Are More Followers Always Better? The Non-Linear Relationship between the Number of Followers and User Engagement on Seeded Marketing Campaigns in Instagram

Suyoung Moon
Student Wisconsin Business School University of Wisconsin-Madison Wisconsin, United States of America

Shijin Yoo
Professor of Marketing Korea University Business School Seoul, Korea, shijinyoo@korea.ac.kr

Follow this and additional works at: https://amj.kma.re.kr/journal

Part of the Advertising and Promotion Management Commons, E-Commerce Commons, Marketing Commons, and the Other Business Commons

Recommended Citation
Moon, Suyoung and Yoo, Shijin (2022) "Are More Followers Always Better? The Non-Linear Relationship between the Number of Followers and User Engagement on Seeded Marketing Campaigns in Instagram," Asia Marketing Journal: Vol. 24 : Iss. 2 , Article 3.
Available at: https://doi.org/10.53728/2765-6500.1589

This Article is brought to you for free and open access by Asia Marketing Journal. It has been accepted for inclusion in Asia Marketing Journal by an authorized editor of Asia Marketing Journal.
Are More Followers Always Better? The Non-Linear Relationship between the Number of Followers and User Engagement on Seeded Marketing Campaigns in Instagram

Suyoung Moon a, Shijin Yoo b,*

a Wisconsin School of Business, University of Wisconsin-Madison, Wisconsin, United States
b Korea University Business School, Seoul, South Korea

Abstract

Seeded marketing campaign (SMC) is a newly created type of marketing activities with the widespread use of social media. Previous research has examined to find out the optimal seeding strategy that yields the best outcome from the campaign. This research explores the relationships between the characteristics of the seeded influencer and user engagement. The data consists of information from 1062 seeded Instagram posts posted in September 2020 in Korea and 778 seeded influencers who posted those contents. Analyzed by negative binomial regression, our quadratic model suggests that the relationship between user engagement and the number of followers of the seeded influencer draws an inverted U-shape, indicating influencers with greater number of followers may not always be the best choice for the marketers. Moreover, this research shows that the negative marginal impact coming from the huge number of followers can be attenuated when the influencer is an expert of the seeded product.

Keywords: Seeded marketing, Viral marketing, Social media, Instagram, User engagement

1. Introduction

The days we are now living are undoubtedly called digital era, where most people can readily access the Internet and use mobile devices. Due to this easy accessibility of the Internet, there has been a widespread use of social media recently; over 3.6 billion people are using social media in 2020, and it is expected to be increased to 4.41 billion in 2025 (Statista 2020a). Following this trend, social media is being used as an important marketing tool, especially as a branding channel (Ashley and Tuten 2015; eMarketer 2013). Marketers utilize social media for branding by managing business-to-consumer communication and providing engagement experiences (Ashley and Tuten 2015).

These activities are called social media marketing. Drury (2008), Cvijikj and Michahelles (2013) defined social media marketing as a usage of the existing social media platforms for increasing the brand awareness among consumers through utilizing word-of-mouth (WOM) principles, supporting two types of promotion: (1) traditional marketing promotion from firms to customers and (2) social promotion communicated among consumers (Mangold and Faulds 2009). The latter, which is a uniqueness of social media platforms, is often conducted by the publication of branded contents that can be generated not only by firms but also by ordinary users. These branded contents generated by users (i.e., user-generated contents or UGCs) are critical in marketing context since they are one of the main sources of consumer-to-consumer WOM. For firms’ point of view, UGCs on social media platforms allow firms to connect with new users, communicate with their followers, and thus widen their opportunities to increase sales (Park et al. 2020). This sheds
light on the new type of marketing activity that social media has brought up: seeded marketing campaigns (SMCs).

SMCs are conducted by firms deliberately seeding a focal product or marketing-relevant information with initially selected subset of consumers. These initial target consumers respond to this marketing activity by generating WOM about the seeded products or information. This type of WOM which is created from UGCs is called amplified WOM (Chae et al. 2017; Libai et al. 2010). This new type of WOM is different from the traditional organic WOM in that it is intentionally triggered by firms at the initial stage.

The evidence that SMCs are considered important in the marketing field can be found in that they are popular and are being conducted by various kinds of firms regardless of the size. According to the American Marketing Association and the Word of Mouth Marketing Association, one-third of marketers plan to run or already have run SMCs sending samples to initially selected consumers, and three-quarters are planning to use or already have used consumers to generate amplified WOM (Chae et al. 2017; WOMMA 2014). The boom of SMCs is not so surprising considering their cost-efficiency compared to traditional mass media marketing. For instance, Hotmail’s campaign which had generated 12 million subscribers in 18 months with $50,000 budget is considered as one of the first successful viral marketing activities. Viral advertisements conducted by Tipp-Ex had generated nearly 10 million clicks in four weeks (Hinz et al. 2011). Filip Tysander, the owner of the Swedish watch company Daniel Wellington, invested just $1500 for kick-starting the company and sold one million watches worldwide for a profit of $220 million in 2015 which is just six years after (Haenlein et al. 2020).

