EFFECTS OF MICROBLOGGING AND THIRD-PARTY WORD OF MOUTH ON PRODUCT SALES: EMPIRICAL STUDY BASED ON MOVIE BOX OFFICE REVENUE

Tong Sun, School of Economics and Management, Tsinghua University, Beijing, China, sunt.12@sem.tsinghua.edu.cn

Tingting Song, School of Economics and Management, Tsinghua University, Beijing, China, songtt.11@sem.tsinghua.edu.cn

Jinghua Huang, School of Economics and Management, Tsinghua University, Beijing, China, huangjh@sem.tsinghua.edu.cn

Abstract

This paper investigates the effects of word of mouth (WOM) through microblogging and third-party platforms on the box office revenue of a movie and the moderating role of movie popularity. The study uses a two-year panel data set of 147 movies, tweets from 120 thousand randomly selected users of Sina Weibo, and reviews from Douban.com discussing these movies in their release period and finds that both the volume of microblogging WOM and the Douban.com WOM rating have significant positive effects on box office. This research further explores the moderating role of movie popularity and finds that the volume of microblogging WOM has a positive effect on box office only for mass movies and the Douban.com WOM rating has a positive effect only for niche movies. This paper is the first one studying both the microblogging and third-party WOM effects on product sales, and the moderating role of product characteristics.

Keywords: Microblog, Third-party platforms, WOM, Popularity, Movie box office revenue.
1 INTRODUCTION

Microblogging has become a commonly used social media tool worldwide since Twitter was launched in 2006. Microblogging enables a unique type of user-generated content (UGC) or electronic word of mouth (e-WOM) and boosts the amount of UGC. In China, the most popular microblogging platform is Sina Weibo, which was launched in 2009. The number of daily active users of Sina Weibo had reached 61.4 million by the end of 2013 at a growth rate of 7.6% (China Internet Association, 2013). Each message in Twitter and Sina Weibo is limited to 140 characters, but the latter contains information four times that in Twitter, thus providing more information to users (Guo et al., 2011).

Online WOM is a critical factor for the purchase decision of consumers. Eighty-eight percent of respondents consult e-WOM before purchasing and refer to two WOM sources on average (Zendesk, 2013). WOM through retailer and third-party platforms considerably influences consumers (Chen et al., 2007; Liu, 2006).

Several studies have focused on the effect of Twitter WOM on product sales (Hennig-Thurau, 2012; Rui et al., 2013). Moreover, researchers have been paying attention to the effect of multiple online WOM sources on product sales (Park et al., 2012). Further, to distinguish the WOM effect on different products in detail, some product characteristics has been tested as moderating roles, for example, popularity(Zhu and Zhang, 2010; Yang et al., 2012; Dewan and Ramaprasad, 2014), brand equity(Ho-Dac et al., 2013) and product involvement(Gu et al., 2012).

However, these studies have some shortcomings. First, these studies use only counts and ratings as proxies for WOM volume and valence, respectively; however, this method is unreliable, thus enlarging the granularity of WOM measurement. The lack of effective sentiment classification techniques also decreases the efficiency of WOM measurement. Second, most studies have covered only a single WOM platform, ignoring multiple ones. Consumers access multiple WOM easily and thus make purchase decisions systematically after reading WOM. Third, current studies on moderating roles of product characteristics are limited. No previous research has examined the moderating role of product characteristics in the effect of multiple platforms’ WOM on product sales, which is a more common and practical consumer purchase scenario. Forth, no research has focused on the effect of microblogging WOM on product sales in the Chinese context because only few proper sentiment analysis techniques are suitable for the complex grammar of Chinese microblog contents. The effect of microblogging WOM on product sales in the Chinese context must be investigated because of the differences among users of Chinese microblogs and Twitter.

The present research aims to study the effect of WOM from multiple sources, particularly microblogging and third-party platforms, on product sales in the Chinese context, with movie popularity as a moderator. We choose movies as products for three reasons. First, experience goods, such as movies, are intangible; only those who have experienced them can tell their true quality. Thus, product-related WOM is useful (Nelson, 1970). Second, real-time media, such as microblogs, may have a greater effect on low-cost products; movies are products that easily turn consciousness into action, that is, purchase (Klein 1998; Villanueva, 2008). Third, most previous studies have examined movies; thus, comparing with their results is convenient (Li, 2011). We choose Sina Weibo because it is the earliest microblogging platform with the largest user base. We choose Douban.com, a website that primarily covers book and movie comments, as the third-party platform because it is the most influential third-party hosted platform among Chinese moviegoers. Douban.com has 72 million registered users and receives 200 visits monthly as of late June 2013 (Chen, 2013). Therefore, we choose Chinese movie box office revenue as the dependent variable and WOM from Sina Weibo and Douban.com as the independent variable. We choose movie popularity as movie character to test previous study (Yang et al., 2012) on multiple platforms.

The rest of this paper is organized as follows. We present a literature review in Section 2. Research hypotheses are introduced in Section 3. We then propose the research methodology according to the
research hypotheses, and empirical research is conducted based on the methodology to test the hypotheses. Finally, we discuss the empirical results and conclude with management implications.

