# Impact and Diffusion of Sentiment in Political Communication – An Empirical Analysis of Political Weblogs

**Linh Dang-Xuan[1] and Stefan Stieglitz[2]**
Department of Information Systems
University of Münster
Leonardo-Campus 11, 48149 Münster, Germany
[1] linh.dang-xuan@uni-muenster.de, [2] stefan.stieglitz@uni-muenster.de

**Abstract**

In this paper, we examine whether sentiment of political blog entries is associated with increased feedback in terms of the quantity of triggered comments and whether the diffusion of sentiment might take place in the political blogosphere. Based on a data set of approximately 17,000 blog entries from the 60 most important German political blogs, we find that blog entries with either more positive or more negative overall sentiment tend to receive significantly more comments compared to sentiment-neutral or mixed-sentiment entries. Furthermore, our results show that positive as well as negative emotions might diffuse in the subsequent comments to corresponding blog entries.

## Introduction

In the past few years, social media have shown a rapid growth of user counts and have been object of scientific analysis (Wigand et al. 2010). Personal publishing modalities such as weblogs, microbloging and social network sites (SNS) have become prevalent (Kaplan and Haenlein 2010). The growing relevance of communication in social media implies a fundamental change in traditional public communication, which has usually been exclusively initiated and managed by specific actors, e.g., politicians, companies as well as journalists (Chadwick 2006). This phenomenon is currently observed by numerous disciplines such as sociology, information communication studies, information systems, political science, and linguistics. Among other fields of interest, it is a common goal to better understand modes of political communication such as political agenda setting or opinion making in social media.

Compared to other social media platforms, blogs have been longer established as a computer-mediated communication (CMC) channel. In particular, blogs have been significantly used in the context of political communication (Wattal et al. 2010, Munson and Resnick 2011). Political blogs offer a reciprocal relationship among their users with bloggers filtering information, proactively seek better information, grasp diverse views, evaluate opinions, and participate in discussions (Blood 2002). Given the controversial nature of politics, political blog postings are characterized by controversy and emotionality, and thus, often exhibit sentiment associated with certain political topics, political parties or politicians.

In this study, we seek to find out (a) whether political blog contributions containing affective dimensions in terms of sentiment would receive more feedback in terms of triggered comments, and (b) whether emotional states or sentiment might spread or diffuse in blog-based political discussions (i.e., in corresponding subsequent comments). To address these questions, we focus on discussions that take place on public political blogs assuming that blogs represent a platform where serious political discussions and interaction could take place.

## Theoretical Background and Hypotheses

Previous studies have dealt with the role of sentiment or emotions in CMC. They found that affective information could be transferred through CMC (Harris and Paradice 2007). Results from studies on social media, discussion forums, online news portals or other contexts indicated that affective dimensions of messages (both positive and negative emotions) could trigger more attention, cognitive involvement (e.g., Smith and Petty 1996), feedback (e.g., Huffaker 2010), and social transmission and sharing behavior (e.g., Berger and Milkman 2012, Stieglitz and Dang-Xuan 2012a, 2012b). Therefore, we derive the following hypothesis:

*H1: The more emotionally-charged (either positive or negative) a political blog entry is, the more comments it* *will trigger, i.e., there is a curvilinear U-shaped relationship between sentiment of blog entries and the quantity of triggered feedback.*

Beyond, studies have shown that not only behaviors and ideas but also emotions might spread through different kinds of social networks in various contexts (e.g., Bono and Ilies 2006, Hill et al. 2010), which is referred to as emotional contagion (Hatfield et al. 1994). In terms of CMC, for example, Huffaker (2010) showed that communication partners sync their wording, which would indicate that messages containing positive (negative) emotion words are likely to receive verbal responses, which also express positive (negative) emotions. Moreover, Huffaker (2010) provided evidence for the concept of language diffusion in online communities: the more often people used words that express affect, the more of the words they used were repeated in subsequent replies. Therefore, we derive the following hypothesis:

*H2: The more positive (negative) the overall sentiment of a political blog entry is, the more positive (negative) is the average overall sentiment of corresponding comments.*

## Data and Methodology

**Data:** For a six-month period from June 01 to December 31, 2011, we tracked all blog entries and corresponding comments from the 60 most important German political blogs according to the average of monthly "Wikio" blog rankings, which is one of the most popular blog ranking in Germany. Note that our sample period (of six months) also covers two *Landtag* (state parliament) elections in the states Berlin and Mecklenburg-Vorpommern in Germany. Both elections took place on September 18, and September 4, 2011, respectively. In total, 16,825 blog entries from 611 different bloggers were obtained including 55,862 corresponding comments.