As the number of social media users increases, the number of social media platforms where users can actively participate increases too. In other words, marketers became to have a variety of options to choose which platform to seed their marketing contents. Despite the wide range of options, SMCs are especially popular on Instagram (Oliveira and Goussevskaia 2020). According to a global survey, 76 percent of responding social media marketers used Instagram to promote their business as of January 2020 (Social Media Examiner 2020a). Furthermore, almost 70 percent of respondents have answered that they will increase the use of Instagram for their marketing activities, which is the largest percentage compared to all the other social media platforms including Youtube, LinkedIn, Facebook, Twitter, Pinterest, Messenger bots, TikTok and Snapchat (Social Media Examiner 2020b). These statistics indicate that Instagram is the first and foremost social media marketing channel, especially in the domain of SMCs. This is due to the increased number of users on Instagram, indicating that large audience of marketing activity will be guaranteed. The number of global users of Instagram has been about 855 million in 2019, and it is expected to surpass 988 million in 2023 (eMarketer 2019). Moreover, Instagram is preferred by both marketers and consumers for its overall high engagement level coming from its visual nature, user stickiness, and better control over spam (Li and Xie 2020).

SMC is in line with social media influencer marketing in that influencers create WOM in exchange for paid contents. An influencer can essentially be anyone with an existing social media following (Esber and Wong 2020; Pittman and Abell 2021). As influencer marketing is a new field for celebrity endorsement, the source credibility model, which is mainly used to evaluate the proper celebrity endorser, should be also considered in the context of SMC (Djafarova and Trofimenko 2019). That is, marketers can select an optimal social media influencer of the campaign using the source credibility model.

There are three main dimensions for the source credibility: attractiveness, trustworthiness, and expertise. Among these dimensions, an expertise is the most critical factor in that it enhances the effects of endorsers (Verhellen, Dens and Pelsmacker 2013). A perceived trustworthiness increases with higher expertise (Erdem and Swait 2004; Wang and Scheinbaum 2018), and the consumers perceive words from expert influencers stronger than those from attractive influencers (Trivedi and Sama 2020). This indicates that the expertise may have bigger weight than other observable and perceivable characteristics of influencers when measuring the effectiveness of SMCs.

This research aims to explore a deeper relationship between initial seeded targets and consumer engagements on Instagram. In other words, this research focuses on how the observable and perceivable characteristics of a marketing campaign’s initial targets are related to the level of engagements as marketing outcome.

There are two main research questions under this research objective:

1. What is the relationship between the number of followers of the seeded target and the user engagement of the seeded content?
2. Does a perceived expertise of the seeded target moderate the above relationship?
2. Related research and hypotheses development

2.1. Seeded marketing campaigns and user engagement

Chae et al. (2017) have categorized extant research on SMCs into four main topics: brand characteristics, incentives, marketing outcomes, and target characteristics. Connecting some brand characteristics to social, emotional, and functional WOM drivers, social and functional drivers are more prominent in online WOM while emotional ones are important in offline WOM (Lovett, Peres and Schachar 2013). Positive externality-based mechanisms are efficient for encouraging or nudging regular consumers’ WOM transmissions such that they were more likely to prefer and select higher-connectivity friends as receivers (Stephen and Lehmann 2016). In terms of the sales outcome, firm-created WOM through SMCs interacts negatively with traditional advertising but positively with promotional activities (Dost et al. 2019). Lastly, seeding to well-connected people is the most successful approach for viral marketing (Hinz et al. 2011).

The matter of whom firms should initially seed their product to is of particular importance for firms in that it is possible for them to select initial subset of consumers that are most likely to maximize user responses to their campaigns (Oliveira and Goussevskaia 2020). The effects of SMCs depend on the characteristics of the seeded targets of the campaigns. Firms can readily exploit some observable network metrics such as the number of followers of the seeded target, thanks to its information availability and accessibility (Hinz et al. 2011). Considering the characteristics of initial target, firms carefully select social media influencers for SMCs. These influencers are commonly acting as endorsers of SMCs, like celebrity endorsers in traditional mass media advertisements.

To capture the total set of behavioral activities and perspectives from consumers toward any marketing action, consumer engagement is critically concerned by firms. User engagement is frequently used to measure the effects of online WOM in social media since the amplified WOM is created directly by the engagement of users. The main goal of social marketing for marketers is customer engagement followed by revenue generation (Oviedo-García et al. 2014). Therefore, the monetary value of a seeded post is normally estimated based on the user engagement (Oliveira and Goussevskaia 2020).

Li and Xie (2020) categorized the user engagement on social media into two types: (1) direct responses such as likes, comments and favorites, and (2) sharing or propagation of the original post. Different types of engagement metrics are used in respective social media. For example, the primary metrics for Facebook and LinkedIn are comments, likes, and shares, while those for Instagram and YouTube are only comments and likes; the primary metrics for engagement on Twitter are favoriting, likes, quoting and sharing (Coelho et al. 2016).

2.2. The number of followers

Influencers on social media are normally identified by their follower’s actions of recognizing, admiring, associating, and aspiring them (Djafarova and Trofimenko 2019; Kutthakaphan and Chokesamritpol 2013). That is, social media influencers are usually identified by the number of followers they have. In terms of the number of followers, which is the observable characteristic of social media influencers, people tend to perceive greater social influence as it increases (Jin and Phua 2014). However, it does not mean that the influencers with larger number of followers is always the better endorsers. Even though larger number of followers guarantees the larger reach of the content, the other criterion than the reach of the message must be considered to measure the success in persuasive communication (Veirman, Cauberghe and Hudders 2017).

Not as an indicator for the reach but for the source credibility, the number of followers is important. The number of followers on social networking sites is a measure of the predictor of social media user credibility (Weismueller et al. 2020; De Veirman et al. 2017). Source credibility model is commonly used to evaluate the appropriate celebrity endorser for its comprehensiveness and applicability to the online context (Djafarova and Trofimenko 2019). Moreover, the impact of electronic WOM(eWOM) depends upon the source credibility; since eWOM may appear less credible than verbal or face-to-face communication due to its anonymity, firms should establish and maintain eWOM credibility in order to encourage recipients to read and accept eWOM (Buttle 1998; Reichelt, Sievert and Jacob 2014). As mentioned in the introduction, attractiveness, trustworthiness, and expertise are three main dimensions in the source credibility model.

