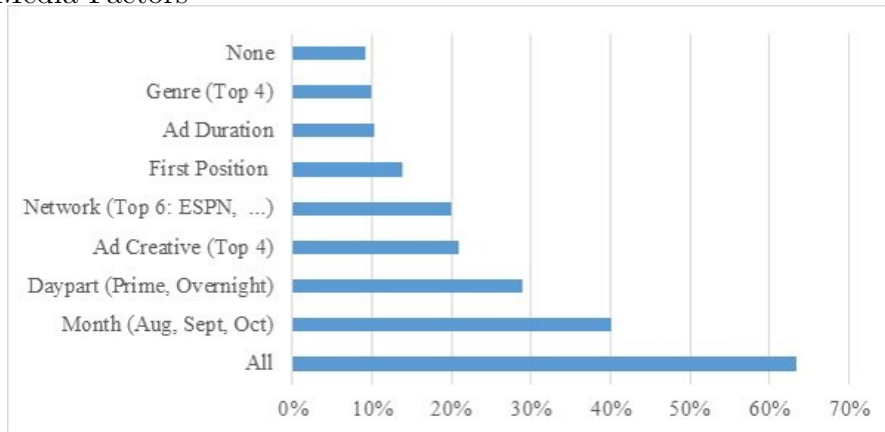
conclusions may hold. The next section replicates and extends these findings in a larger sample from a mature product category.

4   Main study: Pick-up trucks

The second empirical context we study is full-size pick-up trucks. This category is mature and well-known, and consumers search online for product information, but much of their shopping and purchasing activities are conducted offline. According to *Motor Intelligence*, Ford F-150 maintained a market share of about 31% in 2016, followed by Chevrolet Silverado at 22% and Ram Trucks at 18%, continuing a 33-year trend of stability in market share rankings (Xu et al. forthcoming). We focus on these three top brands for this study.

4.1   Data

Data were drawn from the same three sources as in the pilot study. The sample period ranged from February 15, 2015 to January 23, 2016.[14]

The data contain 493,920 minute-level observations for each brand. Overall, Ford spent $1.4B on TV ads in this category and was searched 36MM times; Silverado spent $2.4B on TV ads and was searched 21MM times; and Ram spent $2.6B on TV ads and was searched

---

[14]These dates focused the sample on ordinary advertisements by avoiding Super Bowl outliers.

19MM times.

The content of TV ads for these three pick-up truck brands varies between national and local media. National TV commercials are purchased exclusively by the manufacturer and, according to a content analysis by Xu et al. (2014), typically carry brand-oriented messages with relatively few price-oriented messages. Local TV advertising is done by both manufacturers and local dealers associations, with both parties designing ads that extensively communicate current market-specific pricing terms and promotions. Although we are not able to observe audience sizes for local ads, we aggregate local ad expenditures to the national level within each minute for each brand for inclusion in the analysis.

In addition to search volume data for the three pick-up truck brands, we also collected the number of Google search queries containing the word "SUV" in each minute. The goal was to improve baseline brand search volume estimates by controlling for unobserved time-varying correlates of consumer search for large automobiles.

Two additional data sources were available for analysis. The first comes from Polk Automotive Intelligence and was integrated into comScore's TV Everywhere platform. Polk aggregates data on all new automobile registrations and integrates its data with comScore and other partners at the household level. Polk uses proprietary algorithms to estimate what fraction of viewers in each ad's audience was contemporaneously in the market for a new pick-up truck. We refer to this variable as *NewPickup*.

We also collected advertisement content ratings from Ace Metrix, a leading provider of competitive intelligence about advertising content. Ace Metrix identifies new national ads and, within 24 hours of their first airing on TV, administers a standard survey to 500 online panelists. We were able to match Ace Metrix survey scores to 185 distinct ad creatives, accounting for 92% of the total ad exposures in the sample. The survey instrument that interests us the most is the item called *AdInfo*: "To what extent did this commercial make you feel 'I learned something?'" Answers were given on a scale of 0-100.[15]

An informal but comprehensive review of ad videos in our sample built confidence in

---

[15]Additional survey responses were available (e.g., attention, liking, etc.) but they are highly correlated with *AdInfo*, with correlations ranging from 0.60 to 0.87. The limited number of ad creatives was insufficient to estimate effects of multiple dimensions of ad content.

the validity of *AdInfo*. For example, the commercial with the highest mean AdInfo score was a 60-second spot for Ford. It communicated attributes related to aluminum alloy body composition, total vehicle weight, hauling capacity, towing capacity, and the number of patents used in its design. At the other extreme, the ad with the lowest mean AdInfo score was a 30-second spot for Ram. The complete script reads: "In a work, work, work world, take time for Sunday. Just know that your truck has a little thing for Monday." The first 23 seconds showed scenes of a workplace and then family members resting together in their home. The truck was not focally portrayed until the final seven seconds, and the brand was not clearly identified until the final two seconds.

4.2   Data Visualizations

Figure 7 graphs the brand search volume across the entire sample period. Similar to the DFS data, the time series of brand search volumes are fairly stable over time, with the exception of spectacular but short-lived spikes. The data show repetitive patterns across and within weeks but limited longer-term trends or seasonality.

Figure 8 visualizes search spikes accruing to a few ads with the highest spending in the sample. The post-ad spikes are unmistakable though less dramatic than the DFS brands pictured in Figure 2. Similar to the DFS category, brand search volume returns to baseline within five minutes after an ad airs, and competitor brands see small positive spillovers in online search.

A quick way to calculate a multiplicative search lift is to calculate a "Post/Pre Index:" we sum brand searches in an ad's starting minute and four following minutes, then divide by total search volume in the five minutes prior to the ad.[16] Figure 9 plots this index versus advertisement audience size, by brand. Consistent with the DFS analysis, there is an unmistakable positive correlation between post-ad search spikes and ad audience sizes.

Figure 10 shows how the relationship between post-ad search spike and ad audience size varies by a median split of *NewPickup* and (separately) *AdInfo*. Audiences with above-normal shares of interested consumers generate significantly more brand searches per viewer.

---

[16]The graphics include all ad insertions, so this index can be less than one. When two proximate ads air, the first ad's search spike inflates the pre-search level of the second ad.

