Information Systems Research

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

The Impact of User Personality Traits on Word of Mouth: Text-Mining Social Media Platforms

Panagiotis Adamopoulos, Anindya Ghose, Vilma Todri

INFORMATION SYSTEMS RESEARCH
Articles in Advance, pp. 1–29
ISSN 1047-7047 (print), ISSN 1526-5536 (online)
http://pubsonline.informs.org/journal/isre/

The Impact of User Personality Traits on Word of Mouth: Text-Mining Social Media Platforms

Panagiotis Adamopoulos,[a] **Anindya Ghose,**[b,c] **Vilma Todri**[a]

[a] Goizueta Business School, Emory University, Atlanta, Georgia 30322; [b] Stern School of Business, New York University, New York, New York 10012; [c] Korea University Business School, Korea University, Seoul 136-701, South Korea

Contact: padamop@emory.edu, http://orcid.org/0000-0002-4289-1767 (PA); aghose@stern.nyu.edu (AG); vtodri@emory.edu, http://orcid.org/0000-0002-8346-7364 (VT)

Received: August 10, 2015
Revised: July 9, 2016; May 31, 2017; October 30, 2017
Accepted: November 6, 2017
Published Online in Articles in Advance: May 14, 2018

https://doi.org/10.1287/isre.2017.0768

Abstract. Word of mouth (WOM) plays an increasingly important role in shaping consumers' behavior and preferences. In this paper, we examine whether latent personality traits of online users accentuate or attenuate the effectiveness of WOM in social media platforms. To answer this question, we leverage machine-learning methods in combination with econometric techniques utilizing a novel quasi-experiment. Our analysis yields two main results. First, there is a positive and statistically significant effect of the level of personality similarity between two social media users on the likelihood of a subsequent purchase from a recipient of a WOM message after exposure to the WOM message of the sender. In particular, exposure to WOM messages from similar users in terms of personality, rather than dissimilar users, increases the likelihood of a postpurchase by 47.58%. Second, there are statistically significant effects of specific pairwise combinations of personality characteristics of senders and recipients of WOM messages on the effectiveness of WOM. For instance, introverted users are responsive to WOM, in contrast to extroverted users. Besides this, agreeable, conscientious, and open social media users are more effective disseminators of WOM. In addition, WOM originating from users with low levels of emotional range affects similar users, whereas for high levels of emotional range, increased similarity usually has the opposite effect. The examined effects are also of significant economic importance, as, for instance, a WOM message from an extrovert user to an introvert peer increases the likelihood of a subsequent purchase by 71.28%. Our findings are robust to several alternative methods and specifications, such as controlling for latent user homophily and network structure roles based on deep-learning models. By extending the characteristics that have been theorized to affect the effectiveness of WOM from the observable to the latent space, tapping into users' latent personality characteristics, and illustrating how companies can leverage the abundance of unstructured data in social media, our paper provides actionable insights regarding the future potential of social media advertising and advanced microtargeting based on big data and deep learning.

History: Yong Tan, Senior Editor; Wenjing Duan, Associate Editor.
Funding: The authors acknowledge the financial support from the Marketing Science Institute.
Supplemental Material: The online appendix is available at https://doi.org/10.1287/isre.2017.0768.

Keywords: word of mouth • social media • personality • quasi-experiment • machine learning • deep learning

1. Introduction

Social media constitute one of the most transformative impacts of information technology on various aspects of our everyday life, including how consumers communicate and interact (KPMG 2013). Because of this transformation, word of mouth (WOM), the most trusted information and advertising source among U.S. Internet users (AYTM 2013, Forrester 2013, Nielsen 2013), nowadays plays an increasingly important role in shaping consumers' online behavior and preferences, as users' opinions, choices, and decisions are frequently shared in social media.

Acknowledging that online consumers' choices can be vastly influenced by electronic WOM, marketers leverage social media to achieve key marketing objectives by incentivizing the spread of positive WOM instances from social media users (Adamopoulos and Todri 2014, 2015b). For instance, referral systems, the nurturing of positive online WOM, the spurring of the creation of online forums and communities (Dellarocas 2006), and postpurchase social shares (Todri and Adamopoulos 2014) are some quintessential and effective means of utilizing social media. Relevant academic work has explored some factors that could make the WOM instances more effective (Ghose 2017). Nonetheless, social media provide the opportunity to gain deeper insights into the users' characteristics that accentuate or attenuate the effectiveness of WOM instances (Godes and Mayzlin 2004, Trusov et al. 2009) and extend the characteristics that have been theorized to affect the effectiveness of WOM from the observable (e.g., sex, age, relationship status) to the latent space of characteristics (e.g., personality). Thanks to the recent advances in deep learning and machine learning, both firms and researchers have the unique opportunity to leverage the abundance of unstructured data in social media (Adamopoulos and Tuzhilin 2015) to identify users' latent characteristics and traits that can impact the effectiveness of WOM.

To this end, in this paper, we draw from established and substantial theories in psychology and social sciences to examine whether personality traits of social media users attenuate or accentuate the effectiveness of WOM. Specifically, using recent advancements in big data and machine-learning techniques to extract information from unstructured textual content, we examine whether and how latent personality characteristics of a user affect purchases of actual products with significant monetary cost made by the social media peers of the user after exposure to WOM messages. To better identify instances of successful WOM and distinguish the effectiveness of WOM from correlated behaviors and homophily among social media peers, we use the variation in the visibility of WOM messages in a quasi-experimental setting where the exposure to such WOM instances is independent of the characteristics of the content or the source and recipient of WOM messages. In addition, we also conduct extensive robustness checks and employ interdisciplinary techniques, including latent variable models and deep-learning techniques, to further control for potential unobserved confounders.

Our analysis yields two main results. First, there is a positive and statistically significant effect of the level of personality similarity between two social media users on the likelihood of a subsequent purchase after exposure to WOM. In particular, exposure to WOM messages from similar users in terms of personality, rather than dissimilar users, increases the likelihood of a post-purchase by 47.58%. Second, moving beyond establishing personality similarity as an important determinant of WOM effectiveness, we also examine pairwise combinations of latent personality traits of senders and recipients of WOM messages, and we find that there are statistically significant effects of specific personality characteristics on WOM effectiveness. WOM originating from users who exhibit high levels of agreeableness, conscientiousness, and openness is more likely to be more effective, whereas for users with low levels of conscientiousness or agreeableness, the opposite effect is more likely. In addition, introverted users are susceptible to WOM, in contrast to extroverted users. Finally, WOM originating from users with low levels of emotional range affects similar users, whereas for high levels of emotional range, increased similarity usually has the opposite effect. These results are robust to a wide variety of alternative econometric specifications and additional controls. The corresponding models also exhibit very good predictive ability based on out-of-sample evaluation results. The examined effects are also of significant economic importance, as, for instance, a WOM message from an extroverted user to an introverted peer increases the likelihood of a subsequent purchase by 71.28%.

Overall, on the basis of a novel combination of machine-learning-based text-mining techniques with more conventional econometric methods and a quasi-experiment, this paper examines how latent personality traits and pairwise characteristics of users in social media platforms can facilitate the effect of WOM and subsequent economic outcomes. This is the first paper to study how personality similarity and specific combinations of personality traits affect users' online purchase behavior and facilitate WOM, paving the way for additional research in the unexplored area of latent user characteristics and their effect on individual behaviors in social media and networks. By examining these effects and illustrating how companies can leverage the abundance of unstructured data in social media and tap into users' personality characteristics, apart from extending the relevant theories, our paper provides actionable insights regarding the future potential of social media advertising and advanced microtargeting based on big data and natural language processing.

2. Theoretical Background and Related Work

In this section, we discuss how this paper is related to various streams of research that span the fields of information systems and marketing, among others, and elucidate how our study extends the existing literature. More specifically, drawing on theories that are deeply rooted in psychology and social sciences, we discuss how latent personality characteristics of users could help us predict and better understand the effects of WOM in social media. Finally, we discuss the feasibility of inferring latent personality characteristics by leveraging the abundance of unstructured data of online communications.

2.1. WOM and User Characteristics

The ability of firms to observe the electronic WOM instances at the granular level of individual interactions offers an attractive opportunity to learn how users' characteristics, such as personality traits, can facilitate or confine the effects of WOM. This is a particularly interesting question, as personality has been found to affect various aspects of individual behavior, including job performance (Barrick and Mount 1991), academic motivation (Komarraju and Karau 2005), and romantic relationships (Shaver and Brennan 1992, Tupes and Christal 1992), as well as attitudes toward computer and information systems (Devaraj et al. 2008, Sigurdsson 1991). Focusing on consumer preferences, prior work has demonstrated that personality characteristics can predict whether people would be more likely to accept a suggested product or service (Horton 1979, Hu and Pu 2011). Likewise, prior literature indicates that personality also affects the human decision-making process, such as the preferences for music (Rentfrow and Gosling 2003), consumers' brand preferences (Lin 2002), and effectiveness of recommendation agents (Adamopoulos and Todri 2015a). Hence, tapping into the idea that personality characteristics affect individual behaviors and drawing on theories that are deeply rooted in psychology and social sciences, we aim at extending the set of users' characteristics (Chang 2004, Forman et al. 2008, v. Wangenheim and Bayón 2004) that have been theorized as affecting the success of WOM instances from the observable to the latent space and, in particular, to include latent personality traits. Understanding the relationship between the established and substantial concepts of WOM in social media and user personality allows us to assimilate the sphere of influence of personality traits and predict the effective dissemination of word of mouth in social media platforms.

Building on established theoretical concepts (see Section 2.2), we focus on how the latent characteristics of the social media users can impact the effectiveness of WOM. In this respect, our paper is related to papers examining the effects of observable user characteristics on WOM. More specifically, Godes and Mayzlin (2009) study how loyalty of the disseminators of WOM toward a brand (e.g., frequency of purchases) facilitates WOM and demonstrate that firms should look for less loyal customers to spread WOM. Moreover, Bapna and Umyarov (2015) find that social media users with fewer friends are more susceptible to WOM influence from their peers. In the same vein, Forman et al. (2008) show that, in the context of WOM in an online community, identity-descriptive information of reviewers is used by consumers to supplement or replace product information when evaluating the helpfulness of online reviews. Similarly, Aral and Walker (2012) examine the impact of self-reported observable characteristics (e.g., gender, age) on identifying susceptible and influential members regarding the adoption of a Facebook app.

Our paper is substantially different from these papers, as we are the first to examine the impact of personality traits and latent personality similarity on WOM. Furthermore, we examine real monetary transactions and actual product purchases, which are more likely to be associated with deliberate and conscious user behaviors compared with commonly studied user actions, such as clicks and retweets. Besides this, we employ an extensive set of individual attributes, extracted from unstructured data (and not exclusively self-reported by the users) as well as several measures of pairwise relationship characteristics, in addition to the actual WOM message content. We also leverage a combination of text-mining and econometric techniques to utilize our data set and gain deeper insights into the factors that affect the effectiveness of WOM. This combination of methods and approaches allows us to extend the set of characteristics that have been theorized to affect the effectiveness of WOM from the observable space of self-reported characteristics (e.g., sex, age, relationship status) to the latent space of personality characteristics, paving the way for additional research in the unexplored area of latent user characteristics and personality traits and their effect on individual behaviors. Finally, we also identify a novel quasi-experimental setting that allows us to better disentangle the effects of personality on WOM from correlated user behaviors and homophily while facilitating future research in social media.

2.2. User Personality Characteristics Taxonomy

Personality has been defined as "the dynamic organization within the individual of those psychological systems that determine his characteristics behavior and thought" (Allport 1961, p. 28), and it has been a topic of intense academic interest across many fields. The study of personality has led to the emergence of personality psychology, which has been an identifiable discipline in social sciences for decades. A large number of researchers in this area have investigated personality constructs in an effort to uncover the underlying factors of personality, leading to taxonomies of personality traits and psychologists achieving congruent views on the structure and concepts of personalities (Barrick and Mount 1991). For instance, nowadays, it is widely accepted that there are five robust factors of personality that can serve as a meaningful taxonomy for classifying personality attributes (Digman 1990, Goldberg 1981, Norman 1963). The most influential taxonomy of personality attributes is admittedly the "Big Five" taxonomy, and it serves as a useful integrative framework for thinking about individual differences at a fairly high level of abstraction (Baumgartner 2002). The prevalence of the five-factor taxonomy framework is justified by the compelling evidence for robustness of the five-factor model across different theoretical frameworks, using different assessment approaches including questionnaires and lexical data, in different cultures, and using ratings obtained from different sources (Barrick and Mount 1991, Barrick et al. 2001, Costa and McCrae 1992, McCrae and Costa 1987).