Consumers associate brands with their celebrity endorser to build brand credibility, adding the dimensions of attractiveness and trustworthiness to the brand (Elberse and Verleun 2012). These newly added dimensions may mediate between the
number of followers and the user engagement. People with great number of followers are considered popular and likable (Veirman, Caubergehe and Hudders 2017). Indegree, which is the number of people who follow a user, is frequently used as a proxy for popularity of the user (Cha et al. 2010; Djafarova and Trofimenko 2019). This sociometric popularity, which is the number of followers in social network, is positively related to the perceived social attractiveness (Jin and Phua 2014). Attractiveness also increases the user awareness perceived by the audience (Miciak and Shanklin 1994). This two-way causality between the number of followers and the perceived attractiveness will cause celebrity's attractiveness to increase with increasing returns to scale as the number of followers increases. Unfortunately, the convexity of increase in the attractiveness does not last permanently. Extant research conducted by Tong et al. (2008) has shown that an increase in the number of friends in social network was evaluated positively at first, but too many friends resulted in less favorable evaluations on the social attractiveness. Thus, when the number of followers exceeds the certain point, the perceived attractiveness will start to get decreasing.

Trustworthiness also grows as the number of followers increases to the certain point. Engaging with an audience, which is the element to evaluate trustworthiness (Djafarova and Trofimenko 2019), cannot be seen from the user with no follower. This engagement will rise as the number of followers, who are the main audience, increases. Unfortunately, the interaction between social media influencer and other users is not so active when the influencer has a lot of number of followers, since it is not easy to respond to every audience when there are too many. Furthermore, while considering people with great number of followers as famous, users do not always trust how those people achieved that followership (Djafarova and Trofimenko 2019). Similarly, when the number of followers reaches too high, the followership might be perceived as only superficial, with its quality called into doubt (Utz 2010). Therefore, perceived trustworthiness also declines when the number of followers exceeds the certain point.

We conducted an experiment with 3 (number of followers: low vs. middle vs. high) x 2 (expert: yes vs. no) between-subjects design to check how perceived attractiveness and trustworthiness change with an increase in the number of followers. A total of 216 female respondents from age 18 to 34 were collected on CloudResearch. The study showed participants an image of Instagram user profile with the following description (Fig. 1):

Please, take a moment to look at the Instagram profile of the poster. For large numbers, Instagram uses k as an abbreviation for thousand and m as an abbreviation for million.

Imagine that while you are exploring random Instagram feeds, you come to reach the post that is endorsing lipsticks of a particular cosmetics brand. This post, which is posted by an individual Instagram user, states that its content is sponsored by the cosmetics brand.

Then we asked them how attractive and trustworthy they perceive the user on a 7-point Likert-scale. One-way ANOVA was conducted with perceived attractiveness and trustworthiness as respective dependent variables and the number of followers as an independent variable. The result showed that perceived attractiveness and trustworthiness not always grow as the number of followers increases; they decline when the user's number of followers gets too high. See Appendix A for specific procedure and results.

To sum up, the larger number of followers leads to the larger reach, but it may also give the negative impressions to the other users. Based on these interactions between the number of followers and the perceived attractiveness and trustworthiness, the first hypothesis comes as follows:

H1. There exists an inverted U-shape relationship between the number of followers of the seeded influencer and the user engagement.

2.3. The expertise of social media influencers

Nevertheless, there is still a way for the influencers with many followers to supplement their negative marginal impact. Influencers who do not possess any specialty have user engagement with diminishing returns to scale while some who are more specialized in their knowledge base have not (Esber and Wong 2020; Pittman and Abell 2021). An influencer's expertise, which is the other important dimension for social media influencer's source credibility, may mitigate the negative marginal impact on trustworthiness caused by too many followers. Among the dimensions of source credibility, expertise is considered the most critical factor that enhances the match-up effects between the endorser and the endorsed product or brand (Verhellen, Dens and Pelsmacker 2013). According to previous studies, the higher the endorser's expertise, the stronger the endorser's trustworthiness perceived by consumers (Erdem and Swait 2004; Wang and Scheinbaum 2018). Therefore, perceived trustworthiness of expert endorsers will less
decrease than that of non-expert endorsers due to the high number of followers.

It is also noteworthy that consumers give more weight to the words of the expert influencers than those from the attractive ones (Trivedi and Sama 2020). Therefore, despite the negative or decreasing returns of perceived attractiveness with respect to the number of followers, expert influencers will have stronger power compared to non-expert influencers. Given the positive effects of expertise on trustworthiness and the stronger weight given by consumers, the second hypothesis comes as follows:

H2. The perceived expertise of social media influencers moderates the relationship between the number of followers and user engagement. That is, when the seeded influencer is perceived as an expert of the seeded product, the inverted U-shaped relationship becomes flatter.

