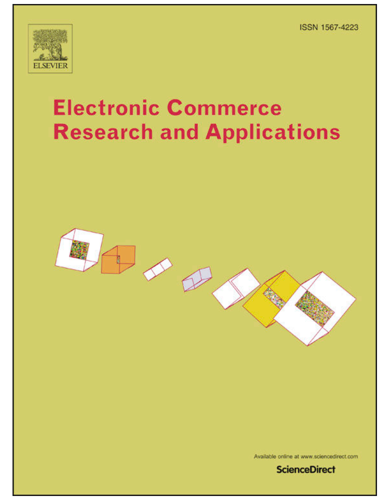
Journal Pre-proofs



Reputational Assets and Social Media Marketing Activeness: Empirical Insights from China

Anders C. Johansson, Zhen Zhu

PII: S1567-4223(23)00070-4
DOI: https://doi.org/10.1016/j.elerap.2023.101305
Reference: ELERAP 101305

To appear in: *Electronic Commerce Research and Applications*

Received Date: 6 October 2022
Revised Date: 1 April 2023
Accepted Date: 22 August 2023

Please cite this article as: A.C. Johansson, Z. Zhu, Reputational Assets and Social Media Marketing Activeness: Empirical Insights from China, *Electronic Commerce Research and Applications* (2023), doi: https://doi.org/10.1016/j.elerap.2023.101305

Reputational Assets and Social Media Marketing Activeness: Empirical Insights from China

Anders C. Johansson

Stockholm School of Economics

Zhen Zhu*

University of Kent

Abstract

We explore the linkages between social media marketing activeness and reputational assets on digital platforms with a unique sample of over 8,000 customer-to-customer (C2C) sellers registered on both Taobao, China's largest C2C online shopping platform, and Sina Weibo, China's largest microblogging platform. A unique collaborative effort between the two platforms enables us to examine whether C2C sellers are motivated to engage in marketing activities on a separate social media platform. Applying machine learning methods, we first classify whether C2C sellers conduct social media marketing on their microblogs or not, which allows the measurement of social media marketing activeness. We then use logistic regression models and find that *earned* reputational assets such as the rating scores and the number of followers are significantly associated with social media marketing activeness on both platforms. However, we identify a conflict of *owned* reputational assets such as the shop age and the paid membership between the two platforms, which provides a potential explanation for the limited success of the cross-platform collaboration.

Keywords: social media marketing, reputational assets, electronic commerce, China.

* Corresponding author, Kent Business School, University of Kent, Parkwood Road, Canterbury, Kent CT2 7FS, United Kingdom. Tel: +44 (0)1227 827726. Email: z.zhu@kent.ac.uk. The authors would like to thank four anonymous reviewers for insightful comments, and Xu Wang and Michelangelo Puliga for assistance during the data collection process. Johansson acknowledges financial support from the Marianne & Marcus Wallenberg Foundation.

1 Introduction

Social media marketing has become a standard practice for businesses in the twenty-first century. It enables people to interact with each other in a variety of ways and for firms and marketers to directly reach out to and engage with both businesses and consumers (Appel et al., 2020). Social media marketing takes advantage of how online relationships function, i.e. in a seamless, omnichannel, highly personalized, and anthropomorphized way (Steinhoff et al., 2019), and has been utilized, both in theory and in practice, as a facilitator of individual expression, a decision support tool and a market intelligence source (Lamberton and Stephen, 2016). Social media marketing incurs lower costs than other traditional offline methods (Du et al., 2015), which is perhaps especially appealing to customer-to-customer (C2C) sellers (Wang et al, 2019). The effectiveness of social media marketing has been well documented in the literature (Chang, et al., 2015; Godey, et al., 2016; Kim and Ko, 2012; Kumar, et al., 2016). Meanwhile, the adoption and usage of social media has been studied in varying contexts such as individuals (Cheung, et al., 2011; Zolkepli and Kamarulzaman, 2015), B2C (business-to-customer) versus B2B (business-to-business) (Dwivedi et al., 2021; Iankova, et al., 2019; López-López and Giusti, 2020; Moore, et al., 2013; Pascucci, et al., 2018), and large corporations versus SMEs (small and medium-sized enterprises) (Abbasi, et al., 2022; Abed, 2020; Chatterjee and Kar, 2020; Heavey, et al., 2020; Sinclaire and Vogus, 2011). The existing literature is primarily focused on various drivers of customer engagement on social media (de Oliveira et al., 2016; Liu et al., 2019; Shi et al., 2016) or factors affecting a firm's decision of adopting and using social media as a marketing channel (Chatterjee and Kar, 2020; Dahnil et al., 2014; Guesalaga, 2016; Michaelidou et al., 2011; Siamagka et al., 2015). However, the usage of social media marketing in the C2C context remains underexplored (Puriwat and Tripopsakul, 2021).

This paper takes a quantitative approach to improve our understanding of the driving forces behind social media marketing activeness. We focus on C2C sellers on Taobao, China's largest C2C online shopping platform. Alibaba, the parent of Taobao, started to invest in Sina Weibo in 2013. After Sina Weibo's IPO on the New York Stock Exchange in 2014, Alibaba increased its ownership in the company to 31.4%, giving it a 14.8% voting power. Sina Weibo was seen as constituting a crucial component in Alibaba's social commerce strategy (Jia and Han, 2020). Alibaba thus saw an opportunity to combine its leading e-commerce platform with the leading microblogging platform Sina Weibo, thereby competing with emerging and increasingly popular social commerce platforms such as Xiaohongshu and Pinduoduo. In that endeavor, the two platforms developed certain features meant to push sellers on Taobao to use Sina Weibo to market their products. While Sina Weibo had undoubtedly become a leading social media marketing channel for large brands and celebrities, its popularity among the mass of C2C sellers remained limited (Jia and Han, 2020). Nevertheless, this collaboration between a leading e-commerce platform and social media platform constitutes a unique opportunity to analyze how e-commerce sellers optimize their marketing on a separate but integrated platform. Therefore, this study aims to shed light on how e-commerce C2C sellers use integrated social media marketing. We believe a quantitative analysis will improve the understanding of the use of social media marketing for e-commerce and, perhaps more importantly, whether a collaboration between e-commerce and social media platforms results in an effective way for e-commerce C2C sellers to use social media marketing in their sales process. To do this, we use a data sample that was collected from August to October in 2016 and consists of 8,637 C2C sellers (hereafter referred to as Taobao-Weibo users). Because of a collaborative effort between the two platforms, we were able to collect information on C2C sellers on Taobao as well as their activities on Sina Weibo. After collecting all the microblogging posts for the C2C seller sample, we then use machine learning and natural language processing methods to deduce if microblogging posts in the sample incorporate social media marketing content.

This study makes several important contributions. First, and to the best of our knowledge, this is the first large-scale study that empirically analyzes how reputation on both e-commerce and social media platforms is associated with social media marketing activeness by e-commerce C2C sellers. The only work we have identified studying the adoption of social media marketing in the C2C context is by Puriwat and Tripopsakul (2021). However, whereas their study relies on self-reported usage of social media platforms such as Facebook and Line, we collect objective metrics for the same individuals from both e-commerce and social media platforms. As mentioned earlier, a unique collaboration between Taobao and Sina Weibo makes it possible for us to trace Taobao sellers' microblogging activeness, thereby providing a unique opportunity to study C2C social media marketing activeness. Second, we provide empirical evidence that improves our understanding of why the collaboration between e-commerce and social media platforms never became successful in terms of attracting the mass of C2C sellers as we identify a conflict between reputational assets on the two platforms. Third, we shed light on one of the world's most important C2C online platforms by examining to what extent sellers on that platform engage in the marketing of their products on social media. Fourth, we contribute to the literature on reputation and social media marketing. Previous studies have primarily focused on if and how social media marketing affects online reputation. In this study, we define reputational assets as the observable reputation-related quantitative or qualitative metrics accumulated on the digital platforms at the user level over time and instead focus on whether the existing reputational assets are factors shaping social media marketing activeness. This has important implications for our understanding of how online C2C actors make decisions about their marketing activities. Moreover, for the reputational assets considered, we differentiate *earned* reputation assets from *owned* reputation assets for both platforms. Fifth, while the ownership of a social media account is often used as a proxy for social media marketing in the literature, we argue that such a proxy is too crude as it fails to capture actions taken by the owner of that account. By applying machine learning and natural language processing methods, we can identify whether or not C2C sellers' microblogging accounts are characterized by social media marketing activities.

The following section provides a brief introduction to the two Chinese online platforms Taobao and Sina Weibo. Section 3 highlights the concept of social media marketing activeness, as opposed to social media marketing adoption. Section 4 then reviews the related literature on reputation factors and develops the research hypotheses. Section 5 presents the empirical methodology by introducing the data and variables, descriptive statistics, and then specifying the empirical model. Section 6 presents and discusses the empirical results, after which the final section concludes the study.

2 C2C E-Commerce and Social Media in China

Since the early 2010s, a number of e-commerce and social media platforms have emerged in China. C2C e-commerce has long been dominated by Alibaba's Taobao. In the social media space, on the other hand, there are a large number of competitors, each with their own specific characteristics. As the access to international social media platforms such as Facebook, Instagram, and Twitter has remained elusive to Chinese netizens, native platforms have been developed and become hugely popular. While many of these platforms were originally very similar to their international counterparts, they have evolved in ways that make the Chinese social media landscape highly unique. In this section, we first introduce Taobao, the leading C2C e-commerce platform in China. We then provide a brief introduction to Sina Weibo, the leading microblogging platform that we focus on in our empirical analysis, as well as the collaboration between the two platforms that was initiated back in 2013.

2.1 Taobao

Taobao (www.taobao.com) was founded by Alibaba Group in 2003. With 836 million active users (Chen, 2020), it is the largest e-commerce platform for online shopping in China (Gao et al, 2016; Li et al, 2008). It provides C2C retailing services to customers in China as well as markets abroad. It also provides business-to-consumer (B2C) services through its sub-platform Taobao Mall or Tmall for short.

Like other e-commerce platforms, Taobao has introduced a series of interactive features to mitigate uncertainty in online exchange relationships (Pavlou et al, 2007). For example, sellers and buyers can rate each other after each transaction and the accumulated ratings are attached to user profiles and are therefore publicly visible (Wang et al, 2016; Ye et al, 2009). Taobao sellers are also provided an instant messenger tool called Aliwangwang to communicate with customers which can be used to build trust and reputation (Cai et al, 2018).

