Marketing Science Institute Working Paper Series 2013
Report No. 13-113

Building the B[r]and: Understanding How Social Media Drives Consumer Engagement and Sales

Yogesh V. Joshi , Liye Ma, William M. Rand, and Louiqa Raschid

Report Summary

Yogesh Joshi , Liye Ma, William Rand, and Louiqa Raschid investigate how activity in digital social media, by new and established brands, relates to engagement with consumers, and eventually, sales.

Their dataset includes two years of Twitter activity and offline concerts for several musical bands, and the corresponding social media activity of the bands' followers. In addition to measuring volume (that is, number of tweets sent per unit of time), the authors use machine learning methods to analyze message sentiment and informational content. They also collect A.C. Nielsen sales data for all albums released by these bands.

They investigate the characteristics and evolution of consumers' engagement (propensity to tweet in response to a band's tweets as well as propensity to send informational or emotional tweets) using a hidden Markov model. They relate engagement to sales via a generalized diffusion model.

Findings

Overall, the authors find that band actions in social media generate interest and change in engagement levels of their followers, and these engagement levels have a positive association with sales. At the same time, there are notable differences in effects for new and established brands (in their context, relatively unknown versus well established bands).

- For followers of new bands, a moderate level of engagement is somewhat unstable: these consumers are prone to migrating to become either more fully engaged or disengaged. For followers of established bands, a moderate level of engagement is relatively stable (i.e., it is a "stickier" state).

- For both new and well-established bands, it is difficult to move followers out of an unengaged state. For new bands, a tweet has about a 10% chance of engaging a follower, with non-informational tweets doing slightly better. For well-established bands, a tweet that is informational and non-emotional has an almost 35% chance of engaging an unengaged follower.

- Messaging propensity is higher for followers of established bands at all levels of engagement. Followers of established brands retain information and emotion from firm messages more than followers of new brands. Interestingly, while for these consumers information is relatively "stickier" than emotion, for consumers of new brands, emotion is stickier than information.

- Higher consumer engagement, as well as a higher degree of informational and emotional content in messages, are all generally positively associated with sales, for both new and well established bands. In addition, while moderate engagement is positively related to sales for the established bands, such an association is not observed for new bands.

Implications
This study is among the first to report nuanced content-level findings in a dynamic framework of social media activity. The key implications are that engaging consumers in digital social media has a significant relation to firm sales. That engagement varies by the nature, frequency, and content of firm messages, as well as with characteristics of the brand.

It is crucial for managers to understand these nuances and design digital marketing strategies accordingly. Managers should experiment with tweet content, taking in account the relative strength of the brand when communicating with specific follower segments.

Yogesh V. Joshi, Liye Ma, and William M. Rand are Assistant Professors of Marketing, and Louiqa Raschid is Professor, Decisions, Operations and Information Technology, all at the Robert H. Smith School of Business, University of Maryland.

Acknowledgments
The authors gratefully acknowledge partial support from the Marketing Science Institute (MSI Award 4-1718) and the National Science Foundation (NSF Award IIS 1018361) for this research; thank students Leanna Gong, Jianyu Li, Radu Machedon, Derek Monner, Yixin Qiu, Hassan Sayyadi, and Shanchan Wu for their help in the data collection efforts; and thank Michel Wedel and Wendy Moe for their helpful comments on previous versions of this manuscript. The usual disclaimer applies.

Introduction

Over the past few years, there has been a significant increase in consumer adoption and use of digital social media. As a direct consequence of this growing interest, firms have been experimenting with various strategies in order to engage meaningfully with consumers within this media. A recent McKinsey survey of 1469 C-level executives reveals that while a majority list digital marketing and social tools as a top ten priority on their strategic agendas and expect social media to deliver value for their firms, only a few have begun any systematic engagement with consumers via these media (Brown and Sikes 2012). The primary reason for this reluctance among managers to fully embrace social media as a platform for marketing investments has been a lack of understanding of how firm actions in social media influence consumer engagement, and in the end, impact sales (Divol, Edelman and Sarrazin 2012). Given this context, in this research we analyze social media dynamics to address the following questions: How do firm messages in digital social media affect consumer engagement? How does this engagement evolve over time? And finally, how does such engagement relate to subsequent performance that generates value, such as sales? To begin answering these questions, in this paper, we empirically analyze how actions taken by music bands in digital media relate to consumer engagement, and subsequently, sales. Since the success of a band depends on the strength of its brand, we anticipate these insights related to the building of a musical brand to be of broader relevance to managers across industries interested in digitally building their own new and existing brands.

In the recent past, there has been growing academic interest as well as research in social media. Research has focused on issues such as understanding the diffusion process in a digital social network given individual connections between its members (Katona, Zubcsek and Sarvary, 2011), the impact of opinion leadership and social contagion within social networks on the adoption of innovations (Iyengar, Van den Bulte and Valente, 2011) the social dynamics of online rating forums (Moe and Trusov, 2011), and the role of connectivity versus activity (i.e. hubs vs. pumps) in online social networks (Stephen, Dover and Goldenberg, 2011). Much of this work combines social network analysis techniques with elements of marketing theory to address substantive issues of interest to managers. Consumer behavior and choices in digital social media differ substantially from traditional media (Leskovec and Kleinberg, 2009, van Doorn et. al. 2010, Hennig-Thurau et. al. 2010), as well as across different types of digital social media, such as blogs (Mayzlin and Yoganarasimhan 2012; Gruhl et al., 2004), YouTube (Oh, Susarla, and Tan, 2008), and Twitter (Huberman et. al., 2009). Consequently, investigating how activities in distinct social media affect consumer engagement and sales has also been a question of interest to managers and researchers alike (Chevalier and Mayzlin 2006).

To address this issue, researchers have started to take a deeper look at some of the online interactions and dynamics in order to provide a richer understanding of how consumers and firms interact in digital social media. For instance, Iyengar, Han and Gupta (2009) investigate how status and the online behavior of friends influence individual behavior. Godes and Mayzlin (2009) investigate firm side actions to see whether firm mediated word of mouth has an impact within a social network. Onishi and Manchanda (2010) study the interactions between traditional media and new media (blogs in their context) to show that both have a positive and interactive effect on sales, with traditional media playing an active role in influencing volume on new media. Stephen and Galak (2012) distinguish between traditional, owned (e.g., websites) and earned media (e.g., publicity) and demonstrated the existence of interactive effects: while all impact sales, social earned media can also play a role in driving activity in traditional media. Sonnier, McAlister and Rutz (2011) use proprietary online interpersonal communications data and automated sentiment scores to show that positive, neutral as well as negative communication has a significant impact on sales.

Along with academic researchers, marketing analytics companies such as comScore are also interested in researching this space: a recent report from comScore and Facebook on the power of "liking" on sites such as Facebook distinguishes between the impact and reach a brand can have on social media platforms by distinctly considering direct reach to fans or followers, and indirect reach to friends of these fans and followers (Lipsman et al. 2012). They conclude that while acquiring fans is important, engaging them in order to reach their friends is equally if not more important. Thus, the value of a brand's presence on social media might be influenced not just by direct reach, but also by engagement, loyalty and as a consequence a brand's indirect reach. However, the link between a brand's activity on social media, the resulting impact on consumer engagement, and the final important consequence on sales has been fairly elusive.

While this is a broad problem in general, in this research we specifically focus on understanding how firm actions in microblogs (a particularly popular digital social medium, e.g., Twitter, Facebook and LinkedIn status updates, Tumblr) influence engagement with potential consumers, and the relationship between this engagement and firms' sales. Defining and measuring consumer engagement is well recognized as a challenging task by researchers (e.g., Schau et. al. 2009, Gupta et. al. 2010, van Doorn et. al. 2010). van Doorn et. al. (2010) define engagement as "*customers' behavioral manifestation toward a brand or firm, beyond purchase, resulting from motivational drivers.*" Accordingly, our study focuses on understanding this underlying behavioral manifestation, i.e., engagement of consumers, given their actions in microblogs along with firm actions therein. Microblogging sites such as Twitter are well suited for our study, given the rapid growth in their adoption by consumers as well as the associated growing interest from both academia and industry, in understanding consumer activities and effective participation of firms on such platforms (Schweidel and Moe 2012, Toubia and Stephen 2013).

In this paper, we analyze the impact of firm actions on consumer engagement and sales within the context of the music industry. This is a particularly interesting industry in terms of social dynamics, since we observe many firms (i.e., music bands) at distinct levels of maturity. In particular, we focus on two distinct types of bands: some that are relatively well established, and others that are relatively unknown. We distinguish between these two types to compare and contrast how consumer engagement in digital social media might differ for new brands versus established ones. Across types, we aim to identify *effective* actions of well performing bands that other firms can learn from in developing their own social media strategies.

We identify the relatively new versus established bands as follows. To identify new bands, we focus on a set of bands that were involved in "rock residencies" – a not-so-uncommon activity among upcoming music bands. A rock residency is an event with a fixed program – typically spread over a pre-defined time period, that allows a relatively unknown musical band to perform for free in established night clubs (Ulaby 2010). Many music clubs across the nation have rock residencies on Monday nights (or other slow week nights) where bands perform in those clubs every Monday for one month. A rock residency serves three purposes: the band obtains a free venue to showcase and develop its talent, the club obtains a free performer, and consumers obtain free live music. Rock residencies are a particularly interesting phenomenon for the purposes of this study because not only are the bands performing within these residencies relatively unknown as a brand, but they also are on a mission to develop themselves into successful professional brands. In addition, given their demographic composition, most bands engaging in residencies tend to be reasonably comfortable with digital social media, and often proactive in using such media to connect with their audience and potential consumers, for building their brand. For instance, most new bands have their own Twitter profile, via which they reach out to their followers, a fan page on last.fm, alongside other digital presences. To study the relatively more established brands, we identify a set of bands that performed at the same music clubs hosting rock residencies, but on the more popular nights of the weeks (e.g., Thursdays/Fridays/Saturdays).[1]

We collect activity data for these bands and their followers on Twitter, as well as information about offline concert performances by the bands during this period. This provides us with a rich set of information to investigate what specific types of messages sent by a band translate into effective engagement with their followers. To better analyze the impact of these messages, we use existing computational techniques for sentiment analysis (i.e., supervised machine learning approaches, such as a Naive Bayes classifier and random trees) to classify the content of these messages along two key distinct dimensions: emotion and information (Brieman 2001, Go, Bhayani and Huang 2009, Celikyilmaz, Hakkani-Tur and Feng 2010, Kouloumpis, Wilson and Moore 2011, Machedon, Rand and Joshi 2013, Monner, Rand and Joshi 2013). Separately, we also collect sales data for the albums released by these bands. Using these sales data, we further investigate to what

[1] By focusing on bands that have performed at the same clubs, we hope to minimize the impact of other extraneous factors on our analysis of interest.

extent the resulting engagement with followers in digital social media translates into monetary benefits for the firm. With this research, we seek to shed some light on (1) how b[r]ands leverage one digital social medium to attract and motivate followers, (2) whether specific b[r]and messages impact followers' engagement levels and subsequent activity, (3) whether offline activity has any significant impact on online activity, (4) what the role of activities in social media is in driving brand sales performance and hence value, and (5) contrast how actions and consequences differ for the relatively newer brands vs. the more established ones.

Extensive past research has shown that consumer behavior progresses through various stages or phases, such as need arousal, information search due to awareness, generation of consideration sets, belief updating, evaluation, purchase and post-purchase reactions (for an excellent review, see Roberts and Lilien 1993, Eliashberg and Lilien 1993, Wedel and Kamakura 2000). At the same time, research on social interactions has highlighted the importance of firm actions on consumer response and engagement (e.g., Godes et. al. 2005). Of late, researchers have used hidden Markov models (HMM) to characterize consumers' different relationship states with firms, and their transitions among states over time (e.g., Netzer, Lattin and Srinivasan 2008, Li, Sun and Montgomery 2011, Shi, Wedel and Pieters 2013). In the social media context, ongoing interactions between firms and consumers allows for the possibility of different levels of engagement for consumers with firms, and for firms to influence this engagement through specific types of actions. Moreover, while firm actions may or may not cause instantaneous changes to consumer engagement levels, consistent actions could lead to long term changes in terms of consumer responses. Hence, HMM serves as a good approach for modeling this problem. Following existing research (e.g., Netzer, Lattin and Srinivasan 2008, Li, Sun and Montgomery 2011, Shi, Wedel and Pieters 2013), we use HMM to understand how firm actions result in various consumer engagement states.