2 LITERATURE REVIEW

We search for literature with the combination of WOM- and sales-related words. The former includes words such as UGC, WOM, online reviews, and specific WOM platforms (e.g., microblogging, Twitter, Facebook), whereas the latter includes sales and revenue. We select the Science Citation Index Expanded (SCI-EXPANDED), Social Sciences Citation Index (SSCI), Arts & Humanities Citation Index (A&HCI), Conference Proceedings Citation Index - Science (CPCI-S) and Conference Proceedings Citation Index - Social Science & Humanities (CPCI-SSH) databases from the Web of Science index database with Topic as the searching domain. The keyword combinations and searching results are listed in Table 1.

<table>
<thead>
<tr>
<th>Keywords</th>
<th>user generated content</th>
<th>UGC</th>
<th>WOM</th>
<th>Word of Mouth</th>
<th>Online Review</th>
<th>Twitter</th>
<th>Micro blog</th>
<th>Faceb ok</th>
<th>YouTube</th>
<th>Subtotal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sales</td>
<td>28</td>
<td>6</td>
<td>35</td>
<td>227</td>
<td>57</td>
<td>10</td>
<td>0</td>
<td>222</td>
<td>10</td>
<td>395</td>
</tr>
<tr>
<td>Revenue</td>
<td>10</td>
<td>0</td>
<td>9</td>
<td>39</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>8</td>
<td>5</td>
<td>78</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>473</td>
</tr>
</tbody>
</table>

Table 1. Literature Search Combination

Only 208 papers are left after we rule out duplicates from the original 473 papers. We classify these papers by their methodologies. Aside from econometric research using second-hand data, the papers are mostly field studies, questionnaires, and case studies. We select econometric studies using second-hand data, a total of 59 papers, to review the most relevant ones. Literature is reviewed based on these 59 papers.

2.1 Effect of Retailer-Hosted, Third-Party-Hosted, and Microblogging WOM on Sales

Previous research has divided e-WOM into retailer-hosted and third-party-hosted platform WOM. The former refers to customer reviews on online platforms selling products directly, such as reviews on Amazon.com. The latter refers to reviews written by consumers or experts after purchase on platforms independent of product selling, such as professional product-review website Cnet.com and movie-review website Yahoo! Movies (Gu et al., 2012).

Previous research on retailer-hosted WOM has focused mostly on Amazon and has agreed on the significant effect of WOM volume on product sales. However, some studies have obtained significant results on WOM valence (Chevalier, 2006), whereas others have not (Lee et al., 2011). Moreover, the unique functions of Amazon also affect product sales. For instance, Amazon’s automatic collaborative recommendation (Chen, 2004), disclosure of reviewer identity (Hu et al., 2008), and review usefulness (Pan et al., 2011) have a significant positive effect on product sales. Such studies have focused mainly on WOM volume and used only the rating of reviewers in studying WOM valence - a defect in sentiment analysis.

Liu (2006), Karniouchina (2011), and Chintagunta et al. (2010) study the effect of Yahoo! Movies WOM on box office revenue and find that WOM volume has a positive effect. However, Liu (2006) finds no significant evidence on WOM valence, whereas Karniouchina (2011) and Chintagunta et al. (2010) find that the effect is significant. The limitation of these studies is that factors influencing box office revenue, especially critical ones, such as release date and screenings, are incomprehensive. In addition, research on restaurant specific third-party platform WOM find that WOM volume and
valence about dishes quality, environment and service have significant impact on consumers’ intention to visit restaurants’ websites (Zhang et al, 2010; Lu et al, 2013).

Succeeding research has examined the association between microblogging WOM and sales. Many studies have used user rating to measure WOM valence, disregarding complex sentiments in WOM texts. In fact, sentiments expressed in e-WOM are rich; only by quantitatively selecting each sentiment in text and adding each sentiment value can we accurately obtain WOM valence (Hu, 2010; Chintagunta, 2010; Sun, 2012). Microblogging WOM studies make up for this limitation by applying text analysis. For example, Hennig-Thurau et al. (2012) find that both volume and valence affect box office revenue, but the two variables have no interaction effect. This paper uses text mining to classify tweets. However, this research focuses only on the effect on opening-week box office revenue and ignores time-varying changes. In addition, whether the relationship is causal is unclear. Rui et al. (2013) find that both the volume of intention tweets and the valence of non-intention tweets affect box office revenue. The contribution of this paper is its classification of the direct and indirect effects of microblogging. However, the research has three limitations. First, a general sentiment analysis is employed, but it is unfit for microblogging text, which is short and unstructured and contains emotion icons, thus decreasing classification accuracy. Second, machine learning is used to classify tweets for massive data and thus produces inaccurate results. Third, this paper leaves out important factors in the econometric model, failing to effectively control endogeneity.

2.2 Effect of Multiple-Source WOM on Sales

Most existing studies have focused only on single-source WOM, ignoring the influence of other sources that may also be references for consumers. The current mixed results may be attributed to the lack of consideration of WOM through multiple sources and their interaction (Godes and Mayzlin, 2004; Liu, 2006).