Figure 7: Fullsize Pick-up Truck Brand Search Volume
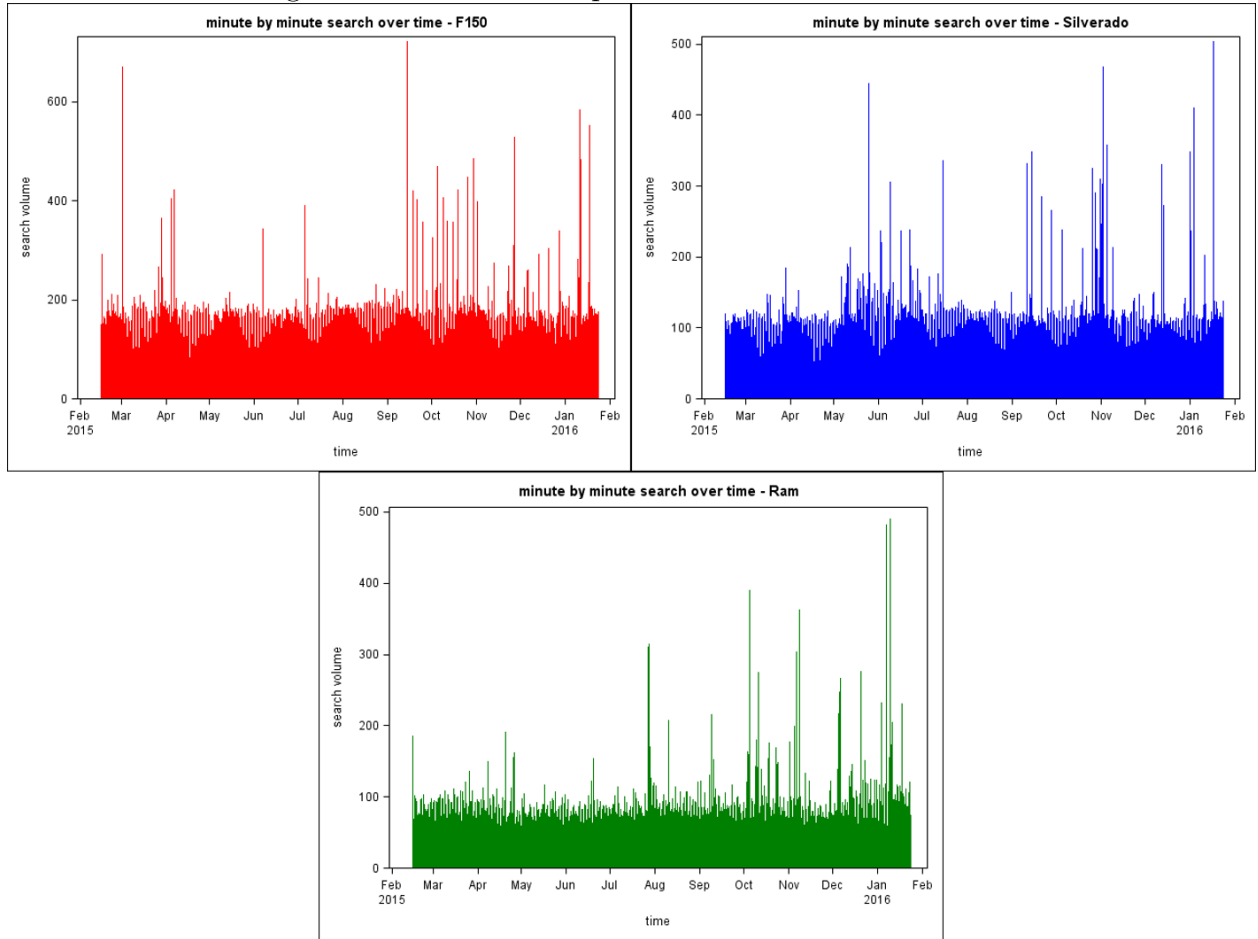
It also shows that more informative ads lead to fewer incremental searches per viewer. We suspect this is because an ad that conveys more information to the viewer may render search less necessary. It also may lead a prospective consumer to bypass the search engine and go to a brand or dealership website to get more information, or it may reduce the chance that the viewer starts multitasking after the ad.

Figure 8: Illustration of Post-ad Search Spikes
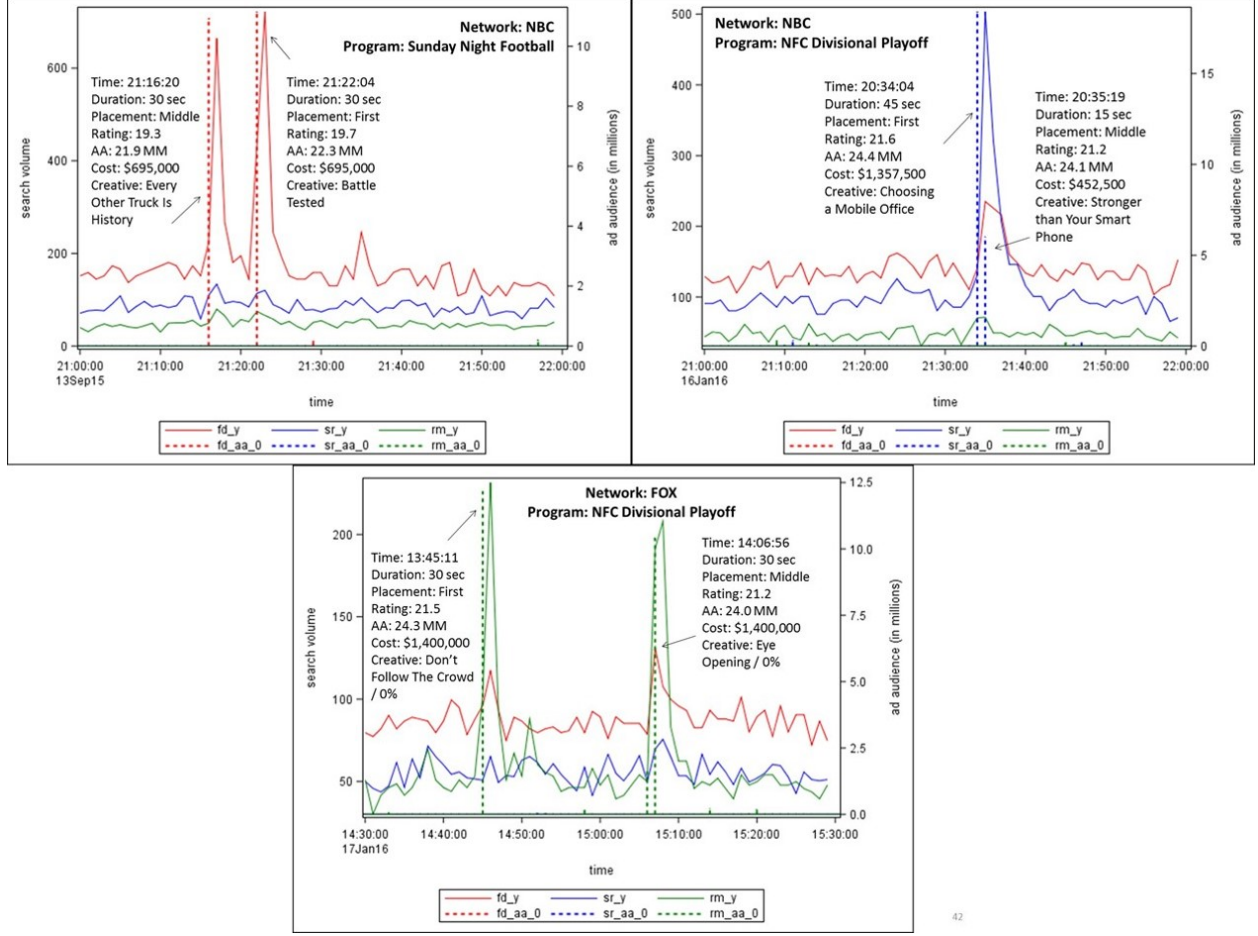
4.3 Model

We assume Google search volume for brand b in minute t, GS_{bt}, can be decomposed into distinct components:

$$
\begin{aligned}
GS_{bt} = \tau_{bt} &+ \sum_{i=0}^{M} [\phi_{bt,t-i} NA_{b,t-i}] + \sum_{i=0}^{N} \sum_{c=1}^{C_b} [\chi_{bct,t-i} NA_{c,t-i}] \\
&+ \sum_{i=0}^{N} [\psi_{bt,t-i} LA_{b,t-i}] + \sum_{i=0}^{N} [\omega_{bt,t-i} DA_{b,t-i}] + \epsilon_{bt}
\end{aligned}
$$

