



## Marketing Science

Publication details, including instructions for authors and subscription information:  
<http://pubsonline.informs.org>

### A Dynamic Model of the Effect of Online Communications on Firm Sales

Garrett P. Sonnier, Leigh McAlister, Oliver J. Rutz,

To cite this article:

Garrett P. Sonnier, Leigh McAlister, Oliver J. Rutz, (2011) A Dynamic Model of the Effect of Online Communications on Firm Sales. Marketing Science 30(4):702-716. <https://doi.org/10.1287/mksc.1110.0642>

Full terms and conditions of use: <https://pubsonline.informs.org/Publications/Librarians-Portal/PubsOnLine-Terms-and-Conditions>

This article may be used only for the purposes of research, teaching, and/or private study. Commercial use or systematic downloading (by robots or other automatic processes) is prohibited without explicit Publisher approval, unless otherwise noted. For more information, contact [permissions@informs.org](mailto:permissions@informs.org).

The Publisher does not warrant or guarantee the article's accuracy, completeness, merchantability, fitness for a particular purpose, or non-infringement. Descriptions of, or references to, products or publications, or inclusion of an advertisement in this article, neither constitutes nor implies a guarantee, endorsement, or support of claims made of that product, publication, or service.

Copyright © 2011, INFORMS

Please scroll down for article—it is on subsequent pages



With 12,500 members from nearly 90 countries, INFORMS is the largest international association of operations research (O.R.) and analytics professionals and students. INFORMS provides unique networking and learning opportunities for individual professionals, and organizations of all types and sizes, to better understand and use O.R. and analytics tools and methods to transform strategic visions and achieve better outcomes.

For more information on INFORMS, its publications, membership, or meetings visit <http://www.informs.org>

# A Dynamic Model of the Effect of Online Communications on Firm Sales

Garrett P. Sonnier, Leigh McAlister

Red McCombs School of Business, University of Texas at Austin, Austin, Texas 78712  
{garrett.sonnier@mcombs.utexas.edu, leigh.mcalister@mcombs.utexas.edu}

Oliver J. Rutz

Yale School of Management, Yale University, New Haven, Connecticut 06521, rutz@uw.edu

Interpersonal communications have long been recognized as an influential source of information for consumers. Internet-based media have facilitated information exchange among firms and consumers, as well as observability and measurement of such exchanges. However, much of the research addressing online communication focuses on ratings collected from online forums. In this paper, we look beyond ratings to a more comprehensive view of online communications. We consider the sales effect of the volume of positive, negative, and neutral online communications captured by Web crawler technology and classified by automated sentiment analysis. Our modeling approach captures two key features of our data, dynamics and endogeneity. In terms of dynamics, we model daily measures of online communications about a firm and its products as contributing to a latent demand-generating stock variable. To account for the endogeneity, we extend the latent instrumental variable technique to account for dynamic endogenous regressors. Our results demonstrate a significant effect of positive, negative, and neutral online communications on daily sales performance. Failure to account for endogeneity results in a severe attenuation of the estimated effects. From a managerial perspective, we demonstrate the importance of accounting for communication valence as well as the impact of shocks to positive, negative, and neutral online communications.

*Key words:* word of mouth; Bayesian estimation; endogeneity; dynamics

*History:* Received: April 2, 2010; accepted: January 17, 2011; Eric Bradlow and then Preyas Desai served as the editor-in-chief and Michel Wedel served as associate editor for this article. Published online in *Articles in Advance* June 6, 2011.

## 1. Introduction

Interpersonal communication, or word of mouth, is recognized as an influential source of information for consumers. For example, Rist (2005) reports that interpersonal communication is the most influential information source among various types of media when making a purchase. As recently as 10 years ago, most interpersonal communication took place in person or over the telephone. Today, billions of people engage in conversations via Internet-based social media such as blogs and online social communities. From the perspective of consumers, online communications are scalable. Rather than one-to-one or one-to-few, online communications tend to be one-to-many or many-to-many, facilitating information exchange and the diffusion of user-generated content (Godes et al. 2005). As the popularity of social media increases with consumers, firms are increasing spending on word-of-mouth marketing. From 2007 to 2008, spending on word-of-mouth marketing rose 14.2% to \$1.54 billion and is expected to reach \$3 billion by 2013 (PQ Media 2009).

From the perspective of the firm, online communications are observable. Observability facilitates the measurement of online communications, which enables the estimation of the impact of online communications on sales and other quantities of interest to the firm. As consumers are influenced by interpersonal communication, there is reason to expect that online communications will influence consumer decisions, especially for search goods (Giese et al. 1996). Regarding valence, consumer behavior theory argues that negative communications may be more diagnostic and thus have greater impact compared with positive communications (Herr et al. 1991). However, behavioral research has also identified situations where positive communications may be more diagnostic. For example, Ahluwalia (2002) finds that positive information is more persuasive in the case of a familiar brand. Although theory points to a potentially stronger effect for positive and negative communications because of their diagnostic nature, communications scored as neutral may help spread awareness and inform consumers about the installed

base of the firm's products. Higher awareness can lead to adoption by innovators, whereas information about the installed base can lead to adoption by imitators (Bass 1969).

Much of the research addressing the impact of online communications has restricted consideration to ratings. Ratings for products such as books and movies are widely available on sites such as Amazon.com or Yahoo! Movies and are also inherently quantitative. Chevalier and Mayzlin (2006) find that differences in the number of ratings (volume) and the average rating (valence) across online book retailers affect relative sales. Dellarocas et al. (2007) demonstrate that the volume and valence of online customer movie ratings are related to future box office revenues. However, other studies have found mixed results in terms of the effect of ratings volume and valence. Duan et al. (2008) find that rating valence does not affect movie box office revenues directly, but rather indirectly through an effect on the volume of ratings. Chintagunta et al. (2010) find that ratings valence explains opening-day movie box office revenues, whereas the volume and variance of ratings does not.

Moe and Trusov (2011) argue that the information content in online ratings may be limited by social dynamics in the ratings environment. They find that consumers' ratings are influenced by previously posted ratings and that the resulting sales effect resulting from this social dynamic is small. Some empirical work addresses the effects of online communications more broadly defined than ratings data collected from a single website or small set of websites.<sup>1</sup> Dhar and Chang (2009) find that the overall volume of blog posts about a music album across the Internet predicts the album's sales rankings. Their data do not contain the valence of the blog posts. A problem with obtaining the valence information in this setting is the sheer magnitude of the data. Godes and Mayzlin (2004) navigate around this problem by restricting their data gathering to comments posted in a single online community and hand-coding sentiment for a random sample of 10% of their data. They find that the dispersion of online conversations regarding new television programs across Usenet forums predicts the program's Nielsen ratings, whereas the volume and valence of conversations do not. Similarly, Liu (2006) restricts his data gathering to text reviews posted on the Yahoo! Movies website. He also hand-codes valence for his sample of text reviews

and finds that volume, but not valence, explains box office revenue.<sup>2</sup>

A central tension in the investigation of the effect of online communications beyond ratings is that considering a wider swath of online communications exacerbates the problem of coding valence. In response, researchers looking beyond ratings have tended to restrict their attention to a single website or online forum when valence is of interest (Godes and Mayzlin 2004, Liu 2006). Contrary to the normative findings of the consumer behavior literature, these studies have not found an effect for valence. The burden of collecting online communications from a broader set of sources than a single website or forum and subsequently hand-coding valence has recently been eased by what a recent *New York Times* article refers to as "sentiment analysis" (Wright 2009). For example, firms such as Visible Technologies and Scout Labs measure the volume and valence of online communications across a variety of social media, including blogs, chat rooms, news sites, YouTube, and Twitter. Sentiment analysis tools hold the promise of assisting companies struggling to assimilate and understand the torrent of communications about their products and services.<sup>3</sup>

The intended contribution of our research is both substantive and methodological. Substantively, we investigate the effect of the volume of positive, negative, and neutral online comments, captured by Web crawler technology and classified by automated sentiment analysis, on daily sales performance. Compared with previous research, our data provide a more comprehensive view of online communications regarding our firm. We are among the first to study the effect of online communications with valence classified by automated sentiment analysis and the first to show that positive, negative, and neutral comments all affect firm sales performance. Methodologically, our modeling approach captures two key features of our online communications data, dynamics and endogeneity. Previous research has shown that the effect of online communications may be dynamic (Godes and Mayzlin 2004, Liu 2006). We

<sup>1</sup> Outside the realm of online communications, Luo (2007, 2009) uses data on customer complaints filed with the U.S. Department of Transportation to show that the number of complaints about an airline is associated with lower stock returns.

<sup>2</sup> In a recent working paper, Tirunillai and Tellis (2010) investigate the effect of online reviews on stock returns. Similar to Liu (2006), they code text reviews from Amazon, Yahoo!, and Epinions as positive or negative. They find an effect for the total volume of reviews and for negative reviews, with the former being stronger.

<sup>3</sup> Two recent working papers make use of automated sentiment analysis data. Shin et al. (2010) find that positive (negative) online comments are a leading indicator of future price increases (decreases) for low price and follower brands (all brands) of MP3 players. Onishi and Manchanda (2010) investigate the effect of blogging during new product introduction and find a positive sales effect for the cumulative volume of blogs and the number of positive blogs but no effect for the number of negative blogs.

view daily measures of online communications about a firm and its products as contributing to a latent dynamic demand-generating stock variable. Similar to goodwill models of advertising (e.g., Nerlove and Arrow 1962, Bass et al. 2007), our model demonstrates a parsimonious means of allowing information contained in past daily online communications to affect sales. Whereas goodwill advertising models have considered the potential endogeneity of advertising in the stock equation (Bass et al. 2007), because of a lack of appropriate instruments, the problem has not been addressed. In the context of our problem and data, endogeneity concerns loom large because of measurement error, omitted variables, and simultaneity. Further complicating matters, our measure of potentially endogenous online communications is multivariate and dynamic. Our methodological contribution stems from the need to address this challenge. We do so by extending the latent instrumental variable (LIV) method (Ebbes et al. 2005, Zhang et al. 2009) to account for a dynamic multivariate endogenous regressor. We are among the first to use the LIV method to address endogeneity and the first to consider the method in the context of time-series data.

