Effects of Feedback and Peer Pressure on Contributions to Enterprise Social Media

Michael J. Brzozowski, Thomas Sandholm, and Tad Hogg
Hewlett-Packard Laboratories
1501 Page Mill Rd, MS 1139, Palo Alto, CA, USA
{mike.brzozowski, thomas.e.sandholm, tad.hogg}@hp.com

ABSTRACT

Increasingly, large organizations are experimenting with internal social media (*e.g.,* blogs, forums) as a platform for widespread distributed collaboration. Contributions to their counterparts outside the organization's firewall are driven by attention from strangers, in addition to sharing among friends. However, employees in a workplace under time pressures may be reluctant to participate–and the audience for their contributions is comparatively smaller. Participation rates also vary widely from group to group. So what influences people to contribute in this environment?

In this paper, we present the results of a year-long empirical study of internal social media participation at a large technology company, and analyze the impact attention, feedback, and managers' and coworkers' participation have on employees' behavior. We find feedback in the form of posted comments is highly correlated with a user's subsequent participation. Recent manager and coworker activity relate to users initiating or resuming participation in social media. These findings extend, to an aggregate level, the results from prior interviews about blogging at the company and offer design and policy implications for organizations seeking to encourage social media adoption.

Categories and Subject Descriptors

H.5.3 [**Information interfaces and presentation**]: Group and organization interfaces; K.4.3 [**Computers and society**]: Organizational impacts

General Terms

Human Factors

Keywords

social media, contributions, attention, blogs, feedback

1. INTRODUCTION

In recent years, social media, such as forums, blogs, microblogs (*e.g.,* Twitter), and bookmarking services, have lowered the barriers to self-publishing on the Web. The prospect of receiving widespread attention motivates contributions to sites like YouTube [18], while attention from particular friends drives posts to Twitter [19]. The abundance of information and opinions freely available online makes readers' attention a scarce resource [30]. It has been said that an *attention economy* drives the Web [12], with myriad contributors competing for readers' attention.

Recently, organizations and researchers have begun experimenting with the use of internal social media in the workplace, hoping to reap the benefits of lightweight informal collaboration among employees. Unlike email, which must be targeted to specific recipients or distribution lists, social media provide a free broadcast platform, allowing authors to circumvent traditional organizational hierarchies and connect with geographically or organizationally distant readers.

Internal blogs can facilitate collaboration and knowledge sharing in an enterprise [20]. But large companies often create disincentives for employees to share knowledge [17]. In an environment where time is money, sharing insights for others' benefit may not be perceived as a good use of one's time. This disparity between benefit and effort required is a common impediment to groupware adoption [16]. Moreover, the attention economy, while effective on the Web, may break down in office settings [37]. Among the possible reasons cited are relatively obscure metrics for attention (while YouTube provides real-time view counts, not all blog or forum servers do) and a lack of management support.

So what leads people to contribute to these media? Previous semi-structured interviews at HP found visible feedback and "management buy-in" to be the top concerns among bloggers [37], which motivate the two hypotheses we set out to evaluate with empirical evidence:

H1 Visible feedback encourages employees to continue contributing to social media.

H2 Visible activity from managers and coworkers motivates employees' contributions to social media.

1.1 Approach

We explored various forms of feedback and reinforcement and the effect they have on observed behavior within a corporate environment using over a year of data on contributions to internal social media at a large global enterprise. We built a tool to monitor employees' contributions across the venues described in Table 1. We cross-referenced them with daily snapshots of the employee directory, giving us information about where authors work, what they do, and who they report to.

GROUP'09, May 10–13, 2009, Sanibel Island, Florida, USA.

Venue	Code	Authors	Posts	Months
Blogs	B	1462	18299	34
Blog comments	C	1692	7395	15
Discussion forums	F	14625	117807	28
Ideas[1]	I	420	926	29
Links[2]	L	159	2889	19
People tags	T	3079	3943	13
Tech reports	R	983	1546	23
Wiki pages	W	525	2456	17

Table 1: Participation observed in various social media venues at HP during the study period.

1.2 Contribution

In this paper, we present two complementary empirical analyses of the dataset described in Section 3. Section 4 discusses a longitudinal time series analysis used to determine how various forms of feedback affect users' future contributions. Section 5 uses an "activity event" model to look at possible influences on users when they started, continued, or stopped participating in each venue.

This paper's contribution is a quantitative analysis relating organizational and social motivators to user behavior across a variety of social media venues, enabling a comparison of various factors influencing user participation in the workplace.

Understanding the relative importance of these factors is valuable in designing an organization's internal social media strategy, and fostering a healthy ecosystem of contributors. With enough participation, blogs can be an effective tool to expose tacit knowledge and support collaboration [20].

2. RELATED WORK

2.1 Online communities

The question of what motivates people to contribute to online communities has inspired a variety of research. Social benefits derived from a sense of community can be powerful motivators for people to participate in a discussion group [8]. Experienced members of a community learn to practice the local "netiquette"– the standards of good behavior that build up members' reputation, trust, and social capital. Markers of reputation are the bedrock of how peer-to-peer commercial transactions function on sites like eBay [29].

The promise of a response is also a potential motivator for contributions. Yet a quarter of all Usenet posts never receive a response [2]. The politeness or rudeness of a message can help or hinder its chances of getting a response, depending on the norms of a forum [7]. Some communities, such as Slashdot, have evolved sophisticated moderation and feedback mechanisms to reward commenters for valued contributions and discourage undesirable behavior [25].

People are more likely to contribute to online communities if they feel their contributions are unique, either because they're explicitly told so [3] or because their opinions differ from the majority [35].

Wang and Fesenmaier [32] investigated factors motivating contribution in an online travel community and found that efficacy (*e.g.* being helpful to others, sharing enjoyment) had the strongest impact followed by relationship building and expectancy (future reciprocation). Status only had a minor effect. They also stressed in their conclusions that group boundaries are important to avoid free-riders. These results were, however, not obtained in an enterprise setting and their main method consisted of correlating participants' expressed motivations to their expressed contributions without any temporal distinction.

2.2 Social media in the enterprise

Lakhani and von Hippel suggest that contributors to the help groups for the Apache web server project may derive intrinsic benefits just from reading the stream of questions and answers posted by other members [24]. These forums can function as collectors for tacit knowledge, much the same way internal corporate blogs do [20]. Benkler even suggests that this "commons-based peer-production" model of collaboration is more efficient at mapping talent to questions and tasks than traditional organizational hierarchies [4].