2.2 Sina Weibo

Chinese citizens are very active online. According to a report by We Are Social (2019), over 1 billion Chinese citizens were social media users in 2019. They spent an average of close to two hours per day on social media. The social media landscape in China differs significantly from that found in most of the rest of the world. One important reason for this is that the Chinese government encourages the use of the internet while maintaining control of information (Feng and Johansson, 2019). Early Chinese adopters of social media started using Twitter soon after its inception in 2006. Soon, several similar microblogging platforms were developed in China (Qin et al., 2017), but none of these gained large traction during this period. Social media platforms quickly became channels where information flowed freely in China. However, after violent riots erupted in the provincial capital Urumqi in Xinjiang in 2009, the Chinese government banned the use of international social media platforms (Sullivan, 2012).

Sina Weibo was subsequently launched in August 2009. Several other microblogging platforms soon emerged. The main competitor is Tencent's platform, but Sina Weibo remains the most popular microblogging platform in China. While many observers tend to call Sina Weibo China's Twitter, there are significant differences between the two platforms. Sina Weibo allowed graphic content to be posted early on, something Twitter incorporated relatively late. One of the main differences between the two is how the platforms are used. Twitter is often seen as a channel used by observers of politics, policies, and social issues. This is not limited to the US, but a phenomenon seen in many other places, including Asian countries such as Indonesia and the Philippines. Weibo users, on the other hand, are often very active on the platform and tend to post more about themselves (Gao et al., 2012).

In terms of the size of the user base, the report by We Are Social (2019) shows that WeChat, a highly successful multi-purpose mobile app developed by Tencent since its release in 2011, remains the most popular social media platform. In a survey, 79% of the respondents said they used WeChat. However, as many as 60% also reported that they use Sina Weibo, suggesting that it remains an important channel for many in China. As noted by Gong et al. (2017), the growth of Weibo has resulted in many businesses exploring it as a channel for their marketing activities.

2.3 Collaboration

On August 5, 2013, Sina Weibo in collaboration with Taobao released a new module specifically designed for Taobao sellers. The new module gives the verified Taobao sellers additional capabilities (compared with other regular Sina Weibo users) to promote their merchandise. More importantly, it grants the verified Taobao sellers an identity of "Tao," which is highlighted in their profiles on the Sina Weibo platform. Moreover, in January 2014, Sina Weibo established cooperation with Alipay (Taobao and Alipay are both subsidiaries of the Alibaba Group) to launch a new platform called Weibo Payment. The new platform makes money transfer much easier for Sina Weibo users. The introduction of these services has encouraged more Taobao sellers to create accounts on Sina Weibo and to make full use of it as a marketing channel for their products. Unlike other social media platforms such as Twitter or Facebook, the explicit collaboration between Sina Weibo and Taobao has made it possible to systematically ascertain whether or not C2C sellers use social media marketing (e.g., via the above-mentioned "Tao" or profile tags such as "Taobao seller"). The gathering of Taobao sellers on Sina Weibo thus gives us a unique opportunity to analyze C2C sellers' marketing behavior on social media.

3 Social Media Marketing Activeness

What drives social media marketing? Previous studies have identified various factors that affect social media adoption by different types of businesses (Dahnil et al., 2014; Dwivedi et al., 2021; Michaelidou et al., 2011; Siamagka et al., 2015). However, to the best of our knowledge, the extant literature has not yet fully investigated the factors behind social media marketing activeness, e.g., whether firms or individual users indeed post to social media after they adopt the social media accounts. Although the post-adoption usage of social media marketing in terms of branded social content has been analyzed and the importance of frequent updates has been highlighted (Ashley and Tuten, 2015), the determinants of the active creation of branded social content are still unclear. One exception is the study by Toubia and Stephen (2013), albeit in a noncommercial setting, which conducts a field experiment on Twitter to identify two types of utilities, intrinsic utility and image-related utility, that motivate users to post to social media. Other more recent efforts of measuring social media marketing activeness include the work by

Moore, et al. (2015) in the B2B and B2C contexts and the one by Puriwat and Tripopsakul (2021) in the C2C context. However, their frequency or intensity measures of social media marketing activeness are based on self-reported survey results rather than objective social media metrics. Therefore, the factors behind social media marketing activeness remain underexplored on a large scale. One practical challenge in the setting of C2C is that while it is straightforward to identify a big company or a famous brand's social media account, it is far more difficult to systematically identify C2C sellers on various social media platforms. Thanks to the above-mentioned collaboration between Taobao and Sina Weibo, a large-scale study of social media marketing activeness in the C2C context becomes feasible because we can cross link the C2C sellers using both platforms.

In a broader sense, as both e-commerce and social media platforms heavily rely on user interactions, the persistence of purchase intention from buyers' perspective as well as social media marketing adoption and usage from sellers' perspective can only be maintained with relationship-building and trust (Chiu, et al., 2012; Herjanto and Amin, 2020; Melnik and Alm, 2002). Moreover, previous studies have long considered reputation systems as the primary approach to establish trust in the sharing economy (Ter Huurne, et al., 2017; Thierer, et al., 2015). We therefore devote the next section to the discussion of reputation factors.

4 Reputational Assets on Digital Platforms

Previous studies have roughly explored three different groups of antecedents of social media marketing adoption, and they are respectively personal, organizational, and external factors (Guesalaga, 2016; Pascucci, et al., 2018). Focusing on B2B organizations, Michaelidou et al. (2011) highlight a large variation in managerial mindsets towards social media and show that this has direct implications for social media adoption. Siamagka et al. (2015) find that features related to perceived usefulness and organizational innovativeness are positively associated with social media adoption among B2B organizations. In the context of SMEs, Dahnil et al., (2014) identify both internal and external factors that could affect the adoption of social media marketing. The internal factors include end users' training and knowledge, organizational resources, leaders' attitude whereas the external factors include business and economic environments. More recently the internal factors have been further specified at individual and organizational levels in the B2C and B2B contexts (Guesalaga, 2016; Pascucci, et al., 2018; Dwivedi et al., 2021).

In the C2C context, the organizational factors become irrelevant. When it comes to the post-adoption social media marketing activeness within the platforms, the initial technical and perceptual barriers should be less of a concern. As stated above, the persistence of social media usage needs to be maintained by relationship and trust building with other users. As a result, we empirically explore the linkages between social media marketing activeness and reputational assets on digital platforms, which are often regarded as the chief means to establish trust in the sharing economy (Ter Huurne, et al., 2017; Thierer, et al., 2015). We focus on reputation as its importance in e-commerce has long been recognized by academics as well as practitioners (Dellarocas, 2003; Melnik and Alm, 2002; Pavlou and Dimoka, 2006). Since most participants on digital platforms are anonymous, reputation signaled by user profiles and other relevant metrics is an invaluable means for customers to mitigate uncertainty when conducting online transactions (Kim and Krishnan, 2015). When developed properly, reputation can help businesses amplify their message and reach out to more customers and in some cases charge a premium for the products and services. On the other hand, a business can be ruined quickly if its customers stay away because of a bad reputation.

Closely related to the image-related utility that motivates users to post to social media (Toubia and Stephen, 2013), reputation is potentially a key determinant of social media marketing activeness. We, therefore, extend the study by Toubia and Stephen (2013) to the commercial setting of C2C social media marketing. We measure the motivating utilities by reputational assets, which are defined as the observable reputation-related quantitative or qualitative metrics accumulated on the digital platforms at the user level over time. As Floreddu et al. (2014) point out, "reputation is a crucial intangible asset formed as a result of past actions and stakeholders' direct experience". For an online C2C seller with a social media account, accumulative reputational assets are generated from both the shopping and social media sites. In this study, we classify these reputational assets into two groups depending on whether they are accumulatively *earned* or *owned* by C2C sellers. A reputational asset is defined as *earned* reputation if it is *received* by C2C sellers and is therefore relatively *beyond their direct control*, for example, the rating scores received from Taobao or the number of followers and the verifications received from Sina Weibo. On the other hand, a reputational asset is defined as *owned* reputation if it is *provided* by C2C sellers and is therefore relatively *under their direct control*, for example, the number of years operating on Taobao or the resulted historical sales and the paid membership on Sina Weibo. The paid membership could be categorized as paid reputation but is rather grouped into owned reputation for simplicity since it is under C2C sellers' direct control. As detailed below, we also consider the non-accumulative measure, gender, as another potential determinant of C2C social media marketing and other non-accumulative measures such as region and price, quantity and category of products as controls.

Previous studies on reputation and social media marketing typically examine how social media activity determines reputation, in particular from a business-to-business (B2B) or business-to-consumer (B2C) perspective. For example, Horn et al. (2015) examine the threat of social media to corporate reputation by customers, employees, and the corporation itself. Liu et al. (2019) analyze how social media marketing affects the reputation of luxury fashion brands and shows that brand exclusivity is not negatively affected by strong social media exposure in China. While studies such as the ones mentioned here help improve our understanding of how social media activity can influence reputation, here we instead focus on how online reputation affects social media activity. C2C platforms often provide measures of activity, sales volumes, or reviews that help customers navigate between alternative providers of the products they seek as the reputation feedback system and web site quality can positively influence C2C e-commerce trust (Jones and Leonard, 2008; Leonard and Jones, 2021; Yang, et al., 2007). How do reputational assets accumulated on the C2C platform influence social media marketing activeness? C2C sellers can be more motivated and confident when doing social media marketing if they have secured good ratings received from previous customers on the C2C platform. That is, a potential customer attracted via social media marketing will eventually find out what ratings the C2C seller has received by following the shopping link embedded. In other words, a cemented reputation received on the C2C platform helps amplify the message that the C2C seller wants to spread across other channels, such as social media platforms. We, therefore, start with the hypothesis that earned reputation is positively associated with social media marketing activeness by C2C sellers.

H1. *Earned reputation on a C2C platform is positively associated with social media marketing activeness by C2C sellers.*

Earned reputation can be built through a variety of channels. Previous studies also find that third-party recognition can positively influence C2C e-commerce trust (Jones and Leonard, 2008; Leonard and Jones, 2021; Yoon and Occena, 2015). In addition to recognition on the e-commerce platform itself, C2C sellers can build up their reputation on other channels, including the social media platforms they opt to use for their marketing efforts. We hypothesize that there is a positive relationship between earned reputation, this time on the social media platform, and social media marketing activeness. The reason for this is similar to why we believe there is a positive and more direct relationship between C2C platform earned reputation and social media marketing activeness: a potential customer attracted via social media marketing can directly observe the C2C seller's earned reputation on the social media platform without following a click-through link.

