Does Offline TV Advertising Affect Online Chatter?
Quasi-Experimental Analysis using Synthetic Control

Seshadri Tirunillai and Gerard J. Tellis

This study befitted from a grant of Don Murray to the USC Marshall Center for Global Innovation.

Seshadri Tirunillai is Assistant Professor at the Bauer College of Business, University of Houston.
Gerard J. Tellis is Neely Chair of American Enterprise, Director of the Center for Global Innovation, and Professor of Marketing and Management & Organization, USC Marshall School of Business.
Does Offline TV Advertising Affect Online Chatter? Analysis of Quasi-Experimental Data Using Synthetic Control

Abstract
This study analyzes the impact of offline television advertising on multiple metrics of online chatter or User-Generated Content (UGC). The context is a quasi-experiment in which a focal brand undertakes a massive advertising campaign for a short period of time. The authors estimate multiple dimensions of chatter (popularity, negativity, visibility and virality) from numerous raw metrics using the content and the hyperlink structure of consumer reviews and blogs. The authors use the method of Synthetic Control to construct a counterfactual (synthetic) brand as a convex combination of the rivals during the pre-advertising period. The gap in the dimensions of chatter between the focal brand and the synthetic brand in the test versus advertising periods assesses the influence of advertising. Offline TV advertising causes a short but significant positive effect on online chatter. This effect is stronger on information-spread dimensions (visibility and virality) than on content-based dimensions (popularity and negativity). Importantly, advertising has a small short-term effect in decreasing negativity in online chatter.

Keywords: User-Generated Content, online chatter, synthetic control, TV advertising, difference-in-difference, quasi-experiments, matching, virality, offline advertising.

Introduction
User Generated Content (UGC) or (online) chatter has become a very important force in contemporary markets for several reasons. First, it has grown enormously in recent years. Second, surveys suggest that it is one of the most important sources of information that consumers trust (A.C. Nielsen Report 2013). Third, it can be easily measured by a number of metrics as several recent studies have shown (e.g., Peters et al 2013; Tirunillai and Tellis 2014; Schweidel and Moe 2014). Fourth, numerous studies have shown its impact on sales (Babic et al 2015) and financial performance (Tirunillai and Tellis 2012). Most important, online chatter is live, spontaneous, passionate, and available at a highly disaggregate temporal level of days, hours, or even minutes. Thus, it can be a powerful diagnostic of consumer sentiment.
This study focuses on chatter in two media that have a relatively high signal to noise ratio: reviews and blogs. Reviews contain chatter on consumers’ evaluation of products from their experience. Blogs are partially complementary to reviews as they reflect the diffusion of information across the web. Each of these sources is quite rich. We extract multiple metrics of chatter from each of them. Because these metrics might have some overlapping information, we use a dynamic factor analysis to derive four main dimensions of chatter: popularity, negativity, virality, and visibility.

With the growth of online advertising, many people assume that the days of offline advertising are numbered. Contrary to this assumption television (TV) advertising still commands about 40% of the advertising budget for firms in the U.S. (Strategic Analytics Report 2015). Other studies also support the dominance of TV advertising in the ad budget (e.g., Joo et al 2014). A big under-researched issue in recent times is whether offline TV advertising can affect online chatter. Offline TV advertising could affect online chatter by stimulating conversations, triggering brand recall, helping to interpret experiences, and refuting negatives (see section on theory). No studies have systematically assessed the effectiveness and dynamics of TV advertising on various metrics of online chatter (see section on literature for details). Such an assessment would have at least three managerial benefits. It would indicate whether and in what way offline TV advertising affects growing online chatter. It would allow firms to reach out to consumers at the point where they are most expressive and likely to influence each other. It would also provide a way to assess how TV advertising is working at a highly disaggregate temporal level, if not live. This is the focus of the current study.

This study aims to address this issue using observational data in a quasi-experimental context. We use the introduction of a big-budget brand TV advertising campaign of a focal brand, Hewlett Packard (HP), to assess if the campaign stimulates chatter. We monitor the results of this campaign daily for about three weeks after the start of the campaign. We monitor the advertising and chatter of HP together with that of control brands that are close rivals of HP in the major markets in which it competes. We link the advertising to multiple metrics of chatter that are based on a large-scale analysis of the content of the text and hyperlink structure of online chatter.
This study tries to assess the effectiveness of advertising on online chatter by the method of Synthetic Control (Abadie and Gardeazabal 2003; Abadie, Diamond, Hainmueller 2010). The essence of this approach is to create a counterfactual (synthetic) brand from all the relevant rivals of HP in the pre-advertising period and compare the gap between HP’s actual chatter and the chatter of the synthetic brand during the advertising period. Further, we analyze the dynamic effect of advertising on the gap in chatter using the Vector Auto-Regressive (VAR) model. In sum, this study seeks answers to the following questions:

- Does offline TV advertising campaign affect online chatter? If so, how strongly?
- Which metrics of online chatter, based on consumer reviews and blogs, are most influenced by TV advertising?
- What are the dynamics of the effects of advertising (“wearin” (buildup), “wearout” (decay), and duration) on the various metrics of chatter?

We find that offline TV advertising causes a short but significant positive effect on online chatter. This effect is stronger on information-spread dimensions (visibility and virality) than on content-based dimensions (popularity and negativity). Importantly, advertising has a small short-term effect in decreasing negativity in online chatter. The detailed discussion of the literature is in Online Appendix A.

The rest of the paper is organized in five sections: theory, research design, method, results, and discussion.

**Theory: How Offline Advertising Affects Online Chatter**

We test whether offline advertising can affect online chatter. Such chatter could subsequently affect sales, revenues and stock prices (Stephen and Galak 2012; Borah and Tellis 2016; Pauwels and Van Ewijk 2014; Tirunillai and Tellis 2012). Offline TV advertising could affect both the generation and propagation of online chatter for at least four reasons. First, advertising might stimulate conversations online (e.g., Srinivasan, Rutz, and Pauwels 2015). Second, advertising might trigger brand recall about user experience that is then reported in new online conversations as suggested in the impression formation literature (e.g., Higgins, Rholes, and Jones 1977). Third, advertising might enable consumers to interpret
information or prior online conversations more favorably to the brand (Hoch and Ha 1986). Fourth, advertising messages could possibly lend greater credibility and persuasiveness (O’Keefe 1999) and consequently refute certain negative conversations that are going on online. These paths are discussed in detail below.

**Stimulating Conversations**

Advertising stimulus can motivate the generation of content by starting a new discussion or invigorating the existing conversations. Advertising increases the awareness of a brand that could increase the interest and attention given to the brand by the customers (Tellis 2003; Srinivasan, Rutz, and Pauwels 2015). Studies suggest that consumers pay more attention to products and services in online media when exposed to advertising. This has been shown empirically in context of consumers searching for information on the web (e.g., Joo et al. 2014; Kireyev, Pauwels and Gupta 2015) where ads trigger or increase the volume of online search (in search engines). Hu, Du and Damangir (2014) further decompose the effect of advertising on online search into its effect impact on pre-purchase information search and on the conversion to final sales. We could extend the logic and attribute similar effects of advertising on generation of chatter due to increased consumer attention to the products. Advertising could draw customers’ attention and prime them to seek for more information about the product or the brand. This increased attention towards a brand leads to an increase in online conversations around the topics related to the brand or its products. In fact, increase in awareness among the influentials (Trusov et al 2009) could not only increase the content generation, but also accelerate the viral propagation of the content on the web. The advertising stimulus could also motivate loyal and passionate customers of the brand to talk about the brand and its usage in online conversations. Studies in self enhancement motives have shown that consumers share information about the brand to make themselves feel good and to look good among others in their social network (Toubia and Stephen 2013). Ultimately, such conversations can subsequently affect the volume of brand mentions and the online visibility of the advertised brand.
Triggering Recall

Secondly, advertising might trigger brand recall about user experience that is then reported in new online conversations. Mention of a brand name triggers recall of past experiences (Tellis 2003). Seeing the brand name in ads could also prime subjects about the brand and make brand recall easier. Specifically, studies suggest that external stimuli activate associated concepts stored in human memory and make them more accessible (Higgins, Rholes, Jones 1977), which in turn increases the chatter around the brand. Users might also be motivated to leverage the greater visibility of brand content due to advertising. Hence, users might create or share content based on users’ readership of others’ postings in social media (Toubia and Stephen 2013).