**Sentiment Analysis:** We used the German version of the Linguistic Inquiry and Word Count (LIWC) software (Pennebaker et al. 2006, Tausczik and Pennebaker 2010) to objectively and systematically analyze blog entries and corresponding comments for emotional components using a psychometrically validated internal dictionary. LIWC is a text-analysis software program that places words from a text file into more than 80 categories based on a series of built-in dictionaries. LIWC has been widely used for academic purposes in psychology and linguistics but also for topics related to political science and communication studies (e.g., Yu et al. 2008, Huffaker 2010; for a comprehensive overview of related studies, see Tausczik and Pennebaker (2010)). In our study, we obtained the occurrences of words belonging to the LIWC categories "positive emotions" and "negative emotions" to profile sentiment in each blog posting. These categories have either been successfully used in previous studies of political text samples or seemed best suited to profile messages in the political domain by covering emotions. Since our sample consists of only German-language postings, we processed our data by using the LIWC German dictionary. The accuracy and robustness of LIWC analysis for German-language text samples have been positively assessed by other studies (e.g., Wolf et al. 2008). To take negations into account, we reverse the interpretation of a word if it is preceded by a negation so that positive words are counted as negative and vice versa. Following Zhang and Skiena (2010), for each blog entry or corresponding comment, we then adapt an overall sentiment measure which determines the direction of the sentiment as well as its strength:

(1) *sentiment* = (*positive* - *negative*) / (*positive* + *negative*)

The measure defined in equation (1), which is also called sentiment polarity (Zhang and Skiena 2010), takes into account the number of LIWC positive-emotion words *positive* as well as the number of LIWC negative-emotion words *negative*. It indicates the percentage of positive sentiment references among total sentiment references. If a blog entry or comment contains neither positive-emotion nor negative-emotion words, *sentiment* is defined as zero. As a consequence, the measure is defined within a continuous range from -1 to 1.

**Regression Analysis:** To test *H1*, we constructed the following variables for each blog entry:

- number of comments: *comm_no*
- overall sentiment (from -1 to 1): *entry_sentiment*

We included word count of a blog entry as a control variable. We also included an indicator for whether or not a URL was included in the blog entry. In addition, we controlled for posting activity of individual blogger as posting frequency might spark more dialogs and discussions (Huffaker 2010). Finally, we included the (average) Wikio rankings of blogs as a proxy for the different degrees of popularity or influence of individual blogs which might influence the quantity of feedback:

- word count (log): *wordcount*
- indicator (dummy variable) for whether or not a URL was included in the post: *url*
- total number of blog entries the blogger has contributed in the six-month period (log): *activity*
- rank of the blog (from 1 to 60, popularity increases in ascending order): *rank*

To test *H2*, we calculated the overall sentiment for each of the corresponding comments to each blog entry. Since each blog entry can trigger multiple comments, we calculated the average overall sentiment over all comments. Note that since not every blog entry has received a positive number of comments (i.e., no comments to be processed by

LIWC at all), the sample is reduced to 5,436 blog entries, each of which has triggered at least one comment:

- average overall sentiment of corresponding comments (from -1 to 1): *comm_avgsentiment*

In *H1*, we hypothesize that the more emotionally-charged (either positive or negative) a political blog entry is, the more comments it will trigger. As the dependent variable *comm_no* represents true-event count data, the Poisson regression model, at the first glance, is recommended for this particular data distribution (Cameron and Trivedi 1998). However, as the standard deviation (and hence the variance) of the dependent variable (*comm_no*) is larger than its mean (see Table 1), the analysis needs to be adjusted for overdispersion. Therefore, we applied the negative binomial regression model assuming the dependent variable to follow the negative binomial distribution (Cameron and Trivedi 1998). The resulting regression model is as follows:

$$(2)\ log(E(comm\_no|*)) = \beta_0 + \beta_1 entry\_sentiment + \beta_2 entry\_sentiment^2 + \beta_3 log(wordcount) + \beta_4 url + \beta_5 log(activity) + \beta_6 rank + \varepsilon,$$

where $E(comm\_no|*)$ is the conditional expectation of *comm_no*. We included the quadratic term $entry\_sentiment^2$ as *H1* would imply a curvilinear U-shaped relationship between the overall sentiment of the blog entry (*entry_sentiment*) which is defined within the range from -1 to 1 and the number of triggered comments (*comm_no*). Note that we log-transformed *wordcount* and *activity* to make our results depend less on outliers. In *H2*, we expect that the more positive (negative) the overall sentiment of a political blog entry is, the more positive (negative) is the average overall sentiment of corresponding comments. Since the dependent variable *comm_avgsentiment* represents average overall sentiment of corresponding comments and thus not true count data, we applied regression analysis using Ordinary Least Square (OLS) estimation to test *H2*. The regression model is as follows:

$$(3)\ log(comm\_avgsentiment) = \beta_0 + \beta_1 entry\_sentiment + \beta_2 log(wordcount) + \beta_3 url + \beta_4 log(activity) + \beta_5 rank + \varepsilon.$$