H2. *Earned reputation on a social media platform is positively associated with social media marketing activeness by C2C sellers.*

Earned reputation is thus likely to act as an amplifier for social media marketing activities. What about owned reputational assets such as overall sales as seen in the form of a ranking on the C2C platform? One could argue that higher sales would lead to C2C being able to allocate more capital to social media marketing activities. It has been shown that there is a direct link between sales (or shop age) and reputation, especially in the case of Taobao (Gao et al., 2016). Moreover, the collaboration between Taobao and Sina Weibo encourages established C2C sellers with high overall sales to use social media as a verified "Tao" identity requires a minimum threshold of overall sales. This would imply a positive relationship between sales and social media marketing activeness. On the other hand, sales, as proxied by an experience ranking on the C2C platform, may instead suggest a measure of effort exclusively dedicated to the platform and that there is less need for omnichannel marketing activities. If this holds, we would thus expect a negative relationship between sales and social media marketing activeness. These two opposing forces make it difficult to determine which of these factors dominates the relationship between sales and social media marketing activeness theoretically. Similarly, the relationship between shop age, as another owned reputational asset, and social media marketing activeness is unclear. We, therefore, leave it to the empirical analysis to determine how owned reputation factors on a C2C platform are associated with social media marketing activeness among C2C sellers.

H3. *Owned reputation on a C2C platform is associated with social media marketing activeness by C2C sellers.*

When it comes to the owned reputational assets on social media platforms such as paid membership and number of follows, their expected relation to social media marketing activeness by C2C sellers is unambiguously positive. Previous studies have established the linkage between the expected social and economic returns from social media platforms and the willingness to pay for memberships and services (Li, et al., 2014; Vock, et al., 2013). Because the social media marketing activity is eventually deployed on a social media platform, a high level of prior owned reputation, which can also be considered as an accumulated investment in terms of both time and money spending, would indicate a strong tendency to use it for marketing activity and investment returns.

H4. *Owned reputation on a social media platform is positively associated with social media marketing activeness by C2C sellers.*

Finally, Figure 1 below summarizes our hypotheses regarding the factors of social media marketing activeness as well as other control variables.

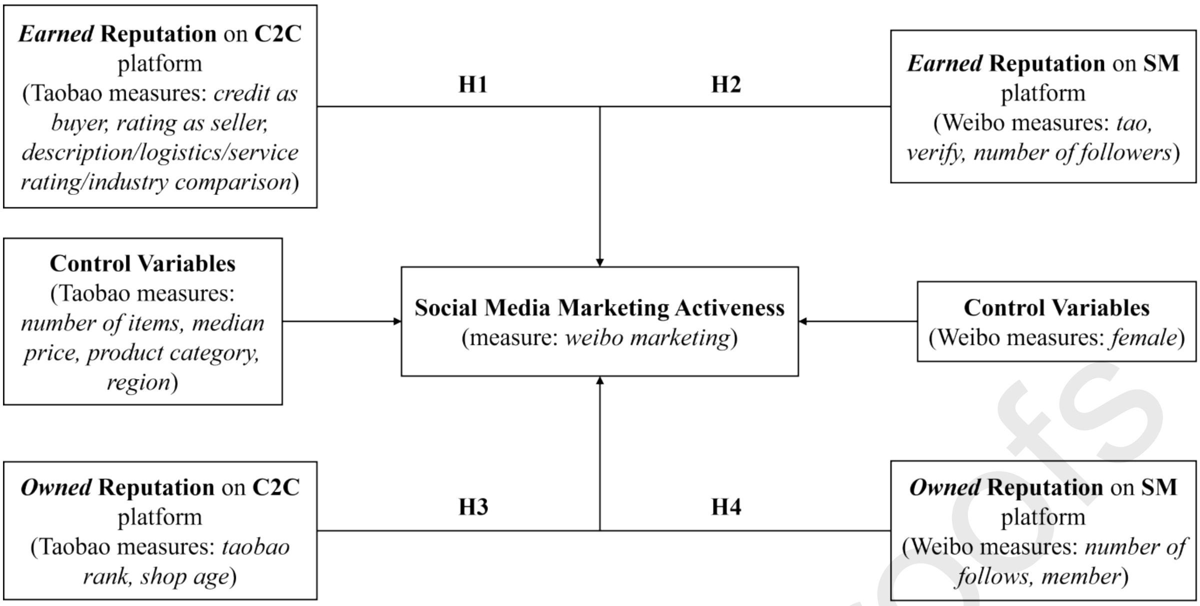


Figure 1: Reputational Assets and Social Media Marketing Activeness

5 Methodology

5.1 Data

We initially identified Taobao sellers on Sina Weibo in November 2014. We select this period because November 11, the largest online shopping day in the world, occurs during this month. As a result, Taobao sellers should have substantial economic incentives to be active during that period. We identified 12,744 sellers who added the links to their Taobao shops. We then gathered data on them from both platforms during the period August to October in 2016, which allowed us to establish reliable measures for variables such as median price and the number of items for each Taobao shop. After removing the ones with invalid links or missing information for key variables, our final dataset contains 8,637 Taobao C2C sellers with a Sina Weibo account.

Figure 1 shows the geographical distribution of our sample of Taobao-Weibo users. The regions with over 500 Taobao-Weibo users are Guangdong (1,103), Zhejiang (758), Shanghai (649), and Beijing (580), which are commonly known as the most e-commerce-active areas in China (Wang et al, 2019). On the other hand, the vast majority of West, Northwest, and Northeast China have Taobao-Weibo users below 50. Geographically speaking, our sample thus provides a relatively good representation of Taobao sellers. Finally, although not shown in Figure 1, 250 users are identifying their location as “Overseas”. The non-trivial number of Taobao sellers outside of China shows the rise of ‘purchasing agents’ for Chinese domestic customers (Wang et al, 2019).

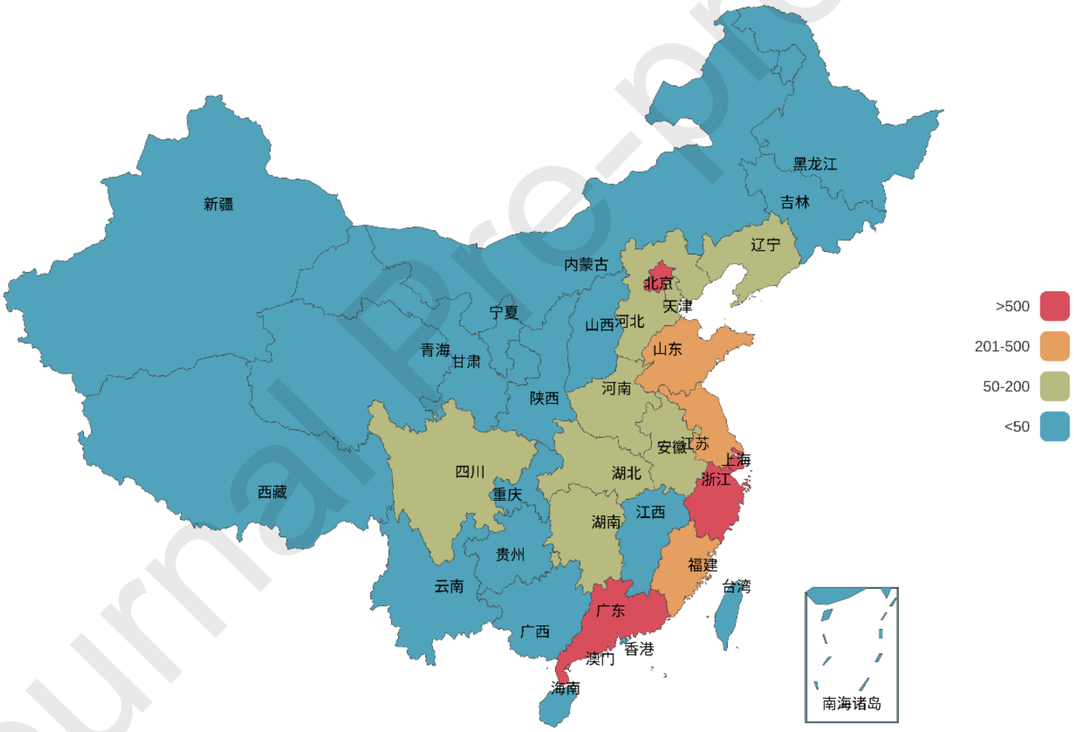


Figure 2: Geographic Distribution of Taobao-Weibo Users

Figure 2 shows the distribution of Taobao-Weibo users by the main product category on Taobao. The top categories to some extent support the notion of a certain level of ‘feminism’ in China's e-commerce industry (Yu and Cui, 2019), as the dominating category is "Clothes and bags" with 3,392 users, followed by "Mother and child" with 996 users and "Beauty care" with 933 users.

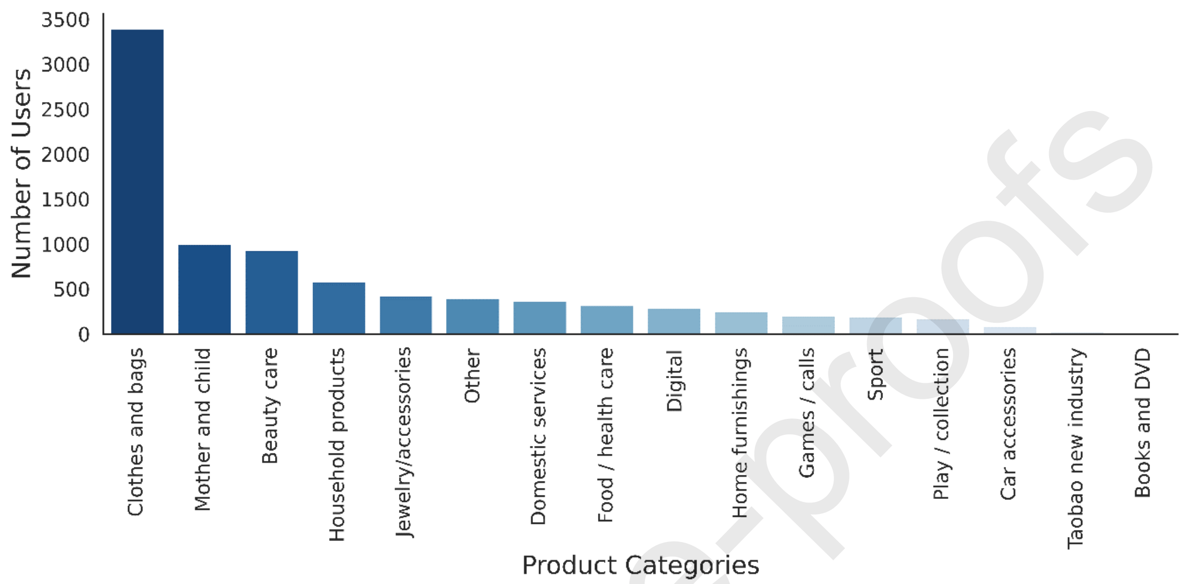


Figure 3: Product Distribution of Taobao-Weibo Users

5.2 Variables

A simple way to measure a Taobao-Weibo user's activeness is to see if the user posted microblogs or not in the observation period. To do so, we define a dummy variable, *weibo use*, equaling 1 if at least one Weibo was posted in the observation period (i.e., from August to October in 2016) or 0 otherwise. However, the postings by individual Taobao-Weibo users may not be marketing-oriented. For instance, a user may post on Weibo to share music or other aspects of his or her social life (Wang et al, 2019). As a result, Weibo activeness does not necessarily imply Weibo marketing activeness. We thus need to take a closer look at the content of Weibo posts to ensure that our key variable captures marketing efforts. To automate the process of classifying Weibo posts into marketing and non-marketing types, we employ a process in which we train a variety of machine learning algorithms and then select the best binary classifier in terms of test performance.

