Rational Spamming

by

Xinyu Cao

B.S. in Mathematics and Physics, Tsinghua University (2011)
M.S. in IEOR, UC Berkeley (2013)

Submitted to the Sloan School of Management
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE IN MANAGEMENT RESEARCH

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

SEPTEMBER 2016

Author .... Signature redacted ......
Sloan School of Management
AUG 05, 2016

Certified by...... Signature redacted ....
Juanjuan Zhang
Epoch Foundation Professor of International Management
Professor of Marketing
Thesis Supervisor

Accepted by ...... Signature redacted ....
Catherine Tucker
Sloan Distinguished Professor of Management
Professor of Marketing
Chair, MIT Sloan PhD Program

Rational Spamming

by

Xinyu Cao

Submitted to the Sloan School of Management
on AUG 05, 2016, in partial fulfillment of the
requirements for the degree of
MASTER OF SCIENCE IN MANAGEMENT RESEARCH

Abstract

Advertising on social media faces a new challenge as consumers can actively choose which advertisers to follow. Tracking company accounts, owned by 93 TV shows on the most popular tweeting website in China, provides evidence that firms advertise intensively, although doing so appears to drive followers away. An analytical model suggests that consumers with limited attention may rationally choose to unfollow a firm. This happens if consumers already know enough about the firm's quality and if the firm advertises too intensely. However, firms might still find intensive advertising the optimal strategy - if a firm is perceived as having a lesser quality offering, it wants to advertise aggressively to change consumers' beliefs about its quality; if a firm is perceived as having a higher quality offering, it also wants to advertise intensively, but in an effort to crowd-out advertising messages from its competitors. [1]

Thesis Supervisor: Juanjuan Zhang
Title: Epoch Foundation Professor of International Management
Professor of Marketing

[1]This essay is based on joint work with John R. Hauser, T. Tony Ke, and Juanjuan Zhang. I gratefully acknowledge the helpful comments from Catherine Tucker, Duncan Simester, Birger Wernerfelt, Sinan Aral, seminar participants at Massachusetts Institute of Technology, and attendees of the 2016 ISMS Marketing Science Conference.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

With the emergence of new communication technologies and platforms, disseminating information to the public has never been more convenient. Firms in particular are able to advertise to consumers in new, more economical ways. For example, firms can create their official accounts on a social media platform, and advertise by posting product-related messages at negligible marginal costs. However, a consequence of such convenience of communication is that consumers are exposed to far more information than they can process. In 2015, every internet user receives 122 emails on average per day,[1] and over 500 million tweets are sent on Twitter per day, among which 71% are ignored and only 23% generate a reply.[2]. In the age of information explosion, a major challenge firms face in advertising is the limited attention of consumers. In this paper, we use the advertising of TV shows on a tweeting website as a concrete setting to investigate how consumers respond to excessive advertising and what competing firms' optimal advertising strategy should be given the limited attention of consumers.

We focus on Weibo, the most popular tweeting website in China. It has similar features as Twitter, and has attracted numerous firms to open business accounts on the platform. In particular, most media companies in China, such as TV channels, newspapers, and news websites, have opened accounts on Weibo to advertise their

[1]http://www.radicati.com/wp/wp-content/uploads/2015/02/Email-Statistics-Report-2015-2019-Executive-Summary.pdf

[2]http://www.digitalinformationworld.com/2015/01/twitter-marketing-stats-and-facts-you-should-know.html

products. We track the Weibo accounts of 93 weekly TV shows that were broadcasted across China from April to October 2015. We find that these company accounts post tweets intensively, and there is no big difference in posting intensity across shows with different levels of popularity. We further look at the change in the number of followers. We find that the effect of posting tweets on the change in the number of followers decreases with posting intensity, and the curve goes down when posting intensity is too high. Since more tweets tend to attract more new followers, it is probable that existing followers choose to unfollow the accounts that post too many tweets.

Motivated by these observations, we develop an analytical model to understand why consumers with limited attention choose to unfollow an account and, more importantly, why firms choose to advertise intensively knowing the limited attention of consumers.

In the model, we treat tweets as a form of informative advertising for a TV show. Following the literature of informative advertising (Erdem & Keane 1996, Erdem *et al.* 2008, etc.), we assume that advertising exposure gives a consumer noisy signals about show quality, and the consumer uses these signals to update his belief about show quality in a Bayesian manner. As his belief becomes more precise, the consumer will face a greater probability of choosing the show with higher true quality, which as a result increases the consumer's expected utility from watching the chosen show. However, the consumer's reading capacity is restricted by his limited attention. If a TV show's account is posting intensively and the consumer already knows about the TV show quite precisely, the consumer would rationally unfollow the account. In this way, he can leave more attention to tweets on shows he knows less about and make a more informed choice of which show to watch eventually.

Understanding the mechanism behind consumers' unfollowing decisions leads to a counterintuitive result - our model indicates that posting intensively can be an optimal strategy for firms. Firms are competing for consumer's limited attention. If a firm posts more tweets, the belief about its show quality may experience a larger change because consumers can receive more information about the show to update their belief (we call this effect the "dispersion effect"); at the same time, information signals from the competitor are crowded out, so the mean belief about the competitor's show quality is likely to change less (we call this effect the "crowd-out effect"). The firm perceived as having higher quality wants to keep its advantage and hence prefers smaller changes in beliefs about itself and about its competitor. Therefore, this firm would like to exploit the crowd-out effect but avoid the dispersion effect. The firm perceived as having lower quality would like to reverse the current situation, hence it favors the dispersion effect but dislikes the crowd-out effect. Both firms need to make a trade-off between these two effects when deciding their posting intensities. When the better-perceived firm has a more precise prior belief among consumers than its worse-perceived competitor, the benefit from the crowd-out effect dominates the damage of the dispersion effect, and the better-perceived firm will post as many tweets as possible. The worse-perceived firm would post intensively in any case because its priority is to improve its perception among consumers through the dispersion effect. As a result, it is rational for both types of firms to post intensively knowing the limited attention of consumers.

There is a growing literature on limited attention. Previous works in psychology provide evidence that people have limited attention in receiving and processing information (see Eppler & Mengis 2004, DellaVigna 2009 for a review). Van Zandt (2004) points out that as the costs of generating and transmitting information fall, the main bottleneck in communication becomes the limited attention of human receivers. He shows that mechanisms that increase the cost of sending messages will shift the task of screening messages from the consumers to firms, who know the contents better, a shift that will benefit both parties. Anderson & De Palma (2009) further consider message examination as an endogenous decision of consumers, and confirm that a higher cost of sending messages will increase the number of messages examined by consumers because the expected average surplus rises. Built on these findings, Anderson & De Palma (2012) model multiple sectors competing for consumersâĂŹ attention, with competition in price within each section. Iyer & Katona (2015) analyze the incentives to be a sender or a receiver for non-commercial users in social networks given that people derive utility from being listened to but receivers have limited attention. Zhu & Dukes (2015) study competing firms' strategy in choosing the product attribute to make prominent when consumers have limited attention, and find that firms emphasizing the same attribute may actually increase product differentiation. Our paper shares a similar context with Anderson & De Palma (2009, 2012) since we all investigate firms' advertising strategies in response to the limited attention of consumers. The difference is that we consider a real-life situation where firms do not pay for consumers' attention, but face the risk of losing or damaging its quality image among attention-constrained consumers. Furthermore, we derive the surprising result that posting intensively can be an optimal strategy for firms in spite of consumers' limited attention.

Our paper also contributes to the literature of competitive advertising strategy. This literature has traditionally focused on advertising expenditure, where firms face the tradeoff between higher market shares (or consideration level) and higher advertising costs (Chintagunta & Vilcassim 1992, Chintagunta & Vilcassim 1994, Erickson 1992, Nguyen & Shi 2006, Villas-Boas 1993, etc.). Past studies have analyzed the profit-maximizing advertising strategies in dynamic competitive settings, using the technique of differential games (Lanchester model) to describe the change in market share as a function of advertising expenditure of two competing firms. However, as mentioned above, one primary challenge of advertising in the information age is consumer's limited attention, and the downside of intensive advertising is not ad costs but the potential loss of audience and a possible decline in reputation. We contribute to the literature of competitive advertising by emphasizing these new perspectives.

Finally, this paper is related to the literature on social media advertising. Stephen & Galak (2012) examine how traditional earned media (e.g., press mentions) and social earned media (e.g., blogs and online community posts) affect sales. They find that, accounting for event frequency, social earned media has a larger effect on sales in aggregate. Toubia & Stephen (2013) point out that as Twitter gets mature, noncommercial users will probably reduce their posting and Twitter may become a media channel where firms broadcast content to consumers. Existing works have shown that social media websites like Twitter and Weibo can be an effective marketing platform. Gong *et al.* (2014) conduct a field experiment on Weibo, showing that posting tweets about TV shows and recruiting influentials to repost these tweets increase show viewing significantly. Liu *et al.* (2014) find that the contents of tweets related to TV shows can make more accurate predictions on show viewing compared to other online data, which at the same time provides evidence that Twitter might be an effective marketing platform for the spread of information. Built on these studies, our paper captures two features of social media advertising (especially tweeting) - low advertising costs and limited consumer attention. We also take a step further to explore competing firms' advertising strategies in an equilibrium setting.

The rest of the paper is organized as follows. In §2, we describe the motivating evidence gathered on Weibo. In §3, we presents the model, analysis, and results. In §4, we discuss the limitations of the model and directions for future research.

Chapter 2

Background and Motivating Phenomena

2.1 Background and Data

Weibo.com is a leading microblogging website in China. Launched by Sina Corporation in August 2009, it has been one of the most popular social media platforms in China. As of June 2015, the number of monthly active users has been 212 million. [1]

The functions and features of Weibo are very similar to Twitter. A registered user can post a "tweet", which is a message with no more than 140 Chinese characters and may include images, short videos, URLs, or a pair of hashtags that indicate the topic of the tweet. A user can "follow" other users to subscribe their tweets so that their future tweets will appear on his homepage automatically, and he can "unfollow" these users later on to cancel this subscribing relationship. Notice that following is a unidirectional relationship: the followee's tweets will appear on the follower's homepage, but not vice versa. A user can "like" a tweet by clicking the "like" button. He can also give a comment under a tweet, or repost a tweet so that it will appear on his timeline as a "retweet". For each user, the tweets posted by those he follows will appear on his homepage according to a reverse chronological order, that is, the most recent tweets will be displayed at the top of the page, and he needs to scroll down

[1] http://tech.sina.com.cn/i/2015-08-19/doc-ifxfxrav2874963.shtml

the screen to see the earlier posts.

The unidirectional relationship and the huge user base make Weibo an ideal platform for wide and fast transmission of information, and hence more and more celebrities, organizations, companies, and media have opened accounts on Weibo to use it as a marketing tool for themselves or their products. Weibo verifies the authenticity of these accounts by adding a mark "V" under the account's profile photo.

We collected data from Weibo to observe how do weekly TV shows advertise on this popular social media platform. There are more than 30 satellite TV channels in the mainland of China, including national and provincial ones, all of which can broadcast their shows to the audience throughout the country. Almost all of these channels have at least one weekly TV show, aired at prime time during weekends. Most of them are reality shows or entertainment shows. Some of them have been very popular and successful, attracting millions of audience. With highly overlapping broadcasting times and similar topics, these TV shows are competing severely. As a major social media platform, Weibo is playing an important role in these TV channels' social media marketing campaign. Usually, a TV channel opens one account for each of its weekly TV show, posting tweets containing information related to the show.