Our estimation shows that consumers exhibit differing levels of engagement with firms; and their levels of engagement also evolve over time, depending on firm actions. We also find notable differences in the nature and dynamics of engagement between followers of new brands and followers of established brands. Specifically, we find that a moderate level of engagement is a somewhat unstable state for followers of new brands, and that such consumers are prone to migrating out of this moderate level of engagement to become either more fully engaged, or disengaged. In contrast, a moderate level of engagement is relatively stable for followers of established brands (i.e., it is a stickier state). Furthermore, we find that the specific content of a firm's communication has distinct effects on consumer engagement; such effects vary for new and established brands; as well as from consumer to consumer, depending on their existing engagement levels. Finally, we find that the level of engagement recovered from our study is strongly associated with the sales of the firm's products, and the nature of this association is also distinct for the new and established brands in a way that is consistent with the characteristics of engagement levels for these brands.

Our research makes a few key contributions. First, we demonstrate the dynamic effects of firm actions in microblogs on consumer engagement. Second, we show that the relationship between firm actions and consumer engagement goes beyond a simple correlation in volume, and that informational and emotional content by firms have distinct effects on consumer engagement, which vary depending on both the type of firm and the status of its consumers. Finally, we show that follower engagement in microblogs is positively linked to sales of a firm's products, thus confirming the importance of consumer engagement on social media websites. To our best knowledge, we are among the first to report such nuanced content level findings at the content level in a dynamic framework, and to specifically relate the impact of firm actions on consumer engagement as well as sales. Consequently, our findings provide direct and timely guidance to practitioners in terms of the nature of messages that can be leveraged to engage with followers, and eventually impact sales.

The rest of this paper is laid out as follows. Below, we first describe the data used in our analysis, and the sources for collection of this data. In this section, we also discuss how we processed available data to generate measures of volume, sentiment, and information content. This is followed by the model section, where we lay out the details of our modeling approach. We then discuss the findings from our analysis and the key insights we derive based on these findings. Finally, we conclude with a summary of our observations and thoughts on future research.

Data

Our dataset consists of music bands that have performed during a two year period (2008-2009) in two clubs in a major metropolitan area in the United States. Among these, 37 bands engaged in rock residencies and were new and up-coming bands, while 20 performed on regular club nights (e.g., Thursday and Friday nights), and were relatively more established bands. Thus, we have 37 focal bands and 20 control bands. For each of these bands, we then collect data from various sources on their social media activities, live performances, and album sales. For social media activity, we focus on the popular micro-blogging site, Twitter.

Twitter data

The Twitter data was collected between October 18, 2010 and January 31, 2011, and spans the 2008-2011 time period. We first identified the number of followers for each band to understand band popularity and reach, and created a list of followers for each band. We retroactively collected all tweets sent by each of these bands between July 1st, 2008 and January 31, 2011, totaling 135 weeks, and all tweets sent by the followers of these bands during the same time period that included the name or twitter ID of any of the bands under consideration.[2] This dataset contains

[2]Note that the band names in our data set are quite unique, hence tweets that contain the band name are highly likely to be about the band.

30,826 tweets sent by the bands, and 7,218 tweets sent by 2,649 of these bands' followers (these bands have more followers, but many of them did not tweet about the bands during the time period). Table 1 shows the summary statistics for the bands and followers activity in Twitter, and reveals some interesting observations.[3] First, focal bands, which are the more up and coming bands, tweet more than the control, established bands. Not only do the focal bands tweet more, but so do their followers. As one would expect, the control bands have more followers, but the difference is not as striking. Finally, the standard deviations are higher than the means, indicating that activity in Twitter is quite skewed across both band types.

Concerts data

While Twitter data provides us with a measure of a band's online activity, we are also interested in measuring a band's offline activity, since offline activity can often also play a role in determining consumer engagement and sales. The best measure for offline activity for these emerging bands is the concerts they played, as bands tend to promote and publicize themselves ahead of and around these events, to the extent possible. Hence, for each band, we collect data about past concerts they performed from Songkick[4] during the same time period as the Twitter data collection. A summary of these data is provided in Table 2. As expected, on average, the more established control bands perform more offline concerts than the focal bands.

Sales data

We collect sales data for each of our bands from AC Nielsen. For each band, we first determine the list of albums released. The sales data contains a total of 34 albums by focal bands, and 29 albums by control bands, that were released during the 135 weeks covered by our tweet data. For each of these albums, the weekly sales numbers are provided by AC Nielsen across all available channels (traditional, as well as non-traditional ones). The descriptive statistics are reported in Table 3. From the table, we see that as might be expected, the average weekly unit sales as well as the number of previous albums are higher for the control bands than for the focal bands, and the standard deviation for sales of focal bands is higher than that for control bands. We include only the albums whose release dates were within the 135 weeks covered by the tweet data, since sales of new media products, such as albums and movies, are usually concentrated in the early weeks right after release.[5]

Content analysis of tweets

[3] All tables and figures are provided at the end of the text, after References.
[4] www.songkick.com
[5] Such a pattern is also very distinctly observed in our data.

Our Twitter data collection exercise provides us with a large set of tweets sent by both the bands and followers (consumers). Such data lend itself to easy measurement of volume (e.g., number of tweets sent per unit of time). However, although volume is an important characteristic of band action in social media, it alone is not sufficient. Each tweet, containing up to 140 characters, likely carries additional diagnostic content. Analyzing the content of these tweets on a large scale is a challenging but important task. To complement the conventional volume metric, and enable us gain a more nuanced understanding of the different firm actions, we use two types of content analysis to generate numerical scores for each tweet.

Conceptually, we view a firm's actions in social media as having an intent of generating engagement on part of the consumer. To this end, tweets act like advertisements. Similar to ads, they might seek to make informational appeals and/or persuasive emotional appeals (e.g., Batra et. al. 1996, Bagwell 2007). Thus, a firm might seek to provide information to the consumer that allows her to learn more about the firm, which may increase awareness about the firm. Alternatively, a firm might primarily seek to persuade a consumer by establishing an emotional connection, hence send messages that are high on emotional content. Hence, for each tweet, we measure informational content as well as emotional content, using existing computational text analysis procedures that use supervised machine learning techniques (e.g., Go, Bhayani and Huang 2009, Celikyilmaz, Hakkani-Tur and Feng 2010, Kouloumpis, Wilson and Moore 2011, Machedon, Rand and Joshi 2013, Monner, Rand and Joshi 2013). To measure informational content, we use supervised machine learning in order to automatically classify whether or not a tweet is informational. We follow the procedure described in Machedon et. al. (2013) to generate a $0/1$ score for each tweet, where 1 indicates that the tweet content is informational. To measure emotional content, we follow the procedure described in Monner et. al. (2013), and based on a set of pretests, we generate a $0/1$ sentiment score for each tweet, where 1 indicates that the tweet content is emotional.[6]

Table 4 summarizes the results of the above two classification rules for band and follower tweets. We emphasize that the emotional classification and informational classification are two separate dimensions, i.e. a tweet can be both emotional and informational, or neither emotional nor informational, etc. From this table, we see that followers in general tend to be more emotional in their tweets than bands: 52.93% of follower tweets are emotional for focal bands and 60.68% for control bands, compared with 32.09% of band tweets being emotional for focal bands and 36.19% for control bands. Further, followers of established bands are relatively more emotional (60.68% vs. 39.32%), whereas followers of new bands are almost equally likely to be emotional or non-emotional in their messages (52.93% vs. 47.07%). In terms of informational content, the descriptive statistics show no major differences between tweets sent by either types of bands or their followers. For both types of bands and their followers, around 30% of tweets are informational.

[6]Note that emotional tweets include tweets that score high on positive as well as negative sentiment.

Finally, the correlation between messages and their content by band across all bands is reported in Table 5. In this table, T_{bands} denotes the number of tweets sent by the band; $\%T_{bands}^{emotional}$ denotes the proportion of a bands tweets that were emotional in nature; etc. A few notable observations can be made from this table. First, the number of band tweets is positively correlated with the number of follower tweets. This serves as preliminary evidence that tweets by a band and tweets by its followers are related, warranting a more detailed analysis in order to better understand this relation. Second, the more a band tweets, the higher its proportion of emotional tweets, and the lower its proportion of informational tweets. Third, the proportion of emotional tweets by a band is positively correlated with the proportion of informational tweets by the band. This positive correlation between informational and emotional tweets is also observed for follower tweets, although the correlation is not as strong as for the band tweets. Finally and more interestingly, the proportion of band emotional tweets is positively correlated with the proportion of follower emotional tweets, and the proportion of band informative tweets is positively correlated with both the proportion of follower informative tweets and that of follower emotional tweets. This suggests that band tweets and follower tweets are related not just at the volume level, but also at content level, and there is clear evidence of the "spillover" of informational and emotional content from bands to their followers. All these observations confirm the close relationship between a band's and followers' tweeting activities, and motivate our modeling effort.

Model

In this section, we describe our modeling approach. We first set up a model for followers' tweeting activities, where our focus is on understanding the behavioral characteristics of different levels of consumer engagement, how such engagement levels change over time, and how bands' actions influence a change in these engagement levels. A conceptual diagram of our model is shown in Figure 1. The centerpiece of our analysis is a hidden Markov model (HMM) of different levels of consumer engagements. As discussed earlier, HMM is a good fit for our study, since it allows for both multiple levels of engagements and the dynamic transition among these different levels over time, depending on firm actions (Netzer, Lattin and Srinivasan 2008). In our model, there are different levels of consumer engagements, each captured as a state. The consumer's overall propensity to tweet, and the propensities to send emotional tweets or informational tweets, all depend on the engagement state. Consumers naturally transition among these engagement states over time. Such transitions are further influenced by the tweets sent by the bands.

A natural limitation of HMM, though, is that conditional on the current state, action is independent from history. In our context, it implies that a band's tweet influences follower's tweets only through a change in the engagement level. Such a formulation might be highly restrictive under certain circumstances. For example, consider a band sending a tweet describing its album. Such a tweet may generate interest in the band among its followers. Furthermore, it may also provide the followers with information so that the follower has new content to tweet about in the subsequent time periods. The former effect is captured by HMM, through states and transition. However, the latter effect is not. Hence, we introduce in our model, in parallel to the engagement states, two "stock" variables that capture the amount of accumulated information and emotion gained through band tweets, as shown in Figure 1. These stock variables change over time depending on band tweets, and will influence the amount and type of tweets sent by the follower.

Along with engagement, we also analyze the relationship between engagement level and the observed sales of bands' albums. To the extent that engagement level inferred from our tweets model can explain band sales, it serves as a validation of the appropriateness of the model, and testifies to the managerial importance of understanding such engagement. This relationship is illustrated in Figure 2, where among other factors, the proportion of followers in the different levels of engagement, followers' information stock, and followers' emotion stock all explain the sales for the bands' albums at the time; along with a bands' offline actions. We model the relation between consumer engagement and album sales using a generalized diffusion model. We now discuss each of these model components in detail.

The tweets model

We first explain the tweets model in detail. There are I twitter users, each following one of the bands. There are T time periods, each representing a week in our context. During any time period, a band may send one or more tweets, or none at all. Similarly, during each time period, a user may send one or more tweets about the band, or none at all.

The content within each tweet is classified using two key dimensions, as discussed earlier. The first is the emotional dimension, where a tweet is classified as being either emotional or non-emotional. The second is the informational dimension, where a tweet is classified as being either informational or non-informational. These are two separate dimensions, thus a tweet can be both emotional and informational, neither emotional nor informational, or emotional but non-informational or vice versa.

