View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by New York University Faculty Digital Archive

NET Institute*

www.NETinst.org

Working Paper #07-32

September 2007

Negative Blogs, Positive Outcomes:
When Should Firms Permit Employees to Blog Honestly

Rohit Aggarwal, Ram Gopal, and Ramesh Sankaranarayanan,
University of Connecticut

* The Networks, Electronic Commerce, and Telecommunications ("NET") Institute, http://www.NETinst.org, is a non-profit institution devoted to research on network industries, electronic commerce, telecommunications, the Internet, "virtual networks" comprised of computers that share the same technical standard or operating system, and on network issues in general.

Negative Blogs, Positive Outcomes: When should Firms Permit Employees to Blog Honestly?

Rohit Aggarwal
Operations and Information Management, University of Connecticut, Storrs, Connecticut 06269,
Ph. No.:- 860-486-6485
Fax No.:- 860-486-4839
rohit.aggarwal@business.uconn.edu

Ram Gopal
Operations and Information Management, University of Connecticut, Storrs, Connecticut 06269,
Ph. No.:- 860-486-2408
Fax No.:- 860-486-4839
ram.gopal@business.uconn.edu

Ramesh Sankaranarayanan
Operations and Information Management, University of Connecticut, Storrs, Connecticut 06269,
Ph. No.:- 860-486-5217
Fax No.:- 860-486-4839
rsankaran@business.uconn.edu

ACKNOWLEDGEMENTS

We thank Nick Economides and the NET Institute (www.NETInst.org) for generous financial assistance. We thank Xinxin Li and participants of *The Sixteenth Annual Workshop on Information Technologies and Systems* for valuable feedback.

ABSTRACT

Weblogs or blogs have recently received a lot of attention, especially in the business community, with a number of firms encouraging their employees to publish blogs to reach out and connect to a wider audience. It is beginning to be recognized that employee blogs can cast a firm in either a positive or a negative light, thereby enhancing or harming the firm's reputation. Paradoxically, under certain conditions negative postings by employees can actually help the overall reputation of the firm. The rationale for this is that negative posts raise the credibility of an employee blog and attract more readers, who then will also be exposed to the positive posts on the blog. Drawing from the literature on customer advocacy and the stage model theory of information processing in cognitive psychology, we develop a model to decipher the relationship between the extent of negative posts and the overall positive Word of Mouth (WOM) generated by the employee blogs for the firm. An empirical model is developed to account for the inherent non-linearities, endogeneity and unobserved heterogeneity concerns, and potential alternative specifications. Our results suggest that negative posts act as a catalyst to increase the readership of an employee blog, with readership increasing exponentially in the initial stages and then stabilizing. The empirical findings are used to generate an analytical framework that firms can use to formulate employee blogging policies. We illustrate the application of the framework using blogging data from Sun Microsystems.

Key words: blog; employee blogs; bloggers; blogging policies; word-of-mouth; customer advocacy; information processing theory; non-linear models;

JEL Classifications: C10, C23, C51, C52, C80, C87, D78, L10, M50, O33.

INTRODUCTION

> If you want to lead, blog…We talk about our successes- and our mistakes. That may seem risky. But it's riskier not to have a blog.
>
> Jonathan Schwarz, CEO of Sun Microsystems, HBR 2005

> Jonathan Schwartz, CEO, Sun Microsystems, says his number 1 priority isn't spending time communicating; it's ensuring that his communications are broadly received. Blogging has become the most efficient form of communication.
>
> Interview by O. Ryan, Fortune Oct. 30, 2006

Weblogs or Blogs are frequently updated, reverse-chronological ordered entries on a single Web page (Scoble et al. 2006). Blogs gained a lot of popularity during 2004 US presidential elections and it was not long after when businesses also realized the importance of blogs as a communication medium. Today many Fortune 500 companies encourage their employees to blog- IBM and Microsoft have over 2000 employee blogs (Byron et al. 2006) and about ten percent of the workforce of Sun Microsystems maintain blogs (Oliver 2006).

The reasons behind this flurry of blog popularity among firms are manifold. A key reason is that employee blogs give a "human face to a company" (Byron et al. 2006; Conlin et al. 2004; Edelmen et al. 2005; Scoble et al. 2006). For example, in the year 2005 Geico filed a lawsuit against Google and eventually lost the case. However, a false rumor spread that Google and not Geico lost the case. In order to quickly and effectively quell the false rumor, a senior executive from Google chose to personally issue a clarification through a blog. Employee blogs also act as an effective marketing tool as they can generate a significant positive Word of Mouth (WOM) for the company and its products (Byron et al. 2006; Conlin et al. 2004; Oliver 2006; Schwartz

2005; Scoble et al. 2006). For instance, a VP of General Motors (GM) utilized a blog to explain and make the case that GM has the most fuel-efficient model in every car segment (Stephens 2005). This positive WOM from employee blogs is important for companies because it may lead to higher sales (Godes et al. 2004) and may lower advertising cost (Scoble et al. 2006).

As an old admonition goes "Loose Lips Might Sink Ships". This is one of the biggest concerns of firms in permitting their employees to blog without restraints. Employees can criticize the firm or its products, say embarrassing things about coworkers, and promote competitors or their products (Ashmore et al. 2006; Byron et al. 2006; Conlin et al. 2004; Edelmen et al. 2005; Halley 2003). For example in the post shown in Figure 1, a Microsoft employee admits that he likes Macintosh computers and encourages his coworkers to also own them. Such a post may create a lot of negative WOM for Microsoft and may lead to lost sales (Mahajan et al. 1984).

Figure 1: Post of a Microsoft Employee

However some believe that a negative post does not always hurt a firm (Edelmen et al. 2005; Scoble et al. 2006).

David Weinberger in a recent HBR article corroborates this belief:

> She [employee blogger] thus avoid[s] the pitfalls that marketing departments repeatedly walk into. Her willingness to admit fallibility-the pace of daily on-line publishing pretty well ensures that Weblogs have the slapdash quality of first drafts-is ironically the very thing that leads her readers to overlook her mistakes and trust her. (Halley 2003)

Director of Communications GM, Mr. Wiley explains in an interview with WSJ:

> A lot of what blogging is about is authenticity, getting beyond corporate speak and PR, and really creating a conversation. Not being thin skinned and accepting the negatives, that's key. (Brian 2005)

The conventional wisdom expressed in the above quotes is that a negative post is not necessarily bad for the firm. If an employee admits flaws in the company then the credibility of the employee's blog may increase and draw more people to the blog. Since more people may come to the blog and also read positive messages, the net effect may be an increase in the overall positive WOM for the firm (Brian 2005; Byron et al. 2006; Edelmen et al. 2005; Halley 2003; Scoble et al. 2006). This raises the natural questions on whether negative posts help firms in any way, and if so, then under what conditions?

The answer to the above questions can help a firm formulate an appropriate blogging policy for its employees. In practice, firms implement two types of blogging policies. One policy prohibits employees from discussing anything related to the firm on their blogs; we term this *No Blogging Policy* (NB). Firms such as Apple (Retail Division) implement this policy because they fear an outbreak of negative messages (Edelmen et al. 2005). A second type of blogging policy allows employees to write both positive and negative things about firm; we term this *Permissible Negative Blogging Policy* (PNB). Many Fortune 500 companies such as Microsoft, Sun Microsystems, IBM, Adobe, Oracle, Yahoo and many others implement this policy (Edelmen et al. 2005; Schwartz 2005). One could also consider a policy under which employees can write positive posts but not negative posts; we term this *No Negative Blogging Policy* (NNB). It is easy to see that NNB dominates NB but in the worst case scenario NNB will not generate any WOM from employee blogs, which is no worse than NB. Therefore, we use NNB as a benchmark policy instead of NB.

The issue of setting the appropriate blogging policy is important because a wrong choice either adversely affects the reputation of the firm or leads to a less than optimal positive WOM for the firm (Ashmore et al. 2006; Edelmen et al. 2005; Scoble et al. 2006). While earlier work suggests that negative messages could, under certain conditions, benefit the firm overall; however, no systematic estimates exist to empirically quantify these conditions. The literature thus far only provides qualitative discussions on the upsides and downsides of negative posts on the employee blogs. To the best of our knowledge ours is the first attempt to empirically quantify the conditions under which a negative post generates higher net positive WOM and the conditions under which PNB should be preferred to NNB.

Measuring net positive WOM from employee blogs is an important challenge that must be overcome in evaluating appropriate firm blogging policies. While much of the attention in the press and in some academic research has been on mechanisms through which employee blogs can create positive or negative WOM for the firm, the WOM generated from employee blogs has not been quantified. This paper represents a first effort to suggest a framework using econometric analysis for estimating the overall positive WOM from an employee blog. Using this framework, a firm can also identify which of its employees' blogs are assets and which of them are hurting its reputation.

