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

Seshadri Tirunillai,[a] Gerard J. Tellis[b]

[a] C. T. Bauer College of Business, University of Houston, Houston, Texas 77204; [b] Marshall School of Business, University of Southern California, Los Angeles, California 90089
Contact: seshadri@bauer.uh.edu, http://orcid.org/0000-0002-6385-7189 (ST); tellis@usc.edu (GJT)

Abstract. This study analyzes the impact of offline television advertising on multiple metrics of online chatter or user-generated content. 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 preadvertising 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 television 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.

Introduction

User-generated content, or (online) chatter, has become a very important force in contemporary markets for several reasons. First, it has grown enormously in the past decade. Second, surveys suggest that it is one of the most important sources of information that consumers trust (Nielsen Reports 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. 2016) 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 43% of the advertising budget for firms in the United States (Kawasaki 2015). Other studies also support the dominance of TV advertising in the ad budget (e.g., Joo et al. 2014). A big underresearched issue 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 the section on Theory: How Offline Advertising Affects Online Chatter). No studies have systematically assessed the effectiveness and dynamics of TV advertising on various metrics of online chatter. 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 advertising and chatter of HP together with those of control brands, that are close rivals of HP in the major markets in which it competes, for three weeks after the start of the campaign. 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 et al. 2010). The essence of this approach is to create a counterfactual (synthetic) brand from all of the relevant rivals of HP in the preadvertising period and compare the gap between HP's actual chatter and the chatter of the synthetic brand during the advertising period. Furthermore, we analyze the dynamic effect of advertising on the gap in chatter using the vector autoregressive (VAR) model. In sum, this study seeks answers to the following questions:

• Does an 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 ("wear-in" (buildup), "wear-out" (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 this 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 et al. 2016). 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 et al. 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 raises the awareness of a brand that could increase the interest and attention given to the brand by the customers (Tellis 2003, Srinivasan et al. 2016). Studies suggest that consumers pay more attention to products and services in online media when exposed to advertising. This has been shown empirically in the context of consumers searching for information on the web (e.g., Joo et al. 2014, Kireyev et al. 2016), where ads trigger or increase the volume of online search (in search engines). Hu et al. (2014) further decompose the effect of advertising on online search into its effect impact on prepurchase information search and on the conversion to final sales. We could extend this argument and attribute similar effects of advertising on generation of chatter to increased consumer attention to the products. Advertising could draw customers' attention and prime them to seek more information about the product or the brand. This increased attention toward a brand leads to a rise 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

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 et al. 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 1998). Furthermore, ads could change the attitudes and beliefs associated with a brand, leading consumers to share their opinions with others to cope with the feeling of dissonance (Festinger 1962). Ultimately, this effect of advertising could translate into lower negatives toward 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 negatives toward the advertised brand.

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 "Let's Do Amazing" advertising campaign by Hewlett-Packard that ran from March 13 through the end of May 2010.[1] This campaign had a budget of about $40 million. 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 (Vranica and Scheck 2010). The campaign portrayed celebrities such as Dr. Dre (music) and Annie Leibovitz (photography) and firms such as UPS (logistics) and Dream-Works (movies) using the company's technology in their daily operations.[2] It did not promote a 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 in Online Appendix B 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 the first observed 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 in Methods titled Identifying Assumptions of Synthetic Control discusses the assumptions of the quasi-experimental design and provides extensive evidence to support these assumptions. We use difference in differences (DiD), synthetic control, and vector autoregression 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 products. 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, those that focus on the content or characteristics of chatter, or viral based, those 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 et al. 2013, Onishi and Manchanda 2012).

Figure 1. Experimental Design
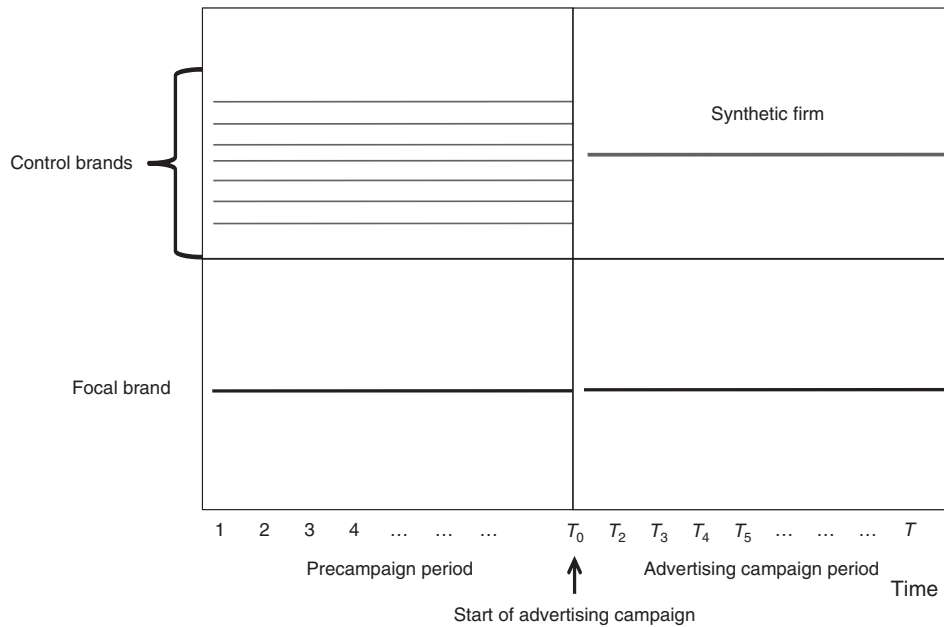
In addition to these, we use polarity, a measure of sentiment divergence. We collect the reviews from consumer reviews and ratings on Amazon.com, Epinions.com, and cnet.com, as we are interested in assessing the impact of TV advertising on product evaluations in reviews. Alternate social media such as Twitter, YouTube, and Facebook had restrictions in terms of availability and drawbacks in measurement of consumer evaluations (because of 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 the brand in chatter. The volume of reviews is based on the number of 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 content in the review. First, we classify the review as positive (or negative) using popular unsupervised machine learning techniques for textual classification of the sentiments—a support vector machine and 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 volume of the positively or negatively classified reviews for a brand in a given time period as positive or negative valence, respectively.

