Marketing Science Institute Working Paper Series 2013
Report No. 13-125

Beyond Likes and Tweets: Marketing, Social Media Content, and Store Performance

Koen Pauwels, E. Craig Stacey, and Andrew Lackman

Report Summary

Social media is often tracked in volume and valence, such as Facebook Likes or positive versus negative Tweets. New monitoring tools also allow marketers to track the *content* of social media conversations, for example, consumers' responses to the brand's marketing, their emotional attachment to the brand or product, or their reports of purchase behaviour related to the brand. Marketers are greatly interested in how marketing activities can affect these social media metrics and to what extent each drives business performance.

Here, Koen Pauwels, Craig Stacey, and Andrew Lackman address the following questions: (1) To what extent does social media conversational content, versus quantity (volume) and sentiment (valence), explain business performance in the short term and the long term? (2) How do paid marketing actions stimulate specific word-of-mouth (WOM) conversational content, and how do these indirect performance effects compare to direct marketing impact?

Their dataset includes marketing, store traffic, and website traffic data from a major U.S. retail brand; WOM data from Crimson Hexagon, a company that specializes in social media monitoring and analysis; and natural search data from Google.com. This yielded complete data for 55 weeks (July 2010 - July 2011). Their vector autoregression (VAR) model estimates how offline and online marketing induce online search and social media conversations and how each of these endogenous variables has direct and indirect effects on business performance.

Their results indicated that different social media content has substantially different performance implications and that marketing actions with small direct effects can have large total effects by stimulating social media conversations. TV, print, and online marketing showed a substantial indirect effect through online search and/or the WOM content metrics of "love for the brand" and "love the ad." The effects of radio were mostly direct on store sales, but then reinforced through "went there/purchased" WOM content.

Content-specific WOM performed better than Facebook and Twitter volume and valence metrics in explaining store traffic and online traffic. Among the conversation topics, "love for the brand" had larger long-term traffic effects, but neutral conversations on "went there/purchased" drove traffic in the short run.

Thus, despite the popularity of Likes and tweets metrics, these findings suggest that social media impact is not just about volume and valence: knowing "what" consumers say matters. Managers who understand how conversations affect their business, and how their marketing strategies influence online conversations, may leverage paid marketing to spark WOM in a way that can positively influence their brands and bottom line.

Koen Pauwels is Professor of Marketing, Ozyegin University. E. Craig Stacey is Director of Research, New York University Stern Center for Measurable Marketing and Founding Partner, Marketing Productivity Group. Andrew Lackman is Business Analyst, Marketing Productivity Group.

1. Introduction

Word of Mouth (WOM) has long been heralded as a more important driver of business success than paid marketing (Kotler 1967, Misner 1999), and the worldwide popularity of online social media (surpassing 1.2 billion users) has provided a new tool for sending and receiving WOM (ComScore 2012). As a result, many companies are moving advertising budgets from offline to online (Forrester 2013), with the $ 1.5B shift of General Motors in 2008 as a notorious example. However, paid advertising, including traditional offline spending, may be key in stimulating (online) word-of-mouth conversations (Godes et al. 2005, Trusov et al. 2009). This observation is central to the influential two-step flow theory in communication (Lazarsfeld, Berelson, & Gaudet, 1948; Katz & Lazarsfeld, 1955) but has been largely overlooked in the current social media effectiveness debate. A key reason is the difficulty in quantifying the relationship between marketing, WOM and business performance, and thus the attribution of performance gains to specific marketing actions and WOM metrics (Godes et al. 2005). Online word-of-mouth, especially in the form of social media, holds much measurement promise as compared to paid offline WOM metrics derived from surveys, experiments or self-report (Dellarocas 2003). Indeed, social media portals provide largely unfiltered forums for conversations about businesses, brands and products. Recent literature has begun to investigate the quantity (i.e., volume) and sentiment (i.e., valence) of social media conversations, including Tweets and Facebook Likes, but has yet to tap into the *topical content* of social media conversations. Managers who understand how conversations affect their business and how their marketing strategies influence online conversations, may leverage paid marketing to spark WOM in a way that can positively influence their brands and bottom line.

This paper is the first to quantify the dynamic effects of offline and online paid marketing actions on the content of social media conversations and the conversion of these conversations into business performance. Our research questions are:

(1) To what extent do WOM conversational topics, versus quantity and sentiment, explain business performance in the short term and the long term?

(2) How do paid marketing actions stimulate specific WOM conversation topics, and how do these indirect performance effects compare to the direct marketing impact?

To address these research questions, we gather data unique in combining typical social media metrics (number and sentiment of Tweets, Facebook likes and comments) with the topic of conversation (love the ad, love the brand, purchased/went to the store) within the context of explaining how offline (TV, radio, print) and online (search, display) paid marketing actions drive both online (i.e., website) and offline (i.e., brick and mortar) store traffic. We show how specific marketing actions induce specific social media conversations, which in turn convert to business performance. The vector-autoregression (VAR) model first estimates the dynamic relationships among paid marketing, social media and business performance, after which impulse response functions track the net performance effects of marketing and social media. A key advantage of the model is that, since all variables are treated as endogenous, the dynamic relationship between all variables is captured without imposing a priori restrictions (Sims 1980). Indeed, we consider the response of store traffic to different social media conversations, but also the direct performance effect of marketing and its indirect effect through stimulating social media conversations. Finally, derived from the VAR estimates, forecast error variance decomposition (FEVD) shows the importance of each variable in driving business.

Our empirical application tracks both online and offline store traffic of a major US retailer, who invests heavily in both offline and online marketing actions. Key results indicate that different social media content has substantially different performance implications and that marketing actions with small direct effects can have large total effects by stimulating social media conversations. As expected, social media conversations have higher short-term performance elasticity (between 0.12 and 0.27) than paid advertising, and enjoy longer carryover effects. Social media content also has higher performance elasticity than typical (volume and valence) measures of online WOM, including Facebook likes and comments, total number of Tweets, and volume of positive versus negative Tweets. Sentimentally-neutral messages such as 'I went to the store' have substantial power to drive traffic. Thus, it is not just about online sentiment: knowing 'what' consumers say matters!

2. Research Background

Word-of-mouth (WOM) has been called "the world's most effective, yet least understood marketing strategy" (Misner 1999). Godes et al. (2005, p. 416-417) define social interactions (a synonym for WOM) as actions that (a) are taken by an individual not actively involved in selling the product or service and that (b) impact other's expected utility from that product or service. From the initial study by Katz and Lazarsfeld (1955), over 70 papers have used self-reports to demonstrate the high sales effectiveness of word-of-mouth (Money et al 1998, Godes and Mayzlin 2004). The current popularity of social media makes it easier for consumers to engage in online WOM and for researchers to track it. WOM proponents tout the benefits of earned media over paid media, which appears to be losing effectiveness (Forrester 2005). Researchers have examined the conditions under which consumers are likely to rely on others' opinions to make a purchase decision, the motivations for different people to spread the word about a product, and the variation in strength of influence people have on their peers in WOM communications. Moreover, customers who self-report being acquired through WOM add more long-term value to the firm than customers acquired through traditional marketing channels (Villanueva, Yoo and Hanssens 2008).

Recent studies have gone beyond self-reports of WOM to analyze Usenet posts (Godes and Mayzlin 2004), product reviews (Chevalier and Mayzlin 2006; Liu 2006), consumer movie reviews (Dellarocas, Zhang and Awad 2007), friend referrals (Trusov et al. 2009), and volume and valence of Tweets (Asur and Huberman, 2010). Godes and Mayzlin (2004) find that offline ratings of new TV shows are driven by the volume and dispersion of conversations across different Usenet groups. Chevalier and Mayzlin (2006) find that better book reviews increase sales, but that negative reviews have greater effects than positive reviews. In contrast, Liu (2006) shows that both negative and positive WOM increase performance (box office revenue). Dellarocas et al. (2007) demonstrate that volume and valence of online movie reviews improve demand forecasting accuracy. Trusov et al (2009) find that new social media via friend referrals are between 20-30 times more effective at driving business that traditional marketing or media appearances. As to the impact of Twitter, Asur and Huberman (2010) find that Tweet rates and valence predict movie box office revenues better than the Hollywood Stock Exhange. Finally, Yamamoto and Matsumura (2011) show the effect of television advertisements, directly and through Twitter posts, on new customer acquisition.

As for methodology, few studies allow for long-term effects and for dynamic interactions among paid, owned and earned media and market performance variables. Asur and Huberman (2010) use a cross-sectional regression of box office performance on Tweets, but do not account for advertising. Yamamoto and Matsumura (2011) capture indirect customer acquisition effects of TV advertising through Twitter with structural equation models. Again, their analysis is cross-sectional instead of over time, and their data did not include different paid media. The only model to capture long-term effects and dynamic interactions is Trusov et al. (2009). However, their data only include events as paid advertising and online friend referrals as social media.

This paper shares above studies' objective of quantifying the effect of WOM and traditional marketing. We aim to move beyond the social media quantity and sentiment by analyzing effects of specific WOM conversation topics. Such different topics include: interactions and experience with the brand/product, emotions toward the brand/product, and conversations directly sparked by the brand. Each of these three conversation types may affect business in different ways. Likewise each conversation type likely has different factors that drive it. We are aim to capture the influences and effects of these three conversation types through Vector Autoregression Modelling, which allows estimation of short-term and long-term, direct and indirect effects of paid marketing and social media on performance.