The sample used for comparing binary classifiers contains 5,000 microblogs randomly selected from 774,429 microblogs collected during the periods of November in 2014 and from July to October in 2016 (Wang et al, 2019). The 5,000 microblogs are manually labeled as either marketing or non-marketing type. The machine learning task is defined as predicting whether a given microblog is of marketing type (1 if true or 0 otherwise) based on its textual content. The procedure can be described as follows:

- The microblogs (which are segmented into words, phrases, etc.) are tokenized, and stop words are removed.
- The microblogs are mapped into feature vectors with the term frequency-inverse document frequency (TF-IDF) method.
- The sample of 5,000 microblogs is randomly divided into a training set of 4,000 microblogs and a test set of 1,000 microblogs.
- A 10-fold cross-validation is applied to tune the hyperparameters of the classifiers.
- The final evaluation with the test set is based on standard performance indicators such as precision, recall, accuracy, and the area under the receiver operating characteristics curve (AUC) score (Lessmann et al, 2008).

The binary classifiers considered include logistic regression, decision tree (CART), random forests, and multinomial naive Bayes (MNB). After running the classification and test procedures for the various algorithms, we find that the MNB classifier performs the best.

The MNB classifier is based on the probability of being in class c given a document d:

$$P(c \mid d) \propto P(c) \prod_{1 \leq k \leq n_d} P(t_k \mid c)$$

$P(t_k \mid c)$ is the probability that feature t_k appearing in a document belonging to class c. Note that document d has n_d features.

The class that a given document d is assigned to is determined by the *maximum a posteriori (MAP)* class c_{map}:

$$c_{map} = \underset{c \in \mathbb{C}}{\operatorname{argmax}} \hat{P}(c \mid d) = \underset{c \in \mathbb{C}}{\operatorname{argmax}} \hat{P}(c) \prod_{1 \leq k \leq n_d} \hat{P}(t_k \mid c)$$

$\hat{P}(c)$ can be simply estimated as the number of documents belonging to class c divided by the total number of documents in the training set, whereas:

$$\hat{P}(t_k \mid c) = \frac{N_{ct_k} + \alpha}{N_c + \alpha n}$$

N_{ct_k} is the fractional counting of TF-IDF feature t_k appears in a sample of class c in the training set and N_c is the fractional counting of all TF-IDF features for class c in the training set. n is the number of TF-IDF features for class c. Finally, $\alpha \geq 0$ is the smoothing prior that accounts for features not present in the training set and prevents zero probabilities in future computations (Manning et al., 2008).

In our case, the MNB classifier's AUC score and precision equal 0.96 and 0.91, respectively. We thus used it to classify all the 774,429 microblogs (Wang et al, 2019), which include the sample of microblogs used in this paper. Once we have run the MNB classifier algorithm on our sample, we find that out of 170,802 microblogs posted by the 8,637 Taobao-Weibo users, 46,326 microblogs are classified as marketing type.

Using the above results, we can define a dummy variable, *weibo marketing*, that equals 1 for a Taobao-Weibo user if the user posted at least one Weibo classified as marketing type by the MNB classifier during the observation period, i.e. from August to October in 2016, or 0 otherwise. Out of the 8,637 Taobao-Weibo users in our sample, 1,454 or 16.8% of the users posted at least one marketing Weibo during the observation period. We further check the heterogeneity of marketing activeness in terms of geographical distribution. Figure A1 in the appendix shows the number of actively marketing Taobao-Weibo users (i.e., *weibo marketing* equalling 1) by region. The region with the most Taobao-Weibo users in our sample, Guangdong (see Figure 2), remains the most marketing active one in Figure A1, with 17.8% of its users (196 out of 1,103) being marketing-active. However, the second most marketing active region becomes Shanghai (131 active users) whereas the third and fourth places are taken by Zhejiang (109 active users) and Beijing (107 active users) respectively.

Similarly, we check the heterogeneity of marketing activeness in terms of product category. Figure A2 in the appendix presents the number of actively marketing Taobao-Weibo users by product category, with a similar ranking to that in Figure 3. Although only 18.6% of its users are classified as actively marketing Taobao-Weibo users (630 out of 3,392), the most populous product category "Clothes and bags" (see Figure 3) still dominates in Figure A2. However, the second most marketing-active category is "Beauty care" (with 222 users being marketing-active) whereas the second most populous category "Mother and child" only takes third place in Figure A2 with 116 users being marketing-active.

When it comes to the potential driving factors behind social media marketing activeness, we consider the reputational assets both *earned* and *owned* by sellers on both C2C and social media platforms. First, there are binary measures of earned reputation on both platforms. Sina Weibo, for example, has both "Tao" and "Verify" identities that provide a signal of trust as the verification of the identities is associated with some thresholds of credibility and prestige (Wang et al, 2019). Second, continuous earned reputation variables include the number of followers on Sina Weibo and rating as seller, credit as buyer (as a C2C seller can also be a buyer on Taobao and be rated by sellers), and specific rating scores at both absolute and relative (i.e., industry comparison) levels regarding product description (i.e., how realistic is the online product description after a buyer receiving the product in hands?), logistics (i.e., how long is the delivery?) and customer services (i.e., how responsive is the seller when a question is raised?). Table A1 in the appendix provides detailed definitions of the variables.

On the other hand, the owned reputational assets include historical sales measured by Taobao rank and shop age on Taobao and paid membership and number of follows on Sina Weibo. We focus on sales here as the other variables are self-explanatory. While it is practically difficult to gather information about sales volume in a traditional sense (e.g., measured by total transaction value or assets) in a virtual online setting, Taobao does provide Taobao rank as a proxy for sales volume (Gao et al, 2016). A similar approach was taken by previous work on Amazon (Chevalier and Mayzlin, 2006). Taobao rank from the lowest to the highest goes from 1 to 20. Each rank corresponds to an accumulated score based on the customer rating after each transaction. After each transaction, a customer can give a rating of “positive” (a +1 score), “neutral” (a 0 score), or “negative” (-1 score). Importantly, if no rating is provided by the buyer, a rating of “positive” (a +1 score) will be assumed by default. Most customers do not leave any ratings after transactions. As a result, most of the ratings are positive and the accumulated score can thus be seen as serving as a proxy for the total number of historical transactions. In other words, the sellers with a larger number of sales transactions would typically have higher ranks. Therefore, Taobao rank effectively serves as a measure of the historical number of sales.

We also consider the binary variable, gender, as a potential determinant of C2C social media marketing activeness as well as control for region, product category, median price, and the number of items displayed when doing empirical analyses. Again, the data of the Taobao-Weibo users were collected for three months to establish reliable measures for variables such as median price and the number of items. As a result, except for static measures such as gender, all the measures are monthly averaged over the observation period from August to October in 2016.

5.3 Descriptive Statistics

Table 1 shows the descriptive statistics of the variables collected from both platforms. Table A1 in the appendix shows the definitions of the variables. In Table 1, the variables 1-8 are from Sina Weibo whereas the ones 9-20 are from Taobao. Again, there are 8,637 Taobao-Weibo users identified in our sample. For the observation period, 32% of them posted at least one microblog (as measured by the *weibo use* variable) and 17% of them posted at least one marketing microblog according to our MNB classifier (as measured by the *weibo marketing* variable). Moreover, 39% and 11% of them have obtained the "Tao" and "Verify" identities respectively, 33% of them have paid memberships, and 75% of them are female according to their profiles. For readers' convenience, we add the variable types in the parentheses next to the variable names, where "d" stands for dependent variables, "c" stands for control variables, "e" stands for earned reputational assets, and "o" stands for owned reputational assets. Table 2 shows the pairwise correlations among the variables. Note that the variables *description industry comparison*, *logistics industry comparison*, and *service industry comparison* are highly correlated with each other and the variables *description rating*, *logistics rating*, and *service rating* are highly correlated with each other (all with the correlation coefficients above 0.9). To avoid the collinearity problem, the two groups of three are averaged to the variables *industry comparison average* and *rating average* respectively when used in the regressions.

Table 1: Descriptive Statistics

CODE	VARIABLES	(1) # Obs	(2) mean	(3) sd	(4) p10	(5) p50	(6) p90
	Weibo Variables						
1	weibo marketing (d)	8,637	0.168	0.374	0	0	1
2	weibo use (d)	8,637	0.322	0.467	0	0	1
3	number of follows (o)	8,637	462.2	531.4	47	273	1,196
4	number of followers (e)	8,637	2,938	18,819	38	410	4,501
5	tao (e)	8,637	0.389	0.488	0	0	1
6	verify (e)	8,637	0.11	0.312	0	0	1

7	member (o)	8,637	0.327	1.274	0	0	0
8	female (c)	8,637	0.746	0.435	0	1	1
	Taobao Variables	8,637					
9	number of items (c)	8,637	114.4	374.3	2	26	273.3
10	median price (c)	8,637	240.9	545.3	18.29	92.59	474.7
11	taobao rank (o)	8,637	8.612	3.038	4	9	12
12	shop age (o)	8,637	5.825	2.299	2.949	5.568	9.13
13	credit as buyer (e)	8,637	619	743.6	40	388	1474
14	rating as seller (e)	8,637	99.11	1.147	97.93	99.44	100
15	description industry comparison (e)	8,637	35.9	38.12	-0.423	29.51	100
16	description rating (e)	8,637	3.408	2.117	0	4.767	5
17	logistics industry comparison (e)	8,637	36.51	38.21	-0.0267	30.98	100
18	logistics rating (e)	8,637	3.42	2.126	0	4.8	5
19	service industry comparison (e)	8,637	37.15	38.47	0	31.79	100
20	service rating (e)	8,637	3.426	2.128	0	4.8	5

"d" stands for dependent variables, "c" stands for control variables, "e" stands for earned reputational assets, and "o" stands for owned reputational assets.