We track 93 accounts, each representing a weekly TV show on a satellite TV channel[2]. We collect data from Weibo API[3] from April 2015 to October 2015 on an hourly base. All of these shows have been broadcasted during the six months, and all accounts have posted tweets during this period. The data we scraped include the number of followers of the account at the scraping time, and all the tweets the account posted during the hour. For each tweet, we know the following characteristics: when it was created, how many times it has been retweeted, commented, or liked, the content of tweets, whether it contains a picture, whether it has a hashtagged topic, whether it is a retweet, the content and poster of the original tweet if it is a retweet. To observe consumer's response to firms' posting behavior, we divide one day into

[2]These are all the active accounts we can find for in-season TV shows in the period. All the shows we track are reality shows or entertainment shows. For each show, different episodes are in the same form but the contents are not necessarily correlated, so the information value of tweets does not decrease very fast as time goes on

[3]http://open.weibo.com/wiki/%E5%BE%AE%E5%8D%9AAPI

six four-hour periods, and calculate the change in the number of followers and the number of tweets for each account in every four-hour period. We also aggregate the characteristics of tweets within every four-hour period.

In order to get supplemental data that reflects the popularity of each show, we scrape data from Baidu Search Index using the name of each show as a keyword[4]. Baidu is the leading search engine in China. Baidu Search Index gives the public interest trend on Baidu for each keyword, and the index is comparable across different keywords. The index we scraped is on a daily base and cover the same time range as the Weibo data.

2.2 Summary Statistics and Motivating Phenomena

Table B.1 gives the summary statistics of each account in the six months, including the average number of followers throughout the period, the total number of tweets each account posts in the period, the maximum number of tweets posted in a four-hour period, and the average daily Baidu search index. We can see that these accounts are overall quite popular, with a mean number of followers more than half million, and the most popular account has more than eight million followers. They are posting tweets actively: the maximum number of tweets an account has posted in a four-hour period is around 20 tweets on average, and the craziest account has posted 81 tweets in a four-hour period (that is, approximately in every 3 minutes one tweet is posted).

In Table B.2, we summarize the dynamics for each account in every four-hour period. We only take active periods into account, that is, periods at which the account posts at least one tweet. We see that their numbers of followers overall increase over time, but losses of followers do happen. The number of tweets posted in a four-hour period varies in a wide range (from 1 to 81).

There are three main features emerging from the data. [5]

[4]http://index.baidu.com/

[5]There is a big difference in the management level across channels. Not all the channels are actively engaged in managing their Weibo accounts. When observing the posting behavior of firms (Table 3~6), we only keep the accounts registered by the top 10 channels in China. The rank of channels comes from their rating in the last five years, see http://www.tvtv.hk/archives/2445.html.

First, these accounts post intensively, especially on the air date of the show and at night.

From Table B.3 and Table B.4 we can see that each account posts three tweets in a four-hour period on average, and the posting intensity is even higher on the air date or at night (8pm-12am). On the air date, an account posts nearly seven tweets in every four hours on average, and at night, an account posts about five tweets in four hours. According to our small-scale qualitative interview with Weibo users[6], a typical user has the patience to read about fifty tweets in one night, and thus the tweets from ten media accounts can take up the user's attention. However, a user usually follows more than one-hundred accounts, and a significant proportion of them are verified accounts (accounts registered by celebrities, media, organizations, etc.). Therefore, the aggregate posting intensity of these accounts is very likely to exceed the reading capacity of consumers.

Second, there is no big difference in posting intensity at peak time for shows with different levels of popularity.

Table B.5 and Table B.6 summarize the posting intensity at night for shows with different levels of popularity. In Table B.5, shows are classified into four groups according to their average Baidu search index, and in Table B.6, shows are classified according to their average number of followers. We find that the difference in posting intensity across different groups is not big, and the posting intensity does not positively correlate with popularity. In Table B.5, the mean posting intensity in the first three quartiles slightly increases, but the mean posting intensity of the fourth quartile is lower than the third quartile. In Table B.6, the first and third quartiles post slightly more than the second and fourth quartiles. We see that the group of most popular accounts is not necessarily the accounts that post the most intensively.

Third, the effect of more tweets on the increase in the number of followers has a decreasing marginal return, and too many tweets are probably driving existing

There are 49 accounts registered by these ten channels.

[6]The qualitative interview was done within the author's friends. About ten people who are frequent users of Weibo did the interview, and they were asked about their using habit as well as how many accounts and verified accounts they were following. We did not find publicly available statistics of the overall behavior of Weibo users.

followers away.

To observe consumer's response to posting intensity, for each account, we regress the change in the number of followers in every four-hour period on the number of tweets the account posts in the period and the quadratic term of the number of tweets. The coefficient of the quadratic term is significantly negative, as shown in column (1) of Table B.7. After we control for the popularity of the show, characteristics of the show, and characteristics of the tweets posted during the period (column (2)), the sign and significance of the quadratic term do not change. We further control for the fixed effects of time of day, date, month, and the fixed effects of the account, the result still holds (column (3)). In column (4)-(6), we add the interaction term of the number of tweets with the indicator of whether it's on the air date of the show, whether it's around the broadcasting time, and whether it's at night. These are the time periods when most accounts post actively. All three interaction terms have significant negative coefficients on a quite large scale, meaning that the effect of a tweet in these time periods are much smaller than in other time periods. However, these time periods should be the time when consumer pay more attention to Weibo and the show account, so it is very likely that too many tweets posted at these time periods have driven followers away.

Figure A-1 and Figure A-2 can help us see the relationship between change in the number of followers and posting intensity more clearly. In Figure A-1, the horizontal axis is the number of tweets posted by each account in a four-hour period, and the vertical axis is the change in the number of followers for each account in the period. Since it looks too messy with all the data points, we only keep the data points of the periods during which the show is broadcasted. We can see that significant loss of followers happens when the number of tweets gets too large, and the average fitting curve has an inverse U-shape. In Figure A-2, we use the data of a popular show called "Dad, Where Are We Going" to observe its trend of daily posting intensity and the corresponding change in the number of followers throughout the broadcasting season of the show. We can see that at most peak times of posting, the change in the number of followers has a trough, particularly in the later stage of the show season.

Motivated by the above phenomena, we would like to ask: when will a rational consumer choose to unfollow an account? Given that consumers have limited attention, what should be competitive firms' optimal posting strategy? We next build a model in which tweets serve as informative advertising and consumers have limited attention. We give an explanation of why consumers may unfollow an intensively posting account rationally. We also find that posting intensively can be the optimal strategy of firms under the competitive setting, even when they know the possible unfollowing action of consumers.

Chapter 3

Analytical Model and Analysis

There are two firms, each having a product (a TV show) and a Weibo account that represents the TV show. Each firm posts tweets through its Weibo account to deliver quality signals of the TV show. A representative consumer follows both accounts at the beginning and will read tweets and then choose a show to watch.

The timing of the game is as follows.

Period 0:

The consumer and two firms share the same prior belief of show quality, $Q_i \sim N(\mu_{i0}, \sigma_{i0}^2),\ i = 1, 2$.

Two firms each decide on the number of tweets to post in the period, $n_i,\ i = 1, 2$, which we call posting intensity.

The consumer gets to know the ratio of the two firms' posting intensity $\psi = n_1/n_2$ by sampling some tweets, and then decides whether to unfollow an account or not.

The consumer reads R_i tweets from each account i that he is still following. Each tweet is an informative but noisy signal of the show quality, and they are independent and identically distributed following a normal distribution, i.e. $S_{ik} \sim N(Q_i, \sigma_s^2)$.

Period 1:

The beliefs about show quality get updated because of the signals conveyed by tweets. The updated belief is $Q_i \sim N(\mu_i, \sigma_i^2)$. We assume the quality signals and belief updating processes are independent across shows.

Then the consumer decides which show to watch according to the posterior belief.

Assume the consumer is risk-neutral, and the utility from watching show i is $U(i) = Q_i$. He will choose the show with higher expected utility, where the expected utility of choice i is $EU(i) = \mu_i$.

There are four key assumptions in the model that we need to clarify.

First, the "quality" of a show measures how closely the content of the show matches consumer's taste. Firms are uncertain about the consumer's taste and how well their shows match the taste, so they have no private information about show quality, and they share the same belief as the consumer. This assumption has been made in previous literatures. For example, Bose *et al.* (2002, 2006) make such assumption to study the optimal pricing decision of a monopoly. By making such assumption, we rule out the possibility that firms can signal their quality via advertising intensity.

Second, we treat tweets as a form of informative advertising. Unlike traditional TV commercials, tweets posted by firm accounts are usually non-repetitive, revealing various information related to the product, so it is natural to consider them as informative to consumers. For example, tweets posted by the TV show accounts usually reveal some contents or short clips of the upcoming episode. It is possible that tweets are also consumption goods which are complements of TV shows, but it is hard to separate these two roles empirically or consider them at the same time. There has been a long debate about the informative versus persuasive (consumption) roles of advertising, see Bagwell (2007) for a comprehensive summary. Ackerberg (2001) empirically distinguishes these two effects using consumer-level data on purchases of a new brand of yogurt, and find that the primary effect of the advertisements was informative. Therefore, we treat tweets as a form of informative advertising in our model, and assume the consumer uses the informative but noisy signals conveyed by tweets to update his quality belief according to Bayes rule, which follows the literature of informative advertising (Erdem & Keane 1996, Erdem *et al.* 2008, etc.).

Third, the consumer has limited attention, and he can read no more than M tweets in each period. When the total number of tweets posted exceeds his reading capacity, the consumer cannot read all the tweets. Since tweets from different accounts are mixed and displayed to the consumer according to a reverse-chronological order, it is very difficult for the consumer to decide how many tweets he will read from each account. Hence we assume the number of tweets the consumer can read from account i is proportional to the number of tweets posted by firm i, that is, when $n_1 + n_2 > M$, $R_i = \frac{n_i}{n_1+n_2}M$, $i = 1, 2$.[1] We assume the consumer can choose to unfollow an account at period 0. There are two types of consumers. A sophisticated consumer may decide to unfollow an account rationally, and a naive one will not unfollow accounts in an effort to optimize the amount of information he can get. The representative consumer is a sophisticated one with probability p, and a naive one with probability $1 - p$. The reason we assume a sophisticated consumer may choose to unfollow an account rather than to optimize the number of tweets to read from each account is twofold. One is that the feature of the platform makes it too costly for the consumer to do so, as we have illustrated above. The other is, Sims (2006) has pointed out that a rational agent with limited attention will react in discrete jumps to signals that a fully rational agent will respond to continuously, and unfollowing an account is a form of discrete reactions to excessive advertising.

Fourth, since we are considering the tweeting and watching decisions for only one episode of each show, the true quality Q_i is the quality for the specific episode, so Q_i is a fixed number, just unknown to firms and the consumer. The variances in prior and posterior beliefs (σ_{i0}^2 and σ_i^2) measure the precision of the beliefs, that is, how much the belief (μ_{i0} and μ_i) might differ from the true quality; they are not the variance of the true quality.

In the following, we first demonstrate the informative value of tweets to the consumer and the two firms. Then we illustrate when a consumer with limited attention will unfollow an account rationally. Lastly, we explore what should be the two competitive firms' advertising strategies given the consumer's limited attention.

[1] Given the short length of tweets, we do not consider the difference between "preview a tweet" and "read a tweet", and it is reasonable to assume that each tweet takes the same amount of attention.