3. Data and measurements

3.1. Seeded marketing campaigns

We collected Instagram posts that contain some keywords that are frequently used in Korea when informing the post as a seeded marketing campaign from September 1st, 2020 to September 30th, 2020. We selected lipstick as a focal seeded product since cosmetics industry heavily rely on influencer marketing (Li and Xie 2020) and consumers are more likely to search for reviews for cosmetics products due to high levels of product diversity and the fact that these products are experience goods (Chae et al. 2017). We obtained this data from Sometrend Biz, a Korean website that collects and analyzes massive data from documents of online platforms including various social media, providing social insights such as the trend of mentioned keywords of interest, related keywords, and analysis of emotional texts. Sometrend Biz shows the posts that meet the conditions set by users, as well as the direct links to the posts. First, we added some search keywords that are frequently used in Korea when informing the post is seeded, to be included in the posts, such as “ad” or “sponsored”. Also, the search keyword “lipstick” was added to be included in the post, which mentions the product category of the seeded product. Then we excluded posts that are just copied and reposted from the original post by excluding posts that contain the search keywords such as “regram” or “repost”. After setting the posted period from September 1st, 2020 to September 30th, 2020, total of 2332 posts were collected. Fig. 3 shows the example of the seeded post we have collected.

By following the links of these collected posts, we crawled the data of the seeded influencer’s profile including the account ID, the number of followers, 

---

*Since September 2020, Korea Fair Trade Commission has legislated a new law that seeded influencer of the brand campaigns must directly mention in the post that the content is seeded, obviously enough for the audience to recognize. As sponsorship presence might affect viewers’ responses to the seeded posts compared to consumer-voluntary posts (Park, Yi and Kang 2019), we only collected posts that mandatorily revealed their sponsorship.*
Imagine that while you are exploring random Instagram feeds, you come to reach the post that is endorsing lipsticks of a particular cosmetics brand. This post, which is posted by an individual Instagram user, states that its content is sponsored by the cosmetics brand.

Fig. 2. Post images.

the number of followings and the introduction bio texts, the content texts and the number of likes and comments of the post. During this process, some data were unable to be collected if the user has deleted one’s account or the post at the moment of crawling. Also, there was only one post of which poster’s number of followers was over 1,000,000, so it was excluded in the final dataset concerning the outlier issue. This collection process resulted in 778 users and their 1062 posts.

The brand-related keywords provided by Some-trend Biz were manually filtered to check the list of specific brands that have participated in SMCs. For example, those that are not the brands selling lipsticks, such as “Dazed” (magazine), “Olive Young” (drugstore), “Insta” (Instagram), were filtered as brand-unrelated keywords. Distinct words that represent the same brand were also checked manually. For example, “Clio” and “Club Clio” representing the same brand were counted as one brand. Next, we checked which brand-related keyword is mentioned in each post, resulting in the final dataset that consists of 65 brands.

3.2. User engagement

The number of likes and the number of comments, which are the primary metrics for engagement on Instagram (Coelho et al. 2016), are used to measure the level of user engagement. Li and Xie

Fig. 3. Example of seeded post.
(2020) measured the number of likes as an engagement on Instagram since sharing is not readily available on Instagram, and liking is a commonly adopted metric that helps audience to show enjoyment, appreciation, or endorsement of the content. Oliveira and Goussevskaia (2020) mentioned a function of the number of likes and comments of the post and the number of followers of the poster as an engagement. Coelho et al. (2016) also built two models with respective dependent variables, which are the number of likes and comments, to evaluate social network relationships on Facebook and Instagram. We collected the number of likes and comments on November 19th, 2020, which is about two months after all the posts are uploaded. The minimum number of likes is 2 and the maximum 11,517; the minimum number of comments is 0 and the maximum 32.

3.3. The number of followers

We also collected the number of followers of the influencers on November 19th, 2020. The minimum number of followers is 7 and the maximum 207,000. The number of followers follows a right-skewed distribution with the median number 10,000 and the average 24,520.11.

3.4. Perceived expertise

The level of interests in a topic is frequently used as the proxy for expertise measurement since there are high correlations between them. Guy et al. (2013) suggested that people are usually interested in topics where they have expertise and vice versa, providing a positive correlation of 0.7 between the expertise and the interest. Adamopoulos, Ghose, and Todri (2018) also measured expertise by capturing the intensity of the specific topics of interest in each user’s discussions on the specific platform. Therefore, in this research, experts will be considered ones who show their intense interest of cosmetics by mentioning related words. We measure the expertise of seeded influencers by counting the number of cosmetics-related keywords.

First, texts from Instagram bios of all 778 seeded influencers were collected. Then we parsed the texts, made word corpus of each bio and manually extracted cosmetics-related words from the corpus and account IDs of respective seeded influencers. Lastly, the number of extracted words was counted.

The maximum number of cosmetics-related words was 12, while the minimum was 0. The proportion of the seeded influencers with no cosmetics-related word contained in their IDs and bios was nearly half (51.8%), with median value of 0. Therefore, we defined the seeded influencer as an expert (Expert = 1, N = 375) if the influencer contains at least one cosmetics-related word in one’s ID or bio, otherwise non-expert (Expert = 0, N = 403).

The experiment we have conducted on CloudResearch also tested whether the bio including cosmetics-related words is perceived as more expert than that including none of the words. We manipulated whether the influencer is perceived as an expert (yes vs. no) by giving influencer’s Instagram bio including cosmetics-related keywords: “Beauty | Lifestyle | Fashion, Beauty Content Creator, Make-Up Artist, (smiling emoji) Beauty doesn’t have a weight limit” and the bio including none of them: “Enjoying my daily life (smiling emoji)”. We asked how they perceive the seeded influencer as an expert with the following question: “I think this Instagram user would be expert/experienced/knowledgeable in the cosmetics field.” (1 = Strongly Disagree, 7 = Strongly Agree). The manipulation succeeded as we expected. See Appendix A for specific procedure and results.