Our data are provided by a well-known technology firm selling a variety of durable goods in an online channel. The firm's products could be characterized as search goods because product features and characteristics are easily evaluated prior to purchase (Nelson 1970). The firm partners with a leading provider of social media-monitoring tools to collect data on the number of times the firm and its products are mentioned on the Internet. The comments in which the firm is mentioned are classified via a proprietary text mining algorithm as positive, negative, and neutral. Rather than ratings data from a single website such as Yahoo! Movies (Duan et al. 2008) or postings from a single forum such as Usenet (Godes and Mayzlin 2004), our data capture a broader view of the online communications about the firm and its products on the Internet. Of interest to managers and academics is the question of whether or not such data explain sales performance.

We use the online communications data along with daily sales data to estimate our model, employing a Bayesian forward-filtering, backward-sampling algorithm to estimate the dynamic parameters. Log Bayes factors and step-ahead forecasts strongly support our proposed model. Our results indicate that positive and neutral comments increase the dynamic stock, whereas negative comments decrease it. Furthermore, the magnitude of the effects is larger for positive and negative online comments compared with neutral online comments. In terms of dynamics, the duration of the effect of a shock to online communications is estimated to be slightly longer than one week. We

find significant correlations in the residual terms of the stock and online communications equations, and we demonstrate that ignoring this endogeneity leads to severely biased estimates of the effects of online communications.

We demonstrate the managerial relevance of our findings by showing the importance of comment valence. We contrast our model with that of an alternative model that uses total comment volume aggregated across valence. A substitute for costly automated sentiment data is volume data available from public sources (e.g., Technorati). We aggregate the online communications into a count of total comment volume analogous to the volume data available from public sources and reestimate our proposed model. We find that the model that accounts for valence provides superior fit. Perhaps more importantly, we find that the effect of online communication on sales is masked when using comment volume aggregated over valence. We also use the model to demonstrate the magnitude of the effect of shocks to positive, negative, and neutral comments. Our results suggest that firms seeking to initiate positive and neutral communications as part of marketing campaign may be justified in doing so (Godes and Mayzlin 2009). Our results also suggest that firms should proactively manage situations leading to negative comments, as such comments can result in a substantial negative impact on sales.

The remainder of our paper is structured as follows. In §2, we present our dynamic model of the effect of online communications on sales. In §3, we discuss some pertinent features of our data and the estimation results. In §4, we discuss some managerial implications of our results. In §5, we summarize and conclude.

## 2. Model

Sales response models have a rich history in the marketing literature. Recently, researchers have considered sales response models through the lens of dynamic linear models (Van Heerde et al. 2004, Bass et al. 2007, Naik et al. 2010). Building on these models, we treat daily sales observations as a function of a latent, dynamic demand-generating stock of information and exogenous covariates (Naik et al. 2010, Rutz and Bucklin 2011). The information stock, in turn, is affected by the previous period's stock and online communications. Let sales and the information stock at time  $t$  be denoted by  $y_t$  and  $S_t$ , respectively. Let the  $K$ -dimensional vector  $x$  contain the exogenous covariates. Finally, let positive, negative, and neutral online communications be denoted by  $c^p$ ,  $c^n$ , and  $c^o$ , respectively. To motivate the specification of our model, consider a model of sales as a function of current and

past online communications. Such a model could be expressed as

$$y_t = x_t' \beta + c_t' \lambda + \sum_{l=1}^L c_{t-l}' \lambda_l + \varepsilon_t, \quad (1)$$

where  $L$  denotes the lag length,  $c_t = [c^p \ c^n \ c^o]_t'$ ,  $[\beta' \ \lambda']'$  are parameters to be estimated, and  $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$  is an error term.

To estimate Equation (1), the researcher must specify  $L$ . With daily time-series data, the number of appropriate lags is potentially large, creating a proliferation of parameters. Lagged regressors are also typically correlated, leading to multicollinearity problems in estimation. Consistent with this issue, the significance of different lagged terms is often sensitive to  $L$ . To allow for a more parsimonious treatment of dynamics, we specify the sales model as

$$\begin{aligned} y_t &= x_t' \beta + S_t + \varepsilon_t, \\ S_t &= \delta S_{t-1} + c_t' \gamma + \omega_t^S. \end{aligned} \quad (2)$$

In Equation (2), sales is a function of exogenous covariates  $x$  and a latent information stock,  $S_t$ . The scalar parameter  $\delta$  measures the carryover effect in the stock equation, and the parameter vector  $\gamma$  measures the effect of positive, negative, and neutral online communications on the information stock  $S_t$ . Assuming  $\omega_t^S \sim N(0, \sigma_\omega^2)$ , the model can be estimated as a Bayesian dynamic linear model (DLM) with the equations for  $y_t$  and  $S_t$  as the observation and state equations, respectively (West and Harrison 1997).

The model specification in Equation (2) accounts for dynamics in the effect of online communications on sales by allowing past online communications to accumulate in a latent stock of demand-generating information. From our perspective, this is appropriate as it seems unlikely that daily online communications are direct calls to action. Given that our technical firm sells durable search goods, consumers are likely seeking information about product features and characteristics prior to purchase. Current and past online communications about the firm and its products can be a source of relevant information for consumers to use during their search process. In a similar setting, Naik et al. (2010) argue that corporate advertising (i.e., advertising for Ford as a brand versus Ford Explorer as a product) does not directly generate sales revenue but rather contributes to a stock of corporate goodwill. Our expectation is that positive online communications will replenish the demand-generating stock, whereas negative online communications will diminish it. With respect to neutral online communications, we argue that they may help spread awareness and product information or inform consumers about the installed base of the firm's products, which may replenish the stock variable (Bass 1969).

In terms of estimation, the full conditional distribution of the dynamic state variable  $S_t$  can be sampled via the forward-filtering, backward-sampling algorithm. Conditional on the current draw of  $S_t$  and assuming online communications are *exogenous*, standard Bayesian regression steps can be used to sample from the full conditional distribution of the  $\delta$  and  $\gamma$  parameters. However, estimation of the  $\gamma$  parameters via standard regression techniques will result in biased estimates if online communications are *endogenous*. We argue that there are three reasons to be concerned about the potential endogeneity of online communications: measurement error, omitted variables, and coevolution between the information stock and online communications.

Measurement error seems to be a likely source of endogeneity when using Web crawler technology to measure online communications. For example, if consumers use jargon or abbreviations (e.g., MS to refer to Microsoft) to refer to the company and its products, the crawler technology may not capture the communications. The crawler technology may also erroneously classify unrelated communications as relevant. Although exclusion terms can be included in the crawler search (e.g., search for Ford, exclude Harrison Ford), measures of online communication are unlikely to be error-free. More importantly, it is very unlikely that either human or automated classification of the valence of an online communication is error-free (Wright 2009). It is well known that measurement error induces correlation between the regression residual and the mismeasured regressor (Greene 2000). In terms of omitted variable bias, it is possible that unobserved firm activity affects goodwill and is also correlated with online communications. Firms have become more engaged in participating in online communications with customers (Godes et al. 2005) and even seeding online communications (Godes and Mayzlin 2009). It is likely that unobserved activity, such as advertising or promotion, coincides with online communications, generating correlation between the stock equation error term and measures of online communication.<sup>4</sup> Finally, it is possible that the information stock and online communications are codetermined. We note that feedback

<sup>4</sup> In most cases, advertising data are available on a weekly or monthly level, but not daily. For example, Taylor Nelson Sofres, the provider of the Leading National Advertisers data, offers weekly advertising data, but not daily. This creates an inherent trade-off for the researcher. The data could be aggregated to a level for which marketing mix data (e.g., advertising) are available. However, daily variation in online communications and sales would be ignored, and the effect of online communications on sales may suffer from aggregation bias. However, if the marketing mix data are not available at the finer level of temporal aggregation, omitted variable bias becomes a greater concern.

effects have been found in previous research considering ratings and online word of mouth (Godes and Mayzlin 2004, Duan et al. 2008).

Empirical applications of advertising goodwill models have considered the possible endogeneity of variables in the stock equation, but they have not addressed this issue because of a lack of instruments (Bass et al. 2007). Indeed, without appropriate instruments, there is little that can be done. Finding appropriate instruments is nontrivial. In the case of endogenous prices in demand models for products with commodity inputs, for example, input prices may be a suitable instrument (Kuksov and Villas-Boas 2008, Musalem et al. 2008). For daily online communications, this is less applicable. The most obvious candidate for an instrument is lagged online communications. However, if online communications are measured with error, or if past online communications are correlated with the contemporaneous values of the omitted variable, lagged values are invalid as an instrument (Angrist and Krueger 2001).<sup>5</sup> Weak or invalid instruments can result in parameter estimates that are more biased than those obtained by simply ignoring the endogeneity problem (Stock et al. 2002, Zhang et al. 2009).