(3)

where

- τ_{bt} denotes baseline search for brand b in minute t, i.e., what the search volume would have been if there had been no TV ads, which we specify as a function of fixed effects, local trends and a covariate, as described further below;

18

Figure 9: Post-ad Search Lift vs. Ad Audience Size, by Brand
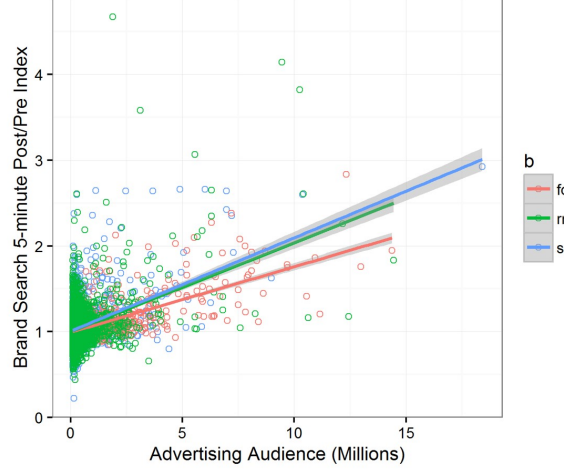
- $NA_{b,t-i}$ and $NA_{c,t-i}$ denote the total audience (in millions) exposed to national TV ads in minute $t-i$ for, respectively, brand b and its two competitors $c \in C_b$;

- $\phi_{bt,t-i}$ and $\chi_{bct,t-i}$ denote the rates at which ad audiences $NA_{b,t-i}$ and $NA_{c,t-i}$, respectively, conduct Google searches for brand b in minute t;

- $LA_{b,t-i}$ and $DA_{b,t-i}$ denote, respectively, the total expenditure (in \$10,000s) on local TV ads by brand b and its dealership associations in minute $t-i$;[17]

- $\psi_{bt,t-i}$ and $\omega_{bt,t-i}$ denote the rates at which \$10,000 in $LA_{b,t-i}$ and $DA_{b,t-i}$, respectively, produce Google searches for brand b in minute t; and

- ϵ_{bt} denotes the error term, which is given a moving-average representation, $\epsilon_{bt} = e_{bt} + \sum_{i=1}^{59} \rho_{bi} e_{b,t-i}$ with $e_{bt} \sim$ i.i.d. $N(0, \sigma_b^2)$, to allow for serial correlation.
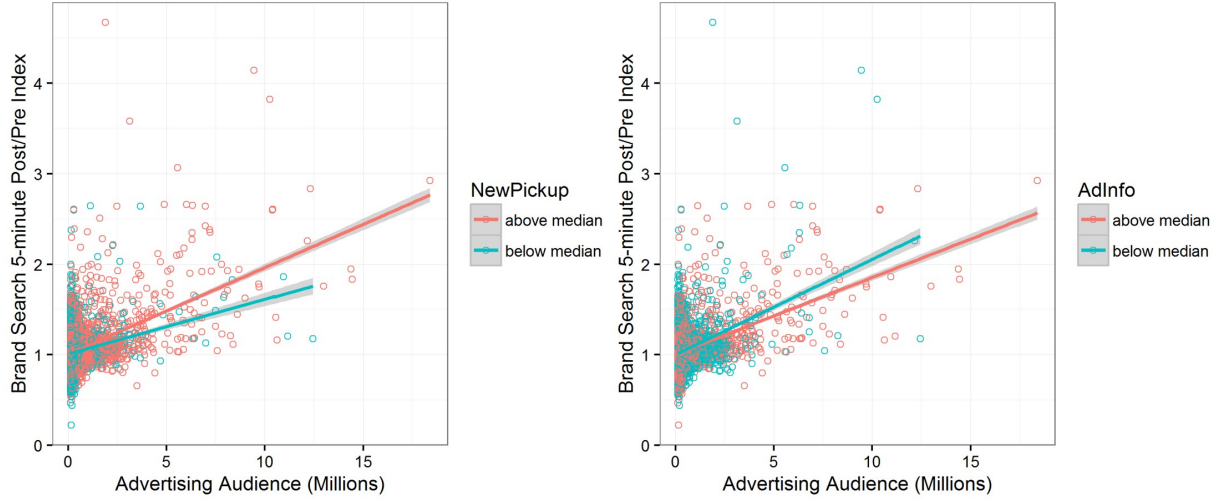
Of key interest is $\phi_{bt,t-i}$, the parameter governing search response to brand b's own national TV ads, which we specify as follows:

$$(4) \qquad \phi_{bt,t-i} = \alpha_{natl,bi} \exp(\gamma_{b,week(t)} + \theta_{b,dow(t)} + \sum_{j=1}^{J} \beta_{bj} X_{bj,t-i})$$

where

---

[17]The average cost of one television ad exposure is about \$.01, so \$10,000 approximates 1MM ad exposures.

Figure 10: Post-ad Search Lift vs. Ad Audience Size, Median Split by NewPickup and by AdInfo

- $\alpha_{natl,bi}$ denotes the baseline rate of brand b search response in minute t to 1 million national TV ad exposures in minute $t - i$;[18]

- $\gamma_{b,week(t)}$ and $\theta_{b,dow(t)}$ represent week and day-of-week fixed effects, respectively, to allow the average search response to TV ads to vary across and within weeks;

- $X_{bj,t-i}$ denotes the jth contextual factor that characterizes brand b's TV ads in minute $t - i$ and $\exp(\beta_{bj})$ captures the multiplier effect of this factor on $\phi_{bt,t-i}$.[19] The factors included are (i) *NewPickup*, (ii) *AdInfo*, (iii) a fixed effect for whether the ad is placed in the first slot of a commercial break, (iv) fixed effects for the two most common dayparts (Prime Time and Overnight), (v) fixed effects for the four most common TV networks (CBS, FOX, ESPN and NBC) and (vi) fixed effects for the three most common program genres (Pro Football, Auto Racing and Sportscast).

The response of brand search to competitor ads, local ads and dealership ads was more muted than response to the brand's own national ads, preventing robust identification of the moderating effects of contextual factors for these ad spots. We adopt simpler specifications

---

[18]We do not constrain the $\alpha_{natl,b,i}$ parameters; the model can capture flexible patterns of delayed response.

[19]For minutes containing multiple TV ads, the contextual factors are audience-weighted averages.

for those response parameters:

$$\chi_{bct,t-i} = \alpha_{natl,bci}$$

(5)
$$\psi_{bt,t-i} = \alpha_{loc,bi}$$

$$\omega_{bt,t-i} = \alpha_{dealer,bi}.$$

It is important to specify the baseline search pattern flexibly to avoid conflating advertising effects with correlated unobservables. To minimize omitted variable concerns, we formulate τ_{bt} as follows:

(6)
$$\tau_{bt} = \mu_{b,hour(t)} + \eta_{b,week(t)}t + \lambda_{b,hour-of-week(t)}t + \kappa_b SUV_t$$

where

- $\mu_{b,hour(t)}$ is a fixed effect for the specific hour containing minute t;

- $\eta_{b,week(t)}$ is a fixed effect that estimates a separate trend in baseline search for each week of the sample period;

- $\lambda_{b,hour-of-week(t)}$ is a fixed effect that estimates a separate trend in baseline search for each hour of the week (i.e., Monday 12 A.M., Monday 1 A.M., ..., Sunday 11 P.M.); and

- SUV_t denotes the level of searches containing the keyword "SUV" in minute t.