3. Modeling Approach

Our modeling approach consists of four steps. First, we test for endogeneity and the possibility for long-term (persistent) effects of WOM, paid marketing, store traffic and online traffic. Next, we specify a Vector Autoregressive (VAR) model that is able to account for endogeneity, dynamic responses and interactions among variables (Dekimpe and Hanssens 1999). Third, we estimate short- and long-run response of store traffic and online traffic to paid marketing and WOM in the form of elasticities and decompositions. Finally, we check the robustness of the estimated elasticities to alternative WOM metrics. Table 1 displays these steps. (Tables and figures follow References.)

The first step of our analysis is to test for endogeneity between paid media (marketing efforts by the retailer), owned media (the retailer's website), earned media (WOM through social networking sites), natural search (Google search) and traffic to the retailers' stores. We expect paid marketing to directly impact store traffic, website traffic, WOM, and natural search. Further, online traffic, WOM, and natural search will also have effects on store traffic, suggesting indirect effects of paid media. We also anticipate that store and website traffic will have feedback effects on WOM and natural search activity. The lagged effects of each relationship are also measured.

The dynamic relationships in Figure 1 are established through Granger causality tests (Granger 1969, Hanssens et al. 2001). Granger causality of a variable Y by a variable X means that we can predict Y substantially better by knowing the history of X than by only knowing the history of Y. This 'temporal causality' is the closest proxy for causality that can be gained from studying the time series of the variables. It does not provide the causality established by manipulating a specific variable in controlled experiments. We perform a series of Granger causality tests on each pair of key variables.[1] If traffic or WOM do Granger cause (some of) the marketing variables, we need to capture the complex interactions of Figure 1 in a full dynamic system of multiple equations.

Next, we test for the potential of permanent effects. No such effects are possible for series that are "stationary", i.e. revert to a stable mean (e.g., Dekimpe and Hanssens 1999). Such mean-reversion is shown through unit-root tests, e.g. the Augmented Dickey-Fuller (ADF) test

[1] We note that a wrong choice for the number of lags in the test may erroneously conclude the absence of Granger causality (e.g., Hanssens 1980). Because we are applying these tests to investigate the need for modelling a full dynamic system, we are not interested in whether variable X causes variable Y at a specific lag, but in whether we can rule out that X Granger causes Y at any lag. Therefore, we run the causality tests for lags up to 13 (i.e. one quarter of 13 weeks) and report the results for the lag that has the highest significance for Granger causality.

and the KPSS test (Kwiatkowski, Phillips, Schmidt and Shin 1992). In contrast, an “evolving” series will not revert back the mean after being shocked; the change will persist into the future. Some of these shocks may be due to the other variables in our dynamic system; as quantified in the estimation step. As for model specification, evolving variables need to be differenced to avoid ‘spurious relation’ problems (Granger and Newbold 1986), unless they are tied in a long-term equilibrium. We test for such equilibrium with cointegration tests (Johansen et al. 2000).

The endogeneity (Granger-Causality), evolution (unit root) and cointegration tests allow us to finalize the specification of equation 1. It is called Vector Autoregression, because the vector of endogenous variables (marketing, WOM, traffic, and natural search) is regressed on the own past of each variable and the past of the other endogenous variables. This model specification thus explains each endogenous variable and allows for dynamic feedback loops (e.g. marketing induces WOM the next week, which stimulates store traffic after two weeks, which in turns stimulates WOM in the same week). The vector of endogenous variables includes: Traffic to the retailer’s stores (T), Traffic to the retailer’s Online site (O), Google search index (G), posts about Loving the brand (L), posts about going to the store or making a Purchase (P), posts about the brand’s Advertisements (A), Television GRPs (TV), number of Circulars distributed (C), Radio GRPs (R), and paid Search impressions (S). Each variable is included in logs, so that the estimated effects are directly interpretable as elasticities (Nijs et al. 2001). The vector of exogenous variables includes for each endogenous variable i) an intercept, ii) a ‘seasonal retail mall index’ to control for seasonality (SRTI), and iii) a holiday dummy variable to capture unusually high marketing efforts and traffic around Christmas. Store traffic, love conversations, and paid search impressions are first differenced (denoted by *d*) in order to correct for evolution. The VAR specification is given by:

$$
\begin{bmatrix} d(T_t) \\ O_t \\ G_t \\ d(L_t) \\ P_t \\ A_t \\ TV_t \\ C_t \\ R_t \\ d(S_t) \end{bmatrix} = \begin{bmatrix} C_{d(T)} \\ C_O \\ C_G \\ C_{d(L)} \\ C_P \\ C_A \\ C_{TV} \\ C_C \\ C_R \\ C_{d(S)} \end{bmatrix} + \begin{bmatrix} \delta_{d(T)} \\ \delta_O \\ \delta_G \\ \delta_{d(L)} \\ \delta_P \\ \delta_A \\ \delta_{TV} \\ \delta_C \\ \delta_R \\ \delta_{d(S)} \end{bmatrix} \times SRTI + D \begin{bmatrix} \theta_{d(T)} \\ \theta_O \\ \theta_G \\ \theta_{d(L)} \\ \theta_P \\ \theta_A \\ \theta_{TV} \\ \theta_C \\ \theta_R \\ \theta_{d(S)} \end{bmatrix} + \sum_{j=1}^{J} \begin{bmatrix} \phi_{1\,1}^{j} & \cdots & \phi_{1\,10}^{j} \\ \vdots & \ddots & \vdots \\ \phi_{10\,1}^{j} & \cdots & \phi_{10\,10}^{j} \end{bmatrix} \times \begin{bmatrix} d(T_{t-j}) \\ O_{t-j} \\ G_{t-j} \\ d(L_{t-j}) \\ P_{t-j} \\ A_{t-j} \\ TV_{t-j} \\ C_{t-j} \\ R_{t-j} \\ d(S_{t-j}) \end{bmatrix} + \begin{bmatrix} \varepsilon_{d(T),t} \\ \varepsilon_{O,t} \\ \varepsilon_{G,t} \\ \varepsilon_{d(L),t} \\ \varepsilon_{P,t} \\ \varepsilon_{A,t} \\ \varepsilon_{TV,t} \\ \varepsilon_{C,t} \\ \varepsilon_{R,t} \\ \varepsilon_{d(S),t} \end{bmatrix} \quad (1)
$$

where *t* indexes weeks, J equals the number of lags, as selected by the Bayesian Information Criterion (Lutkepohl 1993), D is the holiday dummy variable, C, δ, θ, γ and ϕ are the parameters to be estimated and ε_t are white-noise disturbances distributed as $N(0,\sum)$. The full residual variance-covariance matrix Σ contains the short-term (same-week) effect of each endogenous variable on the others, as interpreted in the third step.

In the third step, we derive from the VAR-estimates the Generalized Impulse Response Functions (Dekimpe and Hanssens 1999, Pesaran and Shin 1998). Note from Equation (1) that VARX models capture immediate as well as lagged, direct as well as indirect interactions among the endogenous variables. Based on all these estimated reactions, the impulse response function estimates the net result of a "shock" to one variable (e.g. TV) on the other variables relative to their baselines (i.e. their expected values in the absence of the marketing shock). To tease out which of the variables is responsible for a same-week shock, researchers have historically imposed a causal ordering (Sims 1980). Instead, we estimate Generalized Impulse Response Functions (GIRFs) with the simultaneous-shocking approach (Pesaran and Shin 1998), in which the information in the residual variance-covariance matrix of Equation (1) is used to derive a vector of *expected* instantaneous shock values. The advantage of this approach is that it does not require selecting a causal ordering among the variables. Short-term and long-term elasticities are obtained by comparing each GIRF estimate with its standard error, and only retaining those with a t-value higher than unity (Pesaran and Lee 1993, Sims and Zha 1999, Pauwels et al. 2002).

While the IRFs allow us to calculate the performance effect of a unit change in a marketing or WOM variable, we also want to know how important all present and past changes in that variable are to performance. Based on the same VAR-estimates, Forecast Variance Error Decomposition (FEVD) measures the relative impact over time of shocks initiated by each of the individual endogenous variables Akin to a ‘dynamic R^2’, it calculates the percentage of variation in performance that can be attributed to present and past changes in each variable of equation (1). Recently, Srinivasan et al. (2010) used FEVD to show that specific offline (survey-based) metrics of the consumer purchase funnel (awareness, consideration and liking) are important sales drivers as compared to marketing activity.

As the final step, we verify the robustness of our estimated elasticities by changing the model specification. In particular, we replace the 3 content-specific metrics of social media WOM with (a) the weekly number of Facebook Likes, (b) the weekly number of total Tweets, (c) positive Tweets and (d) negative Tweets. The benefit of these alternative model specifications is twofold. Important to managers, we verify the robustness of our marketing and WOM-elasticities on business performance. Important to researchers, we compare the goodness-of-fit across alternative model specifications to show that the model with content-specific social media conversations fits better than the volume and valence metrics of Facebook and Twitter, currently the most popular social media portals. Specifically, we compare models based on the Akaike Information Criterion (AIC), which balances model accuracy with parsimony (Akaike 1974, Burnham and Anderson 2004).