Denote the tweets sent by user i at time t as $C_{it} = \{C_{it}^{NE,NI}, C_{it}^{E,NI}, C_{it}^{NE,I}, C_{it}^{E,I}\}$, where the superscripts NE/E represent non-emotional and emotional, respectively, and NI/I non-informational and informational, respectively. In a slight abuse of notation, we also use C_{it} to represent the total number of tweets where it does not cause confusion. We model the tweets using a count model, assuming that the number of tweets in each time period comes from a Poisson distribution:

$$C_{it} \sim Poisson(\lambda_{it}) \tag{1}$$

Furthermore, we denote the probability of a tweet being emotional as $p_{E,it}$ and that being informational as $p_{I,it}$. Treating these as orthogonal dimensions, we then model the number of tweets of each type also using Poisson distributions:

$$
\begin{aligned}
C_{it}^{NE,NI} &\sim Poisson(\lambda_{it} \cdot (1 - p_{E,it}) \cdot (1 - p_{I,it})) && (2)\\
C_{it}^{E,NI} &\sim Poisson(\lambda_{it} \cdot p_{E,it} \cdot (1 - p_{I,it})) && (3)\\
C_{it}^{NE,I} &\sim Poisson(\lambda_{it} \cdot (1 - p_{E,it}) \cdot p_{I,it}) && (4)\\
C_{it}^{E,I} &\sim Poisson(\lambda_{it} \cdot p_{E,it} \cdot p_{I,it}) && (5)
\end{aligned}
$$

We seek to understand a user's level of engagement with the band she follows through her tweets and the band's tweet over time. Similar to the user, the band can also send a number of tweets of different types in each time period. We denote the band tweets received by the user i at time t as $B_{it} = \{B_{it}^{NE,NI}, B_{it}^{E,NI}, B_{it}^{NE,I}, B_{it}^{E,I}\}$.

We use a finite number of states to represent users' levels of engagement with the bands at different point in time. Each state represents a specific level of engagement. For example, a user may be in a state of "low engagement," when she is not engaged with the band at all, or be in a state of "moderate engagement" where she shows some interest in the band, etc. Denote the state of user i at time t as $s_{it}, s_{it} \in \{1, ..., S\}$, where S is the total number of states. Since the states represent different levels of engagement, tweeting behaviors are expected to differ by state.

In addition to the different levels of engagement, different users may have different amount of information on the bands they follow, and such information may evolve over time based on the volume and content of tweets sent by the bands. If a band frequently sends out informational tweets, for example, a follower may know more about the band, i.e. have more information, than does a user who follows a band which seldom sends such tweets. We capture this amount of information using an information stock variable. We denote the information stock of user i at time t as I_{it}. Similarly, a follower may also harbor an emotion stock of the band, which depends on the number of emotional tweets the band sends over time. We denote this emotion stock variable of user i at time t as E_{it}.

The engagement state as well as the information and emotion stock of a user are likely to affect both the amount and type of tweets sent by that user. For example, we may expect that a more engaged user may be more likely to tweet, and more information may make a user more likely to tweet. Furthermore, we may expect that the more information a user has, the more likely are her tweets to be informational. To account for such dependencies, we model the tweeting rate parameters as follows:

$$
\begin{aligned}
ln(\lambda_{it}) &= ln(\lambda_{i,0,s_{it}}) + \beta_{I,s_{it}} I_{it} + \beta_{E,s_{it}} E_{it} && (6)\\
logit(p_{E,it}) &= \gamma_{0,s_{it}}^{E} + \gamma_{I,s_{it}}^{E} I_{it} + \gamma_{E,s_{it}}^{E} E_{it} && (7)\\
logit(p_{I,it}) &= \gamma_{0,s_{it}}^{I} + \gamma_{I,s_{it}}^{I} I_{it} + \gamma_{E,s_{it}}^{I} E_{it} && (8)
\end{aligned}
$$

The equations show that the propensity to tweet, and the propensity for a tweet to be emotional, and to be informational, are all state-dependent and are influenced by the information and emotion stock of the user at the time. The coefficients for the information and emotion stocks are also state dependent, as users in different states may respond to information and emotion level differently. More information, for example, may not sway the tweet propensity of an unengaged user by much, whereas it may make an active user much more likely to tweet.

Users naturally transition among the states over time, making this a hidden Markov model (HMM). Band tweets are expected to influence such state transitions. An enthusiastic announcement about an upcoming album, for example, may get a follower engaged very quickly. We denote the state transition matrix as $A_{it} = \left[a_{it,jk}\right]_{j,k=1,\ldots,S}$. That is, the state transition matrix is an $S \times S$ matrix, where element $a_{it,jk}$ denotes the probability of user i transitioning from state j to state k at time t. Starting from a state j, the user can transition to any of the other states in the next time period. We model this using the multinomial logit setup:

$$U_{it,jk} = \alpha_{jk,0} + \alpha_{jk,1} B_{it}^{NE,NI} + \alpha_{jk,2} B_{it}^{E,NI} + \alpha_{jk,3} B_{it}^{NE,I} + \alpha_{jk,4} B_{it}^{E,I} \tag{9}$$

$$a_{it,jk} = exp(U_{it,jk}) / \sum_{l=1}^{S} exp(U_{it,jl}) \tag{10}$$

That is, there is an intrinsic probability of transition from state j to state k, captured by $\alpha_{jk,0}$, and the probability is altered by the number of different tweets the band sends at the time. For example, an emotional and informational tweet by the band may make the user more likely to move from state j to state k, and this will be reflected as a positive sign for the coefficient $\alpha_{jk,4}$. We normalize $U_{it,jj} = 0$ for each j for identification.

In addition to states, the information stock and emotion stock of a user also evolves over time according to the band tweets. For the information stock, we expect only the informational tweets from the band, $B_{it}^{I} = B_{it}^{NE,I} + B_{it}^{E,I}$, to matter, and we model the evolution of information stock as follows:

$$I_{it} = \delta_i^I I_{it-1} + B_{it}^I \tag{11}$$

For the emotion stock, we expect only the emotional tweets from the band, $B_{it}^{E} = B_{it}^{E,NI} + B_{it}^{E,I}$, to matter, and we model the evolution of emotion stock as follows:

$$E_{it} = \delta_i^E E_{it-1} + B_{it}^E \tag{12}$$

That is, receiving informational tweets from the band increases the user's information stock, and receiving emotional tweets from the band increases the user's emotional stock, while the information stock depreciates over time as captured by the factor $\delta_i^I \in (0,1)$, and the emotion stock depreciates over time with a rate $\delta_i^E \in (0,1)$.

We assume the intrinsic tweet propensity parameter $\lambda_{i,0,s_{it}}$, and the two depreciation parameters δ_i^I and δ_i^E for each individual user to be draws from corresponding population level Normal distributions, as follows:

$$ln(\lambda_{i,0,s_{it}}) \sim Normal(ln(\overline{\lambda}_{0,s_{it}}), \sigma^2_{\lambda,s_{it}}) \tag{13}$$

$$logit(\delta_i^I) \sim Normal(logit(\overline{\delta^I}), \sigma^2_{\delta^I}) \tag{14}$$

$$logit(\delta_i^E) \sim Normal(logit(\overline{\delta^E}), \sigma^2_{\delta^E}) \tag{15}$$

This completes the setup of the tweet model as a hierarchical Bayesian model, which lends itself to estimation using Markov-Chain Monte Carlo methods.

The album sales model

We now describe the album sales model used in our study. Since the purpose here is simply to assess the relation between the engagement level derived from our tweet model and the actual sales of album, we follow the well established diffusion approach for modeling album sales by using a Generalized Bass Model (GBM) (Bass 1969, Bass, Krishnan and Jain 1994, Bass, Jain and Krishnan 2000). Specifically, let there be A albums. For an album a, its sales in time t, where t indexes the time period since the release of the album, is:

$$R_{a,t} = M_a f_a(t) \tag{16}$$

where $f_a(t)$ is the hazard of album a at time t, which evolves as per a generalized Bass model that incorporates the effects of engagement and offline actions, as follows:

$$\frac{f_a(t)}{1-F_a(t)} = (p_a + q_a F_a(t)) exp(\overrightarrow{\beta}\mathbf{X}_{a,t}) \tag{17}$$

and

$$ln(M_a) = \gamma_0 + \overrightarrow{\gamma}\mathbf{Z}_a \tag{18}$$

In the model, p_a and q_a are the coefficients of innovation and imitation for album a, respectively. $\mathbf{X}_{a,t}$ are the time-varying covariates of album a that affect its sales. M_a is the market potential for album a, and $\mathbf{Z}_a$ are the covariates that affect such potential. To relate album sales to follower engagement, the covariates $\mathbf{X}_{a,t}$ include: the proportion of followers of the band in each engagement state at the time, the average information stock of the followers at the time, the average emotion stocks of the followers at the time, and the number of concerts performed by the band at the time. Except for the last variable, the others are measures derived from the tweet model. The last variable is a proxy for the important control variable for bands' offline promotion activities, since bands often tend to advertise and promote themselves and their albums during and through live performing concerts. The covariates $\mathbf{Z}_a$ include the number of albums previously released by the band and the number of twitter followers of the band. The former is included to account for the effect of experience, while the latter is a proxy for the band reach.

Furthermore, we assume the coefficients of innovation and imitation for each album a are drawn from population level distributions, as follows:

$$logit(p_a) \sim Normal(logit(\overline{p}), \sigma_p^2) \quad (19)$$

$$ln(q_a) \sim Normal(ln(\overline{q}), \sigma_q^2) \quad (20)$$

This completes the hierarchical Bayesian setup for the album sales model, which lends it to estimation through MCMC.

Results

We posit that the activities of followers of focal bands would differ meaningfully from those of followers of control bands, and so would album sales. Consequently, for both the tweet model and the album sales model, we estimate two separate sets of parameters, for focal bands and for control bands, even though the structure of the model is identical. In the discussion below, we report estimates for focal bands alongside those for control band, and compare and contrast the two sets of estimates.

We adopt a Bayesian approach and estimate our model using Markov Chain Monte Carlo (MCMC) methods. The details of the MCMC procedure are specified in the Appendix. We run 40,000 iterations, with the first 20,000 discarded for burn-in and the remaining used for inference. We run multiple MCMC chains from different starting values. The results indicate that our estimates are not sensitive to starting values.

Followers for estimation

Our original tweet dataset contains 2,649 followers of the focal and control bands who have tweeted about the bands. Our model focuses on the engagement levels and the transition among them over time for individual followers. To ensure adequate data variation for estimation, we excluded those followers who have tweeted only once during the 135 weeks. Research on consumer dynamics often focuses on heavy users (e.g. Guadagni and Little 1983, Erdem and Keane 1996). In our dataset, however, we would be left with too few followers if we increase the exclusion threshold. Furthermore, our hierarchical Bayesian model setup allows us to include units for which we have only a few observations. The dataset used for estimation thus contains followers who have tweeted at least twice. There are 999 such followers. The descriptive statistics of their tweets, by band type, are reported in Table 6. As the table shows, consistent with the findings reported earlier in the data section, followers of focal bands on average tweet slightly more than followers of control bands.

Model selection

We estimated the model for different numbers of latent states, ranging from 1 to 5 (the 1-state model is equivalent to a model with all other components of our proposed model but without latent engagement states). We use log-marginal density as the model selection criteria (Newton and Raftery 1994, Chib 1995). These log marginal densities are reported in Table 7. As the table shows, all the multiple-state models (from 2-state to 5-state) have much better model fit than the 1-state model. This indicates that it is important to account for the different levels of underlying engagement. For the multiple-state models, log-marginal density is maximized at 3-states. Examining the parameter estimates shows that this is reasonable: for both the 4-state and 5-state model, there are multiple states where the propensity to tweet is very low such that the consumer is unlikely to tweet at all, thus adding these additional states does not improve model fit or lead to better understanding of consumer activity. Based on these results, we choose the 3-state model for our subsequent analysis. We denote these three states, starting with the lowest state, as followers having "low" (S_L), "moderate" (S_M), or "high" (S_H) level of engagement with the band. We will use this notation in the subsequent discussion to interpret model results. During this interpretation, we interchangeably refer to follower in a state of low engagement as being unengaged, those in a state of moderate engagement as being somewhat engaged, and followers in the high engagement state to be highly engaged, with the band.