3.1 Informative Value of Reading Tweets

For a risk-neutral consumer, the informative value of reading more tweets is that it increases the expected utility from watching the show that will be chosen at period 1. We will interpret this effect in two ways, and then show that these two interpretations are actually equivalent.

1. Reading more tweets reduces the posterior uncertainty in quality, and hence increases the probability of choosing the show with higher true quality.

Before the consumer reads the tweets, the belief about firm i's show quality is $Q_i \sim N(\mu_{i0}, \sigma_{i0}^2)$. After the consumer reads R_i tweets from firm i's account with each tweet conveying a signal $S_{ik} \sim N(Q_i, \sigma_s^2)$, the belief gets updated to $Q_i \sim N(\mu_i, \sigma_i^2)$ according to the Bayes rule:

$$\mu_i = \frac{\sigma_s^2}{\sigma_s^2 + R_i\sigma_{i0}^2}\mu_{i0} + \frac{R_i\sigma_{i0}^2}{\sigma_s^2 + R_i\sigma_{i0}^2}\bar{S}_i, \ (\bar{S}_i = \frac{\sum_{k=1}^{R_i} S_{ik}}{R_i}) \tag{3.1}$$

$$\sigma_i^2 = \frac{\sigma_{i0}^2\sigma_s^2}{\sigma_s^2 + R_i\sigma_{i0}^2}. \tag{3.2}$$

The posterior uncertainty σ_i^2 decreases with R_i, the number of tweets that the consumer will read. It means that when the consumer can read more tweets from account i, the posterior belief μ_i gets closer to the true quality Q_i. When the sum of posterior variances gets smaller $(\sigma_1^2 + \sigma_2^2)$, it means that overall the consumer knows about the true quality of the two shows more precisely. As a result, the probability that the consumer can choose the show with higher true quality to watch will increase. The consumer's expected utility from watching the show that will be chosen equals $\mathbb{E}[U] = \mathbb{E}[Q_1P(\mu_1) + Q_2P(\mu_2)]$. When the consumer has a larger chance to select the show with higher true quality, his expected utility from watching the chosen show will get larger. We formalize this conclusion in Lemma 1.

Lemma 1 *As the consumer reads more tweets from account i (R_i increases), the posterior uncertainty in firm i becomes smaller (σ_i^2 decreases). As the sum of posterior uncertainty $\sigma_1^2+\sigma_2^2$ gets smaller, the probability that the consumer will choose the show* *with higher true quality gets larger, and as a result the expected utility from watching the show that will be chosen increases.*

(See Appendix C1 for proof)

Notice that if the consumer is risk-averse, a smaller σ_i^2 has an additional benefit, i.e. making the certainty equivalence larger, but it's different from the informative value we are considering here.

2. Reading more tweets makes the potential change in beliefs larger. As a result, the expectation of the maximum of posterior means increases, which equals the expected utility from watching the show that will be chosen.

As we have illustrated, before the consumer reads the tweets, the belief about firm i's show quality is $Q_i \sim N(\mu_{i0}, \sigma_{i0}^2)$, and after the consumer reads R_i tweets from firm i's account with each tweet conveying a signal $S_{ik} \sim N(Q_i, \sigma_s^2)$, the belief gets updated to $Q_i \sim N(\mu_i, \sigma_i^2)$. At the time when the consumer has not read the tweets, due to the uncertainty in the upcoming signals S_{ik} and the uncertainty in true quality Q_i, the value of posterior belief μ_i is unknown, following a normal distribution $N(\mu_{i0}, V_i)$,[2] where

$$V_i = \frac{R_i \sigma_{i0}^4}{\sigma_s^2 + R_i \sigma_{i0}^2}. \tag{3.3}$$

(See Appendix C2 for proof.) V_i measures how much the posterior belief μ_i can differ from the prior belief μ_{i0}. It is easy to see that V_i increases in R_i, meaning that by reading more tweets, consumers can obtain more information to update the belief, so that the potential change in belief gets larger.[3] An increase in V_i means that the possible values of posterior belief μ_i gets more dispersed around μ_{i0}.

[2]Notice that $N(\mu_{i0}, V_i)$ is not the posterior belief of product quality, but the consumer's anticipated distribution of the posterior belief before he reads the tweets.

[3]You may notice that V_i also increases with the variance of prior belief, σ_{i0}^2, for two reasons: the consumer will put more weights on the information signals when the prior belief is not precise, and a higher σ_{i0}^2 also means that the upcoming information signals can be more volatile. If the consumer reads no tweets ($R_i = 0$), V_i equals 0, meaning that the mean belief will not change ($\mu_i = \mu_{i0}$ for sure); if the consumer reads very large number of tweets ($R_i \to \infty$), the posterior mean belief will approaches the true quality Q_i, and thus the change in the mean belief equals the difference between true quality Q_i and prior belief μ_{i0}, which follows $N(0, \sigma_{i0}^2)$, consistent with the fact that $V_i \to \sigma_{i0}^2$ as $R_i \to \infty$.

At the time when the consumer has read the tweets and is deciding which show to watch, he will choose the show with higher value of μ_i, and his expected utility from watching the show will be $\max\{\mu_1, \mu_2\}$. Before the consumer actually reads the tweets, the values of μ_1 and μ_2 are unknown, so at this time his expected utility from watching the show that will be chosen is $\mathbb{E}[\max\{\mu_1, \mu_2\}]$. Knowing the normal distribution of μ_i, we can calculate the expected maximum of these two independent normal random variables according to Nadarajah & Kotz (2008):

$$\mathbb{E}[\max\{\mu_1, \mu_2\}] = \mu_{10}\Phi(\frac{\mu_{10} - \mu_{20}}{\sqrt{V_1 + V_2}}) + \mu_{20}\Phi(\frac{\mu_{20} - \mu_{10}}{\sqrt{V_1 + V_2}}) + \sqrt{V_1 + V_2}\phi(\frac{\mu_{10} - \mu_{20}}{\sqrt{V_1 + V_2}}). \tag{3.4}$$

where $\Phi(\cdot)$ and $\phi(\cdot)$ represent the CDF and PDF of standard normal distribution respectively, and in fact the right-hand-side of formula (3.4) is a strictly increasing function of $V_1 + V_2$. Lemma 2 formalizes the conclusion.

Lemma 2 *As the consumer reads more tweets from account i (R_i increases), the potential change in belief about firm i will be larger (V_i gets larger). Consumer's expected utility from watching the show that will be chosen is given by equation (3.4), and it strictly increases with the total potential change in beliefs ($V_1 + V_2$).*

(See Appendix C3 for proof.)

Intuitively, an increase in V_i means that the possible values of show i's posterior belief (μ_i) get more dispersed. As the possible values of both μ_1 and μ_2 get more dispersed, the expectation of the maximum of μ_1 and μ_2 gets larger. Since V_i increases with the number of tweets that will be read (R_i), Lemma 2 shows that the consumer can have a higher expected utility if he reads more tweets.

To summarize, when the consumer reads more tweets from account i, the posterior uncertainty (σ_i^2) gets smaller, and the potential change in belief (V_i) becomes larger. In fact, for any increment in the number of tweets to read, the increase in V_i is always equal to the decrease in σ_i^2, and both of them result in an increase in the expected utility from watching the show that will be chosen. (see Appendix C4 for proof).

They are just two ways to understand the informative values of tweets. From the consumer side, it is equivalent to consider the problem in either way. However, from the firm side, the two firms mainly care about the potential change in belief (V_1+V_2). The reason is that firms do not know whether their true quality is the higher one or the lower one, and thus it is not necessarily good for a firm that consumers know the true quality more precisely. Knowing the relative positions of their prior beliefs, the firms care about the potential change in the beliefs. For consistency, we will use the potential change in beliefs $(V_1 + V_2)$ to analyze the decision of the consumer and firms.

3.2 Rationality of Unfollowing

As we have illustrated in the assumption part, with limited attention, the consumer can only read a proportion of tweets from each account when the accounts post intensively $(n_1 + n_2 > M)$, and he may choose to unfollow an account at period 0 so as to better allocate his attention. Obviously, it must be suboptimal for the consumer to unfollow both accounts, since he will waste the attention at period 0 and obtain no information. If the consumer unfollows account 1, he cannot see any tweets from account 1, $R_1 = 0$, and then $R_2 = \min\{M, n_2\}$. According to Lemma 2, the consumer will have a higher expected utility from watching the show that will be chosen when $V_1 + V_2$ is larger. Hence when deciding whether to unfollow account 1, the consumer can simply compare the value of $V_1 + V_2$ under the case of following both accounts ($R_1 = \frac{n_1}{n_1+n_2}M$ and $R_2 = \frac{n_2}{n_1+n_2}M$) and that of unfollowing account 1 ($R_1 = 0$ and $R_2 = \min\{M, n_2\}$). If the latter one leads to a higher value of $V_1 + V_2$, then the consumer will be better off to unfollow account 1. The decision of whether to unfollow account 2 is symmetric. The condition for the consumer to unfollow account i $(i = 1, 2)$ is given in Proposition 1. For the tractability of the solution, here we only consider the case when both accounts post more than the consumer's reading capacity $(n_1, n_2 \geq M)$ [4].

Proposition 1 *When the prior belief about firm i is more precise ($\sigma_{i0}^2 < \sigma_{j0}^2$), the consumer will be better off to unfollow account i if he has very limited attention ($M < M_0 \equiv \frac{\sigma_s^2(\sigma_{j0}^2 - \sigma_{i0}^2)}{\sigma_{i0}^2 \sigma_{j0}^2}$) or account i posts too intensively ($\frac{n_i}{n_j} > \frac{M - M_0}{M_0}$). The consumer will not benefit from unfollowing the account with less precise prior belief.*

The intuition of this result is, if the consumer already knows more about show i, the benefit he can get from reading more tweets of show i is smaller compared to reading more tweets of the other show, which he knows less about. Hence he may want to unfollow account i so that he can make room for more information from the other account, and then he can make a more informed decision when choosing the show to watch. When his attention is very limited, he will do so for sure because he would like to allocate his scarce attention to the more useful information. When his attention is not that limited, he will unfollow account i if it posts too intensively and crowds out too much information from the other account.

3.3 Firms' Competitive Advertising Strategy Given Limited Attention of Consumers

Now we ask, knowing that consumers have a limited reading capacity, what should be firms' advertising strategy. We find that competitive firms can be strategic in advertising intensively.

The ultimate goal of a firm should be to maximize the number of people who will watch the show. To achieve this goal, the firm needs to inform more people of the show, and also to maximize the probability that an informed consumer would choose its show to watch. We should notice that the followers of a firm on social media are already informed of the show. These followers may repost the tweets from the firm,

[4]When $n_1+n_2 \leq M$, the consumer has no incentive to unfollow. When $n_1+n_2 > M$, $0 < n_j < M$, and $\sigma_{i0}^2 < \sigma_{j0}^2$, the consumer may unfollow account i when n_i is large and n_j is not very small, but the boundary condition is extremely complicated to express. In §3.3 we show the case that $n_1+n_2 > M$, $0 < n_j < M$ will not happen in equilibrium.

and their retweets can let more people know about the show, but this process is very complicated, and the effect is hard to quantify. Those who are not following the firm's account may also be aware of the show. Also, followers' quality perception of the show can spread out through word-of-mouth and affect the decision of others. Therefore, regarding advertising on social media, we focus on the role of tweets in influencing consumer's quality perception of the show and then affecting the probability that an informed consumer will choose the show to watch. The marginal cost of posting a tweet is negligible, so when deciding on how many tweets to post, the firm is not considering the trade-off between the returns and expenditures, but the impact of tweets on consumer's beliefs.