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

$$Entropy(k) = -\sum_{l=0}^{5} P(\eta = l)\log_2 P(\eta = l). \qquad (1)$$

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 metrics 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 webpages using hyperlinks are one of the core metrics in search algorithms. Given the richness of the citations in blogs, they have been used to understand the diffusion of information over the web (e.g., Leskovec et al. 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

et al. 2011, Trusov et al. 2009, Roos et al. 2016) delve into these metrics, though the network structure of the citations across the Internet has been proven to be a very important metric in information science (e.g., Leskovec et al. 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 this measure similar to the volume of the reviews following the prior literature (e.g., Onishi and Manchanda 2012, Gopinath et al. 2013). We measure the volume of blog posts based on the tags associated with them, tags with any words related to the products of the brand, or presence of the mention of the brand (or products) in the title or content.

In-degree (links) of the brand website. We use the hyperlink 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 et al. 2016). Using the graphical network structure of hyperlinks among the blogs, we measure the number of hyperlinks pointing to uniform resource locators (URLs) containing the home domain of the brand (referred to 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 the in-degree of the home domain site of the brand of interest within our blog data. Details on the measurement are 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 also be linked back (cited) by other blogs, websites, or social media, depending on the value of the content in the post. Thus, an 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[3] 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 blog 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 webpages 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* versus *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 webpage 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.[4] 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 the count of posts about the brand that underwent a change in the ranking in the top-tier blog posts from the presampling time period.

Dimensions of Chatter. 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:

$$
\begin{aligned}
Y_t &= \xi f_t + \epsilon_t, \\
f_t &= \Psi f_{t-1} + \eta_t,
\end{aligned}
\tag{2}
$$

where ξ represents the vector of factor loadings; ϵ is the idiosyncratic error; and η 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

Table 1. Mean Factor Loadings of the Dynamic Factor Analysis

	Estimated dimensions			
	Popularity	Virality	Visibility	Negativity
Volume of reviews	**0.8103**	0.0834	0.3258	0.307
Positive valence (reviews)	0.1035	0.0781	0.3381	−0.3258
Negative valence (reviews)	−0.3547	−0.0746	0.1936	**0.7381**
Polarity (entropy of ratings in reviews)	−0.0124	0.2485	0.4825	**0.6936**
Volume of blogs	**0.6514**	0.3417	**0.6135**	0.4825
In-degree of the brand (blog based)	0.0034	0.4456	**0.5283**	0.2134
In-degree links (blog based)	0.0241	**0.8129**	0.1461	0.283
Volume of blogs that gain rank	0.0187	**0.5523**	0.4682	0.461

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

that gained rank and the in-degree of the blogs. It reflects the relative rate of the spread of information about a brand across the Internet. The more such blog posts transmit the message of the brand, the greater 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, *virality*, 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 webpages (e.g., product pages) on the website in their own blog or webpages. 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 to either the volume or the content of chatter. Similarly, we refer to the last two dimensions as viral based because of their focus on the brand-related information spreading across the web. We use the estimated value on each of these four dimensions across all of 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 on television on any given day (also referred to as "placements" or "TV spots"). We focus on television ads, as the campaign was targeted mainly in the form of 30-second ad spots with

interviews with 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).[5]

Apart from the TV advertising, we also collect information on the online advertising and 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.

Method
This section presents the method of synthetic control in terms of motivation, intuition, specification, sampling, identifying assumptions, and the analysis of the dynamics.

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 is 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 United States, 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 et al. 2015). This is similar to some of the other matching-based program evaluation methods (Imbens and Wooldridge 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 preadvertising 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 preintervention 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 the predicted chatter of the synthetic control and the actual chatter of HP during the advertising period. If the gap is significantly different from zero 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 et al. (2010, 2015). 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 let the remaining brands $j \in [2, \ldots, J + 1]$ represent 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 preintervention period and $T_0 \leq t < T$ constitutes the intervention period (refer to 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 let 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, α_{it}, as the difference between the treated brand and the counterfactual or synthetic brand, which is

$$\alpha_{it} = Y_{it}^I - Y_{it}^N, \quad \forall t \in [1, T]. \tag{3}$$

During the preintervention 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_{1t}^N, of the treated brand had it not undergone the intervention. Thus, the gap in the treatment period can be represented as