Interpreting Experience

Ads also aid in interpreting current or past experiences with brands especially when consumers face ambiguous evidence of product quality (Hoch and Ha 1986). Experimental studies have shown that ads could motivate consumers by reducing the perceived ambiguity of their experience and strengthening positive beliefs driven by objective evidence presented to them (Deighton and Schindler 1988). Further, ads could change the attitudes and beliefs associated with a brand leading consumers to share their opinions with others in order to cope with the feeling of dissonance (Festinger, Riechen and Schacter 1956). Ultimately, this effect of advertising could translate into lower negatives towards the advertised brands.

Refuting Negatives

Ads can help to negate bad word-of-mouth in online conversations (Tellis 2003). This process may work in two ways. First, refutational ads may directly negate negative arguments or negative evidence about the brand, possibly convincing some marginal consumers, who then may no longer participate in negative online conversations. Second, some of the refutational ads may be particularly persuasive to loyal customers (O’Keefe 1999), who then may refute those arguments and use that evidence in online conversations to
refute negatives by other participants. Ultimately, this effect of advertising could translate into lower negative towards the advertised brands.

In sum, offline advertising could affect online conversations through initiating a new thread, triggering positive recall about the brand, interpreting ambiguous experiences positively, or refuting ongoing negative conversations.

Research Design

This section describes the empirical setting, quasi-experimental design, metrics of chatter, and measures of advertising.

Empirical Setting

This study focuses on the advertising campaign, “Let’s Do Amazing” by Hewlett Packard that ran from March 13th through the end of May 2010. This campaign had a budget of about $40 million dollars. Its main goal was to increase awareness about the company and its technologies. The creatives revolve around the main message conveying the depth and breadth of the brand’s capabilities and services. The campaign portrayed celebrities such as Dr. Dre (music), Anne Leibowitz (photography) and firms such as UPS (logistics) and DreamWorks (movies) using the company’s technology in their daily operations. It did not promote specific product or service. Figure B1 (in Online Appendix B) outlines the daily duration of the ad insertions and the daily dollar amount spent during the campaign. Figure B2 specifies the day-part ad placements.

Quasi-Experimental Design

Figure 1 presents the quasi-experimental design. It shows the availability of data before and after the start of the TV ad campaign, which we treat as the intervention. We use the date of first observed

---

1 This period represents the period of high intensity and frequency of the advertisement. After this, there is very low intensity sparse advertising observed sporadically (many of them with a few ads spaced across many weeks).


3 Some of the videos associated with the campaign can be found at [https://www.youtube.com/watch?v=1kwz7TjsOU&index=3&list=PLXjvElQ9mLe0x-1okKDssJIsmvJv4SYup](https://www.youtube.com/watch?v=1kwz7TjsOU&index=3&list=PLXjvElQ9mLe0x-1okKDssJIsmvJv4SYup)
instance of the TV ads of the campaign, March 13 2010, as the start of TV advertising. We use the first 20 days, after this date, as the intervention (or advertising) period. The section titled “Identification of Assumptions” discusses the assumptions of the quasi-experimental design and provides extensive evidence to support these assumptions. We use Difference-in-Difference, Synthetic Control, and Vector-Auto-Regression (VAR) to assess how HP TV advertising affects the chatter of HP over and above the chatter of rival brands.

**Metrics of Chatter**

Chatter can be characterized by a variety of metrics (see Peters et al 2013 for a detailed discussion on various social media metrics). We restrict our analysis to metrics that are relatively easy to compute and provide practicing managers with immediate insight. We derive our metrics from a set of raw online measures obtained from product reviews and blogs on the brands or their product. These raw metrics are typically accessible to managers. We then extract dimensions that represent the commonality across these metrics using dynamic factor analysis. The details of these metrics are below. We classify our metrics as either content-based that focus on the content or characteristics of chatter, or viral-based that focus on the propagation (information-spread) of the brand chatter.

**Raw Metrics from Reviews**

Research in marketing has used various content-based metrics such as volume and valence from product reviews (e.g. Chevalier and Mayzlin 2006; Gopinath et al 2014,) and blogs (e.g. Gopinath, Chintagunta and Venkataraman 2013, Onishi and Manchanda 2012). In addition to these, we use polarity, a measure of sentiment divergence. We collect the reviews from consumer reviews and ratings on amazon.com, epinions.com and cnet.com, as we are interested in assessing the impact of TV advertising on product evaluations in review. Alternate social media such as Twitter, YouTube or Facebook had restrictions in terms of availability and drawbacks in measurement of consumer evaluations (due to lack of data) of the brand during the time period under consideration.
**Volume:** Volume refers to the total number of new reviews of a brand generated in a given time period. This measure reflects the intensity of coverage about brands in chatter. The volume of reviews is based on the number of the reviews of the products of the brand.

**Valence:** Valence refers to whether the overall review is positive or negative reflecting prevailing sentiment. We derive the valence of a review by analyzing the textual contents in the reviews. First, we classify the review as positive (or negative) using popular unsupervised machine learning techniques for textual classification of the sentiments - Support Vector Machine, Naïve Bayesian Classification following the prior literature (e.g., Tirunillai and Tellis 2012). Based on the agreement of the results of these classification algorithms, we choose the overall tone or valence of the textual content of a given review as positive or negative. We refer to the count of the volume of the positively or negatively classified reviews for a brand in a given time period as either positive or negative valence.

**Polarity:** Polarity (Entropy) of reviews measures the dispersion of the ratings across a brand in a given day reflecting the diversity of opinion. If the probability of a rating (measured on a five-star scale) is given by \( P \), measured as the relative frequency at which a given rating level \( (l) \) occurs among the reviews, then the entropy for the brand \( k \) in a given day could be calculated using the Shannon Entropy index (Shannon 1948):

\[
Entropy(k) = - \sum_{l=0}^{5} P(\eta = l) \log_2 P(\eta = l)
\]  

As can be seen, the polarity (entropy) increases with the increase in dispersion of the ratings.

**Raw Metrics from Blogs**

Blogs (or Weblogs) are a popular avenue to share opinions and ideas, pursue open conversations, and discuss topics on similar interests to a wide audience. Blogs were once popular as online journals. Later they evolved to become a popular form of social media for individuals and companies to express opinions, provide commentaries, and document ideas (Aggarwal et al 2012). Research in marketing highlights the importance of blogs as a word of mouth medium (e.g., Mayzlin and Yoganarasimhan 2012). Studies in marketing show that metric of blogs (e.g., volume) have an impact on sales (e.g., Onishi and Manchanda 2012). Apart from the volume base metric, we also exploit the hyperlink structure in the
blogs to derive various metrics for diffusion of information about the brand across the web. Citation structures across the web pages using hyperlinks are one of the core metrics in search algorithms (e.g., Klienberg 1999). Given the richness of the citations in blogs, they have been used to understand the diffusion of information over the web (e.g., Leskovec, Backstrom, and Klienberg 2009). Empirical research in marketing using the link structure as a network metric of chatter is still nascent. Few studies in marketing (e.g., Katona, Zubcsek and Sarvary 2011, Trusov et al 2009; Roos, Mela and Shachar 2013) delve into these metrics, though exploiting the network structure of the citations across the Internet has been proven to be a very important metric in information science (e.g., Leskovec, Backstrom, and Klienberg 2009). We extract four metrics of blogs, volume, in-degree of brand website, in-degree of blogs, and volume of blogs that gain rank.