4. Empirical Analysis

4.1. Data Description

The data comes from three sources. First, a major US retail brand provided marketing data, store traffic data and website traffic data. Second, Crimson Hexagon, a company that specializes in social media monitoring and analysis, provided WOM data. We collected natural search data from Google.com. The combination yields complete data for July 4th 2010 to July 31st 2011 (55 weeks). Table 1 shows summary statistics for the variables included in the model.

Table 1 shows substantial over-time variation in all variables of interest, which allows us to estimate their effects. As for measurement specifics, store traffic data was provided at the financial store level on a daily basis, and then aggregated to the national level on a weekly basis in order to be consistent with other measures. Television and radio advertisements were measured in Gross Rating Points (GRPs); circular advertising as the number of circulars distributed, and paid search advertising in impressions.

As for the online WOM data, we aim to obtain comprehensive coverage of different online platforms, including Twitter, Facebook, blogs, forum messages, etc. This approach fits the paper's purpose of providing a general, non-platform specific assessment of the performance effects of volume, valence and topic of eWOM. To illustrate the benefits of such comprehensive approach, we compare our model with alternatives using either Twitter or Facebook information.

Our data provider is Crimson Hexagon, Inc. , whose ForSight™ platform is based on Hopkins and King (2010). These authors developed an automated content analysis for social scientists, who are typically less interested in optimizing any individual document's classification (the goal of previous approaches) than in accurately generalizing about the population of documents, such as the proportion in a given category. By directly optimizing for this social science goal, Hopkins and King's (2010) method yields approximately unbiased estimates of category proportions even when the optimal classifier performs poorly. Their content analysis software is publicly available at the website of Harvard professor Gary King, (http://gking.harvard.edu/readme) The data provider indexes over 200 million social media posts per day, including blog posts[2], forum messages, Facebook posts, and the entire Twitter fire hose of Tweets (http://www.crimsonhexagon.com/new-data-sources-added-to-social-media-library/) .The technology does not rely on keyword counts, but on statistical classification algorithms that identify patterns of words used in conversations and then recognize information from the conversation that is relevant to the user's chosen topics. The information the platform provides includes positive and negative sentiment, key themes, the relative size of the themes and how they change over time. The reported accuracy rate on these classifications is 97% with a margin of error of +/- 3%. Thus, the technology's coding matches human coding in more than 90% of cases analyzed,. This accuracy rate is confirmed in independent tests by Pew Research Center, who "spent more than 12 months testing CH and its own tests comparing coding by humans and the software came up with similar results" (Pew Research Center 2013, p.2).

We chose three conversational themes based on of their expected differential effect on traffic and paid media's ability to impact the conversations. The first conversation bucket is about loving the brand (Love). Second were posts about going to the store or making a purchase (Went/Purchased). The former represent positive brand recommendations, the latter corresponds to an individual's actions influencing others' expected utility (Godes et al. 2005). The last

[2] The data provider starts its list with 8 public and 2 private blog directories, including www.globeofblogs.com, http://truthlaidbear.com, www.nycbloggers.com and http://dir.yahoo.com/Computers , a list of blogs provided by blogrolling.com, and 1.3 million additional blogs made available to them by Blogpulse.com. The technology then continuously crawl out from the links or "blogroll" on each of these blogs, adding seeds along the way from Google and other sources, to identify the target population of blogs.

conversation bucket is posts about the brand's advertisements or commercials (Ads). We display the time series for these conversation buckets in Figure 2.

Two observations stand out from Figure 2. First, all three conversation threads show spikes around major events, such as Black Friday (November 2010) and retailer actions (May and July 2011). However, the conversation threads do not always follow each other: 'Ad WOM' shows a much higher peak than the others for the second event, while 'Love WOM' peaks at the third event, and is trending upwards in the data period.

How do these content WOM metrics compare to the Facebook Likes and Tweets analyzed in recent literature? Figure 3 shows these times series, while Table 3 displays the correlation among WOM content, Facebook Likes and Tweets.

The most striking observation is the different patterns in Facebook Likes and Tweets; these metrics are virtually uncorrelated. Instead, positive and negative tweets are highly correlated with each other and with the 'Ad WOM' metric (> 0.85). These high correlations mean that we can not include both negative and positive tweets, or both Ad WOM and a Tweet variable in the same model. Instead, we replace these variables with each other to obtain alternative model specifications. Among the WOM content variables, all correlations are below 0.6, which indicate that multicollinearity is unlikely to pose problems in our proposed model.

Finally, we gathered 'natural search' data through Google Analytics, an online tool that provides data on search volume for key terms. Search data was provided in a weekly index of total number of searches through Google.com for the retailer, or the retailer's webpage. Natural search data showed a large spike around Christmas – a clear indication we should allow for seasonality, which we do through the seasonal retail mall index and the holiday dummy variable.

4.2. Results

4.2.1. Model Specification Test: Evolution, Cointegration and Granger Causality

The first analysis step consists of tests for Granger causality, evolution and cointegration. First, the Granger Causality tests[3] show substantial endogeneity among the variables of interest. For instance, TV is Granger causing all traffic and online WOM variables, but is itself also Granger caused by Love WOM conversations and by Store Traffic. We infer that managers observe the Love and Store Traffic bumps induced by a TV campaign, which motivates them to spend more on TV. Such 'performance feedback' (Dekimpe and Hanssens 1999) is also observed for Search Engine Marketing (SEM) and online traffic, and for Circulars and Love WOM. Finally, dual causality is also observed among WOM conversations (Love and Purchase). Overall, the results show substantial endogeneity among and within each type of variable (performance, online WOM, marketing). This indicates the need for a dynamic system model (such as the VAR) to explain each variable by the past of other variables.

How should each variable be included in the dynamic system: in levels or in differences? The evolution tests consistently showed that Online Traffic, Google Search, Advertisement conversations, Went/Purchased conversations, Television GRPs, Radio GRPs, and Circulars are all stationary. In contrast, Store Traffic, Love conversations, and Paid Search Impressions show the presence of a unit root (the series is evolving over time). Evolving performance and marketing variables have been demonstrated and interpreted in previous papers (e.g. Dekimpe and Hanssens 1999), but what does our new finding on Love Conversations mean? Changes to the amount of Love conversations are persistent; i.e. online Brand Love maintains itself at the

[3] Detailed results are available upon request; we do not include them here for space considerations.

new level without the need for further stimuli. Thus, if a marketing action succeeds in increasing online Love conversations for the brand, the social dynamics of online consumers keep it up.

The cointegration tests do not detect a long-term equilibrium among the evolving variables. Therefore, we include them in first differences in the model. This is important for the interpretation of our results: impulse response functions show the effect of a variable on Store Traffic Growth, and a significant effect on Store Traffic Growth thus implies a persistent effect on Store Traffic (see e.g. Trusov et al. 2009). The managerial interpretation is that (part of) the traffic gains persist into the future, most likely due to repeat purchase (Pauwels et al. 2002). Management judged it unlikely that such benefits would continue in the indefinite future, and agreed on a period of 13 weeks (i.e. one quarter) to accumulate persistent effects into a long-term elasticity, which can be compared to the long-term elasticities of stationary variables in the same period (Slotegraaf and Pauwels 2008).

4.2.2. Model Estimation

We estimated the VAR model in equation (1) with one lag to balance forecasting accuracy and model parsimony, with the acceptable model fit (Akaike Information Criterion) of 18.87. This model fit compares favorably to the model with Facebook Likes (19.58) and the model with Tweets (19.84). The model explains 0.39 of store traffic growth, which translates into a high explanatory power for store traffic. Figure 4 displays actual and predicted values for log of store traffic. Note that our model does not capture a few peaks in traffic; conversations with management revealed these coincided with an online coupon deal and a PR event. Rather than including these events ex post in our model, we stick with the a priori specification of equation 1, which in general is more than adequate in tracking store traffic.

4.2.3. Short-Term and Long-Term Traffic Elasticities of WOM and Paid Marketing

Based on the VAR-model, impulse response functions track the response of one endogenous variable to a shock on another variable. When the response variable is evolving and therefore included in first differences (e.g. store traffic growth), the accumulated impulse response function translates the effects back to the variable in levels (e.g. store traffic). To illustrate the difference in over-time effects of a marketing action and its related WOM content, Figure 5 displays the response of store traffic to TV ads versus ad-related WOM conversations. First, note that the effect in week 1 (the short-term impact of the shock) is substantially larger for Ad Conversations in Social Media than for TV advertising. Second, both TV ads and Ad conversation show carry-over effects that do not go to zero, as store traffic is an evolving variable. However, the effect of TV ads becomes insignificant after 2 weeks, while store traffic stays significantly higher than baseline more than 10 weeks after an impulse in Ad conversations. Adding each significant effect (the area under the curve) therefore yields a much higher long-term elasticity for Ad Conversations versus TV advertising.

Summarizing all impulse response functions, Tables 4 and 5 show respectively the Short-Term (same-week) and Long-Term (13 weeks, i.e. 1 quarter) elasticities.