The five-factor taxonomy proposes a comprehensive theoretical framework of five factors necessary and sufficient to represent human personality in terms of traits—it is a framework for distinguishing; ordering; and naming the behavioral, emotional, and experiential characteristics of individuals (John and Srivastava 1999). The latent personality dimensions can be generally defined as follows. The first dimension is titled *agreeableness* and captures a person's tendency to be compassionate and cooperative toward others. Agreeableness is associated with altruism, cooperation, trustfulness, empathy, and compliance. The second dimension is *conscientiousness* and describes a person's tendency to act in an organized or thoughtful way. Individuals characterized by high levels of conscientiousness tend to be driven, deliberate, organized, persistent, and self-assured. The third dimension of the Big Five–factor model is the *extraversion* dimension that refers to a person's tendency to seek stimulation in the company of others. Extraversion consists of outgoingness, sociability, assertiveness, and excitement-seeking behaviors. The fourth dimension is *emotional range*, which describes the extent to which a person's emotions are sensitive to the individual's environment. The tendency of an individual to be worried, depressed, self-conscious, and hedonistic is captured by the aforementioned dimension. Finally, the fifth dimension is that of *openness*, which refers to the extent to which a person is open to experiencing a variety of activities. Adventurousness, intellect, creativity, and liberalism define the openness to experience of individuals.

While leveraging personality characteristics constitutes a promising pathway toward understanding online consumers' behaviors and WOM effects, there exist significant challenges that have so far prevented the exploitation of personality characteristics as predictors or determinants of user behaviors on a large scale. Such challenges emerge from the inherent difficulty of identifying and measuring latent personality characteristics. Specifically, the traditional way of measuring personality characteristics requires the completion of long questionnaires, and hence, it has been particularly burdensome, if not impossible, to obtain such information on a large scale. However, there are currently a few studies that have successfully attempted to automatically derive and assess personality traits from text, based on the established relationship between word use and personality (Fast and Funder 2008, Hirsh and Peterson 2009, Yarkoni 2010). Exploring the feasibility of deriving personality traits from social media text, Mairesse and Walker (2006) demonstrate that computational models based on derived personality traits perform better than models using self-reported personality traits. In addition, tapping into the recent advances of data mining and predictive models, Chen et al. (2015) demonstrate the effectiveness of personality traits derived via a lexicon-based approach (Pennebaker et al. 2007, Yarkoni 2010) and found that predicted personality traits had the same effects as the personality traits measured by traditional personality questionnaires. Hence, drawing on the psychology of language and data-mining algorithms, we automatically infer personality characteristics leveraging the abundance of unstructured data of digital communications and demonstrate the feasibility of such analysis with low cost and on a large scale.

2.3. Interpersonal Personality Similarity

According to the theory of social comparison, people are characterized by the tendency to compare their attitudes and capabilities with those of others and are often inclined to alter their individual opinions and behaviors as a result of such comparisons with other people (Festinger 1954). This inherent tendency to compare oneself with another person increases significantly as this person is seen to be similar to oneself, primarily because of the implicit assumption of individuals that similar people have similar needs and preferences (Feldman and Spencer 1965, Festinger 1954). Increased similarity between two parties (e.g., in communication style, attitudes, cognitive processes, demographic characteristics, religious background, political orientation, physical appearance, socioeconomic status, other dimensions) also promotes social attraction, according to the *theory of interpersonal similarity* (Byrne and Griffitt 1969, Byrne et al. 1968, Singh and Ho 2000). Similarly, differences between the two parties can lead to dislike and avoidance, according to the cognitive consistency theories (Singh and Ho 2000, Tan and Singh 1995).

Examining the effects of interpersonal similarity in terms of personality, prior research has shown, for instance, that people are inclined to make decisions regarding personal relationships based on interpersonal similarity across the dimensions of agreeableness, conscientiousness, extraversion, emotional stability, and openness to experience (Botwin et al. 1997). Drawing on the aforementioned theories and going beyond relationship decisions, related research in psychology and communication has accumulated significant evidence suggesting that higher levels of interpersonal similarity between two parties (i.e., sender and recipient of a message) based on various other characteristics increase the ease of communication and enhance the predictability of various behaviors. For instance, synthesizing findings from a large number of past studies, Lichtenthal and Tellefsen (2001) suggest that similarity in internal characteristics based on composite measures (e.g., combining together personality,[1] education, attitudes, perceptions, political views, and values; Crosby et al. 1990) can increase a consumer's willingness to trust a salesperson in an offline retail setting and follow his or her guidance, whereas similarity in observable characteristics, such as physical attributes, has a much smaller effect on consumers' perceptions or a salesperson's effectiveness. In this paper, we focus on WOM messages among peers in social media going beyond in-person private seller-to-consumer communication, an offline retail setting, and a simple composite measure of similarity across different characteristics. In particular, we examine the impact of the latent interpersonal personality similarity between two parties on the effectiveness of electronic WOM and on the economic behavior of the recipients of WOM messages. Hence, the first research question we examine in this study is as follows:

- *Research Question* 1: Does personality similarity between the source and recipient of a WOM message affect the economic behavior of the recipient after exposure to the WOM message?

2.4. Information Processing and Personality Characteristics

In addition, prior research suggests at a theoretical level that message recipients use information about the source of the message as a heuristic device, drawing on their assessment of the information provider as a simple and convenient shortcut to help them reach judgments and guide actions, and hence, the attributes of an information source can have powerful effects on the way people react to messages as well (e.g., Chaiken 1980, 1987; Hass 1981; Kelman 1961; Mackie et al. 1990). The *information processing literature* has indeed identified this "messenger bias" at an empirical level, demonstrating in various research settings that attributes of a message source often exert direct effects on message recipients' attitudes and behaviors, independent of the message content that is broadcasted by the sender (e.g., Chaiken and Maheswaran 1994, Chang 2004, Cohen 2003, Forman et al. 2008, Kang and Herr 2006, Menon and Blount 2003, Pornpitakpan 2004, Simpson et al. 2000, v. Wangenheim and Bayón 2004). However, no prior academic work has thoroughly examined and scientifically documented the effect of specific personality traits and characteristics on the effectiveness of WOM in social media. In this paper, in addition to our first research question, we also study the impact of corresponding personality characteristics of the sender of a WOM message in combination with the personality characteristics of the recipient to provide richer findings regarding the effect of specific personality traits on WOM effectiveness. For instance, we move from studying interpersonal similarity on the extraversion dimension of personality toward studying whether an extrovert-to-extrovert WOM communication or an introvert-to-introvert WOM communication is more effective. Therefore, the second research question we examine in our study is this:

- *Research Question* 2: Do specific pairwise combinations of personality characteristics of the sender and recipient of a WOM message affect the economic behavior of the recipient after exposure to the WOM message?

3. Experimental Setting and Data Description

In the following section, we present the social commerce venture that was launched in the microblogging (social media) platform of Twitter. This social commerce venture constitutes our empirical setting for studying our research questions pertaining to WOM effects for real-world monetary transactions and how the personality characteristics of the users enable or constrain these WOM effects.

3.1. Empirical Context

The social commerce venture under study is an exemplary business model of leveraging users' connections in social media to stimulate WOM (Todri and Adamopoulos 2014). It is a service that enables customers to make a "frictionless" purchase within a platform while automatically spreading the word about the product and the service to their social media peers; the specific platform is the leading social commerce venture in terms of sales and engagement (AddShoppers 2013). Regarding the data-generating process of the specific social commerce transactions, the social commerce service provider (i.e., American Express) first broadcasts a short message in the platform (i.e., Twitter) announcing the list of participating merchants and the products that are available for sale. In particular, as illustrated in Figure A1 in the online appendix, this announcement provides details about the product offerings (e.g., product, respective sale price) and provides the designated hashtags (i.e., a word or phrase preceded by a hash sign (#)) consumers must employ to make a purchase. Consumers who are interested in making a purchase must have a microblogging account and synchronize their social commerce provider account with their microblogging account. Once the social commerce provider announces the available products from the participating vendors, users can purchase them by posting a short message (i.e., tweet) and including the designated token (i.e., hashtag). In addition to the required hashtag, consumers can choose to add additional content and personalize the purchasing tweet messages they share with their social media peers. Typically, such messages are publicly posted on the profile of the user, and the user's friends (i.e., peers who follow the stream/timeline of the specific user) will automatically receive the corresponding message on their own newsfeed (please see Section 4.2.2 for details). At the same time, the social commerce service provider tracks the tweets that use the designated hashtag on the social media platform and matches them to the desired product. After the purchase is confirmed, the social commerce service provider bills the customers and ships the product within one to five business days.

3.2. Empirical Data

Our database contains all the transactions that were generated through the aforementioned process on the social media platform. Each transaction is committed from a user account in the social media platform and is associated with a specific product offering. The data span all the confirmed transactions that took place from the second calendar week of February 2013 until the first calendar week of March 2013. Each transaction in our database consists of the original message of the user, the message ID, the exact date and time that the message was posted, the user account ID, the designated hashtag, and whether the message would be rendered visible to each of the followers of the user (see Section 4.2.2). Moreover, our database also contains users who were eligible to make a purchase but chose not to do so. Additionally, we have access to user-specific information, such as the user's screen name on the platform, the set of followers and the set of friends (or followees, as they are also called), all the posted messages of the user on the platform, and the self-reported description of the user's profile, etc. We further complement our data set with rich unstructured data from users' profiles and timelines by leveraging text-mining and machine-learning techniques to extract the user personality traits (see Section 4.2.1) as well as to enhance our identification strategy and control for latent similarity (beyond personality traits) among the users (see Section 4.2.3).[2] The main variables of interest and the corresponding machine-learning models are described in detail in the following section.

Additionally, we have information about all the product offerings. The social commerce service provider collaborated with well-known retailers and offered in total eight different products available for purchase (e.g., Figure A1 in the online appendix). The products, which were offered at a reduced retail price, were available for purchase only for a specific period of time. The featured products belong to a wide variety of categories and all of them are mainstream products. In particular, the products correspond to video game consoles and related accessories, electronics and sports equipment (e.g., high-definition tablet, sports and action cameras with related equipment), general-purpose gift cards, and fashion accessories (e.g., designer bracelet, luxury handbags). We should note that the particular set of offerings from the social commerce service provider was available for purchase at a reduced price (about a 25% discount, with an average retail price of US$125) only through the specific platform. Hence, our study does not suffer from sample selection bias issues that would arise, for instance, if the users could choose from which platform or network to make the purchase and we had analyzed only the transactions and the WOM instances that took place on the social media platform. Finally, we control for observed and unobserved heterogeneity at the product level by introducing product-level fixed effects in our model specifications. As robustness checks, we also test different specifications controlling for potentially unaccounted or unobserved correlations among the users, as discussed in detail in the following sections.

4. Empirical Methodology

To formally characterize our econometric model, we model user purchase decisions after being exposed to WOM messages in terms of both message and user characteristics, including latent personality traits and attributes extracted from unstructured textual content using machine-learning techniques. To better control for any unobservable and latent confounders, we utilize the variation in the visibility of WOM messages employing a quasi-experimental research design where the exposure to such WOM instances is independent of the characteristics of the content or the source and recipient of WOM messages; we also conduct extensive robustness checks and allow the visibility of messages to be endogenous (see Section 5.2) while controlling for several potential confounders. The rest of this section is organized as follows: we provide a brief sketch of our main econometric model specification (Section 4.1), the empirical identification of latent personality traits and characteristics (Section 4.2.1), the experimental research design (Section 4.2.2), and a discussion of our model features beyond personality traits (Section 4.2.3).