Park et al. (2012) examine the difference between the effect of WOM through Amazon and Cnet.com on sales of digital cameras. Retailer-hosted WOM differs from third-party-hosted WOM in two ways. First, retailer-hosted WOM is the most accessible to retail customers. Second, retailer-hosted WOM provides not only product information but also service information. Third-party-hosted WOM is often perceived as more trustworthy than retailer-hosted WOM, despite the latter’s usefulness and convenience. Both types of WOM have advantages and disadvantages. Thus, the relative influence of the WOM of these sources must be assessed. Therefore, this paper proposes a pair of competing hypotheses. They find that the valence of third-party-hosted WOM has a greater influence on retail sales than that of retailer-hosted WOM. This paper explains that risk uncertainty increases although search cost is low in retailer-hosted platforms; thus, consumers tend to consult third-party-hosted WOM. This is the first study that compares the effect of WOM on two platforms.

Gu et al. (2012) compare the effect of WOM between Amazon and third-party-hosted platforms. They collect sales rank data on 148 camera products from Amazon for four months and corresponding online review data from Amazon and third-party-hosted platforms in the same period. Third-party-hosted WOM has a significant effect on high-involvement product sales, whereas retailer-hosted WOM does not.

Previous studies have used only digital cameras as research objects. Moreover, only retailer-hosted and third-party-hosted platforms have been studied, and other important social media, especially microblogs, have been ignored.

2.3 Product Character as a Moderating Role

The most common product category is mainstream/mass/popular vs. niche/non-popular. Yang et al. (2012) find a significant effect of WOM valence on box office revenue only in the case of non-mainstream movies and the effect of WOM volume on box office revenue is greater for mainstream
movies. Dewan and Ramaprasad (2014) find that blog buzz has a negative effect on music sales only for the non-popular ones. Zhu and Zhang (2010) find that less popular video games’ user number is subject to WOM. Ho-Dac et al. (2013) find that positive WOM number increases the sales of DVD models of weak brand. A unique study regards the time that a consumer costs when choosing for a product as a classification indicator and finds third-party WOM has a stronger effect on high-involvement product sales (Gu et al., 2012). Another study takes network externality as a classification indicator and finds consumers pay more attention to WOM when consuming products with network externality (Yang and Mai, 2010).

3 RESEARCH HYPOTHESES

The comparison between Douban.com and Sina Weibo on their forms and contents is listed in Table 2.

<table>
<thead>
<tr>
<th>Platform</th>
<th>Form</th>
<th>Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>Douban.com (Third-party review platform)</td>
<td>1. Objective</td>
<td>informativeness</td>
</tr>
<tr>
<td></td>
<td>2. Either aggregate information or single one</td>
<td></td>
</tr>
<tr>
<td>Sina Weibo (social media)</td>
<td>1. Social tie (weak or strong)</td>
<td>New&amp; real time</td>
</tr>
<tr>
<td></td>
<td>2. Generally single information</td>
<td></td>
</tr>
</tbody>
</table>

*Table 2. Comparison: Douban.com vs. Sina Weibo*

As a third-party review platform, Douban.com can provide more objective and comprehensive movie reviews. Viewers can acquire aggregate information provided by the platform or a great number of short reviews from other users and some experts. Indeed, viewers usually come to this platform with certain goals, i.e., browse and obtain useful information of movies. However, as a social-media platform, users on Sina Weibo actually do not have a specific goal. They may unintentionally receive real-time movie information from the users they follow (strong ties or weak ties).

Arndt (1967) explains WOM as a form of face-to-face communication about a commercial entity or offering between consumers, emphasizing its unbiased and personal nature. WOM has two different effects, awareness effect and persuasive effect (Neelamegham and Chintagunta, 1999). The former increases sales by transmitting and sharing product information, making more potential consumers know about the product. The latter may change the attitude of consumers toward a product. Volume and valence (Lee, 2010) are variables for the awareness and persuasive effects, respectively. In this paper, third-party WOM volume means the number of reviews for one movie on Douban.com and third-party WOM rating/valence means ratings for each review. Some researchers argue that text rather than numerical rating is recommended when measuring third-party WOM valence because reviewers tend to write multiple aspects of a product and only give a rating number based on partial evaluation (Chintagunta et al., 2010). However, text reviews are short on Douban.com, usually in one sentence. Under such circumstances, rating is a proper proxy of WOM valence (Bruce et al., 2012).

The volume of microblogging WOM for a movie is the number of microblog elements containing the movie name explicitly displayed in the microblog, including the original tweets and the reposted tweets with comments. The valence of microblogging WOM for a movie is the average value of positive, negative or neutral emotion expressed in all the microblogging elements for a movie.

Hypotheses are proposed according to the above analysis.

3.1 Effect of Third-Party WOM on Movie Box Office Revenue

As our analysis above, Douban.com, as an objective and intensive free-writing WOM platform, attracts many moviegoers sharing their reviews. Potential moviegoers regard WOM on Douban.com as an important reference and intend to acquire movie information from other users.
Godes and Mayzlin (2004) point out that the more frequent consumers talk about a product, the higher the possibility other consumers get to know that product, thus increasing sales. The awareness effect of WOM influences the purchase intention of potential consumers by informing them about the existence of a certain product. The more reviews about a movie on Douban.com, the higher the possibility that the movie is known, enhancing the awareness of consumers regarding that movie. Therefore, we propose the following:

H1a: The volume of third-party movie WOM has a significant positive effect on movie box office revenue.