Even if pick-up truck brands chose to advertise in particular hours of the sample period based on unobserved, hour-specific determinants of brand search, search response parameters would still be unbiased, since the hour-of-sample fixed effects $\mu_{b,hour(t)}$ control for the average search level in each given hour of the sample period.

One might also be concerned about unobservables that have within-hour trends and are correlated with a brand's ad insertions. This is why we include the fixed trend parameters $\eta_{b,week(t)}$ and $\lambda_{b,hour-of-week(t)}$, as well as SUV_t to control for consumers' tendency to search for large automobiles in any given minute of the sample period. We are further comforted by the knowledge that TV advertisers are not able to select particular minutes for their advertisements to start, as standard TV advertising contracts do not pre-specify the minute

21

in which the commercial break will begin, and typically do not pre-specify the slot of the break in which an ad will air (Liaukonyte et al. 2015, Wilbur et al. 2013).

We estimate the model by using pairs of consecutive minutes within each hour to difference out the hour-of-sample fixed effects $\mu_{b,hour(t)}$. Applying the first-differencing operator, i.e. $\Delta x_t = x_t - x_{t-1}$, transforms the model into the following mathematically equivalent representation:

$$\Delta GS_{bt} = \eta_{b,week(t)} + \lambda_{b,hour-of-week(t)} + \kappa_b \Delta SUV_t$$

(7)
$$+ \sum_{i=0}^{M} [\phi_{bt,t-i} NA_{b,t-i} - \phi_{b,t-1,t-1-i} NA_{b,t-1-i}] + \sum_{i=0}^{N} \sum_{c=1}^{C_b} \alpha_{natl,bci} \Delta NA_{c,t-i}$$

$$+ \sum_{i=0}^{N} \alpha_{local,bi} \Delta LA_{b,t-i} + \sum_{i=0}^{N} \alpha_{dealer,bi} \Delta DA_{b,t-i} + \Delta e_{bt} + \sum_{i=1}^{59} [\rho_{bi} \Delta e_{b,t-i}]$$

We estimate Equation 7 using nonlinear least squares with serially correlated residuals. For all three brands, brand search responses become statistically undetectable eight minutes after own national TV ads (i.e., M=8), and four minutes after competitor national ads, local ads and dealership ads (i.e., N=4).

4.4   Results

Figure 11 graphs the point estimates and standard errors of the average own and competitive effects of TV advertising on brand search, by minute following an ad insertion. The size of the own effects, averaged across each brand's ads, follows the ordering of brands by market share: Ford ads on average garner the biggest response, followed by Chevy and Ram.

All three brands' TV ads produce positive spillovers on their rivals' brand search. But here, again, effect magnitudes vary across competitors. Ford search is increased the most by its rivals' advertising. Ford ads also seem to produce greater spillovers; Chevy search increases more after Ford ads than after Ram ads, and Ram search increases more after Ford ads than after Chevy ads.

Although the three brands differ in the number of incremental searches generated per million ad viewers, the own effects follow a remarkably consistent temporal pattern. Figure 12 graphs the percentage of cumulative own effect realized by minute after the ad. About 15% of the cumulative own effect is realized in the minute the ad begins, followed by about

22

Figure 11: Main Effects of Brand Advertising and Competitive Spillovers
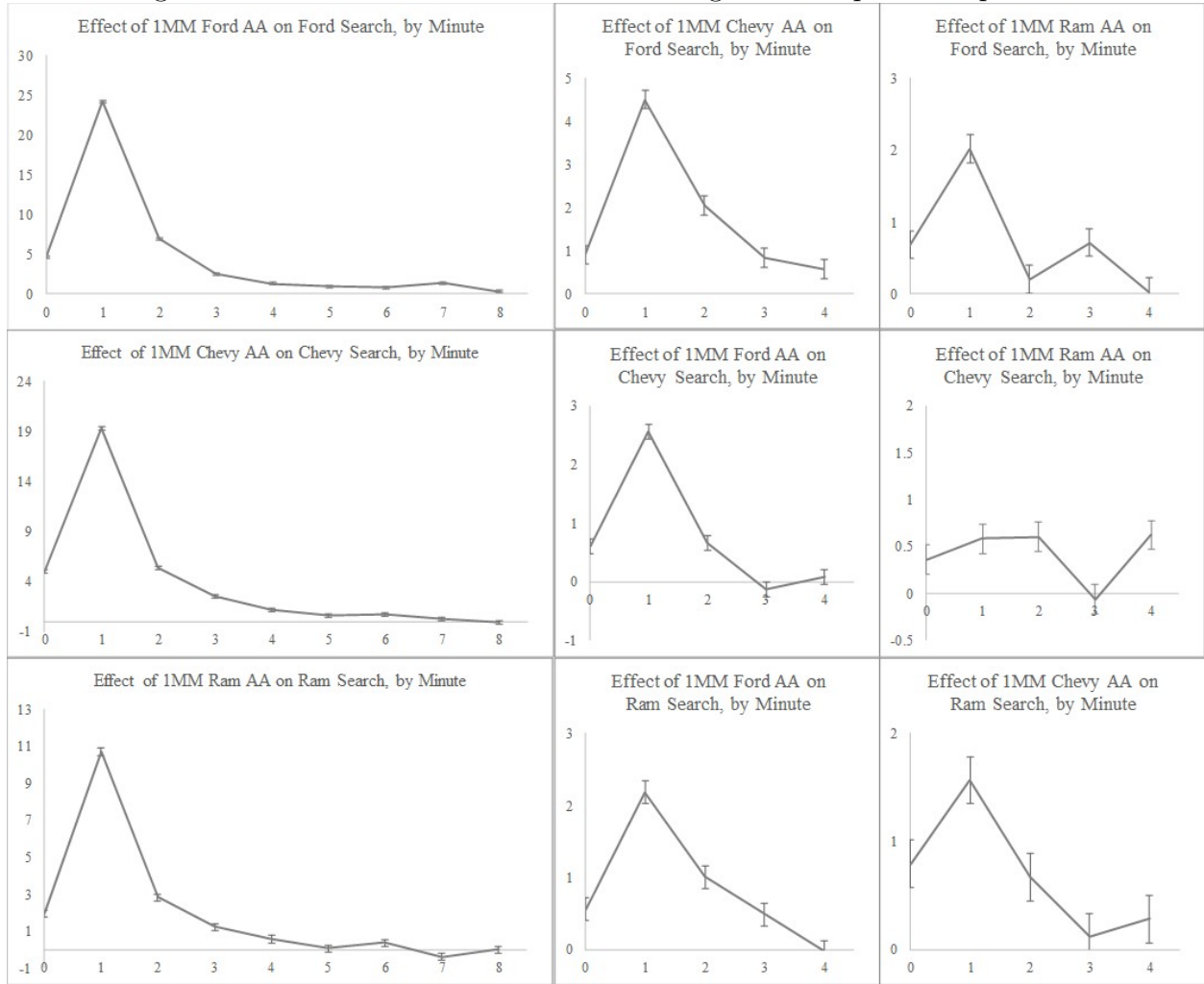
60% in the following minute, and nearly all of the remainder in the second and third minute after the ad.