In sum, variable operationalization and selected descriptive statistics can be found in Table 1.

4. Analysis

Since dependent variables (i.e., the number of likes or the number of comments) in this research are count data, Poisson regression may be applied to analyze the current dataset. However, Poisson regression with count data assumes an equidispersion, which means that the mean and the variance is the same. We conducted an over-dispersion test for dependent variables by calculating the dispersion parameter $\theta$ using the `dispersiontest` function in R. If $\theta$ is bigger than 0, the data is overdispersed. The results of dispersion tests for Poisson models showed that the data are overdispersed for number of likes ($\theta = 0.81, p < 0.01$) and for number of comments ($\theta = 0.48, p < 0.01$), rejecting the null hypothesis of equidispersion.

Therefore, we use negative binomial regression for the analysis of our model to overcome the problem of overdispersed count data with additional dispersion parameter $\theta$. Here, the variance is larger than the mean as follows:

\[ \text{Var}(Y) = \mu + \mu^2 \theta \]

[2] Although automated method could be preferred for categorizing and extracting cosmetics-related words, this manual process overcomes the limitations and complexities of automated text analysis, such as entities with multiple meanings, slangs and abbreviations in social media (Berger et al. 2020).
Table 1. Main variables and descriptive statistics.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Description</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Engagement$_ik$</td>
<td>the number of likes of the post uploaded by influencer $i$ and seeded by brand $k$</td>
<td>350.10</td>
<td>740.07</td>
</tr>
<tr>
<td>Follower$_i$</td>
<td>the number of comments of the post uploaded by influencer $i$ and seeded by brand $k$</td>
<td>7.94</td>
<td>7.12</td>
</tr>
<tr>
<td>Expert$_i$</td>
<td>the number of followers of influencer $i$ (in thousands)</td>
<td>24.58</td>
<td>34.61</td>
</tr>
<tr>
<td></td>
<td>whether the cosmetics-related word is included in influencer $i$’s ID or bio</td>
<td>0.53</td>
<td>0.50</td>
</tr>
</tbody>
</table>

Note. SD stands for Standard Deviation.

$Var(y_{ik}) = E(y_{ik}) + \theta E(y_{ik})^2,$

where $y_{ik}$ is the dependent variable (Cameron and Trivedi 1990, 2013; Cvijikj and Michahelles 2013). Since $y_{ik}$ follows the negative binomial distribution with mean parameter $\mu_{ik}$ and dispersion parameter $\theta$, it has a conditional probability mass function such that:

$$f(y_{ik} | \mu_{ik}, r) = \left( \frac{y_{ik} + r - 1}{y_{ik}} \right)^r \left( \frac{\mu_{ik}}{\mu_{ik} + r} \right)^{y_{ik}} \frac{\Gamma(y_{ik} + r)}{\Gamma(r) \Gamma(y_{ik})},$$

where $r$ is the reciprocal of $\theta$.

Thus, the mean parameter $\mu_{ik}$ can be expressed as follows:

$$\log(\mu_{ik}) = \beta_0 + \beta_1 \text{Follower}_i + \beta_2 \text{Follower}^2_i + \beta_3 \text{Expert}_i + \beta_4 \text{Follower}^*_i \text{Expert}_i + \beta_5 \text{Follower}^2_i \text{Expert}_i + \beta_6 \text{Brand}_k,$$

where $\text{Brand}_k$ is the factor variable to control any fixed effect of seeding brand $k$ that are participating in the SMCs. The coefficient $\beta_2$ will test if there is an inverted U-shape relationship between the number of followers of the seeded influencer and the user engagement, and the coefficient $\beta_3$ will test the moderation effect of the influencer’s expertise.

5. Results

Table 2 and Table 3 present the results of the different specifications of our models for user engagement. In particular, Model 1 measures only the brand effects on user engagement. Model 2 introduces the linear effect of number of followers, while Model 3 adds the quadratic effect of it. Lastly, Model 4 includes the effect of an influencer’s expertise and its interaction effects. Considering the Akaike information criterion (AIC) and the information of the log-likelihood of each model, Model 4 fits the data the best, followed by Model 3, 2, and 1.

Regarding the number of likes, the results show that the coefficient estimate for the linear form of the variable $\text{Follower}_i$ is significantly positive ($\beta_1 = 0.0378$, $p < 0.01$). This is not so surprising in that it is consistent not only with common intuition but also with extant literature suggesting that seeding to the influencers with high number of followers is the best strategy for SMCs (Hinz et al. 2011). However, what we should pay attention to is the parameter for the quadratic term of $\text{Follower}_i$; it is negative and significant ($\beta_2 = -0.0002$, $p < 0.01$). The negative sign indicates that the relationship between the number of followers and the user engagement exists in the inverted U-shape. Therefore, $H_2$ is accepted for the number of likes.

Similar with the results for the number of likes, the coefficient estimate for the linear form of the variable $\text{Follower}_i$ is significantly positive regarding the number of comments ($\beta_1 = 0.0103$, $p < 0.01$). However, the parameter for the quadratic term of $\text{Follower}_i$ is also negative but with comparatively low significance for the number of comments ($p < 0.15$).

It turns out that the relationships between the number of followers and both the number of likes and comments draw an inverted U-shape, but only significantly regarding the number of likes. This may be because liking the posts requires low

Table 2. Estimation results for number of likes.