Our results indicate that *ceteris paribus*, the effect of negative posts on the page-views is best explained by the Michaelis-Menten model of chemical kinetics (Bates et al. 1988; Michaelis et al. 1913; Seber et al. 2003). The Michaelis-Menten model is widely used to study chemical reactions where, as the concentration level of a catalyst increases, the reaction rate increases exponentially initially, and then stabilizes after a point. We find that negative posts act as a catalyst to page-views; with an increase in negative posts the page-views increase exponentially initially and then stabilize after a point. Our results suggest that due to this initial exponential increase in page-views there are conditions under which negative posts generate more net positive WOM or create more positive reputation of a firm than if there were no negative posts. To illustrate our suggested framework, we calculate the average net positive WOM from a sample data of employee bloggers of Sun Microsystems under both PNB and NNB. Further, our analysis suggests that Sun Microsystems could benefit from employees posting more negative messages than they currently do.

THEORY AND LITERATURE

Many studies on blogs suggest that readers expect employee bloggers to promote their firms. But if employees blog about the failures of their employer or promote positive aspects of competitors then it may generate more interest and drive up the readership (Brian 2005; Byron et al. 2006; Edelmen et al. 2005; Halley 2003; Scoble et al. 2006). Research on customer advocacy and information processing theory in cognitive psychology lends support to these findings.

The literature on customer advocacy suggests that posting negative information may engender trust in the audience. If an employee provides positive information as well as some negative information, then customers place more trust on the employee and may also recommend that blog to their friends (Krapfel 1985; Urban 2004; Urban 2005; Urban et al. 2000). Negative information here refers to information on how competitor's products are better than the employer's product, drawbacks in the employer's products, and so on. A recent study on a virtual recommender agent that was cosponsored by GM validates the above claim (Urban 2004). This agent recommended cars to customers according to their needs and these recommendations were not limited to cars from the GM portfolio. This online agent on occasion provided negative information about GM to customers by recommending competing products, and customers rewarded this agent by placing more trust on it and by recommending it to their friends. Other research also validates a part of the findings that the increase in trust on a site leads to increase in adoption of the site (McKnight et al. 2002; Pavlou et al. 2006).

Literature in cognitive psychology suggests that unexpected and interesting information will trigger better assimilation in the reader. To the extent that negative posts by a company's employee are unexpected or interesting, such posts are more likely to capture the attention of readers. The following passage elucidates the cognitive psychology of a customer (Lynch et al. 1982):

> Information that is novel or unexpected seems to capture one's attention, is processed more extensively, and subsequently is much more likely to be recalled than information that is redundant or expected to appear in a given context.

This view is also supported by the stage model theory of information processing, which suggests that the probability of getting information from the first stage of sensory memory to short term memory will be higher if the information is interesting to the reader (Argyres 1999; Atkinson et al. 1968; Francalanci et al. 1998; Gattiker et al. 2005) . This transfer of information to second stage is important because without this step, the information may not make any influence on the reader.

From the above survey, it is reasonable to expect that if an employee publishes some negative posts as well as positive posts, then (i) that employee's blog is likely to attract more readers, and (ii) those readers are more likely to process the information on the blog. Readers are then exposed to the positive as well as negative posts on the blog, which generates positive as well as negative word of mouth (WOM) for the firm. We present a novel approach to measuring the relationship between negative posts on a blog and readership of the blog. We then model the WOM generated by employee blogs as a function of the impact of the blog on a reader and the readership of the blog, and explore the conditions under which a firm can maximize the net positive WOM from its employee blogs.

Previous literature estimates WOM by measuring the participation of people in the discussion of a certain product (Brown et al. 1987; Godes et al. 2004; Mayzlin 2006). Specifically, prior work captures the online WOM for a product on two dimensions: volume (number) of posts about the product, and the dispersion (variance or entropy) of posts across communities/groups (Godes et al. 2004). Such measures are inadequate for our purposes, because both volume and entropy are aggregate level variables and are incapable of capturing the blog-level dynamics of the effect of negative posts on the readership of a blog. Moreover, variables such as volume and entropy do not measure the impact of these posts, because some posts may discuss important aspects of a product and have higher impact, whereas other posts may discuss trivial details with little or no impact.

MODELING WORD OF MOUTH

Given that a firm seeks the maximum possible positive exposure (the “Net Positive WOM”) from its employees’ blogs, this goal can be decomposed into two parts: (a) the firm seeks to positively influence each reader of an employee’s blog, and (b) the firm seeks to attract the maximum possible readership for each blog. We first use empirical analysis to derive a functional form relating the impact of posts in a blog to the extent of readership (number of readers) of the blog. Positive WOM is then modeled as a product of the extent of readership (using this empirically derived functional form) and the impact of a blog on each reader. In this context, we first define a few essential terms.

The net impact of a blog

The extent of positive influence a blog has on a reader depends on the positive and negative posts on that blog. One way to measure *positive* (*negative*) *impact of a blog* can be the number of positive (negative) posts. But then not all posts are similar in terms of their impact: some posts might be on trivial matters, and other posts might be on more significant matters. For example, consider figure 2, which contains two posts by a Microsoft employee.

Post A:

Technorati saves Maryam's day

Google didn't help Maryam (my wife). MSN Search didn't help. What did?

Technorati!

Post B:

Yes, our Web site branding sucks

It's not nice to tell our branding folks that their work sucks, but sorry, it's so obvious now that I just can't pretend to like it. Dare Obasanjo's post demonstrates what's wrong very well.

Figure 2: Posts of a Microsoft Employee

Both of the above posts are negative for Microsoft, but both posts will likely generate different levels of negative opinions about Microsoft. In post A, which is somewhat trivial, the blogger promotes the idea of selling goods through blogs and discusses how his wife found tickets for an event through Technorati (a blog search engine) but not through MSN (Microsoft's search engine). In post B, the same blogger criticizes the branding strategies of Microsoft in strong language. It is likely that post B will generate higher interest among readers and hurt the reputation of Microsoft more than post A.

To capture this notion that different posts generate different levels of interests (and influence) among readers, we use the term "*impact of a post*". The higher the impact of a post, the more is the interest generated by that post among readers. We formulate an objective measure of this impact. Consider an individual employee blogger and her blog. This blog contains several posts (and typically, one new post appears per day). Every post has its unique URL (called its "permalink") that other sites can cite or link to. Each such citation or link from an external site is an "inbound link." Search engines such as Technorati provide a count of the number of such inbound links to each permalink. We use the number of inbound links to each post (permalink) as a measure of the impact of that post. However, not all external sites are equal; some are more influential than others. A measure of an external site's influence is the number of (other) external sites that link to it ("2^{nd} level links"). Therefore, in order to measure the *impact of a post*, we consider the inbound links, along with the 2^{nd} level links to each of these inbound links. The intuition behind this measure is that a post of higher impact is likely to get higher inbound links. Google and other search engines also use the same logic in ranking the importance of an URL (Google.com 2004). As Google founders, Brin and Page explain:

> Intuitively, pages that are well cited from many places around the web are worth looking at. Also, pages that have perhaps only one citation from something like the Yahoo! homepage are also generally worth looking at. If a page was not high quality, or was a broken link, it is quite likely that Yahoo's homepage would not link to it.

Based on the above definition of the impact of a post, the *impact of positive post* is the level of positive opinion formed by a reader after reading a positive post, and is measured by the summation of inbound links and 2^{nd} level links to a positive post. The total impact of all positive posts in a blog (for blogger *i* in period *t*) is denoted by $I_{i,t}^{+}$.

Analogously, the *impact of negative post* refers to the level of negative opinion formed after reading a negative post, and is measured by the summation of inbound links and 2^{nd} level links to a negative post. The total impact of all negative posts in a blog is denoted by $I^{-}_{i,t}$. We define the total impact of neutral posts in a blog similarly and denote it as $I^{0}_{i,t}$.

The *net positive impact of a blog* is the difference between the total *impact of positive posts* and the total *impact of negative posts.* This tries to capture the level of net positive opinion formed by a reader after reading an employee blog, which may consist of positive, negative and neutral posts. The net positive impact of a blog is given by: $I^{P}_{i,t} = I^{+}_{i,t} - I^{-}_{i,t}$.