The short-term store traffic elasticities of WOM conversations are between 0.13 and 0.26; which are all higher than any paid marketing effort, even Search Engine Marketing (0.07). Thus, our results reflect the notion that WOM has a higher traffic effect than paid marketing does. The first unique insight is that the content of the WOM conversation is key: doubling the WOM buzz about the company's ads yields only half the store traffic benefit (13%) of doubling of WOM buzz about actual store visits (26%). As for online traffic, WOM regarding love for the brand is most effective (0.27), with 'Purchase' and 'Ad' WOM tied in second place – all considerably higher than paid marketing actions. Reflecting the dual causality among WOM and performance (Trusov et al. 2009), higher store and online traffic in turn generates more word-of-mouth on all three topics. Interestingly, increase in WOM buzz still get the company less traffic than increase in natural (Google) search. The high elasticity likely reflects the self-selection of consumers who initiate search (Wiesel et al. 2011). Note however that part of this search is also driven by WOM, with elasticities between 0.24 and 0.38. Likewise, specific marketing actions are driving specific WOM topics, with TV and radio having an immediate impact on Purchase WOM, and Circulars on Ad and Love WOM. Of course, marketing-WOM effects may take more than a w eek to materialize, so we turn to the performance benefits in the longer term of 13 weeks (one quarter).

Long-term traffic elasticities show an even larger benefit of WOM conversations over marketing, reflecting the longer carry-over of WOM observed in Figure 5. Consistent with empirical generalizations (Tellis 2009), marketing communication elasticities of store traffic are between 0.042 and 0.108. The long-term store traffic benefits of WOM are an order of magnitude higher, from 1.635 for Ad WOM to 3.549 for Love WOM. In contrast, long-term WOM benefits on online traffic are considerably lower. They are identical to the short-term elasticity as we find no significant carry-over benefits. Search Engine Marketing has a similar long-term elasticity as Ad WOM and Purchase WOM, and only slightly less than Love WOM. The likely explanation is that visiting a physical store requires considerable effort from consumers, and WOM is particularly powerful in helping them decide the effort may be worth it. In contrast, the online store is just a click away, and a marketing push is all that is needed for many consumers to take a look (Verhoef et al. 2007, Wiesel et al. 2011)

In sum, the elasticity calculations show that a 1% change to a WOM creates a much higher traffic lift than a 1% change in marketing activity. Especially conversations about Love for the brand pay off, with a short-term elasticity 3 times higher, and a long-term elasticity 30 times higher than the most effective paid marketing effort. Of course, it may be a lot easier for managers to change paid marketing (e.g. spend 10% more) than to increase WOM conversations by the same amount. Likewise, some variables show a lot more changes than others, as evident from Table 2. Thus, we next analyze the importance of each variable's full variation (all changes) over the data period in driving performance variation.

4.2.4. Relative importance of WOM and paid marketing as a traffic driver

Forecast Error Variance Decomposition (FEVD) measures the relative importance of all present and past changes to a variable in driving traffic. Given our conceptual framework in Figure 1, we order the variables as paid marketing, social media conversations, Google search, online traffic, and store traffic, and obtain the decomposition after 13 weeks (one quarter). Figures 6 and 7 display the FEVD of respectively Store Traffic and Online Traffic.

Of all variation in store traffic, 44 % is driven by its own past changes; thus establishing the 'baseline' in the absence of other influences. The next big bucket is paid marketing, which drives 34.6% of store traffic. The three WOM conversations together account for 23.6% of the variation in traffic. Thus, while any given change in WOM has a high elasticity, the many changes to paid search impressions, circulars, radio and TV makes marketing responsible for a large part of variation in store traffic. As to individual variables, paid Circulars drive most store traffic variation, followed by Love WOM conversations. In contrast, online traffic and Google search explain only a small amount of store traffic variation.

Online traffic variation is mostly driven by own past changes (56% baseline), followed by Google Search (17.2%). Contrary to store traffic, we find that WOM drives more online traffic variation (14.1%) than paid marketing does (12%). This is intuitive as online WOM occurs in the same channel as the retailer's website, while most paid marketing is offline. Conversations about advertisements are the most important WOM driver of online traffic while TV is the main paid marketing driver.

In sum, the FEVD analysis shows a more nuanced picture of the total influence of WOM and paid marketing: both are an important driver of store and online traffic. Consistent with the elasticity results, Love of the brand is most important among the WOM conversations, while different marketing actions drive store and online traffic to a different degree. This leaves us with the question: to what extent do different marketing actions affect traffic directly, versus working indirectly by stimulating search and WOM conversations?

4.2.5. Indirect versus Direct Performance Effects of Paid Marketing

In addition to the direct effect marketing efforts have on traffic, they can also have an indirect effect through WOM or search. For example, as shown in Table 4, paid Circulars stimulate conversations about advertisements. These conversations then drive people to go to the store. Multiplying these effects, we obtain the indirect traffic effect of Circulars through Ad WOM. Table 6 shows the results of our comparison of these indirect effects.

Television, with a direct store traffic elasticity of 0.008, also obtains significant indirect effects by stimulating online traffic and Purchase WOM conversations, for a total effect of 0.012. Thus, the indirect effects are about half of the direct effect and constitutes 32% of the total effect. In contrast, radio only has the indirect effect through Purchase WOM, which adds 22% to the direct effect. These differences are likely due to consumer multitasking: 75% of Americans surf the web while watching television, while 73% are listening to radio while they are driving

(Shields 2010). Circulars work indirectly through online traffic, Love and Ad WOM, yielding 37% of that total effect. Circulars are present in consumers' homes, and thus may offer a visual stimulus for connected consumers to go to the retailer's website and engage in online WOM. Finally, search marketing works through all online and WOM variables, whose indirect effects exceed the direct effect on store traffic. This finding is consistent with Wiesel et al. (2011), who report that most benefits from paid Google advertising materialize in the offline channel. A likely explanation is that the low search costs stimulate consumers to first look online, after which they go offline to purchase in the physical store; also known as the 'research shopper' phenomenon (Verhoef, Neslin and Vroomen 2007).

On average, the direct store traffic effect of paid marketing is 62% of the total effect, and the highest indirect effect goes to WOM conversation concerning consumers going to the store and/or purchasing there. Thus, marketing effectiveness is underestimated if only the direct traffic effects are considered. A key goal and benefit of paid marketing is its power to stimulate conversations around a brand or product, which then causes a ripple effect that ultimately increases business performance. This complex web of effects is uncovered by the dynamic system nature of the model.

4.2.6. Alternative WOM Measures

As a final step, we compare the fit and findings of the proposed WOM content model with alternative specifications using variables currently popular in social media research. Past studies have used either Facebook.com measures (such as number of comments or likes), or the volume and valence of Tweets as a proxy for WOM. Given the correlations shown in Table 3, we estimate 4 alternative models; replacing the WOM content metrics by 1) daily comments and likes on Facebook.com, 2) total number of Tweets, 3) positive Tweets, 4) negative Tweets.

Across these models, Table 7 compares the store traffic elasticity estimates in the short-term (we obtain similar findings for long-term elasticities).

With the exception of the Facebook model, store traffic elasticities for each common variable are very close across alternative model specifications. TV, radio, circulars and paid search marketing have the same direct effect on store traffic, while the effects of online traffic and search show only minimal differences. In contrast, the different WOM metrics themselves show vast differences in effectiveness. In the Facebook model, neither Facebook likes nor comments have a significant effect on store traffic. In contrast, Tweets have significant effects on store traffic. However, their elasticity of 0.16 is lower than that for Purchase and Love WOM conversations. Including the sentiment of Tweets (either positive or negative as they are highly correlated) does not improve the model. As for model fit, the proposed model with WOM conversation topics has a better AIC (18.87) than the model with Facebook Likes (19.58), total Tweets (19.84), positive Tweets (20.02), and negative Tweets (20.24).

How about online traffic? Table 8 shows the short-term elasticities across models.

Online traffic elasticities of store traffic, natural search, radio and paid search impression are similar across model specifications. However, the proposed WOM content model shows a significant effect of TV and circulars and a higher elasticity of the WOM content metrics than Facebook likes (0.108) and Tweets (0.167). Thus, online traffic is different to the extent that Facebook Likes have a significant performance effect. However, it is similar to store traffic in the higher elasticity of conversation topics over other WOM metrics. Together with the better fit of our proposed model, we conclude that using the content metrics for WOM is preferred over using volume and valence metrics straight from Facebook or Twitter.

5. Conclusion

WOM is an important driver of business performance, and the popularity of online social media has made it easier than ever for consumers to engage in it, and for companies to track it. A key question though is which type of metrics are the best summary of the online WOM firehose. This paper presents an econometric analysis with rich data on online WOM volume, valence and content of conversation. We find that content-specific WOM performs better than Facebook and Twitter volume and valence metrics in explaining store traffic and online traffic for a major US retailer. Among the conversation topics, 'Love for the Brand' has the larger long-term traffic effects, but neutral conversations on 'Went There / Purchased' drive traffic in the short run.