$$\alpha_{1t} = Y_{1t}^I - Y_{1t}^N, \quad \forall t \in [T_0, T]. \tag{4}$$

While we observe the outcome of the focal brand Y_{1t}^I, the outcome of the synthetic brand Y_{1t}^N is not observed and has to be estimated. Following Abadie et al. (2010), Y_{it}^N can be written using a factor model given by

$$Y_{it}^N = \delta_t + \boldsymbol{\theta}_t \mathbf{Z}_i + \boldsymbol{\lambda}_t \boldsymbol{\mu}_i + \epsilon_{it}, \tag{5}$$

where δ_t represents the unknown common factor with constant factor loadings across all of the units; $\mathbf{Z}_i$ represents the $(r \times 1)$ vector of observed covariates not influenced by the intervention; $\boldsymbol{\theta}_t$ represents the $(1 \times r)$ vector of unknown parameters of $\mathbf{Z}_i$; the $\boldsymbol{\mu}_i$ vector $(F \times 1)$ captures the factor loadings of $\boldsymbol{\lambda}_t$, the $(1 \times F)$ vector

of unobserved common factors; and ϵ_{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 preintervention period. The covariates (Z_i) could be time invariant (during the sample period) or time varying (with varying frequencies). This specification enables us to account for any time-varying unobservable heterogeneity among the brands during the period. In fact, Equation (3) reduces to fixed-effects difference in differences if we impose the restriction that λ_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 the variables are in the Sampling section.

Abadie et al. (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 $W = (w_1, w_2, \ldots, w_i, \ldots, w_J)$ 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

$$\hat{\alpha}_{it} = Y_{1t} - \sum_{j=2}^{J+1} w_t^* Y_{jt}. \tag{6}$$

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 they minimize the difference between the preintervention characteristics of the treated unit and the synthetic control, $X_1 - X_0 W$, during the preintervention period, where X_1 represents the vector of preintervention characteristics and may include the outcomes, i.e., $X_1 = (Z_1', \bar{Y}_1^{K_1}, \ldots, \bar{Y}_1^{K_M})'$, and X_0 is 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 preintervention characteristics of the treated unit on a given variable m and X_{0m} is the vector of values of the J donor pool units for the corresponding characteristic, we obtain W^* by minimizing

$$\sum_{m=1}^{k} v_m (X_{1m} - X_{0m} W)^2, \tag{7}$$

where v_m is a weight that represents the relative importance of the mth variable in measuring the discrepancy between X_{1m} and $X_{0m} W$ (Abadie et al. 2015). The variables with large predictive power on the outcome would be assigned larger weights in the calculation

of v_m. In practice, the choice of v_m is derived from the data by choosing the weights v_m to closely approximate the trajectory of the treated unit in the preintervention period such that the root mean square prediction error (RMSPE) of the outcome variable is minimized during the preintervention period. Details of the implementation[6] are in Abadie et al. (2010, 2011, 2015).

We need to assess whether the trajectory of the gaps after the intervention can be attributed to the advertising undertaken and not to random chance or 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 whether 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 techniques. We trace the steps used in identification of the gaps for each of the brands in our sample. The goodness of fit can be assessed using the calculated pre- and post-treatment RMSPE for all of the brands. The distribution of the postintervention RMSPE to the preintervention RMSPE ratio indicates the probability of observing the outcome by chance. If the observed postintervention RMPSE/preintervention RMSPE ratio of the focal brand compared to most of the placebo brands is large, we can conclude that the observed gap in the postintervention 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 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 Corporation, Seagate Technology, Western Digital Corporation, SanDisk Corporation, Lexmark International, Inc., Netapp, Inc., and Logitech International. For the time sampling, we choose 70 days of data during the preintervention period to construct the synthetic brand that is the closest representation of the focal brand (HP). For the intervention period, we choose the first 20 days of advertising for presenting the results. We do so 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* (quarterly), (4) *Compounded annual growth rate (CAGR) in revenue* (annual), (5) *Media coverage* (daily), (6) *Marketing* (XSGA, quarterly), (7) *Advertising* (XAD, annual), (8) *Online advertising* (monthly), (9) *Total revenue* (annual), and (10) *Total number of employees* (annual).

Identifying Assumptions of Synthetic Control

The validity and strength of synthetic control depend on the extent to which the following assumptions are met.

1. One assumption is that the donor brands are a good match for the synthetic control (Abadie et al. 2010). This assumption can be ascertained by comparing the gap in chatter between the synthetic control and HP in the pretreatment period. Ideally this gap should be 0. Figure H1 (discussed in detail in the robustness check results in Online Appendix H) shows that this gap is close to 0 for the pretreatment 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 of the brands in the analysis from Kantar Media. Figure 2 shows that the donor