Figure 13 reproduces the own effects of national TV ads for comparison to the estimated own effects of local and dealership TV ads. The average cost to reach one television viewer is about $.01, so $10,000 in local advertising approximates the effect of reaching 1 million ad viewers. The effects of local and dealership TV ads, which are predominantly price-focused, show a similar pattern in that the largest effects are realized in the minute after the ad, but the cumulative impact on post-ad brand search is only 13-17% as large as own national TV ads, which are predominantly brand-oriented. This is consistent with the intuition that

Figure 12: Percentage of Own Effects Realized by Minute
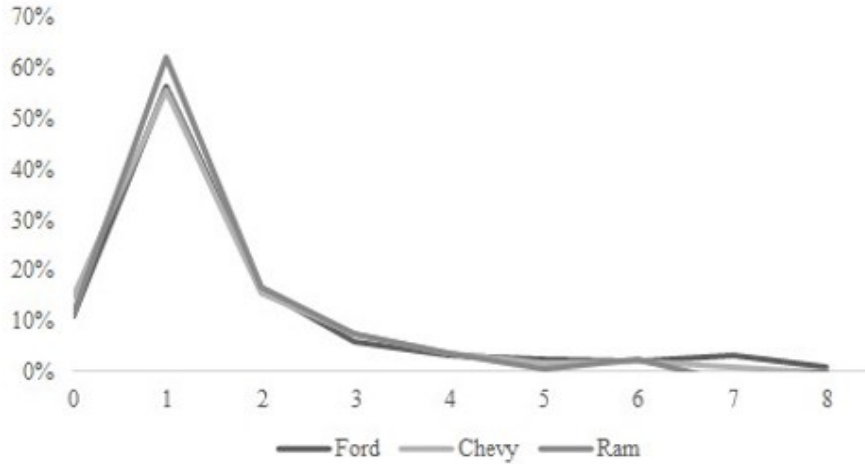
advertising to prospective consumers at lower funnel stages is likely to produce smaller search spikes, although it may be perfectly consistent with the marketer's advertising objective.

Figure 14 displays the estimated contextual factor multipliers. A multiplier of 1 indicates no impact on incremental searches per viewer. For the category leader, Ford, ad informativeness and first slot in a break do not influence the search spike much. Important search spike predictors are daypart, TV network, and program genre. More specifically, Prime Time and Overnight yield 50% more incremental searches per viewer on average; traditional broadcast networks yield fewer incremental searches per viewer than average, whereas ESPN produces about 84% more; Pro Football doubles the incremental searches per viewer whereas sportscast reduces it by 80%.

For Chevy and Ram, increasing *AdInfo* by one standard deviation reduces incremental searches per viewer by 33-44%. First slot in a break increases incremental searches per viewer by nearly 50%. Other multipliers are less similar between Chevy and Ram. Chevy ads generated far more incremental searches per viewer during Prime Time and Overnight, on FOX, and during Pro Football. Ram ads produced more incremental searches per viewer during Overnight, on FOX, CBS and ESPN.

Overall, the brand-specific media contextual effects seem to suggest that consumers sort into TV audiences based on factors that correlate with brand preferences or knowledge. We view the differences across brands as evidence that brands need to investigate their own data to discover how executional elements of their media plan influence multitaskers' online

24

Figure 13: Main Effects of Brand Advertising and Price Advertising
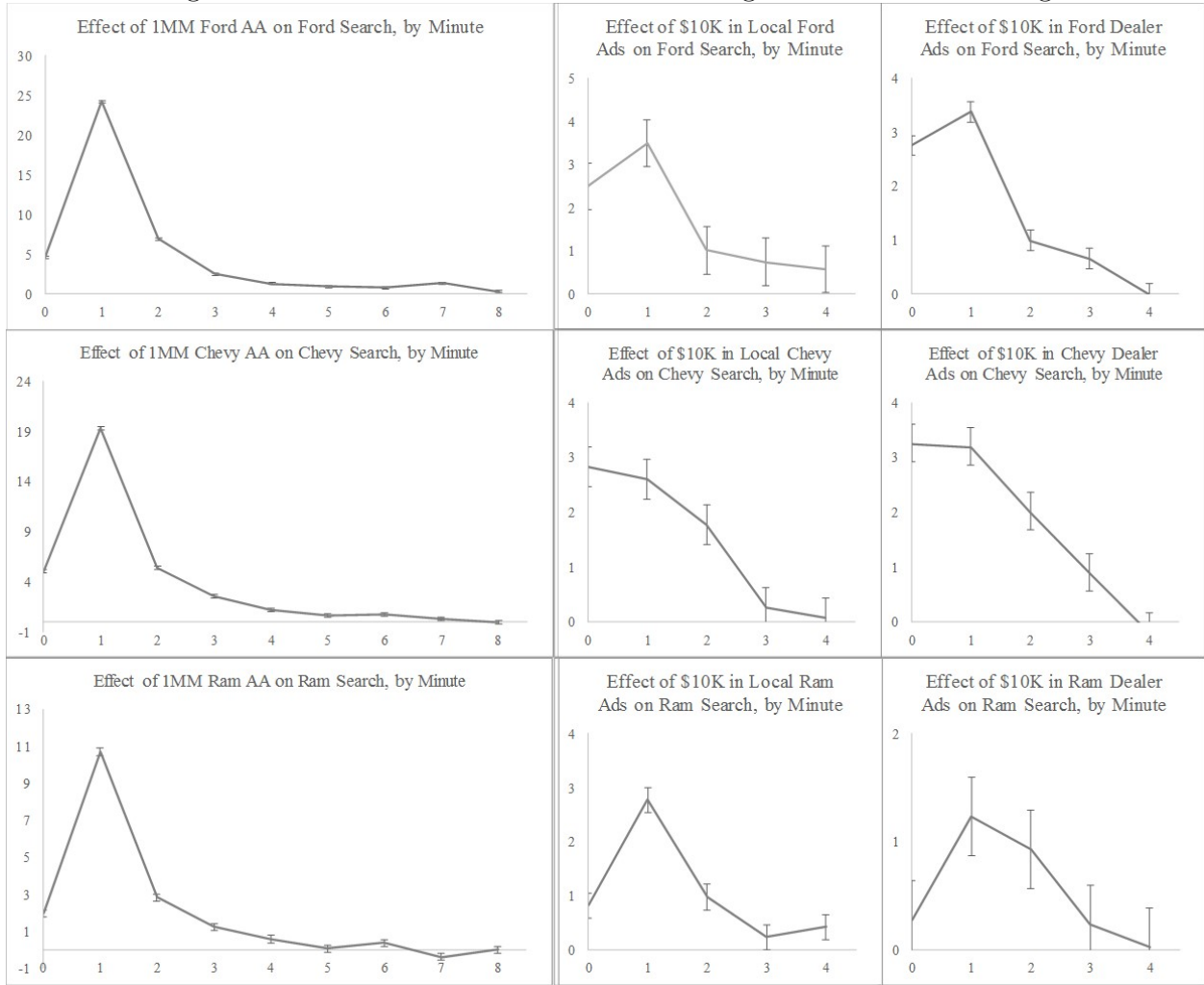

response to TV ads.