Consistent with Trusov et al. (2009), WOM has longer carry-over and a much larger performance elasticity (up to 30 times higher) than the most effective paid marketing. However, companies have more control over changes to paid marketing. Our analysis shows that such marketing changes drive online WOM and are a key business driver.. Especially paid search, circulars and TV succeed in stimulating WOM conversations, while radio has mostly a direct effect of traffic and then on Went/Purchased WOM. A likely explanation is that consumers listen to the radio on their way to a shopping occasion: they either react directly by going into the store, or they do not react at all. Indeed, the retailer's circulars and TV ads are often discussed in online WOM, in contrast to its radio ads. Different advertising can thus induce specific conversations, and should be partly evaluated on its power to do so. This power may persist into the future: we find that Love WOM is evolving in our data period, and that several one-shot marketing campaigns create persistently higher conversations on this topic. Thus, paid marketing can help build (love for) the brand online by starting and perpetuating the conversation among consumers.

When comparing this model to previous studies that used other WOM measures such as Facebook likes/comments or Tweets, we find conversations have a greater ability to explain store and website traffic. Facebook Likes and Comments did not significantly affect Store Traffic, and increased Website Traffic to a lesser extent than do Tweets and WOM conversation topics. Tweets showed a higher store traffic elasticity than paid marketing, but still had an elasticity of only 60-80% of that of conversation topics. Beyond this superior prediction power, we believe there are also less tangible benefits for managers to track and understand exactly what consumers are saying about their product/brand. As another example from our WOM content provider, we pulled the data for Apple Ipad on October 4th 2011. While 54% of all Apple Ipad online mentions were negative, many of those concerned the data plan/provider instead of the Ipad itself. This insight is important for managers to diagnose the root cause of problems and to remedy them. In this case, Apple can work on fixing issues with data plan providers, and then relate these solutions back to consumers through social media and get their reactions. Thus, tracking social media content does not just inform companies about consumer perception of their business model and value chain, it helps them transform it.

The performance effects of specific content WOM topics deserve further discussion. Importantly, consumers appear not only influenced by conversations about love for the ad or the brand, but also by the knowledge that other consumers have purchased. This "observation of others" is an important and often overlooked aspect of WOM (Godes et al. 2005). In the physical world, such observations are limited to one's immediate surroundings. In contrast, the opportunities to observe friends' and strangers' behavior with social media tools appear endless.

Which role can companies play in this social media world? Godes et al. (2005) distinguish four possible roles: observer, moderator (e.g. by establishing online communities or referral systems), mediator and participant (Godes et al. 2005). They note that the last three roles involve perceived bias and ethical issues, which may form a key obstacle for firms contemplating getting out of the merely passive observer role. However, the current paper implies a fifth role: managers can use specific paid marketing actions to stimulate specific WOM content conversations. For instance, when an immediate sales boost is needed, radio drives store traffic and 'Purchased' WOM, which in turn quickly translates into traffic. In contrast, when longer-term benefits are important, circulars stimulate Love WOM, which has the highest long-term store traffic elasticity. Note the difference between this role of using paid marketing to stimulate WOM and the current practice of directly "seeding" WOM , which runs the risk of consumer backlash (Trusov et al. 2009).

We acknowledge several limitations of our empirical work. First, the data come from one large US retailer, and further research is needed to verify whether the specific findings generalize to other settings. Second, we did not have data on marketing by and online WOM for competitors. Such knowledge would increase the explanatory power of the model (Srinivasan et al. 2010) and enable us to investigate the effect of competitive marketing on the company's online WOM. Third, our metrics do not capture offline word-of mouth. This is no issue if online and offline WOM are highly correlated in volume, valence and content. It does become one if opinions diverge among offline and online WOM populations. Thus, an important question for future research is the extent to which volume, valence and content of online WOM are a good proxy for their offline analogs (Godes et al. 2005). Fourth, while our one-step-ahead forecasting shows good fit (figure 4), the data period of 55 weeks does not allow extensive hold-out sample tests. This increases uncertainty about the value of the estimated coefficients going forward. Therefore, we advise the data provider to regularly update the model and use the newly estimate coefficient values Fifth, we lack data on the retailer margin derived from store and online traffic through different acquisition channels. As a result, we could not investigate the interesting question of whether customers acquired with different WOM topics and marketing actions yield different monetary benefits (Villanueva, Yoo and Hanssens 2008). We note that the proposed model is more general than our specific substantive findings, as it has proven successful in analyzing effectiveness across countries and categories, and as it can readily incorporate data on competitors and margin differences across consumers. One limitation of the methodology is its reduced-form nature: it identifies and quantifies the dynamic interactions among paid marketing, WOM and performance based on past data. Extrapolation into the future is subject to the assumption that the basic data-generating process does not change. The second limitation of an econometric model is its focus on the 'what', 'when' and 'how much' of consumer actions instead of the 'why' of engaging in and reacting to WOM. Finally, the current paper explored several potential interactions rather than testing specific hypotheses or developing new theory. We hope these new findings stimulate further conceptual development and inquiry.

In sum, our quantitative analysis shows that WOM conversation topics are relevant for managers aiming to predict business performance and evaluate the indirect effects of paid marketing. As social media is engaging larger audiences across countries and age groups (Comscore 2012), its importance for marketing and business is only expected to grow.

References

Akaike, H . 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* **19** (6): 716–723.

Asur, S. , B.A. Huberman. 2010. *Predicting the Future with Social Media*. HP.

Burnham, K. P., D.R. Anderson. 2004. Multimodel inference: understanding AIC and BIC in model selection. *Sociological Methods and Research*. 33: 261-304.

Chevalier, J., D. Mayzlin. 2006. The Effect of Word of Mouth On Sales: Online Book Reviews. *Journal of Marketing Research*, 43 (August): 345-354.

Dekimpe, M., D. M. Hanssens. 1999. Sustained Spending and Persistent Response: a New Look at Long-term Marketing Profitability. J*ournal of Marketing Research*, 36 (November): 397-412.

Dellarocas, C. 2003. The digitization of word-of-mouth: promise and challenges ofonline feedback mechanisms. Management Science. 49(10), 1407-1424.

Dellarocas, C., X. Zhang, N. Awad. 2007. Exploring the value of online product reviews in forecasting sales: The case of motion pictures. Journal of Interactive Marketing, 21(4), 23-45.

Forrester Research. 2005. *What's The Buzz on Word-Of-Mouth Marketing? Social Computing and Consumer Control Put Momentum into Viral Marketing*. www.forrester.com, May.

Forrester Research. 2013. *Trends For The B2C CMO To Watch In 2013 – A Digital Disruption Report. Embrace Digital Innovation To Deliver Compelling Brand Experiences,* www.forrester.com, January 9.

Godes, D., D. Mayzlin. 2004. Using Online Conversations to Study Word-of-Mouth Communication. *Marketing Science* **23** (4): 545-560.

Godes, D., D. Mayzlin, Y. Chen, S. Das, C. Dellarocas, B. Pfeiffer, B. Libai, S. Sen, M. Shi, P. Verlegh. 2005. Firm's Management of Social Interactions. *Marketing Letters* **16** (3/4): 415-428.

Granger, C. W.J. 1969. Investigating Causal Relations by Econometric Models and Cross-Spectral Methods. *Econometrica* 37: 424–438.

Granger, C. W.J. P. Newbold. 1986. *Forecasting Economic Time Series*. Academic Press, 2nd edition, Orlando, FL.

Hanssens, D.M., L.J. Parsons, R. L. Schultz. 2001. *Market Response Models: Econometric and Time-Series Research*, 2nd ed. Kluwer Academic Publishers, Boston, MA.

Hopkins, D., and G. King. 2010. A Method of Automated Nonparametric Content Analysis for Social Science. American Journal of Political Science 54 (10, 229–247.

Johansen, S., R. Mosconi, B.Nielsen. 2000. Co-integration Analysis in the Presence of Structural Breaks in the Deterministic Trend. *Econometrics Journal* 3: 1-34.

Katz, E., P. F. Lazarsfeld. 1955. *Personal Influence; The Part Played by People in the Flow of Mass Communications*. Free Press, Glencoe, Ill.

Kotler, P. 1967. Marketing Management: Analysis, Planning and Control. Prentice-Hall, Englewood Cliffs, NJ.

Kwiatkowski, D., P. C. B. Phillips, P.Schmidt, Y. Shin. 1992. Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root: How Sure Are We that Economic Time Series Have a Unit Root?. *Journal of Econometrics* **54**(1-3): 159-178.