Figure 2. Brand Advertising Intensity of the Firms
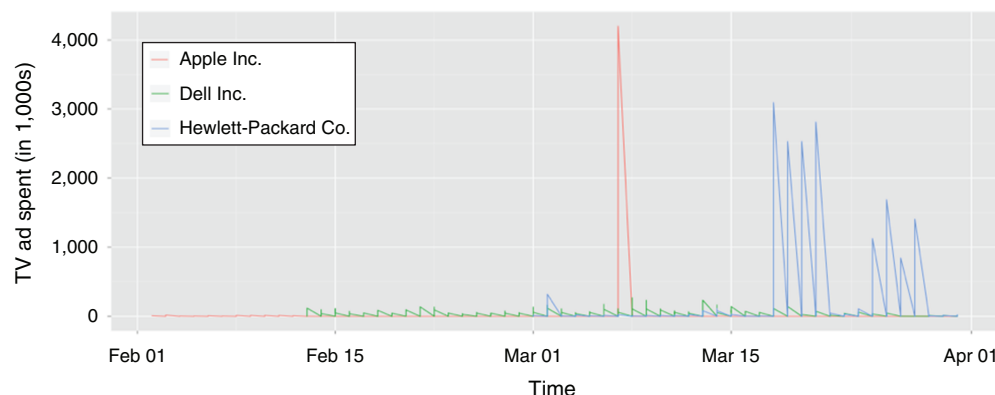
Notes. This figure depicts brand advertising only. Subsequent graphs include product or promotional advertising.

**Figure 3.** Average Weekly TV Advertising of HP vs. Other Top Brands
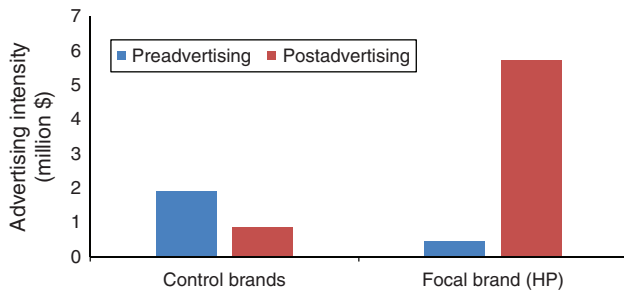
Note. Top brands include Apple, Canon, Dell, and Lexmark.

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 date of the intervention, 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 7 in the pretreatment period is for commercials for the forthcoming launch of the iPad. It represents an aberration that is taken into account when computing the synthetic control. Figure 3 shows the average advertising spending for TV by the control brands and the focal brand 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, and Internet display and search) from Kantar Media. We use weekly level data for comparison, as that was the 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 HP's 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. The figure also shows the average spending of other brands on non-TV advertising. The change in advertising spending is highest for HP's TV advertising. 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 LexisNexis and Factiva for announcements in U.S. newspapers regarding major changes in

Figure 4. Average Weekly Advertising Across Different Media of HP
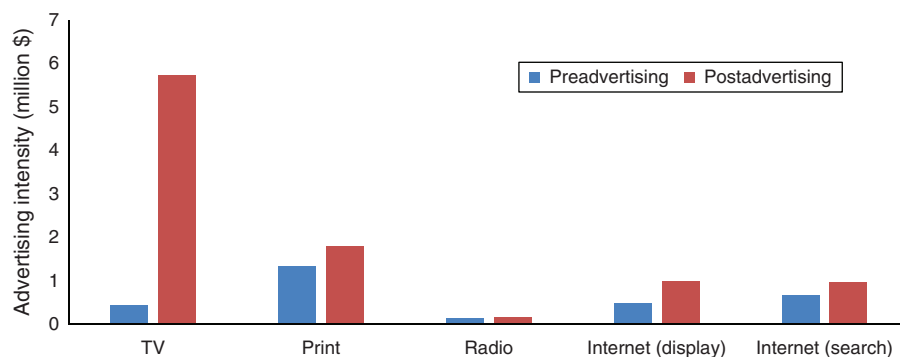
Figure 5. Average Weekly Advertising by Top Brands on Non-TV Media vs. HP TV
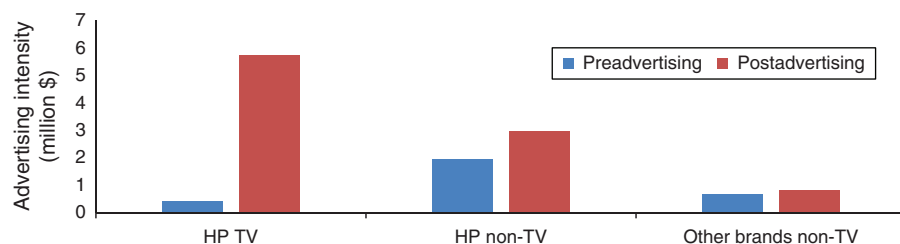
Note. "Non-TV" includes print (including Sunday, national, local, and Hispanic newspaper and magazine and business-to-business), radio (national, local, and spot), Internet display, and outdoor display advertising.

HP marketing between January 1 and April 3, 2010, in the United States. 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," and "channel." Such searches of announcements and press reports were 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 during the advertising period retrieved through an extensive search of Factiva and LexisNexis news databases 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 offline TV versus online YouTube views. We got the YouTube data for the time period from Visible Measures.[7] We got the TV viewership data from syndicated reports (e.g., TV guide websites[8] 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 of offline TV advertising dominate views of online YouTube video advertising. There was a page[9] on HP's website that linked to the videos of the campaign. As shown previously, the views of YouTube videos are much lower than the predicted views of offline TV ads.

Quasi-experimental designs are not perfect and may have some shortcomings. For example, Abadie et al. (2010, p. 501) acknowledged four weaknesses in their design.[10] Thus, discussion of the above assumptions

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

Notes. The television viewership of 205 million is the direct estimate and is only for one channel, and hence is 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.