Suppose firm 1's show is believed to have higher quality than that of firm 2 *a priori* ($\mu_{10} > \mu_{20}$).[5] Since the consumer will choose the show with higher posterior mean (μ_i) to watch and the distribution of μ_i is $N(\mu_{i0}, V_i)$, the probability that the consumer will watch show 1 is

$$P(\mu_1 > \mu_2) = \Phi(\frac{\mu_{10} - \mu_{20}}{\sqrt{V_1 + V_2}}), \tag{3.5}$$

where $\Phi(\cdot)$ denotes the CDF standard normal distribution, and the value of V_i ($i = 1, 2$) is given by equation (3.3). We notice that $\Phi(\cdot)$ is an increasing function and $\mu_{10} - \mu_{20} > 0$, so the probability that the consumer will watch show 1 as given by equation (3.5) decreases with $V_1 + V_2$. Therefore, firm 1 prefers $V_1 + V_2$ to be smaller, and firm 2 prefers $V_1 + V_2$ to be larger. The intuition is as follows. Recall that V_i measures how much the posterior belief about show i can be different from the prior belief. Firm 1 wants to keep its advantage in consumer's perception of show quality, so it wants the potential changes in the beliefs of both shows to be smaller ($V_1 + V_2$ smaller); firm 2 seeks to reverse the current situation, so it prefers the potential changes in beliefs to be larger ($V_1 + V_2$ larger).

Now we know, given the value of μ_{10} and μ_{20}, the probability that the consumer will choose show 1 or show 2 to watch depends on the value of $V_1 + V_2$. The objective

[5]Notice this does not necessarily mean show 1 has higher true quality.

firm 1 can be translated as minimizing $V_1 + V_2$, whereas firm 2's objective is to maximize $V_1 + V_2$. Therefore, to find the firms' advertising strategy in equilibrium, we first need to know how the two firms' advertising intensities affect the value of $V_1 + V_2$. Lemma 3 summarizes how $V_1 + V_2$ varies with the ratio of posting intensities of the two firms under different conditions.

Lemma 3 *1. When the total number of tweets exceeds the reading capacity of the consumer ($n_1 + n_2 \geq M$) and both firms post positive number of tweets ($n_1, n_2 > 0$), the relationship between V_1+V_2 and the ratio of posting intensity n_1/n_2 under different conditions is summarized as follows (denote $M_0 = \left|\frac{(\sigma_{10}^2 - \sigma_{20}^2)\sigma_s^2}{\sigma_{10}^2\sigma_{20}^2}\right|$):*

	Conditions		Relationship between $V_1 + V_2$ and $\frac{n_1}{n_2}$
a.	$\sigma_{10}^2 > \sigma_{20}^2$	$0 < M \leq M_0$	*$V_1 + V_2$ increases in $\frac{n_1}{n_2}$*
b.	$\sigma_{10}^2 > \sigma_{20}^2$	$M > M_0$	*$V_1 + V_2$ increases in $\frac{n_1}{n_2}$ when $0 < \frac{n_1}{n_2} \leq \frac{M+M_0}{M-M_0}$, and $V_1 + V_2$ decreases in $\frac{n_1}{n_2}$ when $\frac{n_1}{n_2} > \frac{M+M_0}{M-M_0}$*
c.	$\sigma_{10}^2 < \sigma_{20}^2$	$0 < M \leq M_0$	*$V_1 + V_2$ decreases in $\frac{n_1}{n_2}$*
d.	$\sigma_{10}^2 < \sigma_{20}^2$	$M > M_0$	*$V_1 + V_2$ increases in $\frac{n_1}{n_2}$ when $0 < \frac{n_1}{n_2} \leq \frac{M-M_0}{M+M_0}$, and $V_1 + V_2$ decreases in $\frac{n_1}{n_2}$ when $\frac{n_1}{n_2} > \frac{M-M_0}{M+M_0}$*
e.	$\sigma_{10}^2 = \sigma_{20}^2$	$M > 0$	*$V_1 + V_2$ increases in $\frac{n_1}{n_2}$ when $0 < \frac{n_1}{n_2} \leq 1$, and $V_1 + V_2$decreases in $\frac{n_1}{n_2}$ when $\frac{n_1}{n_2} > 1$*

2. When the total number of posts does not exceed the reading capacity of the consumer ($n_1+n_2 < M$), the value of V_1+V_2 increases when either n_1 or n_2 increases until $n_1 + n_2$ exceeds M.

3. When either firm posts zero tweets in the period ($n_i = 0$ for some i), $V_i = 0$, and V_j ($j \neq i$) increases in n_j until n_j exceeds M. When $n_j \geq M$, the value of V_j keeps the same.

(See Appendix C5 for proof)

Knowing that firm 1 aims at minimizing V_1+V_2 and firm 2 would maximize V_1+V_2, we can find the optimal posting intensities of the two firms based on Lemma 3. We assume that for each firm i ($i = 1, 2$), the number of tweets it can post in a time period ranges from 0 to $\bar{n}_i$, where $\bar{n}_i$ is a firm-specific upper bound determined by factors like working efficiency, the volume of informative contents available in that period, etc. We assume each upper bound $\bar{n}_i$ is greater than the reading capacity M so that we can focus on firms' strategies given limited attention of consumers and keep the conclusion concise. Proposition 2 gives the equilibrium posting intensities of the two firms.

Proposition 2 *Suppose firm 1 is perceived as having a higher quality offering than firm 2 a priori ($\mu_{10} > \mu_{20}$), and the number of tweets firm i can choose to post in a time period ranges from 0 to $\bar{n}_i$. When the prior belief about firm 1 is more precise ($\sigma_{10}^2 < \sigma_{20}^2$) and firm 1 is able to post enough number of tweets to crowd out the messages from the competitor ($\bar{n}_1 > \bar{n}_2\left(\frac{M}{M_0} - 1\right)$), both firms will post intensively to the upper bound ($n_1 = \bar{n}_1$, $n_2 = \bar{n}_2$). Otherwise, firm 1 will not post in the period ($n_1 = 0$), and firm 2 will still post intensively by choosing any $n_2 \in [M, \bar{n}_2]$. Under the current framework, the two firms have no incentive to deviate from the equilibrium considering possible unfollowing of consumers.*

(See Appendix C6 for proof)

The intuition of proposition 2 is as follows. Given limited attention of the consumer, for each firm i, posting more tweets has two effects. First, it makes the consumer read more tweets from its own account (R_i increases), and thus the potential change in the quality belief of show i becomes larger since V_i is larger. We call this "dispersion effect". Second, when the total number of posted tweets exceeds the consumer's limited reading capacity, more tweets from firm i can crowd out the tweets from its competitor, making the potential change in the belief about the competitor smaller as V_j becomes smaller. We call this "crowd-out effect". Firm 1 is perceived as having a higher quality offering *a priori*. It wants to keep the current advantage and hence benefits from smaller $V_1 + V_2$, so firm 1 favors crowd-out effect which makes V_2 smaller, but dislikes dispersion effect which makes V_1 larger. Firm 2 is perceived as having a lesser quality offering. It wants to reverse the current situation. Therefore, firm 2 wants the value of $V_1 + V_2$ to be larger, meaning that firm 2 likes the dispersion effect which leads to a larger V_2, but dislikes crowd-out effect which makes V_1 smaller.

Recall that

$$V_i = \frac{R_i \sigma_{i0}^4}{\sigma_s^2 + R_i \sigma_{i0}^2} = \begin{cases} \frac{\frac{n_i}{n_1+n_2} M \sigma_{i0}^4}{\sigma_s^2 + \frac{n_i}{n_1+n_2} M \sigma_{i0}^2} & n_1 + n_2 \geq M \\ \frac{n_i \sigma_{i0}^4}{\sigma_s^2 + n_i \sigma_{i0}^2} & n_1 + n_2 < M \end{cases}. \tag{3.6}$$

We first consider the situation of $n_1 + n_2 \geq M$. We can see that when n_i changes, the values of both V_i and V_j are affected, but the change in V_i is more salient when σ_{i0}^2 is larger. If the prior belief for firm 1 is more precise than that for firm 2 ($\sigma_{10}^2 < \sigma_{20}^2$), the change in V_2 is more salient than the change in V_1 when either firm posts more tweets. When firm 1 posts more tweets, the change in V_2 being more salient means that the favored crowd-out effect dominates the un-favored dispersion effect, and thus firm 1 tends to post intensively. When firm 2 post more tweets, the change in V_2 being more salient means that the favored dispersion effect dominates the un-favored crowd-out effect, and thus firm 2 also tends to post intensively. Therefore, the equilibrium that both firms post intensively up to their upper bound exists. However, when $\frac{\bar{n}_1}{\bar{n}_2} \leq \frac{M}{M_0} - 1$, meaning that firm 1 does not have the capability to post enough tweets and crowd out enough tweets from the competitor, the overall benefit from crowd-out effect cannot beat the concern of dispersion effect, so the concern of dispersion effect will make firm 1 give up posting in the period. When the prior belief for firm 1 is less precise than that for firm 2 ($\sigma_{10}^2 > \sigma_{20}^2$), the change in V_1 is more salient than that of V_2. For firm 1, the concern of dispersion effect dominates the favor of crowd-out effect, so firm 1 will not post in the period. Given that firm 1 does not post, there is no concern of crowd-out effect for firm 2, and thus firm 2 will post intensively and take up the consumer's attention.

When $n_1 + n_2 < M$, change in n_i only affects the value of V_i, meaning that only dispersion effect exists. Without the concern of crowd-out effect, firm 2 always has the incentive to post more tweets before reaching the reading capacity of the consumer, so the situation that $n_1 + n_2 < M$ will not happen in equilibrium.

We further examine whether the two firms have the incentive to deviate from the equilibrium considering possible unfollowing of consumers. When the equilibrium is $n_1 = 0$ and $n_2 = \bar{n}_2$, consumers may only unfollow firm 2, but they would not do that because firm 1 is not posting and they will get no benefits by unfollowing firm 2. When the equilibrium is $n_1 = \bar{n}_1$ and $n_2 = \bar{n}_2$, since it only happens when $\sigma_{10}^2 < \sigma_{20}^2$, from Proposition 1 we know that consumers may only choose to unfollow account 1. There are two types of consumers. A sophisticated consumer may choose to unfollow an account rationally, and a naive one will not choose to unfollow accounts to optimize the amount of information he can get. Suppose the firms only know that the representative consumer is a sophisticated one with probability p, and a naive one with probability $1 - p$. If the representative consumer is a sophisticated one, to keep him following the account, firm 1 needs to reduce posting intensity such that $n_1 \leq \bar{n}_2(\frac{M}{M_0} - 1)$. However, that will make the value of $V_1 + V_2$ greater than or equal to that of firm 1 being unfollowed. Since firm 1 aims at minimizing the value of $V_1 + V_2$, and a lost follower still knows the existence of the show, firm 1 gains no benefit in retaining the sophisticated consumer by reducing posting. Considering the possibility that the consumer may also be a naive one who will not unfollow the account even if it is posting intensively, firm 1 will not deviate from the strategy.