To address the endogeneity of online communications in the stock equation, we make use of the LIV technique (Ebbes et al. 2005). As with frugal instrumental variable techniques, which exploit the moments of the distribution of the endogenous regressor, the LIV approach alleviates the problem of finding valid instruments. The LIV approach functions much like the standard instrumental variable approach that expresses an endogenous regressor as a function of some observed instrument presumed to be correlated with the endogenous regressor but orthogonal to the error term. In the LIV approach, a latent variable model is used to account for dependencies between the endogenous covariate and the error by introducing unobserved discrete binary variables. These latent variables are used to decompose the endogenous covariate into a systematic part that is uncorrelated with the error and one that is possibly correlated with the error. This allows for unbiased estimation of the effect of endogenous covariates, such as online communications, in the stock equation.

To elaborate our approach, we begin by expressing the vector of communications at time  $t$  as follows:

$$\begin{bmatrix} c^p \\ c^n \\ c^o \end{bmatrix}_t = \begin{bmatrix} \tilde{c}^p \\ \tilde{c}^n \\ \tilde{c}^o \end{bmatrix}_t + \begin{bmatrix} \omega^p \\ \omega^n \\ \omega^o \end{bmatrix}_t, \quad (3)$$

<sup>5</sup> Note that such a problem could also plague the standard vector autoregression approach as the impulse response function could be biased by the errors in variables.

where  $\tilde{c}_t = [\tilde{c}^p \ \tilde{c}^n \ \tilde{c}^o]_t'$  is the systematic component of the communication variables and is independent of  $\omega_t^s$ , the goodwill residual, and  $\omega_t^c = [\omega^p \ \omega^n \ \omega^o]_t'$ , the random component of the communication variables. Let  $\omega_t = [\omega^s \ \omega^c]_t'$ . We assume  $\omega_t$  follows a multivariate normal distribution,  $\omega_t \sim \text{MVN}(0, \Omega)$ , where  $\Omega$  is a full  $4 \times 4$  matrix. The three off-diagonal elements in the first row of  $\Omega$  are the covariance terms between the goodwill and online communication residuals. Nonzero covariance terms imply the conditional distribution of  $\omega_t^s$ , given  $c_t = [c^p \ c^n \ c^o]_t'$ , depends on  $c^p$ ,  $c^n$ , and/or  $c^o$ . Thus, estimation of the structural parameter  $\gamma$  (the effect of online communications on the information stock) via a standard Bayesian regression step will result in biased estimates (Rossi et al. 2006).<sup>6</sup>

If we observe a set of variables that are correlated with online communications but orthogonal to  $\omega_t^s$ , we can use these observed instruments in estimation. Denote such an instrument for online communications by the vector  $z_t = [z^p \ z^n \ z^o]_t'$ . We may then express the systematic portion of online communications as

$$\begin{bmatrix} \tilde{c}^p \\ \tilde{c}^n \\ \tilde{c}^o \end{bmatrix}_t = \begin{bmatrix} z_t^{p'} \alpha^p \\ z_t^{n'} \alpha^n \\ z_t^{o'} \alpha^o \end{bmatrix}, \quad (4)$$

where the  $\alpha$  parameters are to be estimated. As noted above, such instruments are difficult to observe in general and especially in our context of daily multivariate observations. To overcome this issue, Ebbes et al. (2005) propose a latent discrete binary variable to decompose a single distributed endogenous regressor into a systematic part uncorrelated with the error term and a part potentially correlated with the error term. Zhang et al. (2009) extend this approach to the case of multiple endogenous regressors. Following Zhang et al. (2009), we can express the systematic part of our vector of endogenous online communications as a function of a latent binary instrument,  $v$ , that divides the online communication measures into discrete categories and  $\lambda$ , the category parameters:

$$\begin{bmatrix} \tilde{c}^p \\ \tilde{c}^n \\ \tilde{c}^o \end{bmatrix}_t = \begin{bmatrix} v_t^{p'} \lambda^p \\ v_t^{n'} \lambda^n \\ v_t^{o'} \lambda^o \end{bmatrix}. \quad (5)$$

We call the model described by Equation (5) the *categorical LIV model*. Each latent instrument  $v_t^l$  (for

<sup>6</sup> We use the term structural to differentiate between the parameters in Equation (2) versus the reduced-form representation achieved by substituting for the endogenous regressors, according to the researcher's specification for  $\tilde{c}_t$ .

$l = p, n, o$ ) follows a  $J$ -dimensional multinomial distribution with probabilities  $\{\pi_1^l, \dots, \pi_J^l\}$ , where  $\pi_j^l$  is the probability that the  $j$ th latent instrument for the  $l$ th type of online communications is one. The properties of the LIV estimator are discussed in Ebbes et al. (2005). As discussed therein, the model is identified by the likelihood, with the requirement that  $J \geq 2$ . For any of the endogenous regressors, with  $J = 1$  the systematic portion is constant, and there is no information with which to estimate the structural parameter. It follows that as any of the  $\pi_j^l$  approach 0 or 1, the latent instrument is weakened in its ability to identify the structural parameter.

Ebbes et al. (2005) apply the categorical LIV model to the classic problem of estimating the effect of education on earnings. They find that ordinary least squares estimates suffer from an omitted variable bias ably controlled for by the categorical LIV approach and that the approach is robust to misspecification of the number of categories. Zhang et al. (2009) apply the categorical LIV method to account for endogeneity in a mediation analysis of the effect of attention to feature advertisements on sales. The data sets used in both these studies make use of *cross-sectional* variation. In contrast, our daily sales data are *time-series* data.

A concern with the categorical LIV model for our application is that the behavior of the estimator in the presence of a dynamic endogenous variable is not well understood. If the latent categories fail to capture much of the correlation with the endogenous regressor, there will be less information to identify the structural parameter. We propose an alternative specification of the LIV model to account for dynamic endogenous regressors. Specifically, we model the endogenous online communications as a function of a dynamic state variable uncorrelated with the stock residual  $\omega_i^s$ . Let the systematic portion of online communications be

$$\begin{bmatrix} \tilde{c}^p \\ \tilde{c}^n \\ \tilde{c}^o \end{bmatrix}_t = \begin{bmatrix} \theta^p \\ \theta^n \\ \theta^o \end{bmatrix}_t, \quad (6)$$

where  $\theta_t = [\theta^p \ \theta^n \ \theta^o]_t'$  is a dynamic state matched to the observed vector of online communications as in Equation (3). The state equation for  $\theta$  is

$$\theta_t = \Upsilon \theta_{t-1} + \xi_t, \quad (7)$$

where  $\Upsilon$  is a diagonal matrix of parameters and the vector of residual terms  $\xi_t$  is distributed as multivariate normal with full covariance matrix  $\Psi$ ,  $\xi_t \sim$

$MVN(0, \Psi)$ . We refer to the model described by Equations (6) and (7) as the *dynamic LIV model*. The model hierarchy for our dynamic LIV model is

$$\begin{aligned} y_t &| x_t, \beta, S_t, \sigma_\varepsilon^2, \\ \beta &| y_t, x_t, S_t, \sigma_\varepsilon^2, \\ \sigma_\varepsilon^2 &| y_t, x_t, S_t, \beta, \\ S_t &| y_t, x_t, c_t, \beta, S_{t-1}, \theta_t, \gamma, \delta, \Omega, \\ c_t &| \theta_t, \Omega, \\ \gamma &| c_t, S_t, S_{t-1}, \delta, \Omega, \\ \delta &| c_t, S_t, S_{t-1}, \gamma, \Omega, \\ \Omega &| c_t, S_t, S_{t-1}, \gamma, \delta, \theta_t, \\ \theta_t &| \theta_{t-1}, \Upsilon, \Psi, \\ \Upsilon &| \theta_t, \theta_{t-1}, \Psi, \\ \Psi &| \theta_t, \theta_{t-1}, \Upsilon. \end{aligned} \quad (8)$$

The model is completed by proper but diffuse prior distributions for the initial conditions of the two dynamic state variables, as well as priors for other model hyperparameters. Details are included in the appendix.

In the categorical LIV model, the instrument is specified via latent categories and the category parameters. In the dynamic LIV model, the latent instrument is specified via an autoregressive process. As with a standard Bayesian DLM, the autoregressive process on  $\theta_t$  allows for separate identification of the error variances  $\Omega$  and  $\Psi$ . Analogous to the case of a single category in the categorical LIV model, if the dynamic LIV is modeled using a process with a constant mean, the variances would not be separately identified.

To assess the performance of our dynamic LIV model and the categorical LIV in the presence of a dynamic endogenous regressor, we simulate three data sets according to the model described by Equation (8) with a single endogenous regressor for ease of exposition. The data sets differ in the data-generating values of the parameter  $\Upsilon$ , which governs the carryover dynamics in the latent instrument. We investigate values of  $\Upsilon = 0.5, 0.7$ , and  $0.9$ . The diagonal elements in  $\Omega$ , which represent the variance of the information stock ( $S_t$ ) and the variance of the endogenous regressor ( $c_t$ ), are set to 0.5 and 1.0, respectively. The off-diagonal element of  $\Omega$ , which represents the covariance between the stock residual ( $\omega_i^s$ ) and the residual of the instrument equations ( $\omega_i^c$ ), is set to  $-0.7$ .<sup>7</sup> Finally, we set the structural parameter  $\gamma$  to 1.

<sup>7</sup> The direction of the bias is determined by the sign of the correlation. A negative correlation will induce a downward bias in the structural parameters, whereas a positive correlation will induce an upward bias.