Net Positive WOM can now be defined more formally as the product of *net positive impact* of a blog on a reader and the *number of readers.*

$$\textit{Net Positive WOM} = \textit{Net positive impact} * \textit{Number of readers} \qquad (1)$$

The firm's goal is to formulate blogging policies (specifically, the extent to which negative blogging may be allowed) that maximize net positive WOM. Based on the earlier discussion of the literature on consumer advocacy and cognitive psychology, if negative posts increase the number of readers of a blog (which we verify empirically), then the firm faces a tradeoff: permitting more negative posts gets more readers, but each reader gets a less favorable overall impression of the firm. At lower levels of negative posts, the net positive WOM increases. But as the proportion of negative posts increases, at some point the increase in readership is offset by the unfavorable impression each reader gets, so that the net positive WOM starts to decrease. The firm's task is to set policies for employee blogging that nudge the mix of positive and negative posts toward this optimum.

Factors that affect the number of readers of a blog

We now address the task of relating blog postings to the number of readers of a blog, to examine empirically whether negative posts increase blog readership. The *number of readers* of the blog in a period (each period lasts a day) is measured by *page-views* to the blog in that period. The *page-views* of a blog can be affected by (a) the popularity of the blog, as well as (b) the content (posts) of the blog.

Regarding (a), while an individual post could attract readers, so could the overall popularity of the blog. To control for a blog's popularity, we take two blog-level variables as measure of popularity of the blog in previous period: *page-views lag* (Battelle 2005) and *XML feed subscription* data (Massie et al. 2006). The rationale for using *page-views lag* is that a more popular blog should have higher page-views during past periods. *Page-views lag* can also serve as a proxy for the employee specific latent variables that make some blog more popular than the other (Wooldridge 2001). "*XML feed subscribers*" refers to the number of people who subscribe to a blogger's posts. Whenever this blogger publishes a post, the headline of the post is sent to the subscriber automatically. The number of people subscribing to a blog is likely to be proportional to the popularity of a blog, and so we use *XML feed subscribers* as a measure of popularity of the blog. *XML feed subscribers* is used by search engines such as Ask.com to measure the popularity of blog (Massie et al. 2006).

Regarding (b), to get a numerical measure of the content of each blog, we define two variables: *impact of a blog* and *ratio of negativity*. We use the term *impact of a blog* to refer to the total interest of a reader on a blog and define it (for blogger *i* in period *t*) as the sum of *impact of posts* on the blog ($I_{i,t} = I^{+}_{i,t} + I^{-}_{i,t} + I^{0}_{i,t}$), without regard to whether the posts are positive, negative or neutral (*impact of a blog* differs from *net positive impact* of a blog). This term is based on the premise that posts on significant topics generate more interest and attract more readers than trivial posts. In order to measure the extent of negative posts in each blog, we define the *ratio of negativity,* which is the ratio of the *impact of negative posts* to the *impact of positive* *posts* ($R_{i,t} = \frac{I^-_{i,t}}{I^+_{i,t}}$). This term measures the proportion of negative to positive posts on a blog.

Table 1 provides a summary of these variables.

Table 1: Definitions and Interpretations of Independent Variables

Variable	Definition	Interpretation
Page-views Lag ($P_{i,t-1}$)	One period lag of page-views	High *page-views lag* indicates higher popularity of a blog
Impact of a blog ($I_{i,t}$)	Sum of *impact of posts* on a blog	High *impact of a blog* indicates that a blog generates higher interest among readers
Ratio of negativity ($R_{i,t}$)	Sum of *impact of negative posts*/ Sum of *impact of positive posts*	High *ratio of negativity* indicates higher proportion of negative content on a blog
Subscribers ($S_{i,t}$)	XML feed subscribers of the blog	More *subscribers* indicate higher popularity of a blog

We find that the effect of negative posts on page-views is best explained by the Michaelis-Menten model of chemical kinetics (Michaelis et al. 1913). Negative posts appear to play a role similar to that of a catalyst in a chemical reaction: when the proportion of negative posts is small, an increase in negative posts appears to act as a catalyst to exponentially increase page-views for the blog, but as the proportion of negative posts increases further, the increase in page-views tapers off and stabilizes.

In the next section, we describe our dataset and present some interesting patterns in the data. We compare alternative model specifications to predict page-views per day using the independent variables discussed earlier, and find that the Michaelis-Menten specification offers the best fit. This suggests a functional form relating positive and negative posts to a blog's readership (page views). We then use this functional form in a model of net positive WOM to analytically derive the optimum employee blogging policy for the firm. Finally, we conclude with a discussion of the limitations of this study and future research possibilities.

ECONOMETRIC SPECIFICATION OF PAGE-VIEWS

Data

We randomly selected blogs of eleven employees of Sun Microsystems, and collected data for thirteen weeks starting from February 2006, which resulted in a total of 1001 observations. We split the data set into two subsets: the first twelve weeks' data (924 observations) was used as the training dataset, to estimate the model parameters; the last week's data (77 observations) was used as a test dataset. We obtained the citation/inbound links data for every post in the data set from Technorati.com; daily *XML feed subscription* data for each of the bloggers from Bloglines.com[1]; and daily *page-views* data for each of the bloggers from www.blogs.sun.com. Three graduate students helped us categorize the posts into three categories: positive posts, negative posts and neutral posts. A post was classified as positive if it promoted Sun Microsystems – the company or any its products; or if it was critical of a competitor or their products. Conversely, a post was classified as negative if it criticized Sun Microsystems or its products; or if it promoted a competitor or their products. Posts that did neither of the above were categorized as neutral. Every post was categorized by two graduate students and in case of a tie the third student's decision was sought. The inter-student reliability for the categorization of posts is 0.92, which suggests a high level of agreement about the category of a post.

Exploratory data analysis suggests that bloggers on average post less than one negative post in every set of fifteen posts. On average a blogger posts once every day and half. A post is archived when fifteen subsequent posts have been made on the blog. Therefore, once a negative post is made it stays on the blog page for about twenty two days. This helped us observe a large

[1] Ask.com also uses the same subscription data set to rank the bloggers according to their popularity Massie, R., and Kurapati, K. "There's Blogs (and Feeds!) In Our Search!," Blog.ask.com, 2006.

number of days when there was a negative post on the blog page. This revealed some interesting facts:

(1) Negative posts have disproportionate impact

We found that the *impact of a negative post* on average is double the *impact of a positive post* (see Table 2). The *impact of negative posts* is therefore disproportionate to their frequency of appearance. The number of negative (positive) posts on average is 0.92 (11.98); however, the *impact of negative (positive) posts* on a blog on average is 2.17 (13.56). The *impact of a positive post* on an average is slightly higher than 1 and the *impact of a negative post* on average is approximately 2.

(2) The *net positive impact of a blog* is highly positive

The overall *impact of positive posts* is higher than the *impact of negative posts*. This can be seen in Table 2 where the mean *impact of positive posts* is 13.56 whereas the mean *impact of negative posts* is 2.17.

Table 2 provides the descriptive statistics for all variables.

Table 2: Descriptive Statistics

Variable	**Mean**	**Min.**	**Max.**	**Std. Dev.**
Page-views $(P_{i,t})$	2402.77	796	4587	820.34
Impact of a blog $(I_{i,t})$	17.19	5.94	36.69	5.74
Impact of Positive Posts $(I^{+}_{i,t})$	13.56	4.09	28.66	4.39
Impact of Negative Posts $(I^{-}_{i,t})$	2.17	0	15.74	2.53
Impact of Neutral Posts $(I^{0}_{i,t})$	1.34	0	6.53	1.48
Ratio of Negativity $(R_{i,t})$	0.15	0	1.86	0.36
No. of Positive Posts	11.98	4	15	2.34
No. of Negative Posts	0.92	0	6	1.08
No. of Neutral Posts	2.09	0	10	2.46
Subscriber $(S_{i,t})$	79.38	51	167	57.14

Econometric specification

Consider the following linear model specification:

$$P_{i,t} = \alpha_0 + \alpha_1 P_{i,t-1} + \alpha_2 I_{i,t} + \alpha_3 I_{i,t} R_{i,t} + \alpha_4 R_{i,t} + \alpha_5 S_{i,t} + \varepsilon_{i,t} \qquad (2)$$

If this specification is appropriate then Pooled OLS will give consistent and efficient estimates. Table 3 presents the estimated results from the training data set and reports that the *XML feed subscription*, *Ratio of negativity* and the *constant* are insignificant. Using a Wald Test, we fail to reject the null hypothesis that these three variables together are non-significant (see bottom of Table 3 for results). Multicollinearity is not a problem, since the maximum Variance Inflation Factor (VIF) in specification (2) is 2.4, which is much less than 10. If a specification has an interaction term and there is a multicollinearity problem between the independent variables and the interaction term, then typically the interaction term comes out to be non-significant (Jaccard et al. 2003). But in this case, it comes out to be strongly significant; hence this serves as another check for the conclusion that multicollinearity is not a problem here. Since it is a good practice to retain the main effect if the interaction term is present, we do not drop $R_{i,t}$ from our specification (Jaccard et al. 2003; Wooldridge 2001). Therefore, the new specification based on the discussion is the following:

$$P_{i,t} = \alpha_1 P_{i,t-1} + \alpha_2 I_{i,t} + \alpha_3 I_{i,t} R_{i,t} + \alpha_4 R_{i,t} + \varepsilon_{i,t} \qquad (3)$$

The above analysis tells us that

- *The Ratio of negativity* does not directly affect the *page-views* but moderates the effect of *impact of a blog* on the *page-views*.
- *The XML Feed Subscription* and the intercept are non-significant. After controlling for the *popularity of a blog* through *page-views lag*, *XML Feed Subscription* does not significantly increase the explanatory power of the model. It is not surprising that the

intercept term is insignificant; having a positive intercept will mean that a blogger will get *page-views* even if she has no post and no previous *page-views*, which does not make sense.