A fair amount of literature has examined social and organizational barriers to knowledge sharing and groupware adoption. Imbalance between who contributes and who benefits, leading to "tragedy of the commons" issues, can pose significant challenges [16]. Pay-for-performance incentives pit employees against each other, discouraging them from spending time on tasks that don't directly impact their evaluations [17]. Finally, people who are not collocated are less likely to voluntarily collaborate [23].

However, other studies suggest motivations for employees to adopt internal social media tools. Employee blogging can be a way for people to engage with their organization [14] and maintain a sense of community with their colleagues [21]. A comprehensive study by DiMicco *et al*. found workers at IBM use internal social networking tools to connect with coworkers personally and to advance their career and project objectives [13]. Temporal analysis of blog consumption suggests blogging may serve these dual social and professional purposes [36].

A variety of studies have examined the use of and reaction to particular social media tools at IBM, including social bookmarking [27], people tagging [15], and internal blogging [20]. Kolari *et al*. modeled commenting behavior on blogs at IBM as a graph and projected these "conversations" onto the organization chart to measure the "reach" of blog posts [22].

2.3 Evaluating social influence

A key challenge when identifying motivation for contributions to social media is disentangling endogenous from exogenous effects. More generally, the issue of distinguishing social influence from selection (behavior triggered by influence from peers as opposed to personal preferences) has been a great concern in the social network literature [26, 1, 11, 31, 6]. Manski [26] first formulated the problem and pointed out difficulties that could only be overcome with prior knowledge of the reference groups that were studied. We tackle this problem by connecting our quantitative results to qualitative results from interviews.

More recently, Anagnostopoulos *et al*. [1] suggested and experimentally evaluated influence in social networks via a couple of statistical tests, one based on shuffling and bootstrap techniques and one based on following directed graphs in the opposite direction. Our statistical tests, on the other hand, are based on regression analysis and correlation patterns to highlight temporal effects. Another recent study by Crandall *et al*. [11] investigated the set of authored pages versus edit page discussions on Wikipedia pages as proxies for exogenous and endogenous processes respectively. They found an elaborate interplay between these factors making it

[1] Services explicitly used to propose and discuss new business ideas.

[2] Lightweight services used to share bookmarks; these services are similar to delicious.com and digg.com.

Country	Rate	Business group	Rate	Job function	Rate
United Kingdom	10.0%	Group E	27.2%	Marketing	17.0%
Germany	7.9%	IT	9.4%	Engineering	13.6%
United States	5.8%	Group B	8.3%	Sales	7.6%
Singapore	2.8%	Shared Services	4.1%	Operations	2.2%
Mexico	1.9%	Group D	2.7%	Administration	2.1%
Japan	1.9%	Group A	2.2%	Finance	2.1%

Table 2: Social media participation in the most over- and under-represented countries, business groups, and job functions.

hard to attribute an observation to a single type of source. Although clean separation is not the primary goal of our study we partially address the concerns by relating time series to externally known events and by studying time series for both aggregate, *macro*, and the individual, *micro*, behaviors.

3. SOCIAL MEDIA AT HP

We studied a large global technology enterprise, Hewlett-Packard (HP). HP has a variety of social media services available to all employees, used for internal collaboration and communication. We polled all posts made to them between February 2006 and December 2008, over periods ranging from 13 to 34 months (depending on when we discovered the services). We categorized them into *venues* according to the type of content shared and effort required to post (see Table 1). A few high-profile people have received approval for externally-facing blogs, but the vast majority of this content is only accessible inside the firewall and so not eligible for attention from outside the company.

As shown in Table 2, participation rates vary widely by country, business group, and job function. While employees speak a variety of languages, the vast majority of the content we found was written in English. It is possible that users not fluent in English are using other collaborative venues we were not able to track, which may partially explain the skewed distribution, but we lack sufficient information about users' language preferences to make informed speculations.

It may be that certain professions, like engineering or marketing, naturally lend themselves more to open collaboration than the comparatively secretive practices of corporate finance and administration. But the wide variation in participation by group may also suggest social or organizational influences. Perhaps participation is "contagious" in organizations or professional communities. Or perhaps managers' behavior exerts pressure on their subordinates, setting an example by their attitude towards social media venues, or by explicitly encouraging or discouraging contributions.

4. TIME SERIES ANALYSIS

As a first step towards understanding what factors impact individuals to contribute we study the timing of contributions versus various *impact factors* that indicate a user's posts are being read, for example clicks or comments on previous contributions. The primary goal of this analysis is to compare impact factors and thereby give guidance to future improvements in social software design. The secondary goal is to establish metrics that can be used in intervention analyses and benchmarks against external systems as well as different interfaces to social media, or before and after major organizational events. Next we outline the method used to evaluate traces of user activity including choice of metrics.

4.1 Method

Our general approach is to apply well-proven techniques from regression and time series analysis to study correlations of different factors over time. The primary goal is not to compose a complete model with all possible explanatory factors to use in actual predictions, but rather to highlight structural differences and patterns in the series that might help us understand which impact factors have more and which have less of an effect on the contribution factor by quantifying and comparing correlation metrics between factors (and systems).

We are particularly interested in comparing hidden versus visible impact factors. In the social media environment we studied, the total readership (or "hit count") of a post is not exposed to content authors, making it a "hidden" factor that nevertheless influences both exogenous feedback (*e.g.*, through off-line channels) and visible feedback. Since users can easily see how many comments they have received and who they are from, we consider them a visible impact factor. It is easy to make different factors visible to users, but if a factor has no impact on the contribution it will just clutter the user interface and obfuscate more effective impact factors. In this paper we compare measurements for hidden clicks on documents authored and published with visible comments on the same documents. Additionally, we are interested in the diversity of the different factors, *i.e.* where the clicks and comments come from. Since our system allows users to authenticate, and we have access to the employee database, we can track various employee attributes of both the person submitting and the person receiving a click or a comment, such as their unique employee ID, location (city and country) and organizational unit. These attributes are hard or impossible to track when only the IP address of users are known, which is a common limitation in public web trace analyses. We could have extracted more attributes, but for the time series analysis presented here the sparseness of the data makes more detailed studies unreliable.

We analyze impact factors both on a *micro* and a *macro* level. The micro analysis studies time series of individual users, and collects summary statistics of fits of different models that are then aggregated on a system level. The macro analysis aggregates the impact and contribution factors for all users into a global time series, and then fits a set of models to this series. The reason we study both is that predicting an individual's future contributions based on her past behavior is different from predicting the sum of all contributions based on aggregated data. Aggregation could mask signals (or heterogeneity in signals), as well as highlight signals not directly affecting individuals.

The factors we measure are summarized in Table 3. A user (employee) belongs to exactly one *organization unit* in a single city and country and may click on a document at most once. However, in a few cases a user's organization is missing in our data. There are 90 organization units in HP, with a median of 761 and an average of 4090 employees in each.