Finally, let's take a look at the comparative statics of the boundary conditions that make posting intensively to the upper bound an optimal strategy for both firms. According to Proposition 2, when the prior belief about firm 1 is more precise ($\sigma_{10}^2 < \sigma_{20}^2$) and firm 1 is able to post enough number of tweets to crowd out the messages from the competitor ($\bar{n}_1 > \bar{n}_2\left(\frac{M}{M_0} - 1\right)$), both firms will post intensively to the upper bound ($n_1 = \bar{n}_1$, $n_2 = \bar{n}_2$). We notice that in the condition

$$\bar{n}_1 > \bar{n}_2\left(\frac{M}{M_0} - 1\right), \tag{3.7}$$

the right-hand side of the inequality increases with $\bar{n}_2$ and M, meaning that when the competitor is able to post more tweets or the consumer's reading capacity becomes larger, firm 1 needs to be able to post more tweets in order to make the benefit of crowd-out effect stronger than the risk brought by dispersion effect. The right-hand side of inequality (3.7) decreases in M_0. Given $\sigma_{10}^2 < \sigma_{20}^2$, we have

$$M_0 = \left| \frac{(\sigma_{10}^2 - \sigma_{20}^2)\sigma_s^2}{\sigma_{10}^2 \sigma_{20}^2} \right| = \sigma_s^2 \left(\frac{1}{\sigma_{10}^2} - \frac{1}{\sigma_{20}^2} \right), \tag{3.8}$$

so the right-hand side of (3.7) decreases in σ_s^2, increases in σ_{10}^2, and decreases in σ_{20}^2. The intuition is that if σ_s^2 is larger, it means the signals conveyed by tweets are less precise of the true quality and the consumer will put less weight on the signals when updating his beliefs. The lower weight on signals makes it easier for firm 1 to keep its advantage in the prior belief, and thus the requirement on firm 1's posting capability becomes looser. If σ_{10}^2 is larger, it means the consumer is less certain about his prior belief about firm 1, and hence firm 1's advantage in prior belief becomes less stable, and the consumer will put more weight on the signals when updating his belief about firm 1. As a result, the potential risk brought by dispersion effect becomes larger, and firm 1 needs to be able to post more intensively to make the crowd-out effect stronger than the dispersion effect. Similarly, if σ_{20}^2 is larger, the consumer is less certain about his prior belief about firm 2, and he will put more weight on the signals when updating his belief about firm 2. Then the benefit from the crowd-out effect becomes stronger, and therefore the threshold of firm 1's capability to make the crowd-out effect stronger than the dispersion effect becomes lower.

3.4 Data Revisited: Implication on Tweet Informativeness

According to our analysis in §3.3, if firms can manipulate how informative the tweets are about the show, then firm 1 may want to post more tweets that are not directly related to the show, as they have the crowd-out effect but suffer less from dispersion effect, and firm 2 would not do so.

We test this implication empirically by comparing the proportion of retweets posted by shows with different levels of popularity (shows are classified into two groups according to Baidu search index). We find the proportion of retweets posted by popular shows is nearly twice as high as that of non-popular shows (M_{pop}=0.26, $M_{non-pop}$=0.14, $t(10719) = 20.2$, $p < 0.001$). A retweet is a reposting of a tweet that has been posted by others or by itself. A retweet from other accounts is usually not directly related to the content of the show, and a retweet from itself is repetitive information, so a retweet is generally less informative of the show compared to an original tweet. Therefore, the fact that popular shows are much more likely to post retweets is consistent with the model's implication, and it provides supporting evidence that firms can be strategic in posting.

Distinguishing retweets and original tweets is a rough way of identifying the informativeness of tweets. We may also use machine learning methods to identify the informativeness of tweets and test if popular shows do tend to post tweets that are less informative of show content.

Chapter 4

Concluding Remarks

In this paper, we study competitive firms' advertising strategy given limited attention of consumers, using the advertising of TV shows on a social media platform as an example. The data we collected from a primary Chinese tweeting website provides evidence that firms advertise intensively, although doing so appears to drive followers away. The analytical model we build suggests that consumers with limited attention may choose to unfollow a firm-owned account rationally if he already knows enough about the firm's product quality and the firm advertises too intensively. However, advertising intensively can be an optimal strategy for competitive firms, since the firm perceived as having a lesser quality offering wants to change consumers beliefs about its quality, and the firm perceived as having a higher quality offering would like to crowd out tweets from the competitor and keep its advantage.

A limitation of our model is that we only consider the role of advertising in influencing consumers' beliefs about product quality, but do not consider the effect of advertising in informing more people of the product. We do this for two reasons. First, usually people who follow a firm account are already aware of the products (or the brands), so the main effect of posts made by firm accounts is to spread product-related information rather than to arouse awareness. Although a follower may repost the messages and make more people informed of the show, the effect is hard to quantify. Second, the current framework allows us to find a clear relationship between the firms' objective functions and their posting intensity, and the problem will be intractable if we also incorporate the role of advertising in informing more people of the show. Even if we take that effect into account, it is likely that the effect will be an additional reason for the firms to advertise intensively.

Appendix A

Figures

Figure A-1: Change in the Number of Followers on the Air Date

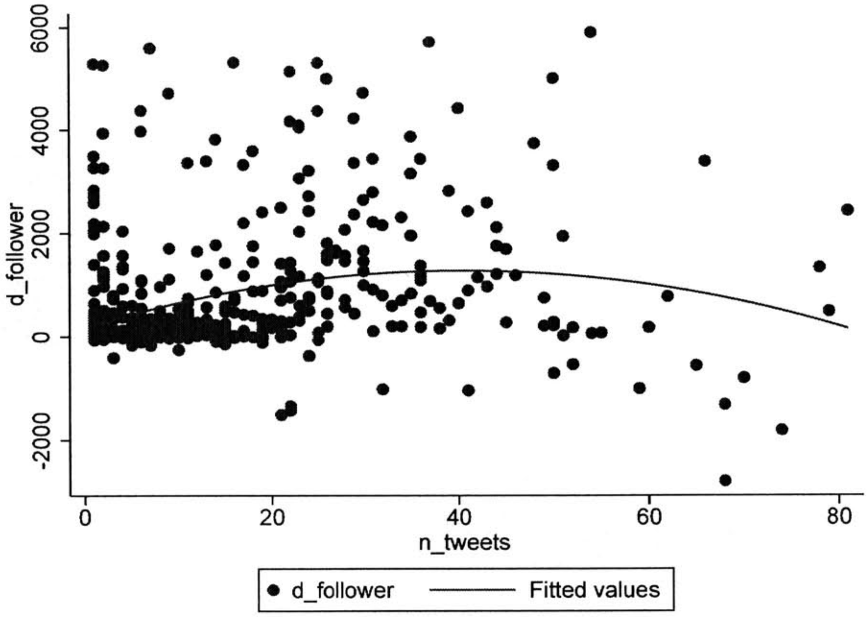
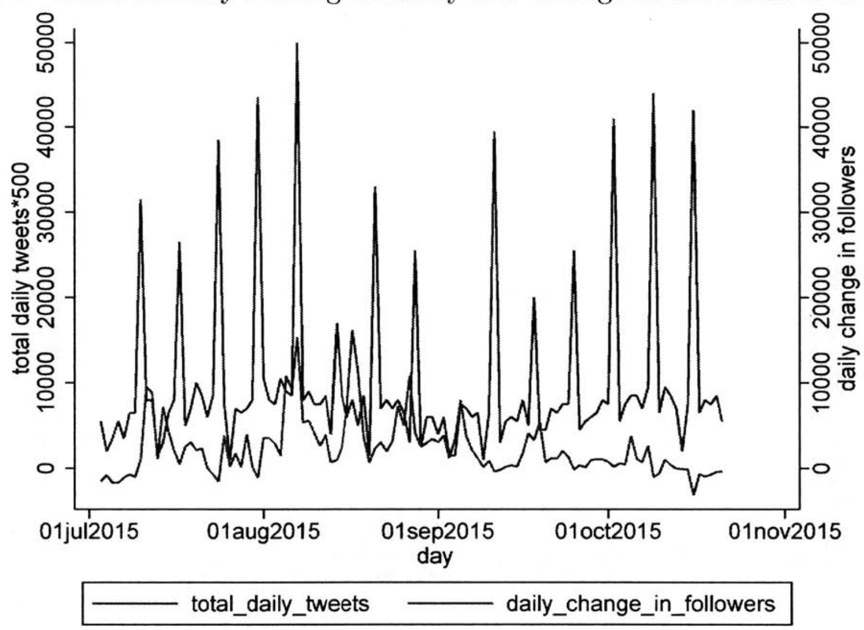


Figure A-2: Trend of Daily Posting Intensity and Change in the Number of Followers



Appendix B

Tables

Table B.1: Summary Statistics: Each Account over the Six-Month Window

	Mean	Std Dev	Min	Max	N
Avg. # of Followers	540,680.43	1,198,894.69	2,141	8,569,499	93
Total # of Tweets	597.18	553.92	30	2,463	93
Max. # of Tweets per 4-hour	20.86	18.14	2	81	93
Avg. Daily Baidu Index	24,827.50	46,906.35	155	283,246	93

"Avg." means average over the six months.

Table B.2: Summary Statistics: Dynamics of Each Account per Four-Hour Window

	Mean	Std Dev	Min	Max	N
Change in # of followers	246.78	720.89	-2932	5977	16999
Percentage Change	8.09	28.18	-125	461	16999
Number of Tweets	2.79	4.16	1	81	16999

Table B.3: Posting Intensity on the Air Date vs. Off

Air Date	Mean	Std Dev	Min	Max	p5	p95	N
0	2.32	1.75	1	35	1	5	8837
1	6.76	9.77	1	81	1	26	1884
Total	3.10	4.70	1	81	1	8	10721
N	10721						

Table B.4: Posting Intensity at Different Times of a Day

Time of Day	Mean	Std Dev	Min	Max	p5	p95	N
12am-4am	1.51	1.17	1	8	1	4	256
4am-8am	1.11	0.32	1	2	1	2	35
8am-12pm	2.26	1.36	1	14	1	5	2815
12pm-4pm	2.80	1.95	1	15	1	7	2869
4pm-8pm	2.72	2.27	1	22	1	7	2613
8pm-12am	5.30	9.53	1	81	1	25	2133
Total	3.10	4.70	1	81	1	8	10721
N	10721						

Table B.5: Posting Intensity at Night for Shows with Different Avg. Baidu Index

Quartile of Baidu Index	Mean	Std Dev	Min	Max	p5	p95	N
1	4.03	6.73	1	55	1	14	633
2	5.19	8.41	1	78	1	23	506
3	6.45	12.23	1	81	1	31	564
4	5.77	10.04	1	66	1	34	430
Total	5.30	9.53	1	81	1	25	2133
N	2133						

Table B.6: Posting Intensity at Night for Shows with Different Avg. Number of Followers

Quartile of Avg. Number of Followers	Mean	Std Dev	Min	Max	p5	p95	N
1	6.16	9.20	1	79	1	24	461
2	4.77	7.77	1	44	1	25	638
3	6.31	11.07	1	78	1	34	433
4	4.47	10.16	1	81	1	16	601
Total	5.30	9.53	1	81	1	25	2133
N	2133						