**Table 1** Coefficient Estimates for a Dynamic Endogenous Regressor Under Different Carryover Parameters for the Dynamic Latent Instrument

Variable	Specification of endogenous regressor		
	Exogenous	Endogenous	
		<i>Categorical LIV</i>	<i>Dynamic LIV</i>
Estimates for $\gamma^b$			
Data-generating value, $\Upsilon^a$		$\Upsilon = 0.5$	
Posterior mean	0.76	0.69	<b>0.97<sup>c</sup></b>
Posterior standard error	(0.03)	(0.18)	(0.03)
95% coverage interval	(0.71, 0.82)	(0.28, 0.95)	(0.90, 1.04)
Data-generating value, $\Upsilon$		$\Upsilon = 0.7$	
Posterior mean	0.70	0.74	<b>0.95</b>
Posterior standard error	(0.04)	(0.11)	(0.05)
95% coverage interval	(0.63, 0.76)	(0.53, 0.94)	(0.85, 1.06)
Data-generating value, $\Upsilon$		$\Upsilon = 0.9$	
Posterior mean	0.72	0.72	<b>0.98</b>
Posterior standard error	(0.03)	(0.07)	(0.04)
95% coverage interval	(0.66, 0.78)	(0.58, 0.86)	(0.91, 1.05)

<sup>a</sup>The parameter  $\Upsilon$  governs the carryover dynamics in the dynamic latent instrument.

<sup>b</sup>The parameter  $\gamma$  is the effect of the endogenous regressor on the stock variable. In all three cases, we generate data with  $\gamma = 1$ .

<sup>c</sup>Bold indicates the 95% coverage interval spans the data-generating value of  $\gamma = 1$ .

The remaining details of the simulation are omitted for brevity but are available from the authors upon request. For each of the three simulated data sets, we estimate the exogenous stock model described by Equation (2), the categorical LIV model, and the dynamic LIV model.

In Table 1, we report the posterior mean and standard error of the  $\gamma$  estimates from the three models, as well as the 95% coverage intervals. Given the data-generating parameters, we expect to see a downward bias in the estimates of the structural parameter that ignores endogeneity. For all three simulated data sets, treating the endogenous regressor as exogenous results in biased estimates of the structural parameter, with posterior mean estimates of  $\gamma$  ranging from 0.70 to 0.76 across the data sets. None of the coverage intervals spans the data-generating value of  $\gamma = 1$ . The categorical LIV estimator does not seem to be able to capture sufficient exogenous variation in the dynamic endogenous regressor to remedy the endogeneity bias, yielding posterior mean estimates of  $\gamma$  ranging from 0.69 to 0.74 across the data sets. As with the exogenous model, none of the coverage intervals spans the data-generating value of  $\gamma = 1$ .<sup>8</sup> For the

<sup>8</sup> For each scenario ( $\Upsilon = 0.5, 0.7, 0.9$ ), we generated multiple data sets and estimated the exogenous model, categorical LIV model, and dynamic LIV model. In the interest of brevity, we do not present a full simulation study. For the exogenous and dynamic models, the results are consistent with the estimates reported in

dynamic LIV model, the posterior mean estimates of  $\gamma$  range from 0.95 to 0.98 across the three data sets. In all cases, the 95% coverage interval spans the data-generating value of  $\gamma = 1$ . We now turn our attention to the performance of the models in our empirical application.

### 3. Data and Estimation Results

Online communication data for our firm were collected by a leading provider of social media-monitoring tools via their proprietary Web crawler technology. The data consist of daily counts of online comments for the firm and its products and services for the period April 1, 2007 to December 7, 2007. The valence of each mention is classified by a proprietary algorithm as positive, negative, or neutral.<sup>9</sup> Table 2 presents some descriptive statistics for the daily counts. The neutral category accounts for the majority of observations.

In addition to the online communication data, our cooperating firm provided us with daily sales revenue data and data on the date of new product launch announcements (covering the same period as the online communications data). To specify our model of sales performance, we include in the vector  $x_t$  a dummy variable for weekends and a dummy variable for new product launch announcements. We estimate four models. Model M1 is described by Equation (2) and treats the online communications as exogenous. Models M2–M4 treat online communications as endogenous. Model M2 uses lagged communications as an instrument, model M3 uses the categorical LIV approach, and model M4 uses our dynamic LIV approach. Table 3 reports the posterior mean estimates of the structural parameters and the parameters on the weekend and new product dummies, along with the standard error of the estimates (in parentheses) across the four models. The asterisks in the tables denote parameter estimates for which the coverage intervals of the estimates do not span zero. We use the coverage intervals to assess the statistical significance of the parameter estimates.

Table 1. However, on occasion, the coverage interval for the categorical LIV model contains the data-generating value for the structural parameter. In these cases, the posterior mean of the standard error is relatively high and the coverage intervals relatively wide. This is similar to the case of estimation with weak instruments, which sometimes alleviate parameter bias but at the expense of higher standard errors.

<sup>9</sup> The provider uses supervised machine learning to perform the automated sentiment analysis. Two broad categories of data features are used: those derived from natural language processing and those derived from social context. The features are derived from titles, bodies, threads, sites, permalinks, authors, publication dates, queries, customers, etc. For example, a simple feature might be the existence of a word such as “love” in a post, which would be statistical evidence in favor of a positive classification.



**Table 2** Descriptive Statistics: Positive, Negative, and Neutral Comments

Valence of comment	Average per day	Median	Standard deviation
Positive	255	199	249
Negative	370	301	333
Neutral	2,287	1,975	1,436

For model M1, which treats online communications as exogenous, the 95% coverage intervals on the coefficients for positive, negative, and neutral online communications all span zero. We find a similar result for model M2, which uses lagged online communications as an instrument. For model M3, the categorical LIV model, the 95% coverage interval on positive online communications does not span zero; the 95% coverage intervals on the negative and neutral coefficients do. However, the posterior mean estimates of the probabilities for the categorical latent instruments are 0.98, 0.02, and 0.04 for positive, negative, and neutral online communications, respectively. This is an important point because these estimates indicate that there is little variation with which to identify the structural parameters (Ebbes et al. 2005). The results for models M2 and M3 suggest that lagged values as instruments, as well as the categorical LIV approach, do little to alleviate endogeneity concerns.

For model M4, the dynamic LIV model, we find that positive and neutral comments have a positive effect on the stock variable, whereas negative comments have a negative effect. In all cases, the 95% coverage intervals of the estimates do not span zero. The posterior mean estimate of the coefficient on positive

comments is larger in magnitude than that of negative comments. This result is consistent with behavioral research that finds positive information is more persuasive in the case of a familiar brand (Ahluwalia 2002). The coefficient for neutral comments is positive, but it is smaller in magnitude than those for positive and negative comments. This is consistent with the idea that neutral comments spread awareness and/or inform consumers about the installed base of the firm's products, the latter of which can be important if adoption is influenced by imitators (Bass 1969).<sup>10</sup>

Under the dynamic structure of the model, the contemporaneous effect of an increase in the  $l$ th category of online comments at time  $t$  on sales at time  $t$  is measured by  $\gamma^l$ .<sup>11</sup> The effect on future sales in period  $k > t$  is given by  $\gamma^l \delta^{k-t}$ . It follows that the cumulative effect up to time  $k$  on sales of an increase in the  $l$ th type of online comments at time  $t$  is  $\sum_{i=1}^k \gamma^l \delta^{i-1}$ , a geometric series. Assuming an infinite horizon, we measure the long-run effect of an increase in the  $l$ th category of online comments at time  $t$  as  $\gamma^l / (1 - \delta)$ .

Table 4 reports the contemporaneous and long-run elasticities of sales with respect to each category of online communications. Positive and negative comments have elasticities that are larger in magnitude compared with neutral comments. Among positive and negative comments, positive comments have a slightly larger elasticity. These results demonstrate the importance of valence in measuring the effect of online communications. Figure 1 presents a plot of the duration of the elasticities with respect to a change in online comments of each type at time  $t$ . The majority of the effects have dissipated within a week. The magnitude of the long-run elasticities and the duration of the elasticities indicate that ignoring the dynamic effects would severely understate the effect of online communications on sales.

For all four models, the coefficient on the weekend dummy is negative, indicating a sales drop on the weekends. The 95% coverage intervals on the estimates do not span zero for any of the four models. In contrast, the new product launch announcements do not appear to have a significant effect on sales.

**Table 3** Parameter Estimates for Positive, Negative, and Neutral Comments and Exogenous Covariates Across Alternative Model Specifications

Variable	Specification of online communications			
	Exogenous		Endogenous	
	M1: Exogenous	M2: Lagged IV	M3: Categorical LIV	M4: Dynamic LIV
Comment valence				
Positive	0.010 (0.005)	0.069 (0.044)	0.022* (0.009)	0.424* (0.117)
Negative	−0.002 (0.004)	−0.022 (0.027)	0.012 (0.007)	−0.250* (0.076)
Neutral	0.001 (0.001)	0.006 (0.006)	0.003 (0.002)	0.024* (0.006)
Exogenous covariates				
Weekend	−14.110* (0.320)	−14.080* (0.320)	−14.190* (0.310)	−13.940* (0.350)
New product launch announcement	−0.680 (0.480)	−0.670 (0.470)	−0.590 (0.480)	−0.690 (0.480)

Note. Cell entries are posterior mean and posterior standard error (in parentheses).

\*95% coverage interval does not span zero.

<sup>10</sup> As discussed in West and Harrison (1997) and Van Heerde et al. (2004), an advantage of the DLM framework is that the model may be specified directly in levels rather than changes, even in the presence of a random walk. However, concerns may linger over the issue of serial correlation in the observation equation errors. Using the Durbin–Watson and Bruesch–Godfrey tests (Greene 2000), we find no evidence of serial correlation in the errors. If serial correlation were present, it could easily be accommodated in the DLM by augmenting the state space to allow for autoregressive error terms. See Rutz and Bucklin (2011) for details.