In particular, when most ratings of one movie are high, that is, the WOM valence of this movie is high, viewers’ attitude towards the movie would be positive. Then, the intention of potential moviegoers is converted into action, thus increasing box office revenue (Chevalier and Mayzlin, 2006; Dellarocas et al., 2007; Li and Hitt, 2008; Eliashberg et al., 2000; Karniouchina, 2011). Based on the characteristics of Douban.com, we would expect that viewers are easily to be influenced by other users’ attitudes (i.e., ratings). Therefore, we propose the following:

H1b: The third-party WOM rating has a significant positive effect on movie box office revenue.

3.2 Effect of Microblogging WOM on Movie Box Office Revenue

In general, users on Sina Weibo often come across some novel information they do not have awareness, such as tweets of new movies.

We infer that the more reviews about a movie on Sina Weibo, the higher the possibility that the movie is known, enhancing the awareness of consumers regarding that movie. Hence, we expect that users may be influenced by the volume of microblogging WOM. First, according to mere exposure effect, potential moviegoers tend to believe that movie quality is higher when they find more movie-related microblogging WOM (RB Zajonc, 1968). The herding effect also points that potential moviegoers believe that this movie is of high quality if they are more users talking about the movie (Bikhchandani et al., 1998). Second, the greater the amount of movie-related WOM, the higher the possibility that potential moviegoers become aware of the movie will be by chance. Microblogging WOM about movies may influence box office revenue because it may influence the purchase intention of potential moviegoers who browse the microblog without a clear intention (Neelamegham and Chintagunta 1999). Third, Chinese microblog users pay more attention to the media attribute of this platform, that is, obtaining more novel information (Wang et al. 2013). Given the larger coverage and deeper influence of microblogging, we propose the following:

H2a: The volume of microblogging WOM has a significant positive effect on movie box office revenue.

In addition, box office revenue can be influenced by the valence of microblogging WOM. First, the valence of WOM with social relationship is more persuasive. Ordinary WOM lacks social ties between the WOM sender and receiver, and a persuasive effect is thus weakened in contrast to that in traditional offline WOM (Chatterjee, 2001). The relationship Follow, which connects microblog accounts, adds social ties. Second, microblogging WOM is objective and independent; hence, the valence of WOM may influence the purchase decision of potential moviegoers. Therefore, we propose the following:

H2b: The valence of microblogging WOM has a significant positive effect on movie box office revenue.

3.3 The Moderating Role of Movie Popularity

The popularity of movies may play a moderating role in the above relationships. Movies can be categorized as mass movies and niche movies according to their popularity (Yang et al., 2012; Dewan
and Ramaprasad, 2014). Mass movies refer to those relatively more commercial movies with the greater marketing budget, which targets the large market segment; while niche movies are less commercial but more artistic with the smaller marketing budget, which targets relatively the smaller market segment. The main difference between mass movies and niche movies is the advertisement budget.

As we mentioned above, viewers on Douban.com are mainly obtaining useful movie reviews. They are easily to be influenced by others’ attitudes (i.e., ratings). We would expect that the persuasive effect of Douban.com would be stronger for niche movies, because viewers may not obtain more information of such movies from other channels, such as social media (Zhu and Zhang et al., 2010). Viewed through the lens of signalling theory (Ackerloff, 1970), both the ads and the persuasive effects of third-party WOM valence may help reduce consumers’ purchasing risk and gain their trust. So for niche movies which are lacking in ads (Yang et al., 2013), consumers may tent to seek for objective information from third-party platform ratings and would be influenced by the persuasive effect. Therefore, we propose the following:

H3a:
The positive effect of valence of third-party WOM on box office revenue is stronger for niche movies.

However as we have analyzed, users on Sina Weibo are easily influenced by the volume of tweets, that is, when more users talk about a movie, users are more easily to have awareness of it. Indeed, users are willing to talk about more popular topics or movies on Sina Weibo. In contrast, if one topic or one movie is niche, then the possibility of posting it is lower. Therefore, we propose the following:

H3b:
The positive effect of volume of microblogging WOM on box office revenue is stronger for popular movies.

4 RESEARCH METHODOLOGY

4.1 Research Design

Prior research has chosen daily or weekly box office revenue as the dependent variable. We use daily data to investigate the effects because the life span of movies in China is about two weeks. Specifically, we use the number and average ratings of daily reviews on Douban.com as the volume and valence of third-party WOM. We use the number of daily tweets, including retweets, as the volume of microblogging WOM; we use the average sentiment of daily reviews as the valence of microblogging WOM. Moreover, we draw from the Chinese microblogging semantic analysis algorithm to calculate the sentiment of tweets (Zhao et al., 2012). The principle of the algorithm is as follows: (i) we categorize emoticons into three categories, namely, positive, neutral, and negative, because users express their emotion with emoticons; (ii) we convert each tweet into a sequence of words {wi}, where wi is a word or emoticon and i is its position in t, and then encode a negative emotion as -1, a neutral emotion as 0, and a positive emotion as 1; (iii) we employ the simple method of naive Bayes to build the classifier, which consumes little training time and rapidly predicts the category; and (iv) we obtain the sentiment correctly when tweets are imported into the classifier. The accuracy of this method is 64.3% when applied to our dataset.