**Volume of Blogs:** We use volume of blogs similar to the volume of reviews similar to prior literature (e.g., Onishi and Manchanda 2012; Gopinath, Chintagunta and Venkataraman 2013). We measure volume of the blog posts based either on the tags associated with them, tags with any words related to the products of the brand, or on presence of the mention of the brand (or products) in the title.

**In-degree (links) of the Brand Website:** We use the hyperlinks structure in the blogs to derive metrics for information propagation across the web (blogs). Studies have researched the importance of message transmission across the Internet and its impact on purchase (e.g., Baker, Donthu and Kumar 2016). Using the graphical network structure of hyperlink among the blogs, we measure the number of hyperlinks pointing to a Uniform Resource Locators (URLs) containing the home domain of the brand (referred as in-degree centrality) in the given time period. This measure could be viewed as the citations received from across the web for the brand. More cited posts tend to have greater influence than less cited ones. Because the hyperlinks are sticky, we observe the cumulative number of hyperlinks in any given time period. To account for only the new citations in any given time period, we calculate the change in the in-degree centrality of the brand’s primary domain in the blog network in a given time period. Because our focus is on measuring the transmission of information regarding a specific brand, we focus...
only on in-degree of the home domain site of the brand of interest within our blog data. Details on the measurement is in Online Appendix C.

**In-degree (links) of blog posts:** As each of the blog posts about the brand accumulates more readership, it may not only transmit the information about the brand to its readers but the given post will also be linked back (cited) by other blogs, websites, or social media depending on the value of the content in the post. Thus, increase in the readership of the blog posts helps in spreading the brand’s presence and influence across the web. Spinn3r records the in-degree\(^4\) of the URLs of the blogs it tracks, which enables us to calculate the in-degree of blog posts of the brand. Since the in-degrees of links are cumulative, we calculate the number of new links gained in the given time period as the change in the aggregate in-degree in the top tier blogs posts about the brand during the consecutive time periods.

**Volume of blogs that gain/lose rank:** The in-degree measure of the focal brand’s domain ignores characteristics of the web pages and the citations that could drive the propagation of information across the Internet. For example, it ignores the popularity or authority of the citing web page (e.g., New York Times vs. Denver post). To control for these variables, we rely on the ranking of the blog relative to all other pages in a given time period. The blog rank captures this information that is missing in the centrality measures. Each blog post tracked by Spinn3r has a rank associated with it relative to the other blogs in the blog index. This rank could be likened to the rank of a web page in a search engine’s index. Thus, the ranking of a given blog URL measures the influence of the specific URL relative to the other blogs in the time period. The ranking of the URLs by Spinn3r is measured through their custom algorithm.\(^5\) The algorithm takes into account various characteristics of the blog while assessing the influence and assigning it a rank. Because Spinn3r indexes most of the blogs across the blogosphere, the ranking of the URLs within the system of Spinn3r can be considered universal and reliable. We calculate

\(^4\) http://spinn3r.com

\(^5\) The ranking is based on the “authority” of the node (URL). Crudely, this can be viewed as analogous to the Page Rank or HITS algorithm. In fact, Spinn3r (http://spinn3r.com) used the Blog Index for creating the once popular blog ranking and trending site “tailrank.com”. We restrict our Internet universe to the blogs (also referred to as blogosphere). Also, we focus only on the spread of messages surrounding the brand and its products.
the count of posts about the brand that underwent a change in the ranking in the top-tier blog posts from the pre-sampling time period.

**Dynamic Factor Analysis**

These raw metrics could have overlapping information. To eliminate collinearity and capture common dimensions underlying these raw metrics, we use dynamic factor analysis (e.g. Du and Kamakura 2012; Stock and Watson 2011). We assume that a few factors ($f_t$) underlie the observed raw observations of the different measures ($y_t$) of reviews (volume of reviews, valence, polarity) and blogs (volume of blogs, in-degree of the website, in-degree of the blog posts, volume of blogs that gain rank). We assume that these measures follow the following model:

\[
Y_t = \xi f_t + \epsilon_t
\]

\[
f_t = \Psi f_{t-1} + \eta_t
\]

$\xi$ represents the vector of factor loadings, the $\epsilon$ is the idiosyncratic error which are assumed to be uncorrelated, and $\eta$ is the white noise with $E(\epsilon_t\eta_{t-k}) = 0$. Table 1 summarizes the mean factor loadings of the metrics across the brands. The results of the dynamic factor analysis suggest four dimensions underlie these metrics. We label these dimensions, popularity, negativity, virality, and visibility, based on their loadings on the raw metrics. The first dimension, *popularity*, loads mainly on the volume of reviews and blogs. It measures the popularity and importance of the brand. This factor may relate to the role of seeding conversations that we discussed in the theory. The second dimension, *negativity*, loads positively on negative valence and polarity and negatively on positive valence. The roles of interpreting experiences and especially refuting negatives, mentioned in our theory, relate to reducing prevalence of this factor. The third dimension, *virality*, loads on the volume of blogs that gained rank and the in-degree of the blogs. It reflects the relative rate of spread of information about a brand across the Internet. The more such blog posts transmit the message of the brand, the more the reach of the brand among consumers. Advertising could stimulate the increased sharing of the content across the web. It could also increase the recall of brands thus enabling the spreading of the conversations, consequently increasing virality. The
fourth dimension, *visibility*, loads on the volume of blogs and the in-degree links of the brand website. This dimension captures the degree of visibility of the brands across the Internet and thus the successful transmission of information through the blog posts. As advertising increases the customer attention to some brands, it could result in customers linking the brand’s website or the web pages (e.g. products pages) on the website in their own blog or web pages. This could increase the visibility of the brand’s focal website. We refer to the first two dimensions as content-based dimensions, as they load on the underlying measures that are related either to the volume or the content of the measures of chatter. Similarly, we refer to the last two dimensions as viral-based due to their focus on the brand-related information spreading across the web. We use the estimated value on each of these four dimensions across all the brands as the dependent variables that could be affected by advertising. In the subsequent discussion, we refer to these factors as dimensions of chatter.

**Measures of Advertising**

We measure the advertising intensity as the number of insertions of the ads placed in television on any given day (also referred to as “placements’ or ‘TV spots’). We focus on television ads as the campaign targeted mainly in form of 30-second ad spots with interviews of various celebrities in their work environment using HP products. To check the robustness of the results, we also use alternate measures of intensity such as the daily advertising expenditure (in dollars) and the total duration of ads broadcast during a given time period (in seconds per day). We also account for the influence of the time-of-the day by including the day-part measures of advertising in our analysis.

Apart from the TV advertising, we also collect information on the online advertising and the ongoing product advertisements during the time period under analysis to explicitly account for possible confounds. We get the details of the advertising measures from Kantar Media (Stradegy).

---

6 The Robustness analysis indicates that these alternate metrics do not change the direction of the results.
Method

This section presents the method of Synthetic Control in terms of motivation, intuition, specification, identification, sampling of firms, statistical inference, choice of methods, and assumptions and limitations.

Motivation

The ideal approach to assess the effectiveness of offline advertising on online chatter is through field experiments or Randomized Control Trials (RCTs). However, these are challenging in most field situations for several reasons. First, conducting RCTs are immensely resource intensive and expensive. Second, designing and implementing an experiment with a real company’s advertising, on a large scale, such as across the whole USA, is challenging. Third, execution of RCTs over an extended period of time, such as several months or a year, may be impractical in certain situations. Fourth, even if implemented, numerous changes among rival brands, consumers, and the focal brand during the test period might undermine the assumption of exogeneity of the treatment (advertising). In most cases we have access to observational data. Our aim is to estimate the effect of offline advertising on metrics of online chatter using observational data and overcome some of the endogeneity concerns (Luan and Sudhir 2010; Tellis 2003). We exploit a quasi-experimental situation, when a focal brand, HP, runs a big advertising campaign for a limited time, using Synthetic Control.