Lazarsfeld, P. F., Berelson, B., H. Gaudet. 1948. *The people's choice*. University Press, Columbia University, New York.
Liu, Y.. 2006. Word-of-Mouth for Movies: Its Dynamics and Impact on Box Office Revenue. *Journal of Marketing.* 70 (July): 74-89.
Lutkepohl, H. 1993. *Introduction to Multiple Time Series Analysis*. Springer-Verlag, Berlin.
Misner, I. R. 1999. *The World's Best Known Marketing Secret: Building Your Business with Word-of-Mouth Marketing*. 2nd ed. Bard Press, Austin, TX.
Money, B. R., M. C. Gilly, J. L. Graham. 1998. Explorations of National Culture and Word-Of-Mouth Referral Behavior in the Purchase of Industrial Services in the United States and Japan. *Journal of Marketing.* **62** (4): 76-87.
Nijs, V., M. Dekimpe, J.-B. E. M. Steenkamp, D. Hanssens. 2001. The Category Demand Effects of Price Promotions. *Marketing Science* **20** (1): 1-22.
Pauwels, K., D. Hanssens, S. Siddarth. 2002. The Long-term Effects of Price Promotions on Category Incidence, Brand Choice and Purchase Quantity. *Journal of Marketing Research* 39 (November): 421-439.
Pesaran, M. Hashem, R. G. Pierse, K. C. Lee. 1993. Persistence, Cointegration, and Aggregation: A Disaggregated Analysis of Output Fluctuations in the U.S. Economy. *Journal of Econometrics* 56 (March): 57-88.
Pesaran, M. H., and Y. Shin. 1998. Generalized Impulse Response Analysis in Linear Multivariate Models. *Economics Letters,* Vol.58: 17-29.
Pew Research Center (2013), "Twitter Reaction to Events often at Odds with Overall Public Opinion", March 4th, accessed March 5th at http://www.pewresearch.org/2013/03/04/twitter-reaction-to-events-often-at-odds-with-overall-public-opinion/
Shields, Mike (2010), "Simultaneous viewing and surfing commonplace", July 8th, accessed March 5th 2013 at http://www.adweek.com/news/television/simultaneous-viewing-and-surfing-commonplace-102771
Sims, C.A. 1980. Macroeonomics and Reality. *Econometrica* **48**(1): 1-48.
Sims, C.A., T. Zha. 1999. Error Bands for Impulse Responses. *Econometrica* **67** (5): 1113-1155.
Slotegraaf, R., K. Pauwels. 2008. The impact of brand equity and innovation on the long-term effectiveness of promotions. *Journal of Marketing Research* 45(June): 293-306.
Srinivasan, S., K. Pauwels, M. Vanhuele. 2010. Mindset Metrics in Market Response Models: An Integrative Approach. *Journal of Marketing Research* **47** (4): 672-684.
Tellis, G. 2009. Generalizations about Advertising Effectiveness in Markets. *Journal of Advertising Research* 49: 240-245.
Trusov, M., R. Bucklin, K. Pauwels. 2009. Effects of Word of Mouth versus Traditional Marketing: Findings for an Internet Social Networking Site. *Journal of Marketing* **73** (5): 90-102.
Verhoef, P., S. Neslin and B. Vroomen. 2007. Multichannel customer management: understanding the 'research-shopper' phenomenon. *International Journal of Research in Marketing* **24** (2): 129-148.

Villanueva, J., S. Yoo, D. Hanssens. 2008. The Impact of Marketing-Induced Versus Word-of-Mouth Customer Acquisition on Customer Equity Growth *Journal of Marketing Research* **45** (1): 48-59.
Wiesel, T., K. Pauwels, J. Arts. 2011. Practice-Prize Paper: Marketing's Profit Impact: Quantifying Online and Offline Funnel Progression. *Marketing Science* **30** (4): 604-611.
Yamamoto, H. and M. Matsumura. 2011. Marketing Ecosystem: The Dynamics of Twitter, TV advertising and Customer Acquisition. *Fifth International Association for Advancement of Artificial Intelligence Conference on Weblogs and Social Media*, June.

Table 1: Analysis steps in the Vector-Autoregression Modeling Approach

Methodological step	Relevant literature	Research question
1. Endogeneity and evolution tests		
Granger Causality Test	Granger (1969)	What is the temporal causality among variables?
Augmented Dickey-Fuller Test	Kwiatkowski, Phillips,	Are variables stationary or evolving?
KPSS unit root test	Schmidt, and Shin (1992)	Are the unit root results robust to null hypothesis?
Cointegration test	Johansen et al. (2000)	Are evolving variables in long-run equilibrium?
2. Model of dynamic interactions		
Vector Autoregressive model	Dekimpe and Hanssens	How do sales and marketing variables interact in
VAR in differences	(1999); Pauwels et al.	the long run and the short run, accounting for the
Vector Error Correction model	(2002)	unit root and cointegration results?
3. Policy simulation analysis		
Generalized Impulse response function	Pesaran and Shin (1998)	What is the immediate and dynamic effect of an impulse, without imposing a causal ordering?
Short-run (immediate) and long-run (cumulative) elasticity	Pauwels et al. (2002)	By how much % does performance change as a result of a 1% marketing change?
Forecast Error Variance Decomposition	Srinivasan et al. (2010)	How much of performance variation is explained by past marketing &WOM changes (dynamic R^2)
4. Robustness checks		
Replace proposed social media variables with alternatives	This paper	How stable are the model parameters to alternative word-of-mouth metrics?
Compare model fit with AIC	Akaike (1974)	Which word-of-mouth metrics yield the best fit?

Table 2: Descriptive Statistics of Variables included in model

	Mean	Median	Standard Deviation	Minimum	Maximum
Store Traffic	681821	682627	153469	407984	1203802
Online Traffic	110206	106476	33999	52653	269087
Google Search	36	35	13	14	100
Went/Purchased	753	645	401	358	2511
Ads	1736	1355	1112	634	5623
Love	538	460	290	230	1678
Television GRPs	8	7	7	0	24
Radio GRPs	3	0	7	0	20
Circulars	384089	0	897440	0	2908555
Paid Search Impressions	672267	645553	212767	283846	1268048
Total Tweets	7790	7276	3170	4056	3170
Positive Tweets	2964	1727	1052	1012	1052
Negative Tweets	1113	967	639	510	639
Face Daily Likes	5057	4298	3162	1267	18516

Table 3: Correlation among alternative online word-of-mouth variables

	d(Love WOM)	Purchase WOM	Ad WOM	Total Tweets	Positive Tweets	Negative Tweets
d(Love WOM)						
Purchase WOM	0.54					
Ad WOM	0.36	0.54				
Total Tweets	0.51	0.67	**0.86**			
Positive Tweets	0.45	0.55	**0.87**	**0.97**		
Negative Tweets	0.43	0.62	**0.97**	**0.91**	**0.89**	
Facebook Likes	0.07	-0.15	-0.06	0.10	0.20	-0.06

Table 4: Short-term (same-week) Elasticities and their standard errors*

	Store Traffic	Online Traffic	Google Search	Love WOM	Ads WOM	Purchased / Went WOM
Store traffic		0.259 (0.123)	0.343 (0.110)	0.649 (0.155)	0.671 (0.211)	0.844 (0.169)
Online traffic	0.159 (0.076)		0.479 (0.155)	0.506 (0.155)	0.677 (0.205)	0.420 (0.177)
Google Search	0.227 (0.073)	0.517 (0.110)		0.146 (0.144)	*0.167* (0.195)	*0.160* (0.177)
Love WOM	0.216 (0.052)	0.274 (0.084)	0.375 (0.072)		0.846 (0.133)	0.836 (0.161)
Ads WOM	0.126 (0.040)	0.207 (0.062)	0.241 (0.055)	0.475 (0.075)		0.415 (0.089)
Purchased/ Went WOM	0.256 (0.051)	0.208 (0.006)	0.375 (0.074)	0.762 (0.095)	0.674 (0.145)	
TV	0.008 (0.004)	0.007 (0.004)	*0.005* (0.006)	*0.007* (0.008)	*0.008* (0.011)	0.011 (0.009)
Circular	0.007 (0.002)	0.004 (0.004)	0.004 (0.003)	0.006 (0.005)	0.010 (0.007)	*0.003* (0.006)
Radio	0.014 (0.005)	*0.003* (0.008)	*0.005* (0.007)	*0.001* (0.010)	*0.002* (0.014)	0.015 (0.012)
SEM	0.070 (0.026)	0.082 (0.042)	0.089 (0.038)	0.154 (0.055)	0.201 (0.073)	0.095 (0.063)

* The elasticities that do not significantly differ from zero are indicated in *italics*

Table 5: Long-term Elasticities and their standard errors*

	Store Traffic	Online Traffic	Google Search	Love WOM	Ads WOM	Purchased / Went WOM
Store traffic	1.594 (0.152)	0.259 (0.123)	0.343 (0.110)	1.762 (0.422)	0.671 (0.211)	1.062 (0.399)
Online traffic	**2.138** (0.982)	1.575 (0.366)	1.211 (0.506)	6.572 (2.018)	0.680 (0.205)	0.420 (0.177)
Google Search	**3.057** (0.946)	0.787 (0.240)	2.103 (0.651)	9.686 (1.867)	2.259 (1.089)	1.327 (0.514)
Love WOM	**3.428** (0.820)	0.274 (0.084)	0.811 (0.348)	10.956 (1.318)	1.580 (0.667)	1.277 (0.387)
Ads WOM	1.635 (0.514)	0.207 (0.062)	0.588 (0.264)	6.181 (0.969)	1.527 (0.476)	0.708 (0.293)
Purchased/ Went WOM	3.334 (0.667)	0.208 (0.088)	0.698 (0.289)	9.911 (1.241)	1.645 (0.887)	1.675 (0.525)
TV	0.105 (0.050)	0.018 (0.014)	*0.005* (0.006)	*0.085* (0.106)	*0.010* (0.011)	0.011 (0.009)
Circular	0.088 (0.030)	0.004 (0.004)	0.004 (0.003)	0.149 (0.127)	0.010 (0.007)	*0.003* (0.006)
Radio	0.181 (0.063)	*0.003* (0.008)	*0.005* (0.007)	*0.205* (0.264)	*0.002* (0.013)	0.015 (0.012)
SEM	0.108 (0.062)	0.210 (0.136)	0.089 (0.039)	0.237 (0.081)	0.201 (0.074)	0.095 (0.059)