4.1. Econometric Model Specification

To estimate the moderating effects of personality traits on WOM effectiveness and consumers' behavior, we use a continuous-time single-failure survival model. In particular, we model *how quickly* users purchase a product, if any, using a proportional hazards model and correcting for censoring of transactions that might had been intended to occur after the observation window (Kalbfleisch and Prentice 2011). In detail, we specify the following survival model:

$$\begin{aligned}\lambda_i(t)=\lambda_0(t)\exp\big(&\boldsymbol{\beta}_s\ \textit{Recipient-Sender Similarity}_{ij}\\ &+\boldsymbol{\beta}_p\ \textit{Recipient-Sender Personality Similarity}_{ij}\\ &+\boldsymbol{\beta}_{pw}\ \textit{Recipient-Sender Personality Similarity}_{ij}\\ &\quad\times \textit{WOM Message Visibility}_{ij}\\ &+\boldsymbol{\beta}_w\ \textit{WOM Message}_j+\boldsymbol{\beta}_e\ \textit{User Expertise}_{ij}\\ &+\boldsymbol{\beta}_l\ \textit{User Leadership}_{ij}+\boldsymbol{\beta}_c\ \textit{Recipient-Sender}\\ &\qquad \textit{and WOM Message Controls}_{ij}\big),\end{aligned} \tag{1}$$

where $\lambda_i(t)$ is the hazard of peer i of consumer j making a social commerce purchase after having been exposed to a WOM message from j, $\lambda_0(t)$ represents the baseline hazard, *Recipient-Sender Similarity* captures the level of similarity between sender j and recipient i of a WOM message using various metrics as discussed in Section 4.2.3, *Recipient-Sender Personality Similarity* measures the personality similarity between a pair made up of recipient and sender of a WOM message as discussed in Section 4.2.1, *WOM Message Visibility* captures the visibility level of the corresponding WOM message as discussed in Section 4.2.2, *WOM Message* measures the intensity and type of the WOM message as discussed in Section 4.2.3, *User Expertise* and *User Leadership* control for latent user expertise and interests as well as social media leadership of senders and recipients as discussed in Section 4.2.3, and the additional *Recipient-Sender and WOM Message Controls* capture additional user activities and characteristics as discussed in Section 4.2.3, such as the number of messages the recipients and senders have posted in the social media platform and the number of interactions between the recipient and the sender of the WOM message.

In addition, as discussed in Section 2.4, we also study the impact of the personality characteristics of a WOM message sender in combination with the personality characteristics of the recipient to provide richer findings regarding the effect of specific personality traits on WOM effectiveness. Therefore, we also specify the following survival model:

$$\begin{aligned}\lambda_i(t) = \lambda_0(t)\exp\big(&\boldsymbol{\beta}_s \textit{ Recipient-Sender Similarity}_{ij}\\ &+\boldsymbol{\beta}_p \textit{ Recipient-Sender Personality Combinations}_{ij}\\ &+\boldsymbol{\beta}_{pw} \textit{ Recipient-Sender Personality Combinations}_{ij}\\ &\quad\times \textit{WOM Message Visibility}_{ij}\\ &+\boldsymbol{\beta}_w \textit{ WOM Message}_j + \boldsymbol{\beta}_e \textit{ User Expertise}_{ij}\\ &+\boldsymbol{\beta}_l \textit{ User Leadership}_{ij} + \boldsymbol{\beta}_c \textit{ Recipient-Sender and}\\ &\qquad \textit{WOM Message Controls}_{ij}\big), \qquad (2)\end{aligned}$$

where *Recipient-Sender Personality Combinations* captures the specific pairwise combinations of personality characteristics for recipients and senders of the WOM message. For instance, such recipient-sender personality combinations allow us to study whether extrovert-to-extrovert WOM communication or an introvert-to-introvert WOM communication is more effective.

4.2. Econometric Model Identification

The following sections discuss the machine-learning and natural-language-processing approaches we employed to analyze a vast amount of the user-generated unstructured data and identify our econometric specifications as well as the quasi-experimental research design we utilized to further distinguish the effect of personality characteristics and personality similarity on WOM effectiveness from unobserved confounders and correlated user behaviors.

4.2.1. Using Text-Mining for Extracting Personality Traits. The personality traits of the sender and the recipient of each WOM message and the corresponding *Recipient-Sender Personality Similarity* and *Recipient-Sender Personality Combinations* are derived based on a textual analysis of unstructured user-generated data. In particular, for each user, we analyzed the content of all the messages that were publicly posted on the social media platform over time as well as the user-defined description of their accounts. From the messages of the users analyzed are excluded any messages that were not written by the particular user each time (e.g., retweets), as those messages do not correspond to the linguistic style of the specific user and, hence, might not reflect his or her personality. In addition, we excluded all the private messages between the users as well as non-English messages. After the preprocessing of the "corpus" of user-generated content, there were on average 21,948 words per user; this number is higher than the typical number of words in other studies employing user personality attributes (e.g., Golbeck et al. 2011) and can lead to more accurate results.

The messages and the rest of the user-generated content of each user are merged into a single "document" as in Golbeck et al. (2011), and the latent personality traits and characteristics of individuals are then derived using linguistic analytics. In particular, following Golbeck et al. (2011) and Mahmud et al. (2013), the tokens of the user-generated content—after some preprocessing of the words, which includes removal of stop-words and non-English words, stemming, and fuzzy matching—are matched with the Linguistic Inquiry and Word Count (LIWC) psycholinguistic dictionary, which has been developed over several years and currently includes almost 4,500 words and word stems associated with one or more personality categories (Pennebaker et al. 2007), to compute relative scores in each dictionary category. Afterward, based on Yarkoni (2010), a weighted combination is estimated based on the coefficient between category scores and characteristics, using coefficients that were derived by comparing personality scores obtained from surveys with LIWC category scores from text (Tausczik and Pennebaker 2009, Yarkoni 2010). Finally, the outcome of these models is that for each one of the users in our data set, a score is estimated reflecting the percentile score for the specific characteristic. We also categorize the users into low-level and high-level groups for each personality characteristic depending on whether the corresponding score is less than or greater than the 50th percentile.

One of the advantages of the employed approach is that automated methods for personality assessment are more efficient and objective (Fast and Funder 2008) than traditional ways of measuring personality. In addition, the traditional way of measuring personality, which requires people to complete personality questionnaires, does not allow for obtaining personality traits at a large scale or low cost for the population of interest (Chen et al. 2015). Finally, user-generated content is more reflective of users' actual personalities, not an "idealized" version of themselves (Back et al. 2010).

4.2.2. Empirical Identification: Exogenous Variation in Message Visibility. This section discusses in detail the identification strategy for better distinguishing personality effects from correlated user behaviors and homophily. Our natural-like experiment research design is enabled by a unique feature of the microblogging platform that allows certain types of public messages to have different levels of visibility to the peers of the central user who is posting the message. Normally, a message posted by a user on the social media platform appears in the timeline of all the followers (i.e., anyone who is following the sender of the message); the timelines of users were not algorithmically curated during our study. Hence, in our context, whenever a user makes a purchase, social media peers of this user are exposed to the advocacy of the user toward the brand/product as the purchase is visible in their timelines, and therefore, their purchasing decisions might be affected through WOM. However, users who are connected in the platform (through a nonreciprocal or reciprocal relationship) tend to have similar preferences and idiosyncrasies. Hence, if one simply employs observational data under the aforementioned research design, it would be difficult to distinguish the actual effect a WOM instance might exert from simple correlations in users' behaviors and homophily; nonetheless, the discovery of correlations among latent personality traits and the effectiveness of WOM would be sufficient for forecasting objectives and practical marketing strategies.

In this study, we employ a research design framework that exploits whether a message broadcasted in the platform was rendered visible to particular users. In Twitter's platform, a publicly broadcasted message may not be visible in the timelines of some followers of a central user either because the particular follower missed viewing the actual tweet in his or her timeline or because it was directly addressed to another user account (i.e., the message began with another Twitter account username following the "@" symbol and there was no other character before this symbol). In particular, messages starting *immediately* with a specific Twitter account username (see Figure A2(b) in the online appendix) are visible only to the corresponding account (i.e., recipient of the message) and the set of followers who follow both the sender and the recipient, and only these users will be able to see the corresponding message in their timelines. In our context, such messages are not visible to the rest of the peers of the central user. On the other hand, if one (or more) characters (e.g., ".") appear before the "@" symbol (see Figure A2(a)), then the message is broadcasted to all the followers of the sender, although it is addressed to a specific account.[3] Hence, this unique feature of Twitter (i.e., the visibility of a message depends not only on whether another user, a brand, or the social commerce provider is mentioned in the message and the local social network of the followers of the sender but also on the absolute position of the "@" character) and the corresponding natural-like experiment induced by the differences in visibility of messages enable us to examine outcome measures for observations in treatment (i.e., visible message) and comparison (i.e., nonvisible message) groups. In this respect, this paper is also related to the stream of work that has leveraged the visibility of advertisements (Ghose and Todri-Adamopoulos 2016) to estimate the causal effect of online ads on consumer behavior.

Such a quasi-experiment creates an exogenous source of variation in the explanatory variables and allows us to identify the various effects related to being exposed to a friend's actual purchase and advocacy; both exposed (i.e., treatment) and nonexposed (i.e., control) group users are connected to a peer who completed a purchase. Hence, differences in purchases between treatment and control groups can then be attributed to the characteristics of their peer who made a purchase and the corresponding WOM messages they received. Thus, the microblogging service in our setting provides an ideal setting for identifying the impact of personality on the effectiveness of WOM. Besides taking advantage of this experimental design, we also avoid any observer biases, as our manipulation is nonintrusive; the subjects are completely unaware of being part of our experiment, and hence, they do not alter their behavior in anticipation of the experiment. The experiment occurred over 20 days during which 46,582 purchases were generated. Nevertheless, despite the variation in the treatment assignment, we control for differences in the pairwise relationships between users using an extensive set of constructs described in detail in Section 4.2.3 and also conduct several robustness checks as described in Section 5.2. Figures A4–A8 in the online appendix show the distribution of personality characteristics on the treatment and control groups illustrating the overlap of the two groups. Finally, we also allow for the visibility of the messages to be endogenous and build latent variable models to further control for potential unaccounted homophily while we also control for numerous potential confounders (see Section 5.2).

4.2.3. Additional Model Features. As illustrated by the outline of the model in Section 4.1, we expect that several factors and variables might affect a user's decision to make a purchase. In Section 4.2.1, we discussed in detail the "personality traits" and the corresponding machine-learning and natural-language-processing algorithms employed to identify the specific effects. In this section, we describe in detail the remaining effects and the corresponding constructs, including observed and latent pairwise user similarity, message advocacy, and user expertise and leadership as well as additional user and message controls.

The factor *Recipient-Sender Similarity* represents the similarity between the sender and the recipient of WOM messages and is measured based on the similarity of the two social media users in terms of overlap of the local communities as captured by the (i) Jaccard similarity coefficient of the sets of their *followers* and (ii) Jaccard similarity coefficient of *friends* as well as (iii) similarity of interests and topics discussed in social media posts based on the results of a latent Dirichlet allocation (LDA) model (Blei et al. 2003). LDA is a probabilistic generative model for natural language processing (NLP), which models every document in the corpus as a distribution over topics and every topic as a distribution over words. In our study, we build the LDA model on the corpus of all the messages of the users in our extended data set using a part-of-speech tagger developed specifically for the platform of Twitter (Owoputi et al. 2013) and following the estimation procedures of Hoffman et al. (2010). In particular, for the implementation of the LDA model, we used 139,850,033 messages. Moreover, we also find the *natural number of topics* that are present in our corpus based on the specific process and measure proposed by Arun et al. (2010), computed in terms of the Kullback–Leibler divergence (Kullback and Leibler 1951); our findings are not sensitive to the number of topics. Additionally, for the hyperparameters of our model, we learn an asymmetric prior directly from our data. Apart from measuring the sender-recipient similarity based on these distinct metrics, we also measure the similarity as a single standardized factor using factor analysis among the different metrics based on the principal factors method. The employed measures of user similarity capture both observed similarity (e.g., the number of interactions in the network) and *latent similarity* (e.g., latent common interests as captured by the topics of the LDA model) to better control for potentially unobserved confounders and homophily.