Intuition

The logic for Synthetic Control is to compare the outcomes (chatter) of the focal brand (HP) undergoing the intervention (TV advertising) with the chatter of a counterfactual or synthetic brand that is similar to the focal brand but does not undergo the intervention (Abadie, Diamond, and Hainmueller 2014). This is similar to some of the other matching based program evaluations method (Imbens and Woolridge 2009). We create the counterfactual brand (“synthetic control”) from a convex combination of rival brands in the same industry that closely resemble HP during the pre-advertising period. To do so, we
estimate weights for each rival brand such that the synthetic brand best approximates the actual brand HP on key characteristics (including chatter) during the pre-intervention period. We then compare the chatter of the focal brand to the weighted chatter of the synthetic control during the intervention (advertising) period. We estimate the effectiveness of the advertising campaign as the Gap between predicted chatter of the synthetic control and the actual chatter of HP during the advertising period. If the Gap is significantly different from 0 during the intervention period, it implies that advertising affects the chatter of the focal brand more than it affects the chatter of rival brands.

**Specification**

For simplicity, we follow the notations of Abadie, Diamond and Hainmueller (2014). Let there be J+1 brands under observation during the sample periods, \( t \in [1, \ldots, T] \). Let \( j=1 \) represent the focal brand that is treated (undertakes the advertising campaign) and the remaining brands \( j \in [2, \ldots, J + 1] \), the potential control units (also referred to as the “donor pool” in the matching methods literature). Let the focal brand be exposed to the intervention (here, the start of the advertising campaign) from period \( T_0 \). Thus, \( 1 \leq t \leq T_0 \) constitute the pre-intervention period and \( T_0 \leq t < T \) the intervention period (refer Figure 1 for the design). Let \( Y_{it} \) represent the outcome variable (here a dimension of chatter) for brand \( i \) at time \( t \). Let \( Y_{it}^N \) represent the outcomes of the brands in the absence of the intervention and \( Y_{it}^I \), represent the outcomes of the focal brand, that had the intervention. The net effect of intervention at any given time period is the Gap, \( \alpha_{it} \), as the difference between the treated brand and the counterfactual or synthetic brand, which is

\[
\alpha_{it} = Y_{it}^I - Y_{it}^N \quad \forall \ t \in [1, T] 
\]  

(3)

During the pre-intervention period, there is no treatment, \( Y_{it}^I = Y_{it}^N \). Hence there should be no gap and \( \alpha_{it} = 0 \). During the intervention period, we can explicitly estimate the gap, \( Y_{it}^N \), of the treated brand, had it not undergone the intervention. Thus, the gap in the treatment period can be represented as:

\[
\alpha_{it} = Y_{it}^I - Y_{it}^N \quad \forall \ t \in [T_0, T] 
\]  

(4)
While we observe the outcome of the focal brand $Y_{1t}$, the outcome of the synthetic brand $Y_{it}^N$ is not observed and has to be estimated. $Y_{it}^N$ can be written using as a factor model given by:

$$Y_{it}^N = \delta_t + \theta_t Z_i + \lambda_t \mu_i + \epsilon_{it} \tag{5}$$

Where $\delta_t$ represents the unknown common factor with constant factor loadings across all the units, $Z_i$ represents the $(r \times 1)$ vector of observed covariates not influenced by the intervention, $\theta_t$ represents the $(1 \times r)$ vector of unknown parameters of $Z_i$, $\mu_i$ vector $(F \times 1)$ captures the factor loadings of $\lambda_t$, the $(1 \times F)$ vector of unobserved common factors and $\epsilon_{it}$ is the error that is assumed to be distributed with a zero mean. The unobserved factors are captured through the inclusion of the outcomes in the pre-intervention period. The covariates $(Z_i)$ could be time-invariant (during the sample period) or time-varying (with varying frequencies). This specification enables us account for any time-varying unobservable heterogeneity among the brands during the period. In fact, Equation 3 reduces to fixed effects difference-in-difference if we impose the restriction that $\lambda_t$ is constant for all time periods. We use brand (or firm) characteristics that are both time-variant measures (vary at a high frequency here, daily) and time-invariant measures that are measured at a relatively lower frequency (quarterly or annual). These variables also include regular brand advertising that takes place outside of the focal ad campaign. We use multiple measures to best match the Synthetic Control with the focal brand. Details of variables are in the Sampling section.

Abadie, Diamond and Hainmuller (2010) show that the treated brand’s outcome during intervention can be calculated using a convex combination of the untreated units, the synthetic control, under the conditions discussed below. Let $W = (w_1, w_2, ..., w_i, ..., w_i)$ be the weight vector ($w_i \geq 0$ and $\Sigma w_i = 1$) that defines the weights of the unit $j$. Using these weights, the synthetic control estimator can be written as:

$$\alpha_{it} = Y_{1t} - \Sigma_{j=2}^{l+1} w_j Y_{jt} \tag{6}$$

Though there could be numerous possible combinations of the weights, we are interested in finding the optimal weights of the untreated units that best approximate the treated unit. We choose the
optimal weights ($W^*$) for the brands such that it minimize the difference between the pre-intervention characteristics of the treated unit and the synthetic control, $X_1 - X_0 W$ during the pre-intervention period, where $X_1$ represents the vector of pre-intervention characteristics and may include the outcomes, i.e. $X_1 = (Z'_1, \bar{Y}_{1,1}^{K_1}, ..., \bar{Y}_{1,1}^{K_M})'$ and $X_0$ be the corresponding matrix of values of the variables of the J donor pool units for the same characteristics.. If $X_{1m}$ is the vector of the pre-intervention characteristics of the treated unit on a given variable $m$ and $X_{0m}$ the vector of values of the J donor pool units for the corresponding characteristic, we obtain $W^*$ by minimizing

$$
\sum_{m=1}^{k} u_m (X_{1m} - X_{0m} W)^2
$$

(7)

where $u_m$ is a weight that represents the relative importance of the $m^{th}$ variable in measuring the discrepancy between $X_{1m}$ and $X_{0m} W$ (Abadie, Diamond, Hainmueller 2014). The variables with large predictive power on the outcome would be assigned larger weights in the calculation of $u_m$. In practice, the choice of $u_m$ is derived from the data by choosing the weights $u_m$ to closely approximate the trajectory of the treated unit in the pre-intervention period such that the Mean Square Prediction Error (RMSPE) of the outcome variable is minimized during the pre-intervention periods. Details of the implementation are in Abadie, Diamond Hainmueller (2010, 2011, and 2014).

We need to assess if the trajectory of the gaps after the intervention can be attributed to the advertising undertaken and not due to random chance or due to factors other than advertising. The conventional standard error calculation commonly used in regression is based on large sample inferential statistics and is not well suited for our aggregate data with a few brands in the control group. Hence, we use “placebo tests” that check the robustness of the method and assess the matching controls. The main aim of the placebo test is to investigate if the observed effect for the brand undertaking the advertising is large relative to the effect estimated for a brand (that did not undertake the advertising) chosen at random. The placebo test is similar to the permutation procedures employed in propensity matching technique. We

---

7 The software can be found at [http://web.stanford.edu/~jhain/synthpage.html](http://web.stanford.edu/~jhain/synthpage.html).
trace the steps used in identification of the gaps for each of the brands in our sample. The goodness of results can be assessed using the calculated pre- and post-treatment Root Mean Square Prediction Error (RMSPE) for all the brands. The distribution of the post-intervention MSPE to the pre-intervention MSPE ratio (post/pre MSPE ratio) indicates the probability of observing the outcome by chance. If the observed post/pre MSPE ratio of the focal brand compared to all the placebo brands is large, we can conclude that the observed gap in the post-intervention period is not by chance, and safely reject the null hypothesis that advertising does not cause the observed change in chatter. The details of these tests are given in online Appendix H.