* The elasticities that do not significantly differ from zero are indicated in *italics*

Table 6: Decomposition of store traffic impact in direct and indirect marketing effects

	Store Traffic	Online Traffic	Google Search	Love	Ads	Went/ Purchased	Total Effect
Tv	68%	10%	0%	0%	0%	23%	0.012
Circular	63%	6%	8%	12%	11%	0%	0.011
Radio	78%	0%	0%	0%	0%	22%	0.018
SEM	38%	7%	11%	18%	14%	13%	0.187
Average	62%	6%	5%	7%	6%	14%	

Table 7: Short-term Store Traffic Elasticity Comparison across online WOM metrics*

	FB Daily Likes	Total Tweets	Positive Tweets	Negative Tweets	Conversation Topics WOM
Store Traffic	0.531	0.588	0.588	0.588	0.588
Online Traffic	x	0.144	0.148	0.145	0.159
Google	0.267	0.220	0.222	0.226	0.227
Love					0.216
Ads					0.126
Purchase					0.256
Likes/Comments	0.000				
Total Tweets		0.161			
Positive Tweets			0.139		
Negative Tweets				0.140	
TV	0.016	0.008	0.008	0.007	0.008
Circulars	0.007	0.007	0.007	0.007	0.007
Radio	0.015	0.013	0.014	0.014	0.014
Paid Search	0.220	0.064	0.065	0.067	0.070

*Insignificant elasticities are represented with an x

Table 8: Short-term Online Traffic Elasticity Comparison across online WOM metrics**

	FB Daily Likes	Total Tweets	Positive Tweets	Negative Tweets	Conversation Buckets
Store Traffic	x	0.242	0.244	0.239	0.259
Online Traffic	0.850	0.896	0.891	0.893	
Google	0.492	0.513	0.521	0.502	0.517
Love					0.274
Ads					0.207
Purchase					0.208
Likes/Comments	0.108				
Total Tweets		0.167			
Positive Tweets			0.150		
Negative Tweets				0.188	
TV	0.011	x**	x**	0.006	0.007
Circulars	x	x	x	x	0.004
Radio	x	x	x	x	x
Paid Search	0.120	0.076		0.072	0.082