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>5.7635*** (0.32)</td>
<td>5.2143*** (0.27)</td>
<td>5.0050*** (0.26)</td>
<td>4.9140*** (0.27)</td>
</tr>
<tr>
<td>Follower$_i$</td>
<td>0.0192*** (0.00)</td>
<td>0.0356*** (0.00)</td>
<td>0.0378*** (0.00)</td>
<td></td>
</tr>
<tr>
<td>Follower$_i^2$</td>
<td></td>
<td>−0.0001*** (0.00)</td>
<td></td>
<td>−0.0002*** (0.00)</td>
</tr>
<tr>
<td>Expert$_i$</td>
<td></td>
<td></td>
<td></td>
<td>0.1176 (0.08)</td>
</tr>
<tr>
<td>Follower$_i$ X Expert$_i$</td>
<td></td>
<td></td>
<td></td>
<td>−0.0058 (0.00)</td>
</tr>
<tr>
<td>Follower$_i^2$ X Expert$_i$</td>
<td></td>
<td></td>
<td></td>
<td>0.0001* (0.00)</td>
</tr>
<tr>
<td>AIC</td>
<td>14,133.00</td>
<td>13,660.00</td>
<td>13,601.00</td>
<td>13,601.00</td>
</tr>
<tr>
<td>2 X log-likelihood</td>
<td>−14,001.24</td>
<td>−13,525.88</td>
<td>−13,465.36</td>
<td>−13,459.19</td>
</tr>
</tbody>
</table>

*p < 0.1; **p < 0.05; ***p < 0.01.

Standard errors in parentheses.
involvement with reactive consumption, while writing comments requires high involvement with proactive contribution (Oviedo-García et al. 2014; Tsai and Men 2012, pp. 10–13). According to Agrawal, Gupta and Yousaf (2018), moreover, liking is more affected by superficial features than commenting which requires much time and effort; message or posts giving information and socio-emotional messages in the posts have a significantly larger impact on the growth of likes compared to that of comments. Considering an informative trait of seeded contents, the difference between results of the number of likes and comments does not seem to be non-sense at all.

Model 4, which has the best fit to our data among all other models, analyzes the effect of an expertise of the seeded influencer and its interaction effects with the number of followers. What we should pay attention to in this model is the interaction term of an expertise and the quadratic form of the number of followers in order to confirm whether an influencer’s expertise alleviates the negative slope of an inverted U-shape between the number of followers and the user engagement. For the number of likes, the coefficient of the interaction term of Follower² and Expert is positive ($\beta_3 = 0.0001$, $p < 0.1$). Therefore, $H_2$ is marginally accepted, indicating that if the influencer is an expert of the seeded product, an inverted U-shape relationship between the number of followers and the number of likes would be flatter than that of non-expert influencers. Analyzing the number of comments as a dependent variable, on the other hand, the interaction term of Follower² and Expert has a positive but low significant parameter value. However, Expert solely has a positive and significant coefficient estimate in terms of the number of followers ($\beta_3 = 0.3276$, $p < 0.01$) while it is not for the number of likes, which indicates that whether the seeded influencer is an expert or not will have positive association with the number of comments of the seeded post.

### 6. General discussion

The goal of this research is to find the relationship between the influencer’s characteristics and the effects of SMCs. This research explores how the number of followers, which is the main observable social metrics on Instagram, interacts with the user engagement. The dataset that we have collected consists of over 1000 distinct posts from about 800 influencers that are seeded in September 2020. We analyzed the user engagement of these posts in the relation to the number of followers that the influencer has, taking into account the different status of seeded influencers in terms of the expertise.

The results show the inverted U-shaped relationship between the number of followers of the seeded influencer and the number of likes. Also, this research finds out that the relationship between the number of followers of the seeded influencer and the number of comments is inverted U-shape, but with low significance. This may be due to the high involvement of writing comments with proactive contribution, in contrast to the low involvement of liking the posts with reactive consumption (Oviedo-García et al. 2014; Tsai and Men 2012, pp. 10–13).

Regarding the moderating effect of the influencer’s expertise, whether the seeded influencer is an expert of the seeded product significantly moderates the inverted U-shaped relationship between the number of followers and the number of likes but does not moderate that between the number of followers and the number of comments. However, the number of comments is in a positive relationship with the seeded influencer’s expertise.

#### 6.1. Theoretical implications

The intrinsic characteristics of social media makes it possible for both researchers and marketers to observe the online WOM instances that provide deeper insights into the association between the seeded influencer’s characteristics and the effectiveness of WOM.
(Adamopoulos, Ghose and Todri 2018; Godes and Mayzlin 2004; Trusov, Bucklin and Pauwels 2009).

Extant research emphasize the importance of seeding to the optimal influencers when conducting SMCs, and this research extends the topic to examining the deeper relationship between the influencer’s characteristics and the user engagement.

This research has theoretical contributions in that it proposes the new idea that the seeded influencer with more followers is not always the better option, which contradicts to many extant literatures that have asserted that well-connected users are the optimal seeding influencers (Hinz et al. 2011; Moehr 2014). The negative coefficient estimate for the quadratic term of the main independent variable shows that the greater number of followers does not always lead to higher engagement of the seeded contents.

Moreover, this research does not only use the numeric social metrics of the influencer’s characteristics but also takes the source credibility model into account, which has been emphasized in prior research of traditional celebrity endorsement. Extending this traditional theory to the newly created type of endorser in social media, this research finds out that the influencer’s status of expert mitigates the negative marginal impact of the number of followers on user engagement. We also check how some dimensions of source credibility model interact with the number of followers of the influencer by conducting a survey experiment, suggesting how traditional source credibility model should be combined with social media influencers, which is the new type of celebrity endorser these days.