**Sampling**

To select the donor pool of brands for the construction of the synthetic control, we filter firms by applying the criteria outlined in the Online Appendix D. We have the following firms constituting the donor pool of rivals for the synthetic control – Apple Inc., Dell Inc., Canon Inc., EMC Corp, Seagate Technology, Western Digital Corp, Sandisk Corp, Lexmark Intl Inc., Netapp Inc., and Logitech International. For the time sampling, we choose 70 days of data during the pre-intervention period to construct the synthetic brand that is the closest representation of the focal brand (HP). For the intervention period, we choose the first twenty days of advertising for presenting the results. We do so in order to have a clean intervention period without any other big advertising campaigns.

We choose the following variables as our predictor variables for constructing the synthetic control. 1) Firm Size (Total Assets, Quarterly); 2) Market Capitalization (Daily); 3) Debt/ Equity Ratio (Daily); 4) Compounded Annual Growth Rate in Revenue (Quarterly); 5) Media coverage (Daily); 6) Marketing (XSGAQ, Quarterly); 7) Advertising spending (XAD, Annual); 8) Online advertising (Monthly) 9) Total Revenue (Quarterly); 10) Total Number of Employees (Annual).

**Identifying Assumptions of Synthetic Control**

The validity and strength of Synthetic Control depends on the extent to which following assumptions are met.
1. One assumption is that the donor brands are a good match for the Synthetic Control (Abadie, Diamond Hainmueller 2010). This assumption can be ascertained by comparing the gap in chatter between the Synthetic Control and HP in the pre-treatment period. Ideally this gap should be 0. Figure H1 (discussed in details in robustness check, Online Appendix H) shows that this gap is close to 0 for the pre-treatment period.

2. A second assumption is that only the focal brand undergoes the treatment during the intervention period and that the donor brands contributing to the synthetic brand do not undergo similar confounding events. To verify these assumptions, we collected TV advertising data across all the brands in the analysis from Kantar Media. Figure 2 shows that the donor brands have no major TV brand advertising during the treatment period, either due to their own plans or in reaction to HP. In this figure, March 13, 2010 is the intervention indicated by the start of the HP TV ad campaign. As can be inferred from the figure, the spending of HP after the start of the brand campaign is significantly larger than that of the other brands. The spending by Apple on March 7th in the pre-treatment period is for commercials for the forthcoming launch of iPad. It represents an aberration that is taken into account when computing the Synthetic Control. Figure 3 shows that the average advertising spending for TV by the control brands versus the focal brands before and during the intervention period. This graph depicts all product and brand advertising on TV across the top spending brands. We include only the top brands, as the other brands do not have major TV advertising during the time period compared to these. As can be inferred from the figure, this change in advertising is highest for the focal brand.

3. A third assumption is that no major changes in advertising on other channels (radio, print, and online-display and search) occur for HP and other brands during the treatment period. We collected advertising data across the various media channels (TV, print, radio, Internet-display and search) from Kantar Media. We use weekly level data for comparison, as that was highest frequency that we could obtain to compare across the different media. Figure 4 shows the advertising of HP during the treatment period. We can see that the major change in advertising during the treatment period is in HPs TV advertising. A smaller increase occurs for online advertising. We control for this difference by including
this variable as a covariate in the VAR analysis. The advertising spending for TV versus other media of HP is in Figure 5. It also shows the average spending of other brands on non-TV advertising. The change in advertising is highest for the HP’s TV. The top spending control brands do not substantially increase advertising in non-TV channels during the intervention period.

4. A fourth assumption is that HP undergoes no other changes in the other marketing variables during the intervention period, such as changes in price, channels, products, and promotions. We searched for the following terms on Lexis Nexis and Factiva for announcements in US newspaper regarding major changes in HP marketing between Jan 2010 and April 3, 2010 in US. We use the base term for the company (e.g. “HP” or “Hewlett Packard”) with marketing related terms such as “product release,” “retailing,” “distribution,” “promotion,” “new product,” “innovation,” “marketing,” “strategy,” “channel.” Such searches of announcements and press reports are also used by other recent articles in *Marketing Science* (e.g., Saboo et al 2016; Borah and Tellis 2014). The detailed table of the most important news items retrieved through an extensive search of Factiva and Lexis Nexis news databases during the advertising period is in Online Appendix E. With one exception, these searches did not reveal any major changes during the intervention period in channels, products, pricing, or promotion. During the intervention period, the only major change in marketing reported was the uploading of the TV campaign on HP’s YouTube channel. YouTube has recently become a platform where advertisers upload TV ads for additional exposure or in the hope of the ad going viral. Nevertheless, we compare the estimated views of offline TV versus online YouTube. We got the YouTube data for the time period from Visible Measures8. We got the TV viewership from syndicated reports (e.g., TV guide websites9 and Kantar Media). Details of the estimation are in Online Appendix F. The results of the viewership are in Figure 6. As can be seen from the figure, during the intervention period, views from offline TV advertising dominate views from

---

8 http://www.visiblemeasures.com
9 http://tvbythenumbers.zap2it.com/2010/
online YouTube video advertising. There was a webpage\textsuperscript{10} on HP’s website that linked to the videos of the campaign. As shown previously the views YouTube videos are much lower than the predicted views of offline TV.

Quasi-experimental designs are not perfect and may have some shortcomings. For example, Abadie and et al (2010, p 501) acknowledge four weaknesses in their design.\textsuperscript{11} Thus, discussion of the above assumptions indicates that this quasi-experiment is reasonable although not perfect.

**Analysis of the Dynamics**

Though the above method of Synthetic Control helps us in assessing the impact of advertising on chatter metrics, yet it does not quantify the dynamic effect of advertising campaign after it is launched. Theory and the literature suggest that that advertising has a dynamic effect on response, exhibiting wear-in (time for the effect of advertising to peak up) and wear-out (time for the effect of advertising to decay). In order to assess the dynamics of the effect of advertising on the different chatter metrics, we adopt the Vector Autoregressive (VAR) framework. The VAR model helps in capturing the dynamics in terms of short run and long run response and the wear-in and wear-out of advertising on response. We use a bivariate VAR with the advertising and the gap estimated for the chatter dimension using the Synthetic Control as the endogenous variables. We use the media citations and online advertising as the exogenous variable in the model. The estimation of the gap controlled for many of the other factors that could influence the dimension of the chatter, thus providing us the advantage of keeping the VAR model parsimonious. We estimate the model for each of the dimensions. The details of the implementation of the VAR model is in Online Appendix G.

\textsuperscript{10}This can be assessed through the Internet Archive Project at https://web.archive.org/web/20100314151549/http://www.hp.com/united-states/do-amazing/index.html?#/Intro

\textsuperscript{11}1) Increase in anti-tobacco sentiment created in California could have spread to other states. 2) Tobacco companies could have diverted planned advertising in other states to California. 3) Rise in tobacco taxes in California increased cross-border smuggling. 4) Rise in tobacco taxes in California increased cross-border purchases.
Results

This section covers the results of difference in difference, estimation of synthetic control and analysis of dynamics. The robustness checks and associated analysis is in Online Appendix H.