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, it does not quantify the dynamic effect of an advertising campaign after it is launched. Theory and the literature suggest 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). To assess the dynamics of the effect of advertising on the different chatter metrics, we adopt the vector autoregressive 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 model 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 variables 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 are in Online Appendix G.

Results

This section covers the results of the difference in differences estimation, estimation of the synthetic control, and analysis of dynamics. The robustness checks and associated analysis are in Online Appendix H.
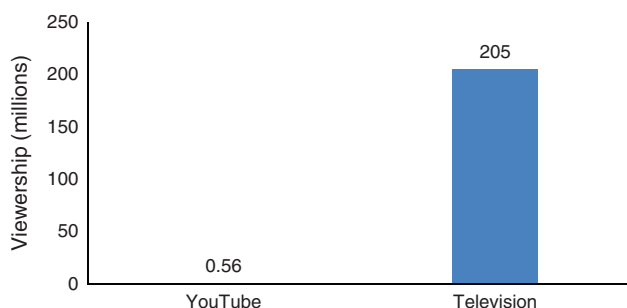
Assessing the Effect of Advertising Using Difference in Differences

DiD is a 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 a commonly used panel data technique to assess the impact of economic programs or interventions (Imbens and Wooldridge 2009). We specify the following fixed-effects least squares regression to assess the impact on each of the outcomes of the chatter metric:

$$Y_{it} = \tau D_{it} + \pi_t + \gamma_i + \epsilon_{it}. \tag{8}$$

Here, Y_{it} is a metric of chatter at time t, D_t is a dummy for the advertising that takes a value of one after the start of advertising for the focal firm and zero otherwise, π_t is the common time effect, γ_i is the brand fixed effect, and ϵ_{it} is the unobservable random error term. The parameter of interest is τ, which captures the impact of advertising. The standard errors are robust standard errors clustered by brand.

Table 2. Assessing the Effect of Advertising with Difference in Differences

	Popularity	*Virality*	*Negativity*	*Visibility*
DiD coefficient	5.57*	16.23*	−0.09*	11.28*
	(1.59)	(5.52)	(0.012)	(5.31)
R-squared	0.942	0.919	0.908	0.867

*Indicates significance at the 0.05 level.

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

Estimating the Synthetic Control

We construct the synthetic control for HP estimating weights for control brands following the method outlined above for the period from January 1, 2010, through March 12, 2010. The daily level data constitute 70 calendar days during the preintervention 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 nonzero weights (0.37, 0.12, and 0.51, respectively) among all of the donor brands. The other brands in the donor pool received 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 the construction of the synthetic control brand is in Table 3. The table 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 preintervention period.

Effect of Advertising on Competing 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 whether HP advertising has any effect on dimensions of HP chatter 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 preintervention period). Figure H1 in Online Appendix H shows that this gap is significantly different from 0 for each dimension of chatter after

Table 3. Comparison of HP and Synthetic Control Brand on Key Preintervention Characteristics

	Time period	Mean	HP	Synthetic
Firm size (quarterly, billion)	Quarter 1 2010	23.1	113.6	40.31
Market capitalization (daily, billion)	January–March 2010	41.68	119.1	35.82
Debt/equity ratio (quarterly)	Quarter 1 2010	0.20	0.33	0.23
CAGR in revenue (annual, %)	2009	−10.8	−3.2	−13.1
Media coverage (daily)	January–March 2010	45	16	62
Marketing (XSGA) (quarterly, million)	Quarter 1 2010	1,063.79	3,648	2,420.81
Advertising (XAD) (annual, million)	2009	397.12	700.0	710
Total revenue (annual, billion)	2009	29.3	114.5	43.0
Total number of employees (annual, thousand)	2010	8.6	324	142
Online advertising (monthly, million)	January–March 2010	0.93	4.9	4.61

Table 4. Elasticities of Focal Brand's Advertising on Rival Brands' Chatter

Dimension of chatter	Focal brand's elasticity	Rival brands' elasticities		
	HP	Dell	Apple	Canon
Popularity	**0.02**	−0.001	0.015	0.06
Virality	**0.02**	0.03	0.08	0.01
Negativity	**−0.01**	0.01	0.001	0.01
Visibility	**0.04**	0.026	0.07	0.02

Note. Bold indicates significance ($p < 0.05$).

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 preadvertising and advertising periods. The figure also depicts the intensity of the campaign in the lower panel. The popularity of HP and the synthetic control brand trace similar paths during the preadvertising 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 preadvertising 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