4.1.1 *Time Series*

We split up and collected the data for weekly time periods (seven days) for all factors over a period of 55 weeks. Posts at HP fol-

factor	type	description
AuthoredDocs	Contribution	number of documents published to all venues
TotalClicks(Comments)	Impact	number of clicks or comments on authored documents
DistinctEmployeeClicks(Comments)	Impact Diversity	number of employees commenting or clicking on authored documents
DistinctDocClicks(Comments)	Impact Diversity	number of authored documents receiving clicks or comments
DistinctCityClicks(Comments)	Impact Diversity	number of cities clicks or comments originate from
DistinctCountryClicks(Comments)	Impact Diversity	number of countries clicks or comments originate from
DistinctOrgUnitClicks(Comments)	Impact Diversity	number of org units clicks or comments originate from

Table 3: Time series analysis factors.

low definite seven-day cycles, with significantly less activity over the weekend; as a result, shorter time periods would be subject to significantly more noise depending on whether they stretch over a weekend. Other multiples of a week could be used, but 85% of all comments and 69% of all clicks occur within one week of the original post, so this interval likely captures most potential feedback a user might receive. However, we found that within-workweek predictions were much more accurate than cross-weekend predictions, so which day was used to demarcate the time series periods affected our macro (but not micro) results. After studying all possibilities, we decided to start new seven-day long periods on Wednesdays to offset the weekend anomaly and to capture the strong within-workweek correlations.

About 130k documents from all venues in the system were authored and published during this time, out of which 61k were comments on previous documents. We tracked 50k clicks from authenticated users. The contribution distribution across users has a long tail, which has also been observed in a wide variety of other on-line communities [34]. Therefore, if we used all the data, the mostly inactive users would dominate the results and would likely give misleading design implications.

Our micro analysis is particularly sensitive to these heavy tails since series of mostly empty values would destroy correlations from weak signals. To circumvent this problem we establish a user contribution threshold of an average of one contribution every other week during the analyzed period. Only users with more contributions are considered. Furthermore the micro analysis is only done for each user from the point where his or her contributions started till the time contribution stopped. If this period is less than a month, the user will only be considered in the macro analysis. These filters led to a study of 295 users (about 16% of all users receiving clicks from authenticated users) in the micro analysis and 931 users (about 50% of all click receivers) in the macro analysis. We also note that only about 10% of all users received clicks on their authored documents from authenticated users, which was the limiting factor of the scope of this analysis, and which is a direct effect of the long tail of contributions.

4.1.2 Metrics

We employ three sets of metrics on all impact and contribution factors: *correlation structure*, *contribution correlation* and *contribution predictability*.

The **correlation structure** is represented by the partial autocorrelation function (PACF) [33] for lag k defined as

$$PACF(k) = Corr(Z_t, Z_{t+k}|Z_{t+1}, ..., Z_{t+k-1}) \quad (1)$$

where Z_t is the datum observed at time t, and $Corr$ is the correlation defined as $Cov(Z_t, Z_{t+k})/\sigma$ where Cov is the covariance and σ the standard deviation. The conditional terms in Eq. 1 distinguish the PACF from the autocorrelation function (ACF). The PACF is useful in that it gives the correlation between two data points with mutual linear dependencies of intervening data points removed. It can also be used as a direct indicator of the number of lags to include in autoregression models [33]. Furthermore, it is convenient in our analysis because the average value across a large number of tests (one for each user) makes intuitive sense.

Standard time series models assume stationarity in variance, *i.e.* variance is assumed not to change over time. From observations of the time series we saw that the variance was proportional to the level, *e.g.* scatter plots are not fully linear but *drop off* for high values. The typical Box-Cox stationarity transformation [5] in this case is to take the square root of the impact factors, which also worked well for us.

The **contribution correlation** metric is defined in terms of a linear regression of an impact factor to the contribution factor (both factors are sampled in the same time interval). The null hypothesis is that the coefficient ϕ_1 is 0 in the fitted model

$$C_t = \phi_0 + \phi_1\sqrt{I_t} + a_t \quad (2)$$

where C_t is the contribution factor at time t, I_t is the impact factor at time t, ϕ_0 is the intercept of the regression (level), and a_t is a stochastic white noise process. To quantify impact we track the R^2 value [28] representing how much of the variance of the contribution factor can be explained by the variance in the impact factor. This yields a value between 0 and 1, where 1 means that all of the variance can be explained by the impact factor. It is defined as

$$R^2 = 1 - \frac{\sum_{t=0}^{T}(Z_t - r_t)^2}{\sum_{t=0}^{T}(Z_t - \phi_0)^2} \quad (3)$$

where Z_t is the observed series datum, r_t is the modelled datum, and T is the total number of periods modelled. Similar to the PACF metric the R^2 metric has an intuitive aggregate interpretation across users, and more importantly it is designed to compare model fits in an unbiased way which is at the core of our method.

The **contribution predictability** metric is also defined in terms of a linear regression. The null hypothesis is that the coefficient ϕ_1, ϕ_2, ϕ_3 are all 0 in the fitted model:

$$C_t = \phi_0 + \phi_1\sqrt{I_{t-1}} + \phi_2\sqrt{I_{t-2}} + \phi_3\sqrt{I_{t-3}} + a_t \quad (4)$$

we again collect the R^2 statistic. This value is referred to as the *predictive power* here and it is later used to order the metrics by level of impact. For this metric we also measure whether the impact is significantly positive or negative by using the p-value of the null hypothesis that individual coefficients ϕ are 0 and applying the t-statistic (ϕ/σ). As an aggregate measure we define:

$$sgn_p = \frac{\phi_k}{|\phi_k|} \quad (5)$$

where ϕ_k is the coefficient of first lag with a significant t-statistic at the 5% significance level. If there are no significant lags sgn_p is 0. Intuitively the *contribution predictability* metric denotes the impact

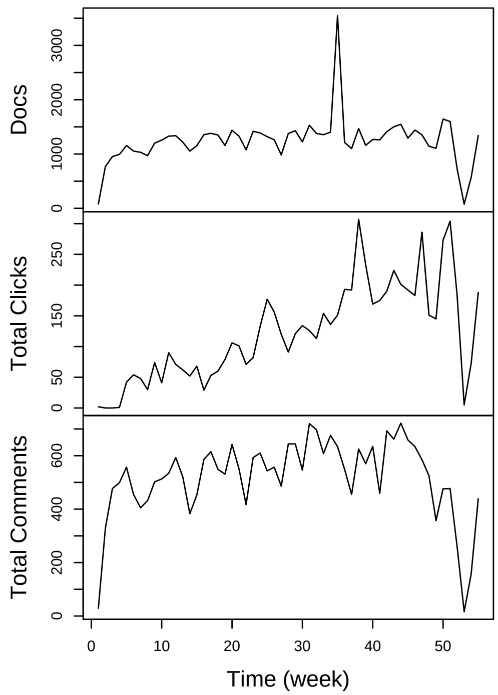

Figure 1: Contribution and impact factor time series.

that the history, up to three weeks back, of clicks and comments has on the number of documents published in the current week.