For the analysis, we use a log-linear model, which has been widely used in prior studies, to investigate the relationship between WOM and retail sales (e.g., Chevalier and Mayzlin 2006). Key variables are shown in Table 2. Moreover, we include three control variables, namely, Weekend, lnScreenings, and lnDays, following previous literature (Elberse and Eliashberg, 2003; Liu, 2006). Specifically, the three variables are the number of screens (lnScreenings), the number of days released (lnDays), and a dummy variable indicating if day t is a weekend (Weekend).
\[
\ln(\text{BoxOfficeRev}_{it}) = \alpha_0 + \alpha_1 \ln(\text{Total Tweets}_{it}) + \alpha_2 \ln(\text{Ave Sentiment}_{it})
\]
\[
+ \alpha_3 \ln(\text{Third Party Volume}_{it}) + \alpha_4 \ln(\text{Third Party Ratings}_{it}) + \alpha_5 \ln(\text{Screenings}_{it})
\]
\[
+ \alpha_6 \text{Weekend}_{it} + \alpha_7 \ln(\text{Days}_{it}) + \epsilon_{it}
\]  

<table>
<thead>
<tr>
<th>Type</th>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent</td>
<td>(\ln(\text{BoxOfficeRev}_{it}))</td>
<td>The log value of movie box office of movie (i) at day (t)</td>
</tr>
<tr>
<td>Independent</td>
<td>(\ln(\text{Total Tweets}_{it}))</td>
<td>The log value of number of tweets of movie (i) at day (t)</td>
</tr>
<tr>
<td></td>
<td>(\ln(\text{Ave Sentiment}_{it}))</td>
<td>Daily average Sina Wei user grade for movie (i) at day (t)</td>
</tr>
<tr>
<td></td>
<td>(\ln(\text{Third Party Volume}_{it}))</td>
<td>Log value of reviews number of movie (i) at day (t) on Douban.com</td>
</tr>
<tr>
<td></td>
<td>(\ln(\text{Third Party Ratings}_{it}))</td>
<td>Daily average Douban.com user grade for movie (i) at day (t)</td>
</tr>
<tr>
<td>Control</td>
<td>(\ln(\text{Screenings}_{it}))</td>
<td>Log value of daily number of screens for movie (i) at day (t)</td>
</tr>
<tr>
<td></td>
<td>(\text{Weekend}_{it})</td>
<td>Dummy variable indicates if (t) is weekend (1 for Saturday or Sunday, and 0 otherwise)</td>
</tr>
<tr>
<td></td>
<td>(\ln(\text{Days}_{it}))</td>
<td>Log value of number of days movie (i) has been released at day (t)</td>
</tr>
</tbody>
</table>

**Table 2. Variables and Descriptions**

### 4.2 Data Collection

We collect data according to the following steps. First, we choose movies, screenings, and box office revenue. A total of 172 movies with complete information are collected from 201 movies released in Wanda Cinemas in 2011 and 2012 from the China Movie Box Office Database (website available at <http://58921.com/boxoffice>). Wanda Cinemas is the most famous cinema in China, taking up more than 20% of the market share. According to Chinese Film Market Review (2013), Wanda Cinemas covers most regions and wins most box office revenue in China since 2009; thus, we regard it as the most appropriate source of box office revenue. We automatically collect the daily box office revenue of 172 movies widely released during their screening period with a crawler program. In the present research, screening period covers the release date to the end of screening. Second, data on microblogging WOM are processed. We collect 1,913,786 tweets mentioning the titles of the 172 movies during their screening period from the microblogging database (containing 120,000 accounts, the database schema includes microblogging content, time of posting, and basic account information). The 120,000 accounts stored in the database are randomly selected from a social network perspective, that is, they are randomly selected from some start points in Sina Weibo users’ social network and expanded to their adjacent points randomly until 120,000 accounts are captured (Zhao et al., 2012). Sentiment analysis is conducted by the Chinese microblogging semantic analysis algorithm (Zhao et al., 2012). Third, we collect data on Douban.com WOM about 172 movies and calculate the daily review number (WOM volume) and average rating (WOM valence) for each movie. Next, twenty-five movies without WOM data during the screening period are ruled out from the dataset. Please note that though both intention and post-purchase WOM are all presented on Douban.com, only the latter one is complete and can be collected, which is different from Sina Weibo. Finally, we obtain 3,812 records on 147 movies. Each record stores daily box office revenue, the number of microblogging contents, the average sentiment value of microblogging contents, the number of Douban.com reviews, and the average rating of Douban.com reviews for each movie. Moreover, screenings and whether the date is a weekend are recorded.
5 EMPIRICAL RESEARCH