Table 2: Correlation Matrix

#	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
1	1																			
2	0.652	1																		
3	0.127	0.112	1																	
4	0.149	0.124	0.054	1																
5	0.135	0.106	0.097	0.032	1															
6	0.146	0.127	0.099	0.21	0.152	1														
7	0.307	0.271	0.122	0.285	0.111	0.191	1													
8	0.07	0.12	-0.075	0.028	-0.022	-0.072	0.059	1												
9	0.068	0.031	0.008	0.009	0.08	0.016	0.05	-0.004	1											
10	0.006	-0.011	0.02	-0.006	-0.029	0.002	0.022	-0.045	0.007	1										
11	0.087	0.01	0.025	0.112	0.302	0.105	0.146	-0.015	0.216	-0.085	1									
12	0.037	0.014	0.002	0.018	0.116	0.021	0.045	0.067	0.13	0.02	0.435	1								
13	0.103	0.136	-0.01	0.058	0.046	0.028	0.13	0.139	0.067	-0.031	0.172	0.302	1							
14	0.058	0.085	-0.049	0.026	-0.004	0.01	0.085	0.097	0.022	0.078	0.026	0.118	0.124	1						
15	0.078	0.08	-0.006	0.006	0.05	0.022	0.064	0.032	0.036	0.041	0.05	0.025	0.077	0.169	1					
16	0.122	0.069	0.025	0.048	0.191	0.075	0.095	-0.042	0.18	-0.078	0.448	0.114	0.074	-0.002	0.61	1				
17	0.083	0.083	-0.016	0.011	0.054	0.016	0.063	0.043	0.033	0.023	0.069	0.026	0.079	0.158	0.904	0.611	1			
18	0.123	0.07	0.024	0.048	0.191	0.073	0.095	-0.041	0.181	-0.08	0.449	0.114	0.074	-0.005	0.598	0.998	0.616	1		
19	0.086	0.089	-0.01	0.011	0.052	0.016	0.064	0.033	0.034	0.025	0.058	0.018	0.072	0.162	0.925	0.616	0.926	0.613	1	
20	0.123	0.07	0.025	0.048	0.19	0.073	0.095	-0.042	0.18	-0.08	0.448	0.113	0.073	-0.005	0.597	0.998	0.608	0.998	0.62	1

5.4 Model Specification

Since the outcome variable *weibo marketing* is binary, we apply a standard logistic regression model to understand the impact of the variables collected from both platforms on Weibo activeness. Our main empirical model is as follows:

$$Pr(weibo\ marketing_i = 1 \mid Weibo_i, Taobao_i) = \frac{exp(\alpha_0 + \beta_1 Weibo_{1i} + \beta_2 Weibo_{2i} + \cdots + \gamma_1 Taobao_{1i} + \gamma_2 Taobao_{2i} + \cdots)}{1 + exp(\alpha_0 + \beta_1 Weibo_{1i} + \beta_2 Weibo_{2i} + \cdots + \gamma_1 Taobao_{1i} + \gamma_2 Taobao_{2i} + \cdots)}$$

where $weibo\ marketing_i = 1$ if user i posted at least one marketing Weibo during the observation period and $Weibo_i$ represents variables collected from the platform of Weibo whereas $Taobao_i$ represent variables collected from the platform of Taobao.

6 Empirical Results

6.1 Baseline Results

Table 3 shows the results of logistic regressions with different model specifications. The first and the second columns include the Weibo variables and the Taobao variables respectively. The third to fifth columns include the variables from both platforms by varying the fixed effects we control for.

As can be seen in the table, the results across different specifications are quite stable. We find that whether a registered Taobao seller is marketing-active in Weibo is associated with factors from both platforms. Looking first at earned reputation, our dependent variable is positively associated with the two verified identities *tao* and *verify* and *number of followers* on the Weibo platform. This result supports hypothesis H2, namely that earned reputation on social media is positively associated with social media marketing activeness. If we instead look at earned reputation on the Taobao platform, we find a similar pattern. Our dependent variable is positively related to both proxies for earned reputation, *credit as buyer* and *rating as seller*. These results provide support for our hypothesis H1, i.e. that earned reputation on the C2C platform is positively associated with social media marketing activeness. We also find that owned reputational assets on Sina Weibo, as measured by *number of follows* and *member*, matter positively, which supports our hypothesis H4.

Next, we focus on the owned reputation variables on the Taobao side of the dataset. Our key explanatory variables here are *taobao rank*, which we argue serves well as a proxy for the number of sales on the Tabao platform, and *shop age*, which measures the number of years between the Taobao shop opening date and the data collection date. As we mentioned in Section 3 of the paper, there are two potential and opposite effects that sales may have on social media marketing activeness. Focusing on the main regression model in Column 5, we see that the coefficient for *taobao rank* is negative and significant. This lends support to the argument that more sales on Taobao are associated with C2C sellers feeling less need to engage in social media marketing activities. The alternative measure, *shop age*, also behaves similarly when no fixed effects of region or product category are controlled for in Column 2. But its significance is crowded out by *taobao rank* once the fixed effects are present. Looking at the other control variables from the Taobao part of the dataset, it appears that *number of items* and *median price* are by and large insignificantly related to Weibo activeness. Note that the significance of *number of items* disappears after the region fixed effects are included. Also recall that, to avoid the collinearity problem, *rating average* is the average across *description rating*, *logistics rating*, and *service rating* and its coefficient is consistently positive and significant. On the other hand, *industry comparison average* is the average across *description industry comparison*, *logistics industry comparison*, and *service industry comparison,* and its coefficient is insignificant.

While hitherto a relatively unexplored topic, previous research has shown that attitudes and behavior on social media differ across gender. In general, women tend to be motivated by being able to uphold close ties and obtain social information while men tend to be motivated by the opportunity to obtain more general information (Krasnova et al., 2017). Differences between males and females can also be found in how they use online information for purchasing decisions. Nadeem et al. (2015) find that the influence of peer recommendations is significantly stronger on the attitudes of females. Women constitute an important customer segment in B2C and C2C e-commerce. This is evident when we look at how social media marketing is evolving. For example, in the quickly growing subdomain of influencer marketing, as many as 77% of influencers are women who most often reach out to and influence other women (Gesenhues, 2019). This again suggests that women tend to place a high value on information from their peers when making purchasing decisions. The geographic focus of this study is China, a country where it is commonly known that women drive online sales and fuel what has been labeled the 'sheconomy'. Even though the country exhibits a significant gender imbalance with about 31.6 million more males, recent official figures show that women account for 55% of online spending (Liu, 2019). A typical example of this phenomenon is the social shopping platform Xiaohongshu (RED), where shoppers go to find shopping information, tips, and reviews in the form of text, images, and video. In 2019, over 80% of its users were female (Lim, 2020). Looking at gender as a potential determinant of social media marketing activeness in our empirical analysis, we find that the coefficient for *female* C2C sellers is significant and positive in both Columns 1 and 3-5. This finding ties into discussions on the issue of 'feminism' in China's e-commerce industry found in the literature and on the emergence of the so-called 'sheconomy'.

Table 3: Logistic Regression on *weibo marketing*

VARIABLES	(1) Weibo	(2) Taobao	(3) Weibo+Taobao	(4) Weibo+Taobao	(5) Weibo+Taobao
number of items (c)		0.000111	0.000258***	0.000238***	0.000124
		(7.89e-05)	(8.62e-05)	(8.89e-05)	(9.25e-05)
median price (c)		0.000104	3.95e-05	3.32e-05	5.54e-05

		(6.66e-05)	(5.73e-05)	(5.93e-05)	(7.19e-05)
taobao rank (o)		0.0205	-0.0445***	-0.0476***	-0.0470***
		(0.0162)	(0.0142)	(0.0143)	(0.0178)
shop age (o)		-0.0370**	-0.00701	-0.00720	-0.0193
		(0.0178)	(0.0155)	(0.0156)	(0.0189)
credit as buyer (e)		0.000260***	0.000179***	0.000177***	0.000183***
		(4.25e-05)	(3.92e-05)	(3.99e-05)	(4.59e-05)
rating as seller (e)		0.175***	0.0974***	0.120***	0.117**
		(0.0478)	(0.0345)	(0.0359)	(0.0480)
industry comparison average (e)		-0.000697	-0.000458	-0.00101	-0.00107
		(0.00137)	(0.00115)	(0.00116)	(0.00144)
rating average (e)		0.165***	0.151***	0.165***	0.147***
		(0.0302)	(0.0246)	(0.0253)	(0.0318)
number of follows (o)	0.000420***		0.000441***	0.000469***	0.000371***
	(5.33e-05)		(5.40e-05)	(5.44e-05)	(6.88e-05)
number of followers (e)	1.13e-05***		1.16e-05***	1.06e-05***	1.33e-05***
	(2.83e-06)		(2.88e-06)	(2.82e-06)	(3.93e-06)
tao (e)	0.535***		0.494***	0.481***	0.590***
	(0.0623)		(0.0660)	(0.0665)	(0.0813)
verify (e)	0.519***		0.514***	0.515***	0.558***
	(0.0891)		(0.0898)	(0.0904)	(0.111)
member (o)	0.348***		0.326***	0.319***	0.324***
	(0.0199)		(0.0204)	(0.0205)	(0.0248)
female (c)	0.504***		0.478***	0.420***	0.342***
	(0.0774)		(0.0791)	(0.0809)	(0.0993)
Constant	-2.725***	-20.21***	-12.61***	-15.23***	-14.82***
	(0.0842)	(4.750)	(3.413)	(3.562)	(4.767)
product category fixed effects (c)	No	Yes	No	Yes	Yes
region fixed effects (c)	No	Yes	No	No	Yes
Observations	8,637	5,522	8,637	8,637	5,522

Standard errors in parentheses; ***, **, and * denote significance at the 1%, 5% and 10% level, respectively.

“c” stands for control variables, “e” stands for earned reputational assets, and “o” stands for owned reputational assets.

6.2 Robustness Checks

We carry out a robustness check in which we instead use *weibo use* as the dependent variable. Here, the dependent variable equals 1 if the Sina Weibo user posted on his or her microblog (not necessarily marketing oriented) during the observation period. As a result, *weibo use* is a more generous measure of social media marketing activeness than *weibo marketing*.