We are not so concerned with causality versus correlation here since we test the predictability of the contribution factor as well. Our main concern is to quantify and compare which predictor is better. Furthermore, the time lags help disambiguate the direction of the impact.

4.2 Results

4.2.1 Impact Visibility Analysis

We start the analysis by just studying the *AuthoredDocs*, *TotalClicks* and *TotalComments* factors. Figure 1 shows the time series of the contribution and impact factors aggregated in weekly periods. Three features stand out. First, the clicks series shows a level shift starting in week 15 which coincides with an internal technology conference that advertised our system. One could argue that this shift should be suppressed with differentiation, but we found that doing so would in fact destroy useful information such as correlations. Since differentiation can be compared to ϕ_1 being close to 1, it is also accommodated for in our model. Another reason to not differentiate is that it may help some users' model fits but destroy others and thus create an unwanted bias in our analysis.

Second, the document spike at week 35 coincides with a major event (acquisition) for HP that was heavily discussed in the blogosphere. This exogenous event also resulted in an unusually large number of new users appearing in many of the venues we tracked, who had no prior history we could use to make predictions. We found it clarifying to remove this spike, since it was not caused by any previous participation factors but it impacted the contribution factor without prior, current or future changes in any of the impact factors, causing anomalies in some of the documents and comments statistics.

Finally, the drop at week 52 corresponds to the Christmas holiday closure. Because the Christmas break affected all metrics the same way we chose not to suppress the results from that week artificially.

Figure 2 shows the scatter plots with smoothed regression lines for the regression factor against the regression response (note that the week 35 contribution anomaly has been suppressed). The regression factors are lagged by one week to represent a predictive regression setup. We can see that the smoothing algorithm (Gaussian least squares) failed to draw a regression line for the clicks series which is a first hint that this factor has less impact. Next we will quantify that visual cue.

The top part of Table 4 shows the micro analysis metrics measured using individual user regressions. The R^2_{cp} or *predictive power* metric is represented by the mean value followed by the first and third quartile in parentheses; all other metrics just show the mean to conserve space. The heavily skewed contribution distribution made the 5% significance bounds very wide, which is why we show the quartile bounds in the table. Looking at the micro metrics, the correlation structure has at most one week of memory for all the factors: only PACF(1) shows significant correlation. Thus, for individual users, only clicks and comments received very recently on authored documents seem to affect contributions. The contribution correlation is substantially higher for the *TotalComments* factor compared to the *TotalClicks* factor. That is, the same period correlation between comments and documents is higher than the same period correlation between clicks and documents.

Finally, the contribution prediction is slightly better for *TotalComments* compared to *TotalClicks*, but the *AuthoredDocs* contribution factor is the best predictor (highest R^2_{cp}). So for individual users, past authored documents matter more than feedback in terms of comments and readership such as clicks, when determining future contributions. Here we also see the first quantitative evidence for comments (visible impact) being more effective than clicks (hidden impact), which we saw qualitatively in Figure 2.

The bottom part of Table 4 shows the macro analysis results. The interpretation of the macro results is somewhat different in that it measures the predictability of the global system contribution given the global impact factor. We see substantially more memory in the time series process correlation structure, in particular for the *AuthoredDocs* and *TotalClicks* metrics. This can be seen in the aggregate (macro) PACF metrics in Table 4 where the document publication two weeks back and the click traffic three weeks back have a non-negligible correlation to the current values. This increase in memory on the macro scale may be attributed to network effects not captured in the egocentric micro analysis. The contribution correlation shows similar patterns compared to the individual micro metrics but the contribution prediction metrics show a more substantial differentiation. *TotalComments* predicts better than *TotalClicks*, but again the *AuthoredDocs* metric is the best predictor by far.

The main result from the visibility analysis is that comments have a greater effect than clicks when determining future document contribution, which was confirmed both on a micro and on a macro scale. This result gives support to the first hypothesis, **H1**, posed in Section 1.

We also note that our system displays the most popular stories on the front page. So for the most influential users, clicks are in some sense at least partially visible. That might explain the fact that the comment and clicks metrics tend to reverse for the top contributing users (not shown here). However, we consider clicks hidden because there is no information on where the clicks originated from, which will play a role in the diversity analysis below, and because only a very limited set of users benefit from the front page popularity exposure.

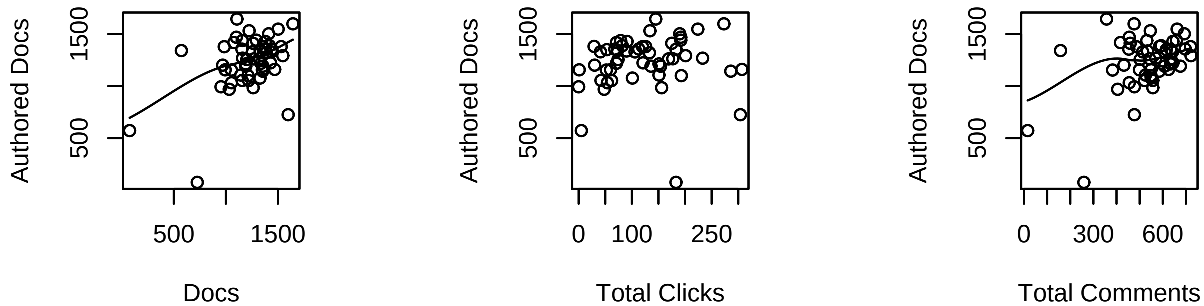

Figure 2: Contribution and impact factor scatter plot with one-week predictive regression lines.

Micro (local) analysis				
	contribution predictability		contribution correlation	correlation structure
	R^2_{cp}	sgn_p	R^2_{cc}	PACF(1,2,3)
AuthoredDocs	$0.14(0.01, 0.39)$	0.30	1.00	$(0.28, 0.06, 0.06)$
TotalComments	$0.12(0.01, 0.38)$	0.21	0.30	$(0.13, 0.01, 0.03)$
TotalClicks	$0.10(0.01, 0.37)$	0.18	0.13	$(0.07, -0.00, 0.02)$
Macro (global) analysis				
AuthoredDocs	0.38	1.00	1.00	$(0.45, -0.21, 0.08)$
TotalComments	0.32	1.00	0.76	$(0.58, -0.07, 0.10)$
TotalClicks	0.16	1.00	0.29	$(0.74, -0.08, 0.36)$

Table 4: Results for visibility analysis ordered by descending predictive power.