<sup>11</sup> Recall that  $l \in \{p, n, o\}$  stands for the three categories of online communications: positive, negative, and neutral.

**Table 4** Contemporaneous and Long-Run Elasticities of Sales with Respect to Online Communications

Valence of comment	Contemporaneous	Long run
Positive	0.64* (0.18)	1.54* (0.34)
Negative	−0.56* (0.17)	−1.34* (0.38)
Neutral	0.34* (0.09)	0.80* (0.11)

*Note.* Cell entries are posterior mean and posterior standard error (in parentheses).

\*95% coverage interval does not span zero.

To compare the models estimated thus far, we compute log Bayes factors using the step-ahead predictive densities (West and Harrison 1997, Van Heerde et al. 2004, Bass et al. 2007). We also compute the mean absolute error (MAE) and mean squared error (MSE) of the step-ahead forecasts (West and Harrison 1997). We treat model M4, the dynamic LIV model, as the null model. A log Bayes factor between 1 and 2 indicates evidence in favor of the null model, whereas a log Bayes factor greater than 2 indicates strong evidence in favor of the null model. Table 5 reports the results. We find strong evidence for the dynamic LIV model over the alternative models. Figure 2 presents a plot of weekday sales along with the step-ahead forecasts and coverage intervals. Figure 3 presents the evolution of the stock variable  $S_t$  for our dynamic LIV model. The posterior mean estimate of the carryover parameter  $\delta$  is 0.58 (with a standard error of 0.10). The 95% coverage interval does not span zero.

In Table 6, we report the posterior mean estimates of the correlation between the stock equation error

**Table 5** Model Comparison

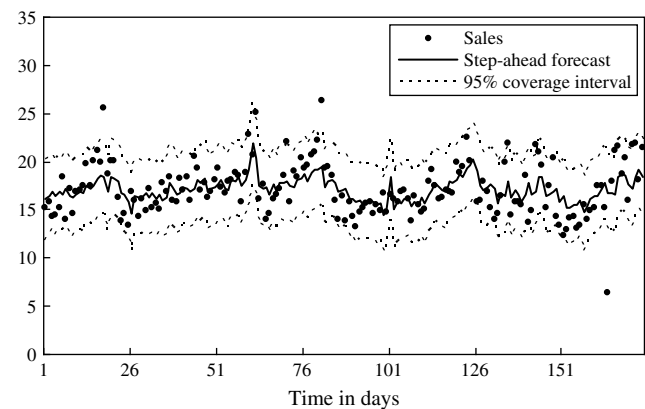
Model	Specification of online communications	Instruments	Log Bayes factor <sup>a</sup>	MAE <sup>b</sup>	MSE <sup>c</sup>
M1	Exogenous	—	17.60	1.80	6.52
M2	Endogenous	Lagged communications	69.42	1.83	7.08
M3	Endogenous	Categorical LIV	11.70	1.76	6.16
M4	Endogenous	Dynamic LIV	—	1.60	4.89

<sup>a</sup>Model M4 is the null model.

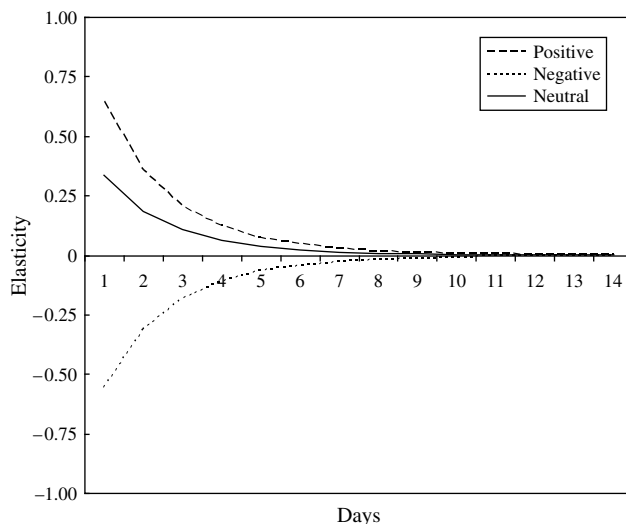
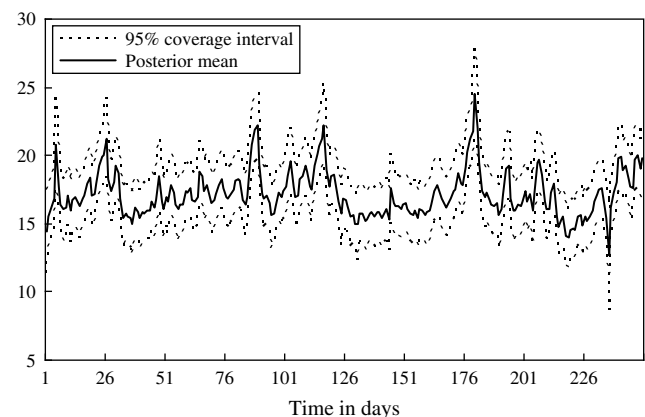
<sup>b</sup>Mean absolute error of step-ahead forecast.

<sup>c</sup>Mean squared error of step-ahead forecast.

terms and the errors from the dynamic LIV equations. We find negative correlations between the stock error and the LIV error for positive and neutral comments (posterior mean estimates of −0.80 and −0.46, respectively). We find positive correlation (posterior mean estimate of 0.32) between the stock error and

**Figure 2** Plot of Weekday Sales Revenues and Step-Ahead Forecast

*Notes.* To aid the graphical presentation of the data, the plot excludes weekends because weekend sales are orders of magnitude lower than weekday sales. In the model, the drop in sales on weekends is ably captured by the weekend dummy.

**Figure 1** Elasticity Wearout for Online Communications**Figure 3** Plot of Posterior Mean of a Stock Variable with a 95% Coverage Interval

**Table 6** Error Correlations for the Stock and Instrument Equations of the Dynamic LIV Model

Correlation estimates	
Variable	$\omega_t^S$
$\omega_t^{\text{pos}}$	−0.80* (0.04)
$\omega_t^{\text{neg}}$	0.32* (0.10)
$\omega_t^{\text{neu}}$	−0.46* (0.08)

*Note.* Cell entries are posterior mean and posterior standard error (in parentheses).

\*95% coverage interval does not span zero.

the LIV error for negative comments. For all three correlations, the 95% coverage interval does not span zero. These results strongly indicate the presence of regressor-error dependencies in the stock equation. Furthermore, given the estimated signs of the coefficients (positive for positive and neutral comments and negative for negative comments) and the correlations (negative for positive and neutral comments and positive for negative comments), the expectation is that ignoring the dependencies in estimation will bias the coefficient estimates toward zero. This is consistent with the estimation results for model M1, which treats online communications as exogenous, giving face validity to our results.

## 4. Managerial Implications

Our results thus far demonstrate that online communications beyond those captured by product ratings and reviews have an effect on firm sales. We now turn our attention to two specific managerial implications of our results. First, we examine the importance of accounting for the sentiment of online communications. Second, we examine the revenue impact of shocks to positive, negative, and neutral online communications.

### 4.1. The Importance of Accounting for Sentiment

Several new sentiment analysis companies are attempting to interest client firms in their ability to help understand the information being exchanged in the online space (Wright 2009). However, automated sentiment analysis data are ostensibly more costly. Indeed, firms can obtain volume data for free from sites such as Technorati. Furthermore, previous research has found an effect for the simple overall volume of communications. We investigate the importance of accounting for sentiment by estimating a restricted model, M5, which sets  $\gamma^p = \gamma^n = \gamma^o = \gamma$ . This tests the null model, M4, against a restricted model where online communications have the same effect across positive, negative, and neutral comments (i.e., a single coefficient on the total volume of comments aggregated over positive, negative, and neutral). For the sake of completeness, we also test a second restricted version of the dynamic LIV model, M6, which sets  $\gamma^p = \gamma^n = \gamma^o = 0$ . This tests the null model against the restricted model, where online communications have no effect.

The log Bayes factor as well as the MAE and MSE of the step-ahead forecasts, reported in Table 7, indicate strong support for model M4, which accounts for sentiment. Furthermore, the 95% coverage interval for  $\gamma$  in model M5, which aggregates over valence, spans zero. Thus, the effect of online communications is masked by ignoring comment valence captured by automated sentiment analysis. This result demonstrates the value of automated sentiment data. Firms that currently rely on free or less costly total volume data may be underestimating the value of online communications.

### 4.2. Revenue Effects of Shocks to Positive, Negative, and Neutral Communications

Firms are increasingly attempting to initiate word-of-mouth communications among existing and potential customers with the goal of generating awareness and spreading positive information (Godes and Mayzlin 2009). Although it seems extremely unlikely that a firm would attempt to trigger negative comments, it is the case that small mistakes by firms can also

**Table 7** The Importance of Accounting for Sentiment

Model	Specification of online communications	Instruments	Restrictions on structural parameters	Log Bayes factor <sup>a</sup>	MAE <sup>b</sup>	MSE <sup>c</sup>
M4	Endogenous	Dynamic LIV	None	—	1.60	4.89
M5 <sup>d</sup>	Endogenous	Dynamic LIV	$\gamma^p = \gamma^n = \gamma^o = \gamma$	14.71	1.78	6.56
M6 <sup>e</sup>	Endogenous	Dynamic LIV	$\gamma^p = \gamma^n = \gamma^o = 0$	14.89	1.83	7.14

<sup>a</sup>Model M4 is the null model.

<sup>b</sup>Mean absolute error of step-ahead forecast.

<sup>c</sup>Mean squared error of step-ahead forecast.

<sup>d</sup>Model M5 aggregates positive, negative, and neutral communications into a total count.