Another possible specification (4) is to decompose the *impact of a blog* into the *impact of positive, negative* and *neutral posts*, and their interaction with the *ratio of negativity*.

$$P_{i,t} = \alpha_1 P_{i,t-1} + \alpha_2 I^{+}_{i,t} + \alpha_3 I^{-}_{i,t} + \alpha_4 I^{0}_{i,t} + \alpha_5 I^{-}_{i,t} R_{i,t} + \alpha_6 I^{0}_{i,t} R_{i,t} + \alpha_7 R_{i,t} + \varepsilon_{i,t} \qquad (4)$$

This specification is more flexible than (3) but the downside is that we need to estimate more number of parameters, which leads to larger standard error in *page-views*. Moreover, we checked for the linear restrictions $\left(\alpha_3 - \alpha_2 - \alpha_5 = 0,\ \alpha_4 - \alpha_2 = 0,\ \alpha_5 - \alpha_6 = 0\right)$ that reduce the specification (4) to specification (3) and we failed to reject these linear restrictions *(p>0.18)*. Therefore, we continued with the specification (3), which is a more parsimonious model.

We now test the assumption that specification (3) has captured all the non-linearity in the data. Using Ramsey's RESET test, we reject the null hypothesis that all non-linearity pattern is captured in the specification (see bottom of Figure 3 for results). Based on the Wald test results the model (3) can be reduced to the following specification (ignoring the error term and the non-significant *ratio of negativity* main effect).

$$P_{i,t} = \alpha_1 P_{i,t-1} + \alpha_2 I_{i,t} + \alpha_3 I_{i,t} R_{i,t}$$

$$\Leftrightarrow \frac{P_{i,t} - \alpha_1 P_{i,t-1}}{I_{i,t}} = \alpha_2 + \alpha_3 R_{i,t} \qquad (5)$$

The interpretation of Equation (5) is that, ceteris paribus, if unit *impact of positive posts* is replaced by unit *impact of negative posts* then the *page-views* should increase linearly with the *ratio of negativity*.

Figure 3 is a plot between L.H.S. of (5) and *Ratio of Negativity*. It shows that there is a non-linear pattern (initial exponential increase followed by stability) instead of a linear line. This suggests that this non-linearity may be the cause of failure of Ramsey's RESET test. We find that such curves are best modeled by the Michaelis-Menten model of chemical kinetics (Bates et al. 1988; Michaelis et al. 1913; Seber et al. 2003). The impact of negative posts on the page views for a blog appears to share some similarities with the process by which a catalyst speeds up a chemical reaction, for which a nonlinear model such as the Michaelis-Menten model is appropriate. Negative posts appear to be most effective when present in small quantities: *page views* increase with the *ratio of negativity* exponentially at first, but at a rapidly declining rate, and stabilize after a point.

Table 3: Effect of negative post on page-views

Parameter	Pooled OLS Estimates
$P_{i,t-1}$	0.765*** (0.066)
$I_{i,t}$	21.467*** (5.966)
$R_{i,t}$	9.956 (10.752)
$I_{i,t}R_{i,t}$	15.357*** (0.929)
$S_{i,t}$	0.359 (0.419)
Adj. R^2	80.1%
N	924
Pr>F	0.000
RMSE	366.135

***=p<0.01, **=p<0.05, *=p<0.1, robust std. errors in parenthesis

Wald Test Results (Ratio, Subscription, Constant are insignificant):
F (3, 918) =1.127, (p>0.33) – cannot reject insignificance

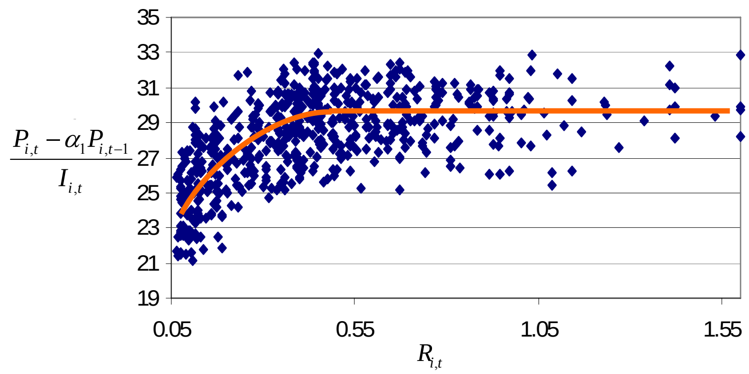


Figure 3: $\frac{P_{i,t} - \alpha_1 P_{i,t-1}}{I_{i,t}}$ vs. $R_{i,t}$

Ramsey Reset Test Results (no non-linearity):
F (3, 915) =3.558, (p<0.014) – there is non-linearity

Based on this observation, we changed our model specification to the following and estimated the model using Pooled Non-linear Least Square (NLS). Results are given in Table 4.

$$P_{i,t} = \beta_1 P_{i,t-1} + \beta_2 I_{i,t} + \beta_3 I_{i,t} \frac{R_{i,t}}{R_{i,t} + \beta_4} + \beta_6 R_{i,t} + \varepsilon_{i,t} \qquad (6)$$

Table 4: Effect of negative post on page-views

Parameter	Pooled NLS Estimates
$P_{i,t-1}$	0.749*** (0.073)
$I_{i,t}$	18.921*** (5.987)
$I_{i,t}R_{i,t}/(R_{i,t}+\beta_4)$	20.532*** (1.857)
β_4	0.120*** (0.027)
$R_{i,t}$	10.893 (12.539)
N	924
Pr>F	0.000
RMSE	243.175

***=p<0.01, **=p<0.05, *=p<0.1, robust std. errors in parenthesis

Ramsey RESET Test Results (no non-linearity): F (3, 915) =1.469, (p>0.22) – there is no non-linearity

DWH Test Results (instrument is *two period lag of impact*): F (3, 893) =2.714, (p<0.045) – reject exogeneity assumption

Using Ramsey's RESET test, this time we failed to reject the null hypothesis *(p>0.22)* that the entire non-linearity in the data is captured (Note: Using Ramsey's test requires us to retain the slope intercept in the model, even though our model specification does not have a slope intercept. See bottom of Table 4 for results.).

We note that there could be a possible simultaneity in *page-views* and *impact of a blog* for the same period: if page-views for a post increase in a period then that post is likely to get more inbound-links, which increases the *impact of the blog*. We require an instrument variable that is correlated with the *impact of a blog* but not with the shock in *page-views* (sudden increase/decrease in page-views: i.e. the idiosyncratic error). A lag in the *impact of a blog* could be a valid instrument if the model is dynamically complete or sequentially exogenous (Wooldridge 2001); i.e. if the shock in *page-views* in a period will not change the *impact of a blog* in past periods. Literature in econometrics has shown that if the dynamic completeness condition is violated then there will be autocorrelation in idiosyncratic errors. We tested for autocorrelation (Wooldridge 2001) but failed to reject the null hypothesis (*p>0.29*), suggesting that the assumption of dynamic completeness in our model is justified. Therefore, we can use *lags of impact of a blog* as valid instruments for *impact of a blog*.