The impact of band actions on follower engagement

Having postulated three underlying levels of consumer engagement, we first analyze the impact of a brand's actions in social media on these three levels of consumer engagement. As outlined in Figure 1, we first focus on the engagement states and consumers' transition across them. The estimates for the hidden Markov model for the focal bands are reported in Table 8.

From Table 8, we see that in the absence of band tweets, unengaged consumers are unlikely to move up in their level of engagement: the intercepts corresponding to movement out of the unengaged state (S_L) are both negative. For somewhat engaged consumers, they exhibit a tendency to move towards the unengaged state but are unlikely to move to the more engaged state in the absence of any band communications. On the other hand, the highly engaged consumers, in the absence of any band action, are more likely to remain in a fully engaged state than to move to either the somewhat engaged or unengaged state.

When a band tweets, the nature of the message has a differential impact on the change in the level of engagement among followers. First, tweets of any type will help move unengaged consumers to a higher level of engagements (all coefficients are positive and statistically significant for transitions from S_L, whether to S_M or S_H). When followers are unengaged, non-informational messages are more likely to improve chances of engagement. An emotional but non-informational tweet has the highest impact among different types of tweets. A potential explanation for this is that for generating engagement, unengaged followers might first need to establish a strong emotional connection with a band's message, before acting on the information being provided by the band.

For somewhat engaged followers, a tweet with any message type lowers their propensity to move to an unengaged state (all four coefficients from S_M to S_L have negative posterior mean, and three of them are statistically significant). Emotional and informational tweets have the highest impact in keeping somewhat engaged followers from being disengaged. Interestingly, only emotional but non-informational tweets move these consumers to a highly engaged state (only coefficient that is positive and statistically significant for transition from S_M to S_H). However, that effect is also weakened by the fact that a tweet of such type is the least effective in preventing followers from transitioning to the unengaged states. Tweets that are non-emotional and non-informational are likely to keep somewhat engaged followers from becoming disengaged, but are not helpful in further increasing their engagement.

For the fully engaged state, in general, a message with any type of content helps keep consumers in the same state (majority of the coefficients from S_H are negative and statistically significant). Transition from this state, if it takes place, is more likely to be to an unengaged state rather than a somewhat engaged state. Somewhat surprisingly, emotional but non-informational tweets might cause fully engaged consumers to transition to an unengaged state (for $S_H \rightarrow S_L$,the coefficient for E,nI is 0.09 and statistically significant). A potential explanation for this observation is that when followers are fully engaged, "empty" emotional messages might no longer be enough: their high level of engagement might lead them to desire to keep being informed by the band about ongoing developments with the band. We do note, however, that the coefficient is quite small, and the increased probability of transitioning to the unengaged state is counter-balanced by the much reduced probability of transitioning to the somewhat engaged state (the corresponding coefficient for $S_H \rightarrow S_M$ is -1.54 and statistically significant).

We now turn our attention to the state-transition estimates for the control bands, as reported in Table 9. First, note that all intercepts except one are negative and statistically significant, indicating that consumers in general are quite likely to stay in the same state of engagement in the absence of any tweets. This is not surprising, since we are now looking at the followers of established bands, who have been around for a longer time, allowing consumers to have a good sense of the fit of the band's music with their tastes, leading to more stable underlying states.

Considering followers in the unengaged state, we see that informational but non-emotional tweets have a strong effect in transitioning them to both somewhat engaged and fully engaged states (the coefficient for nE,I is positive and statistically significant for both $S_L \to S_M$ and $S_L \to S_H$). Understandably, given that the bands have been around for some time, followers may need actual information in order to become engaged, instead of just emotional appeals. Note, however, that emotional but non-informational tweets do help followers transition directly to the fully engaged state (the coefficient for E,nI is positive and statistically significant for $S_L \to S_H$), although this mostly reflects just a substitution between the somewhat engaged and fully engaged states (the coefficient is negative and statistically significant for $S_L \to S_M$). Overall, this shows that for unengaged followers of control bands, while an informational tweet is the most effective, an emotional tweet may somewhat help in transitioning directly to the fully engaged state instead of the somewhat engaged state.

For followers in the somewhat engaged state, tweets of any type reduce the probability of transitioning to the unengaged state (coefficients for $S_M \to S_L$ are negative, and except one, all statistically significant). However, while tweets help prevent these followers from becoming disengaged, they do not seem to be effective in transitioning them into the fully engaged states (only the coefficient for E,nI is positive for $S_M \to S_H$, and even then it is not statistically significant). This suggests that, compared to the case of focal bands, for followers of control bands the somewhat engaged state is less of an interim state, lacking "upward mobility", i.e. followers may transition from the unengaged state to either the somewhat engaged or fully engaged state. But if they get to the somewhat engaged state, they may lack the interest to become further engaged.

Finally, for followers in the fully engaged states, band tweets in general reduce their probability of transitioning to the unengaged state. However, such tweets do not appear to be effective on preventing followers from transitioning to the somewhat engaged state, for which the default probability is quite high. This suggests that for control bands, the fully engaged state of followers is less sticky than that for focal band. This is not surprising, since for an existing brand, it is more challenging to keep customers engaged, since excitement is harder to generate in the absence of the "newness" factor.

To facilitate a more intuitive understanding of a band's actions in social media on state transitions, based on the estimates in Tables 8 and 9, we compute the probabilities of transitioning across states given a single tweet of a specific type. These transition probabilities are shown in Table 10. Note that these transition matrices are derived from the parameter estimates that we discussed above. Therefore, they do not present new interpretations and new insights per se. Instead of repeating the detailed interpretation of the parameters, we discuss here a few salient observations in a more substantive way. (Also, some patterns are more visible here since an individual parameter tells only part of the story. For example, the coefficient for an emotional but non-informational tweet increases the probability of a focal band's follower transitioning from somewhat engaged to fully engaged state, but it only slightly decreases the probability of transitioning to unengaged state. Thus, the overall effect is clearer in the calculated transition matrix.)

First, for both focal bands and control bands, it is quite difficult to move followers out of an unengaged state, i.e. to get them engaged to a certain level. The probability of staying in the S_L state is very high for both types of band, with or without tweets. For focal bands, a tweet in general has approximately a 10% chance of getting a follower engaged, with non-informational tweets doing slightly better. In contrast, for control bands only informational and non-emotional tweets have any discernible effect. An informational and non-emotional tweet does have a strong effect for the control band though, as the change of getting a follower engaged is almost 35%. This suggests that followers of control bands, since they have presumably been around longer, are not easily persuaded by emotional arguments any more, and need a strong reason based on information alone, to cause any change in behavior.

Secondly, for focal bands, the somewhat engaged state is much less stable than for control bands: i.e., the probability of remaining in that same state is much lower for followers of focal bands, with or without tweets. This suggests that the somewhat engaged state is a transient one for focal bands, but not so for control bands: followers of control bands may already know the bands well enough to decide that they may get engaged but not too engaged, while followers of focal bands are likely still in the burn-in process where moderate level of engagement is of temporary nature. Consistent with this, we also note that the fully engaged state is more stable for focal bands than for control bands, suggesting that followers of focal bands may prefer a more extreme level of engagement while those of control bands are harder to stay highly engaged for long.

Finally, informational but non-emotional tweets seem to be more effective for control bands in general. For control bands, informational tweets are both more effective in transitioning followers out of the unengaged state, i.e. getting them engaged to a certain extent, and more effective in keeping them in the fully engaged state. For focal bands, somewhat interestingly, tweets that are both emotional and informational, and tweets that are neither emotional nor informational, seem to be more effective than those that are either emotional or informational. The former two types of tweets both are more likely to transition followers into fully engaged state and more likely to keep them there. The difference is on the margin, as the transition probabilities show that tweets in general are helpful for focal bands.

In summary, the impact of a firm's messages varies depending on the state of engagement of followers, as well as depending on whether the firm is a new one or an established one. Next, we take a look at how a firm's message, the consequent follower state, and information and emotion content relate to observed follower behavior.

The nature of follower actions

In this section, we report on the link between estimated follower engagement, as well as their informational and emotional knowledge level, on their observed actions, i.e., the volume, emotion and information of tweets sent by followers. When followers tweet about a band, this action reflect the follower engagement as well as interest in the band. Since followers are potential consumers, it is important for a firm to understand what their actions imply about their underlying state. In addition, in social media, followers' messages reach out to their followers, thus constituting advertising on behalf of the band to a wider audience which might perhaps not have been reached by the band's tweet. Such action could also lead to tweets about the band from the followers of followers, and beyond, thus further amplifying the reach of the band. Consequently, understanding this link is of strong interest to researchers and managers alike.

We first report on the retention rates of information and emotion knowledge for followers in Table 11. From this analysis, we make the following observations. First, overall, retention of information and emotion knowledge is higher for control bands than for focal bands. This could be because of the long term nature of relationship with the control bands than the focal bands. Further, while followers of control bands are more likely to retain information stock than emotion stock, in contrast, followers of focal bands are more likely to retain emotion stock than information stock. The managerial implications of these observed decay rates are that while knowledge in general is "stickier" in established relationships than in nascent ones, emotion is relatively "stickier" within nascent relationships, whereas information is relatively "stickier" within established ones.

We then report on follower actions. Our results of this analysis are shown in Table 12. From this table, first note that the intrinsic propensity to tweet differs significantly across states. For followers in the unengaged state for both focal bands and control bands, the propensity to tweet is so low as to be nonexistent. For somewhat engaged followers, the propensity to tweet is still quite low. For fully engaged followers, tweeting is highly likely, averaging one every 2-3 periods for followers of focal bands and one every 1-2 periods for followers of control bands. Across all three states, followers of control bands are more likely to tweet than followers of focal bands. On the first look, this may be surprising since the descriptive statistics of the data show that focal bands have more follower tweets than control bands. To understand this apparent contradiction, recall that we have shown that fully engaged followers of focal bands are more likely to remain fully engaged than their counterparts in control bands. Furthermore, focal bands tweet more than control bands do, and would thus likely have more of their followers in fully engaged states than control bands do. Therefore, although in any given state a follower of a control band is more likely to tweet, followers of focal bands have more tweets overall, since more of them are in the fully engaged state than followers of control bands are. Such in-depth and nuanced understanding of the tweeting behavior is enabled by our modeling of both different levels of engagement states and the different behavior in each state.

We next look at the effect of information and emotion stocks on the propensity to tweet.

Since the intrinsic propensity to tweet is extremely low for unengaged state, we focus more on the somewhat engaged and fully engaged state. We see that in general, higher information and emotion stocks make a follower more likely to tweet, although the effect is not always statistically significant (it is for both stocks for the fully engaged state for focal bands, and for emotion stock for the somewhat engaged state for control bands).

We next turn to the type of tweets sent by followers. Again we focus on the somewhat engaged and fully engaged states. We note that by default, follower tweets are equally likely to be emotional or non-emotional (only the coefficient for somewhat engaged state for control bands is positive and statistically significant). In contrast, follower tweets are more likely to be non-informational than informational (all intercept coefficients are negative, and with the exception of the somewhat engaged state for focal bands, statistically significant). This is understandable, as followers need to be informed before they can be informational. Some interesting observations can be made regarding the impact of information and emotion stock on the type of follower tweets. First, more information stock makes a follower's tweet more likely to be informational (the coefficients for information are positive for both states and both band types, and statistically significant for three of them), whereas more emotion makes a follower's tweet less likely to be informational (three of the corresponding coefficients are negative and two statistically significant). This suggests that followers are more likely to tweet informationally when they have more information, and when they are less emotional, an understandable but notable finding. Furthermore, for fully engaged followers of focal band, higher information stock makes the follower tweet less likely to be emotional, while higher emotion stock make it more likely so, a finding similar to that for sending informational tweets. Interestingly, however, the same is not true for fully engaged followers of control bands. For these followers, higher informational stock makes the tweet **more** likely to be emotional, while higher emotional stock does not have such impact. This suggests that followers of control bands likely have pre-established emotion towards the band, which is triggered by new information about the band, but is not affected by the emotion embodied in tweets from the band.