4.2.2 Impact Diversity Analysis

We now drill deeper into the diversity of the impact by studying attributes of users who click and comment on documents. By *impact diversity* we mean, for instance, how many different cities or countries the comments originated from and how well that predicts the future contribution of the user receiving the comments. For this analysis we restrict the presentation to the macro metrics because of space considerations and because the differentiation was clearer between the attribute metrics on the macro scale. Table 5 shows the results.

Based on the contribution predictability results, the number of documents that comments are made on is more effective (gives stronger signal of the level of contribution the following week) than the number of documents that have been clicked on (the worst predictor). One could argue that this may partly be explained by there being a higher effort involved in commenting on a new document than adding more comments, something that does not hold true for clicks. Comparing geographic metrics with organizational ones we can see that the country and city metrics are slightly better than the corresponding organization unit metrics both for clicks and comments. So geographic diversity tends to be more important than organizational diversity to motivate contribution. We also note again that the impact metrics for comments are consistently higher than those for clicks. The fact that the macro analysis agrees with the micro analysis in this regard is a sign of stability in the result that hypothesis **H1** is (quantitatively) supported by the data.

4.3 Summary

The key findings in the time series analysis as to what motivated user contribution were: comments are more effective than clicks (**H1**); diversity does matter, in particular geographic diversity; and previous contributions play a greater role for individual users than feedback and readership factors. Promoting attention from colleagues across geographic barriers, would thus seem to be the most effective way of nurturing and motivating contributions in the enterprise social media that we study according to this analysis. In terms of sensitivity of the results, we note that omitting a week of data during a contribution anomaly (week 35), resulted in better predictive power across all factors, and also served to differentiate the results better. The main quantitative difference if the week would have been kept in the data was that the *AuthoredDocs* factor would not have been the best predictor in the macro metrics, and would thus have contradicted the micro analysis, which remained unaffected. In general the macro metrics were more sensitive to different treatments of the data, whereas the micro analysis showed less factor differentiation but remained stable.

5. ACTIVITY EVENT ANALYSIS

Activity regression considers how activity in a person's workgroup relates to a person becoming active and continuing activity. For the purposes of our analysis, an employee's *workgroup* consists of a person, his or her manager, and all direct reports to that manager. These workgroups are a different grouping of employees than the organization unit discussed in Section 4.

5.1 Method

For each venue, we defined a series of activity events. Defining whether a person is active is arbitrary. As a simple measure, which captures most of the continued activity of the more active users in our data, we consider a person "active" if they had posted to the venue within the previous 30 days, and "inactive" otherwise. Using a period of 30 days is a commonly used criterion for lack of activity. In our data, about 5% of a user's posts are more than 30 days after a prior post in the same venue by that user.

We could also consider other measures of inactivity, such as when people have significantly larger gap between posts then their individual prior history rather than a fixed time (30 days) for everyone. This would test whether the analysis significantly confounds two distinct processes: an active user deliberately deciding to become inactive vs. users who have a continuing, but low, participation rate. The former (explicit decisions to change prior behavior) may have more significance, and these two cases may require different methods to increase participation.

Macro (global) analysis				
	contribution predictability		contribution correlation	correlation structure
	R^2_{cp}	sgn_p	R^2_{cc}	PACF(1,2,3)
DistinctCountryComments	0.37	1.00	0.74	$(0.58, -0.10, 0.07)$
DistinctCityComments	0.37	1.00	0.77	$(0.58, -0.09, 0.11)$
DistinctEmployeeComments	0.36	1.00	0.76	$(0.58, -0.10, 0.14)$
DistinctOrgUnitComments	0.35	1.00	0.80	$(0.56, -0.06, 0.06)$
DistinctDocComments	0.34	1.00	0.76	$(0.56, -0.05, 0.12)$
DistinctCountryClicks	0.23	1.00	0.30	$(0.73, -0.10, 0.38)$
DistinctCityClicks	0.20	1.00	0.26	$(0.77, -0.10, 0.41)$
DistinctEmployeeClicks	0.20	1.00	0.26	$(0.77, -0.10, 0.40)$
DistinctOrgUnitClicks	0.19	1.00	0.27	$(0.79, -0.08, 0.38)$
DistinctDocClicks	0.13	1.00	0.37	$(0.68, -0.06, 0.26)$

Table 5: Results for macro (global) diversity analysis ordered by descending predictive power.

For the series of activity events, we denote each post as *continuing* or *new* activity according to whether the person had previously posted in that venue within the last 30 days. We ignore *new* events within the first 30 days a service is observed, because we don't know for certain that the user was inactive for the previous month. We add an *inactive* event for a user after 30 days of no posts.

A large reorganization at HP made the manager relation unstable in the directory, and so for this section we restricted our data set to events up through September 2008, removing about three months' worth of data.

We examined the relation between people becoming active and the activity of their managers by recording, for each *new* event whether their manager was active in the past 30 days in that venue, and p_{AM}, the fraction of employees with active managers at the time of the event. Under the null hypothesis that there is no relation between these properties, the fraction of *new* events with active managers should be close to that expected by random selection with probability

$$p_{AM} = \frac{\text{\# employees with active managers}}{\text{\# employees}} \tag{6}$$

Conversely, we compare manager activity of users who become inactive (i.e, *inactive* events) with the fraction of *active* employees who have active managers at the time of the event. If employees who are already active choose to become inactive with no relation to the activity of their manager, we expect the fraction of *inactive* events with active managers should be close to that expected by random selection with probability

$$p_{IM} = \frac{\text{\# active employees with active managers}}{\text{\# active employees}} \tag{7}$$

These probabilities vary with time. For our analysis we compare the observed events with these corresponding probabilities at the time of the event.

5.2 Results

5.2.1 *Manager and Coworker Activity*

For employees who become active, Figure 3 compares the observed and expected fractions of active managers for the different venues.