## Empirical Results

Descriptive statistics for each variable of interest are shown in Table 1. On average, a blog entry received about three comments. However, the standard deviation is high with 9.79 and almost two third of all blog entries did not get any feedback at all. The average overall sentiment or sentiment polarity of a blog entry is measured at 0.20 indicating a slightly more positive sentiment in blog entries. This also holds for corresponding comments ($M = 0.22$).

| **Variable** | **Mean** | **SD** |
|---|---|---|
| Dependent Variables | | |
| *comm_no* | 3.30 | 9.79 |
| *comm_avgsentiment* | 0.22 | 0.39 |
| Independent Variables | | |
| *entry_sentiment* | 0.20 | 0.54 |
| *wordcount* | 337 | 519 |
| *url* | 0.10 | 0.30 |
| *activity* | 28 | 96 |

*Table 1. Descriptive Statistics*

**Sentiment and Triggered Feedback:** In *H1*, we hypothesize a curvilinear U-shaped relationship between the overall sentiment of the blog entry and the number of triggered comments, i.e., the more emotionally-charged (either positive or negative) a political blog entry is, the more comments it will trigger. Results of the negative binomial regression (see Table 2) support our hypothesis as the coefficient of the quadratic term $entry\_sentiment^2$ is positive and highly statistically significant at one-percent level ($b = 0.21$, $p < 0.01$). In other words, clearly articulated sentiment (either positive or negative) in blog entries positively affects their likelihood to receive feedback. This result also implies that blog entries which are rather sentiment-neutral (i.e., neither positive nor negative references) or of mixed-sentiment nature (i.e., similar quantities of positive and negative references which would cancel out each other according to the sentiment polarity measure as defined in equation (1)) tend to trigger less feedback.

| **Independent Variables** | **Dependent Variable: *comm_no*** |
|---|---|
| | ***b (SE)*** |
| *entry_sentiment* | -0.01 (0.04) |
| $entry\_sentiment^2$ | 0.21*** (0.06) |
| *wordcount (log)* | 0.48*** (0.02) |
| *url* | -1.23*** (0.14) |
| *activity (log)* | -0.21** (0.02) |
| *rank* | 0.02*** (0.01) |
| Note that *b* denotes the estimated coefficients and *SE* the estimated robust standard errors in parentheses. Furthermore, *, ** and *** indicate significance level at 10%, 5% and 1%, respectively. The data set includes 16,825 observations (i.e., blog entries). | |

*Table 2. Negative Binomial Regression Output*

**Sentiment Diffusion:** Results of the OLS regression predicting the average overall sentiment of comments to blog entries are shown in Table 3. *H2* predicts that he more positive (negative) the overall sentiment of a political blog entry is, the more positive (negative) is the average overall sentiment of corresponding comments. We find a significant positive relationship between the overall sentiment of blog entries (*entry_sentiment*) and the average overall sentiment of corresponding comments (*comm_avgsentiment*), which supports *H2*. The coefficient of *entry_sentiment* is positive and highly statistically significant at one-percent level ($b$ = 0.24, $p$ < 0.01). This result suggests that positive as well as negative emotions might diffuse in the following discussion in the way that blog entries containing positive (negative) emotion words are likely to trigger verbal responses, which also express positive (negative) emotions.

| **Independent Variables** | **Dependent Variable: *comm_avgsentiment*** |
| --- | --- |
| | ***b (SE)*** |
| *entry_sentiment* | 0.24*** (0.03) |
| *wordcount (log)* | -0.06*** (0.01) |
| *url* | 0.13* (0.07) |
| *activity (log)* | -0.02* (0.01) |
| *rank* | -0.01*** (0.00) |
| $R^2$ = 0.13 | |
| Note that *b* denotes the estimated coefficients and *SE* the estimated robust standard errors in parentheses. Furthermore, *, ** and *** indicate significance level at 10%, 5% and 1%, respectively. The data set includes 5,436 observations (i.e., blog entries which have received at least one comment). | |