5.1 Descriptive Statistics

We present the descriptive statistics of the sample data in Table 3. The average total box office revenue is RMB 32.8 million; the minimum revenue is RMB 0.2422 million, and the maximum revenue is RMB 256 million. Thus, the movies have significant differences. Moreover, the ratio of the box office revenue during the opening week to the total box office revenue is up to 60%. The average release days of movies are about two weeks, from 10 d to 65 d. Thus, the life span of movies is short. Hence, using daily box office revenue as the dependent variable is more meaningful than using weekly data. Moreover, the average number of daily tweets is 79.78; the minimum number is 1, and the maximum number is 6367. The number of daily reviews of different movies in Douban.com is quite different; the minimum number is 0, and the maximum number is 9136.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Box office (aggregate)</td>
<td>¥ 32800000</td>
<td>¥ 50100000</td>
<td>¥ 242200</td>
<td>256000000</td>
</tr>
<tr>
<td>Ratio of opening week to aggregate box office</td>
<td>0.67</td>
<td>0.19</td>
<td>0.03</td>
<td>0.99</td>
</tr>
<tr>
<td>Box office (daily)</td>
<td>¥ 861622.6</td>
<td>¥ 1813830</td>
<td>¥ 0</td>
<td>¥ 21200000</td>
</tr>
<tr>
<td>TotalTweets (daily)</td>
<td>79.78</td>
<td>449.47</td>
<td>1</td>
<td>6367</td>
</tr>
<tr>
<td>TweetsAveSentiment (daily)</td>
<td>0.24</td>
<td>0.46</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>ThirdPartyVolume (daily)</td>
<td>267.71</td>
<td>607.12</td>
<td>0</td>
<td>9136</td>
</tr>
<tr>
<td>ThirdPartyRating (daily)</td>
<td>5.74</td>
<td>2.33</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>Screenings (daily)</td>
<td>342.03</td>
<td>442.52</td>
<td>1</td>
<td>3415</td>
</tr>
<tr>
<td>Weekend</td>
<td>0.29</td>
<td>0.45</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Days (aggregate)</td>
<td>15.18</td>
<td>10.39</td>
<td>10</td>
<td>65</td>
</tr>
</tbody>
</table>

Note: Total number of movies = 147. Total number of WOM messages analyzed = 3812

Table 3. Summary Statistics of the Movie Sample

When comes to the movie popularity categorization, we divide movies as mass ones and niche ones according to the market role of distributors (Gemser et al., 2007; Yang et al., 2012; Dewan and Ramaprasad, 2014). According to 2012-2013 Chinese Movie Industry Report (available at <http://www.entgroup.com.cn/research/detail.aspx?id=15601&typeid=34>), there are 193 movie distributors in mainland China and the top 5 of them take up 82.68% market share in 2013. Besides, the top 5 distributors have remained their leading status for at least 6 years. Therefore, if a movie is distributed by one of the top 5 distributors, it can be regarded as mass movie and if not, we define it as niche one. In the case of our dataset, 105 movies are assumed to be mass movies and 42 are niche ones.

5.2 Empirical Model Test

Multicollinearity among WOM variables is of particular concern because our research measures the WOM effect from multiple sources. We check the correlation among key independent variables in Table 4, following previous studies (Gu et al., 2012; Duan et al., 2008a) and all correlations are less than 0.6. Besides, we also calculate Variance Inflation Factor, and the result is 1.20 which is far smaller than 10.0, indicating that multicollinearity is not a serious problem in our analysis.
Hausman provides a specification test commonly used as a criterion in choosing between the fixed effects (FE) and random effects (RE) models. The p-value is 0.00; hence, we should choose the less restrictive FE model to test our hypotheses. Table 5 presents the results; Column 1 in Table 5 shows that the coefficients of the control variables are significant. Specifically, the coefficient of Weekend is positive (α6= 0.396, p<0.01), indicating that the box office revenue on weekends is higher than that on weekdays. The coefficient of lnScreenings is positive (α5= 0.964, p<0.01), showing that the box office revenue increases as the number of screens increases. The coefficient of lnDays is negative (α7= -0.354, p<0.01), illustrating that the box office revenue decreases as time passes.

Douban.com rating is strongly positively associated with box office revenue (α4= 0.074, p<0.05), whereas the volume of Douban.com reviews is not significant, indicating that the box office revenue does not have a significant relationship with the number of Douban.com reviews.

The coefficient of the microblogging WOM volume is significant and positive, indicating that the number of tweets has a significant effect on box office revenue. In other words, the more tweets talking about a movie on the Sina Weibo, the higher the box office revenue. In addition, the coefficient of microblogging WOM valence is not significant and its possible effect on movie box office revenue is not supported.

Afterward, we consider the moderating effect of popularity and report the results in column 2 and column 3. Specifically, when the popularity of movies is low, as is shown in column 2, only the coefficient of ratings of Douban.com on movie box is significant positive, indicating that the positive effect of valence of Douban.com WOM on box office revenue is stronger for niche movies. Hence, H3a is supported. However, when the popularity is high, only the coefficient of volume of tweets on movie box is significant positive, indicating that the positive effect of volume of Weibo WOM on box office revenue is stronger for mass movies. Hence, H3b is supported.