Of the people who become active, about $5 - 10\%$ have active managers at the time; compared to about a tenth that number for employees as a whole. We quantify this difference with randomization tests [9]. Specifically, for each *new* event e we record x_e

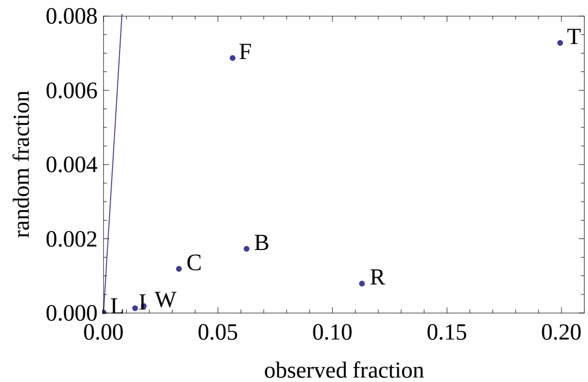

Figure 3: Fraction of *new* events with active managers: observed and expected if the events were selected at random from among employees at the time. The labels for the points indicate the venue (see the codes in Table 1). The line indicates equal values for the observed and expected fractions, which is well to the left of most of the observed values.

equal to 1 or 0 if the manager was active or not at the time, respectively. We also record the fraction $p_{AM}(e)$, from Eq. 6, at the time of the event. The observed number of events with an active manager is $n_{AM} = \sum_e x_e$.

To compare with the null hypothesis that there is no relation between these events and manager activity, we generate $N = 1000$ samples $s_1, \ldots, s_N$. A sample consists of selecting random values, 0 or 1, for each x_e where the probability for $x_e = 1$ is $p_{AM}(e)$. We denote the sum of these randomly generated values for sample i as s_i, which is a sample for the number of active managers for the events under the null hypothesis. The set of samples estimates the distribution of number of active managers we would observe under the null hypothesis. If the the actual observed value, n_{AM}, differs from most of these samples, the null hypothesis is unlikely to account for the observation.

This procedure shows that for all venues except Links, the high fraction of active managers is unlikely to arise from random selection among the population of employees (p-value less than 10^{-3}). The Links venue is inconclusive, with only 141 *new* events, none of which had active managers.

A similar analysis of *inactive* events shows employees who become inactive are slightly *less* likely than random to have active managers. The difference between observed values and selection

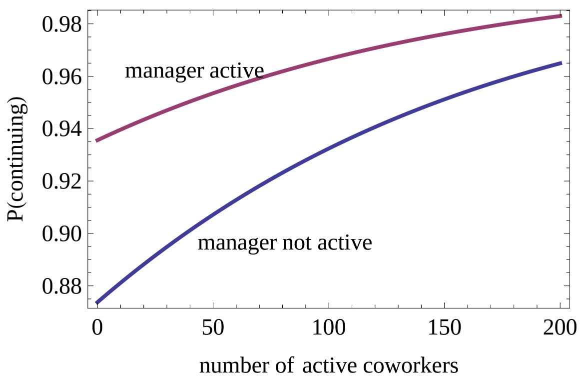

Figure 4: Proportion of events by active employees (i.e., the *continuing* and *inactive* events) that are *continuing* events, i.e., continuing participation by the employee, as a function of number of coworkers active in the prior 30 days and whether the manager is active, according to a logistic regression fit for the Blog venue.

parameter	value	confidence interval
β_0	2.7	2.4, 2.9
β_m	-0.74	$-0.98, -0.50$
β_c	0.007	0.004, 0.01

Table 6: Parameters of the logistic regression model of Eq. 8 for blogs, and their 95% confidence intervals.

with the null hypothesis is significant only for some venues: Blogs, Comments, Forums and Tags (all with p-value less than 10^{-3} except 10^{-2} for Comments). The other venues are consistent with the null hypothesis of no relation between manager activity and a user becoming inactive.

As a specific model of how activity of the workgroup relates to continued participation, we fit a logistic regression to the probability an event by an active employee is another post (*i.e.*, *continuing*) rather than becoming inactive, based on whether the manager was active at the time and the number of active coworkers. A logistic model is appropriate to model binary outcomes (in this case whether the event is *continuing* or not) [10]. In contrast to the regression model for a continuous outcome in Section 4, with binary outcomes the noise model involves independent Bernoulli trials, with the probability for success ranging between 0 and 1, depending on the model parameters. Specifically, we fit a model of the form

$$\Pr(\textit{continuing}|c, m) = \frac{1}{1 + \exp(-(\beta_0 + \beta_c c + \beta_m m))} \quad (8)$$

where c is the number of active coworkers and m is an indicator variable for inactive managers, i.e., is 0 or 1 according to whether the manager was active or not. A positive value of β_c means this probability increases as the number of active coworkers increases, i.e., the person is more likely to continue activity rather than become inactive. A negative value for β_m means *continuing* events are more likely when the manager is active.

For blogs, Figure 4 shows the regression model, with parameters in Table 6, relating the probability an event by an active employee is *continuing* rather than *inactive*. The estimated value for β_m is about 100 times larger in magnitude than β_c, so this model indicates whether the manager is active has far more influence than the activity of a single coworker in the workgroup.

event	number of events	number of prior posts mean	std. deviation
	blogs		
continuing	16010	141.5	241.6
inactive	1888	9.5	30.4
	forums		
continuing	90504	59.7	130.6
inactive	22219	6.4	11.7

Table 7: Number of prior posts for *continuing* and *inactive* events.

event	number of events	number of recent replies mean	std. deviation
	blogs		
continuing	10891	4.9	7.7
inactive	1370	0.7	1.9
	forums		
continuing	89424	4.4	7.6
inactive	22211	0.5	1.3

Table 8: Number of recent replies for *continuing* and *inactive* events.

5.2.2 Feedback and History

Using the list of activity events described above, we compared *continuing* and *inactive* events. As a measure of feedback, visible to the user, for each event we determine the number of recent replies to that person's posts in the venue (*i.e.*, within the last 30 days). The feedback is a measure of community interest in the user's participation. We also count the user's prior posts to the venue, as a measure of the level of the user's participation.

Table 7 compares the average number of prior posts in two venues, Blogs and Forums, with a large number of events. In both venues, users involved in *continuing* events tend to have longer history in the venue, as measured by the number of posts. To quantify the significance of the differences between prior posts seen in the table, we apply a permutation randomization test. That is, under the null hypothesis of no relation between type of event (either *continuing* or *inactive*) and number of prior posts, we generate samples by randomly permuting the number of prior posts among all these events. We then compare the observed difference in the averages with the distribution of these samples. These randomization tests for difference in means in prior posts between *continuing* and *inactive* events indicate the differences seen in the tables are unlikely to arise by chance if there were no difference between these event types (p-value less than 10^{-3}).

As discussed in Section 4, recent feedback to the user correlates with continued activity in terms of *number* of posts. Table 8 provides another view of that relation. The *continuing* events are associated with larger number of recent replies. Thus not only is the *amount* of subsequent activity related to feedback, but so is a user's choice to become inactive. The permutation randomization test described above indicates the differences are unlikely to arise by chance if there were no difference between these event types (p-value less than 10^{-3}).