Moreover, the vector *WOM Message* represents the intensity and type of the WOM message and is captured by the *sentiment* of the tweet and whether the message was *personalized* (i.e., explicit rather than implicit advocacy). The sentiment of the message (measured in a continuous scale between -1 and $+1$) provides a richer metric of the intensity of the advocacy of the sender compared with other naïve metrics (e.g., lexicon-based scores). The main approach we employed uses a publicly available commercial sentiment analysis mechanism based on deep learning (AlchemyApi 2012).[4]

Furthermore, *User Expertise* is measured based on the standardized similarity of the timeline of a user with the timelines of the corresponding vendor and product using the probabilistic NLP model we previously described. In other words, this metric of latent user expertise and interests captures the intensity of the specific topics of interest in each user's discussions on the specific platform. The motivation for this metric is that, for instance, users who frequently broadcast messages about technological trends and topics similar to those in the discussions in the social media accounts of the specific technological products and their vendors are more likely to be perceived by their social media peers as experts in the area of technological products. We should note that although the *natural number of topics* is used, based on our empirical results, the findings of our study are not sensitive to the number of topics or the hyperparameters of the employed NLP model of latent user expertise.

The *User Leadership* is measured in terms of the additive smoothed *ratio of followers to followees*. The additive smoothed ratio is commonly applied in empirical studies to prevent the corresponding popularity or leadership metric from being oversensitive to small changes in the numbers of friends (followees) and followers. In addition to the social media leadership of the users, we also control for the *number of followers of the users* in the platform, whether each user has an officially *verified* account on the platform, and the number of public lists ("endorsements") in which other social media users have included the sender, as described in the following paragraphs.

Finally, the additional vector of *Recipient-Sender and WOM Message Controls* include the *number of messages* the recipients and senders have posted on the social media platform, the reciprocity of the relationship between the recipient and the sender of the WOM message, the number of interactions between the recipient and the sender, the number of *public lists* of which the sender (or recipient) are members, whether he or she still has a *default profile* in the platform, and whether the user has an officially *verified* account on the platform to further control for the popularity of the user and his or her level of engagement with the platform. We should note here that in the empirical identification of our econometric model, we employ additional user, pairwise relationship, message, and product controls as discussed in the following section.

Table 1 summarizes the main variables that were used in the analysis and shows the corresponding descriptive statistics. In our analysis, we only use observations corresponding to dyadic (pairwise) relationships and social media users who did not receive messages from multiple senders, similar to prior work (e.g., Aral and Walker 2012).

Table 1. Main Variables and Descriptive Statistics

Variable	Description	Median/Mean	SD	Min	Max
Purchase	Whether the recipient of the message made a purchase	0.015	0.12	0	1
Visible message	Whether the message was visible to the "recipient"	0.77	0.42	0	1
Number of followers	Number of followers	342	1,010,000	0	376,000
Number of friends	Number of followees	996	12,064	0	115,000
Number of messages	Number of messages posted	997	48,900	10	411,000
Number of list memberships	Number of lists the user is a member of	5	2,682	0	11,100
Default profile	Whether the user has a default profile	0.21	0.41	0	1
Verified account	Whether the user has a verified profile	0.01	0.10	0	1
Sentiment of message	Intensity of message advocacy	0.21	0.35	−1	1
Reciprocal relationship	Whether the relationship between the users is reciprocal	0.08	0.27	0	1
Number of peer-to-peer interactions	Number of interactions between users	0.26	6.10	0	1,612
Personalized message	Whether the message was personalized by the sender	0.82	0.38	0	1
User expertise	Level of expertise of the user with specific product/vendor	0.36	0.40	0	1
User reference	Whether a user is mentioned in the message	0.29	0.46	0	1
Agreeableness	Level of agreeableness in the personality of a user	31.32	28.23	0	100
Conscientiousness	Level of conscientiousness in the personality of a user	67.39	19.89	0	100
Extraversion	Level of extraversion in the personality of a user	36.08	28.64	0	100
Emotional range	Level of emotional range in the personality of a user	24.68	17.99	0	100
Openness	Level of openness in the personality of a user	75.34	18.05	0	100

Notes. For the variables *Number of followers, Number of friends, Number of messages*, and *Number of list memberships*, we report the median instead of the mean. The values of the variables *Number of followers, Number of friends, Number of messages, Number of list memberships, Default profile*, and *Verified account* correspond to the time of transmission of the WOM message.

5. Empirical Results

We estimate the effects of various user personality traits and characteristics of pairwise relationships on WOM messages by aggregating many individual experiments, in which the visibility of messages varies within and across the social media peers of the original consumers in the context of a social commerce venture with real physical products and monetary transactions of significant cost. We next present our key results on the effects of combinations of personality traits and personality similarity on the effectiveness of WOM. Then, we discuss the economic impact of our results and show robustness to a variety of alternative specifications and models.

5.1. Main Results

5.1.1. Main Results for Personality Similarity. Table 2 presents the results of the different specifications of our WOM effectiveness model for economic transactions and social commerce purchases (see Equation (1)). In particular, Model 1 constitutes our baseline specification and includes the constructs of dyadic similarity and strength of relationship between the recipient and sender of the WOM message (i.e., pairwise similarity between peers, reciprocity of relationship, and number of user interactions), WOM message advocacy (i.e., sentiment of message and personalized message), sender and recipient expertise and leadership, as well as additional sender and recipient controls (e.g., number of followers and officially verified profile). Model 2 introduces the notion of personality similarity based on the information of whether the sender and the recipient of the WOM message share the same main personality type (e.g., whether the dimension of extraversion has the highest level among the five personality dimensions for both the recipient and the sender), while Model 3

Table 2. Survival Analysis (Personality Similarity)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
User similarity	1.0994***	1.1003***	1.0974***	1.0953***	1.2962***	1.2942***	1.2967***	1.2797***
	(0.0043)	(0.0044)	(0.0044)	(0.0044)	(0.0219)	(0.0219)	(0.0216)	(0.0213)
Reciprocal relationship	7.3565***	7.3720***	6.7628***	6.8198***	6.9464***	6.9192***	6.9508***	6.3147***
	(0.3460)	(0.3470)	(0.3203)	(0.3229)	(0.3083)	(0.3072)	(0.3094)	(0.2891)
Number of peer-to-peer interactions	1.0008	1.0009	1.0006	1.0006	1.0006	1.0006	1.0008	1.0005
	(0.0008)	(0.0008)	(0.0009)	(0.0009)	(0.0008)	(0.0008)	(0.0008)	(0.0009)
Sentiment of message	1.6083***	1.6077***	1.5281***	1.5825***	1.6483***	1.6242***	1.5613***	1.4458***
	(0.1074)	(0.1074)	(0.1008)	(0.1054)	(0.1037)	(0.1025)	(0.0987)	(0.0947)
Personalized message	1.1084***	1.1088***	1.1079***	1.1057***	1.0559***	1.0540***	1.0486***	1.0387***
	(0.0052)	(0.0052)	(0.0052)	(0.0052)	(0.0039)	(0.0039)	(0.0052)	(0.0053)
User expertise (Sender)	1.2499***	1.2471***	1.2616***	1.2690***	1.2000***	1.1917***	1.2148***	1.1957***
	(0.0290)	(0.0291)	(0.0295)	(0.0297)	(0.0256)	(0.0254)	(0.0261)	(0.0265)
User leadership (Sender)	1.0138***	1.0138***	1.0110***	1.0108***	1.0098***	1.0099***	1.0100***	1.0048**
	(0.0017)	(0.0017)	(0.0017)	(0.0017)	(0.0018)	(0.0018)	(0.0018)	(0.0018)
Personality similarity (Main personality type)		0.9506						
		(0.0452)						
Personality similarity			1.4758***					
			(0.0480)					
Personality similarity (Agreeableness)				1.1074*	1.4608***	1.4591***	1.4427***	1.4369***
				(0.0494)	(0.1141)	(0.1141)	(0.1137)	(0.1132)
Personality similarity (Conscientiousness)				1.0681	0.9697	0.9746	0.9876	1.0121
				(0.0400)	(0.0474)	(0.0476)	(0.0483)	(0.0493)
Personality similarity (Extraversion)				1.3087***	0.9761	0.9649	0.9280	0.9005
				(0.0617)	(0.0658)	(0.0653)	(0.0635)	(0.0621)
Personality similarity (Emotional range)				1.0400	1.0061	1.0031	1.0143	1.0123
				(0.0417)	(0.0399)	(0.0399)	(0.0405)	(0.0406)
Personality similarity (Openness)				1.0486	0.9115	0.9084	0.9348	0.9199
				(0.0460)	(0.0656)	(0.0653)	(0.0667)	(0.0653)
Visible message = 1					1.3100***	1.2334**	1.1818*	1.4369***
					(0.0946)	(0.0902)	(0.0867)	(0.1075)
Visible message = 1 × *Personality similarity (Agreeableness)*					0.7616**	0.7720**	0.7735**	0.7633**
					(0.0682)	(0.0691)	(0.0696)	(0.0690)
Visible message = 1 × *Personality similarity (Conscientiousness)*					1.0860	1.0834	1.0690	1.0165
					(0.0666)	(0.0664)	(0.0659)	(0.0626)
Visible message = 1 × *Personality similarity (Extraversion)*					1.3490***	1.3487***	1.3952***	1.4503***
					(0.1106)	(0.1108)	(0.1156)	(0.1215)
Visible message = 1 × *Personality similarity (Emotional range)*					0.9993	0.9992	0.9925	1.0272
					(0.0559)	(0.0560)	(0.0558)	(0.0579)
Visible message = 1 × *Personality similarity (Openness)*					1.1875*	1.1970*	1.1526	1.1657
					(0.1000)	(0.1008)	(0.0966)	(0.0974)
Sender/recipient popularity controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Message controls	No	No	No	No	No	Yes	Yes	Yes
Product controls	No	No	No	No	No	No	Yes	Yes
Additional sender/recipient controls	No	No	No	No	No	No	No	Yes
Log-likelihood	−23,726.8	−23,726.3	−23,646.8	−23,650.8	−31,823.7	−31,810.8	−31,785.2	−31,592.7
BIC	47,573.06	47,583.85	47,424.87	47,480.72	63,915.88	63,902.27	63,887.75	63,612.58
AIC	47,473.69	47,474.55	47,315.56	47,331.67	63,691.39	63,667.58	63,622.45	63,255.45
χ^2	3,580.761	3,581.907	3,740.889	3,732.779	4,119.118	4,144.934	4,196.058	4,581.064
p	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
N	152,730	152,730	152,730	152,730	199,563	199,563	199,563	199,563

Notes. Semiparametric survival analysis with Cox proportional hazards model. The HRs represent the percent increase (HR > 1) or decrease (HR < 1) in postpurchase hazards associated with each attribute.

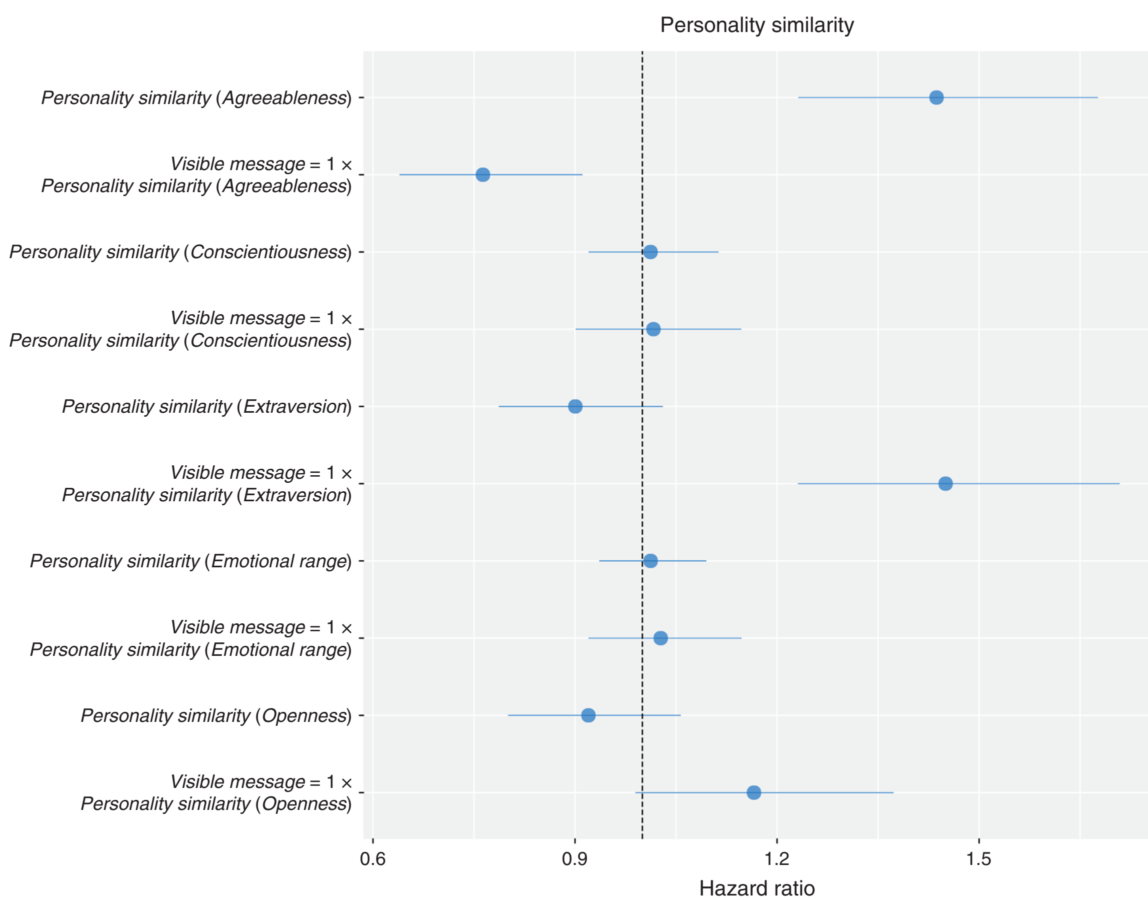
$^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