Table B.7: Change in the Number of Followers every Four Hours

	(1)	(2)	(3)	(4)	(5)	(6)
	d_follower	d_follower	d_follower	d_follower	d_follower	d_follower
n_tweets	72.77***	36.58***	35.77***	33.54***	37.79***	30.98***
	(1.128)	(1.487)	(1.416)	(1.589)	(1.455)	(1.648)
n_tweets_sqr	-0.981***	-0.530***	-0.482***			
	(0.0295)	(0.0320)	(0.0301)			
Show Characteristics						
search_index_yesterday		0.000541***	0.000411***	0.000402***	0.000409***	0.000404***
		(0.0000193)	(0.0000198)	(0.0000198)	(0.0000198)	(0.0000198)
avg_search		0.00177***	0.409***	0.418***	0.414***	0.412***
		(0.0000467)	(0.0820)	(0.0821)	(0.0819)	(0.0821)
show_day		18.84***	17.92***	38.21***		
		(4.550)	(4.327)	(4.254)		
show_time		7.552	4.614		102.5***	
		(11.26)	(11.18)		(11.04)	
lifetime_of_show		-0.0285***	0.386**	0.405**	0.395**	0.394**
		(0.00158)	(0.179)	(0.179)	(0.179)	(0.179)
Tweet Characteristics						
total_reposted		0.0658***	0.0509***	0.0484***	0.0491***	0.0467***
		(0.00308)	(0.00289)	(0.00288)	(0.00288)	(0.00288)
avg_comment		0.700***	0.649***	0.658***	0.654***	0.667***
		(0.0216)	(0.0211)	(0.0211)	(0.0211)	(0.0211)
percent_pic		-61.92***	-68.04***	-62.33***	-65.73***	-59.66***
		(8.936)	(8.655)	(8.651)	(8.634)	(8.646)
percent_retweet		67.91***	26.83***	33.70***	22.25**	35.35***
		(9.885)	(9.492)	(9.498)	(9.513)	(9.524)
percent_retweet_self		-63.56***	-28.56**	-21.45*	-29.60**	-22.11*
		(13.23)	(12.43)	(12.43)	(12.43)	(12.45)
percent_topic		69.79***	56.14***	61.66***	55.31***	61.24***
		(7.848)	(7.881)	(7.874)	(7.872)	(7.885)
avg_length		-0.0891***	-0.0666**	-0.0543**	-0.0728***	-0.0472*
		(0.0272)	(0.0264)	(0.0264)	(0.0264)	(0.0264)
show_day=1 × n_tweets				-18.01***		
				(1.587)		
show_time=1 × n_tweets					-26.22***	
					(1.546)	
night						49.12***
						(5.563)
night=1 × n_tweets						-13.99***
						(1.638)
Month	No	No	Yes	Yes	Yes	Yes
Day of Week	No	No	Yes	Yes	Yes	Yes
Hour	No	No	Yes	Yes	Yes	Yes
Show	No	No	Yes	Yes	Yes	Yes
Observations	57745	57745	57745	57745	57745	57745
Adjusted R^2	0.081	0.237	0.350	0.349	0.351	0.348

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Appendix C

Proofs

C1. Proof of Lemma 1

1. We will prove that when $\sigma_1^2 + \sigma_2^2$ is smaller, the probability of choosing the show with higher true quality will be larger.

When the consumer decides which show to watch, his belief about the true quality of show i is $Q_i \sim N(\mu_i, \sigma_i^2)$. The interpretation of this belief is that, the true quality Q_i lies somewhere around μ_i, and the difference between the true quality Q_i and μ_i is a normally distributed random variable, follows $N(0, \sigma_i^2)$. We can also understand it as, μ_i lies somewhere around the true quality with a random error following $N(0, \sigma_i^2)$, that is, $\mu_i \sim N(Q_i, \sigma_i^2)$.

Suppose the reality is that show 1 has higher true quality than show 2 ($Q_1 > Q_2$), then the consumer can make a right choice if and only if $\mu_1 > \mu_2$ since he would choose the show according to the value of μ_i. Therefore, the probability that the consumer can make a right choice is $P(\mu_1 > \mu_2) = P(\mu_1 - \mu_2 > 0)$.

Notice that we assume the beliefs and signals of the two shows are independent, then $\mu_1 - \mu_2$ is also a normal random variable following $N(Q_1 - Q_2, \sigma_1^2 + \sigma_2^2)$. Then

$$\begin{aligned} P(\mu_1 - \mu_2 > 0 | Q_1, Q_2, \sigma_1^2, \sigma_2^2) &= P(\frac{(\mu_1 - \mu_2) - (Q_1 - Q_2)}{\sqrt{\sigma_1^2 + \sigma_2^2}} > -\frac{Q_1 - Q_2}{\sqrt{\sigma_1^2 + \sigma_2^2}}) && \text{(C.1)} \\ &= \Phi(\frac{Q_1 - Q_2}{\sqrt{\sigma_1^2 + \sigma_2^2}}), && \text{(C.2)} \end{aligned}$$

where $\Phi(\cdot)$ is the CDF of standard normal distribution. Equation (C.2) comes from the fact that $\frac{(\mu_1-\mu_2)-(Q_1-Q_2)}{\sqrt{\sigma_1^2+\sigma_2^2}}$ follows standard normal distribution. Since $\Phi(\cdot)$ is an increasing function and $Q_1 - Q_2$ is supposed to be positive here, the probability $P(\mu_1 > \mu_2)$ decreases with $\sigma_1^2 + \sigma_2^2$.

Symmetrically, if show 2 has higher true quality ($Q_2 > Q_1$), the probability of making the right choice is $P(\mu_2 > \mu_1) = \Phi(\frac{Q_2-Q_1}{\sqrt{\sigma_1^2+\sigma_2^2}})$, which also decreases with $\sigma_1^2 + \sigma_2^2$.

2. Intuitively, a larger probability of choosing the show with higher true quality will lead to a higher expected utility from watching the show that will be chosen, and we will prove this in now.

Standing at the beginning of period 0, with prior belief $Q_i \sim N(\mu_{i0}, \sigma_{io}^2)$, the expected utility from the show that will be watching is

$$\begin{aligned} E[Q_i : \mu_i > \mu_j] &= \int [P(\mu_1 > \mu_2)Q_1 + P(\mu_2 > \mu_1)Q_2]dG(Q_1)dG(Q_2) \\ &= \int [\Phi(\frac{Q_1 - Q_2}{\sqrt{\sigma_1^2 + \sigma_2^2}})Q_1 + \Phi(\frac{Q_2 - Q_1}{\sqrt{\sigma_1^2 + \sigma_2^2}})Q_2]dG(Q_1)dG(Q_2) \quad \text{(C.3)} \end{aligned}$$

where $G(Q_i)$ denotes the CDF of prior belief, $N(\mu_{i0}, \sigma_{io}^2)$. We separate the integrals in equation (C.3) into two parts, cases of $Q_1 > Q_2$ and of $Q_2 > Q_1$, then we have

$$\begin{aligned} E[Q_i : \mu_i > \mu_j] &= \int_{Q_1>Q_2} [\Phi(\frac{Q_1 - Q_2}{\sqrt{\sigma_1^2 + \sigma_2^2}})Q_1 + \Phi(\frac{(Q_2 - Q_1)}{\sqrt{\sigma_1^2 + \sigma_2^2}})Q_2]dG(Q_1)dG(Q_2) \\ &\quad + \int_{Q_2>Q_1} [\Phi(\frac{Q_1 - Q_2}{\sqrt{\sigma_1^2 + \sigma_2^2}})Q_1 + \Phi(\frac{(Q_2 - Q_1)}{\sqrt{\sigma_1^2 + \sigma_2^2}})Q_2]dG(Q_1)dG(Q_2) \\ &= \int_{Q_1>Q_2} [Q_2 + (Q_1 - Q_2)\Phi(\frac{Q_1 - Q_2}{\sqrt{\sigma_1^2 + \sigma_2^2}})]dG(Q_1)dG(Q_2) \\ &\quad + \int_{Q_2>Q_1} [Q_1 + (Q_2 - Q_1)\Phi(\frac{Q_2 - Q_1}{\sqrt{\sigma_1^2 + \sigma_2^2}})]dG(Q_1)dG(Q_2) \quad \text{(C.4)} \end{aligned}$$

Let's observe equation (C.4). In the first part of the integral, $Q_1 - Q_2 > 0$, and $G(Q_1), G(Q_2)$ is fixed with the prior belief given, so the integrand decreases with $\sigma_1^2 + \sigma_2^2$. In the second part of the integral, $Q_2 - Q_1 > 0$, so similarly we have the integrand decreasing with $\sigma_1^2 + \sigma_2^2$. Therefore, if $\sigma_1^2 + \sigma_2^2$ gets smaller, the expected utility from the show that will be chosen gets larger.

C2. $\mu_i \sim N(\mu_{i0}, V_i)$

Recall equation (3.1) in the main part,

$$\mu_i = \frac{\sigma_s^2}{\sigma_s^2 + R_i\sigma_{i0}^2}\mu_{i0} + \frac{R_i\sigma_{i0}^2}{\sigma_s^2 + R_i\sigma_{i0}^2}\bar{S}_i, \ (\bar{S}_i = \frac{\sum_{k=1}^{R_i} S_{ik}}{R_i}). \tag{C.5}$$

At the time when the consumer has not read the tweets, μ_{i0}'s are known, $S_i \sim N(Q_i, \sigma_s^2)$, so $\bar{S}_i \sim N(Q_i, \sigma_s^2/R_i)$. Given that the belief about Q_i is $N(\mu_{i0}, \sigma_{i0}^2)$, we know the belief about $\bar{S}_i$ is $N(\mu_{i0}, \sigma_{i0}^2 + \frac{\sigma_s^2}{R_i})$. According to equation (3.1), we know the posterior belief μ_i follows a normal distribution. The mean is

$$\mathbb{E}\left[\frac{\sigma_s^2}{\sigma_s^2 + R_i\sigma_{i0}^2}\mu_{i0} + \frac{R_i\sigma_{i0}^2}{\sigma_s^2 + R_i\sigma_{i0}^2}\bar{S}_i\right] = \frac{\sigma_s^2}{\sigma_s^2 + R_i\sigma_{i0}^2}\mu_{i0} + \frac{R_i\sigma_{i0}^2}{\sigma_s^2 + R_i\sigma_{i0}^2}\mu_{i0} = \mu_{i0}, \tag{C.6}$$

and the variance is

$$Var\left[\frac{\sigma_s^2}{\sigma_s^2 + R_i\sigma_{i0}^2}\mu_{i0} + \frac{R_i\sigma_{i0}^2}{\sigma_s^2 + R_i\sigma_{i0}^2}\bar{S}_i\right] = \left(\frac{R_i\sigma_{i0}^2}{\sigma_s^2 + R_i\sigma_{i0}^2}\right)^2 \left(\sigma_{i0}^2 + \frac{\sigma_s^2}{R_i}\right)^2 \tag{C.7}$$

$$= \frac{R_i\sigma_{i0}^4}{\sigma_s^2 + R_i\sigma_{i0}^2}. \tag{C.8}$$

C3. Proof of Lemma 2

Denote $X \equiv \sqrt{V_1 + V_2}$, then

$$\begin{aligned} h(X) &\equiv E[\max\{\mu_1, \mu_2\}] \\ &= \mu_{10}\Phi(\frac{\mu_{10} - \mu_{20}}{\sqrt{V_1 + V_2}}) + \mu_{20}\Phi(\frac{\mu_{20} - \mu_{10}}{\sqrt{V_1 + V_2}}) + \sqrt{V_1 + V_2}\phi(\frac{\mu_{10} - \mu_{20}}{\sqrt{V_1 + V_2}}) \\ &= \mu_{10}\Phi(\frac{\mu_{10} - \mu_{20}}{X}) + \mu_{20}\Phi(\frac{\mu_{20} - \mu_{10}}{X}) + X\phi(\frac{\mu_{10} - \mu_{20}}{X}). \end{aligned} \tag{C.9}$$

Calculate the derivative of $h(X)$ with respect to X, we get $f'(X) = \phi(\frac{\mu_{10} - \mu_{20}}{X})$, which is always positive. Therefore, $E[max\{\mu_1, \mu_2\}]$ increases with $V_1 + V_2$.