5.3 Summary

Activity within a user's workgroup correlates with participation. Manager and coworker activity are correlated with employees becoming active in the venues we studied, supporting the second hypothesis, **H2**, in Section 1. On the other hand, *lack* of manager activity is only modestly correlated with employees becoming inactive, and only in some of the venues.

We suspect managers' participation is more important in venues that imply discussion (*e.g.*, blogs, blog comments, forums) than in venues that more naturally serve as memory archives (*e.g.*, links, wikis). Prior interviews with bloggers found that they seek external validation for their invested time [37]. We believe that managers "leading by example" has a positive impact on getting their direct reports to try participating in enterprise social media.

A long history of posts correlates with continued activity, as do having many recent replies. This correlation between continued activity and history occurs in a variety of web sites where users contribute content, including Digg and Wikipedia [34].

6. CONCLUSIONS

Through an analysis of temporal relations among measures of user activity, feedback and workgroup involvement, we found robust correlations between these measures and activity. These results match the factors people emphasized in prior interview studies of a small set of social media users in this company. In terms of predicting future participation in a venue, the number of prior posts and of recent posts account for the most variation among the factors we studied for individual users. This mirrors other findings that the more people contribute to an online community, the more likely they are to continue posting [34]. As additional factors, we found comments the user received were more predictive than total readership. Thus users' knowledge that their contributions are of interest to others in the organization relates to further participation, in accordance with factors mentioned in the interview studies:

> "That's one of the big weaknesses of it, the only way you know if anybody is reading it is if they take the trouble to reply. Without that you have no clue who people are...it's largely unidirectional." [37]

In terms of readership, measured by clicks, and attention, measured by comments, we find that diversity, particularly geographic diversity, is an important predictor of future participation. Since readership in general, and its diversity in particular, are not directly visible to users, this observed relationship suggests the readership measures are a proxy for the relevance of a user's posts to the community. This study was situated in a large global enterprise; in smaller, more collocated organizations, the diversity of visibility may be less of a motivation.

We also found activity within a user's workgroup correlates significantly with participation, but has a weaker relation with users becoming inactive. This suggests direct exposure to social media, via people the user interacts with regularly, is important to encourage people to start using the media. But once they start, feedback from throughout the organization becomes significant.

In short, it seems that managers' participation is a key motivator in getting people to *start* contributing to enterprise social media, while comments and a diverse readership are key in getting them to sustain their contributions. This echoes comments made in interviews by employees at HP:

> "After starting my blog, it was amazing to me how quickly I met other people, especially across different business groups... you know, I would post something on my blog and a week later I'd get an invitation to present on it." [37]

Corporate culture may be a factor in what motivates employees. At HP, as in many large organizations, employees' performance is evaluated principally by their direct managers, so managers have considerable influence. This effect may be smaller at companies that emphasize peer-review performance metrics or have matrix-style management structures. However, these findings are likely generalizable to all large companies with manager-driven performance evaluations.

6.1 Design Implications

These findings suggest that organizations seeking to reap the benefits of widespread social media usage should encourage managers to "lead by example" or at least support the practice. Indeed, a senior vice president at HP received high praise for engaging with individual contributors using his internal blog, and inspired a number of people in his division to experiment with blogs.

Once people have invested time in creating a blog post or social network profile, tools should provide some feedback to content authors that their content is being seen by others. Making measures of attention visible to users will help sustain participation by those whose posts are seen to be most interesting to the community. This could take the form of simple hit counters or deeper analytics about readers; for example, Flickr shows how many people viewed each photo, and YouTube now provides details about the pages linking to a video. But moreover, this work shows that in corporate environments, it's not just the raw quantity of attention an author receives but also *who* the readers are. This might suggest providing some basic demographics highlighting the diversity of an author's readership like Google Analytics provides for websites, or samples of readers' details similar to how LinkedIn shows the professions of visitors to a user's profile.

Another surprising finding in this work is that activity at the end of a week is relatively less likely to influence behavior the following week, suggesting that to some extent people may forget about the previous week's content over the weekend. While more work is needed to quantify this effect, it may encourage designers of corporate social media to help remind users of the context of conversations the previous week on Mondays.

6.2 Future Work

Our results, and their correspondence with the interview studies, suggest factors that may causally influence people to start and continue participating in social media. An important direction for future work is to test the extent to which the observed correlations are in fact causal. For social media within the organization, we can explicitly observe how changes in the organizational structure and the information presented to users affects participation. A further possibility is intervention experiments where different subgroups of people are provided with different feedback, which can give stronger confidence for causal relationships and suggestions of how to improve the number of participants and design feedback to encourage them to provide information relevant to others in the organization.

One direction for elaborating relationships between employees and coworkers or managers is to extend the linear regression models to consider nonlinear relations and whether there is a "dose/response" relation between participation and, say, how active a manager is in terms of number of posts rather than just a yes or no feature of whether the manager is active. This could distinguish whether manager activity acts mainly as a positive example or whether a manager's perceived disapproval of employees' participation is a more important issue, as expressed in some of the interviews.

We also plan to conduct intervention studies to explore whether exposing readership data to authors affects their behavior. While this could certainly motivate people whose content is widely read, it remains to be seen whether revealing that their content is unpopular might discourage other authors. Another area that merits exploration is whether the presence of an explicit "follower" network further encourages users to contribute as it does on Twitter [19].

Acknowledgments

We thank Sarita Yardi and Bernardo Huberman for helpful comments.

7. REFERENCES

[1] A. Anagnostopoulos, R. Kumar, and M. Mahdian. Influence and correlation in social networks. In *KDD '08: Proc. of the 14th ACM int'l conf. on Knowledge discovery and data mining*, pages 7–15, New York, NY, USA, 2008. ACM.

[2] J. Arguello, B. S. Butler, E. Joyce, R. Kraut, K. S. Ling, C. Rosé, and X. Wang. Talk to me: foundations for successful individual-group interactions in online communities. In *CHI '06: Proc. of the ACM conf. on Human Factors in computing systems*, pages 959–968, New York, NY, USA, 2006. ACM.

[3] G. Beenen, K. Ling, X. Wang, K. Chang, D. Frankowski, P. Resnick, and R. E. Kraut. Using social psychology to motivate contributions to online communities. In *CSCW '04: Proc. of the ACM conf. on Computer supported cooperative work*, pages 212–221, New York, NY, USA, 2004. ACM.

[4] Y. Benkler. Coase's penguin, or linux and the nature of the firm. *The Yale Law Journal*, 112:369–446, 2002.

[5] G. Box and D. R. Cox. An analysis of transformations. *J. of the Royal Statistical Society*, B(26):211–252, 1964.

[6] Y. Bramoulle, H. Djebbari, and B. Fortin. Identification of peer effects through social networks. Technical report, University of Laval, 2006.