<table>
<thead>
<tr>
<th>Category</th>
<th>Variable</th>
<th>Model 1</th>
<th>Model 2 (popularity=0)</th>
<th>Model 3 (popularity=1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Douban.com Platform</td>
<td>Volume</td>
<td>-0.003</td>
<td>-0.008</td>
<td>0.009</td>
</tr>
<tr>
<td></td>
<td>Ratings</td>
<td>0.074***</td>
<td>0.206***</td>
<td>-0.020</td>
</tr>
<tr>
<td>Microblogging Platform</td>
<td>Tweets</td>
<td>0.058***</td>
<td>0.059</td>
<td>0.050**</td>
</tr>
<tr>
<td></td>
<td>AveSentiment</td>
<td>0.042</td>
<td>0.052</td>
<td>0.061</td>
</tr>
<tr>
<td>Control Variables</td>
<td>Weekend</td>
<td>0.396***</td>
<td>0.315***</td>
<td>0.477***</td>
</tr>
<tr>
<td></td>
<td>Screenings</td>
<td>0.964***</td>
<td>0.944***</td>
<td>1.012***</td>
</tr>
<tr>
<td></td>
<td>Days</td>
<td>-0.354***</td>
<td>-0.443***</td>
<td>-0.254***</td>
</tr>
<tr>
<td></td>
<td>R-sq</td>
<td>0.88</td>
<td>0.88</td>
<td>0.89</td>
</tr>
<tr>
<td></td>
<td>Number of observations</td>
<td>1790</td>
<td>713</td>
<td>1077</td>
</tr>
</tbody>
</table>

Table 5. Analysis of Sina Weibo and Douban.com WOM Effects on movie box office

5.3 Robustness Check

Prior studies (Duan et al., 2008b; Sonnier et al., 2011) have pointed out a positive feedback loop between WOM and product sales: WOM increases product sales and thus generates more WOM. Similarly, we adopt the two-stage least squares (2SLS) estimation used by Duan et al. (2008b) to check the robustness of our results to account for dynamic relationships between WOM through microblogging and Douban.com and box office revenue. In addition to Equation (1), the following two equations are added to the system of equations. The two equations include the lagged box office revenue.
revenue and WOM volume (i.e., the volume of Douban.com and microblogging WOM) in each equation to capture their possible effects. In addition, we control for restaurant-specific fixed effects (\( \text{Hi} \)) and the effect of the weekend in each equation.

We run 2SLS to estimate the three simultaneous equations. The results are presented in Table 5. The static and dynamic models have no significant difference, which indicates the robustness of our estimation results after the dynamic relationships are incorporated.

\[
\ln \text{Total Tweets}_{it} = \gamma_0 + \gamma_1 \ln \text{Box Office Rev}_{it-1} + \gamma_2 \ln \text{Total Tweets}_{it-1} + \gamma_3 \text{Weekend}_{it} + u_i + \phi_{it} \\
\ln \text{Third Party Volume}_{it} = \delta_0 + \delta_1 \ln \text{Box Office Rev}_{it-1} + \delta_2 \ln \text{Third Party Volume}_{it-1} + \delta_3 \text{Weekend}_{it} + u_i + \pi_{it}
\]  

<table>
<thead>
<tr>
<th>Category</th>
<th>Variable</th>
<th>Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Douban.com Platform</td>
<td>Volume</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>Ratings</td>
<td>0.025**</td>
</tr>
<tr>
<td>Microblogging Platform</td>
<td>Tweets</td>
<td>0.108***</td>
</tr>
<tr>
<td></td>
<td>AveSentiment</td>
<td>0.044</td>
</tr>
<tr>
<td>Control Variables</td>
<td>Weekend</td>
<td>0.489***</td>
</tr>
<tr>
<td></td>
<td>Screenings</td>
<td>1.110***</td>
</tr>
<tr>
<td></td>
<td>Days</td>
<td>-0.125***</td>
</tr>
</tbody>
</table>

Note: **p<0.05, ***p<0.01

Table 6. Results of the Dynamic Model

6 DISCUSSION AND CONCLUSION

6.1 Discussion

Parts of our hypotheses are supported after statistical analysis. The results are listed in Table 7.

<table>
<thead>
<tr>
<th>Hypotheses</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1a: The volume of third-party movie WOM has a significant positive</td>
<td>nonsupport</td>
</tr>
<tr>
<td>effect on movie box office revenue.</td>
<td></td>
</tr>
<tr>
<td>H1b: The third-party WOM rating has a significant positive effect on</td>
<td>support</td>
</tr>
<tr>
<td>movie box office revenue.</td>
<td></td>
</tr>
<tr>
<td>H2a: The volume of microblogging WOM has a significant positive</td>
<td>support</td>
</tr>
<tr>
<td>effect on movie box office revenue.</td>
<td></td>
</tr>
<tr>
<td>H2b: The valence of microblogging WOM has a significant positive</td>
<td>nonsupport</td>
</tr>
<tr>
<td>effect on movie box office revenue.</td>
<td></td>
</tr>
<tr>
<td>H3a: The positive effect third-party WOM rating on box office</td>
<td>support</td>
</tr>
<tr>
<td>revenue is stronger for niche movies.</td>
<td></td>
</tr>
<tr>
<td>H3b: The positive effect of volume of microblogging WOM on box</td>
<td>support</td>
</tr>
<tr>
<td>office revenue is stronger for popular movies.</td>
<td></td>
</tr>
</tbody>
</table>