employs a richer metric of personality similarity, as it measures the similarity across the five factors of the personality model taxonomy (rather than just whether they share the same main personality type). Besides this, Model 4 decomposes the metric of personality similarity between the sender and the recipient of the message and contains detailed measures of the pairwise personality similarity for all five factors. Model 5 leverages the quasi-experimental design utilizing the information of message visibility to distinguish the effects of WOM through the social media posts from correlated behaviors and homophily among users as described in Section 4.2.2. Then, Model 6 captures whether a user was mentioned in the WOM message, Model 7 controls for the specific product mentioned in the WOM message, and Model 8 introduces additional sender and recipient controls (number of messages posted on the platform, number of lists the user is a member of, etc.).

To enhance the validity of our models, we also test the proportional-hazards assumption for all our models. On the basis of the results of the tests, we cannot reject the null hypothesis of zero slope for any of the models. The rejection of the null hypothesis of a zero slope would indicate a deviation from the proportional-hazards assumption. In other words, our models and the corresponding covariates do not violate the proportional-hazards assumption. We have also conducted various tests of collinearity and no issue was detected. In addition, all the employed models provide a very good fit to our data based on the information of the log-likelihood, the Akaike information criterion (AIC) and Bayesian information criterion (BIC) metrics, and the χ^2 statistic. We should also note that Models 5–8, which leverage the additional information of the message's visibility, better fit the data as indicated by the χ^2 statistic.

Figure 1 as well as Figure A9 in the online appendix provide a graphical representation of the results. In particular, Figure A9 shows the effects of user attributes, pairwise relationship characteristics, and user personality similarity on the effectiveness of WOM displaying the hazard ratios (HRs) representing the percent increase (HR > 1) or decrease (HR < 1) in postpurchase hazards associated with each attribute as well as the corresponding 95% confidence intervals (whiskers).

Figure 1. (Color online) Effects of Similarity of User Personality on Dyadic WOM Effectiveness



Notes. Effects are shown with 95% confidence intervals (whiskers). The figure displays HRs representing the percent increase (HR > 1) or decrease (HR < 1) in purchase hazards associated with each attribute. The econometric model controls for sender and recipient characteristics, message-related attributes, and product fixed effects.

Figure 1 shows only the effects of the main variables of interest.

From the hazard ratios shown in Table 2 and Figures 1 and A9, we see that the increased similarity and strength of relationship between users (*User similarity*: 1.0994, $p < 0.001$; *Reciprocal relationship*: 7.3565, $p < 0.001$) as well as more intense WOM message advocacy (*Sentiment of message*: 1.6083, $p < 0.001$; *Personalized message*: 1.1084, $p < 0.001$) are associated with higher levels of purchases after exposure to WOM messages. Similarly, users (i.e., senders of messages) with higher product expertise (*User expertise* (*Sender*): 1.2499, $p < 0.001$) and leadership (*User leadership* (*Sender*): 1.0138, $p < 0.001$) are also associated with higher likelihood of purchases of their peers after being exposed to their advocacy (i.e., recipients of messages). Then, from Models 2 and 3, we see that increased personality similarity across the five personality factors is associated with larger positive WOM effects (*Personality similarity*: 1.4578, $p < 0.001$); the less rich metric of binary similarity only on the prevalent personality type (*Personality similarity* (*Main personality type*)) did not identify a statistically significant result, providing evidence for the multidimensionality of the personality of users. This finding is in accordance with prior literature indicating that similarity on internal characteristics, based on a composite measure of various characteristics, enhances the effectiveness of sellers in traditional offline retail settings (Crosby et al. 1990, Lichtenthal and Tellefsen 2001). Decomposing the personality similarity into five constructs in Model 4, we see that similarity on agreeableness and extraversion (*Personality similarity* (*Agreeableness*): 1.1074, $p < 0.05$—*Personality similarity* (*Extraversion*): 1.3087, $p < 0.001$) are associated with statistically significant higher levels of purchases after exposure to WOM messages; similar users in terms of agreeableness and extraversion compared with dissimilar users are associated with a 10.74% and 30.87% increase in the likelihood of a purchase, respectively. This finding is in accordance with prior literature suggesting that individuals make decisions about relationships based on their similarity on such personality traits (Botwin et al. 1997).

We then distinguish the effects of WOM through the social media posts from correlated behaviors and homophily among users; the effect of WOM is transmitted through visible messages, whereas homophily is present even with nonvisible messages. From the results of Model 5, we see that the correlation of behaviors among peers is higher with higher user personality similarity at the levels of agreeableness (1.4608, $p < 0.001$), whereas WOM effects increase with lower similarity on agreeableness (*Visible message* = 1 × *Personality similarity* (*Agreeableness*): 0.7616, $p < 0.01$). On the other hand, WOM effects increase with higher similarity on extraversion (*Visible message* = 1 × *Personality similarity* (*Extraversion*): 1.3490, $p < 0.001$) as well as openness (*Visible message* = 1 × *Personality similarity* (*Openness*): 1.1875, $p < 0.05$); the effect of WOM for similar users in terms of extraversion and openness is accentuated by 34.90% and 18.75%, respectively. This finding is in accordance with prior literature, as it has been found that greater similarity in terms of extraversion and openness between jurors and expert witnesses in mock trials was correlated with increased perceived witness confidence and credibility (Gardner et al. 2013). We should note here that Model 5 exhibits a statistically significant increase in the fit to the data. Model 6 also controls for differences between WOM messages that mention a social media user and messages that do not mention a user. Finally, Models 7 and 8 corroborate these findings and further refine our estimates using additional controls for the products as well as the sender and the recipient of the WOM message. The above results and the corresponding effects are also of significant economic importance, as, for instance, a WOM message from a similar user in terms of personality (rather than a dissimilar one) increases the likelihood of a purchase by 47.58%.

5.1.2. Main Results for Combinations of Personality Characteristics. Table 3 extends our previous findings presenting the results of our model capturing the effect of pairwise personality characteristics on WOM effectiveness (see Equation (2)). These results go beyond the personality similarity of the sender and the recipient of the WOM message and introduce combinations of the specific personality characteristics of both the sender and the recipient of the WOM message. In particular, *Model* 1 is our baseline model measuring the effects of personality similarity for each of the five factors of the Big Five personality taxonomy. *Models* 2–5 extend this model and measure the effect of the pairwise personality characteristics of the sender and the recipient on the effectiveness of the WOM message. All models include detailed controls for both the sender and the recipient of the message, the message itself, and the products. The baseline case for Models 2–6 represents dyads in which the WOM message was not rendered visible, as described in Section 4.2.2.

We test the proportional-hazards assumption for these models as well. As before, on the basis of the results of the tests, we cannot reject the null hypothesis of zero slope for any of the models. Besides this, all employed models provide a very good fit to our data.

Figures A10–A14 in the online appendix provide a graphical representation of the results of Table 3 and show the effects of pairwise combinations of user personality characteristics on WOM effectiveness. These figures display hazard ratios representing the percent increase (HR > 1) or decrease (HR < 1) in purchase hazards associated with each attribute and the corresponding 95% confidence intervals (whiskers). In particular,

Table 3. Survival Analysis (Personality Characteristics)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Personality similarity (Agreeableness)	1.4427*** (0.1137)		1.3810*** (0.1068)	1.3232*** (0.0935)	1.4276*** (0.1121)	1.5220*** (0.1117)
Personality similarity (Conscientiousness)	0.9876 (0.0483)	1.1021* (0.0525)		1.0379 (0.0507)	1.1335** (0.0524)	1.0336 (0.0477)
Personality similarity (Extraversion)	0.9280 (0.0635)	1.0526 (0.0708)	0.9217 (0.0638)		0.8069** (0.0565)	0.9515 (0.0653)
Personality similarity (Emotional range)	1.0143 (0.0405)	1.0710 (0.0455)	1.0852* (0.0433)	0.9787 (0.0400)		0.9562 (0.0379)
Personality similarity (Openness)	0.9348 (0.0667)	1.0941 (0.0763)	1.0107 (0.0689)	0.8564* (0.0592)	0.9707 (0.0635)	
Visible message = 1 × *Personality similarity (Agreeableness)*	0.7735** (0.0696)		0.7669** (0.0686)	0.8583 (0.0701)	0.7470** (0.0677)	0.7316*** (0.0620)
Visible message = 1 × *Personality similarity (Conscientiousness)*	1.0690 (0.0659)	0.9832 (0.0597)		1.0234 (0.0631)	1.0106 (0.0554)	1.0274 (0.0612)
Visible message = 1 × *Personality similarity (Extraversion)*	1.3952*** (0.1156)	1.1737* (0.0942)	1.4223*** (0.1194)		1.5592*** (0.1321)	1.3837*** (0.1139)
Visible message = 1 × *Personality similarity (Emotional range)*	0.9925 (0.0558)	0.9241 (0.0541)	0.9852 (0.0514)	1.0302 (0.0587)		1.0348 (0.0572)
Visible message = 1 × *Personality similarity (Openness)*	1.1526 (0.0966)	0.9637 (0.0808)	1.0880 (0.0890)	1.1923* (0.0979)	1.1559 (0.0917)	
Low Agreeableness (Sender) × *Low Agreeableness (Recipient)*		0.6159*** (0.0491)				
Low Agreeableness (Sender) × *High Agreeableness (Recipient)*		0.5593*** (0.0555)				
High Agreeableness (Sender) × *Low Agreeableness (Recipient)*		1.3050** (0.1295)				
High Agreeableness (Sender) × *High Agreeableness (Recipient)*		1.0843 (0.1376)				
Low Conscientiousness (Sender) × *Low Conscientiousness (Recipient)*			0.6902* (0.1151)			
Low Conscientiousness (Sender) × *High Conscientiousness (Recipient)*			0.9953 (0.1026)			
High Conscientiousness (Sender) × *Low Conscientiousness (Recipient)*			1.6170*** (0.1975)			
High Conscientiousness (Sender) × *High Conscientiousness (Recipient)*			1.7278*** (0.1485)			
Low Extraversion (Sender) × *Low Extraversion (Recipient)*				1.1884* (0.0912)		
Low Extraversion (Sender) × *High Extraversion (Recipient)*				0.5808*** (0.0773)		
High Extraversion (Sender) × *Low Extraversion (Recipient)*				1.7128*** (0.2014)		
High Extraversion (Sender) × *High Extraversion (Recipient)*				0.9471 (0.1921)		
Low Emotional range (Sender) × *Low Emotional range (Recipient)*					1.6119*** (0.1244)	
Low Emotional range (Sender) × *High Emotional range (Recipient)*					1.0822 (0.1391)	
High Emotional range (Sender) × *Low Emotional range (Recipient)*					0.9103 (0.0958)	
High Emotional range (Sender) × *High Emotional range (Recipient)*					0.5369** (0.1034)	
Low Openness (Sender) × *Low Openness (Recipient)*						0.9529 (0.3461)
Low Openness (Sender) × *High Openness (Recipient)*						0.7611 (0.1391)
High Openness (Sender) × *Low Openness (Recipient)*						1.7139*** (0.2695)

Table 3. (Continued)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
High Openness (*Sender*) × *High Openness* (*Recipient*)						1.3596*** (0.1062)
Visible message = 1	1.1818* (0.0867)					
Sender/recipient controls	Yes	Yes	Yes	Yes	Yes	Yes
Message controls	Yes	Yes	Yes	Yes	Yes	Yes
Product controls	Yes	Yes	Yes	Yes	Yes	Yes
Log-likelihood	−31,785.2	−31,749.4	−31,731.7	−31,717	−31,647.6	−31,747.1
BIC	63,887.75	63,852.65	63,817.28	63,787.94	63,649.13	63,848.09
AIC	63,622.45	63,556.74	63,521.37	63,492.03	63,353.21	63,552.18
χ^2	4,196.058	4,267.771	4,303.139	4,332.482	4,471.298	4,272.33
p	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
N	199,563	199,563	199,563	199,563	199,563	199,563

Notes. Semiparametric survival analysis with Cox proportional hazards model. The HRs represent the percent increase (HR > 1) or decrease (HR < 1) in postpurchase hazards associated with each attribute. The baseline case for Models 2–6 represents dyads in which the WOM message was not visible.