We choose a two period lag of *impact of a blog* as an instrument because it satisfies both the conditions for being an instrument. First of all, two period lag is correlated with the present *impact of a blog* because both periods have many common posts. Secondly, the two period *lag of impact of a blog* may not be correlated with the shock in *page-views* because the shock in *page-views* in the present period may not change the two period lag of *impact of a blog.* This is because the conventional wisdom of bloggers suggests that if posts are old then the readers typically have come across the information from some other source; hence, they may not want to cite an old post (Byron et al. 2006; Scoble et al. 2006). In our data too we observe that most of the citations (92.1%) to a post are received in the first two days of the posting. Therefore, we tested the exogeneity assumption of *impact of a blog* using two period lag as instrument for the *impact of a blog* in an application of Durbin-Wu-Hausman test based on Gauss-Newton regression (Davidson et al. 1993) (see Appendix B for further details). The test rejected the exogeneity assumption (See bottom of Table 4 for results) and hence, we use the method of instrument variables (IV) as a solution to the problem of endogeneity. Two period *lag of impact of a blog* acts as an instrument variable for *impact of a blog* to yield Pooled nonlinear instrument variable (NLIV) estimates. If we compare the coefficients between Pooled NLS and Pooled NLIV estimates in Tables 4 and 5 (columns 1), we observe that the effect of *ratio of negativity* has gone down and the effect of *impact of a blog* alone has gone up. Clearly, if we had not controlled for endogeneity due to simultaneity then we will have overestimated the role of *negative post* and our analysis would have been biased towards the importance of having higher negative posts.

Testing for the stability of the estimates

Page-views may depend, ceteris paribus, on the time; we therefore introduce three time dummies for the third, sixth and the ninth week and blog dummies for each employee blog. Using Wald test, we fail to reject the null hypothesis that the time and blog dummies together are insignificant (see bottom of Table 5 for results). Table 5 compares the stability of estimates with (Table 5, column 2) and without (Table 5, column 1) the presence of control variables. Using the Lagrange-Multiplier (LM) test we failed to reject any significant effect of unobserved heterogeneity (see bottom of Table 5 for results) (Wooldridge 2001); but for comparing the stability of estimates we also find First-Differencing (FD) IV estimates (Table 5, column 3). The parameters and the standard errors in all three cases are similar suggesting that our model (6) is good to predict *page-views* and we do not need to add any more control variables.

Table 5: Effect of negative post on hits

Parameter	Pooled NLIV Estimates	Pooled NLIV Estimates	FD (IV) Estimates
$P_{i,t-1}$	0.797*** (0.031)	0.770*** (0.118)	0.783*** (0.121)
$I_{i,t}$	21.716*** (2.385)	22.215*** (2.479)	23.129*** (3.547)
$I_{i,t}R_{i,t}/(R_{i,t}+\beta_4)$	9.994*** (0.869)	10.040*** (1.149)	9.421*** (1.394)
β_4	0.131*** (0.012)	0.137*** (0.022)	0.185*** (0.07)
$R_{i,t}$	12.184 (15.082)	21.024 (34.403)	59.287 (96.095)
Control Variables	--	Time & Blog	--
N	902	902	891
Pr>F	0.000	0.000	0.000
RMSE	243.191	245.111	18.364++

***=p<0.01, **=p<0.05, *=p<0.1, robust std. errors in parenthesis
++ RMSE under FD is not comparable to others because the dependent variable is the difference in consecutive *page-views*

Wald Test Results (Time & Blog are insignificant):
F (14, 883) =1.154, (p>0.30) – cannot reject insignificance

LM Test (Unobserved heterogeneity insignificant):
$\chi^2(1)=0.83$, (p>0.36) – cannot reject insignificance

Comparing base model with other potential models over test data

We consider the following three model forms and contrast the results.

1. An initial exponential increase followed by stability

This is the case when, *ceteris paribus*, the increase in the *impact of negative posts* leads to exponential increase in *page-views* initially, and after a certain point, *page-views* do not change. We take two model specifications for this case- a base model that uses Michaelis-Menten model and a model that uses the exponential function (Bates et al. 1988; Seber et al. 2003).

2. An initial exponential increase followed by a decrease

This is the case when, *ceteris paribus*, the increase in the *impact of negative posts* leads to exponential increase in *page-views* initially, and after a certain point, *page-views* start decreasing as the extent of negative blogs increases. Such a case may happen if readers start to view the blogger as a complainer because of high negative content on a blog and may no longer consider the blog trustworthy. We consider two model specifications for this case – a refinement of the base model that uses Michaelis-Menten model, and the quadratic growth model (Bates et al. 1988; Seber et al. 2003).

3. Linear increase

This will be the case when, ceteris paribus, the increase in the *impact of negative posts* leads to linear increase in *page-views.*

We also test for logarithmic specification but that model performs the worst of all the considered specifications. Therefore, for brevity we do not report the model based on logarithmic specification. The final selection criterion is the out-of-sample criterion for the models: Root Mean Square Error (RMSE) and Mean Absolute Deviation (MAD). We do not use in-sample criteria such as Adjusted R^2 or Akaike's Information Criterion (AIC) because they have been criticized in the literature for selecting models that may have been tuned to the noise in the data (Davidson et al. 1993; Wooldridge 2001).

Model comparisons shown in Table 6 indicate that our base model (6) based on the Michaelis-Menten model of chemical kinetics offers the best fit. In the next section we will model how the functional form suggested by the Michaelis-Menten specification can help a firm in setting the optimal blogging policy.

		Table 6: Comparison of Potential Models				
		(Ceteris Paribus) With the increase in *Impact of Negative Posts, Page-Views* initially increase				
		exponentially & after a point stabilize		**exponentially & after a point fall**		**linearly**
Model		$P_{i,t} = \beta_1 P_{i,t-1} + \beta_2 I_{i,t} + \beta_3 I_{i,t} \frac{R_{i,t}}{R_{i,t} + \beta_4} + \beta_6 R_{i,t}$	$P_{i,t} = \beta_1 P_{i,t-1} + \beta_2 I_{i,t} + \beta_3 I_{i,t} \exp(\beta_4 R_{i,t}) + \beta_6 R_{i,t}$	$P_{i,t} = \beta_1 P_{i,t-1} + \beta_2 I_{i,t} + \beta_6 R_{i,t} + \beta_3 I_{i,t} \frac{R_{i,t}}{R_{i,t} + \beta_4 + \beta_5 R_{i,t}^2}$	$P_{i,t} = \beta_1 P_{i,t-1} + \beta_2 I_{i,t} + \beta_3 I_{i,t} R_{i,t} + \beta_4 I_{i,t} R_{i,t}^2 + \beta_6 R_{i,t}$	$P_{i,t} = \beta_1 P_{i,t-1} + \beta_2 I_{i,t} + \beta_3 I_{i,t} R_{i,t} + \beta_6 R_{i,t}$
IV Estimates (Robust Standard Errors)	β_1	0.797*** (0.031)	0.785*** (0.042)	0.795*** (0.037)	0.801*** (0.045)	0.804*** (0.037)
	β_2	21.716*** (2.385)	30.244*** (4.157)	21.926*** (2.404)	22.834*** (2.947)	24.061*** (2.536)
	β_3	9.994*** (0.869)	-8.344*** (1.009)	10.085*** (0.821)	15.403*** (2.545)	6.301*** (0.637)
	β_4	0.131*** (0.012)	-5.917*** (0.32)	0.143*** (0.024)	-7.127*** (0.887)	
	β_5			0.013 (0.055)		
	β_6 ++	12.184 (15.082)	34.248 (45.819)	17.853 (21.175)	21.116 (32.31)	39.172 (47.281)
Training Data Set	**RMSE**	243.191	287.425	261.476	315.297	356.724
	MAD	225.473	278.305	252.740	302.656	341.992
Testing Data Set	**RMSE**	390.607	465.614	402.456	494.961	576.776
	MAD	373.232	458.153	398.396	482.003	562.551
***=p<0.01, **=p<0.05, *=p<0.1						
++ Main effect of *Ratio of Negativity* is insignificant (p>0.25) in all models						

POLICY ANALYSIS

A firm can choose a policy of *No Negative Blogging* (NNB) that completely forbids any negative blogging by employees, or a less restrictive policy *Permissible Negative Blogging* (PNB) that allows for some negative blogging. If a firm chooses PNB, the net positive WOM for the firm is given by:

$$Net\ Positive\ WOM = \left(I_{i,t}^{+} - I_{i,t}^{-}\right) * P_{i,t} \tag{7}$$

In this equation, $P_{i,t}$ is given by specification (6). Using specification (6) and the definitions given in Table 1, we get Equation (8), which gives us *page-views* received by a blogger in a period:

$$P_{i,t} = \beta_1 P_{i,t-1} + \beta_2\left(I_{i,t}^{+} + I_{i,t}^{-} + I_{i,t}^{0}\right) + \beta_3\left(I_{i,t}^{+} + I_{i,t}^{-} + I_{i,t}^{0}\right)\frac{I_{i,t}^{-}}{I_{i,t}^{-} + \beta_4 I_{i,t}^{+}} + \varepsilon_{i,t} \tag{8}$$

In the Michaelis-Menten functional form, each variable applies to an individual blogger on a daily basis. However, for simplicity we now replace each variable in equations (7) and (8) with its average value across bloggers across time to get *average net positive WOM* from PNB. This averaging of variables could be justified considering that as a practical matter, under a policy that permits partial negative blogging, bloggers could overshoot or undershoot the optimal level of negative posts from time to time. As long as the average level of negative posts across bloggers and across several periods is close to the optimal level, the purpose of the policy is served. Therefore, we have

$$Average\ Net\ Positive\ WOM \text{ from BW, } WOM_{PNB}^{+} = \left(\overline{I}_{i,t}^{+} - \overline{I}_{i,t}^{-}\right) * \overline{P} \tag{9}$$

Where

$$\overline{P} = \frac{1}{1-\beta_1}\left(\beta_2\left(\overline{I}^{+} + \overline{I}^{-} + \overline{I}^{0}\right) + \beta_3\left(\overline{I}^{+} + \overline{I}^{-} + \overline{I}^{0}\right)\frac{\overline{I}^{-}}{\overline{I}^{-} + \beta_4 \overline{I}^{+}}\right) \tag{10}$$

We also suppress the main effect of *ratio of negativity* because it is insignificant in all the models. By ignoring this term we underestimate the page-views under PNB and eventually, *net positive WOM* under PNB. This makes our analysis more conservative.