The relation between consumer engagement and sales

Finally we analyze whether the estimated follower engagement states and follower knowledge (i.e., emotional and informational stock) impact final market sales for the focal as well as established brands. To do so, we estimate the generalized diffusion model specified earlier. The model is estimated using MCMC, the detail of which is specified in the appendix. Most of the album sales happen in the early weeks after release. For our estimation, we used the first 12 weeks of sales of each album (we also estimated the model using 16, 20, and 30 weeks, and find that the estimates are not sensitive to this choice). The estimated results are reported in Table 13. From this table, we make the following observations.

First, note that the estimate for the average coefficient of innovation across all albums for the focal bands is 0.058 ($e^{-2.781}/(1+e^{-2.781})$), and for the control bands is 0.106; and the corresponding average coefficient of imitation is 0.0002 for focal bands and essentially 0 for control bands. This primarily indicates that album sales are exponential for both types of bands. Further, they take place at a much faster rate for the more established, control bands, and at a slower pace for the newer, focal bands.

The impact of follower engagement on sales follows an interesting pattern. First, note that for both types of bands, the coefficient for the proportion of fully engaged followers is positive with a large magnitude and is statistically significantly different from 0. This shows that the more fully engaged followers that a band has at a point in time, the higher the sales. However, for focal bands the coefficient for the proportion of somewhat engaged followers is negative, while for control band that coefficient is positive and statistically significant. This contrast is interesting. Recall that when we discuss the estimates of the tweet and engagement model, we find that the somewhat engaged state is somewhat transient for followers of focal bands, but more stable and final for followers of control bands. Their relationship with sales here further validates that observation. For followers of focal bands, the somewhat engaged state indicates that they are trying to make up their mind on the band, and are in no mood to purchase, yet. But for followers of control bands, the somewhat engaged state indicates moderate level of interest, learned over time since the bands have been around, and thus it has positive albeit small impact on sales. Again, such nuanced understanding is enabled by our modeling of the different levels of underlying engagement for consumers and their transition activities.

The coefficients for information stock, emotion stock, and offline concert play are all positive and statistically significant. A notable exception, however, is the coefficient for the information stock for control bands is negative. A possible explanation might be that somehow the content of the information is not recommending purchase. However, this point should be investigated further in future studies hopeful with the availability of more data. The other coefficients are all positive, confirming the positive impact of the corresponding factors. That the offline concert has positive impact on sales is reassuring, since we expect promotion activities to help sales in the first place. The positive effect of information and emotion stock on sales further confirm the impact of customer engagement on social media.

In terms of the dynamic market potential, we observe that both number of albums and followers has a positive impact. Albums have a relatively stronger impact on market potential for control bands, whereas the follower base has a relatively stronger impact for new bands. This result suggests that firms indeed are likely to benefit from attempts at building follower bases within social media, as such bases are associated with a positive impact on sales.

In summary, relating the engagement states and knowledge and information stock that are recovered using our tweet model to actual album sales serves as a strong validation to our tweet model. That these engagement levels and knowledge and informations stocks are generally positively related to sales shows that engaging in consumers in social media has significant implications on firm's sales. It is therefore crucial for managers to understand the nuances of such engagement on social media and to devise marketing strategies accordingly. As one of the earliest studies in that area, our research contributes to this understanding.

Conclusions

In this research, we have analyzed the link between firm actions, consumer engagement, and sales in digital social media. Our key contributions in this work are as follows. First, we identify the effect of firm actions in microblogs on follower engagement within these microblogs. Second, we show that follower engagement in microblogs is positively linked to firm sales. Third, we show that these effects vary for new brands vis-a-vis existing brands. For each effect, we further identify the differential roles of emotional and informational message content on generating, growing and sustaining consumer engagement, their consequent actions in microblogs, and sales.

By deriving these results, we provide empirical support for the hypothesis that firm actions in social media can generate interest and changes in engagement levels of their followers, and that such engagement levels can have a positive bottom line impact, on sales. Demonstrating these effects is of interest not only to researchers, but also to managers struggling with the question of how best to manage and drive their firms' social media efforts. For instance, we do observe that overall it is indeed challenging for firms' messages in microblogs to cause followers that are in a low state of engagement to start becoming engaged, both for new as well as established brands. However, certain types of messages are more likely to be more effective than others. Our results indicate that for new brands, sending messages, and in particular, sending messages that might be light on informational content but heavier on emotional content tend to be relatively more effective. In contrast, we find that for established brands the messages that are likely to have the most effect are ones that are high on informational content, and low on emotion. A direct implication for managers is to experiment with their tweet content appropriately, by taking into account the relative strength of their brand when communicating with specific follower segments in digital social media.

Our results also suggest that it is important to take into account the possibility that follower engagement states might be relatively more transient for new brands, when anticipating the impact of digital social campaigns. Our observation - that even moderate engagement on the part of followers of established brands relates positively to brand sales, but that this is not the case for new brands - suggests that new brands might have to first engage substantially with their online followers using specific message content, to generate changes in engagement levels that can translate to sales. An implication of this finding is that expecting a direct immediate impact of a digital campaign might hence, in some sense, be more challenging for new brands, than existing ones.

While our research begins to shed some light on how firms can engage with consumers in social media while having an impact on their bottom line, it also points to several interesting areas for future research. The first set of challenges we faced during this work were associated with processing and handling the data available via this media. Data challenges by themselves present many interesting avenues for further research. While in this work, we relied on extracting binary measures for information and emotion for individual messages, future research can investigate defining multi-dimensional continuous measures that will allow us to understand unstructured text message content at a far more detailed level than is currently feasible. While we considered a set of new and established brands within the music industry, it would be interesting to further explore how firm actions impact consumer behavior and sales across a host of other industries and product categories. Finally, by collecting more data at various stages of the consumption process such as awareness levels, consideration sets, as well as repeat behavior, future research can investigate better the linkage between engagement levels and sales. Our work, as well as such potential future investigations would hopefully enable managers to take more evidence based actions for making their firm investments in social media more effective in generating value.

Appendix: Model Estimation Procedure

We estimate our model using a Markov Chain Monte Carlo (MCMC) procedure, by taking conditional draws of parameters as specified below and iterating until convergence is achieved. When a closed-form expression for the posterior does not exist, we take draws using a Metropolis-Hastings procedure with a random walk. Our estimation procedure utilizes data augmentation, where we take explicit draws for the state of each consumer in each time period. We use $g(\cdot)$ to denote the *p.m.f.* of the Poisson distribution and use $f(\cdot)$ to denote a generic *p.d.f.*.

The Tweets Model

- Draw $\left\{\lambda_{i,0,s}\right\}_{s=1,...,S}$ for each user i, $i = 1,...,I$:

  $f(\{\lambda_{i,0,s}\}_{s=1,..,S} | \left\{\overline{\lambda}_{0,s}, \sigma^2_{\lambda,s}\right\}_{s=1,..,S}, \{\beta_{I,s}, \beta_{E,s}\}_{s=1,..,S}, \left\{\gamma^{\mathrm{E}}_{0,\mathrm{s}}, \gamma^{\mathrm{E}}_{\mathrm{I,s}}, \gamma^{\mathrm{E}}_{\mathrm{E,s}},\right\}_{\mathrm{s=1,..,S}}, \left\{\gamma^{\mathrm{I}}_{0,\mathrm{s}}, \gamma^{\mathrm{I}}_{\mathrm{I,s}}, \gamma^{\mathrm{I}}_{\mathrm{E,s}},\right\}_{\mathrm{s=1,..,S}}, \delta^{\mathrm{I}}_{\mathrm{i}}, \delta^{\mathrm{E}}_{\mathrm{i}}) \propto$
  $\left(\prod_{t=1}^{T} \tilde{G}\right) \left(\prod_{s=1}^{S} \varphi\left(log(\lambda_{i,0,s}); \overline{\lambda}_{0,s}, \sigma^2_{\lambda,s}\right)\right)$,
  where $\tilde{G} = g\left(C_{it}^{NE,NI}; \lambda_{it} \cdot (1 - p_{E,it}) \cdot (1 - p_{I,it})\right) g\left(C_{it}^{E,NI}; \lambda_{it} \cdot p_{E,it} \cdot (1 - p_{I,it})\right)$
  $g\left(C_{it}^{NE,I}; \lambda_{it} \cdot (1 - p_{E,it}) \cdot p_{I,it}\right) g\left(C_{it}^{NE,NI}; \lambda_{it} \cdot p_{E,it} \cdot p_{I,it}\right)$
  In the equation, $\varphi(\cdot)$ is the *p.d.f.* of the Normal distribution. λ_{it}, $p_{E,it}$, $p_{I,it}$ are as specified in Equations (6)-(8) in the main text. States in HMM models are invariant to permutation. To ensure identification, we impose the restriction $\lambda_{i,0,s} < \lambda_{i,0,s+1}$for each user i, $i = 1,...,I$, and each state s, $s = 1,...,S-1$. That is, states indexed by higher numbers have higher intrinsic propensities to tweet and correspond to higher engagement levels.

- Draw $\overline{\lambda}_{0,s}$ for each state s, $s = 1,...,S$:

  $log(\overline{\lambda}_{0,s}) | \left\{\lambda_{i,0,s}\right\}_{i=1,...,I}, \sigma^2_{\lambda,s} \sim$
  $$\text{Normal}\left(\left(\frac{1}{\sigma_0^2} + \frac{I}{\sigma^2_{\lambda,s}}\right)^{-1}\left(\frac{\mu_0}{\sigma_0^2} + \frac{\sum_{i=1}^{I} log(\lambda_{i,0,s})}{\sigma^2_{\lambda,s}}\right), \left(\frac{1}{\sigma_0^2} + \frac{I}{\sigma^2_{\lambda,s}}\right)^{-1}\right)$$
  We choose a diffuse conjugate Normal prior for$\overline{\lambda}_{0,s}$.

- Draw $\sigma^2_{\lambda,s}$ for each state s, $s = 1,...,S$:

  $$\sigma^2_{\lambda,s} | \left\{\lambda_{i,0,s}\right\}_{i=1,...,I}, \overline{\lambda}_{0,s} \sim \text{Inv-Gamma}\left(v_0 + \frac{I}{2}, s_0 + \sum_{i=1}^{I}\left(log(\lambda_{i,0,s}) - log(\overline{\lambda}_{0,s})\right)^2\right)$$
  We choose a diffuse conjugate Inverse-Gamma prior for $\sigma^2_{\lambda,s}$.

- Draw $\{\beta_{I,s}, \beta_{E,s}\}_{s=1,...,S}$:

$f(\{\beta_{I,s},\beta_{E,s}\}_{s=1,\dots,S} \,|\, \{\lambda_{i,0,s}\}_{s=1,\dots,S}, \left\{\gamma^{\mathrm{E}}_{0,\mathrm{s}},\gamma^{\mathrm{E}}_{\mathrm{I,s}},\gamma^{\mathrm{E}}_{\mathrm{E,s}},\right\}_{\mathrm{s=1,\dots,S}}, \left\{\gamma^{\mathrm{I}}_{0,\mathrm{s}},\gamma^{\mathrm{I}}_{\mathrm{I,s}},\gamma^{\mathrm{I}}_{\mathrm{E,s}},\right\}_{\mathrm{s=1,\dots,S}}, \delta^{\mathrm{I}}_{\mathrm{i}}, \delta^{\mathrm{E}}_{\mathrm{i}}) \propto$
$\left(\prod_{t=1}^{T}\tilde{G}\right)\left(\prod_{s=1}^{S} f(\beta_{I,s})f(\beta_{E,s})\right)$
In this equation, λ_{it}, $p_{E,it}$, $p_{I,it}$ are as specified in equations (6)-(8) in the main text. We use a diffuse prior for each $\beta_{I,s}$ and $\beta_{E,s}$.