Figure 7. Effect of Advertising on Dimensions of Chatter
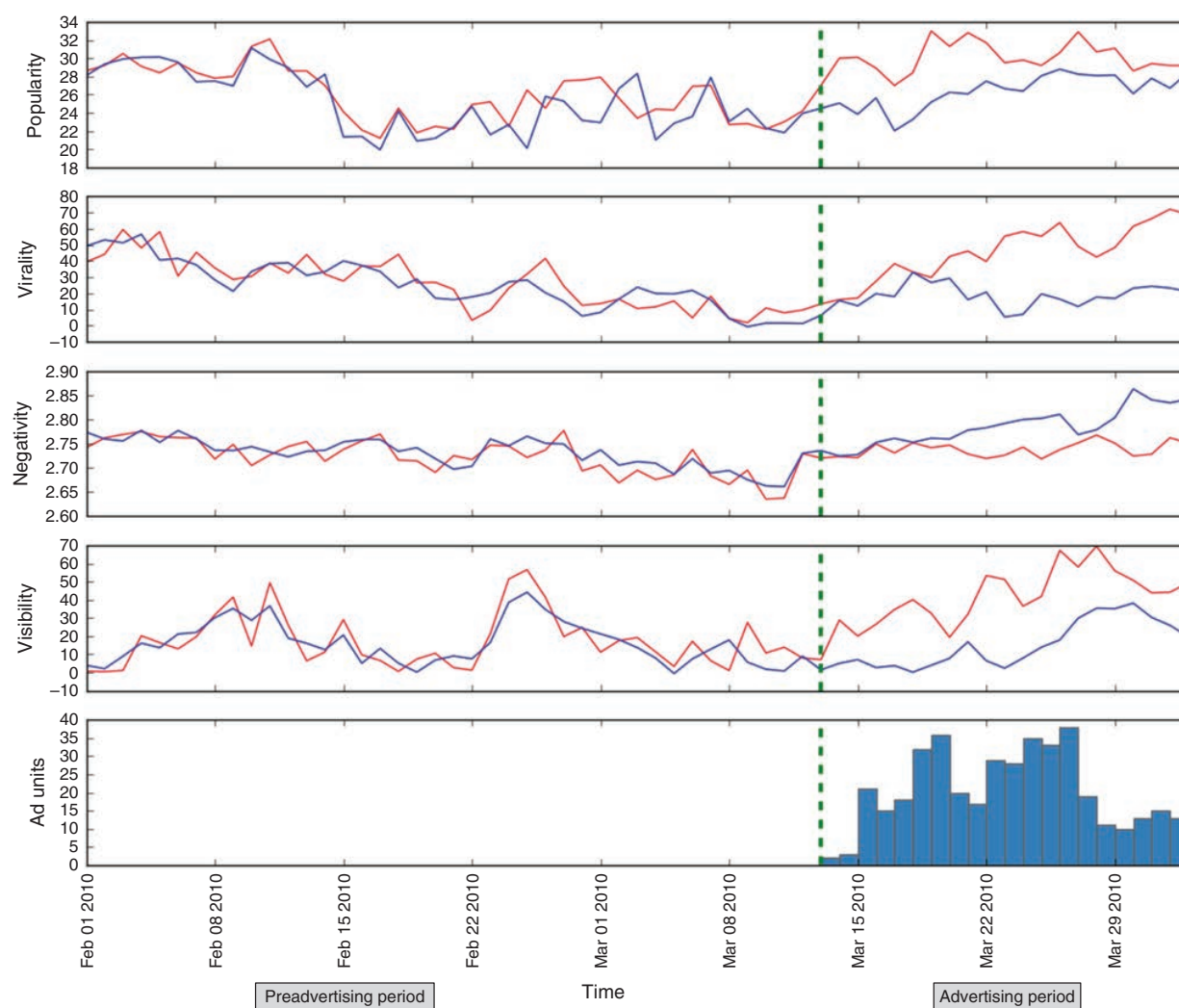
popularity goes up by about 13% on average for the days immediately following the launch of the advertising. The gap reaches a peak of about 19% and then tapers down a couple of days after the start of the campaign. Overall, the volume of chatter of HP relative to that of the synthetic control increases about 15% during this time period. Thus, based on the existence of a marked gap between the trajectory of HP and that of the synthetic control, we can infer that the advertising has a definite impact on popularity. Since the popularity is based on the 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, 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 effects during the advertising period, the preadvertising RMSPE, and the advertising period to preadvertising period RMPSE 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 postadvertising to preadvertising ratio of RMSPE. The ratio of the advertising RMSPE to preadvertising RMSPE is highest for HP when compared to most of the placebo brands, suggesting that the effect of advertising is strong and significant for HP.

We repeat the analysis for negativity, on 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

Table 5. The Effects of TV Advertising on Dimensions of Chatter of HP

Outcome variable	Estimates	Preadvertising RMSPE	RMSPE ratio
Popularity	5.36*	2.41	2.23
Virality	19.85*	2.36	3.12
Negativity	−0.05*	0.22	2.18
Visibility	25.18*	3.24	3.27

Notes. The estimates here represent the mean gap (difference between the brand and the synthetic control) during the intervention period and the preadvertising RMSPE. The RMSPE ratio is the ratio of the advertising RMSPE to that of the preadvertising RMSPE of the focal brand.

*Indicates significance of the exact p-value at the 0.1 level, which is the highest value that can be obtained given the small number of brands in the sample (details in Online Appendix H).

suggests that advertising increases positive chatter and decreases negative chatter. As discussed in the section Theory: How Offline Advertising Affects Online Chatter, this could be attributed to the change in the customer expectations due to advertising, resulting in them interpreting their experiences with the product of the advertised brand favorably, thus generating positive conversations and refuting negative conversation. The third panel of Figure 7 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 the gap of the negativity dimension over the period. We compare the sizes 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 of 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 webpages 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 7 depict the temporal evolution of the impact of advertising on the virality (second panel) and visibility (fourth panel). 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 website. 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 and Schwartz 2011) that suggests that providing cues (about the product or brand) externally helps in aiding word of mouth. The influence on the sharing of the content

Table 6. Dynamics of Offline Advertising on Chatter

	Short-term elasticity to advertising	Cumulative elasticity to advertising	Wear-in (duration, in days)	Wear-out (duration, in days)	Total duration of impact (days)
Popularity	0.08	0.15	2	6	8
Virality	0.05	0.21	6	8	14
Negativity	−0.03	−0.06	4	3	7
Visibility	0.07	0.17	3	7	10

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, fourth panel) 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 increase 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 webpage 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 Analysis of Postintervention Dynamics section.