*Table 3. OLS Regression Output*

## Conclusion

The main contribution of our paper is to examine the role of sentiment in discussions in the political blogosphere. We find that clearly articulated sentiment (either positive or negative) in blog entries positively affects their likelihood to receive feedback. This finding suggests that, in the political blogosphere, people might tend to participate more in emotionally-charged (either positive or negative) discussions. Furthermore, our results indicate that positive as well as negative emotions might diffuse in the following discussion. This way, not only information but also sentiment in political context could be disseminated, which, to a certain extent, might influence political opinion-making processes. Therefore, our work implies that it is important for politicians and political parties to identify influential users and follow the discussions including sentiment occurring within their peers. For that, political parties and politicians might follow the approach of social media monitoring, which has more widely been used in the corporate context. As a limitation, our study relies on a data sample, which is restricted to German political blogs. As future work, we aim at extending our study to a larger scale (e.g., longer time periods of data collection, other countries and languages) and more general contexts, i.e., we will not limit our investigation only to political communication.

## References


Berger, J., and Milkman, K. 2012. What Makes Online Content Viral?. *Journal of Marketing Research*, forthcoming.

Blood, R. 2002. *The weblog handbook: practical advice on creating and maintaining your blog*. Cambridge, MA: Perseus.

Bono, J. E. and Ilies, R. 2006 Charisma, positive emotions and mood contagion. Leadership Quarterly, 17: 317-334.

Cameron, A., and Trivedi, P. 1998. *Regression analysis of count data*. Cambridge University Press.

Chadwick, A. 2006. *Internet Politics: States, Citizens, and New Communications Technologies*. NY: Oxford University Press.

Harris, R. B., and Paradice, D. 2007. An investigation of the computer-mediated communication of emotion. *Journal of Applied Sciences Research* 3: 2081-2090.

Hatfield, E., Cacioppo, J., and Rapson, R. L. 1994. *Emotional contagion*. NY: Cambridge University Press.

Hill, A. L., Rand, D. G., Nowak, A., and Christakis, N. A. 2010. Emotions as infectious diseases in a large social network: the SISa model. In *Proceedings of the Royal Society B*.

Huffaker, D. 2010. Dimensions of Leadership and Social Influence in Online Communities. *Human Communication Research* 36(4): 593-617.

Kaplan, A. M., and Haenlein, M. 2010. Users of the world, unite! The challenges and opportunities of Social Media. *Business Horizons* 53(1): 59-68.

Munson, S., and Resnick, P. 2011. The Prevalence of Political Discourse in Non-Political Blogs. In *Proceedings of ICWSM-11*.

Pennebaker, J. W., Booth, R. J., and Francis, M. E. 2006. *Linguistic inquiry and word count: LIWC*. Austin, TX: Erlbaum.

Smith, S. M., and Petty, R. E. 1996. Message framing and persuasion: A message processing analysis. *Personality and Social Psychology Bulletin* 22(3): 257-268.

Stieglitz, S., and Dang-Xuan, L. 2012a. Political Communication and Influence through Microblogging - An Empirical Analysis of Sentiment in Twitter Messages and Retweet Behavior. In *Proceedings of HICSS-45*, 3500-3509.

Stieglitz, S., and Dang-Xuan, L. 2012b. Impact and Diffusion of Sentiment in Political Communication - An Empirical Analysis of Public Political Facebook Pages. In *Proceedings of ECIS 2012*.

Tausczik, Y. R., and Pennebaker, J. W. 2010. The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology* 29: 24-54.

Wattal, S., Schuff, D., Mandviwalla, M., and Williams, C. 2010. Web 2.0 and Politics: The 2008 U.S. Presidential Election and an E-Politics research agenda, *MIS Quarterly* 34(4): 669-688.

Wigand, R. T., Wood, J. D., and Mande, D. M. 2010. Taming the Social Network Jungle: From Web 2.0 to Social Media. In *Proceedings of AMCIS 2010*, Paper 416.

Wolf, M., Horn, A., Mehl, M., Haug, S., Pennebaker, J. W., and Kordy, H. 2008. Computergestützte quantitative Textanalyse: Äquivalenz und Robustheit der deutschen Version des Linguistic Inquiry and Word Count. *Diagnostica* 2: 85-98.

Yu, B., Kaufmann, S., and Diermeier, D. 2008. Exploring the characteristics of opinion expressions for political opinion classification. In *Proceedings of dg.o 2008*, 82-91.

Zhang, W., and Skiena, S. 2010. Trading Strategies To Exploit Blog and News Sentiment. In *Proceedings of ICWSM-10*.