Since Instagram policy prohibits collecting data from the platform directly, it is hard to collect massive, detailed data from Instagram even though it is the primary social media marketing platform. Due to this limitation, majority of prior research have studied social media marketing on limited platforms such as Twitter or Facebook, which seem quite old-fashioned in the field of SMC. This research collects and analyzes the data consisting of more than 1000 distinct posts from about 800 seeded influencers on Instagram. Detailed data that can be observed on Instagram including the influencer’s bio, social connectedness, and content texts are also collected and analyzed in this paper.

6.2. Managerial implications

While many marketers agree that promoting SMCs is the foremost marketing activity that they should develop more on, it is not so easy for them to decide which influencer might fit the best for their campaigns. There is no particular influencer to seed who absolutely fits the best with every brand and every SMC. Just paying a huge amount of their budget to the influencers with millions of followers is not always the best plan for the brand. What they should do is to thoroughly plan to create strategic partnerships with certain influencers on social media (Mynatt 2020). This research provides the guideline for marketers that suggests how to utilize observable influencer characteristics to decide the optimal initial influencer when conducting SMCs.

For example, let us consider that Brand A from our dataset plans their seeded marketing campaign to get high user engagement. Since likes and comments of the post, which are the main metrics for user engagement on Instagram, require different involvements to the users and therefore have different implications to the firms, the brand should set their strategies respectively with consideration of what they aim at. Suppose Brand A tries to get higher number of likes from the seeded content, when seeding to influencers with low number of followers under 27,484, it is better for Brand A to seed its product to expert influencers to get more likes. When targeting influencers with followers between 27,485 and 76,900, seeding to non-expert influencers rather than experts will produce greater number of likes. In other words, the number of likes yielded by non-experts exceeds the number of likes yielded by experts in some follower range. However, if Brand A has affordable budget to seed its product to influencers holding greater number of followers than 76,901, they should target experts again, getting the highest number of likes with the expert influencer with 160,706 followers, reaching 2669 likes. Fig. 4A illustrates this description regarding the number of likes.

If Brand A’s objective is to get high-involved engagement, which is comment, it is always better for the brand to target expert influencers than non-experts. For non-expert influencers, the number of comments will start to decrease from the influencer with 115,921 followers. However, if they choose to target experts, the number of followers that starts to reduce the number of comments is bigger, which is 167,596. This non-expert influencer with 167,596 followers will gain the greatest number of comments, which is 14, among all other expert and non-expert influencers. Fig. 4B illustrates this description regarding the number of comments.

Besides, not only marketers but also potentially to-be-seeded influencers, who are so-called influencers on social media, can strategically accept the suggestions of SMCs by firms, considering their
number of followers, their expertise and the type of the product to be seeded in order to gain more engagement.

6.3. Limitations and future research

This research contains some areas that can be furthered by future research. In this research, we analyzed only one product category, which is lipstick. It is true that SMCs on Instagram are especially conducted within fashion and cosmetics industries due to its visual nature (Statista 2020b).

Nevertheless, extending the analysis to the various product categories will be contributing to generalize the models and the results proposed in this paper. In addition, this research conducts a cross-sectional study. However, considering the dynamic feature of the social metrics on social media, the extension of the models with time dynamic variables might help researchers and marketers gain deeper insight of dynamic online WOM effects on social media. Also, the results suggest that the number of likes and comments have different weights in terms of the role of user engagement.
Extending the model with different weights of dependent variables will contribute to both academic and managerial field of marketing.

Moreover, our analysis does not account for some variables related to the text content. We did not add WOM valence as a control variable because reliable sentiment analysis algorithms are not available for the Korean language, and scoring it manually seems infeasible given the amount of data (Chae et al. 2017). Also, considering the purpose of seeded contents, we thought there would be no significant variation in valence across our data. In order to check the likelihood, we asked a coder who is a native Korean speaker working in the marketing field to manually read the posts and judge their valence as described in the research from Chae et al. (2017), providing 100 posts randomly sampled from our 1062 posts. The coder found most of the posts were “positive” (94%); posts judged as “mixed/neutral” were only 6% and none of the posts were judged as “negative”.

The amount of information in the post could also be taken into the model. However, Li and Xie (2020) have found out that text content including the length of the post does not affect engagement on Instagram. To check the applicability of this finding, we analyzed our model adding the length variable, which we define as the number of characters in the post, with randomly sampled 100 posts. The result showed that it does not have any significant effect on the number of likes and comments (p > 0.1). Finally, it would be an interesting prescriptive study to develop an algorithm to find an optimal influencer to seed with the optimal number of followers among the candidates.

Conflict of interest

The authors have no conflicts of interest to declare.

Appendix.

Appendix A. Number of followers, perceived attractiveness and trustworthiness

To check the interactions assumed in the hypotheses, we conducted an experiment with female MTurk panel workers on CloudResearch, which is an advanced participant-sourcing platform with selected workers. We employed female workers from age 18 to 34 because the experiment provides Instagram post about sponsored lipstick, which is same as our main data, and this age group is the dominant users of Instagram among other ages (Statista 2021). Also, it is in accordance with extant research to focus on female participants who use Instagram (Lee and Eastin 2020; Pittman and Abell 2021).

We investigated the effects of seeded influencer’s number of followers on perceived attractiveness and trustworthiness and compared the results between expert seeded influencers and non-expert influencers. We manipulated whether the influencer is perceived as an expert (yes vs. no) by giving influencer’s Instagram bio including cosmetics-related keywords and the bio including none of them.