- Draw $\left\{\gamma^{E}_{0,s},\gamma^{E}_{I,s},\gamma^{E}_{E,s},\right\}_{s=1,\dots,S}$:
  $f(\left\{\gamma^{E}_{0,s},\gamma^{E}_{I,s},\gamma^{E}_{E,s},\right\}_{s=1,\dots,S} \,|\, \{\lambda_{\mathrm{i},0,\mathrm{s}}\}_{\mathrm{s=1,\dots,S}}, \{\beta_{\mathrm{I,s}},\beta_{\mathrm{E,s}}\}_{\mathrm{s=1,\dots,S}}, \left\{\gamma^{\mathrm{I}}_{0,\mathrm{s}},\gamma^{\mathrm{I}}_{\mathrm{I,s}},\gamma^{\mathrm{I}}_{\mathrm{E,s}},\right\}_{\mathrm{s=1,\dots,S}}, \delta^{\mathrm{I}}_{\mathrm{i}}, \delta^{\mathrm{E}}_{\mathrm{i}}) \propto$
  $\left(\prod_{t=1}^{T}\tilde{G}\right)\left(\prod_{s=1}^{S} f(\gamma^{E}_{0,s})f(\gamma^{E}_{I,s})f(\gamma^{E}_{E,s})\right)$
  In this equation, λ_{it}, $p_{E,it}$, $p_{I,it}$ are as specified in equations (6)-(8) in the main text. We use a diffuse prior for each $\gamma^{E}_{0,s}$, $\gamma^{E}_{I,s}$ and $\gamma^{E}_{E,s}$.

- Draw $\left\{\gamma^{I}_{0,s},\gamma^{I}_{I,s},\gamma^{I}_{E,s},\right\}_{s=1,\dots,S}$:
  $f(\left\{\gamma^{I}_{0,s},\gamma^{I}_{I,s},\gamma^{I}_{E,s},\right\}_{s=1,\dots,S} \,|\, \{\lambda_{i,0,s}\}_{s=1,\dots,S}, \{\beta_{\mathrm{I,s}},\beta_{\mathrm{E,s}}\}_{\mathrm{s=1,\dots,S}}, \left\{\gamma^{\mathrm{E}}_{0,\mathrm{s}},\gamma^{\mathrm{E}}_{\mathrm{I,s}},\gamma^{\mathrm{E}}_{\mathrm{E,s}},\right\}_{\mathrm{s=1,\dots,S}}, \delta^{\mathrm{I}}_{\mathrm{i}}, \delta^{\mathrm{E}}_{\mathrm{i}}) \propto$
  $\left(\prod_{t=1}^{T}\tilde{G}\right)\left(\prod_{s=1}^{S} f(\gamma^{I}_{0,s})f(\gamma^{I}_{I,s})f(\gamma^{I}_{E,s})\right)$
  In this equation, λ_{it}, $p_{E,it}$, $p_{I,it}$ are as specified in equations (6)-(8) in the main text. We use a diffuse prior for each $\gamma^{I}_{0,s}$, $\gamma^{I}_{I,s}$ and $\gamma^{I}_{E,s}$.

- Draw $\{\delta^{I}_{i},\delta^{E}_{i}\}$ for each user i, $i=1,\dots,I$:
  $f(\delta^{I}_{i},\delta^{E}_{i} \,|\, \{\lambda_{i,0,s}\}_{s=1,\dots,S}, \{\beta_{\mathrm{I,s}},\beta_{\mathrm{E,s}}\}_{\mathrm{s=1,\dots,S}}, \left\{\gamma^{\mathrm{E}}_{0,\mathrm{s}},\gamma^{\mathrm{E}}_{\mathrm{I,s}},\gamma^{\mathrm{E}}_{\mathrm{E,s}},\right\}_{\mathrm{s=1,\dots,S}}, \left\{\gamma^{\mathrm{I}}_{0,\mathrm{s}},\gamma^{\mathrm{I}}_{\mathrm{I,s}},\gamma^{\mathrm{I}}_{\mathrm{E,s}},\right\}_{\mathrm{s=1,\dots,S}}) \propto$
  $\left(\prod_{t=1}^{T}\tilde{G}\right)\varphi\left(logit\left(\delta^{I}_{i}\right); logit\left(\overline{\delta^{I}}\right), \sigma^{2}_{\delta^{I}}\right)\varphi\left(logit\left(\delta^{E}_{i}\right); logit\left(\overline{\delta^{E}}\right), \sigma^{2}_{\delta^{E}}\right)$
  In this equation, $\varphi(\cdot)$ is the *p.d.f.* of the Normal distribution. λ_{it}, $p_{E,it}$, $p_{I,it}$ are as specified in equations (6)-(8) in the main text, and δ^{I}_{i} and δ^{E}_{i} enter the equations through $I_{i,t}$ and $E_{i,t}$ as defined in equations (11) and (12).

- Draw $\overline{\delta^{I}}$:
  $logit\left(\overline{\delta^{I}}\right) | \{\delta^{I}_{i}\}_{i=1,\dots,I}, \sigma^{2}_{\delta^{I}} \sim \text{Normal}\left(\left(\frac{1}{\sigma^{2}_{0}}+\frac{I}{\sigma^{2}_{\delta^{I}}}\right)^{-1}\left(\frac{\mu_0}{\sigma^{2}_{0}}+\frac{\sum_{i=1}^{I} logit\left(\delta^{I}_{i}\right)}{\sigma^{2}_{\delta^{I}}}\right), \left(\frac{1}{\sigma^{2}_{0}}+\frac{I}{\sigma^{2}_{\delta^{I}}}\right)^{-1}\right)$
  We choose a diffuse conjugate Normal prior for $\overline{\delta^{I}}$.

- Draw $\overline{\delta^{E}}$:
  $logit\left(\overline{\delta^{E}}\right) | \{\delta^{E}_{i}\}_{i=1,\dots,I}, \sigma^{2}_{\delta^{E}} \sim \text{Normal}\left(\left(\frac{1}{\sigma^{2}_{0}}+\frac{I}{\sigma^{2}_{\delta^{E}}}\right)^{-1}\left(\frac{\mu_0}{\sigma^{2}_{0}}+\frac{\sum_{i=1}^{I} logit\left(\delta^{E}_{i}\right)}{\sigma^{2}_{\delta^{E}}}\right), \left(\frac{1}{\sigma^{2}_{0}}+\frac{I}{\sigma^{2}_{\delta^{E}}}\right)^{-1}\right)$
  We choose a diffuse conjugate Normal prior for $\overline{\delta^{E}}$.

- Draw $\sigma^{2}_{\delta^{I}}$:

$\sigma^2_{\delta^I} | \left\{\delta^I_i\right\}_{i=1,...,I}, \overline{\delta^I} \sim$ Inv-Gamma $\left(v_0 + \frac{I}{2}, s_0 + \sum_{i=1}^{I} \left(logit\left(\delta^I_i\right) - logit\left(\overline{\delta^I}\right)\right)^2 \right)$
We choose a diffuse conjugate Inverse-Gamma prior for $\sigma^2_{\delta^I}$.

- Draw $\sigma^2_{\delta^E}$:
  $\sigma^2_{\delta^E} | \left\{\delta^E_i\right\}_{i=1,...,I}, \overline{\delta^E} \sim$ Inv-Gamma $\left(v_0 + \frac{I}{2}, s_0 + \sum_{i=1}^{I} \left(logit\left(\delta^E_i\right) - logit\left(\overline{\delta^E}\right)\right)^2 \right)$
  We choose a diffuse conjugate Inverse-Gamma prior for $\sigma^2_{\delta^I}$.

- Draw $\left\{\vec{\alpha}_{st} = \left(\alpha_{st,0}, \alpha_{st,1}, \alpha_{st,2}, \alpha_{st,3}, \alpha_{st,4}\right)\right\}_{t=1,...,S,t\neq s}$ for each state s, $s = 1,...,S$:
  $f(\left\{\vec{\alpha}_{st} = \left(\alpha_{st,0}, \alpha_{st,1}, \alpha_{st,2}, \alpha_{st,3}, \alpha_{st,4}\right)\right\}_{t=1,...,S,t\neq s} | \{\mathrm{s_{it}}\}_{\mathrm{i=1,...,I,t=1,...,T}}) \propto$
  $\left(\prod_{i=1}^{I}\prod_{t=1}^{T-1}\left(a_{it,s_{it}s_{i,t+1}}\right)^{\mathbb{I}\{s_{it}=s\}}\right)\left(\prod_{t=1,t\neq s}^{S} f\left(\vec{\alpha}_{st}\right)\right)$
  In this equation, $a_{it,s_{it}s_{i,t+1}}$ is as specified in equations (9) and (10) in the main text. $\mathbb{I}\{\cdot\}$is the indicator function. We use a diffuse prior for each $\vec{\alpha}_{st}$.

- Draw $s_{i,t}$ for each user i, $i = 1,...,I$ at each time period t, $t = 1,...,T$: This is the data augmentation step of the estimation. We draw the state of each customer at each time using the forward-backward algorithm developed in Chib (1996): for each i, we first draw the state at the final time period based on the data likelihood, then draw the state in the previous time period, until the first, based on the data likelihood and the subsequent state which is drawn in the earlier step.

The Album Sales Model

We follow the approach developed in Moe and Fader (2002) to estimate the album sales model. Specifically, we derive the probability of a purchase happening in each specific time period, conditional on the sale happening in the first T time periods. This translates the sales of an album into draws from a multinomial distribution (per Equation (7) in Moe and Fader (2002)):
$R_{a,t} \sim \text{Multinomial}(p_a(1), p_a(2), ..., p_a(T); \sum_{t=1}^{T} R_{a,t})$
where $R_{a,t}$ is the sales of album a at time t, and the probability of a sale happening in a time period, $p_a(t)$, is derived from the diffusion model:
$p_a(t) = f_a(t) / \sum_{\tau=1}^{T} f_a(\tau)$
In the equations, $f_a(t)$ is as specified in equation (17) of the main text. This leads to the likelihood function:
$L\left(\{R_{a,t}\}_{t=1,...,T} | \Phi\right) = \left(\prod_{t=1}^{T} p_a(t)^{R_{a,t}}\right) \varphi\left(log\left(M_a\right) - log\left(\sum_{t=1}^{T} R_{a,t} / \sum_{t=1}^{T} f_a(t)\right); 0, \sigma^2_m\right)$
In the equation, $\varphi(\cdot)$ is the *p.d.f.* of Normal distribution, andσ^2_m is a nuisance parameter that measures the variance of album market potentials. To simplify notation, we use Φ to denote all parameters of the album sales model where it does not cause confusion to do so.
Our conditional draws for the album sales model are as follows:

- Draw p_a and q_a for each album a, $a = 1,...,A$:

  $f(p_a, q_a|\Phi) \propto L\left(\{R_{a,t}\}_{t=1,...,T}|\Phi\right) \varphi\left(logit(p_a); logit(\bar{p}), \sigma_p^2\right) \varphi\left(log\left(q_q\right); log(\bar{q}), \sigma_q^2\right)$
  In this equation, $\varphi(\cdot)$ is the *p.d.f.* of Normal distribution.

- Draw $\bar{p}$ and $\bar{q}$:

  $logit(\bar{p})\,|\,\{p_a\}_{a=1,...,A}, \sigma_p^2 \sim \text{Normal}\left(\left(\frac{1}{\sigma_0^2}+\frac{A}{\sigma_p^2}\right)^{-1}\left(\frac{\mu_0}{\sigma_0^2}+\frac{\sum_{a=1}^{A} logit(p_a)}{\sigma_p^2}\right), \left(\frac{1}{\sigma_0^2}+\frac{A}{\sigma_p^2}\right)^{-1}\right)$

  $log(\bar{q})\,|\,\{q_a\}_{a=1,...,A}, \sigma_q^2 \sim \text{Normal}\left(\left(\frac{1}{\sigma_0^2}+\frac{A}{\sigma_q^2}\right)^{-1}\left(\frac{\mu_0}{\sigma_0^2}+\frac{\sum_{a=1}^{A} log(q_a)}{\sigma_q^2}\right), \left(\frac{1}{\sigma_0^2}+\frac{A}{\sigma_q^2}\right)^{-1}\right)$
  We choose a diffuse conjugate Normal prior for $\bar{p}$ and $\bar{q}$.