C4. Proof of Equivalence

In order to prove that two interpretations are equivalent, we just need to show, for any (R_1, R_2), (R_1', R_2'), $\sigma_1^2(R_1) + \sigma_2^2(R_2) < \sigma_1^2(R_1') + \sigma_2^2(R_2')$ is equivalent to $V_1(R_1) + V_2(R_2) > V_1(R_1') + V_2(R_2')$.

In fact, for any R_i and R_i' $(i = 1, 2)$, $\sigma_i^2(R_i) < \sigma_i^2(R_i')$ is equivalent to $V_i(R_i) > V_i(R_i')$, because

$$
\begin{aligned}
& \sigma_i^2(R_i) < \sigma_i^2(R_i') \\
\Leftrightarrow\quad & \frac{\sigma_{i0}^2\sigma_s^2}{\sigma_s^2 + R_i\sigma_{i0}^2} < \frac{\sigma_{i0}^2\sigma_s^2}{\sigma_s^2 + R_i'\sigma_{i0}^2} \\
\Leftrightarrow\quad & \frac{\sigma_{i0}^4\sigma_s^2(R_i - R_i')}{(\sigma_s^2 + R_i'\sigma_{i0}^2)(\sigma_s^2 + R_i\sigma_{i0}^2)} > 0,
\end{aligned}
$$

and

$$
\begin{aligned}
& V_i(R_i) > V_i(R_i') \\
\Leftrightarrow\quad & \frac{R_i\sigma_{i0}^4}{\sigma_s^2 + R_i\sigma_{i0}^2} > \frac{R_i'\sigma_{i0}^4}{\sigma_s^2 + R_i'\sigma_{i0}^2} \\
\Leftrightarrow\quad & \frac{\sigma_{i0}^4\sigma_s^2(R_i - R_i')}{(\sigma_s^2 + R_i'\sigma_{i0}^2)(\sigma_s^2 + R_i\sigma_{i0}^2)} > 0.
\end{aligned}
$$

Therefore, for any (R_1, R_2), (R_1', R_2'), $\sigma_1^2(R_1) + \sigma_2^2(R_2) < \sigma_1^2(R_1') + \sigma_2^2(R_2')$ is equivalent to $V_1(R_1) + V_2(R_2) > V_1(R_1') + V_2(R_2')$.

C5. Proof of Lemma 3

1. Suppose $n_1 + n_2 \geq M$, and $n_1, n_2 > 0$. Then $R_i = \frac{n_i}{n_1+n_2}M$ when the consumer is following both accounts. Denote the ratio of posting intensity is $\psi = \frac{n_1}{n_2}$ $(\psi > 0)$, and

$$
f(\psi) = V_1 + V_2 = \frac{\frac{\psi}{1+\psi}M\sigma_{10}^4}{\sigma_s^2 + \frac{\psi}{1+\psi}M\sigma_{10}^2} + \frac{\frac{1}{1+\psi}M\sigma_{20}^4}{\sigma_s^2 + \frac{1}{1+\psi}M\sigma_{20}^2}.
$$

Table C.1: The Relationship Between $f = V_1 + V_2$ and $\psi = \frac{n_1}{n_2}$

Conditions			Relationship between f and ψ
a.	$\sigma_{10}^2 > \sigma_{20}^2$	$0 < M \le M_0$	f increases in ψ
b.	$\sigma_{10}^2 > \sigma_{20}^2$	$M > M_0$	f increases in ψ when $0 \le \psi \le \frac{M+M_0}{M-M_0}$, and f decreases in ψ when $\psi > \frac{M+M_0}{M-M_0}$
c.	$\sigma_{10}^2 < \sigma_{20}^2$	$0 < M \le M_0$	f decreases in ψ
d.	$\sigma_{10}^2 < \sigma_{20}^2$	$M > M_0$	f increases in ψ when $0 \le \psi \le \frac{M-M_0}{M+M_0}$, and f decreases in ψ when $\psi > \frac{M-M_0}{M+M_0}$
e.	$\sigma_{10}^2 = \sigma_{20}^2$	$M > 0$	f increases in ψ when $0 \le \psi \le 1$, and f decreases in ψ when $\psi > 1$

The derivative of f with respect to ψ is

$$f'(\psi) = \frac{M(1+\psi)\sigma_s^2[M\sigma_{10}^2\sigma_{20}^2 + (\sigma_{10}^2+\sigma_{20}^2)\sigma_s^2]}{(M\psi\sigma_{10^2} + \psi\sigma_s^2 + \sigma_s^2)^2} \tag{C.10}$$

$$\cdot \frac{(1+\psi)(\sigma_{10}^2 - \sigma_{20}^2)\sigma_s^2 - M(\psi-1)\sigma_{10}^2\sigma_{20}^2}{(M\psi\sigma_{20^2} + \psi\sigma_s^2 + \sigma_s^2)^2} \tag{C.11}$$

The denominator is positive and $M(1+\psi)\sigma_s^2[M\sigma_{10}^2\sigma_{20}^2 + (\sigma_{10}^2+\sigma_{20}^2)\sigma_s^2] > 0$, so the sign of $f'(\psi)$ is the same as that of $[(1+\psi)(\sigma_{10}^2-\sigma_{20}^2)\sigma_s^2 - M(\psi-1)\sigma_{10}^2\sigma_{20}^2]$. Denote

$$h = (1+\psi)(\sigma_{10}^2-\sigma_{20}^2)\sigma_s^2 - M(\psi-1)\sigma_{10}^2\sigma_{20}^2 \tag{C.12}$$

$$= \left(\frac{(\sigma_{10}^2-\sigma_{20}^2)\sigma_s^2}{\sigma_{10}^2\sigma_{20}^2} - M\right)\psi + \left(\frac{(\sigma_{10}^2-\sigma_{20}^2)\sigma_s^2}{\sigma_{10}^2\sigma_{20}^2} + M\right), \tag{C.13}$$

and $M_0 = \left|\frac{(\sigma_{10}^2-\sigma_{20}^2)\sigma_s^2}{\sigma_{10}^2\sigma_{20}^2}\right|$. Notice that $\psi = \frac{n_1}{n_2} > 0$.

When $\sigma_{10}^2 > \sigma_{20}^2$, $h = (M_0 - M)\psi + (M_0 + M)$. When $0 < M \le M_0$, $h > 0$ for any ψ. When $M > M_0$, $h > 0$ if and only if $\psi < \frac{M+M_0}{M-M_0}$.

When $\sigma_{10}^2 < \sigma_{20}^2$, $h = (M - M_0) - (M + M_0)\psi$. When $0 < M \le M_0$, $h < 0$ for any ψ. When $M > M_0$, $h > 0$ if and only if $\psi < \frac{M-M_0}{M+M_0}$.

Then we can get the relationship between $f = V_1 + V_2$ and $\psi = \frac{n_1}{n_2}$, as summarized in Table C.1.

2. When $n_1 + n_2 < M$, $R_i = n_i$.

$$V_1 + V_2 = \frac{n_1\sigma_{10}^4}{\sigma_s^2 + n_1\sigma_{10}^2} + \frac{n_2\sigma_{20}^4}{\sigma_s^2 + n_2\sigma_{20}^2}.$$

We can see that now a marginal increase in n_i only leads to an increase in V_i without affecting the value of V_j ($j \neq i$). Therefore, when either n_1 or n_2 increases, $V_1 + V_2$ increases, until $n_1 + n_2$ exceeds M.

3. When either $n_1$or n_2 equals 0. Without loss of generalizability, assume $n_1 = 0$, then $V_1 = 0$,

$$f = V_2 = \frac{\min\{n_2, M\}\sigma_{20}^4}{\sigma_s^2 + \min\{n_2, M\}\sigma_{20}^2}.$$

When $n_2 < M$, a marginal increase in $n_2$will lead to an increase in f. When $n_2 \geq M$, the value of f will keep at $\frac{M\sigma_{20}^4}{\sigma_s^2+M\sigma_{20}^2}$.

C6. Proof of Proposition 2

Given that $\mu_{10} > \mu_{20}$, as we have analyzed in the main part, the objective of firm 1 is to minimize $V_1 + V_2$, and the objective of firm 2 is to maximize $V_1 + V_2$. Knowing the objectives of the two firms, we analyze their posting intensities in equilibrium under each condition as listed in Table C.1.

First of all, when $n_1 + n_2 < M$, $V_1 + V_2$ strictly increases with either $n_1$or n_2 increases. Since firm 2 always prefers $V_1 + V_2$ to be, it always has incentive to increase posting before reaching the upper bound. Given that $\bar{n}_2 > M$, the situation that $n_1 + n_2 < M$ will never exist in equilibrium.

a. $\sigma_{10}^2 > \sigma_{20}^2$, and $0 < M \leq M_0$.

As we have shown, $V_1 + V_2$ strictly increases with $\frac{n_1}{n_2}$ strictly when $n_1 + n_2 \geq M$ and $n_1, n_2 > 0$, so firm 1 will reduce posting intensity to make $\frac{n_1}{n_2}$ smaller, and firm 2 will also reduce posting intensity to make $\frac{n_1}{n_2}$ larger. Since both firms have incentives to keep reducing their posting intensities, the situation that $n_1 + n_2 \geq M$ and $n_1, n_2 > 0$ will not happen in equilibrium.

As we have said, when both firms reduce posting intensities such that $n_1 + n_2 < M$, $V_1 + V_2$ will decrease when either n_1 or n_2 decreases. Firm 1 still has the incentive to reduce posting intensity, so in the end firm 1 will post $n_1 = 0$. Given that firm 1 posts zero in the period, according to our conclusion in Lemma 3, scenario 3, $V_1 + V_2$ will increase with n_2 when $n_2 < M$, and then keep the same value when $n_2 \geq M$.

Therefore, firm 2 will at least post up to M, and it is indifferent between posting M and posting up to $\bar{n}_2$.

b. $\sigma_{10}^2 > \sigma_{20}^2$, and $M > M_0$.

As we have shown in Lemma 3, when $n_1+n_2 \geq M$ and $n_1, n_2 > 0$, $f(\frac{n_1}{n_2}) = V_1+V_2$ increases in $\frac{n_1}{n_2}$ when $0 < \frac{n_1}{n_2} \leq \frac{M+M_0}{M-M_0}$, and and $f(\frac{n_1}{n_2}) = V_1 + V_2$ decreases in $\frac{n_1}{n_2}$ when $\frac{n_1}{n_2} > \frac{M+M_0}{M-M_0}$. In other words, on $\frac{n_1}{n_2} \in (0, \frac{M+M_0}{M-M_0}]$, the $f(\frac{n_1}{n_2})$ approaches its minimum when $\frac{n_1}{n_2} \to 0$, and on $\frac{n_1}{n_2} \in (\frac{M+M_0}{M-M_0}, \infty)$, $f(\frac{n_1}{n_2})$ approaches its minimum when $\frac{n_1}{n_2} \to \infty$.