Simplifying from equations (9) and (10) gives the *average net positive WOM* for a firm under PNB in terms of freely available data.

$$WOM_{PNB}^{+} = \left(\overline{I}_{i,t}^{+} - \overline{I}_{i,t}^{-}\right) * \frac{1}{1-\beta_1}\left(\beta_2\left(\overline{I}^{+} + \overline{I}^{-} + \overline{I}^{0}\right) + \beta_3\left(\overline{I}^{+} + \overline{I}^{-} + \overline{I}^{0}\right)\frac{\overline{I}^{-}}{\overline{I}^{-} + \beta_4 \overline{I}^{+}}\right) \quad (11)$$

Under NNB, a blogger may not put any negative post, and so the *average net positive WOM* for a firm under NNB can be calculated by putting the *average impact of negative posts* equal to zero, i.e. $\overline{I}^{-} = 0$, in Equation(11).

$$WOM_{NNB}^{+} = \overline{I}_{i,t}^{+} * \frac{1}{1-\beta_1}\left(\beta_2\left(\overline{I}^{+} + \overline{I}^{0}\right)\right) \quad (12)$$

Proposition: There exists a cutoff value of the average impact of negative posts $\overline{I}^{-*}$ such that (a) for all values of the *average impact of negative posts* $\overline{I}^{-} \varepsilon \left(0, \overline{I}^{-*}\right)$, the Permissible Negative Blogging (PNB) policy creates more *average net positive WOM* for the firm than the No Negative Blogging (NNB) policy, whereas (b) for $\overline{I}^{-} > \overline{I}^{-*}$, the NNB policy offers a higher *average net positive WOM* than PNB.

Proof: See Appendix C

This proposition suggests that disallowing negative posts is not always the best strategy. The basic intuition is that when there is a small amount of negative posts on a blog, it attracts more readers, who are also exposed to the (more numerous) positive posts on the blog. As more negative posts are added, more readers are attracted at a decreasing rate (as shown empirically in the earlier section), and each of these readers is exposed to a lesser proportion of positive posts. At some point, the positive WOM benefit from having negative posts starts to decrease. Eventually, having a large body of negative posts makes for a worse WOM than having no negative posts at all, making the NNB policy better. Please note that a sufficient condition for the above proposition to hold is that $\beta_3 > \beta_2\beta_4$. Figure 4 shows a comparison of the WOM generated by the PNB and NNB policies.

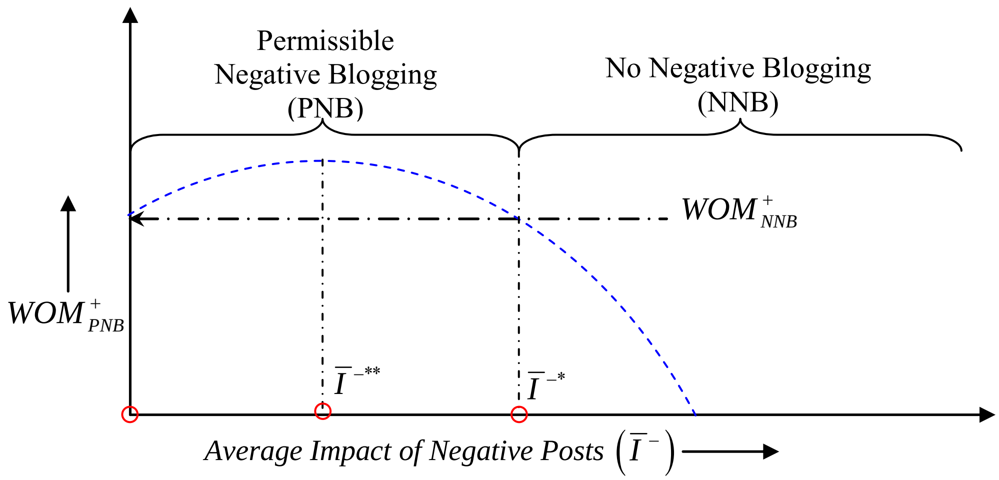


Figure 4: Policy Decisions corresponding to *average impact of negative posts*

Policy recommendations

Based on the proposition, it is clear that a firm should adopt a more lenient blogging policy (PNB) if $\bar{I}^{-} \varepsilon \left(0, \bar{I}^{-*}\right]$. If $\bar{I}^{-} > \bar{I}^{-*}$, then the firm should adopt a policy restricting any form of negative blogging (NNB). Further, if $\bar{I}^{-} \varepsilon \left(0, \bar{I}^{-*}\right]$ and a firm already has a PNB policy, then (i) if $\bar{I}^{-} \varepsilon \left(0, \bar{I}^{-**}\right)$ then the firm should further relax its blogging policy and encourage its employees to put more negative posts than usual; (ii) if $\bar{I}^{-} \varepsilon \left(\bar{I}^{-**}, \bar{I}^{-*}\right)$ then the firm should restrict its blogging policy and discourage its employees from some forms of negative blogging. This will help the firm encourage a level of blogging that is closer to the optimal level $\bar{I}^{-**}$.

Illustration of the policy implementation to the case of Sun Microsystems

Firms that encourage blogging by their employees can easily access the kind of data required for the above analysis. We now apply our analysis to data from Sun Microsystems, and illustrate how suitable recommendations can be made for the optimal blogging policy. Using specification (6) we calculate the upper limit of *average impact of negative posts*, $\bar{I}^{-*} = 6.70$ ($R = 0.49$). Beyond this point, employee blogs may still generate *positive WOM (under PNB),* but can generate more *positive WOM* if employees refrain from writing anything negative about the firm (NNB). In our data set we find that the actual *average impact of negative posts* maintained is $\overline{I}^{-} = 2.17$ ($R = 0.16$), which is far less than the calculated upper value. Therefore, the sample data set and the results based on our suggested framework suggest that the Permissible Negative Blogging policy is the apt policy for Sun Microsystems. Further, the optimal *average impact of negative posts* for the whole data set, $\overline{I}^{-**} = 2.58$ ($R = 0.19$), which is more than the present *average impact of negative posts*. Therefore, in order to move closer to the optimal solution, Sun Microsystems should encourage its employees to increase the extent of negativity on their blogs. The exact nature of policy changes could pertain to specific areas (e.g. permit more negative blogging on product policy, but not on HR policy), but that is beyond the scope of this paper.

CONCLUSIONS & FUTURE RESEARCH

Our results indicate that negative posts do not always harm a firm and can under some conditions create more net positive word of mouth for firms. This result suggests that a No Negative Blogging policy or No Blogging policy need not always be the best solution for firms. Firms can decide a suitable blogging policy using our suggested framework; using a sample from Sun Microsystems employee blogs, this framework indicates that Permissible Negative Blogging policy is the best policy for Sun Microsystems. Further, our analysis suggests that an increase in the proportion of negative posts from its employees is in fact beneficial for Sun Microsystems. Though we have considered only the case of uniform blogging policy across employees; our suggested framework is fairly general and the blogging policy decision can be customized for bloggers in a firm. One of our main empirical finding is that, ceteris paribus, negative posts act as a catalyst to exponentially increase the *page-views* (reaction rate) initially with the increase in their *impact* (concentration of catalyst) and after a point the *page-views* do not increase (reaction rate stabilizes).

One basic problem with policy comparison analysis is that the behavior of an economic agent may not remain same under different policies. In our case, under restrictive policies such as No Negative Blogging policy, an employee may not publish negative posts as per policy, but she may also publish fewer positive posts because of dissatisfaction created from the restriction (Ashmore et al. 2006; Scoble et al. 2006). In addition, we have not accounted for the potential reputation losses involved in the implementation of a restrictive policy such as No Negative Blogging. But all this make our analysis conservative because if we were to consider both of these factors then it will only increase the region of Permissible Negative Blogging.