A.1 Procedure

At the very first question, we asked the familiarity of Instagram on a 7-point Likert-scale and those who answered that they are familiar (greater point than 3) could proceed with the rest of the survey. After removing responds from the participants who have failed the attention check, responses from 216 female participants (Mage = 29) were collected in total. Study had a 3 (number of followers: low vs. middle vs. high) x 2 (expert: yes vs. no) between-factorial design. Participants were randomly assigned to one of the six cells.

Next, we showed participants an image of virtual seeded influencer’s Instagram profile. The profile includes ID, the number of posts, followers and followings, and bio. The number of followers for low condition was 559, for middle 51k (51,000) and for high 4.2m (4,200,000). These numbers were chosen based on the general guide to the types of influencers and the abbreviation k and m that Instagram uses for thousand and million (Influencer Marketing Hub 2021). Then the participants were asked to choose the correct number of followers for the attention check, considering the purpose of the study. As a manipulation check, those who succeeded attention check were asked to answer how they perceive the seeded influencer as an expert: “I think this Instagram user would be expert/experienced/knowledgeable in the cosmetics field.” (1 = Strongly Disagree, 7 = Strongly Agree).

Next, we showed them the image of seeded Instagram post endorsing a lipstick, stating that it is sponsored by the brand with hashtag #sponsored. The number of likes and comments were deliberately removed (Pittman and Abell 2021). Then the participants were asked to answer how attractive and trustworthy they perceive the influencer: “I think this Instagram user would be attractive/stylish/classy.”
(1 = Strongly Disagree, 7 = Strongly Agree); “I think this Instagram user would be trustworthy/reliable/honest.” (1 = Strongly Disagree, 7 = Strongly Agree). The survey items were referred by previous research (Djafarova and Trofimenko 2019; Ohanian 1990, 1991; Poyry et al. 2019; Veirman et al. 2017). Table A1 shows the survey constructs and items with their scale reliability.

Table A1. Survey constructs and items.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Item</th>
<th>M</th>
<th>SD</th>
<th>Cronbach’s α</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attractiveness</td>
<td>Attractive</td>
<td>4.53</td>
<td>1.41</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td>Stylish</td>
<td>5.15</td>
<td>1.4</td>
<td>4.3</td>
</tr>
<tr>
<td></td>
<td>Classy</td>
<td>3.69</td>
<td>1.45</td>
<td></td>
</tr>
<tr>
<td>Trustworthiness</td>
<td>Trustworthy</td>
<td>3.4</td>
<td>1.49</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Reliable</td>
<td>4.01</td>
<td>1.5</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td>Honest</td>
<td>4.1</td>
<td>1.45</td>
<td>4.4</td>
</tr>
<tr>
<td>Expertise</td>
<td>Expert in the cosmetics field</td>
<td>3.98</td>
<td>1.8</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td>Experienced</td>
<td>4.54</td>
<td>1.85</td>
<td>4.4</td>
</tr>
<tr>
<td></td>
<td>Knowledgeable</td>
<td>4.6</td>
<td>1.83</td>
<td></td>
</tr>
</tbody>
</table>

A.2 Results

The expert influencer’s profile was perceived as significantly more expert than the non-expert influencers (M_expert = 5.05 vs. M_non-expert = 3.74, p = 0.00). Therefore, the manipulation has been successfully checked.

We conducted one-way ANOVA with perceived attractiveness as a dependent variable and the number of followers as an independent variable for both expert and non-expert influencers. For non-expert influencers, there were significant difference between the follower groups: low vs. middle (p = 0.00); low vs. high (p < 0.02); middle vs. high (p < 0.12). For expert influencers, the differences of the follower groups were as follows: low vs. middle (p < 0.02); low vs. high (p < 0.21); middle vs. high (p < 0.62). Specifically, perceived attractiveness for experts increased from low-level followers to middle-level followers, but decreased from middle to high (MLow x Nonexpert = 3.81, MMiddle x Nonexpert = 5.21, MHigh x Nonexpert = 4.63). This pattern was shown the same for the non-experts but with slighter amount of decrease (MLow x Expert = 4.45, MMiddle x Expert = 5.28, MHigh x Expert = 4.99). Accordingly, it shows that the perceived attractiveness starts to get negative returns to scale from the certain point, not always increasing.

Again, one-way ANOVA was conducted with perceived trustworthiness as a dependent variable and the number of followers as an independent variable for both expert and non-expert influencers. For non-expert influencers, the difference between the low-level and the middle-level follower group were significant and the difference from low vs. high and middle vs. high-level follower groups were marginally significant: low vs. middle (p < 0.04); low vs. high (p < 0.14); middle vs. high (p < 0.87). For expert influencers, the differences of the follower groups had low significance: low vs. middle (p < 0.6); low vs. high (p < 0.88); middle vs. high (p < 0.89). Perceived trustworthiness for the experts increased from low-level followers to middle-level followers, but decreased from middle to high (MLow x Nonexpert = 3.36, MMiddle x Nonexpert = 4.14, MHigh x Nonexpert = 3.98). This pattern was shown the same for the non-experts (MLow x Expert = 4.14, MMiddle x Expert = 4.47, MHigh x Expert = 4.31). In conclusion, it suggests that the perceived trustworthiness not always increases, but declines at the certain point.
References


Utz, Sonja (2010), “Show me your friends and I will tell you what type of person you are: How one’s profile, number of friends, and type of friends influence impression formation on social network sites,” Journal of Computer-Mediated Communication, 15(2), 314–35.