Assessing Effect of Advertising using Difference in Difference

Difference in Difference (DiD) is commonly adopted panel model to examine the effects of interventions in marketing (e.g., Liaukonyte et al 2015; Chevalier and Mayzlin 2006; Goldfarb and Tucker 2011; Chiou and Tucker 2012). It is commonly used panel data technique to assess the impact of economic programs or interventions (Imbens and Woolridge 2009). We specify the following fixed-effect least square regression to assess the impact on each of the outcomes of the chatter metric.

\[ Y_{it} = \tau D_{it} + \pi_t + \gamma_i + \epsilon_{it} \] (8)

Here, \( Y_{it} \) is a metric of chatter at time \( t \), \( D_{it} \) is a dummy for the advertising which takes a value of 1 after the start of advertising for the focal firm and zero otherwise, \( \pi_t \) is the common time effect, \( \gamma_i \) is the brand fixed effect, and \( \epsilon_{it} \) is the unobservable random error term. The parameter of interest is \( \tau \) that captures the impact of advertising. The standard errors are robust standard errors clustered by brands.

Table 2 depicts the coefficient of the effect as reflected in the advertising dummy and the associated robust standard error in parentheses along with the adjusted R-square. As can be inferred from the table, the campaign has a significant positive impact on the popularity, virality and the visibility dimensions of chatter with varying intensity. The advertising campaign has a negative effect on Negativity. Difference in Difference assumes that the difference between the advertising and the control brands stay constant over the course of advertising.

Estimating Synthetic Control

We construct the Synthetic Control for HP estimating weights for control brands following the method outlined above for the period from 1\(^{st}\) January 2010 through March 12\(^{th}\) 2010. The daily level data constitute 70 calendar days during the pre-intervention period. The estimation of weights follows the
optimization procedure outlined above. We illustrate the results with the dimension of popularity in detail and then summarize the results for the other three dimensions of chatter.

The weights estimated for the synthetic control resulted in Dell, Apple and Canon having the highest non-zero weights (0.37, 0.12 and 0.51 respectively) among all the donor brands. The other brands in the donor pool receive relatively very low (near zero) weights and hence were ignored. The estimated weights seem reasonable based on our prior knowledge of similarity of rival brands to HP. These brands also have similarities in terms of size, customer segments, marketing investments, and product categories. Dell is a primary rival of HP and has lots of products competing in similar consumer and business segments. Apple is another rival in the consumer market especially that of personal computing devices. Canon and HP compete in similar business-to-business and consumer markets during the time period under investigation.

The summary of the predictor variables used in construction of the synthetic control brand is in Table 3. It also shows the comparison of HP and the synthetic control on the different predictor variables. The synthetic control is similar to HP in many of the underlying characteristics over the pre-intervention period.

**Effect of Advertising on Competition Brands**

A fundamental concern in treatment-control approaches is whether the advertising of HP affects the chatter of the rival brands. We can ascertain this effect, if any, in two ways.

a) We test the cross elasticity of HP advertising on the dimensions of chatter of HP and rival brands during the intervention period. Table 4 presents these elasticities. The elasticities of HP advertising on dimensions of HP chatter are significantly different from 0. On the other hand, the elasticities of HP advertising on dimensions of rival brands’ chatter are not different from 0. This is a sufficient condition for the use of Synthetic Control.

b) We test if HP advertising has any effect on dimensions of HP chapter beyond that on the chatter of rival brands. To do so, we examine the gap or difference between HP actual chatter and that of
the Synthetic Control chatter during the intervention period. The fundamental approach of Synthetic Control is to control for the chatter of rival brands during the intervention period. Indeed, the Synthetic Control is the weighted average of the chatter of rival brands during the intervention period (where the weights are estimated in the pre-intervention period). Figure H1 (in Online Appendix H) shows that this gap is significantly different from 0 for each dimension of chatter post-intervention. The placebo tests also confirm this finding. Thus, HP advertising triggers HP chatter above that for rival brands. This is the necessary condition for successful use of Synthetic control.

Effect of Advertising on the Dimensions of Chatter

We consider the impact of advertising separately for the content-based dimensions (popularity and negativity) and the information-based dimensions (virality and visibility) of chatter.

Impact of Advertising on the Content-Based Dimensions

Figure 7 depicts the trajectory of evolution of the popularity of HP (red) against the popularity of the Synthetic Control (blue) during the sample time horizon, which includes the pre-advertising and advertising period. The figure also depicts the intensity of the campaign in the lower panel. The popularity of HP and the Synthetic Control brand trace a similar path during the pre-advertising time period. This suggests that the optimization algorithm was able to converge on weights for the Synthetic Control brand such that the resulting Synthetic brand closely resembles HP during the pre-advertising period. However, during the advertising period, a marked divergence emerges in the trajectory of popularity between the Synthetic Control and HP. This result suggests that HP’s advertising stimulated chatter about HP during the advertising period reflecting its increasing popularity. The gap in popularity goes up by about 13% on an average for the days immediately following the launch of the advertising. The gap reaches a peak of about 19% and then tapers down a couple of days after the start of the campaign. Overall, volume of chatter of HP relative to that of the Synthetic Control increases about 15% during this time period. Thus, based on the existence of a marked gap between the trajectory of HP and
the Synthetic Control, we can infer that the advertising has a definite impact on popularity. Since the popularity is based on creation of new content, we can infer that advertising stimulated conversations around the brand, which is in line with the theory. Though the gap is prominent during the start of the advertising period, yet it tapers out as the advertising campaign intensity diminishes suggesting that the effect of advertising on the popularity of chatter is only transitory and not permanent.

Table 5 shows the estimated effect during the advertising period, the pre-advertising RMSPE (Root Mean Square Prediction Error), and the advertising period to pre-advertising period MPSE ratio. The effect size of popularity shows a mean lift size of about 5.36 posts during the advertising period. We infer the significance of the effect using the RMSPE ratio and the associated exact p-value. The details of these calculations are in Online Appendix H. Following the conventional standard (Abadie et al 2010) we compare the gap of the focal brand, HP, with that of the placebo brands using the post advertising to pre-advertising ratio of RMSPE. The ratio of the advertising RMSPE to pre-advertising RMSPE is highest for HP when compared to all the placebo brands suggesting that the effect of advertising is strong and significant for HP.

We repeat the analysis for negativity, on an average across the brands, the factor that loads positively on negative valence and polarity and negatively on positive valence (Table 1). Advertising seems to negatively impact the negativity dimension. This result suggests that advertising increases positive chatter and decreases negative chatter. As discussed in the theory section, this could be attributed to the change in the customer expectations due to advertising, resulting in them interpreting their experience with the product favorably for the advertised brand, thus generating positive conversations and refuting negative conversation. Figure 7, panel 3, depicts the impact of advertising on the negativity of reviews over time. The impact of advertising immediately after the start of advertising is not discernible. However, after a few days, the effect increases and then wanes as seen by the slow increase in gap of the negativity dimension over the period. We compare the size of the effects between the metrics in terms of elasticity using an advertising response (VAR) model that is discussed below.

**Impact of Advertising on the Information-Spread Dimensions**
We repeat the analysis for the information-spread dimensions – virality and visibility - for all the brands in the sample during the sample time period. These two dimensions are primarily driven by the hyperlink structure in the blogs as can be seen in Table 1. Virality is driven by the in-degree citations to the blog posts and the volume of the blog posts that gain rank in the time period. Visibility is driven by the in-degree citations to the brand’s web pages and the overall volume of the blog posts in the given time period. Advertising has a strong and significant impact on both these dimensions (Table 5). The plots for these dimensions in Figure 4 depict the temporal evolution of the impact of advertising on the virality (Panel 2) and visibility (Panel 4). There is a noticeable impact of advertising for the various information-spread (diffusion) dimensions of chatter over the time period. The magnitude of the gap, between the focal brand and the Synthetic Control on both these dimensions, increases during the time period of advertising. There is a strong immediate impact on visibility and the effect is sustained for a few time periods. Whereas the gap between the synthetic brand and the focal brand increases immediately for visibility, virality shows a slow increase in the magnitude of the gap as the campaign progresses. These results are in line with the inference on the estimates of these dimensions (Table 5).