Figure 15 indicates the search spike enhancements to an increase of one standard deviation in *NewPickup*, i.e. the percentage of ad audience currently in the market for a new pick-up truck. Having a larger share of ad viewers who are interested in the product category generates far larger post-ad search spikes for Chevy and Ram than for Ford. Ford is the long-standing market leader in the category and presumably the brand consumers are most familiar with. The *NewPickup* result indicates that consumers nearing purchase are disproportionately more likely to seek out information about category followers Chevy and Ram after being exposed to a TV ad. This seems to suggest that consumers who are nearing

Figure 14: Brand-specific Search Spike Multipliers
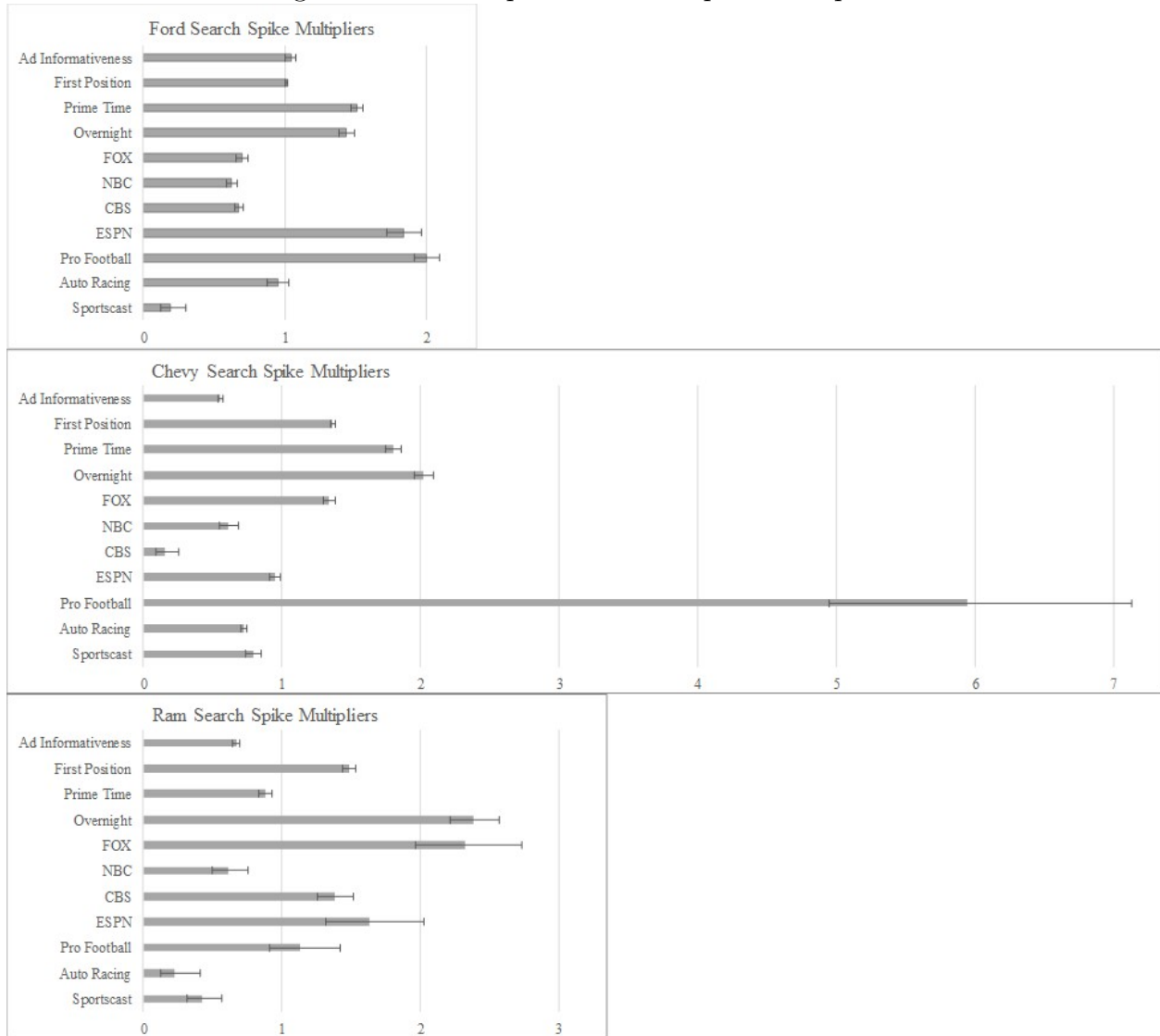
their purchase are especially prone to search category followers.

Figure 15: Brand-specific Search Spike Multipliers
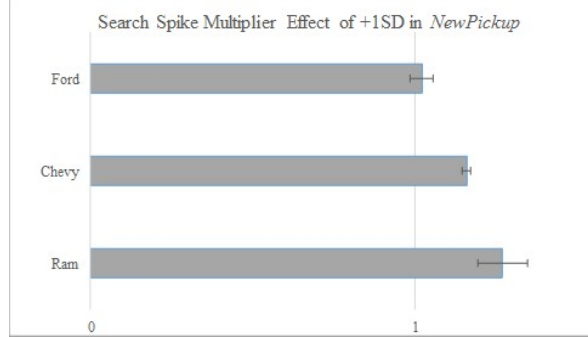
We close by reporting parameter estimates of the control variables in the baseline search function (Equation 6): hour-of-week trends ($\lambda_{b,hour-of-week(t)}$) and minute by minute "SUV" keyword search (κ_b). Figure 16 plots the brand-specific hour-of-week trend estimates (on the left).[20] We then calculate the implied average baseline search for each minute of the week (up to an intercept), as shown on the right. For all three brands, baseline search patterns look similar within each day of the week, though some intraday patterns differ between brands. Baseline search peaks during the morning hours and again during the evening prime time for Ford and Chevy; whereas Ram shows more muted variation within each day. Table 1 reports the brand-specific estimates of the control variable SUV_t which, as expected, is a positive (non-causal) predictor of pick-up truck brands' baseline search by minute. Finally, Figure 17 displays the estimates of the moving-average error process for each brand, indicating similar patterns of serial correlation in the residuals that fades away quickly within one hour.

Table 1: Effect of SUV Control Variable by Brand

	Estimate	Standard Error	
Ford	0.30	0.04	**
Chevy	0.16	0.04	**
Ram	0.29	0.04	**

Note: ** 99% significance level

---

[20]Week-specific trend parameters are reasonable but not particularly informative, so we exclude them for brevity.