For future research, it will be interesting to see how the company specific characteristics affect our framework, but such analysis will require *page-views* data of employee blogs of different companies during the same time-period. The other interesting direction would be to study how writing different aspects of the content of a post attribute to the *impact of a post.* Such study would help firms to move closer to the optimal solution. For example, if the outcome of the study is that promoting a product of a firm's competitor has much higher influence on average than criticizing about a firm's product, and firm wants to encourage slight negative content on employee blogs, then it may ask its employees not to promote competitor's products but encourage them to discuss drawbacks of firm's products.

REFERENCES

Argyres, N.S. "The impact of information technology on coordination: Evidence from the B-2 "Stealth" bomber," *Organization Science* (10:2) 1999, p 162.

Ashmore, R.W., and Brian, M.H. "http ://managing the risk of employee blogging," *Risk Management* (53:4) 2006, p 40.

Atkinson, R.C., and Shiffrin, R.M. "Human memory: A proposed system and its control processes," *The psychology of learning and motivation* (2) 1968, pp 89-195.

Bates, D.M., and Watts, D.G. *Nonlinear Regression Analysis and Its Applications* Wiley, 1988, p. 384.

Battelle, J. *The Search: How Google and Its Rivals Rewrote the Rules of Business and Transformed Our Culture* Penguin Group, 2005, p. 320.

Brian, S. "Corporate Marketers Try Out Blogs," in: *Wall Street Journal*, 2005, p. B.8.

Brown, J.J., and Reingen, P.H. "Social Ties and Word-of-Mouth Referral Behavior," *Journal of Consumer Research* (14:3) 1987, p 350.

Byron, D., and Broback, S. *Publish and Prosper: Blogging for Your Business* New Riders Press, 2006, p. 208.

Conlin, M., and Andrew, P. "Blogging with The Boss's Blessing," *Business Week*:3889) 2004, p 100.

Davidson, R., and MacKinnon, J.G. *Estimation and Inference in Econometrics* Oxford University Press, 1993, p. 896.

Drezner, D.W., and Farrell, H. "The power and politics of blogs," American Political Science Association., 2004.

Edelmen, and Intelliseek "Talking From The Inside Out: The Rise of Employee Bloggers."

Francalanci, C., and Galal, H. "Information technology and worker composition: Determinants of productivity in the life insurance industry," *MIS Quarterly* (22:2) 1998, p 227.

Gattiker, T.F., and Goodhue, D.L. "What Happens after ERP Implementation: Understanding the Impact of Interdependence and Differentiation on Plant-Level Outcomes," *MIS Quarterly* (29:3) 2005, p 559.

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

Google.com "Our Search: Google Technology," 2004.

Halley, S. "HBR Case Study: A blogger in their midst," *Harvard Business Review* (81:9) 2003, p 30.

Jaccard, J., and Turrisi, R. *Interaction Effects in Multiple Regression*, (Second ed.) Saga Publications, 2003, p. 92.

Krapfel, R.E., Jr. "An Advocacy Behavior Model of Organizational Buyers' Vendor Choice," *Journal of Marketing* (49:4) 1985, p 51.

Lynch, J.G., Jr., and Srull, T.K. "Memory and Attentional Factors in Consumer Choice: Concepts and Research Methods," *Journal of Consumer Research* (9:1) 1982, p 18.

Mahajan, V., Muller, E., and Kerin, R.A. "Introduction Strategy for New Products with Positive and Negative Word-of-Mouth," *Management Science* (30:12) 1984, p 1389.

Massie, R., and Kurapati, K. "There's Blogs (and Feeds!) In Our Search!," Blog.ask.com, 2006.

Mayzlin, D. "Promotional Chat on the Internet," *Marketing Science* (25:2) 2006, p 155.

McKnight, D.H., Choudhury, V., and Kacmar, C. "Developing and validating trust measures for e-commerce: An integrative typology," *Information Systems Research* (13:3) 2002, p 334.

Michaelis, L., and Menten, M.L. "Die Kinetik der Invertinwirkung," in: *Biochem. Z.*, 1913, pp. 333-369.

Oliver, R. "Blogger in Chief," *Fortune* (154:10) 2006, p 51.

Pavlou, P.A., and Fygenson, M. "Understanding and Predicting Electronic Commerce Adoption: An Extension of the Theory of Planned Behavior," *MIS Quarterly* (30:1) 2006, p 115.

Schwartz, J. "If You Want to Lead, Blog," *Harvard Business Review* (83:11) 2005, p 30.

Scoble, R., and Israel, S. *Naked conversations: How Blogs are Changing the Way Businesses Talk with Customers* John Wiley, 2006.

Seber, G.A.F., and Wild, C.J. *Nonlinear Regression* Wiley-Interscience, 2003.

Shirky, C. "Power Laws, Weblogs, and Inequality," in: *Exposure: From Friction to Freedom*, Aula, 2003.

Stephens, T. "Myths and Facts about Fuel Economy," GM FastLane Blog, 2005.

Urban, G.L. "The Emerging Era of Customer Advocacy," *MIT Sloan Management Review* (45:2) 2004, p 77.

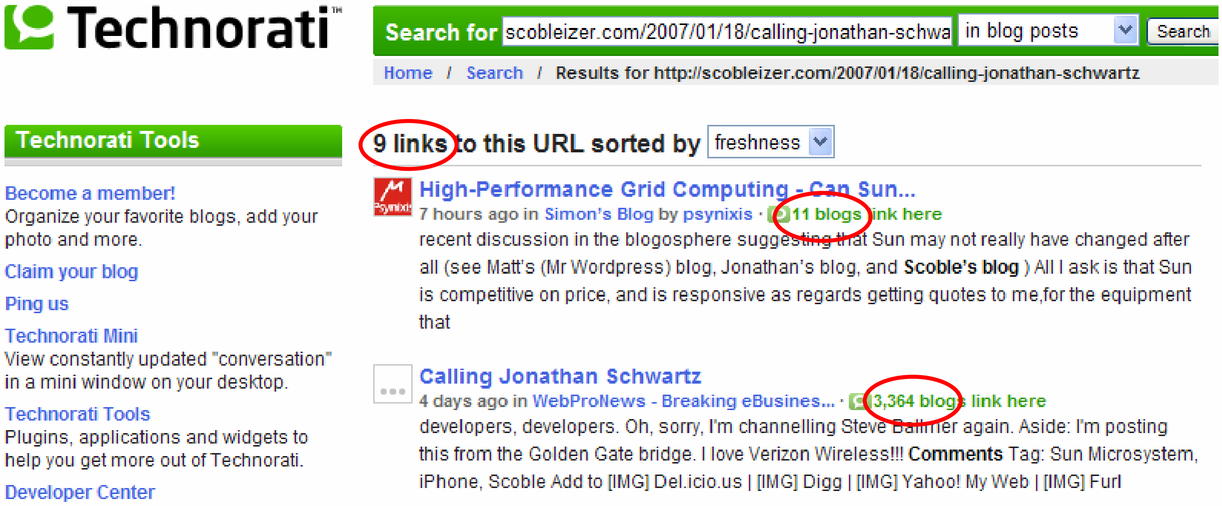
Urban, G.L. "Customer Advocacy: A New Era in Marketing," *Journal of Public Policy & Marketing* (24:1) 2005, p 155.

Urban, G.L., Sultan, F., and Qualls, W.J. "Placing Trust at the Center of Your Internet Strategy," *MIT Sloan Management Review* (42:1) 2000, p 39.

Wooldridge, J.M. *Econometric Analysis of Cross Section and Panel Data*, (1 ed.) The MIT Press, 2001, p. 740.

APPENDIX A: MEASURING THE IMPACT OF A POST

Consider the links to a post as shown below.



$Impact\ of\ a\ post,\ S \equiv \#of\ \ inbound\ \ links + \sum \ln(\#\ of\ \ 2^{nd}\ \ level\ \ inbound\ \ links)$

For example, *impact* of the above post, $S \equiv 9 + \ln(11) + \ln(3364) + \ldots$.

We have taken log of 2^{nd} level links because many studies show that inbound links to a blog follow skewed distribution (Drezner et al. 2004; Shirky 2003). Literature widely uses logarithmic transformation for such skewed distribution data (Wooldridge 2001).

Our results are robust to the choice of method used for measuring *impact of a post*. We have also done the analysis without considering 2^{nd} level links. Our results remain same with the only difference that the standard errors without considering 2^{nd} level links came out to be higher than the standard errors with 2^{nd} level links.