$^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

Figure A10 illustrates the effects of the level of the agreeableness personality trait (for both the sender and the recipient of the message) on WOM effectiveness. Figures A11–A14 show the effects of the levels of conscientiousness, extraversion, emotional range, and openness, respectively.

Based on the results shown in Table 3, as discussed above, the correlation in user behaviors is higher with higher user personality similarity at the levels of agreeableness (1.4427, $p < 0.001$), whereas WOM effects increase with lower similarity on agreeableness (0.7735, $p < 0.01$). On the other hand, WOM effects increase with higher similarity on extraversion (1.3952, $p < 0.001$) and openness (1.1526, $p < 0.1$). We further extend these results capturing the effects of combinations of personality characteristics of the sender and the recipient of a WOM message. In particular, based on the results shown in Table 3 and Figures A10–A14, WOM originating from senders with personality with high levels of agreeableness is effective (*High Agreeableness* (*Sender*) × *Low Agreeableness* (*Recipient*): 1.3050, $p < 0.01$), whereas users with low levels of agreeableness exhibit opposite effects (*Low Agreeableness* (*Sender*) × *Low Agreeableness* (*Recipient*): 0.6159, $p < 0.001$; *Low Agreeableness* (*Sender*) × *High Agreeableness* (*Recipient*): 0.5593, $p < 0.001$). This result finds support in prior literature indicating that agreeableness consists of tendencies of individuals to be sympathetic, trusting, and trustworthy (Costa and McCrae 1992) as well as that agreeableness is associated with motives to maintain positive interpersonal relations (Jensen-Campbell and Graziano 2001). Prior literature also shows that the personality trait of agreeableness is a predictor of successful transformational management (Judge and Bono 2000); transformational leaders gain support by inspiring and engaging followers. Besides this, our findings also reveal that senders with personalities characterized by high levels of conscientiousness are more likely to be effective advocates (*High Conscientiousness* (*Sender*) × *Low Conscientiousness* (*Recipient*): 1.6170, $p < 0.001$; *High Conscientiousness* (*Sender*) × *High Conscientiousness* (*Recipient*): 1.7278, $p < 0.001$), whereas the pair *Low Conscientiousness* (*Sender*) × *Low Conscientiousness* (*Recipient*) seems to have the opposite effect (0.6902, $p < 0.05$). This result is supported by prior literature, as the conscientiousness trait represents the tendency of an individual to be reliable, responsible, and self-assured (Barrick and Mount 1991). Prior literature has also positively linked conscientiousness to sales performance (Barrick et al. 1993). Furthermore, our findings indicate that recipients characterized by low levels of extraversion are more responsive to WOM effects (*Low Extraversion* (*Sender*) × *Low Extraversion* (*Recipient*): 1.1884, $p < 0.01$; *High Extraversion* (*Sender*) × *Low Extraversion* (*Recipient*): 1.7128, $p < 0.001$), while recipients with high levels of extraversion seem not to be responsive to WOM effects (*Low Extraversion* (*Sender*) × *High Extraversion* (*Recipient*): 0.5808, $p < 0.001$). The relationship between this personality trait and WOM effectiveness is an interesting and unexpected finding. In particular, Ogunlade (1979) reports that extroverts are more susceptible to contagion than introverts based on an experiment that defines contagion as imitation of another person's behavior. However, this divergent finding may well be because contagion in that experimental design involved disobeying rules (Ogunlade 1979) and that extroverts defy rules more often than introverts (Parish et al. 1965). Additionally, when a pair of users is characterized by low levels of emotional range, then strong WOM effects are more likely (*Low Emotional range* (*Sender*) × *Low Emotional range* (*Recipient*): 1.6119, $p < 0.001$); when both peers exhibit high levels of emotional range, the opposite effect is more likely (*High Emotional* *range* (*Sender*) × *High Emotional range* (*Recipient*): 0.5369, $p < 0.01$). The fact that high levels of emotional range are not linked to the effective dissemination of WOM is in accordance with prior research, as it has been shown that high levels of emotional range signal a lack of self-confidence and self-esteem (McCrae and Costa 1991). It is worth noting that high levels of emotional range (also referred as neuroticism) have been associated with negative behaviors, such as the inability to maintain personal relationships (Karney and Bradbury 1997). Similarly, low levels of emotional range are linked to successful advising and friendships (Klein et al. 2004), providing support for our findings. Finally, our study also reveals that WOM originating from senders with higher levels of openness is more likely to be effective (*High Openness* (*Sender*) × *Low Openness* (*Recipient*): 1.7139, $p < 0.001$; *High Openness* (*Sender*) × *High Openness* (*Recipient*): 1.3596, $p < 0.001$). High levels of openness indicate that a person is more curious, creative, and open to novel ideas and viewpoints, and they have been found to enhance the perceived persuasiveness of an individual (Oreg and Sverdlik 2014). However, prior literature has not identified an actual change of behavior (of recipients) beyond perceived persuasion (of senders) (Oreg and Sverdlik 2014). Overall, these results illustrate the economic importance of the effects for both marketers and social media companies, as, for instance, a positive WOM message from an extroverted user to an introverted peer is associated with an increase of 71.28% in the likelihood for a subsequent purchase.

5.1.3. Predictive Ability. To assess the out-of-sample performance of our models and validate our aforementioned findings, we employ a holdout evaluation scheme with an 80/20 random split of data and evaluate each model in terms of concordance (i.e., the probability that predictions and outcomes are concordant). In particular, Tables 4 and 5 present for each model the Harrell's C concordance coefficient, which measures the likelihood of correctly ordering survival times for pairs of dyads of senders and recipients of WOM messages. The concordance measure is similar to the Mann-Whitney-Wilcoxon test statistic as well as the area under the ROC curve metric and takes values between 0 and 1, with a value of 0.5 indicating baseline performance and a value of 1 perfect predictive discrimination. Figures 2 and 3 provide a graphical illustration of the out-of-sample performance of the employed models and specifications. Based on the results, in addition to very good explanatory power, all the employed models exhibit very good out-of-sample performance, outperforming the baseline by a large margin and illustrating the predictive ability of the models. It is interesting to note in Figure 2 the increased out-of-sample performance that is exhibited by the models that leverage the information of the visibility of the message. This difference is also statistically significant. Finally, we should also note that Model 8 achieved the best out-of-sample predictive performance.

5.2. Robustness Checks

5.2.1. Alternative Estimators. Enhancing the robustness of our findings, we expand our analysis and also model *whether* a user will purchase a product after being exposed to the advocacy of the WOM message (rather than *how quickly* a peer will purchase a product). We model this decision of a user (i.e., recipient of message) to make a purchase or not employing discrete choice models in a hedonic-like framework. The model platform is an underlying random utility model or latent regression model (McFadden 1973), $y^* = \mathbf{x}\boldsymbol{\beta} + \epsilon$, in which $\mathbf{x}\boldsymbol{\beta}$ contains the constructs discussed in Section 4.1 and the continuous latent utility or measure, y^*, is observed in discrete form through a censoring mechanism:

$$\begin{aligned} \mathit{purchase} &= 1 \quad \text{if } y^* > 0, \quad \text{and} \\ \mathit{purchase} &= 0 \quad \text{if } y^* \le 0. \end{aligned}$$

Tables 6 and 7 present the results of the models for the effect of personality similarity of the sender and the recipient of the WOM message and the effects of

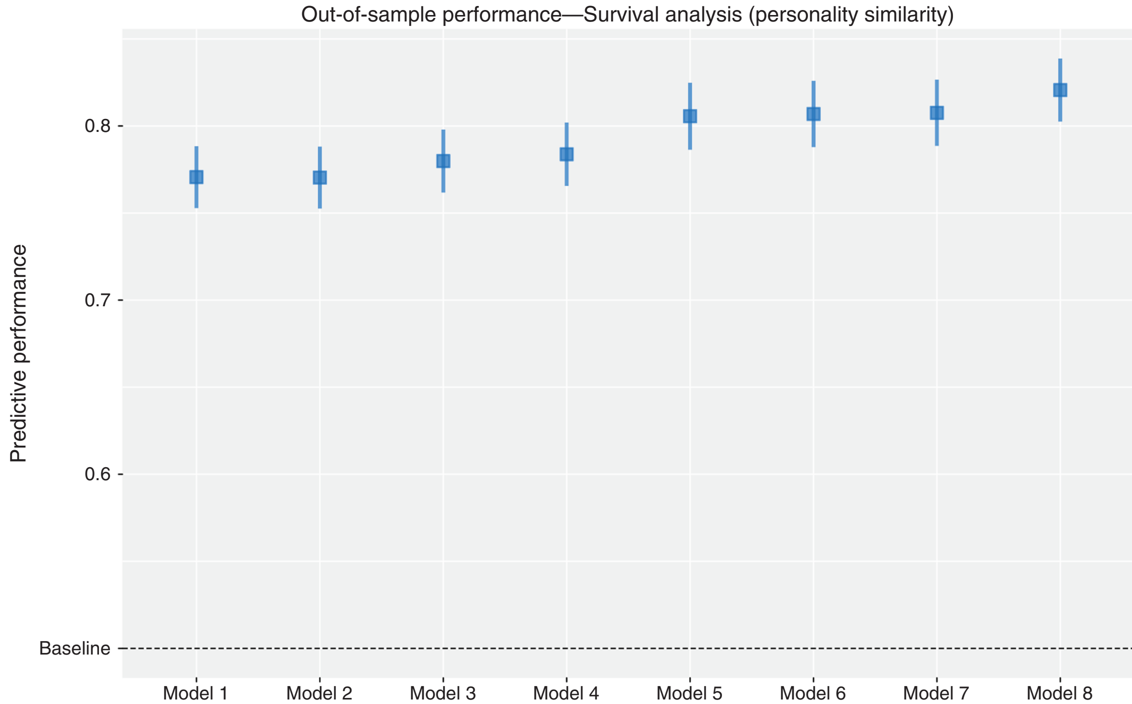
Table 4. Out-of-Sample Predictive Power—Survival Analysis (Personality Similarity)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Coefficient	0.7706	0.7704	0.7798	0.7838	0.8056	0.8069	0.8076	0.8206
Jackknife SE	0.0091	0.0091	0.0093	0.0093	0.0099	0.0097	0.0097	0.0093
T	84.7874	84.5486	84.2728	84.3143	81.7640	82.7620	82.9960	88.6130
$P > \lvert t \rvert$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Table 5. Out-of-Sample Predictive Power—Survival Analysis (Personality Characteristics)