<sup>e</sup>Model M6 restricts the effect of online communications to zero.

quickly become large problems. For example, in 2006, a Comcast repair technician fell asleep during a service call to a customer's home. The customer shot a video of the sleeping technician and posted it to a blog. The video generated more than 200,000 views and was discussed by the technology blog Gizmodo and the mainstream news, including MSNBC and the *New York Times* (Belson 2006). Using our model, we investigate the effect of a shock to positive, negative, and neutral online communications.

Under our dynamic LIV model, a shock to the  $l$ th category of online comments at time  $t$  will carry over to future values. To capture this dynamic, we use the forecast distributions of the observation and state variables in our DLM (West and Harrison 1997) to generate 14-day forecasts of online comments and sales. Letting  $T$  represent the terminal period of our data, the posterior distributions of the dynamic state variables in the model become the prior distribution for the  $T + 1$ st period. We first compute a baseline forecast and subsequently shock the  $T + 1$ st prior, in turn, for each of the  $l \in \{p, n, o\}$  categories and recompute the forecasts. Figure 4 presents the cumulative revenues over the forecast period for the baseline and the three scenarios where positive, negative, and neutral comments receive a 10% shock in the initial forecast period. For positive comments, the shock results in an 18% increase in cumulative revenues over the forecast period. For negative comments, the shock results in an 11% decrease in cumulative revenues. For neutral comments, the shock results in a 7% increase in cumulative revenues.

The results on positive and neutral communications complement the existing literature that shows that word-of-mouth campaigns can be an effective tool for the firm (Godes and Mayzlin 2009) while also underscoring the importance of managing situations that may result in negative communications. For example, given the expected revenue decline following a negative shock to online communications,

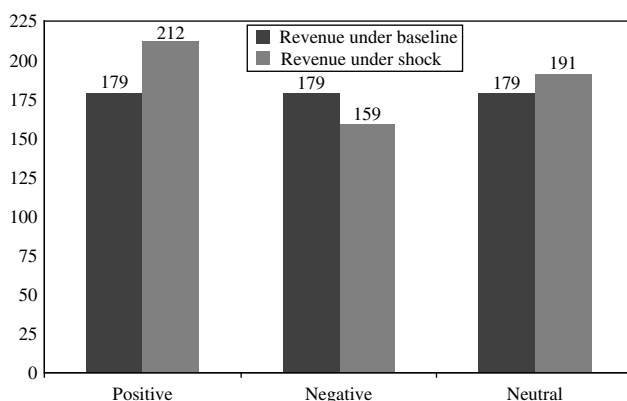
the firm may want to increase advertising and promotional spending to offset the decline. On the other hand, following a positive shock to online communications, the firm may want to decrease spending to boost profitability. A full treatment of this issue would require spend data, which we unfortunately do not have.

## 5. Summary and Conclusions

The rise of Internet-based social media has facilitated the spread of online communications as well as the observability and measurement of the same. In particular, ratings and reviews permeate the online space for a wide range of categories. The academic marketing literature has begun to explore the role of online ratings in explaining firm outcomes. We argue that although interesting and important, ratings data are only a narrow slice of online communications about a firm and its products. Web crawler technology has enabled firms to measure the number of mentions regarding a firm and its products or services across a much wider swath of the online space. Such data are currently available to firms via public and private sources. Furthermore, private firms are investing in automated classification of online comments as positive, negative, or neutral (Wright 2009). In this paper, we investigate whether and how such information affects sales and develop an estimation methodology for uncovering such an effect.

Our measures of online communications consist of daily counts of comments for a durable goods firm selling via an online channel. The data are captured by Web crawler technology. Rather than ratings data from a single website such as Yahoo! Movies (Liu 2006, Duan et al. 2008) or postings from a single forum such as Usenet (Godes and Mayzlin 2004), our data capture a broader view of the online communications about the firm and its products across the Internet. Via a proprietary technique for conducting automated sentiment analysis, the comments are classified as positive, negative, and neutral. With these data, we model the effect of online communications on sales. Our model of the effect of online communications contains two important features. First, we account for dynamic effects of online communications by modeling sales as arising from a demand-generating dynamic stock variable. Positive, negative, and neutral online communications contribute to the evolution of the dynamic stock. The model is inspired by goodwill models of advertising. However, such models typically assume exogeneity of the advertising variables in the stock equation, prompting the second important feature of our model. We account for potential endogeneity in the estimation of the effect of online communications on the stock variable by

**Figure 4** Fourteen-Day Cumulative Revenue Forecasts Under Baseline and 10% Shocks



extending the latent instrumental variable estimator to account for dynamic endogenous regressors.

We argue that endogeneity of online communications is a concern for at least three reasons. First, Web crawler technology operates by searching for keywords. For some firms and products, keywords may not be unique. Although searches can be narrowed and refined by exclusion terms, the possibility of measurement error looms. Furthermore, the algorithms that classify mentions as positive, negative, and neutral are unlikely to be error-free, which also raises the potential for measurement error. Second, we do not have data on daily advertising or promotional activities of the firm. In addition, firms may engage in efforts to mediate, moderate, or seed online communications (Godes et al. 2005). These unobserved activities are likely to be correlated with observed measures of online communications, causing an omitted variables problem. Finally, research on the effect of ratings and reviews on sales has demonstrated feedback effects in sales and ratings data. Similarly, the stock variable and our measures of online communication may coevolve.

The essential problem with an endogenous regressor is correlation between the regressor and the error term. With a valid observed instrumental variable in hand, correcting for endogeneity bias is straightforward. However, the search for good instruments is a classic problem in applied econometrics. With weak or invalid instruments, the solution may aggravate rather than ameliorate the problem. Recent work in marketing has considered a new estimator to correct for endogeneity bias based on latent instruments. The LIV estimator (Ebbes et al. 2005) divides the endogenous regressor into systematic and random components. The systematic component is then modeled as arising from a set of latent categories with category-specific means. Most, if not all, of the existing applications of the LIV method make use of data with cross-sectional variation. The properties of this estimator are not well explored in the context of time-series data. In simulated time-series data as well as our actual data, we find the categorical LIV method is unable to correct the endogeneity bias. To overcome the problem, we extend the LIV method by modeling the systematic component of the endogenous regressor as a latent dynamic state matched to the observed endogenous online communications variable and, by construction, orthogonal to the stock equation error term.

Our empirical results show that endogeneity bias hampers estimates of the effect of online communications on sales. Models that treat online communications as exogenous or use lagged communications as an instrument suggest no effect. In

contrast, our dynamic LIV model finds a significant impact of online communications on sales. Log Bayes factors and step-ahead forecast performance indicate strong support for our proposed dynamic LIV model. We are among the first to document the impact of online communications measured by Web crawler technology and scored with automated sentiment analysis. To the best of our knowledge, we are the first to document an effect for positive, negative, and neutral comments in this domain. Our findings underscore the importance of accounting for dynamics and endogeneity in the model. We find that the effect sizes for positive and negative comments are larger than that of neutral comments and that the effect size for positive comments is larger than that of negative. In terms of the duration of the effects, we find that the majority of the effects have dissipated after about a week.

Important for managers, we find that aggregating online comments over valence masks the effect on sales. Thus, although obtaining the sentiment analysis data is costly relative to publicly available data on the total volume of comments, the sentiment analysis data improve model fit and lead to a fundamentally different conclusion regarding the effect of online communications. Previous research has noted the increasing propensity of firms to attempt to initiate word-of-mouth communications among existing and potential customers (Godes and Mayzlin 2009). Our forecasting exercise demonstrates that firms have good reason to focus on increasing positive and neutral comments, but it also demonstrates that events precipitating negative comments can adversely affect revenues.

A limitation of our study is that our data are for a single firm. However, our goal here is to provide an initial step toward understanding the effect of online communications on sales performance. To this end, we develop a method for coping with the challenges of dynamics and endogeneity inherent in this new data. Future research should strive to replicate this result across other firms and industries. Additional future research topics in this area are plentiful. As noted in §1, firms are increasing spending on word-of-mouth marketing. We do not have data to address issues such as the relationship between advertising and promotional spending and the generation of online communications. Our data were collected by our cooperating firm with an intentionally broad focus. Future research may investigate a more refined classification of online communications, such as brand comments, product comments, or service comments. Such refinements could lead to important strategic findings in terms of differential effects across types of online communication. Finally, an interesting issue to consider is the strategic role of

the firm. Beyond listening to and measuring online communications, firms may take a more active role in shaping or seeding online conversation. Indeed, our results suggest that firms have an incentive to involve themselves in the process. However, as firms become more engaged in the process, consumers' perceptions regarding the authenticity of online communications may be harmed, eventually dampening or eliminating the effect. Future research should seek to understand effective strategies for firm involvement that preserve the authenticity of the information and its ultimate effect.

### Acknowledgment

O. J. Rutz's current affiliation is the Foster School of Business, University of Washington, Seattle, WA 98195.

### Appendix

We detail the sampler used to estimate the dynamic LIV model described in Equation (8):

$$\begin{aligned} & y_t | x_t, \beta, S_t, \sigma_\varepsilon^2, \\ & \beta | y_t, x_t, S_t, \sigma_\varepsilon^2, \\ & \sigma_\varepsilon^2 | y_t, x_t, S_t, \beta, \\ & S_t | y_t, x_t, c_t, \beta, S_{t-1}, \theta_t, \gamma, \delta, \Omega, \\ & c_t | \theta_t, \Omega, \\ & \gamma | c_t, S_t, S_{t-1}, \delta, \Omega, \\ & \delta | c_t, S_t, S_{t-1}, \gamma, \Omega, \\ & \Omega | c_t, S_t, S_{t-1}, \gamma, \delta, \theta_t, \\ & \theta_t | \theta_{t-1}, \Upsilon, \Psi, \\ & \Upsilon | \theta_t, \theta_{t-1}, \Psi, \\ & \Psi | \theta_t, \theta_{t-1}, \Upsilon. \end{aligned}$$

Where convenient, we drop the subscript  $t$  for ease of exposition.