APPENDIX B: APPLICATION OF DURBIN-WU-HAUSMAN TEST BASED ON GAUSS-NEWTON REGRESSION

$$X(\beta) = \frac{\partial x_{i,t}(\beta)}{\partial \beta} = \frac{\partial \left(\beta_1 H_{i,t-1} + \beta_2 I_{i,t} + \beta_3 I_{i,t} \frac{R_{i,t}}{R_{i,t} + \beta_4} + \beta_6 R_{i,t} \right)}{\partial \beta}$$

$$= \begin{pmatrix} H_{i,t-1} & I_{i,t} & \dfrac{I_{i,t} R_{i,t}}{R_{i,t} + \beta_4} & \dfrac{-\beta_3 I_{i,t} R_{i,t}}{\left(R_{i,t} + \beta_4\right)^2} & R_{i,t} \end{pmatrix}$$

$$I_{i,t} \sim H_{i,t-1}, I_{i,t-2}, \frac{I_{i,t-2} R_{i,t}}{R_{i,t} + \hat{\beta}_4}, \frac{-\hat{\beta}_3 I_{i,t-2} R_{i,t}}{\left(R_{i,t} + \hat{\beta}_4\right)^2}, R_{i,t}$$

$$\frac{I_{i,t} R_{i,t}}{R_{i,t} + \hat{\beta}_4} \sim H_{i,t-1}, I_{i,t-2}, \frac{I_{i,t-2} R_{i,t}}{R_{i,t} + \hat{\beta}_4}, \frac{-\hat{\beta}_3 I_{i,t-2} R_{i,t}}{\left(R_{i,t} + \hat{\beta}_4\right)^2}, R_{i,t}$$

$$\frac{-\hat{\beta}_3 I_{i,t} R_{i,t}}{\left(R_{i,t} + \hat{\beta}_4\right)^2} \sim H_{i,t-1}, I_{i,t-2}, \frac{I_{i,t-2} R_{i,t}}{R_{i,t} + \hat{\beta}_4}, \frac{-\hat{\beta}_3 I_{i,t-2} R_{i,t}}{\left(R_{i,t} + \hat{\beta}_4\right)^2}, R_{i,t}$$

$$\hat{u}_{i,t} \sim X(\hat{\beta}), \hat{u}_{i,t}^{(1)}, \hat{u}_{i,t}^{(2)}, \hat{u}_{i,t}^{(3)}$$

(Davidson et al. 1993)

Null Hypothesis: Coefficients of errors are jointly insignificant

APPENDIX C: PROOF OF PROPOSITION

We show that the shape of WOM^{+}_{PNB} with respect to the *average impact of negative post* $\overline{I}^{-}$ is concave (this is the basis for Figure 4), by showing that the second derivative is negative:

$$\frac{\partial^2 WOM^{+}_{PNB}}{\partial \overline{I}^{-^2}} = -\frac{2\left(\beta_2\overline{I}^{-^3} + 3\beta_2\beta_4\overline{I}^{-^2}\overline{I}^{+} + 3\beta_2\beta_4^2\overline{I}^{-}\overline{I}^{+^2} + \beta_2\beta_4^3\overline{I}^{+^3} + \beta_3\overline{I}^{-^3} + 3\beta_3\beta_4\overline{I}^{-^2}\overline{I}^{+} + 3\beta_3\beta_4^2\overline{I}^{-}\overline{I}^{+^2} + \beta_3\beta_4^2\overline{I}^{0}\overline{I}^{+^2} + \beta_3\beta_4\overline{I}^{+^3} + \beta_3\beta_4\overline{I}^{+^2}\overline{I}^{0}\right)}{\left(\overline{I}^{-} + \beta_4\overline{I}^{+}\right)^3} \qquad \textit{(13)}$$

In the above expression note that all values in the numerator and denominator are positive, and the expression overall has a negative sign.

$WOM^{+}_{PNB} = WOM^{+}_{NNB}$ under two conditions: (1) when $\overline{I}^{-} = 0$, because when there are no negative posts, WOM^{+}_{PNB} and WOM^{+}_{NNB} are identical, and (2) $\overline{I}^{-} = \overline{I}^{-*}$, where $\overline{I}^{-*}$ solves $WOM^{+}_{PNB} = WOM^{+}_{NNB}$. This expression is quadratic in $\overline{I}^{-}$, and so, solving for $\overline{I}^{-*}$ yields (ignoring the negative root):

$$\overline{I}^{-*} = \frac{\sqrt{\beta_2^2\overline{I}^{0^2} + 2\beta_2\beta_3\overline{I}^{0^2} - 2\beta_2^2\beta_4\overline{I}^{0}\overline{I}^{+} + \beta_3^2\overline{I}^{0^2} - 2\beta_2\beta_3\beta_4\overline{I}^{0} + \beta_2^2\beta_4^2\overline{I}^{+^2} + 4\beta_2\beta_3\overline{I}^{+^2} + 4\beta_2\beta_3\overline{I}^{0}\overline{I}^{+} + 4\beta_3^2\overline{I}^{+^2} + 4\beta_3^2\overline{I}^{0}\overline{I}^{+}} - \left(\beta_2\overline{I}^{0} + \beta_3\overline{I}^{0} + \beta_2\beta_4\overline{I}^{+}\right)}{2\left(\beta_2 + \beta_3\right)}$$

We find that $\beta_3 > \beta_2\beta_4$ is a sufficient condition for $\overline{I}^{-*}$ to exist and be positive. From our empirical analysis, the data satisfies the condition $\beta_3 > \beta_2\beta_4$. Let

$$a = \sqrt{\beta_2^2\overline{I}^{0^2} + 2\beta_2\beta_3\overline{I}^{0^2} - 2\beta_2^2\beta_4\overline{I}^{0}\overline{I}^{+} + \beta_3^2\overline{I}^{0^2} - 2\beta_2\beta_3\beta_4\overline{I}^{0} + \beta_2^2\beta_4^2\overline{I}^{+^2} + 4\beta_2\beta_3\overline{I}^{+^2} + 4\beta_2\beta_3\overline{I}^{0}\overline{I}^{+} + 4\beta_3^2\overline{I}^{+^2} + 4\beta_3^2\overline{I}^{0}\overline{I}^{+}}$$

and $b = \beta_1\overline{I}^{0} + \beta_2\overline{I}^{0} + \beta_1\beta_3\overline{I}^{+}$.

When $\beta_3 > \beta_2\beta_4$, it can be seen that:

$$a =$$
$$\sqrt{\beta_2^2\overline{I}^{0^2} + 2\beta_2\beta_3\overline{I}^{0^2} - 2\beta_2^2\beta_4\overline{I}^0\overline{I}^+ + \beta_3^2\overline{I}^{0^2} - 2\beta_2\beta_3\beta_4\overline{I}^0 + \beta_2^2\beta_4^2\overline{I}^{+^2} + 4\beta_2\beta_3\overline{I}^{+^2} + 4\beta_2\beta_3\overline{I}^0\overline{I}^+ + 4\beta_3^2\overline{I}^{+^2} + 4\beta_3^2\overline{I}^0\overline{I}^+}$$
$$= \sqrt{\left(4\beta_2\beta_3 + 4\beta_3^2 + \beta_2^2\beta_4^2\right)\overline{I}^{+^2} + \left(-2\beta_2^2\beta_4\overline{I}^0 + 4\beta_2\beta_3\overline{I}^0 - 2\beta_2\beta_3\beta_4\overline{I}^0 + 4\beta_3^2\overline{I}^0\right)\overline{I}^+ + \beta_2^2\overline{I}^{0^2} + 2\beta_2\beta_3\overline{I}^{0^2} + \beta_3^2\overline{I}^{0^2}}$$
$$> \sqrt{\left(4\beta_2\beta_3 + 4\beta_3^2 + \beta_2^2\beta_4^2\right)\overline{I}^{+^2} + \left(2\beta_2^2\beta_4\overline{I}^0 + 2\beta_2\beta_3\beta_4\overline{I}^0\right)\overline{I}^+ + \beta_2^2\overline{I}^{0^2} + 2\beta_2\beta_3\overline{I}^{0^2} + \beta_3^2\overline{I}^{0^2}} > 0$$

This proves that a exists and is positive. Moreover, b is also positive. Now, we show that $a^2 - b^2$ is positive:

$$a^2 - b^2 = 4\left(\beta_2 + \beta_3\right)\overline{I}^+\left(\beta_3\overline{I}^+ + \left(\beta_3 - \beta_2\beta_4\right)\overline{I}^0\right) > 0$$

Algebraically, if $a^2 - b^2 > 0$ *and* $a + b > 0$, then $a - b > 0$. Hence, $\overline{I}^{-*} > 0$ (Proved).