- Draw σ_p^2 and σ_q^2:

  $\sigma_p^2|\,\{p_a\}_{a=1,...,A}, \bar{p} \sim \text{Inv-Gamma}\left(\nu_0+\frac{A}{2}, s_0+\sum_{a=1}^{A}\left(logit(p_a)-logit(\bar{p})\right)^2\right)$

  $\sigma_q^2|\,\{q_a\}_{a=1,...,A}, \bar{q} \sim \text{Inv-Gamma}\left(\nu_0+\frac{A}{2}, s_0+\sum_{a=1}^{A}\left(log(q_a)-log(\bar{q})\right)^2\right)$
  We choose a diffuse conjugate Inverse-Gamma prior for σ_p^2 and σ_q^2.

- Draw $\vec{\beta}$:

  $f(\vec{\beta}|\Phi) \propto L\left(\{R_{a,t}\}_{t=1,...,T}|\Phi\right) f\left(\vec{\beta}\right)$. We use a diffuse prior for $\vec{\beta}$.

- Draw γ_0 and$\vec{\gamma}$:

  $f(\gamma_0, \vec{\gamma}|\Phi) \propto L\left(\{R_{a,t}\}_{t=1,...,T}|\Phi\right) f(\gamma_0, \vec{\gamma})$. We use a diffuse prior for γ_0 and$\vec{\gamma}$.

References

Bagwell, Kyle (2007), "The Economic Analysis of Advertising," *The Handbook of Industrial Organization*, Chapter 28. Eds. M. Armstrong and R. Porter, Elsevier San Deigo, CA, USA, 1701–1844.

Bass, Frank M (1969), "A new product growth model for consumer durables," *Management Science* 15 215–227.

—, Trichy V. Krishnan, and Dipak C. Jain (1994), "Why the Bass model fits without decision variables," *Marketing Science* 13 119–130.

—, Dipak C. Jain, and Trichy V. Krishnan (2000), "Modeling the marketing- mix influence in new product diffusion," V. Mahajan, E. Muller, Y. Wind, eds. *New Product Diffusion Models*. Kluwer Academic Publishers, Norwell, MA, 123–140.

Batra, Rajeev, John G. Myers, and David A. Aaker (1996), "Advertising Management," Prentice Hall, Upper Saddle River, NJ.

Breiman, Leo (2001), "Random forests," *Machine Learning* 45.1 5-32.

Brown, Brad, Johnson Sikes. 2012. Minding Your Digital Business: McKinsey Global Survey Results. McKinsey and Company Research Report.

Celikyilmaz, Asli, Dilek Hakkani-Tür, and Junlan Feng (2010), "Probabilistic Model-Based Sentiment Analysis of Twitter Messages," *IEEE Workshop on Spoken Language Technologies* 79-84.

Chevalier, Judith A. and Dina Mayzlin (2006), "The Effect of Word of Mouth on Sales: Online Book Reviews," *Journal of Marketing Research* 43 (3) 345-354.

Chib, Siddhartha (1996), "Calculating posterior distributions and modal estimates in Markov mixture models," *Journal of Econometrics* 75 79-97.

Divol, Roxane, David Edelman, Hugo Sarrazin. 2012. Demystifying Social Media. McKinsey Quarterly, April.

Eliashberg, Jehoshua and Gary L. Lilien (1993), "Marketing: Handbooks in Operations Research and Management Science," Volume 5. Elsevier Science Publishers B.V. The Netherlands.

Erdem, Tulin and Michael P. Keane (1996), "Decision-Making under Uncertainty: Capturing Dynamic Brand Choice Processes in Turbulent Consumer Goods Markets," *Marketing Science* 15(1) 1-20.

Go, Alec, Richa Bhayani, and Lei Huang (2009), "Twitter Sentiment Classification Using Distant Supervision," *CS224N Project Report*, Stanford University.

Godes, David, Dina Mayzlin, Yubo Chen, Sanjiv Das, Chrysanthos Dellarocas, Bruce Pfeiffer, Barak Libai, Subrata Sen, Mengze Shi, and Peeter Verlegh (2005), “The Firm’s Management of Social Interactions,” *Marketing Letters* 16(3/4) 415-28.

— and Dina Mayzlin (2009), “Firm-Created Word-of-Mouth Communication: Evidence from a Field Test,” *Marketing Science* 28(4) 721-739.

Gruhl, Daniel, R. Guha, David Liben-Nowell, and Andrew Tomkins (2004), “Information Diffusion Through Blogspace,” *Proceedings of the 13th international conference on World Wide Web (WWW ’04)*. ACM, New York, NY, 491-501.

Guadagni, Peter M. and John D.C. Little (1983), “A Logit Model of Brand Choice Calibrated on Scanner Data,” *Markting Science* 2(3) 203-238.

Gupta, Sunil, Kristen Armstrong, Zach Clayton. 2010. Social Media. Harvard Business School Publishing, Boston MA, Item 9-510-095.

Hennig-Thurau, Thorsten, Edward C. Malthouse, Christian Friege, Sonja Gensler, Lara Lobschat, Arvind Rangaswamy, and Bernd Skiera (2010), “The Impact of New Media on Customer Relationships,” *Journal of Service Research* 13 311-330.

Huberman, Bernardo A., Daniel M. Romero, and Fang Wu (2009), “Social networks that matter: Twitter under the microscope,” *First Monday*, 14(1).

Iyengar, Raghuram, Sangman Han, and Sunil Gupta (2009), “Do Friends Influence Purchases in a Social Network?” *Harvard Business School Marketing Unit Working Paper* No. 09-123.

—, Christophe Van den Bulte, and Thomas W. Valente (2011), “Opinion Leadership and Social Contagion in New Product Diffusion,” *Marketing Science* 30(2) 195-212.

Katona, Zsolt, Peter P. Zubcsek, and Miklos Sarvary (2011), “Network Effects and Personal Influences: Diffusion of an Online Social Network,” *Journal of Marketing Research* 48(3) 425-443.

Koulumpis, Efthymios, Theresa Wilson, and Johanna Moore (2011), “Twitter Sentiment Analysis: The Good, The Bad and the OMG,” *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media.*

Leskovec, Jure, Lars Backstrom, and Jon Kleinberg (2009) “Meme-tracking and the dynamics of the news cycle,” *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD ’09)* ACM, New York, NY, 497-506.

Li, Shibo, Baohong Sun, and Alan L. Montgomery (2011), “Cross-Selling the Right Product to the Right Customer at the Right Time,” *Journal of Marketing Research* 48(4) 683-700.

Lipsman, Andrew, Graham Mudd, Mike Rich, and Sean Bruich (2012), "The Power of Like: How Brands Reach and Influence Fans Through Social Media Marketing," *Journal of Advertising Research*, 52(1) 40.

Machedon, Radu, William M. Rand, and Yogesh V. Joshi (2013), "Automatic Classification of Social Media Messaging using Multi-Dimensional Sentiment Analysis and Crowdsourcing," Working Paper.

Mayzlin, Dina and Hema Yoganarasimhan (2012), "Link to Success: How Blogs Build an Audience by Promoting Rivals," *Management Science* 58(9) 1651–1668.

Moe, Wendy W. and Peter S. Fader (2002), "Using Advance Purchase Orders to Forecast New Product Sales," *Marketing Science* 21(3), 347-364

— and Michael Trusov (2011), "The Value of Social Dynamics in Online Product Ratings Forums," *Journal of Marketing Research* 48(3) 444-456.

Monner, Derek, William M. Rand, and Yogesh V. Joshi (2013), "Prediction With Twitter," Working Paper.

Netzer, Oded, James M. Lattin, and V. Srinivasan (2008), "A Hidden Markov Model of Customer Relationship Dynamics," *Marketing Science* 27(2) 185-204.

Newton, Michael A. and Adrian E. Raftery (1994), "Approximate Bayesian Inference by the Weighted Likelihood Bootstrap (with discussion)," *Journal of the Royal Statistical Society*, Ser. B 56 3-48.

Oh, Jeongha, Anjana Susarla, and Yong Tan (2008), "Examining the Diffusion of User-Generated Content in Online Social Networks," Working paper.

Onishi, Hiroshi and Puneet Manchanda (2010), "Marketing Activity, Blogging and Sales," *Marketing Science Institute Working Paper Series* Report No. 10-113.

Roberts, John H. and Gary L. Lilien (1993), "Explanatory and Predictive Models of Consumer Behavior," Chapter 2, *Marketing*, Eds. J. Eliashberg, G. L. Lilien. Elsevier Science Publishers B.V., The Netherlands.

Schau, Hope J., Albert M. Muniz Jr., and Eric J. Arnould (2009), "How Brand Community Practices Create Value," *Journal of Marketing* 73(Sep) 30-51.

Schweidel, David A. and Wendy M. Moe (2012) "The Perils of 'Selective Listening' in Social Media Monitoring: Sentiment and Venue Choice in Social Media Posting Behavior," Working paper, available at SSRN: http://ssrn.com/abstract=1874892.

Shi, Wei, Michel Wedel, and Rik Pieters (2013), "Information Acquisition during Online Decision-Making: A Model-Based Exploration Using Eye-Tracking Data," *Management Science Forthcoming*.

Sonnier, Garrett P., Leigh McAlister, and Oliver J. Rutz (2011), "A Dynamic Model of the Effect of Online Communications on Firm Sales," *Marketing Science,* 30(4) 702-716.

Stephen, Andrew T., Yaniv Dover, and Jacob Goldenberg (2011), "A Comparison Of The Effects Of Transmitter Activity And Connectivity On The Diffusion Of Information Over Online Social Networks," Working Paper, INSEAD.

— and Jeff Galak (2012), "The Effects of Traditional and Social Earned Media on Sales: A Study of a Microlending Marketplace," *Journal of Marketing Research,* 59(10) 624-639.

Toubia, Olivier and Andrew T. Stephen (2013), "Intrinsic Versus Image-Related Motivations in Social Media: Why Do People Contribute Content to Twitter," *Marketing Science Forthcoming*.

Ulaby, Neda (2010), "No-Cover Mondays: L.A.'s Rock Residency Scene. NPR broadcast," Online version available at http://www.npr.org/templates/story/story.php?storyId=128188020.

van Doorn, Jenny, Katherine N. Lemon, Vikas Mittal, Stephan Nass, Doreén Pick, Peter Pirner, and Peter C. Verhoef (2010), "Customer Engagement Behavior: Theoretical Foundations and Research Directions," *Journal of Service Research* 13 253-266.

Wedel, Michel and Wagner A. Kamakura (2000) "Market Segmentation: Conceptual and Methodological Foundations," Kluwer Academic Publishers.

Tables

Bands	Twitter Activity	Mean	SD	Min	Max
Focal ($n_{bands} = 37$)	Number of Band Tweets	650.68	775.27	42	3189
	Number of Follower Tweets	139.14	267.58	2	1014
	Number of Twitter Followers	1276.95	2799.21	12	17266
Control ($n_{bands} = 20$)	Number of Band Tweets	337.55	369.84	2	1190
	Number of Follower Tweets	103.50	252.26	2	1148
	Number of Twitter Followers	1513.96	2603.46	33	11460

Table 1: Summary Statistics for Twitter Data

Number of Concerts	Mean	SD	Min	Max
Focal Bands	12.26	14.17	0	53
Control Bands	17.38	16.08	0	51

Table 2: Summary Statistics for Band Concerts

Bands	Albums	Mean	SD	Min	Max
Focal ($n_{albums} = 34$)	Weekly Unit Sales	1237.65	4122.62	5	23245
	# Previous Albums	3.32	2.27	0	8
Control ($n_{albums} = 29$)	Weekly Unit Sales	1951.35	3227.05	11	11373
	# Previous Albums	9.28	7.81	0	25

Table 3: Summary Statistics for Album Sales

Band	Tweets	Emotional Content		Informational Content	
Type	by	Non-Emotional, %	Emotional, %	Non-Informative, %	Informative, %
Focal	Bands	67.91	32.09	72.10	27.90
	Followers	47.07	52.93	69.46	30.54
Control	Bands	63.81	36.19	70.85	29.15
	Followers	39.32	60.68	72.37	27.63

Table 4: Summary of Tweet Content Characteristics

	$T_{followers}$	$\%T_{bands}^{emotional}$	$\%T_{bands}^{informational}$	$\%T_{followers}^{emotional}$	$\%T_{followers}^{informational}$
T_{bands}	.227	.117	- .377	.055	.019
$T_{followers}$		- .028	- .144	- .274	- .126
$\%T_{bands}^{emotional}$			.208	.250	.026
$\%T_{bands}^{informational}$				.262	.327
$\%T_{followers}^{emotional}$					.134

Table 5: Correlation Between Band and Follower Tweets, by Content.