1. Generate  $\beta | y, X, S, \sigma_\varepsilon^2$ .

$$\begin{aligned} \beta | y, X, S, \sigma_\varepsilon^2 & \sim \text{MVN}\left(b, \left[\frac{1}{\sigma_\varepsilon^2} X'X + \Sigma_\beta^{-1}\right]^{-1}\right), \\ b & = \left[\frac{1}{\sigma_\varepsilon^2} X'X + \Sigma_\beta^{-1}\right]^{-1} \left[\frac{1}{\sigma_\varepsilon^2} X'\tilde{y} + \Sigma_\beta^{-1}\tilde{\beta}\right], \\ \tilde{y} & = y - S, \\ \Sigma_\beta & = 10^6 \times I_K, \\ \tilde{\beta} & = 0_K. \end{aligned}$$

2. Generate  $\sigma_\varepsilon^2 | y, X, S, \beta$ .

$$\begin{aligned} \sigma_\varepsilon^2 | y, X, S, \beta & \sim \text{IG}\left(a_\sigma + \frac{1}{2}(\tilde{y} - X\beta)'(\tilde{y} - X\beta), b_\sigma + \frac{T}{2}\right), \\ a_\sigma & = 2 + K, \\ b_\sigma & = 1. \end{aligned}$$

3. Generate  $S_t | y_t, x_t, c_t, \beta, S_{t-1}, \theta_t, \gamma, \delta, \Omega$ .

We estimate the information stock via a dynamic linear model (West and Harrison 1997). The observation equation of our DLM matches observed sales with the latent dynamic variable of interest—in our case, the stock variable. The observation and state equations of the DLM are given by

$$\begin{aligned} y_t & = x_t'\beta + S_t + \varepsilon_t, \\ S_t & = \delta S_{t-1} + c_t'\gamma + \omega_t^S, \end{aligned} \quad (9)$$

where  $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$ . The latent instrument equation for the vector of online communications is  $c_t = \theta_t + \omega_t^c$  and  $\omega_t \sim \text{MVN}(0, \Omega)$ , where  $\Omega$  is a full  $4 \times 4$  matrix. We partition  $\Omega$  such that  $\Omega = \begin{bmatrix} \Omega_{ss} & \Omega_{sc} \\ \Omega_{cs} & \Omega_{cc} \end{bmatrix}$ , where  $\Omega_{ss}$  is the scalar variance of  $\omega_t^S$ ,  $\Omega_{cc}$  is the  $3 \times 3$  covariance matrix of  $\omega_t^c$ , and  $\Omega_{sc} = \Omega_{cs}'$  is a  $3 \times 1$  vector of covariance terms for  $\omega_t^S$  and  $\omega_t^c$ . We use the forward-filtering and backward-smoothing algorithm (West and Harrison 1997) to sample  $S_t$ . The posterior at  $t-1$  is

$$S_{t-1} | D_{t-1}^S \sim N(m_{t-1}^S, C_{t-1}^S). \quad (10)$$

The prior at time  $t$  is

$$S_t | D_{t-1}^S \sim N(a_t^S, R_t^S), \quad (11)$$

where  $a_t^S = \delta m_{t-1}^S + c_t'\gamma + \Omega_{sc}\Omega_{cc}^{-1}[c_t - \theta_t]'$  and  $R_t^S = \delta C_{t-1}^S \delta' + \Omega_{ss} - \Omega_{sc}\Omega_{cc}^{-1}\Omega_{cs}$ .

The step-ahead forecast distribution is

$$y_t | D_{t-1}^S \sim N(f_t^S, Q_t^S), \quad (12)$$

where  $f_t^S = a_t^S + x_t'\beta$  and  $Q_t^S = R_t^S + \sigma_\varepsilon^2$ .

Finally, the posterior at time  $t$  is

$$S_t | D_t^S \sim N(m_t^S, C_t^S), \quad (13)$$

where  $m_t^S = a_t^S + A_t^S(y_t - f_t^S)$ ,  $C_t^S = R_t^S - A_t^S Q_t^S A_t^S$ , and  $A_t^S = R_t^S(Q_t^S)^{-1}$ .

We use backward sampling to obtain draws of  $S_t$  for all  $t = 1, \dots, T$ . First, we sample  $S_T | D_T^S$  from its conditional distribution as described in (13). Second, for each  $t = T-1, T-2, \dots, 0$ , we sample from  $(S_t | S_{t+1}, D_t^S) \sim N(g_t^S, H_t^S)$ , where  $g_t^S = m_{t+1}^S + B_{t+1}^S(S_{t+1} - a_{t+1}^S)$ ,  $H_t^S = C_{t+1}^S - B_{t+1}^S R_{t+1}^S B_{t+1}^S$ , and  $B_{t+1}^S = C_{t+1}^S \delta (R_{t+1}^S)^{-1}$ .

4. Generate  $\delta, \gamma | c_t, S_t, S_{t-1}, \Omega$ .

Let  $\tilde{\beta} = [\delta' \gamma']'$ ,  $\tilde{y}_t = S_t - \Omega_{sc}\Omega_{cc}^{-1}[c_t - \theta_t]'$ ,  $\tilde{x}_t = [S_{t-1} \ c_t']$ , and  $\tilde{\Omega}_{ss} = \Omega_{ss} - \Omega_{sc}\Omega_{cc}^{-1}\Omega_{cs}$ . Then

$$\begin{aligned} \tilde{\beta} | \tilde{y}, \tilde{X}, \Omega & \sim \text{MVN}\left(\tilde{b}, \left[\frac{1}{\tilde{\Omega}_{ss}} \tilde{X}'\tilde{X} + \Sigma_{\tilde{\beta}}^{-1}\right]^{-1}\right), \\ \tilde{b} & = \left[\frac{1}{\tilde{\Omega}_{ss}} \tilde{X}'\tilde{X} + \Sigma_{\tilde{\beta}}^{-1}\right]^{-1} \left[\frac{1}{\tilde{\Omega}_{ss}} \tilde{X}'\tilde{y} + \Sigma_{\tilde{\beta}}^{-1}\tilde{\beta}\right], \\ \Sigma_{\tilde{\beta}} & = 10^6 \times I_4, \\ \tilde{\beta} & = 0_4. \end{aligned}$$

5. Generate  $\Omega | c_t, S_t, S_{t-1}, \gamma, \delta, \theta_t$ .

$$\begin{aligned} \Omega | c_t, S_t, S_{t-1}, \gamma, \delta, \theta_t \\ \sim \text{IW}\left(q^\Omega + T, \left(V^\Omega + \sum_{i=1}^T (\phi_i - \bar{\phi})(\phi_i - \bar{\phi})'\right)\right), \end{aligned}$$

$$\begin{aligned}\phi_t &= [S_t \ c_t']', \\ \bar{\phi}_t &= [\delta S_{t-1} + c_t' \gamma \ \theta_t']', \\ q^\Omega &= 6, \\ V^\Omega &= I_4.\end{aligned}$$

6. Generate  $\theta_t | \theta_{t-1}, \Upsilon, \Psi$ .

To generate  $\theta_t$ , we transform to the reduced form of the model. To accommodate the transformation, we define the following terms.

Let

$$\begin{aligned}\tilde{S}_t &= S_t - \delta S_{t-1}, \quad \tilde{y}_t = [\tilde{S}_t \ c_t']', \quad F_t = \begin{bmatrix} \gamma^p & \gamma^n & \gamma^o \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \\ \tilde{\omega}_t &= [\omega_t^S + \omega_t^{c'} \gamma \ \omega_t^{c'}]', \quad L = \begin{bmatrix} 1 & \gamma^p & \gamma^n & \gamma^o \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \\ \text{and } \tilde{\Omega} &= L\Omega L'.\end{aligned}$$

The observation equation of the reduced form matches the stock variable and online communications with the latent dynamic instrument. The observation and state equations of this DLM are given by

$$\begin{aligned}\tilde{y}_t &= F_t \theta_t + \tilde{\omega}_t, \\ \theta_t &= \Upsilon \theta_{t-1} + \xi_t,\end{aligned}\tag{14}$$

where  $\xi_t \sim \text{MVN}(0, \Psi)$ . We use the forward-filtering and backward-smoothing algorithm (West and Harrison 1997) to sample  $\theta_t$ . The posterior at  $t-1$  is

$$\theta_{t-1} | D_{t-1}^\theta \sim \text{N}(m_{t-1}^\theta, C_{t-1}^\theta).\tag{15}$$

The prior at time  $t$  is

$$\theta_t | D_{t-1}^\theta \sim \text{N}(a_t^\theta, R_t^\theta),\tag{16}$$

where  $a_t^\theta = \Upsilon m_{t-1}^\theta$  and  $R_t^\theta = \Upsilon C_{t-1}^\theta \Upsilon' + \Psi$ .

The step-ahead forecast distribution is

$$\tilde{y}_t | D_{t-1}^\theta \sim \text{N}(f_t^\theta, Q_t^\theta),\tag{17}$$

where  $f_t^\theta = F_t a_t^\theta$  and  $Q_t^\theta = F_t R_t^\theta F_t' + \tilde{\Omega}$ .