**Television has a significant second period effect

Figure 1: Conceptual framework of marketing effects through WOM and Search

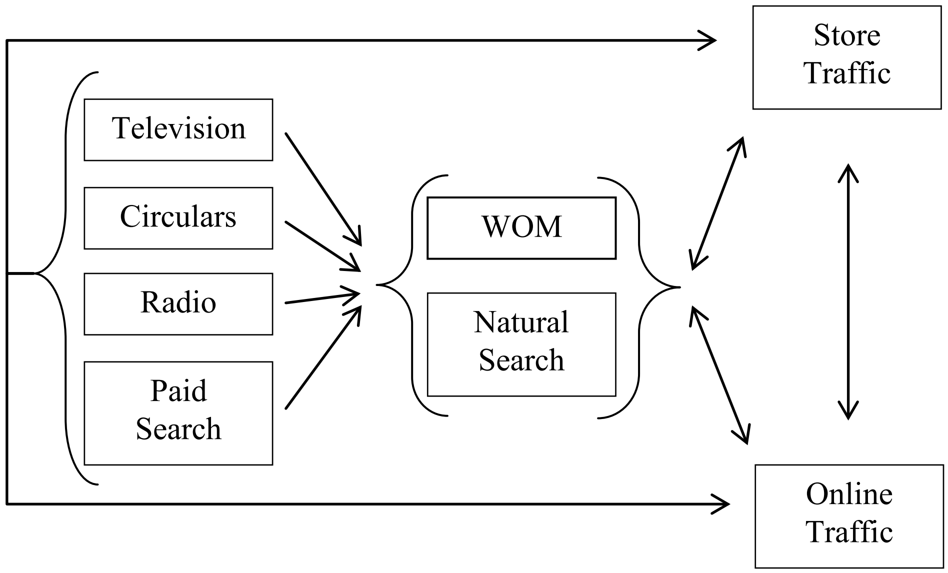


Figure 2: Content of social media word-of-mouth conversations

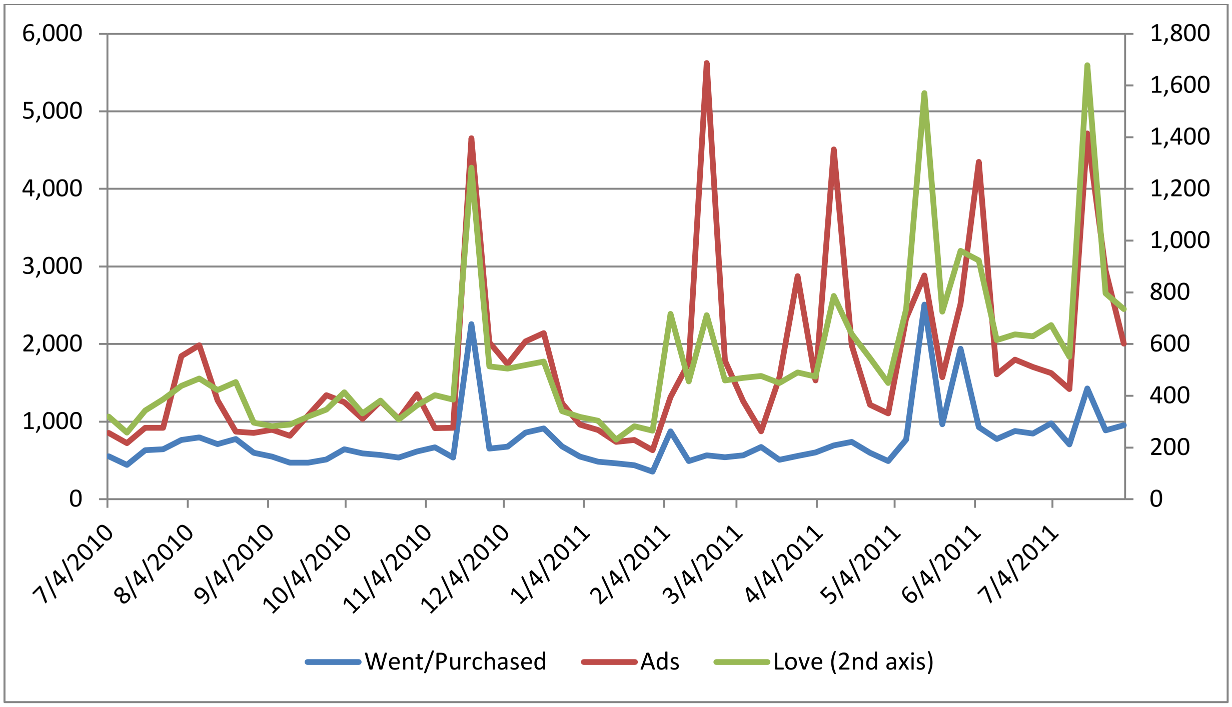


Figure 3: Facebook Likes and positive versus negative Tweets

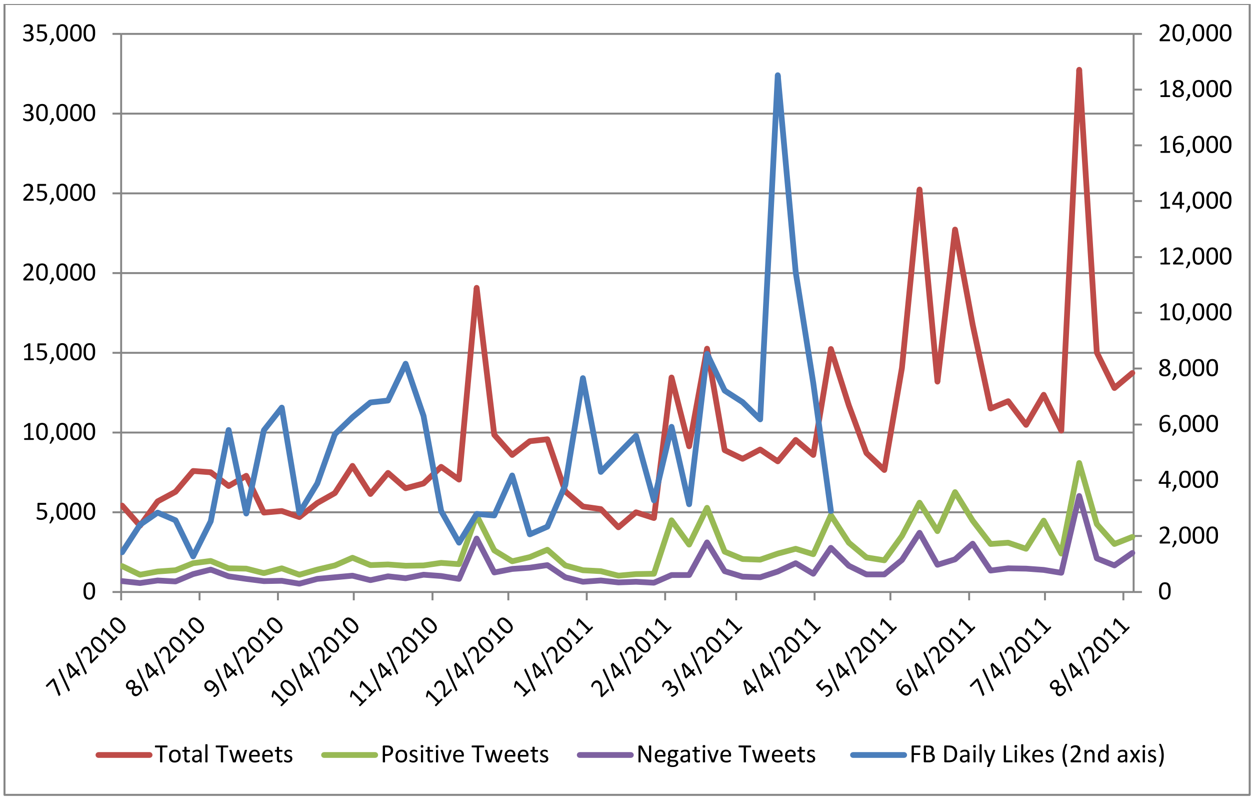


Figure 4: Predicted (blue) versus actual (green) values for log of store traffic

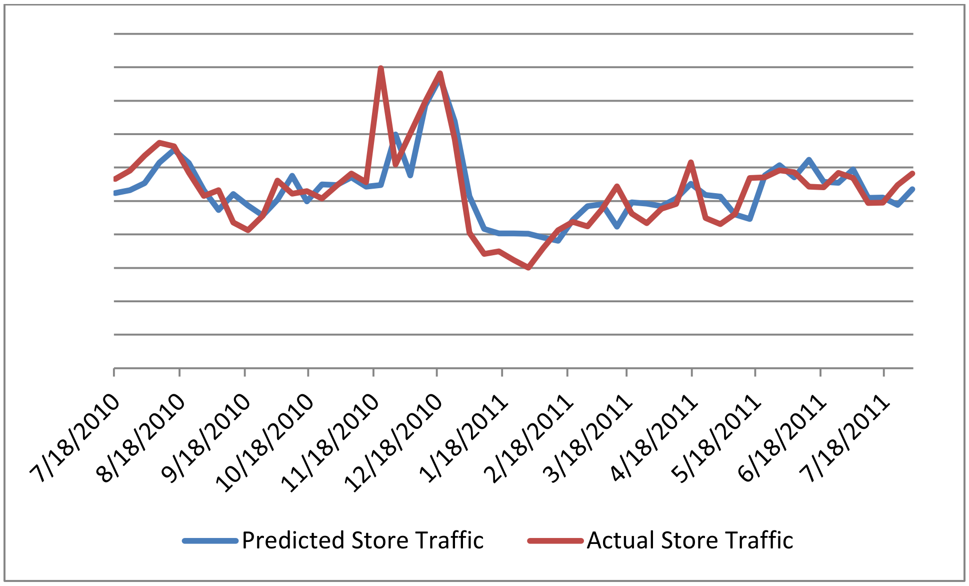
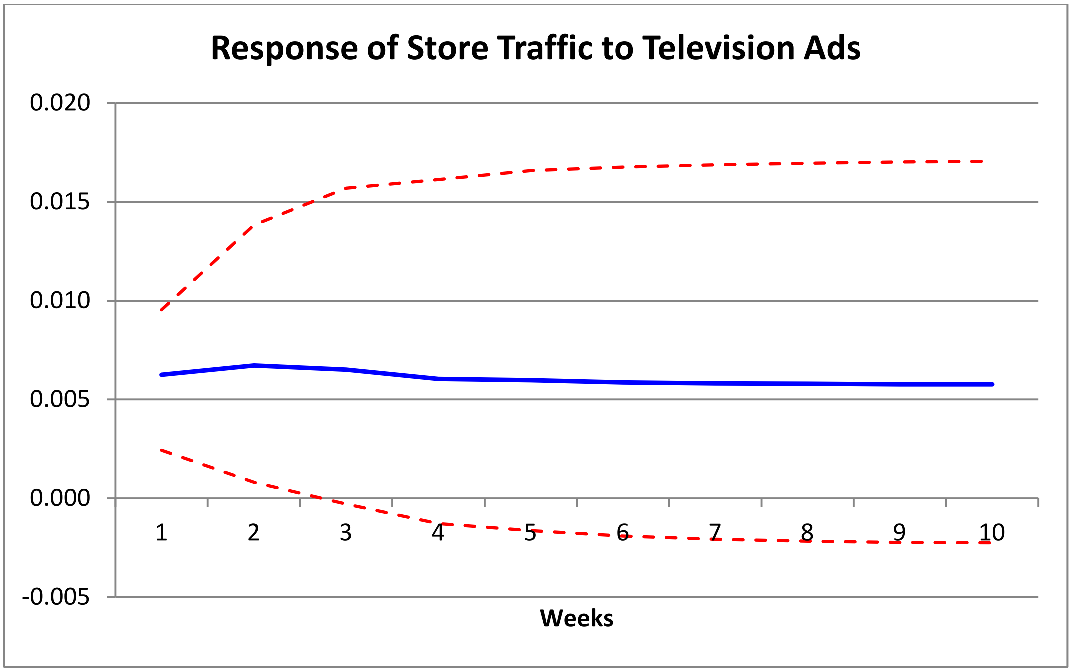


Figure 5: Response of Store Traffic to a shock in TV ads versus Ad conversations



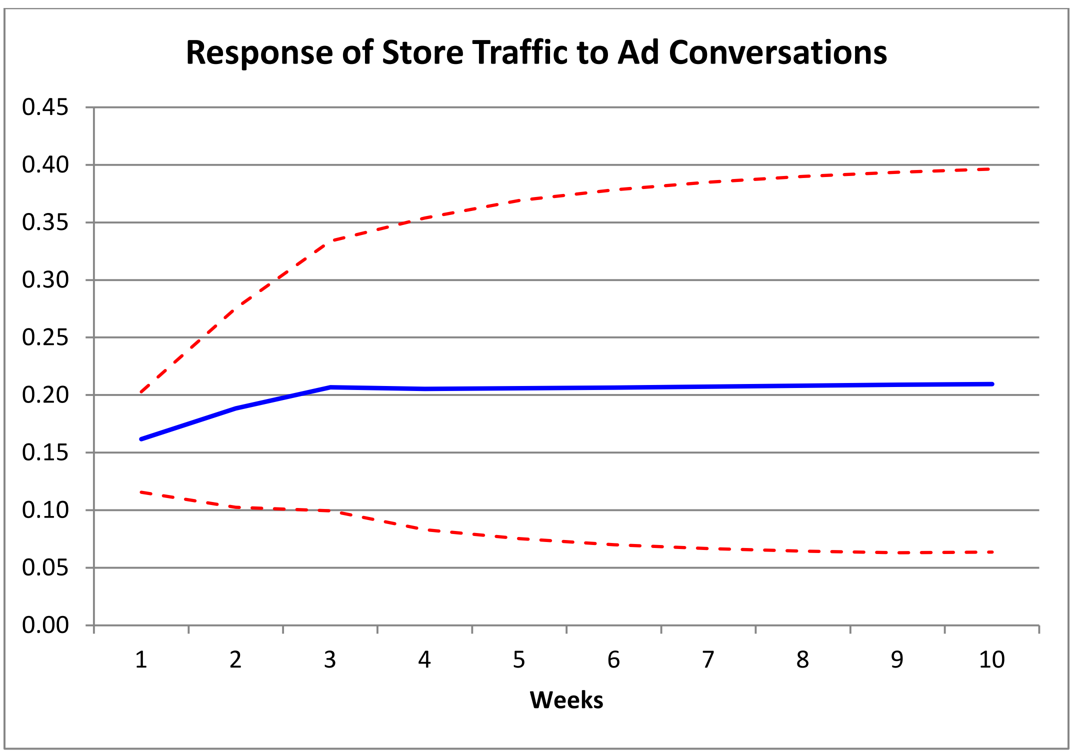


Figure 6: Forecast Error Variance Decomposition (FEVD) of Store Traffic growth

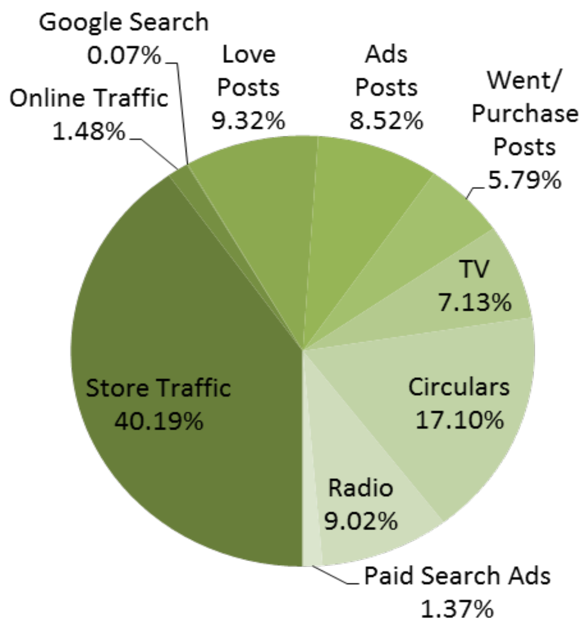


Figure 7: Forecast Error Variance Decomposition (FEVD) of Online Traffic growth