Band Type	Number of Followers	Tweets per Follower			
		Mean	SD	Min	Max
Focal	709	5.72	7.96	2	66
Control	290	5.21	13.85	2	213

Table 6: Tweets Used for Estimation.

Number of HMM States	Log Marginal Density
1	-39691.61
2	-26575.90
3	-23823.17
4	-25466.76
5	-29008.96

Table 7: Estimation Results: LMD and HMM Selection

From	To	No Tweets	nE, nI	E, nI	nE, I	E, I
S_L	S_M	- 3.4011(*)	1.4515(*)	1.5539(*)	1.1071(*)	.8107(*)
S_L	S_H	- 6.1137(*)	1.4830(*)	1.5539(*)	1.1430(*)	1.0862(*)
S_M	S_L	1.1787(*)	- .9436(*)	- .1316	- .6088(*)	- 1.4681(*)
S_M	S_H	- 3.5402(*)	- .0567(*)	.0993(*)	.0224	- .0055
S_H	S_L	- .8211(*)	- .0384(*)	.0902(*)	- .1210(*)	.0504
S_H	S_M	- .5628(*)	- 2.3759(*)	- 1.5405(*)	- .1312	- 2.6772(*)

(* The 95% posterior credible interval does not include zero.)

Table 8: Estimation Results: State Transition for Focal Bands

From	To	No Tweets	nE, nI	E, nI	nE, I	E, I
S_L	S_M	- 6.2181(*)	- .1165	- .7809(*)	5.5617(*)	- 1.3013(*)
S_L	S_H	- 8.0216(*)	- 1.9407(*)	1.1741(*)	3.5018(*)	.3980
S_M	S_L	- 2.0703(*)	- 2.7524(*)	- 1.5387(*)	- .2224	- .8780(*)
S_M	S_H	- 2.8779(*)	- .0060	.0480	- .0180	- .1014
S_H	S_L	- 2.1103(*)	- 3.4171(*)	.4178	- 1.1456(*)	- 2.0573
S_H	S_M	.2638	.0159	.1364	- .2053	- .4837(*)

(* The 95% posterior credible interval does not include zero.)

Table 9: Estimation Results: State Transition for Control Bands

No Tweet	Focal Bands			Control Bands		
% From \ To	S_L	S_M	S_H	S_L	S_M	S_H
S_L	96.57	3.22	.21	99.77	.20	.03
S_M	75.95	23.37	.68	10.67	84.57	4.76
S_H	21.89	28.35	49.76	5.0	53.73	41.27

Non-Emotional, Non-Informational Tweet

% From \ To	S_L	S_M	S_H	S_L	S_M	S_H
S_L	86.8	12.35	.85	99.82	.18	0
S_M	55.18	43.62	1.2	.76	93.99	5.26
S_H	28.68	3.59	67.74	.17	56.85	42.98

Emotional, Non-Informational Tweet

% From \ To	S_L	S_M	S_H	S_L	S_M	S_H
S_L	85.61	13.5	.9	99.8	.09	.11
S_M	73.41	25.76	.83	2.49	92.07	5.43
S_H	30.03	7.61	62.36	6.88	55.75	37.37

Non-Emotional, Informational Tweet

% From \ To	S_L	S_M	S_H	S_L	S_M	S_H
S_L	90.27	9.1	.63	65.38	33.91	.71
S_M	63.2	35.74	1.06	8.73	86.49	4.78
S_H	20.63	26.44	52.93	1.84	50.52	47.65

Emotional, Informational Tweet

% From \ To	S_L	S_M	S_H	S_L	S_M	S_H
S_L	92.46	6.93	.61	99.9	.05	.05
S_M	42.12	56.26	1.62	4.75	90.64	4.61
S_H	30.81	2.61	66.58	.85	44.15	55

Table 10: HMM State Transition Matrices (Calculated from Posterior Mean)

Retention Rates	Focal Bands	Control Bands
Emotion	.0552	.1343
Information	.0365	.2064

Table 11: Estimation Results: Retention Rates for Follower Knowledge.

Propensity to Tweet

	Focal Bands				Control Bands		
State	Intercept	Information	Emotion	State	Intercept	Information	Emotion
S_L	0	.1572(*)	.0292	S_L	.0001	- .5132	- 1.1994(*)
S_M	.0035	.0196	.0468	S_M	.0140	.0223	.0684(*)
S_H	.3531	.0312(*)	.0176(*)	S_H	.6873	.0056	- .0395

Propensity for a Tweet to be Emotional

State	Intercept	Information	Emotion	State	Intercept	Information	Emotion
S_L	- 5.0459(*)	1.9088(*)	- 2.0756(*)	S_L	- 1.1086	- .7211	.2081
S_M	.2605	.0395	- .0187	S_M	.4393(*)	- .0134	.0692
S_H	- .0971	- .0327(*)	.0335(*)	S_H	.0975	.1509(*)	- .0594

Propensity for a Tweet to be Informational

State	Intercept	Information	Emotion	State	Intercept	Information	Emotion
S_L	- .9985	- 2.6815(*)	- 1.5851	S_L	1.0011	.9755	.4261
S_M	- .2284	.0681(*)	- .1476(*)	S_M	- 1.1675(*)	.0918	- .0899
S_H	- 1.0590(*)	.0281(*)	.0043	S_H	- 1.1960(*)	.1682(*)	- .0787(*)

(* The 95% posterior credible interval does not include zero.)

Table 12: Estimation Results: State Dependent Follower Actions

	Focal Bands				Control Bands			
	Posterior		Post. Quantiles		Posterior		Post. Quantiles	
	Mean	SD	2.5%	97.5%	Mean	SD	2.5%	97.5%
	The Diffusion Curve							
$\overline{logit}(p)$	- 2.781	.295	- 3.372	- 2.219	- 2.135	.221	- 2.585	- 1.716
$\overline{ln}(q)$	- 8.678	1.768	- 12.498	- 5.758	- 14.941	2.426	- 19.733	- 10.504
σ_p^2	2.434	.736	1.381	4.186	1.169	.374	.645	2.089
σ_q^2	24.940	12.269	8.201	54.723	53.885	22.892	19.355	107.979
β_E	- .374	.063	- .498	- .247	.161	.058	.040	.271
β_{E_+}	2.133	.120	1.898	2.358	1.810	.120	1.578	2.033
$\beta_{Information}$	.099	.018	.063	.135	- .103	.013	- .128	- .078
$\beta_{Emotion}$	.084	.012	.062	.107	.121	.013	.097	.145
$\beta_{Concerts}$	.035	.007	.021	.048	.033	.006	.021	.046
	The Market Potential							
$ln(\gamma_0)$	5.312	.424	4.475	6.107	5.732	.404	4.887	6.486
$\gamma_{\#albums}$	.018	.738	- 1.496	1.318	.755	.454	- .149	1.697
$\gamma_{\#followers}$	.713	.305	.116	1.305	.515	.418	- .317	1.324

Table 13: Estimation Results: The Album Sales Model

Figures

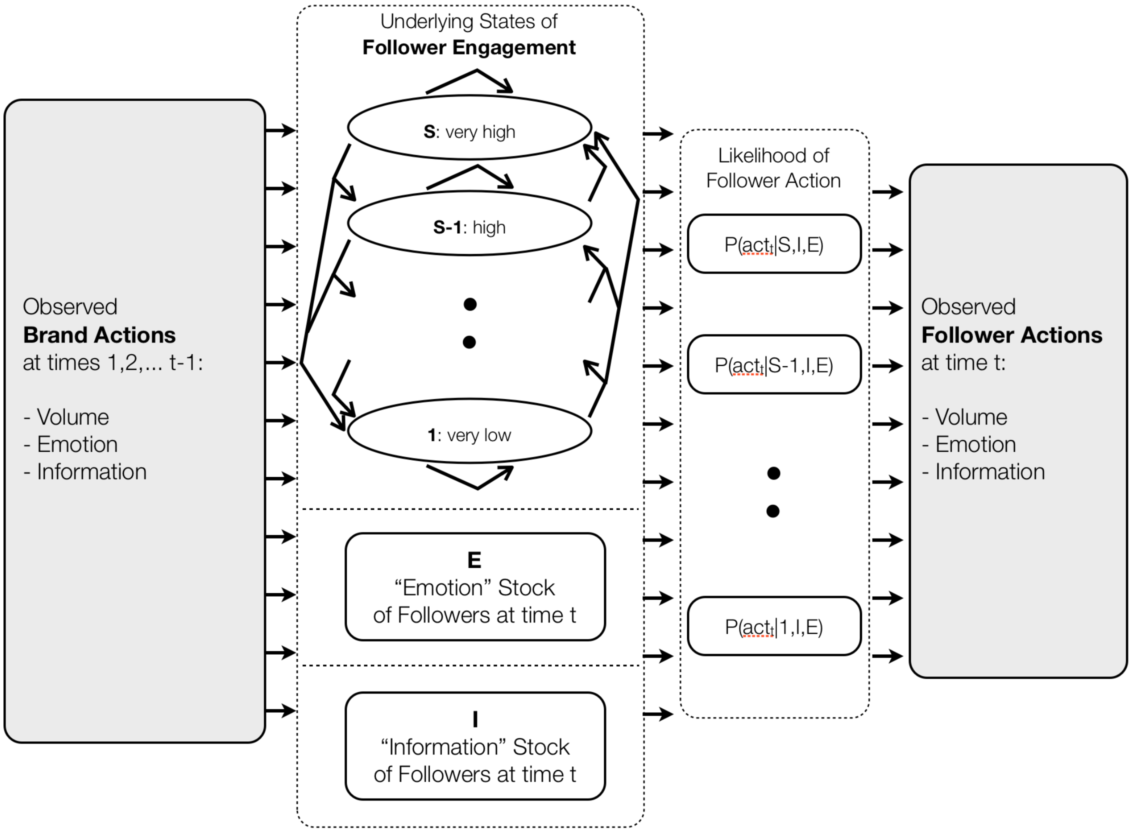

Figure 1: A Hidden Markov Model for Understanding Consumer Engagement in Digital Social Media.

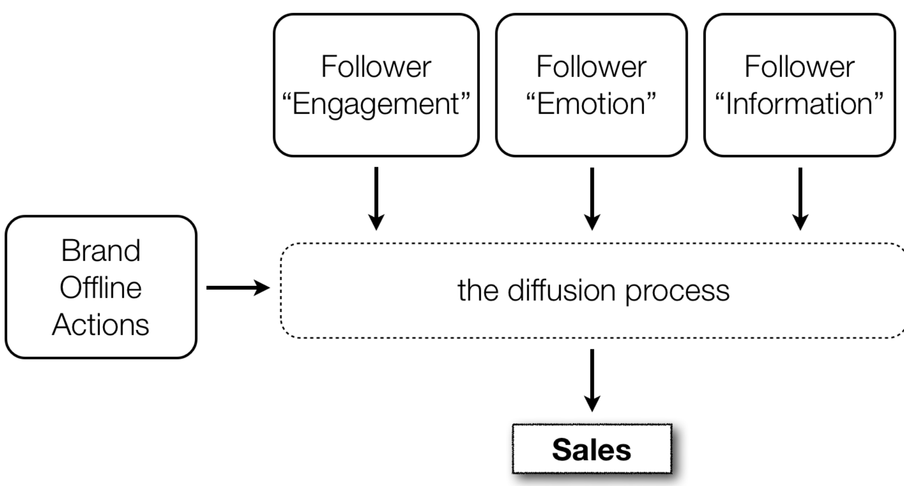

Figure 2: A Generalized Diffusion Model for Incorporating the Impact of Consumer Engagement on Sales.