The results of the new set of logistic regressions are presented in Table 4. The main findings remain qualitatively unchanged. In particular, the proxies for earned reputation on the two platforms are still highly significant and positively associated with *weibo use*. Similarly, the coefficient for gender, *female*, is still significant and positive. Finally, the proxy for sales, *taobao rank*, is once more negatively associated with the dependent variable. These results lend further support to our initial findings and our working hypotheses presented in Section 3.

Table 4: Logistic Regression on *weibo use*

VARIABLES	(1) Weibo	(2) Taobao	(3) Weibo+Taobao	(4) Weibo+Taobao	(5) Weibo+Taobao
number of items (c)		1.76e-05	0.000140*	0.000127	1.02e-05
		(7.13e-05)	(7.95e-05)	(8.13e-05)	(7.92e-05)
median price (c)		-2.48e-05	-8.35e-05*	-8.62e-05*	-7.17e-05
		(5.76e-05)	(4.86e-05)	(4.99e-05)	(6.18e-05)
taobao rank (o)		-0.0297**	-0.0839***	-0.0875***	-0.0870***
		(0.0132)	(0.0113)	(0.0114)	(0.0144)
shop age (o)		-0.0257*	-0.0120	-0.0124	-0.0118
		(0.0146)	(0.0125)	(0.0126)	(0.0153)
credit as buyer (e)		0.000353***	0.000284***	0.000281***	0.000275***
		(3.93e-05)	(3.47e-05)	(3.53e-05)	(4.09e-05)
rating as seller (e)		0.174***	0.117***	0.135***	0.131***
		(0.0360)	(0.0254)	(0.0263)	(0.0363)
industry comparison average (e)		0.000682	0.00125	0.000791	0.000335
		(0.00113)	(0.000933)	(0.000945)	(0.00118)
rating average (e)		0.0886***	0.0604***	0.0751***	0.0748***
		(0.0233)	(0.0187)	(0.0192)	(0.0244)
number of follows (o)	0.000334***		0.000356***	0.000369***	0.000296***
	(4.62e-05)		(4.69e-05)	(4.72e-05)	(6.01e-05)
number of followers (e)	2.36e-05***		2.53e-05***	2.37e-05***	2.55e-05***
	(4.37e-06)		(4.50e-06)	(4.42e-06)	(5.60e-06)
tao (e)	0.301***		0.395***	0.386***	0.513***
	(0.0502)		(0.0536)	(0.0539)	(0.0663)
verify (e)	0.429***		0.440***	0.446***	0.398***
	(0.0799)		(0.0811)	(0.0814)	(0.101)
member (o)	0.378***		0.365***	0.359***	0.362***
	(0.0243)		(0.0246)	(0.0246)	(0.0297)
female (c)	0.677***		0.598***	0.546***	0.526***
	(0.0608)		(0.0623)	(0.0636)	(0.0800)
Constant	-1.772***	-18.66***	-13.02***	-15.06***	-14.79***
	(0.0645)	(3.575)	(2.516)	(2.610)	(3.605)
product category fixed effects (c)	No	Yes	No	Yes	Yes
region fixed effects (c)	No	Yes	No	No	Yes
Observations	8,637	5,538	8,637	8,637	5,538

Standard errors in parentheses; ***, **, and * denote significance at the 1%, 5% and 10% level, respectively.

“c” stands for control variables, “e” stands for earned reputational assets, and “o” stands for owned reputational assets.

Another robustness check we have tried is to recompute the variable *weibo marketing* by randomly reclassifying some microblogs. To do so, we randomly select 10% of the microblogs and reclassify them into marketing type if they are non-marketing type according to our MNB classifier and vice versa. Even though our MNB classifier has achieved impressive performance scores (AUC score and precision equaling 0.96 and 0.91 respectively), there might still be some room for improvement with other machine learning models, such as neural-network-based ones. Therefore, this additional check will test how sensitive our results react to the accuracy of our MNB classifier. The results are presented in Table A2 in the appendix. Again, our findings are robust as the results remain qualitatively unchanged.

7 Conclusions

Social media marketing has become a standard tool for businesses of all sizes. In this study, we contribute to the literature on social media marketing by examining a unique dataset of C2C sellers' social media marketing activeness on China's leading microblogging platform. Based on a sample of over 8,000 C2C sellers registered on both Taobao, China's largest C2C online shopping platform, and Sina Weibo, China's largest microblogging platform, this paper empirically explores the factors influencing C2C sellers' social media marketing use. We find that whether a registered Taobao seller is actively marketing in Weibo is associated with factors on both platforms. In particular, it relates to reputational assets accumulated on both the C2C e-commerce platform as well as the microblogging platform. Interestingly, we also find that female users are more likely to be active, which further confirms the 'feminism' in China's e-commerce industry and the advancement of the so-called 'sheconomy'. Finally, owned reputational assets such as Taobao rank, a proxy for its turnover, are negatively associated with social media marketing activeness. One potential reason for this is that owners of highly ranked Taobao shops feel that they are less dependent on social media marketing efforts.

We contribute to the literature in multiple ways. First of all, we identify several new factors that potentially influence social media marketing activeness, including reputational assets classified into earned and owned reputational assets on the C2C as well as the microblogging platform and gender. Our finding that earned reputation on both platforms is positively associated with social media marketing activeness is particularly interesting, as the relationship is often discussed in terms of how social media marketing affects reputation, not the other way around. Perhaps more importantly, we also shed light on the limitations of collaborative efforts between C2C e-commerce and social media platforms as our empirical analysis provide evidence suggesting that there is an inherent conflict between reputation on the different platforms, something that has a direct effect on C2C sellers' willingness to engage in marketing activities on social media platforms. This is important as the collaboration between a C2C e-commerce platform and a social media platform is a potential alternative to vendors such as Xiaohongshu and Pinduoduo that provide an integrated social commerce experience on the same platform. Furthermore, we shed new light on the important and quickly growing C2C market in China. Finally, our empirical approach allows us to develop a more suitable measure for social media marketing than that found in most related studies. In most of the existing literature that has analyzed factors that may affect the adoption of social media marketing and its potential effects, social media marketing is often measured by whether a business has a social media account or not. However, given the fact that a social media account can be easily created, it is more relevant to know to what extent the social media account is used for marketing purposes. Through the adoption of machine learning and natural language processing techniques, we are able to develop a measure for social media activeness instead of merely social media adoption, thereby shedding new light on the driving forces behind social media marketing by C2C sellers.

To the best of our knowledge, this is the first study that examines the interplay between two leading C2C e-commerce and social media platforms as they experiment with a new collaborative effort aimed at helping C2C sellers marketing their products on social media. As mentioned throughout the text, the quickly growing social commerce space in China shows the importance of understanding how e-commerce interacts with social media. A successful integration of C2C e-commerce and social media platforms could theoretically provide a powerful alternative to integrated social commerce platforms such as Xiaohongshu and Pinduoduo. Our analysis suggests that the conflict between different reputational assets on the two platforms constitutes one important factor for the limited success of the collaboration that was initiated in 2013. Finally, the empirical relationships examined in the paper are at most correlational, rather than causal. It is impossible to deny the possibilities that unmeasurable internal factors such as Taobao-Weibo users' intrinsic willingness to engage customers are behind the empirical relationships found in this study and that external factors such as direct competition from WeChat, Xiaohongshu, Pinduoduo, and other platforms represent plausible reasons for the limited success of Taobao and Sina Weibo's social commerce collaboration. We therefore leave the analysis of causality to future studies once more data and research designs become available.

References

Abed, S.S., 2020. Social commerce adoption using TOE framework: An empirical investigation of Saudi Arabian SMEs. International Journal of Information Management, 53, p.102118.

Ali Abbasi, G., Abdul Rahim, N.F., Wu, H., Iranmanesh, M. and Keong, B.N.C., 2022. Determinants of SME's social media marketing adoption: competitive industry as a moderator. Sage Open, 12(1), p.21582440211067220.

Appel, G., Grewal, L., Hadi, R., and Stephen, A.T., 2020. The future of social media marketing. Journal of Academy of Marketing Science 48, 79-95.

Ashley, C. and Tuten, T., 2015. Creative strategies in social media marketing: An exploratory study of branded social content and consumer engagement. Psychology & marketing, 32(1), pp.15-27.

Cai, Y.J., Wang, Y., and Zhang, J., 2018. Enhancing e-platform business by customer service systems: a multi-methodological case study on Ali Wangwang instant message's impacts on TaoBao. Annals of Operations Research, pp.1-23.

Chang, Y.T., Yu, H., and Lu, H.P., 2015. Persuasive messages, popularity cohesion, and message diffusion in social media marketing. Journal of Business Research, 68(4), pp.777-782.

Chatterjee, S. and Kar, A.K., 2020. Why do small and medium enterprises use social media marketing and what is the impact: Empirical insights from India. International Journal of Information Management, 53, p.102103.

Chen, W., 2020. Studious shoppers can now find online classes on Alibaba's Taobao. KrAsia (https://kr-asia.com/studious-shoppers-can-now-find-online-classes-on-alibabas-taobao).

Cheung, C.M., Chiu, P.Y. and Lee, M.K., 2011. Online social networks: Why do students use facebook?. Computers in human behavior, 27(4), pp.1337-1343.

Chevalier, J.A. and Mayzlin, D., 2006. The effect of word of mouth on sales: Online book reviews. Journal of Marketing Research, 43(3), pp.345-354.

Chiu, C.M., Hsu, M.H., Lai, H. and Chang, C.M., 2012. Re-examining the influence of trust on online repeat purchase intention: The moderating role of habit and its antecedents. Decision Support Systems, 53(4), pp.835-845.

Dahnil, M.I., Marzuki, K.M., Langgat, J., and Fabeil, N.F., 2014. Factors influencing SMEs' adoption of social media marketing. Procedia-Social and Behavioral Sciences, 148, pp.119-126.

Dellarocas, C., 2003. The digitization of word of mouth: Promise and challenges of online feedback mechanisms. Management Science, 49(10), pp.1407-1424.

de Oliveira, M.J., Huertas, M.K.Z. and Lin, Z., 2016. Factors driving young users' engagement with Facebook: Evidence from Brazil. Computers in Human Behavior, 54, pp.54-61.

Du, R.Y., Hu, Y. and Damangir, S., 2015. Leveraging trends in online searches for product features in market response modeling. Journal of Marketing, 79(1), pp.29-43.

Dwivedi, Y.K., Ismagilova, E., Rana, N.P. and Raman, R., 2021. Social media adoption, usage and impact in business-to-business (B2B) context: A state-of-the-art literature review. Information Systems Frontiers, pp.1-23.