Figure 16: Brand-Specific Hour-of-Week and Minute-of-Week Seasonality in Baseline Search
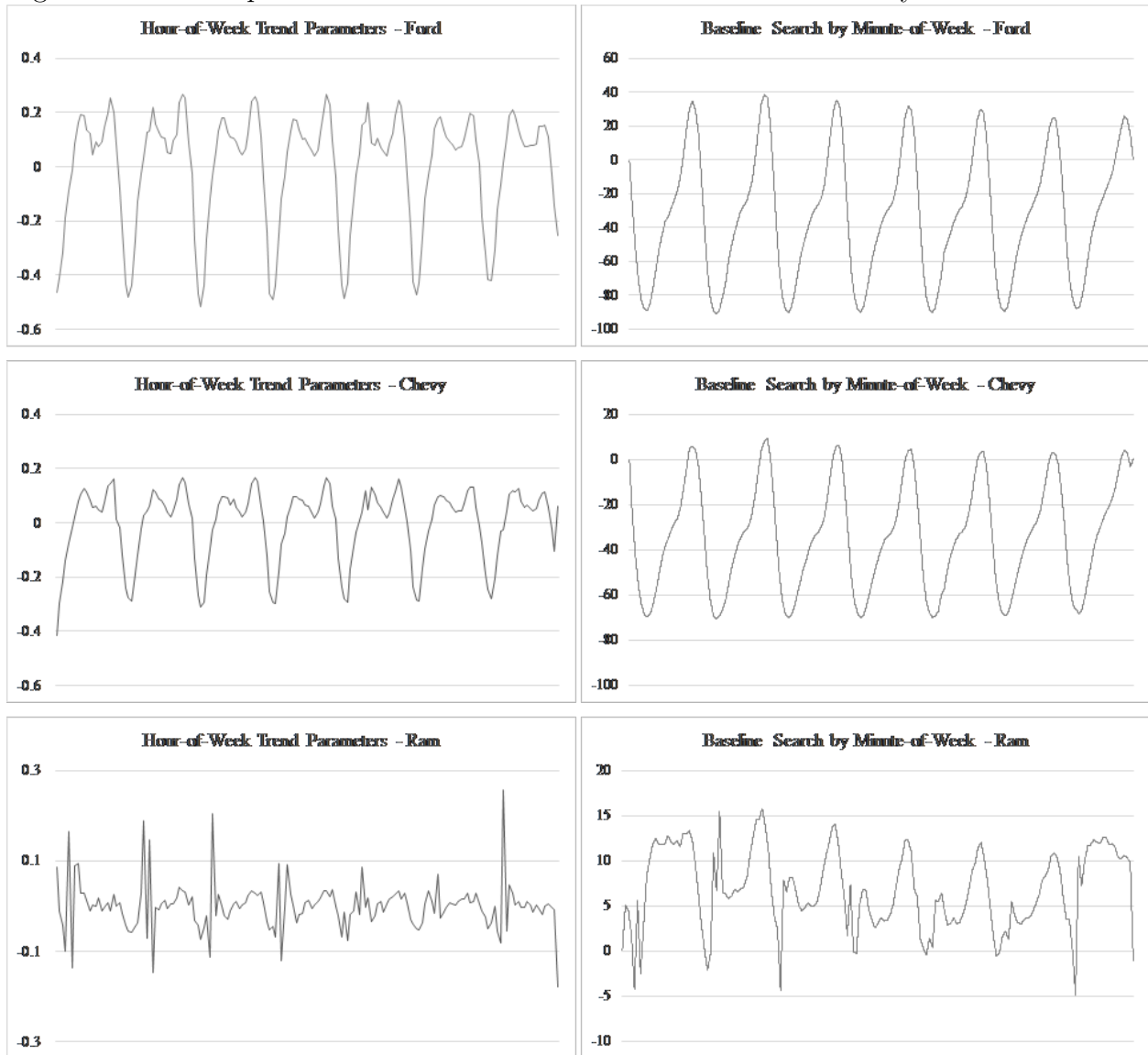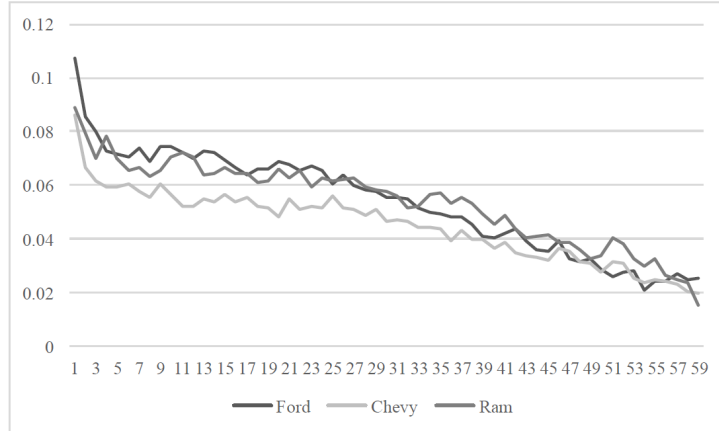
Figure 17: Moving Average Error Process Parameter Estimates by Brand

5   Discussion

Digital metrics such as online search volume can offer advertisers immediate feedback on how people respond to TV ads. We have found convincing evidence that some brands can now detect post-ad spikes from minute-by-minute brand search data, even following "ordinary" TV ad insertions with small to medium-sized audiences.

Empirical analyses in two disparate product categories produced many convergent findings. First, both studies found a strong positive relationship between ad audience size and search response size. Brand search spikes peak in the first minute after the ad insertion, last five minutes or less, and show no evidence of "stealing" search volume from following minutes, at least over immediate time horizons.

Second, TV ads consistently produced positive spillovers to competitor brand search. The spillovers are appreciable in size and asymmetric across brands. The pick-up truck analysis showed that Ford, the category leader, benefited the most from its rivals' advertising; and that it produced larger spillovers for each of its rivals than the third brand.

Third, contextual media factors such as TV network, program genre, and daypart play important roles in predicting variation in post-ad search spikes; but their effects are brand-specific, rather than category-specific. This is consistent with a model in which consumers' brand preferences or knowledge correlate with their TV program preferences. It suggests that brands need to investigate data generated by their own ad campaigns to uncover the most impactful media strategies; peer observations may be of limited value. This result

further suggests a positive equilibrium consequence for TV advertisers, in that it may reduce competitive clutter as brands increasingly make more differentiated advertising purchases.

Fourth, post-ad search spikes vary as a function of advertisement content. For fantasy sports brands, search spikes depend on ad creative and advertised products. In trucks, search spikes were nearly an order of magnitude smaller after price-focused commercials than after brand-focused advertising. Search spikes were smaller after more informative national ads than after less informative national ads.

The fact that post-ad spikes are reliably estimable from minute-by-minute brand search data is likely to be seen as good news by many TV advertisers, especially given the continuing calls for marketing accountability, along with recent literature highlighting the severe difficulty of estimating causal impacts of advertising on sales. However, the moderating effects of ad content on post-ad search spikes lead us to caution marketers against optimizing TV ads based solely on brand search spikes, except perhaps for campaigns explicitly intended to maximize consumer brand curiosity and information search. Advertisers seeking to optimize for other lower-funnel metrics (such as product knowledge, price awareness, purchase consideration or store visits) may be led astray by a TV ad campaign designed to maximize online search response to the exclusion of other metrics. In extreme cases, after controlling for audience sizes and media factors, larger post-ad search spikes might even be a negative indicator of TV ad performance, for example when the marketer's goal is to improve consumer knowledge of product attributes by delivering information via the TV ad itself.

For advertisers whose objectives reach beyond curiosity generation, we suggest considering multitaskers' online response holistically in the course of evaluating advertising tactics. For example, an ad that maximizes brand search response may be found inferior to a different ad that leads more consumers directly to the brand's website, thereby reducing cost-per-click expenditures. It also may be that a more informative ad is the best choice for moving prospects farther down the funnel, but yields a smaller post-ad search spike. The main guidance here would be to select metrics for campaign evaluation that reflect campaign objectives rather than expedience; and not to assume that all digital metrics are highly, or even positively, correlated. We hope the results of this paper can offer some guidelines for practitioners to think more deeply about the desirability of post-ad search spikes as a

measure of TV ad effectiveness.

One key advantage of using minute-by-minute brand search data is that it can be collected for multiple competing brands. That said, many brands could also benefit from analyzing more granular, second-by-second brand website traffic data to estimate advertising response by traffic source, such as direct navigation, social networking websites and search engine referrals. Following post-ad site visitors through to purchase or other types of transaction would be another critically important step forward from an attribution standpoint.