Analysis of Postintervention 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 model (e.g., Hewett et al. 2016, Kireyev et al. 2016, Srinivasan et al. 2016). The stability of the dimensions, their appropriateness for the specification of the model, and the model details are in Online 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 and the strongest cumulative impact (short plus long term) on virality and visibility. In general, the cumulative effect of advertising is higher for all of 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. By 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 of 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 the 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 setup 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 et al. (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 et al. (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 of past research has been 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 affecting sales, advertising 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 the 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 metrics in the results above, advertising seems to have a magnifying effect, suggesting amplification of information propagation across the Internet, observed as the widening of the gaps along these dimensions. Advertising seems to accelerate the propagation of information of the brands and at the same time increase 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. 2016). 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 analysts has been to aggregate the data. However, we highly recommend analyzing the data at the daily level 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 to assess 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 response 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 differences, panel factor models (e.g., interactive fixed effects, common correlated effects), and instrumental variables. Nevertheless, subject to the limitations in the Identifying Assumptions of Synthetic Control 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.

Endnotes

[1] This period represents the period of high intensity and frequency of the advertisement. After this, there was very low-intensity sparse advertising observed sporadically.

[2] Some of the videos associated with the campaign can be found at https://www.youtube.com/watch?v=1kwz7tTjsOU&index=3&list=PLXjvElQ9mIe0x-1okKDssJIsmvJV4SYup.

[3] See http://spinn3r.com (accessed August 2015).

[4] The ranking is based on the "authority" of the node (URL). Crudely, this can be viewed as analogous to the PageRank or HITS algorithm. We restrict our Internet universe to the blogs (also referred to as the blogosphere). Also, we focus on only the spread of messages surrounding the brand and its products.

[5] The robustness analysis indicates that these alternate metrics do not change the direction of the results.

[6] The software can be found at http://web.stanford.edu/~jhain/synthpage.html.

[7] See http://www.visiblemeasures.com.

[8] See http://tvbythenumbers.zap2it.com/2010/.

[9] 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 (accessed July 2016).

[10] (1) Increase in antitobacco 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.

References

Abadie A, Gardeazabal J (2003) The economic costs of conflict: A case study of the Basque Country. *Amer. Econom. Rev.* 93(1):113–132.

Abadie A, Diamond A, Hainmueller J (2010) Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *J. Amer. Statist. Assoc.* 105(490):493–505.

Abadie A, Diamond A, Hainmueller J (2011) Synth: An R package for synthetic control methods in comparative case studies. *J. Statist. Software* 42(13):1–17.

Abadie A, Diamond A, Hainmueller J (2015) Comparative politics and the synthetic control method. *Amer. J. Political Sci.* 59(2):495–510.

Aggarwal R, Gopal R, Sankaranarayanan R, Singh PV (2012) Blog, blogger, and the firm: Can negative employee posts lead to positive outcomes? *Inform. Systems Res.* 23(2):306–322.

Babic A, Sotgiu F, de Valck K, Bijmolt TH (2016) The effect of electronic word of mouth on sales: A meta-analytic review of platform, product, and metric factors. *J. Marketing Res.* 53(3):297–318.

Baker AM, Donthu N, Kumar V (2016) Investigating how word-of-mouth conversations about brands influence purchase and retransmission intentions. *J. Marketing Res.* 53(2):225–239.

Berger J, Schwartz EM (2011) What drives immediate and ongoing word of mouth? *J. Marketing Res.* 48(5):869–880.

Borah A, Tellis GJ (2014) Make, buy, or ally? Choice of and pay-off from announcements of alternate strategies for innovations. *Marketing Sci.* 33(1):114–133.

Borah A, Tellis GJ (2016) Halo (spillover) effects in social media: Do product recalls of one brand hurt or help rival brands? *J. Marketing Res.* 53(2):143–160.

Chevalier JA, Mayzlin D (2006) The effect of word of mouth on sales: Online book reviews. *J. Marketing Res.* 43(3):345–354.

Chiou L, Tucker C (2012) How does the use of trademarks by third-party sellers affect online search? *Marketing Sci.* 31(5):819–837.

Deighton J, Schindler RM (1998) Can advertising influence experience? *Psych. Marketing* 5(2):103–115.

Du RY, Kamakura WA (2012) Quantitative trendspotting. *J. Marketing Res.* 49(4):514–536.

Festinger L (1962) *A Theory of Cognitive Dissonance*, Vol. 2 (Stanford University Press, Stanford, CA).

Goldfarb A, Tucker C (2011) Online display advertising: Targeting and obtrusiveness. *Marketing Sci.* 30(3):389–404.

Gopinath S, Chintagunta PK, Venkataraman S (2013) Blogs, advertising, and local-market movie box office performance. *Management Sci.* 59(12):2635–2654.