Feng, X. and Johansson, A.C., 2019. Top executives on social media and information in the capital market: Evidence from China. Journal of Corporate Finance, 58, pp. 824-857.

Gao, Q., Abel, F., Houben, G.J., and Yu, Y., 2012. A comparative study of users' microblogging behavior on Sina Weibo and Twitter. In: Masthoff, J., Mobasher, B., Desmarais, M., Nkambou, R. (Eds.), User Modeling, Adaptation, and Personalization: 20th International Conference, UMAP 2012, Montreal, Canada, July 16–20, 2012 Proceedings. Springer Berlin Heidelberg, Heidelberg, pp. 88–101.

Gao, B., Chan, W.K.V., Chi, L., and Deng, X.N., 2016. Size and growth dynamics of online stores: A case of China's Taobao. Com. Electronic Commerce Research and Applications, 17, pp.161-172.

Gesenhues, 2019. Women make up majority of influencer community, still earn less than male influencers. Marketing Land, June 6.

Godey, B., Manthiou, A., Pederzoli, D., Rokka, J., Aiello, G., Donvito, R., and Singh, R., 2016. Social media marketing efforts of luxury brands: Influence on brand equity and consumer behavior. Journal of Business Research, 69(12), pp.5833-5841.

Gong, S., Zhang, J., Zhao, P., and Jiang, X., 2017. Tweeting as a marketing tool: A field experiment in the TV industry. Journal of Marketing Research 54, pp. 833-850.

Guesalaga, R., 2016. The use of social media in sales: Individual and organizational antecedents, and the role of customer engagement in social media. Industrial Marketing Management, 54, pp.71-79.

Heavey, C., Simsek, Z., Kyprianou, C. and Risius, M., 2020. How do strategic leaders engage with social media? A theoretical framework for research and practice. Strategic Management Journal, 41(8), pp.1490-1527.

Herjanto, H. and Amin, M., 2020. Repurchase intention: the effect of similarity and client knowledge. International Journal of Bank Marketing, 38(6), pp.1351-1371.

Iankova, S., Davies, I., Archer-Brown, C., Marder, B. and Yau, A., 2019. A comparison of social media marketing between B2B, B2C and mixed business models. Industrial Marketing Management, 81, pp.169-179.

Jia, L., and Han, X., 2020. Tracing Weibo (2009-2019): The commercial dissolution of public communication and changing politics. Internet Histories 4, 304-332.

Jones, K. and Leonard, L.N., 2008. Trust in consumer-to-consumer electronic commerce. Information & management, 45(2), pp.88-95.

Keinänen, H. and Kuivalainen, O., 2015. Antecedents of social media B2B use in industrial marketing context: customers' view. Journal of Business & Industrial Marketing.

Kim, A.J., and Ko, E., 2012. Do social media marketing activities enhance customer equity? An empirical study of luxury fashion brand. Journal of Business Research, 65(10), pp.1480-1486.

Kim, Y. and Krishnan, R., 2015. On product-level uncertainty and online purchase behavior: An empirical analysis. Management Science, 61(10), pp.2449-2467.

Kumar, A., Bezawada, R., Rishika, R., Janakiraman, R. and Kannan, P.K., 2016. From social to sale: The effects of firm-generated content in social media on customer behavior. Journal of Marketing, 80(1), pp.7-25.

Lamberton, C. and Stephen, A.T., 2016. A thematic exploration of digital, social media, and mobile marketing: Research evolution from 2000 to 2015 and an agenda for future inquiry. Journal of Marketing, 80(6), pp.146-172.

Leonard, L.N. and Jones, K., 2021. Trust in C2C electronic commerce: Ten years later. Journal of Computer Information Systems, 61(3), pp.240-246.

Lessmann, S., Baesens, B., Mues, C., and Pietsch, S., 2008. Benchmarking classification models for software defect prediction: A proposed framework and novel findings. IEEE Transactions on Software Engineering, 34(4), pp.485-496.

Li, D., Li, J., and Lin, Z., 2008. Online consumer-to-consumer market in China–a comparative study of Taobao and eBay. Electronic Commerce Research and Applications, 7(1), pp.55-67.

Li, G., Liu, H. and Li, G., 2014. Payment willingness for VIP subscription in social networking sites. Journal of Business Research, 67(10), pp.2179-2184.

Lim, S., 2020. Emerging platforms: What is Xiaohongshu and how can brands leverage it? The Drum, 16 March.

Liu, D., 2019. China's $670 billion 'sheconomy' is growing like crazy. NBC News online, 3 May.

Liu, S., Perry, P., and Gadzinski, G., 2019. The implications of digital marketing on WeChat for luxury fashion brands in China. Journal of Brand Management 26, 395-409.

Liu, X., Shin, H. and Burns, A.C., 2019. Examining the impact of luxury brand's social media marketing on customer engagement: Using big data analytics and natural language processing. Journal of Business Research.

López-López, D. and Giusti, G., 2020. Comparing digital strategies and social media usage in B2B and B2C industries in Spain. Journal of Business-to-Business Marketing, 27(2), pp.175-186.

Manning, C.D., Raghavan, P. and Schütze, H., 2008. Introduction to Information Retrieval. Cambridge: Cambridge University Press.

Melnik, M.I. and Alm, J., 2002. Does a seller's ecommerce reputation matter? Evidence from eBay auctions. Journal of Industrial Economics, 50(3), pp.337-349.

Michaelidou, N., Siamagka, N.T., and Christodoulides, G., 2011. Usage, barriers, and measurement of social media marketing: An exploratory investigation of small and medium B2B brands. Industrial Marketing Management, 40(7), pp.1153-1159.

Moore, J.N., Hopkins, C.D. and Raymond, M.A., 2013. Utilization of relationship-oriented social media in the selling process: A comparison of consumer (B2C) and industrial (B2B) salespeople. Journal of Internet Commerce, 12(1), pp.48-75.

Moore, J.N., Raymond, M.A. and Hopkins, C.D., 2015. Social selling: A comparison of social media usage across process stage, markets, and sales job functions. Journal of Marketing Theory and Practice, 23(1), pp.1-20.

Pascucci, F., Ancillai, C. and Cardinali, S., 2018. Exploring antecedents of social media usage in B2B: a systematic review. Management Research Review.

Pavlou, P.A. and Dimoka, A., 2006. The nature and role of feedback text comments in online marketplaces: Implications for trust building, price premiums, and seller differentiation. Information Systems Research, 17(4), pp.392-414.

Pavlou, P.A., Liang, H., and Xue, Y., 2007. Understanding and mitigating uncertainty in online exchange relationships: A principal-agent perspective. MIS Quarterly, pp.105-136.

Puriwat, W. and Tripopsakul, S., 2021. Explaining social media adoption for a business purpose: an application of the UTAUT model. Sustainability, 13(4), p.2082.

Qin, B., Strömberg, D., and Wu, Y., 2017. Why does China allow freer social media? Protests versus surveillance and propaganda. Journal of Economic Perspectives, 31, pp. 117-140.

Shi, S., Chen, Y. and Chow, W.S., 2016. Key values driving continued interaction on brand pages in social media: An examination across genders. Computers in Human Behavior, 62, pp.578-589.

Siamagka, N.T., Christodoulides, G., Michaelidou, N., and Valvi, A., 2015. Determinants of social media adoption by B2B organizations. Industrial Marketing Management, 51, pp.89-99.

Sinclaire, J.K. and Vogus, C.E., 2011. Adoption of social networking sites: an exploratory adaptive structuration perspective for global organizations. Information Technology and Management, 12, pp.293-314.

Steinhoff, L., Arli, D., Weaven, S., Kozlenkova, I.V., 2019. Online relationship marketing. Journal of Academy of Marketing Science 47, 369-393.

Sullivan, J., 2012. A tale of two microblogs in China. Media, Culture and Society, 34, 773-783.

Ter Huurne, M., Ronteltap, A., Corten, R. and Buskens, V., 2017. Antecedents of trust in the sharing economy: A systematic review. Journal of Consumer Behaviour, 16(6), pp.485-498.

Thierer, A., Koopman, C., Hobson, A. and Kuiper, C., 2015. How the internet, the sharing economy, and reputational feedback mechanisms solve the lemons problem. U. Miami L. Rev., 70, p.830.

Toubia, O. and Stephen, A.T., 2013. Intrinsic vs. image-related utility in social media: Why do people contribute content to Twitter? Marketing Science, 32(3), pp.368-392.

Vock, M., Dolen, W.V. and Ruyter, K.D., 2013. Understanding willingness to pay for social network sites. Journal of Service Research, 16(3), pp.311-325.

Wang, X., Baesens, B., and Zhu, Z., 2019. On the optimal marketing aggressiveness level of C2C sellers in social media: Evidence from China. Omega, 85, pp.83-93.

Wang, Z., Li, H., Ye, Q. and Law, R., 2016. Saliency effects of online reviews embedded in the description on sales: Moderating role of reputation. Decision Support Systems, 87, pp.50-58.

We Are Social, 2019. Digital in 2019: China.

Yang, J., Hu, X. and Zhang, H., 2007. Effects of a reputation feedback system on an online consumer-to-consumer auction market. Decision Support Systems, 44(1), pp.93-105.

Ye, Q., Li, Y., Kiang, M., and Wu, W., 2009. The Impact of Seller Reputation on the Performance of online sales: evidence from TaoBao buy-it-now (BIN) data. ACM SIGMIS Database: the DATABASE for Advances in Information Systems 40, pp.12-19.

Yoon, H.S. and Occena, L.G., 2015. Influencing factors of trust in consumer-to-consumer electronic commerce with gender and age. International journal of information management, 35(3), pp.352-363.

Yu, H., and Cui, L., 2019. China's E-Commerce: Empowering Rural Women? The China Quarterly 238, pp.1-20.

Zolkepli, I.A. and Kamarulzaman, Y., 2015. Social media adoption: The role of media needs and innovation characteristics. Computers in human behavior, 43, pp.189-209.