Notice that

$$\lim_{\frac{n_1}{n_2} \to 0} f(\frac{n_1}{n_2}) = \frac{M\sigma_{20}^4}{\sigma_s^2 + M\sigma_{20}^2} < \frac{M\sigma_{10}^4}{\sigma_s^2 + M\sigma_{10}^2} = \lim_{\frac{n_1}{n_2} \to \infty} f(\frac{n_1}{n_2})$$

when $\sigma_{10}^2 > \sigma_{20}^2$, so on $(0, \infty)$ $f(\frac{n_1}{n_2})$ approaches its minimum when $\frac{n_1}{n_2} \to 0$.

When $n_1 + n_2 < M$, $V_1 + V_2$ strictly increases when either n_1 or n_2 increases.

Therefore, whether $n_1 + n_2 > M$ or $n_1 + n_2 \leq M$, firm 1 will always benefit from reducing posting till $n_1 = 0$. Similar as in condition a., given that firm 1 posts zero, firm 2 will post intensively to take up all the attention of the consumer, and it is indifferent between sending M to $\bar{n}_2$ tweets.

c. $\sigma_{10}^2 < \sigma_{20}^2$, and $0 < M \leq M_0$.

As we have shown, $V_1 + V_2$ strictly decreases with $\frac{n_1}{n_2}$ when $n_1 + n_2 \geq M$ and $n_1, n_2 > 0$, so firm 1 will increase posting intensity to make $\frac{n_1}{n_2}$ larger, and firm 2 will also increase posting intensity to make $\frac{n_1}{n_2}$ smaller. Hence both firm 1 and firm 2 will post up to their upper bound, i.e. $n_1 = \bar{n}_1$, and $n_2 = \bar{n}_2$.

As we have shown, the scenario that $n_1 + n_2 < M$ will not happen in equilibrium. Given $n_1 + n_2 \geq M$, for any n_2, a positive $\frac{n_1}{n_2}$ will always lead to smaller value of $V_1 + V_2$ compared to that of $n_1 = 0$, and thus firm 1 will never choose $n_1 = 0$. Firm 2 will not choose $n_2 = 0$ since any positive n_2 will lead to a larger value of $V_1 + V_2$.

Therefore, the only possibe equilibrium under the condition $\sigma_{10}^2 < \sigma_{20}^2$, and $0 < M \leq M_0$ is that $n_1 = \bar{n}_1$ and $n_2 = \bar{n}_2$.

d. $\sigma_{10}^2 < \sigma_{20}^2$, and $M > M_0$.

When $n_1 + n_2 \geq M$ and $n_1, n_2 > 0$, as we have shown, $V_1 + V_2$ increases in $\frac{n_1}{n_2}$

when $0 < \frac{n_1}{n_2} \le \frac{M-M_0}{M+M_0}$, and $V_1 + V_2$ decreases in $\frac{n_1}{n_2}$ when $\frac{n_1}{n_2} > \frac{M-M_0}{M+M_0}$, so firm 1 will either choose $n_1 \to 0$ or $n_1 = \bar{n}_1$, whichever leads to smaller $V_1 + V_2$.

Suppose$n_1 = \bar{n}_1$, firm 2 will get maximum value of $V_1 + V_2$ when $n_2 = \frac{M+M_0}{M-M_0}\bar{n}_1$, but when $\frac{M+M_0}{M-M_0}\bar{n}_1 > \bar{n}_2$, it is not achievable. Given that $V_1 + V_2$ increases in $\frac{n_1}{n_2}$ when $0 < \frac{n_1}{n_2} \le \frac{M-M_0}{M+M_0}$, firm 2 will post $n_2 = \bar{n}_2$ then. So $n_2 = \min\{\frac{M+M_0}{M-M_0}\bar{n}_1, \bar{n}_2\}$. (i) When $\frac{M+M_0}{M-M_0}\bar{n}_1 < \bar{n}_2$, we have $n_2 = \frac{M+M_0}{M-M_0}\bar{n}_1$and $n_1 = \bar{n}_1$, so $\frac{n_1}{n_2} = \frac{M-M_0}{M+M_0}$. However, knowing that

$$f(\frac{M - M_0}{M + M_0}) > \lim_{\frac{n_1}{n_2} \to 0} f(\frac{n_1}{n_2}),$$

firm 1 has incentive to deviate to $n_1 \to 0$, so this will not be the equilibrium. (ii) When $\frac{M+M_0}{M-M_0}\bar{n}_1 \ge \bar{n}_2$, we have $n_2 = \bar{n}_2$. Now

$$\frac{n_1}{n_2} = \frac{\bar{n}_1}{\bar{n}_2} \ge \frac{M - M_0}{M + M_0}.$$

Firm 1 has no incentive to deviate to $n_1 \to 0$ as long as

$$f(\frac{n_1}{n_2}) \le \lim_{\frac{n_1}{n_2} \to 0} f(\frac{n_1}{n_2}) = \frac{M\sigma_{20}^4}{\sigma_s^2 + M\sigma_{20}^2},$$

which is equivalent to

$$\frac{n_1}{n_2} \ge \frac{M - M_0}{M_0}.$$

Therefore in order to have $n_1 = \bar{n}_1$ and $n_2 = \bar{n}_2$ as the equilibrium, we need to have

$$\frac{\bar{n}_1}{\bar{n}_2} \ge \max\left\{\frac{M - M_0}{M + M_0}, \frac{M - M_0}{M_0}\right\} = \frac{M - M_0}{M_0} = \frac{M}{M_0} - 1.$$

To summarize, when $\bar{n}_1/\bar{n}_2 \ge \frac{M}{M_0} - 1$, both firms post up to their upper bounds ($n_1 = \bar{n}_1$ and $n_2 = \bar{n}_2$) in equilibrium. Otherwise, firm 1 will post zero ($n_1 = 0$), and the same as in condition a, b, firm 2 may choose any n_2 between $[M, \bar{n}_2]$.

e. $\sigma_{10}^2 = \sigma_{20}^2$, $M > 0$.

When $n_1 + n_2 \ge M$ and $n_1, n_2 > 0$, $V_1 + V_2$ increases in $\frac{n_1}{n_2}$ when $0 < \frac{n_1}{n_2} \le 1$, and $V_1 + V_2$ decreases in $\frac{n_1}{n_2}$ when $\frac{n_1}{n_2} > 1$, so firm 1 will either choose $n_1 \to 0$ or $n_1 = \bar{n}_1$, whichever leads to smaller $V_1 + V_2$. We notice that

$$\lim_{\frac{n_1}{n_2}\to 0} f(\frac{n_1}{n_2}) = \frac{M\sigma_{20}^4}{\sigma_s^2 + M\sigma_{20}^2} = \frac{M\sigma_{10}^4}{\sigma_s^2 + M\sigma_{10}^2} = \lim_{\frac{n_1}{n_2}\to\infty} f(\frac{n_1}{n_2}).$$

Since $f(\frac{n_1}{n_2})$ increases $\frac{n_1}{n_2}$ when $0 < \frac{n_1}{n_2} \leq 1$and decreases in $\frac{n_1}{n_2}$ when $\frac{n_1}{n_2} > 1$, for any finite value of $\frac{n_1}{n_2}$,

$$\lim_{\frac{n_1}{n_2}\to 0} f(\frac{n_1}{n_2}) = \lim_{\frac{n_1}{n_2}\to\infty} f(\frac{n_1}{n_2}) < f(\frac{n_1}{n_2}).$$

Hence firm 1 always has incentive to reduce posting until $n_1 = 0$.

When $n_1 = 0$, the same as before, firm 2 may choose any n_2 between $[M, \bar{n}_2]$.

Bibliography

Ackerberg, Daniel A. 2001. Empirically distinguishing informative and prestige effects of advertising. *RAND Journal of Economics*, 316–333.

Anderson, Simon P, & De Palma, André. 2009. Information congestion. *The RAND Journal of Economics*, **40**(4), 688–709.

Anderson, Simon P, & De Palma, André. 2012. Competition for attention in the information (overload) age. *The RAND Journal of Economics*, **43**(1), 1–25.

Bagwell, Kyle. 2007. The economic analysis of advertising. *Handbook of industrial organization*, **3**, 1701–1844.

Bose, Subir, Orosel, Gerhard O, & Vesterlund, Lise. 2002. Optimal pricing and endogenous herding.

Bose, Subir, Orosel, Gerhard, Ottaviani, Marco, & Vesterlund, Lise. 2006. Dynamic monopoly pricing and herding. *The RAND Journal of Economics*, **37**(4), 910–928.

Chintagunta, Pradeep K, & Vilcassim, Naufel J. 1992. An empirical investigation of advertising strategies in a dynamic duopoly. *Management science*, **38**(9), 1230–1244.

Chintagunta, Pradeep K, & Vilcassim, Naufel J. 1994. Marketing investment decisions in a dynamic duopoly: A model and empirical analysis. *International Journal of Research in Marketing*, **11**(3), 287–306.

DellaVigna, Stefano. 2009. Psychology and Economics: Evidence from the Field. *Journal of Economic Literature*, **47**(2), 315–372.

Eppler, Martin J, & Mengis, Jeanne. 2004. The concept of information overload: A review of literature from organization science, accounting, marketing, MIS, and related disciplines. *The information society*, **20**(5), 325–344.

Erdem, Tülin, & Keane, Michael P. 1996. Decision-making under uncertainty: Capturing dynamic brand choice processes in turbulent consumer goods markets. *Marketing science*, **15**(1), 1–20.

Erdem, Tülin, Keane, Michael P, & Sun, Baohong. 2008. A dynamic model of brand choice when price and advertising signal product quality. *Marketing Science*, **27**(6), 1111–1125.

Erickson, Gary M. 1992. Empirical analysis of closed-loop duopoly advertising strategies. *Management Science*, **38**(12), 1732–1749.

Gong, Shiyang, Zhang, Juanjuan, Zhao, Ping, & Jiang, Xuping. 2014. Tweets and sales. *Available at SSRN 2461370.*

Iyer, Ganesh, & Katona, Zsolt. 2015. Competing for Attention in Social Communication Markets. *Management Science.*

Liu, Xiao, Singh, P., & Srinivasan, Kannan. 2014. A Structured Analysis of Unstructured Big Data Leveraging Cloud Computing. *In: Carnegie Mellon University Working paper.*

Nadarajah, Saralees, & Kotz, Samuel. 2008. Exact distribution of the max/min of two Gaussian random variables. *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on*, **16**(2), 210–212.

Nguyen, Dung, & Shi, Lei. 2006. Competitive advertising strategies and market-size dynamics: A research note on theory and evidence. *Management Science*, **52**(6), 965–973.

Sims, Christopher A. 2006. Rational inattention: Beyond the linear-quadratic case. *The American economic review*, **96**(2), 158–163.

Stephen, Andrew T, & Galak, Jeff. 2012. The effects of traditional and social earned media on sales: A study of a microlending marketplace. *Journal of Marketing Research*, **49**(5), 624–639.

Toubia, Olivier, & Stephen, Andrew T. 2013. Intrinsic vs. image-related utility in social media: Why do people contribute content to twitter? *Marketing Science*, **32**(3), 368–392.

Van Zandt, Timothy. 2004. Information Overload in a Network of Targeted Communication. *RAND Journal of Economics*, 542–560.

Villas-Boas, J Miguel. 1993. Predicting advertising pulsing policies in an oligopoly: A model and empirical test. *Marketing Science*, **12**(1), 88–102.

Zhu, Yi, & Dukes, Anthony. 2015. Prominent Attributes. *Available at SSRN 2633851.*