Gopinath S, Thomas JS, Krishnamurthi L (2014) Investigating the relationship between the content of online word of mouth, advertising, and brand performance. *Marketing Sci.* 33(2):241–258.

Hewett K, Rand W, Rust RT, van Heerde HJ (2016) Brand buzz in the echoverse. *J. Marketing* 80(3):1–24.

Higgins ET, Rholes WS, Jones CR (1977) Category accessibility and impression formation. *J. Experiment. Soc. Psych.* 13(2):141–154.

Hoch SJ, Ha Y-W (1986) Consumer learning: Advertising and the ambiguity of product experience. *J. Consumer Res.* 13(2):221–233.

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

Imbens GW, Wooldridge JM (2009) Recent developments in the econometrics of program evaluation. *J. Econom. Literature* 47(1):5–86.

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

Katona Z, Zubcsek PP, Sarvary M (2011) Network effects and personal influences: The diffusion of an online social network. *J. Marketing Res.* 48(3):425–443.

Kawasaki L (2015) The United States tops ad spending per capita in 2014 at USD$567 per person. Press release, Strategy Analytics, Boston (February 26), https://www.strategyanalytics.com/strategy-analytics/news/strategy-analytics-press-releases/strategy-analytics-press-release/2015/02/26/the-united-states-tops-ad-spending-per-capita-in-2014-at-usd$567-per-person.

Kireyev P, Pauwels K, Gupta S (2016) Do display ads influence search? Attribution and dynamics in online advertising. *Internat. J. Res. Marketing* 33(3):475–490.

Leskovec J, Backstrom L, Kleinberg J (2009) Meme-tracking and the dynamics of the news cycle. *Proc. 15th ACM SIGKDD Internat. Conf. Knowledge Discovery Data Mining* (ACM, New York), 497–506.

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

Luan YJ, Sudhir K (2010) Forecasting marketing-mix responsiveness for new products. *J. Marketing Res.* 47(3):444–457.

Mayzlin D, Yoganarasimhan H (2012) Link to success: How blogs build an audience by promoting rivals. *Management Sci.* 58(9):1651–1668.

Nielsen (2013) Global trust in advertising and brand messages. http://www.nielsen.com/us/en/insights/reports/2013/global-trust-in-advertising-and-brand-messages.html.

O'Keefe DJ (1999) How to handle opposing arguments in persuasive messages: A meta-analytic review of the effects of one-sided and two-sided messages. *Ann. Internat. Comm. Assoc.* 29(1):209–249.

Onishi H, Manchanda P (2012) Marketing activity, blogging and sales. *Internat. J. Res. Marketing* 29(3):221–234.

Pauwels K, van Ewijk B (2014) Do online behavior tracking or attitude survey metrics drive brand sales? An integrative model of attitudes and actions on the consumer boulevard. Report 13-118, Marketing Science Institute, Cambridge, MA.

Peters K, Chen Y, Kaplan AM, Ognibeni B, Pauwels K (2013) Social media metrics—A framework and guidelines for managing social media. *J. Interactive Marketing* 27(4):281–298.

Roos JMT, Mela CF, Shachar R (2016) The effect of links and excerpts on Internet news consumption. Working paper, Erasmus University, Rotterdam, Netherlands.

Saboo AR, Chakravarty A, Grewal R (2016) Organizational debut on the public stage: Marketing myopia and initial public offerings. *Marketing Sci.* 35(4):656–675.

Schweidel DA, Moe WW (2014) Listening in on social media: A joint model of sentiment and venue format choice. *J. Marketing Res.* 51(4):387–402.

Sethuraman R, Tellis GJ, Briesch RA (2011) How well does advertising work? Generalizations from meta-analysis of brand advertising elasticities. *J. Marketing Res.* 48(3):457–471.

Srinivasan S, Rutz OJ, Pauwels K (2016) Paths to and off purchase: Quantifying the impact of traditional marketing and online consumer activity. *J. Acad. Marketing Sci.* 44(4):440–453.

Stephen AT, Galak J (2012) The effects of traditional and social earned media on sales: A study of a microlending marketplace. *J. Marketing Res.* 49(5):624–639.

Stock JH, Watson MW (2011) Dynamic factor models. Clements MP, Hendry DF, eds. *Oxford Handbook of Economic Forecasting* (Oxford University Press, New York), 35–59.

Tellis GJ (2003) *Effective Advertising: Understanding When, How, and Why Advertising Works* (Sage Publications, Thousand Oaks, CA).

Tirunillai S, Tellis GJ (2012) Does chatter really matter? Dynamics of user-generated content and stock performance. *Marketing Sci.* 31(2):198–215.

Tirunillai S, Tellis GJ (2014) Mining marketing meaning from online chatter: Strategic brand analysis of big data using latent Dirichlet allocation. *J. Marketing Res.* 51(4):463–479.

Toubia O, Stephen AT (2013) Intrinsic vs. image-related utility in social media: Why do people contribute content to Twitter? *Marketing Sci.* 32(3):368–392.

Trusov M, Bucklin RE, Pauwels K (2009) Effects of word-of-mouth versus traditional marketing: Findings from an Internet social networking site. *J. Marketing* 73(5):90–102.

Vranica S, Scheck J (2010) H-P to brandish its technology credentials. *Wall Street Journal* (March 11), http://online.wsj.com/news/articles/SB10001424052748704655004575114010908592910.