Appendix

Table A1: Variable Description

VARIABLES	DESCRIPTION
Weibo Variables	
1 weibo marketing (d)	A dummy variable that equals 1 if at least one marketing Weibo during the 3-month observation period.
2 weibo use (d)	A dummy variable that equals 1 if at least one Weibo during the 3-month observation period.
3 number of follows (o)	The number of follows in Sina Weibo (3-month average).
4 number of followers (e)	The number of followers in Sina Weibo (3-month average).
5 tao (e)	A dummy variable that equals 1 if the user is verified as a Taobao seller in Sina Weibo.
6 verify (e)	A dummy variable being 1 if the user is identity-verified.
7 member (o)	A dummy variable being 1 if the user has paid for a Sina Weibo VIP membership.
8 female (c)	A dummy variable being 1 if the user claims as female in Sina Weibo.
Taobao Variables	
9 number of items (c)	The number of items listed in the Taobao shop (3-month average).
10 median price (c)	The median price of all the listed items (3-month average).
11 taobao rank (o)	Taobao rank (1-20), the higher the value the better rating points accumulated. It is effectively a measure of sales.
12 shop age (o)	The number of years between the Taobao shop opening date and the data collection date. The days are converted to decimal years (3-month average).
13 credit as buyer (e)	The credit level as a buyer, the higher the better ratings received from sellers (3-month average).
14 rating as seller (e)	The percentage of good ratings out of all ratings received from buyers (3-month average).
15 description industry comparison (e)	The relative difference between the Taobao shop's description rating (defined below) and the industry average (3-month average). Its range goes from -100 to 100.
16 description rating (e)	The average rating (1-5) received in terms of description matching product delivered, the higher the better (3-month average).
17 logistics industry comparison (e)	The relative difference between the Taobao shop's logistics rating (defined below) and the industry average (3-month average). Its range goes from -100 to 100.
18 logistics rating (e)	The average rating (1-5) received in terms of delivery speed, the higher the better (3-month average).
19 service industry comparison (e)	The relative difference between the Taobao shop's service rating (defined below) and the industry average (3-month average). Its range goes from -100 to 100.
20 service rating (e)	The average rating (1-5) received in terms of customer service, the higher the better (3-month average).
industry comparison average (e)	The average of the above description industry comparison, logistics industry comparison, and service industry comparison.
rating average (e)	The average of the above description rating, logistics rating, and service rating.

"d" stands for dependent variables, "c" stands for control variables, "e" stands for *earned* reputational assets, and "o" stands for *owned* reputational assets.

Table A2: Logistic Regression on randomly reclassified *weibo marketing*

VARIABLES	(1) Weibo	(2) Taobao	(3) Weibo+Taobao	(4) Weibo+Taobao	(5) Weibo+Taobao
number of items (c)		7.82e-05	0.000197**	0.000185**	7.86e-05
		(7.35e-05)	(8.40e-05)	(8.66e-05)	(8.09e-05)
median price (c)		9.49e-05	4.59e-06	5.46e-06	6.14e-05
		(6.43e-05)	(5.52e-05)	(5.68e-05)	(6.77e-05)
taobao rank (o)		0.0107	-0.0439***	-0.0481***	-0.0409**
		(0.0155)	(0.0133)	(0.0134)	(0.0167)
shop age (o)		-0.0346**	-0.00320	-0.00276	-0.0226
		(0.0170)	(0.0146)	(0.0147)	(0.0177)
credit as buyer (e)		0.000274***	0.000224***	0.000220***	0.000205***
		(4.15e-05)	(3.69e-05)	(3.75e-05)	(4.35e-05)
rating as seller (e)		0.202***	0.0925***	0.110***	0.151***
		(0.0465)	(0.0320)	(0.0331)	(0.0465)
industry comparison average (e)		0.000653	0.000895	0.000366	0.000421
		(0.00132)	(0.00109)	(0.00111)	(0.00137)
rating average (e)		0.124***	0.104***	0.117***	0.107***
		(0.0285)	(0.0229)	(0.0235)	(0.0295)
number of follows (o)	0.000304***		0.000325***	0.000344***	0.000309***
	(5.16e-05)		(5.23e-05)	(5.26e-05)	(6.55e-05)
number of followers (e)	7.70e-06***		7.87e-06***	7.26e-06***	7.48e-06**
	(2.34e-06)		(2.39e-06)	(2.30e-06)	(3.00e-06)
tao (e)	0.415***		0.397***	0.384***	0.495***
	(0.0590)		(0.0626)	(0.0630)	(0.0764)
verify (e)	0.392***		0.390***	0.391***	0.447***
	(0.0873)		(0.0881)	(0.0884)	(0.107)
member (o)	0.308***		0.285***	0.277***	0.260***
	(0.0193)		(0.0198)	(0.0199)	(0.0238)
female (c)	0.587***		0.534***	0.487***	0.435***
	(0.0742)		(0.0757)	(0.0773)	(0.0948)
Constant	-2.479***	-22.41***	-11.79***	-13.66***	-17.77***
	(0.0794)	(4.627)	(3.168)	(3.286)	(4.617)
product category fixed effects (c)	No	Yes	No	Yes	Yes
region fixed effects (c)	No	Yes	No	No	Yes
Observations	8,637	5,531	8,637	8,637	5,531

Standard errors in parentheses; ***, **, and * denote significance at the 1%, 5% and 10% level, respectively.

"c" stands for control variables, "e" stands for earned reputational assets, and "o" stands for owned reputational assets.

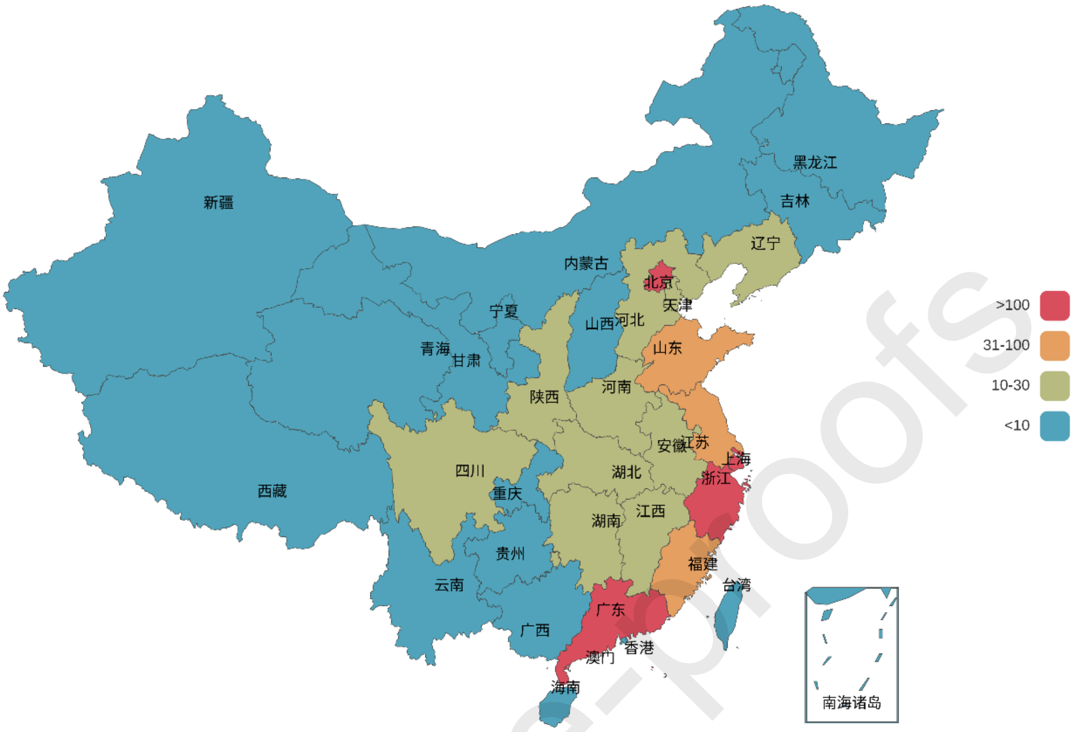


Figure A1: Geographic Distribution of Marketing-Active Taobao-Weibo Users

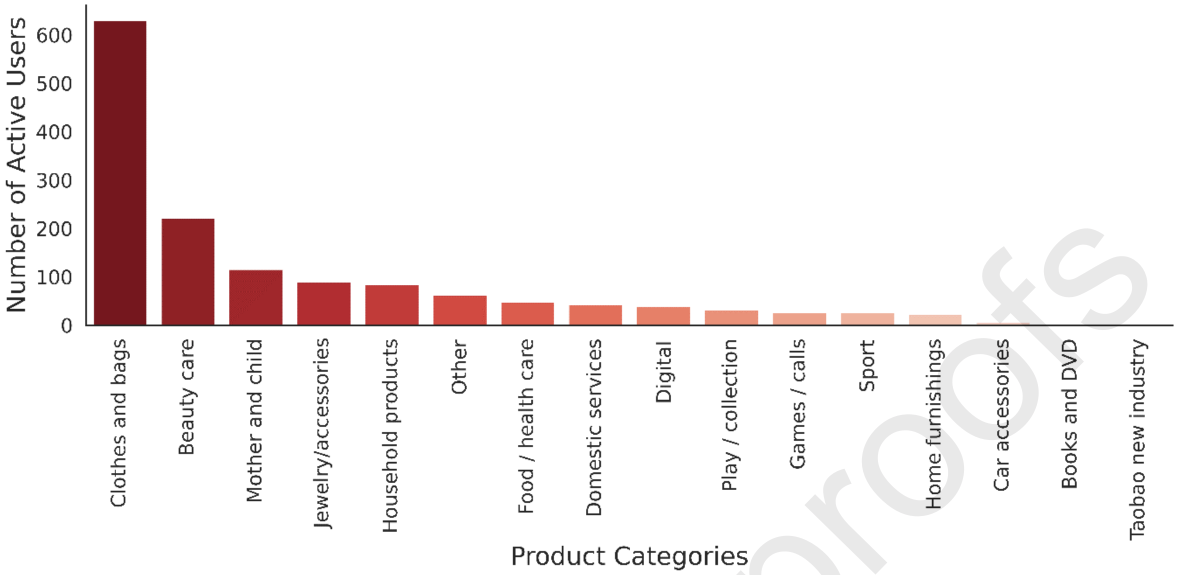


Figure A2: Product Distribution of Marketing-Active Taobao-Weibo Users

Anders C. Johansson: Conceptualization, Methodology, Writing, Reviewing and Editing.

Zhen Zhu: Conceptualization, Methodology, Software, Data Curation, Writing, Reviewing and Editing.

Submission of "Reputational Assets and Social Media Marketing Activeness: Empirical Insights from China"

Highlights

- Our study exploits the unique collaborative effort between Sina Weibo and Taobao in China.
- We measure social media marketing activeness with machine learning and natural language processing methods.
- We differentiate between earned and owned reputation factors accumulated on both platforms and test their relationships to social media marketing activeness.

- While the earned reputation factors are associated with social media marketing activeness, the owned reputation factors are observed with conflicts, which provides a potential explanation for the limited success of the cross-platform collaboration.

1