A deep understanding of digital response to TV ads opens up a number of emerging practices that warrant further research. The richer data available from server logs could be used to estimate different decay rates for different ad insertions. Data fusion technologies could be leveraged to link new site visitors after TV ads to offline actions such as in-store purchases, closing the offline-online-offline customer journey. Some marketers have attempted to adjust retargeting tactics based on new site visitors' time of arrival, for example, customizing retargeting content for post-ad visitors according to the television network on which the TV ad aired. Conversions accruing from retargeting efforts are then attributed back to the TV ad that produced the site visitor spike. Sophisticated offline-online attribution such as this requires the marketer to have a strategy that can identify cross-device users to avoid double-counting. For example, a TV ad might prompt a multitasking viewer to first search on a mobile device, then visit the brand website a few minutes later on a desktop or laptop to complete a purchase.

Although our paper is about how marketers may better leverage digital metrics such as online search in assessing TV ad effectiveness, we by no means advise marketers to abandon traditional measures of advertising response, such as offline sales, store traffic, or brand mindset metrics. Although post-ad brand search spikes offer a highly responsive metric for TV advertising attribution, these spikes are typically generated by less than one out of every 1,000 TV ad viewers. By contrast, although offline sales and brand mindset metrics resist estimation of causal ad effects, they are needed to generate a complete understanding of advertising response across all segments in the marketplace.

Finally, the dependable and sizable influence of TV ads on online brand searches cautions marketers against the use of simplistic "last-touch" attribution strategies, as they may

overestimate the effect of search engine marketing and social media advertising, and understate the generative influence of television advertising. Traditional and digital advertising budgets are still commonly divided between siloed agencies with little or no coordination among them. The between-media dependence observed in this study renews the call for holistic integration and evaluation of advertising campaigns and cross-media synergies (e.g., Naik and Peters 2009, Kim and Hanssens forthcoming).

To conclude, we see our research as part of a larger effort to understand how measurable funnel actions such as online search correspond to the reach, timing, placement and content of TV advertising and other promotional activities. We are confident that the drive for marketing accountability will continue forward and that multitaskers' immediate digital responses to traditional advertising will feature prominently as marketers continue to refine their understanding of how advertising affects the customer journey.

References

Anderson ET, Simester D (2013) Advertising in a competitive market: The role of product standards, customer learning, and switching costs. *Journal of Marketing Research* 50(4):489–504.

ComScore (2016) The new digital ecosystem: Navigating the cross-platform reality. Accessed January 2017, http://www.comscore.com/Insights/Presentations-and-Whitepapers/2016/The-New-Digital-Ecosystem-Navigating-the-Cross-Platform-Reality.

Draganska M, Hartmann WR, Stanglein G (2014) Internet versus television advertising: A brand-building comparison. *Journal of Marketing Research* 51(5):578–590.

Fossen BL, Schweidel DA (forthcoming) Television advertising and online word-of-mouth: An empirical investigation of social tv activity. *Marketing Science* .

Godes D, Mayzlin D (2004) Using online conversations to study word-of-mouth communication. *Marketing Science* 23(4):545–560.

Gong S, Zhang J, Zhao P, Jiang X (forthcoming) Tweeting as a marketing tool – field experiment in the tv industry. *Journal of Marketing Research* .

He C, Klein TJ (2016) Advertising as a reminder: Evidence from the dutch state lottery.

Hitwise (2016) Mobile search: Topics and themes Accessed January 2017, http://hitwise.connexity.com/rs/371-PLE-119/images/hitwise-mobile-search-report-us.pdf.

Hu Y, Du RY, Damangir S (2014) Decomposing the impact of advertising: Augmenting sales with online search data. *Journal of Marketing Research* 51(3):300–319.

Joo M, Wilbur KC, Cowgill B, Zhu Y (2014) Television advertising and online search. *Management Science* 60(1):56–73.

Joo M, Wilbur KC, Gauri D (2016) Temporal distance and price elasticity: Empirical investigation of the cruise industry. Mimeo, Ohio State University. http://ssrn.com/abstract=2630858.

Kannan PK, Li HA (2016) Digital marketing: A framework, review and research agenda. *International Journal of Research in Marketing* .

Kim H, Hanssens D (forthcoming) Advertising and word-of-mouth effects on pre-launch consumer interest and initial sales of experience products. *Journal of Interactive Marketing* .

Kitts B, Bardaro M, Au D, Lee A, Lee S, Borchardt J, Schwartz C, Sobieski J, Wadsworth-Drake J (2014) Can television advertising impact be measured on the web? Web spike response as a possible conversion tracking system for television. *Proceedings of the Eighth International Workshop on Data Mining for Online Advertising* 1–9.

Lamberton C, Stephen AT (2016) A thematic exploration of digital, social media and mobile marketing research's evolution from 2000 to 2015 and an agenda for future research. *Journal of Marketing* 80(6):146–172.

Lewis RA, Nguyen D (2015) Display advertising's competitive spillovers to consumer search. *Quantitative Marketing and Economics* 13(2):93–115.

Lewis RA, Rao JM (2015) The unfavorable economics of measuring the returns to advertising. *Quantitative Journal of Economics* 1941–1973.

Lewis RA, Rao JM, Reiley DH (2011) Here, there, and everywhere: Correlated online behaviors can lead to overestimates of the effects of advertising. *20th International World Wide WEb Conference (WWW20)* (April):157–166.

Lewis RA, Reiley DH (2013) Down-to-the-minute effects of super bowl advertising on online search behavior. *Proceedings of fourteenth ACM conference on Electronic commerce* 639–656.

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

Mayzlin D, Shin J (2011) Uninformative advertising as an invitation to search. *Marketing Science* 30(4):666–685.

Naik PA, Peters K (2009) A hierarchical marketing communications model of online and offline media synergies. *Journal of Interactive Marketing* 23:288–299.

Nielsen (2016) The nielsen comparable metrics report. Accessed January 2017, http://www.nielsen.com/content/dam/corporate/us/en/reports-downloads/2016-reports/q2-comparable-metrics-report-oct-2016.pdf.

Sahni N (2016) Advertising spillovers: Evidence from online field experiments and implications for returns on advertising. *Journal of Marketing Research* 53(4):459–478.

Sethuraman R, Tellis GJ, Briesch R (2011) How well does advertising work? generalizations from a meta-analysis of brand advertising elasticity. *Journal of Marketing Research* 48(3):457–471.

Shapiro B (forthcoming) Positive spillovers and free riding in advertising of prescription pharmaceuticals: The case of antidepressants. *Journal of Political Economy* .

Tirunillai S, Tellis GJ (forthcoming) Does offline tv advertising affect online chatter? quasi-experimental analysis using synthetic control. *Marketing Science* .

Video Advertising Bureau (VAB) (2016) National TV measurement. Accessed January 2017, http://www.thevab.com/national-tv-measurement.

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

Xu L, Wilbur KC, Siddarth S, Silva-Risso J (2014) Price advertising by manufacturers and dealers associations. *Management Science* 60(11):2816–2834.

Xu L, Wilbur KC, Silva-Risso J (forthcoming) Dynamic quality ladder model predictions in nonrandom holdout samples. *Management Science* .

Zigmond D, Stipp H (2010) Assessing a new advertising effect. *Journal of Advertising Research* 50(2):162–168.

Zigmond D, Stipp H (2011) Vision statement: Multitaskers may be advertisers' best audience. *Harvard Business Review* (January).