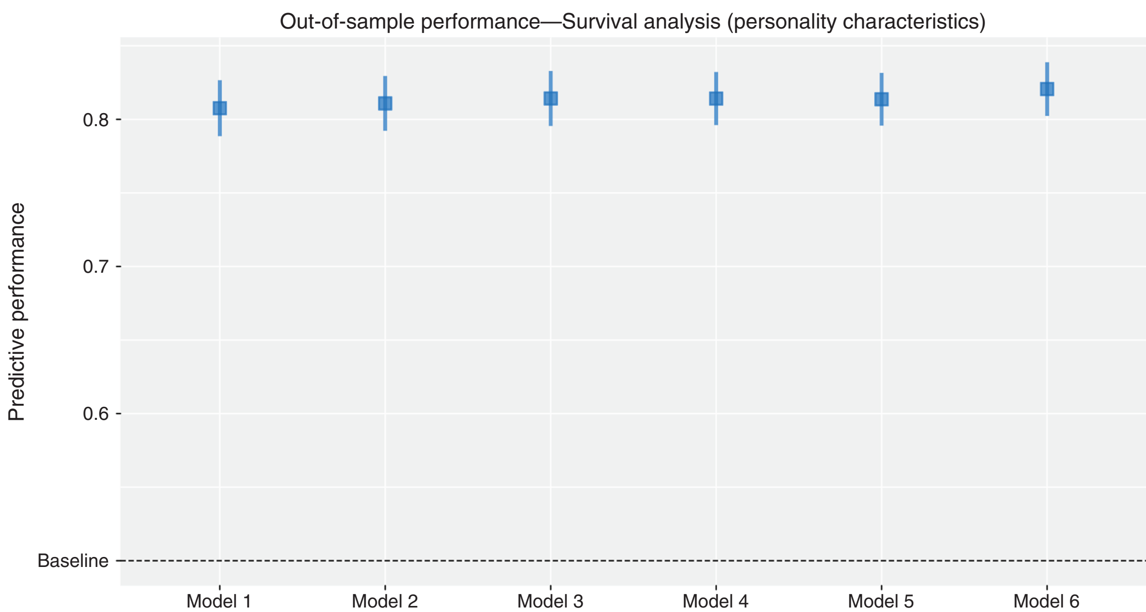
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Coefficient	0.8076	0.8108	0.8142	0.8141	0.8136	0.8206
Jackknife SE	0.0097	0.0095	0.0096	0.0092	0.0092	0.0093
t	83.0052	84.9336	85.1285	88.0351	88.6176	88.0250
$P > \lvert t \rvert$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Figure 2. (Color online) Out-of-Sample Performance of Models of Dyadic WOM Effectiveness



Notes. Performance is shown with 95% confidence intervals (whiskers). The figure displays Harrell's *C* concordance coefficient representing the number of times that we can correctly order survival times for pairs of dyadic relationships. The baseline case represents randomly indicating the order of a pairs of dyadic relationships.

Figure 3. (Color online) Out-of-Sample Performance of Models of the Effect of Personality Types on Dyadic WOM Effectiveness



Notes. Performance is shown with 95% confidence intervals (whiskers). The figure displays Harrell's *C* concordance coefficient representing the number of times that we can correctly order survival times for pairs of dyadic relationships. The baseline case represents randomly indicating the order of a pairs of dyadic relationships.

combinations of the specific personality characteristics on WOM effectiveness, respectively. All models employ the same constructs and controls as the models presented in Section 5.1. The results further corroborate our previous findings. Hence, both combinations of personality characteristics and personality similarity of peers affect whether a user will conduct a purchase as well as how quickly he or she will purchase a product.

To control for potentially confounding effects and nonrandom visibility of messages, we also employ the

Table 6. Discrete Choice Model (Personality Similarity)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
User similarity	1.2535***	1.2557***	1.2385***	1.2356***	1.4305***	1.4279***	1.4400***	1.4173***
	(0.0131)	(0.0133)	(0.0130)	(0.0129)	(0.0358)	(0.0357)	(0.0366)	(0.0358)
Reciprocal relationship	5.8839***	5.8926***	5.5813***	5.6300***	5.6088***	5.5785***	5.5942***	5.1470***
	(0.3152)	(0.3158)	(0.2988)	(0.3013)	(0.2796)	(0.2785)	(0.2802)	(0.2634)
Number of peer-to-peer interactions	0.9995	0.9995	0.9993	0.9993	0.9992	0.9992	0.9994	0.9991
	(0.0012)	(0.0012)	(0.0012)	(0.0012)	(0.0012)	(0.0012)	(0.0012)	(0.0012)
Sentiment of message	1.5841***	1.5848***	1.5074***	1.5621***	1.6310***	1.6124***	1.5396***	1.4452***
	(0.1099)	(0.1100)	(0.1036)	(0.1083)	(0.1066)	(0.1058)	(0.1012)	(0.0989)
Personalized message	1.1087***	1.1091***	1.1087***	1.1066***	1.0547***	1.0526***	1.0455***	1.0361***
	(0.0056)	(0.0056)	(0.0056)	(0.0057)	(0.0041)	(0.0041)	(0.0054)	(0.0056)
User expertise (*Sender*)	1.2207***	1.2169***	1.2332***	1.2438***	1.1738***	1.1674***	1.1892***	1.1581***
	(0.0298)	(0.0298)	(0.0303)	(0.0305)	(0.0264)	(0.0261)	(0.0269)	(0.0271)
User leadership (*Sender*)	1.0130***	1.0129***	1.0104***	1.0101***	1.0090***	1.0092***	1.0092***	1.0043*
	(0.0017)	(0.0017)	(0.0017)	(0.0018)	(0.0018)	(0.0018)	(0.0018)	(0.0019)
Personality similarity (*Main personality type*)		0.934						
		(0.0461)						
Personality similarity			1.4447***					
			(0.0486)					
Personality similarity (*Agreeableness*)				1.0923	1.4566***	1.4545***	1.4306***	1.4239***
				(0.0498)	(0.1164)	(0.1163)	(0.1152)	(0.1146)
Personality similarity (*Conscientiousness*)				1.0544	0.9622	0.9671	0.9765	1.0045
				(0.0407)	(0.0484)	(0.0486)	(0.0491)	(0.0502)
Personality similarity (*Extraversion*)				1.3009***	0.9658	0.9540	0.9144	0.8877
				(0.0629)	(0.0668)	(0.0663)	(0.0642)	(0.0628)
Personality similarity (*Emotional range*)				1.0394	1.0078	1.0049	1.0175	1.0154
				(0.0428)	(0.0410)	(0.0409)	(0.0416)	(0.0417)
Personality similarity (*Openness*)				1.0469	0.9109	0.9082	0.9404	0.9231
				(0.0470)	(0.0670)	(0.0667)	(0.0685)	(0.0669)
Visible message = 1					1.2471**	1.1687*	1.1129	1.3727***
					(0.0953)	(0.0905)	(0.0867)	(0.1086)
Visible message = 1 × *Personality similarity* (*Agreeableness*)					0.7536**	0.7660**	0.7682**	0.7492**
					(0.0691)	(0.0702)	(0.0707)	(0.0694)
Visible message = 1 × *Personality similarity* (*Conscientiousness*)					1.0818	1.0791	1.0711	1.0121
					(0.0683)	(0.0680)	(0.0679)	(0.0642)
Visible message = 1 × *Personality similarity* (*Extraversion*)					1.3481***	1.3473***	1.3993***	1.4685***
					(0.1133)	(0.1134)	(0.1189)	(0.1263)
Visible message = 1 × *Personality similarity* (*Emotional range*)					0.9972	0.9977	0.9889	1.0277
					(0.0572)	(0.0573)	(0.0571)	(0.0596)
Visible message = 1 × *Personality similarity* (*Openness*)					1.1831	1.1921*	1.1410	1.1568
					(0.1019)	(0.1026)	(0.0977)	(0.0988)
Constant	0.0059***	0.0060***	0.0051***	0.0050***	0.0047***	0.0048***	0.0034***	0.0036***
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0004)	(0.0004)	(0.0003)	(0.0003)
Sender/recipient popularity controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Message controls	No	No	No	No	No	Yes	Yes	Yes
Product controls	No	No	No	No	No	No	Yes	Yes
Additional sender/recipient controls	No	No	No	No	No	No	No	Yes
Log-likelihood	−9,371.1	−9,370.1	−9,304.9	−9,307.7	−12,473	−12,459.8	−12,433	−12,239
BIC	18,873.51	18,883.52	18,753.02	18,806.32	25,226.64	25,212.55	25,195.58	24,917.32
AIC	18,764.21	18,764.28	18,633.79	18,647.34	24,991.95	24,967.66	24,920.07	24,549.98
χ^2	3,765.623	3,767.552	3,898.046	3,892.493	4,319.754	4,346.05	4,399.636	4,787.724
p	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
N	152,730	152,730	152,730	152,730	199,563	199,563	199,563	199,563

Notes. WOM effectiveness analysis with logistic regression model. The log odds (LOs) represent the percent increase (LO > 1) or decrease (LO < 1) in postpurchase likelihood associated with each attribute.
$^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

Table 7. Discrete Choice Model (Personality Characteristics)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Personality similarity (Agreeableness)	1.4306*** (0.1152)		1.3689*** (0.1088)	1.3112*** (0.0945)	1.4147*** (0.1139)	1.5201*** (0.1146)
Personality similarity (Conscientiousness)	0.9765 (0.0491)	1.0516 (0.0518)		1.0297 (0.0519)	1.1252* (0.0539)	1.0318 (0.0493)
Personality similarity (Extraversion)	0.9144 (0.0642)	1.0202 (0.0701)	0.9060 (0.0646)		0.7949** (0.0573)	0.9450 (0.0666)
Personality similarity (Emotional range)	1.0175 (0.0416)	1.0663 (0.0458)	1.0890* (0.0448)	0.9846 (0.0411)		0.9569 (0.0388)
Personality similarity (Openness)	0.9404 (0.0685)	1.1001 (0.0778)	1.0122 (0.0709)	0.8610* (0.0608)	0.9742 (0.0654)	
Visible message = 1 × *Personality similarity (Agreeableness)*	0.7682** (0.0707)		0.7611** (0.0699)	0.8463* (0.0707)	0.7426** (0.0691)	0.7165*** (0.0625)
Visible message = 1 × *Personality similarity (Conscientiousness)*	1.0711 (0.0679)	1.0228 (0.0642)		1.0178 (0.0648)	1.008 (0.0574)	1.0165 (0.0627)
Visible message = 1 × *Personality similarity (Extraversion)*	1.3993*** (0.1189)	1.1925* (0.0981)	1.4324*** (0.1236)		1.5744*** (0.1371)	1.3778*** (0.1161)
Visible message = 1 × *Personality similarity (Emotional range)*	0.9889 (0.0571)	0.9305 (0.0557)	0.9816 (0.0527)	1.0266 (0.0600)		1.0333 (0.0587)
Visible message = 1 × *Personality similarity (Openness)*	1.1410 (0.0977)	0.9393 (0.0804)	1.0848 (0.0912)	1.1743 (0.0984)	1.1460 (0.0935)	
Low Agreeableness (Sender) × *Low Agreeableness (Recipient)*		0.6710*** (0.0559)				
Low Agreeableness (Sender) × *High Agreeableness (Recipient)*		0.5914*** (0.0615)				
High Agreeableness (Sender) × *Low Agreeableness (Recipient)*		1.3549** (0.1415)				
High Agreeableness (Sender) × *High Agreeableness (Recipient)*		1.1250 (0.1494)				
Low Conscientiousness (Sender) × *Low Conscientiousness (Recipient)*			0.6411* (0.1122)			
Low Conscientiousness (Sender) × *High Conscientiousness (Recipient)*			0.9729 (0.1046)			
High Conscientiousness (Sender) × *Low Conscientiousness (Recipient)*			1.5839*** (0.2047)			
High Conscientiousness (Sender) × *High Conscientiousness (Recipient)*			1.6994*** (0.1571)			
Low Extraversion (Sender) × *Low Extraversion (Recipient)*				1.1253 (0.0914)		
Low Extraversion (Sender) × *High Extraversion (Recipient)*				0.5327*** (0.0730)		
High Extraversion (Sender) × *Low Extraversion (Recipient)*				1.5673*** (0.1933)		
High Extraversion (Sender) × *High Extraversion (Recipient)*				0.8769 (0.1816)		
Low Emotional range (Sender) × *Low Emotional range (Recipient)*					1.5475*** (0.1268)	
Low Emotional range (Sender) × *High Emotional range (Recipient)*					1.0105 (0.1376)	
High Emotional range (Sender) × *Low Emotional range (Recipient)*					0.8661 (0.0964)	
High Emotional range (Sender) × *High Emotional range (Recipient)*					0.4807*** (0.0977)	
Low Openness (Sender) × *Low Openness (Recipient)*						0.8289 (0.3129)

Table 7. (Continued)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Low Openness (Sender)						0.6833^{*}
× *High Openness (Recipient)*						(0.1299)
High Openness (Sender)						1.6568^{**}
× *Low Openness (Recipient)*						(0.2694)
High Openness (Sender)						1.2915^{**}
× *High Openness (Recipient)*						(0.1070)
Visible message = 1	1.1129					
	(0.0867)					
Constant	0.0034^{***}	0.0100^{***}	0.0083^{***}	0.0053^{***}	0.0013^{***}	0.0021^{***}
	(0.0003)	(0.0010)	(0.0013)	(0.0006)	(0.0002)	(0.0004)
Sender/recipient controls	Yes	Yes	Yes	Yes	Yes	Yes
Message controls	Yes	Yes	Yes	Yes	Yes	Yes
Product controls	Yes	Yes	Yes	Yes	Yes	Yes
Log-likelihood	−12,433	−12,383.7	−12,376.5	−12,367.9	−12,301.9	−12,392.4
BIC	25,195.58	25,133.57	25,119.06	25,101.89	24,969.9	25,151.01
AIC	24,920.07	24,827.46	24,812.94	24,795.77	24,663.78	24,844.89
χ^2	4,399.636	4,498.252	4,512.764	4,529.936	4,661.925	4,480.816
P	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
N	199,563	199,563	199,563	199,563	199,563	199,563

Notes. WOM effectiveness analysis with logistic regression model. The log odds (LOs) represent the percent increase (LO > 1) or decrease (LO < 1) in postpurchase likelihood associated with each attribute. The baseline case for Models (2)–(6) represents dyads in which the WOM message was not visible.