Table 7. Results of Hypotheses Test
We begin our analysis by estimating the influence of third-party WOM on box office revenue. Hypotheses 1a is rejected, but Hypotheses 1b is supported. The nonsignificance of Hypotheses 1a is inconsistent with the result of most previous studies. The following reasons may shed light on this phenomenon. First, the WOM awareness effect has been tested many times, but it is subject to the display mode of information and the viewing intention of website users. WOM on Douban.com is not disseminated by sharing; it is stored on the website, ready for information search by consumers. This kind of WOM display fails to expand coverage and enhance awareness of the product. In other words, more movie reviews cannot reach more potential consumers. Douban.com users intend to search for detailed and objective movie WOM after knowing about the movie. Therefore, the WOM display interface on Douban.com and the viewing intention of users hinder the formation of the WOM awareness effect. Second, more proper econometric methods and control variables are used in the present study, preventing potential endogeneity. Third, movies are a disposable commodity, in contrast to ordinary products or services. Thus, the number of previous moviegoers does not inform the degree of recognition of a movie.

Moreover, we test the influence of microblogging WOM. Hypothesis 2a is supported, whereas Hypothesis 2b is rejected. Hypothesis 2b is rejected for the following reasons: First, all the extreme emotions on Sina Weibo, positive or negative, will all intrigue viewers and contribute to box office revenue. Actually, it is the buzz itself, not the valence embedded in the buzz that attracts viewers’ attention and shapes their attitudes (Berger et al., 2010). Second, microblog is more a media than a social tool in China and users of Sina Weibo tend to receive real time information to keep pace with fashion rather than recommendations. Third, the insignificant conclusion can be inferred by previous studies on Twitter. Rui et al. (2013) find the percentage of both the positive and negative Twitter WOM have a significant effect on box office revenue and Hennig-Thurau et al. (2012) have a similar conclusion saying the number of negative Twitter WOM has a greater negative effect on box office than positive WOM’s positive effect. It’s quite clear that their metrics are different from this paper. We use average valence to indicate the macro level emotion (Yu et al., 2013), while they focus on positive and negative WOM, respectively. In other words, they didn’t separate valence from volume. Previous studies find that the volume of positive WOM is always greater than the negative one and the negative one may have a greater effect on box office revenue. So it’s obvious that no clear effect would happen to box office.

I addition, the results also prove that the above relationships can be altered by the moderating role of movie popularity. Hypothesis 3a is supported suggesting that the third-party rating plays a more critical role when positively influencing box office revenue for niche movies because potential moviegoers have to consult to post-moviegoers opinions from online channels to reduce risks when judging for a niche movie which is lacking in public information. Hypothesis 3b is supported suggesting that the microblogging WOM volume plays a more critical role when positively influencing box office revenue for mass movies because new and popular news are prevalent information on social media platforms like microblog and viewers on such a platform may be informed and get aware of popular movie information.

6.2 Implications and Limitations

The findings of this study have a number of important implications for research on the effect of e-WOM on product sales.

First, this paper uses a sentiment analysis technique specially designed for Chinese microblogging contents to analyse texts at word level. Most previous studies have regarded user rating as a measurement of WOM valence; these ratings ignore the subtle sentiment in WOM and hence are inapplicable to microblogging WOM. We use a high-accuracy machine-learning sentiment classifier to ensure the validity of the tests.

Second, this study examines WOM on microblog and third-party platforms simultaneously. Most online platforms useful for potential moviegoers have been considered, keeping the model reliable.
Third, this paper not only talks about the moderating role of movie popularity in influencing the effect of online WOM on product sales, but also focuses such issue on two platforms – microblog and third-party platform. This brand new contribution makes this paper the first study researching on popularity as a moderating role on more than one platform, indicating that the moderating role may be different on multiple platforms when comparing to one single platform.

Forth, the effect of WOM on box office revenue in the Chinese context is discussed in this paper. Research results on Twitter WOM may be different from Chinese microblogging WOM because Chinese microblogging has started late and its user base is unique. We devise hypotheses according to the difference between Chinese microblogging and Twitter.

The results also have valuable practical implications. Sellers, especially movie distributors, can regard our results as a guide on where to conduct a marketing campaign and how to do it. For instance, sellers can increase movie-related discussions on a microblog by sending out incentives or paying attention to Douban.com movie ratings. Besides, movie distributors should pay more attention to microblogging WOM volume for mass movie marketing and third-party platform rating is more important for niche movies.

However, this research has several limitations. First, Sina Weibo has some restrictions in massive data downloading; thus, we have access only to microblogging contents. Data incompleteness decreases the accuracy of our model. Second, we download only microblogging data, but the third-party data we have crawled are complete; thus, the interaction effect cannot be tested. Future research with access to complete data may consider focusing on the interaction effect between platforms.

Acknowledgement

This research was supported by the National Natural Science Foundation of China (71272028) and MOE Project of the Key Research Institute of Humanity and Social Sciences at Universities (13JJD630008). Besides, microblogging data we use in this paper is provided by courtesy of Prof. Wu Junjie, Beihang University.

References