These results suggest that advertising has a strong immediate influence on users, motivating them to discuss the advertised product in online media as well as direct readers to the brand’s websites. Strong impact on virality suggests acceleration of the propagation of the brand related information due to advertising. This is in line with prior research (e.g. Berger ad Schwartz 2011) that suggest that providing cues (about the product or brand) externally helps in aiding the word-of-mouth. The influence on the sharing of the content is relatively slow, reflecting a weak positive feedback effect of advertising on the citation behavior among websites. The impact of advertising on visibility (Figure 7, panel 4) is significant and the magnitude of the impact is sustained for a longer period than for virality. It also shows some periods of contraction and expansion in the gap over time for visibility. The gap in chatter correlates with the intensity and timing of advertising – it seems to increases with the increase in intensity of advertising and also with advertising during the prominent day parts (e.g., prime time). This increase in the in-degree citations to the brand in visibility reflects the increase in attention to the brand’s web page and the
associated products. This supports our proposition that advertising could trigger recall thus motivating
users to share or cite content across the web. This increase in visibility could increase the traffic to the
brand related sites and subsequent conversion or sales. The temporal evolution of the effect is explored in
further detail in the following section on dynamics section below.

**Analysis of Post-Intervention Dynamics**

We assess the short-term and long-term dynamics between the metrics of chatter and advertising
using the Impulse Response Functions using a Vector AutoRegressive (VAR) model (e.g., Hewett et al
2016; Kireyev, Pauwels and Gupta 2015; Srinivasan, Rutz and Pauwels 2015). The stability of the
dimensions, their appropriateness for the specification of the model and the model details are in Appendix
G. The summary of results is in Table 6. The elasticity and the duration of the wear-in and wear-out vary
for different dimensions of chatter. Assessing the impact in terms of elasticity enables comparison of the
impact of advertising across dimensions. As shown in the table, advertising has the strongest short-term
impact on the popularity of the blogs while the strongest cumulative impact (short plus long-term) on
virality and visibility. In general, the cumulative effect of advertising is higher for all the information-
spread dimensions than for the content dimensions. This effect could be because advertising takes longer
to spread through the system or that the digital trail of the information spread (through the hyperlinks in
the articles) is not easily erased and consequently lingers in the collective memory of the social media
ecosystem for a relatively long period. In contrast to the other dimensions, the negativity of the focal
brand decreases as reflected in the negative elasticity both in the short and long term.

The time taken for wear-in and wear-out of advertising suggests that the median value of the
wear-in across all the dimensions of chatter is 3.5 days and the corresponding value for wear-out is 6.5
days. In terms of the time taken for the wear-in, the shortest duration of impact is for the popularity metric
and the longest duration is for virality metric. The duration of wear-out for virality and visibility
(information-spread dimensions) is longer than that for popularity and negativity (content-based metrics).
Discussion

This section summarizes the main findings, discusses generalizability, and lists some implications and limitations.

Summary

The main findings of the study are the following:

1. Offline television TV advertising has a positive effect on many metrics of chatter as can be ascertained by Synthetic Control.
2. In terms of the cumulative effect, advertising seems to affect the information-spread dimensions (virality and visibility) much more than content based dimensions (popularity and negativity).
3. Among the information-spread based dimensions, advertising affects visibility the most, followed by virality.
4. In terms of the elasticity of the effects, advertising has a strong effect on popularity in the short term while it has the most impact on virality in the long term. Advertising also has a small temporary effect in decreasing negativity.
5. Among the information-spread dimensions, the wear-in is quickest for popularity and the wear-out is longest for visibility and virality. The accumulated effect persists for the longest time for virality.

Generalizability

This study is unique in that it assesses the effect of offline TV advertising on different dimensions of online chatter. Because of the depth of the dependent variables and the quasi-experimental set up of advertising, it was restricted to one category. To get a sense of generalizability, we compare the estimated effects of advertising in this study versus those from a meta-analysis of over 400 estimates in Sethuraman, Tellis and Briesch 2011. Overall, the estimated short-term elasticities from this study range from -0.03 (negativity dimension) to 0.08 with a mean of 0.04. This number compares well with the mean of 0.12 from the meta-analysis of Sethuraman, Tellis and Briesch 2011. The cumulative effect of advertising in this study is about twice the short-run effect, similar to the finding in Sethuraman et al 2011. These results suggest that the estimated effects of advertising are similar to the past findings and may generalize across categories. However, the results here are on online chatter, while the focus on past research is on sales. Moreover, this study examines the effect of advertising in a quasi-experimental context relatively
minimizing some of the endogeneity concerns. In addition, this study also examines the wear-in and wear-out of offline advertising on different dimensions of online chatter.

**Implications**

These findings have four implications for analysts and managers.

First, while advertising expenditures are moving rapidly from traditional to new media, managers need to realize that besides the effect of advertising on sales, it may also affect online chatter. Testing that effect may enable advertisers to assess the effectiveness of advertising more quickly than they could by using sales as the dependent variable. Social media intelligence could be an effective tool in forecasting success of marketing activities. A deep analysis of the network structure in the blogs and social media could help in social media marketing through seeding – identifying and targeting the authoritative chatter generating nodes and influential communities in the network.

Second, when carrying out such analyses, analysts need to consider a rich variety of chatter metrics, to capture the variety of effects that TV advertising may have on online chatter. Based on the importance of the virality and visibility metric in the results above, advertising seems to have a magnifying effect, suggesting amplification of the information propagation across the Internet, observed as the widening of the gap along these dimensions. Advertising seems to accelerate the propagation of information of the brands and at the same time increases the visibility of the brands. This increase in the visibility could translate to awareness about the products and brands among new readers, which could further translate to future sales as suggested by some of the past research in online word of mouth (Babic et al 2015). The trajectory of the propagation of information among the network suggests that the timing of the advertising could be manipulated to sustain the momentum of the information spread across the social media network.

Third, observational data in the form of chatter metrics are readily available through social media intelligence agencies (e.g., Radian6, Crimson Hexagon, BrandWatch). Analysts need to fully exploit the richness of these new chatter data. Most chatter metrics are available at an aggregate (brand or firm) level.
and at a high temporal frequency (e.g., daily, hourly and in some cases by the minute). In the past, the tendency of analysis has been to aggregate the data. However, we highly recommend analyzing the data at the daily level so as to extract rich, detailed, and timely recommendations for managers.

Fourth, though the method of synthetic control is applied to assess the impact of advertising, it could be extended for assessing the impact of other marketing interventions, such as changes in price, sales promotions, or press releases. As in the case of advertising done here, the purpose of doing so is to use chatter metrics as an instantaneous or live diagnostic of the performance of changes in the marketing mix. As such, analysts and managers do not have to wait for monthly or quarterly sales reports.

**Limitations and Future Research**

This study suffers from several limitations. First, this study related offline advertising to online chatter. The explosion of media necessitates the testing of other more complex paths through which advertising could work. One particular possibility is to ascertain how offline advertising affects respond to online advertising and vice versa. The substantive findings of this study suggest that the offline and online worlds are connected in intimate ways that deserve scrutiny and exploitation.

Second, this study is limited to one category during one-time period. Also messages in the campaign catered to humorous elements in the creatives. Future research would need to test the generalizability of the results over other categories of products and services, other time periods, and a variety of creatives besides humor.