Finally, the posterior at time  $t$  is

$$\theta_t | D_t^\theta \sim \text{N}(m_t^\theta, C_t^\theta),\tag{18}$$

where  $m_t^\theta = a_t^\theta + A_t^\theta(\tilde{y}_t - f_t^\theta)$ ,  $C_t^\theta = R_t^\theta - A_t^\theta Q_t^\theta A_t^\theta$ , and  $A_t^\theta = R_t^\theta (Q_t^\theta)^{-1}$ .

We use backward sampling to obtain draws of  $\theta_t$  for all  $t = 1, \dots, T$ . First, we sample  $\theta_T | D_T^\theta$  from its conditional distribution as described in (18). Second, for each  $t = T-1, T-2, \dots, 0$ , we sample from  $(\theta_t | \theta_{t+1}, D_t^\theta) \sim \text{N}(g_t^\theta, H_t^\theta)$ , where  $g_t^\theta = m_t^\theta + B_t^\theta(\theta_{t+1} - a_{t+1}^\theta)$ ,  $H_t^\theta = C_t^\theta - B_t^\theta R_{t+1}^\theta B_t^\theta$ , and  $B_t^\theta = C_t^\theta \Upsilon' (R_{t+1}^\theta)^{-1}$ .

7. Generate  $\Upsilon | \theta_t, \theta_{t-1}, \Psi$ .

Let

$$\tilde{\beta} = \text{diag}[\Upsilon], \quad \tilde{y}_t = \theta_t', \quad \tilde{X}_t = \begin{bmatrix} \theta_{t-1}^p & 0 & 0 \\ 0 & \theta_{t-1}^n & 0 \\ 0 & 0 & \theta_{t-1}^o \end{bmatrix}, \quad \text{and}$$

$$\tilde{\Psi} = I_T \otimes \Psi,$$

$$\tilde{\beta} | \tilde{y}, \tilde{X}, \tilde{\Psi} \sim \text{MVN}(\tilde{b}, [\tilde{X}' \tilde{\Psi}^{-1} \tilde{X} + \Sigma_{\tilde{\beta}^{-1}}]^{-1}),$$

$$\tilde{b} = [\tilde{X}' \tilde{\Psi}^{-1} \tilde{X} + \Sigma_{\tilde{\beta}^{-1}}]^{-1} [\tilde{X}' \tilde{\Psi}^{-1} \tilde{y} + \Sigma_{\tilde{\beta}^{-1}}^{-1} \tilde{\beta}],$$

$$\Sigma_{\tilde{\beta}} = 10^6 \times I_3,$$

$$\tilde{\beta} = 0_3.$$

8. Generate  $\Psi | \theta_t, \theta_{t-1}, \Upsilon$ .

$$\Psi | \theta_t, \theta_{t-1}, \Upsilon$$

$$\sim \text{IW}\left(q^\Upsilon + T, \left(V^\Upsilon + \sum_{t=1}^T (\theta_t - \Upsilon \theta_{t-1})(\theta_t - \Upsilon \theta_{t-1})'\right)\right),$$

$$q^\Upsilon = 5,$$

$$V^\Upsilon = I_3.$$

## References

- Ahluwalia, R. 2002. How prevalent is the negativity effect in consumer environments? *J. Consumer Res.* **29**(2) 270–279.
- Angrist, J. D., A. B. Krueger. 2001. Instrumental variables and the search for identification: From supply and demand to natural experiments. *J. Econom. Perspect.* **15**(4) 69–85.
- Bass, F. M. 1969. A new product growth for model consumer durables. *Management Sci.* **15**(5) 215–227.
- Bass, F. M., N. Bruce, S. Majumdar, B. P. S. Murthi. 2007. Wearout effects of different advertising themes: A dynamic Bayesian model of the advertising-sales relationship. *Marketing Sci.* **26**(2) 179–195.
- Belson, K. 2006. Your call is important to us. Please stay awake. *New York Times* (June 26), <http://www.nytimes.com/2006/06/26/technology/26comcast.html>.
- Chevalier, J. A., D. Mayzlin. 2006. The effect of word of mouth on sales: Online product reviews. *J. Marketing Res.* **43**(3) 345–354.
- Chintagunta, P. K., S. Gopinath, S. Venkataraman. 2010. The effects of online user reviews on movie box office performance: Accounting for sequential rollout and aggregation across local markets. *Marketing Sci.* **29**(5) 944–957.
- Dellarocas, C., M. Zhang, N. F. Awad. 2007. Exploring the value of online product reviews in forecasting sales: The case of motion pictures. *J. Interactive Marketing* **21**(4) 23–45.
- Dhar, V., E. A. Chang. 2009. Does chatter matter? The impact of user-generated content on music sales. *J. Interactive Marketing* **23**(4) 300–307.
- Duan, W., B. Gu, A. B. Whinston. 2008. The dynamics of online word-of-mouth and product sales—An empirical investigation of the movie industry. *J. Retailing* **84**(2) 233–242.
- Ebbes, P., M. Wedel, U. Böckenholt, T. Steerneman. 2005. Solving and testing for regressor-error (in)dependence when no instrumental variables are available: With new evidence for the effect of education on income. *Quant. Marketing Econom.* **3**(4) 365–392.
- Giese, J. L., E. R. Spangenberg, A. E. Crowley. 1996. The effect of product-specific word-of-mouth communication on product category involvement. *Marketing Lett.* **7**(2) 87–99.

- Godes, D., D. Mayzlin. 2004. Using online conversations to study word-of-mouth communication. *Marketing Sci.* **23**(4) 545–560.
- Godes, D., D. Mayzlin. 2009. Firm-created word-of-mouth communication: Evidence from a field test. *Marketing Sci.* **28**(4) 721–739.
- Godes, D., D. Mayzlin, Y. Chen, S. Das, C. Dellarocas, B. Pfeiffer, B. Libai, S. Sen, M. Shi, P. Verlegh. 2005. The firm's management of social interactions. *Marketing Lett.* **16**(3/4) 415–428.
- Greene, W. H. 2000. *Econometric Analysis*, 4th ed. Prentice-Hall, Upper Saddle River, NJ.
- Herr, P. M., F. R. Kardes, J. Kim. 1991. Effects of word-of-mouth and product-attribute information on persuasion: An accessibility-diagnostics perspective. *J. Consumer Res.* **17**(4) 454–462.
- Kuksov, D., J. M. Villas-Boas. 2008. Endogeneity and individual consumer choice. *J. Marketing Res.* **45**(6) 702–714.
- Liu, Y. 2006. Word of mouth for movies: Its dynamics and impact on box office revenue. *J. Marketing* **70**(3) 74–89.
- Luo, X. 2007. Consumer negative voice and firm-idiosyncratic stock returns. *J. Marketing* **71**(3) 75–88.
- Luo, X. 2009. Quantifying the long-term impact of negative word of mouth on cash flows and stock prices. *Marketing Sci.* **28**(1) 148–165.
- Moe, W. W., M. Trusov. 2011. The value of social dynamics in online product ratings forums. *J. Marketing Res.* **48**(3) 444–456.
- Musalem, A., E. T. Bradlow, J. S. Raju. 2008. Who's got the coupon: Estimating consumer preferences and coupon usage from aggregate information. *J. Marketing Res.* **45**(6) 715–730.
- Naik, P. A., K. Raman, S. Srinivasan. 2010. Managing corporate and product brands. Working paper, University of California, Davis, Davis.
- Nelson, P. 1970. Information and consumer behavior. *J. Political Econom.* **78**(2) 311–329.
- Nerlove, M., K. J. Arrow. 1962. Optimal advertising policy under dynamic conditions. *Economica* **29**(114) 129–142.
- Onishi, H., P. Manchanda. 2010. Marketing activity, blogging and sales. Working paper, University of Michigan, Ann Arbor.
- PQ Media. 2009. Despite worst recession in decades, brands increased spending on word-of-mouth marketing 14.2% to \$1.54 billion in 2008. (July 29), <http://www.pqmedia.com/about-press-20090729-wommf.html>.
- Rist, P. 2005. BIGresearch releases SIMM VII; word of mouth most influential, other media vary by demos and product categories. BIGresearch (December 20), <http://www.bigresearch.com/news/big122005.htm>.
- Rossi, P., G. Allenby, R. McCulloch. 2006. *Bayesian Statistics and Marketing*. John Wiley & Sons, West Sussex, UK.
- Rutz, O. J., R. E. Bucklin. 2011. From generic to branded: A model of spillover in paid search advertising. *J. Marketing Res.* **48**(1) 87–102.
- Shin, H. S., D. M. Hanssens, B. Gajula. 2010. The impact of positive vs. negative online buzz on retail prices. Working paper, Long Island University, Brookville, NY.
- Stock, J., J. H. Wright, M. Yogo. 2002. A survey of weak instruments and weak identification in generalized method of moments. *J. Bus. Econom. Statist.* **20**(4) 518–529.
- Tirunillai, S., G. Tellis. 2010. Does chatter matter? The impact of online consumer generated content on a firm's financial performance. Working paper, University of Southern California, Los Angeles.
- Van Heerde, H. J., C. F. Mela, P. Manchanda. 2004. The dynamic effect of innovation on market structure. *J. Marketing Res.* **41**(2) 166–183.
- West, M., J. Harrison. 1997. *Bayesian Forecasting and Dynamic Models*. Springer-Verlag, New York.
- Wright, A. 2009. Mining the Web for feelings, not facts. *New York Times* (August 23), <http://www.nytimes.com/2009/08/24/technology/internet/24emotion.html>.
- Zhang, J., M. Wedel, R. Pieters. 2009. Sales effects of visual attention to feature ads: A Bayesian mediation analysis. *J. Marketing Res.* **46**(5) 669–681.