[7] M. Burke and R. Kraut. Mind your ps and qs: the impact of politeness and rudeness in online communities. In *CSCW '08: Proc. of the ACM conf. on Computer supported cooperative work*, pages 281–284, New York, NY, USA, 2008. ACM.

[8] B. Butler, L. Sproull, S. Kiesler, and R. Kraut. Community effort in online groups: Who does the work and why? In S. Weisband and L. Atwater, editors, *Leadership at a Distance*. Erlbaum, 2002.

[9] P. R. Cohen. *Empirical Methods for Artificial Intelligence*. MIT Press, Cambridge, MA, 1995.

[10] D. Collett. *Modelling Binary Data*. CRC Press, Boca Raton, FL, 2nd edition, 2003.

[11] D. Crandall, D. Cosley, D. Huttenlocher, J. Kleinberg, and S. Suri. Feedback effects between similarity and social influence in online communities. In *KDD '08: Proc. of the 14th ACM int'l conf. on Knowledge discovery and data mining*, pages 160–168, New York, NY, USA, 2008. ACM.

[12] T. H. Davenport and J. C. Beck. The attention economy. *Ubiquity*, 2(14):1, 2001.

[13] J. DiMicco, D. R. Millen, W. Geyer, C. Dugan, B. Brownholtz, and M. Muller. Motivations for social networking at work. In *CSCW '08: Proc. of the ACM conf. on Computer supported cooperative work*, pages 711–720, New York, NY, USA, 2008. ACM.

[14] L. Efimova and J. Grudin. Crossing boundaries: A case study of employee blogging. In *HICSS '07: Proc. of the 40th Annual Hawaii int'l conf. on System Sciences*, page 86, Washington, DC, USA, 2007. IEEE Computer Society.

[15] S. Farrell, T. Lau, S. Nusser, E. Wilcox, and M. Muller. Socially augmenting employee profiles with people-tagging. In *UIST '07: Proc. of the 20th annual ACM symposium on User interface software and technology*, pages 91–100, New York, NY, USA, 2007. ACM.

[16] J. Grudin. Groupware and social dynamics: eight challenges for developers. *Commun.ACM*, 37(1):92–105, 1994.

[17] P. J. Hinds and J. Pfeffer. Why organizations don't "know what they know": Cognitive and motivational factors affecting the transfer of expertise. In M. S. Ackerman, V. Pipek, and V. Wulf, editors, *Sharing Expertise: Beyond Knowledge Management*, pages 3–26. MIT Press, Cambridge, MA, 2003.

[18] B. A. Huberman, D. M. Romero, and F. Wu. Crowdsourcing, attention and productivity. *CoRR*, abs/0809.3030, 2008.

[19] B. A. Huberman, D. M. Romero, and F. Wu. Social networks that matter: Twitter under the microscope. *First Monday*, 14(1), January 2009.

[20] J. Huh, L. Jones, T. Erickson, W. A. Kellogg, R. K. E. Bellamy, and J. C. Thomas. Blogcentral: the role of internal blogs at work. In *CHI '07: extended abstracts on Human factors in computing systems*, pages 2447–2452, New York, NY, USA, 2007. ACM.

[21] A. Jackson, J. Yates, and W. Orlikowski. Corporate blogging: Building community through persistent digital talk. In *HICSS '07: Proc. of the 40th Annual Hawaii int'l conf. on System Sciences*, page 80, Washington, DC, USA, 2007. IEEE Computer Society.

[22] P. Kolari, T. Finin, Y. Yesha, Y. Yesha, K. Lyons, S. Perelgut, and J. Hawkins. On the Structure, Properties and Utility of Internal Corporate Blogs. In *Proc. of the int'l conf. on Weblogs and Social Media (ICWSM 2007)*, March 2007.

[23] R. E. Kraut, S. R. Fussel, S. E. Brennan, and J. Siegel. Understanding effects of proximity on collaboration: Implications for technologies to support remote collaborative work. In P. Hinds and S. Kiesler, editors, *Distributed Work*, pages 137–162. MIT Press, Cambridge, MA, 2002.

[24] K. R. Lakhani and E. von Hippel. How open source software works: "free" user-to-user assistance. *Research Policy*, 32(6):923 – 943, 2003.

[25] C. Lampe and E. Johnston. Follow the (slash) dot: effects of feedback on new members in an online community. In *GROUP '05: Proc. of the int'l ACM conf. on supporting group work*, pages 11–20, New York, NY, USA, 2005. ACM.

[26] C. F. Manski. Identification of Endogenous Social Effects: The Reflection Problem. *Review of Economic Studies*, 60:531–542, 1993.

[27] D. R. Millen, J. Feinberg, and B. Kerr. Dogear: Social bookmarking in the enterprise. In *CHI '06: Proc. of the ACM conf. on Human Factors in computing systems*, pages 111–120, New York, NY, 2006. ACM.

[28] J. S. Milton and J. C. Arnold. *Introduction to Probability Theory and Statistics*. McGraw Hill, 2003.

[29] P. Resnick, K. Kuwabara, R. Zeckhauser, and E. Friedman. Reputation systems. *Commun. ACM*, 43(12):45–48, 2000.

[30] H. A. Simon. Designing organizations for an information-rich world. In M. Greenberger, editor, *Computers, communications, and the public interest*, pages 37–72. Johns Hopkins Press, 1971.

[31] C. Steglich and T. A. Snijders. Dynamic networks and behavior: Separating selection from influence. Technical report, Interuniversity Center for Social Science Theory and Methodology, University of Groningen, 2007.

[32] Y. Wang and D. R. Fesenmaier. Assessing Motivation of Contribution in Online Communities: An Empirical Investigation of an Online Travel Community. *Electronic Markets*, 13:33–45, 2003.

[33] W. W. S. Wei. *Time Series Analysis, Univariate and Multivariate Methods*. Pearson Addison Wesley, 2006.

[34] D. M. Wilkinson. Strong regularities in online peer production. In *EC '08: Proc. of the 9th ACM conf. on Electronic commerce*, pages 302–309, New York, NY, USA, 2008. ACM.

[35] F. Wu and B. A. Huberman. How public opinion forms. In *Proc. WINE*, volume 5385 of *Lecture Notes in Computer Science*, pages 334–341. Springer, 2008.

[36] S. Yardi, S. Golder, and M. Brzozowski. The pulse of the corporate blogosphere. In *Conf. Supplement of CSCW 2008*. ACM, November 8-12 2008.

[37] S. Yardi, S. A. Golder, and M. J. Brzozowski. Blogging at work and the corporate attention economy. In *Proc. CHI 2009*. ACM, to appear April 4-9 2009.