Third, Synthetic Control is just one method for the analysis of quasi-experiments. Future research in marketing could test the strengths and weaknesses of this method against other methods such as difference-in-difference, panel factor models (e.g. interactive fixed effects, common correlated effects), and instrumental variables. Nevertheless, subject to the limitations in the Identifying Assumptions section, Synthetic Control provides a simple but elegant method for the analysis of quasi experiments involving dynamics.
Fourth, in this study, competitors did not respond aggressively to HP’s advertising campaign. In other circumstances, such as duopolies (e.g., Coke-Pepsi), reaction to intervention is likely to occur. Such reactions enrich the phenomenon and increase the opportunities for research.
References


Table 1: Mean Factor Loadings of the Dynamic Factor Analysis

<table>
<thead>
<tr>
<th></th>
<th>Estimated Dimensions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Popularity</td>
</tr>
<tr>
<td>Volume of reviews</td>
<td>0.8103</td>
</tr>
<tr>
<td>Positive Valence (reviews)</td>
<td>0.1035</td>
</tr>
<tr>
<td>Negative Valence (reviews)</td>
<td>-0.3547</td>
</tr>
<tr>
<td>Polarity (Entropy of ratings in reviews)</td>
<td>-0.0124</td>
</tr>
<tr>
<td>Volume of blogs</td>
<td>0.6514</td>
</tr>
<tr>
<td>Indegree of the brand (blog based)</td>
<td>0.0034</td>
</tr>
<tr>
<td>Indegree links (blog based)</td>
<td>0.0241</td>
</tr>
<tr>
<td>Volume of Blog that gain rank</td>
<td>0.0187</td>
</tr>
</tbody>
</table>

Note: The measures having the highest factor loadings are in bold.

Table 2: Assessing Effect of Advertising with Difference in Difference

<table>
<thead>
<tr>
<th>DID Coefficient</th>
<th>Popularity</th>
<th>Virality</th>
<th>Negativity</th>
<th>Visibility</th>
</tr>
</thead>
<tbody>
<tr>
<td>R Square</td>
<td>5.57* (1.59)</td>
<td>16.23* (5.52)</td>
<td>-0.09* (0.012)</td>
<td>11.28* (5.31)</td>
</tr>
</tbody>
</table>

Note: * indicates significance at 0.05 level.

Table 3: Comparison of HP and Synthetic Control Brand on Key Pre-Intervention Characteristics

<table>
<thead>
<tr>
<th></th>
<th>Time Period</th>
<th>Mean</th>
<th>HP</th>
<th>Synthetic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Firm Size (Quarterly, billion)</td>
<td>Q1 2010</td>
<td>23.1</td>
<td>113.6</td>
<td>40.31</td>
</tr>
<tr>
<td>Market Capitalization (daily, billion)</td>
<td>January-March 2010</td>
<td>41.68</td>
<td>119.1</td>
<td>35.82</td>
</tr>
<tr>
<td>Debt/ Equity Ratio (Daily)</td>
<td>January-March 2010</td>
<td>1.28</td>
<td>0.63</td>
<td>3.37</td>
</tr>
<tr>
<td>CAGR in EBDITA (Annual, %)</td>
<td>2009</td>
<td>-10.8</td>
<td>-3.2</td>
<td>-13.1</td>
</tr>
<tr>
<td>Media coverage (Daily)</td>
<td>January-March 2010</td>
<td>49</td>
<td>79</td>
<td>68</td>
</tr>
<tr>
<td>Marketing (XSGAQ) (Quarterly, million)</td>
<td>Q1 2010</td>
<td>1063.79</td>
<td>3648</td>
<td>2420.81</td>
</tr>
<tr>
<td>Advertising (XAD) (Annual, million)</td>
<td>2009</td>
<td>397.12</td>
<td>700.0</td>
<td>710</td>
</tr>
<tr>
<td>Total Revenue (Annual, billion)</td>
<td>2009</td>
<td>29.3</td>
<td>114.5</td>
<td>43.0</td>
</tr>
<tr>
<td>Total Number of Employees (Annual, Thousands)</td>
<td>2010</td>
<td>8.6</td>
<td>324</td>
<td>142</td>
</tr>
<tr>
<td>Online Advertising (Monthly, million)</td>
<td>January-March 2010</td>
<td>0.93</td>
<td>4.9</td>
<td>4.61</td>
</tr>
</tbody>
</table>
### Table 4: Elasticities of Focal Brand’s Advertising on Rival Brands’ Chatter

<table>
<thead>
<tr>
<th>Dimension of Chatter</th>
<th>Focal Brand Elasticity</th>
<th>Rival Brands’ Elasticities</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>HP</td>
</tr>
<tr>
<td>Popularity</td>
<td>0.02</td>
<td>-0.001</td>
</tr>
<tr>
<td>Virality</td>
<td>0.02</td>
<td>0.03</td>
</tr>
<tr>
<td>Negativity</td>
<td>-0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>Visibility</td>
<td>0.04</td>
<td>0.026</td>
</tr>
</tbody>
</table>

Note: Bold Indicates significance (p <0.05)

### Table 5: The Effects of TV Advertising on Dimensions of Chatter of HP

<table>
<thead>
<tr>
<th>Outcome Variable</th>
<th>Estimates</th>
<th>Pre-advertising RMSPE</th>
<th>RMSPE Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Popularity</td>
<td>5.36*</td>
<td>2.41</td>
<td>2.23</td>
</tr>
<tr>
<td>Virality</td>
<td>19.85*</td>
<td>2.36</td>
<td>3.12</td>
</tr>
<tr>
<td>Negativity</td>
<td>-0.05*</td>
<td>0.22</td>
<td>2.18</td>
</tr>
<tr>
<td>Visibility</td>
<td>25.18*</td>
<td>3.24</td>
<td>3.27</td>
</tr>
</tbody>
</table>

Note: The estimates here represent the mean gap (difference between the brand and the synthetic control) during the intervention period and the pre-advertising RMSPE (Root Mean Square Prediction Error). The RMSPE ratio is the ratio of the advertising RMSPE to that of the pre-advertising RMSPE of the focal brand * indicates significance of the exact-p value at 0.1 levels which is the highest value that can be obtained given the small number of brands in the sample. (details in the Online Appendix H).

### Table 6: Dynamics of Offline Advertising on Chatter

<table>
<thead>
<tr>
<th></th>
<th>Short Term Elasticity to Advertising</th>
<th>Cumulative Elasticity to Advertising</th>
<th>Wear-in (Duration, in Days)</th>
<th>Wear-out (Duration, in Days)</th>
<th>Total Duration of Impact (Days)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Popularity</td>
<td>0.08</td>
<td>0.15</td>
<td>2</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>Virality</td>
<td>0.05</td>
<td>0.21</td>
<td>6</td>
<td>8</td>
<td>14</td>
</tr>
<tr>
<td>Negativity</td>
<td>-0.03</td>
<td>-0.06</td>
<td>4</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>Visibility</td>
<td>0.07</td>
<td>0.17</td>
<td>3</td>
<td>7</td>
<td>10</td>
</tr>
</tbody>
</table>
Figure 1: Experimental Design

Figure 2: Brand Advertising Intensity of the Firms

Note: This figure depicts the brand advertising only. Subsequent graphs include product or promotional advertising.
Figure 3: Average Weekly TV Advertising of HP Versus Other Top Brands
(top brands include Apple Inc, Canon, Dell, Lexmark)

Figure 4: Average Weekly Advertising across Different Media of HP
Figure 5: Average Weekly Advertising by Top Brands on Non-TV media Versus HP TV

Note: Non-TV includes Print (including Sunday, national, local and Hispanic newspapers and magazine, B2B advertising), Radio (national, local and spot), Internet Display, Outdoor Display.

Figure 6: Reach of YouTube Videos and Television Advertisements in the Intervention Period

Note: The viewership of 205 million is the direct estimate and is only for one channel, hence a low estimate for the viewership. The overall viewership estimate is around 722 million. The details about the data collection and estimates are in Online Appendix F.
Figure 7: Effect of Advertising on Dimensions of Chatter