$^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

propensity score matching technique using as covariates the similarity measures for the users verifying that there are no statistically significant differences between observations in the treatment and control groups; the results are similar for alternative matching rules and covariates. Tables 8 and 9 present the results of the models for the WOM effectiveness over the matched sample.

5.2.2. Deep Learning. In addition, to enhance even further our identification strategy and provide additional evidence of the robustness of our findings, we also control for the latent characteristics of the users tapping into the social network structure and the advances of deep learning. In particular, we use the method of DeepWalk, a deep-learning method for graphs (Perozzi et al. 2014), to learn the latent representations of the users and their similarity and control for both latent user homophily and network structure roles. Tables 10 and 11 present the corresponding results. As shown in these tables, the results including these additional controls further corroborate our findings. Similar results are also achieved employing the method of node2vec (Grover and Leskovec 2016).

5.2.3. Additional Robustness Checks. Moreover, apart from the semiparametric survival analysis that was presented in the previous section, we also estimate the previously described models using a parametric survival analysis proportional hazards model of log relative-hazard form. Table A1 in the online appendix shows the effects of personality similarity on WOM effectiveness. Then, Table A2 in the online appendix presents the effects of combinations of the specific personality characteristics of the sender and recipient of WOM messages on WOM effectiveness. The results corroborate our previous findings indicating that our findings are robust across various econometric specifications; very similar results are also obtained employing an accelerated failure-time model as well.

Furthermore, we also account for potentially unobserved or unaccounted correlations among user dyads (Lin and Wei 1989). Tables A3 and A4 in the online appendix present the results of the models for the WOM effectiveness under this alternative specification. The results corroborate our findings. Besides this, Tables A5 and A6 control whether the sender decided to make the message visible using a common norm among the users of Twitter (i.e., start a message with a dot as the first character). Then, Tables A7 and A8 show the corresponding results after further accounting for user homophily by controlling for user similarity in intrinsic tastes and preferences based on the overlap in brands that each user follows in the social network platform and the related industries of these brands (e.g., electronics, foods). Finally, to further control for potentially unaccounted homophily, we build latent variable models (i.e., structural models) where the similarity between the sender and the recipient is latent and measured based on the features described in Section 4.2.3 including information at the LDA topic level. Tables A9 and A10 show the corresponding results. The results of all the aforementioned alternative specifications and models are highly consistent and further corroborate our findings.

Table 8. Propensity Score Matching (Personality Similarity)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
User similarity	1.0994***	1.1003***	1.0974***	1.0953***	1.2928***	1.2909***	1.2930***	1.2748***
	(0.0043)	(0.0044)	(0.0044)	(0.0044)	(0.0235)	(0.0234)	(0.0231)	(0.0228)
Reciprocal relationship	7.3565***	7.3720***	6.7628***	6.8198***	7.0126***	6.9875***	7.0195***	6.4013***
	(0.3460)	(0.3470)	(0.3203)	(0.3229)	(0.3131)	(0.3120)	(0.3142)	(0.2949)
Number of peer-to-peer interactions	1.0008	1.0009	1.0006	1.0006	1.0006	1.0006	1.0008	1.0005
	(0.0008)	(0.0008)	(0.0009)	(0.0009)	(0.0008)	(0.0009)	(0.0008)	(0.0009)
Sentiment of message	1.6083***	1.6077***	1.5281***	1.5825***	1.6398***	1.6148***	1.5564***	1.4412***
	(0.1074)	(0.1074)	(0.1008)	(0.1054)	(0.1039)	(0.1026)	(0.0991)	(0.0951)
Personalized message	1.1084***	1.1088***	1.1079***	1.1057***	1.0653***	1.0633***	1.0571***	1.0468***
	(0.0052)	(0.0052)	(0.0052)	(0.0052)	(0.0041)	(0.0041)	(0.0054)	(0.0055)
User expertise (Sender)	1.2499***	1.2471***	1.2616***	1.2690***	1.2059***	1.1973***	1.2201***	1.1993***
	(0.0290)	(0.0291)	(0.0295)	(0.0297)	(0.0260)	(0.0258)	(0.0265)	(0.0269)
User leadership (Sender)	1.0138***	1.0138***	1.0110***	1.0108***	1.0099***	1.0100***	1.0102***	1.0052**
	(0.0017)	(0.0017)	(0.0017)	(0.0017)	(0.0018)	(0.0018)	(0.0018)	(0.0018)
Personality similarity (Main personality type)		0.9506						
		(0.0452)						
Personality similarity			1.4758***					
			(0.0480)					
Personality similarity (Agreeableness)				1.1074*	1.4372***	1.4353***	1.4186***	1.4117***
				(0.0494)	(0.1247)	(0.1247)	(0.1242)	(0.1235)
Personality similarity (Conscientiousness)				1.0681	0.9845	0.9897	1.0039	1.0299
				(0.0400)	(0.0528)	(0.0530)	(0.0539)	(0.0550)
Personality similarity (Extraversion)				1.3087***	1.0065	0.9949	0.9542	0.9245
				(0.0617)	(0.0748)	(0.0742)	(0.0720)	(0.0703)
Personality similarity (Emotional range)				1.0400	1.0084	1.0052	1.0155	1.0117
				(0.0417)	(0.0444)	(0.0444)	(0.0451)	(0.0452)
Personality similarity (Openness)				1.0486	0.9313	0.9278	0.9543	0.9376
				(0.0460)	(0.0736)	(0.0733)	(0.0748)	(0.0731)
Visible message = 1					1.3602***	1.2842***	1.2256**	1.4779***
					(0.1018)	(0.0973)	(0.0932)	(0.1145)
Visible message = 1 × *Personality similarity (Agreeableness)*					0.7726**	0.7827*	0.7858*	0.7756**
					(0.0752)	(0.0761)	(0.0768)	(0.0761)
Visible message = 1 × *Personality similarity (Conscientiousness)*					1.0715	1.0682	1.0528	1.0006
					(0.0699)	(0.0696)	(0.0690)	(0.0655)
Visible message = 1 × *Personality similarity (Extraversion)*					1.3064**	1.3072**	1.3562***	1.4107***
					(0.1147)	(0.1150)	(0.1204)	(0.1265)
Visible message = 1 × *Personality similarity (Emotional range)*					1.0015	1.0015	0.9955	1.0303
					(0.0592)	(0.0593)	(0.0592)	(0.0615)
Visible message = 1 × *Personality similarity (Openness)*					1.1575	1.1674	1.1241	1.1395
					(0.1046)	(0.1055)	(0.1010)	(0.1020)
Sender/recipient popularity controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Message controls	No	No	No	No	No	Yes	Yes	Yes
Product controls	No	No	No	No	No	No	Yes	Yes
Additional Sender\Recipient controls	No	No	No	No	No	No	No	Yes
Log-likelihood	−23,726.8	−23,726.3	−23,646.8	−23,650.8	−30,360.2	−30,348.7	−30,325.8	−30,142.5
BIC	47,573.06	47,583.85	47,424.87	47,480.72	60,988.01	60,977.22	60,967.73	60,710.70
AIC	47,473.69	47,474.55	47,315.56	47,331.67	60,764.45	60,743.49	60,703.52	60,355.04
χ^2	3,580.761	3,581.907	3,740.889	3,732.779	4,138.573	4,161.527	4,207.497	4,573.983
P	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
N	152,730	152,730	152,730	152,730	191,355	191,355	191,355	191,355

Notes. Semiparametric survival analysis with Cox proportional hazards model on matched sample. The HRs represent the percent increase (HR > 1) or decrease (HR < 1) in postpurchase hazards associated with each attribute.

$^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

Table 9. Propensity Score Matching (Personality Characteristics)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Personality similarity (Agreeableness)	1.4186***		1.3494***	1.3046***	1.3954***	1.5044***
	(0.1242)		(0.1155)	(0.1010)	(0.1216)	(0.1229)
Personality similarity (Conscientiousness)	1.0039	1.1397*		1.0545	1.1509**	1.0542
	(0.0539)	(0.0599)		(0.0566)	(0.0575)	(0.0534)
Personality similarity (Extraversion)	0.9542	1.0830	0.9494		0.8261*	0.9883
	(0.0720)	(0.0798)	(0.0723)		(0.0637)	(0.0745)
Personality similarity (Emotional range)	1.0155	1.0661	1.1027*	0.9785		0.9598
	(0.0451)	(0.0504)	(0.0483)	(0.0446)		(0.0424)
Personality similarity (Openness)	0.9543	1.0930	1.0418	0.8775	0.9933	
	(0.0748)	(0.0834)	(0.0778)	(0.0662)	(0.0712)	
Visible message = 1 × *Personality similarity (Agreeableness)*	0.7858*		0.7831*	0.8726	0.7642**	0.7414**
	(0.0768)		(0.0758)	(0.0764)	(0.0751)	(0.0682)
Visible message = 1 × *Personality similarity (Conscientiousness)*	1.0528	0.9510		1.0081	0.9927	1.0081
	(0.0690)	(0.0615)		(0.0661)	(0.0577)	(0.0637)
Visible message = 1 × *Personality similarity (Extraversion)*	1.3562***	1.1435	1.3791***		1.5216***	1.3329**
	(0.1204)	(0.0978)	(0.1238)		(0.1380)	(0.1173)
Visible message = 1 × *Personality similarity (Emotional range)*	0.9955	0.9325	0.9738	1.0332		1.0373
	(0.0592)	(0.0579)	(0.0537)	(0.0625)		(0.0607)
Visible message = 1 × *Personality similarity (Openness)*	1.1241	0.9600	1.0521	1.1619	1.1238	
	(0.1010)	(0.0859)	(0.0919)	(0.1016)	(0.0950)	
Low Agreeableness (Sender) × *Low Agreeableness (Recipient)*		0.6338***				
		(0.0521)				
Low Agreeableness (Sender) × *High Agreeableness (Recipient)*		0.5853***				
		(0.0597)				
High Agreeableness (Sender) × *Low Agreeableness (Recipient)*		1.3064**				
		(0.1316)				
High Agreeableness (Sender) × *High Agreeableness (Recipient)*		1.1069				
		(0.1430)				
Low Conscientiousness (Sender) × *Low Conscientiousness (Recipient)*			0.6994*			
			(0.1176)			
Low Conscientiousness (Sender) × *High Conscientiousness (Recipient)*			1.0189			
			(0.1072)			
High Conscientiousness (Sender) × *Low Conscientiousness (Recipient)*			1.6768***			
			(0.2087)			
High Conscientiousness (Sender) × *High Conscientiousness (Recipient)*			1.8158***			
			(0.1613)			
Low Extraversion (Sender) × *Low Extraversion (Recipient)*				1.2293**		
				(0.0981)		
Low Extraversion (Sender) × *High Extraversion (Recipient)*				0.5872***		
				(0.0796)		
High Extraversion (Sender) × *Low Extraversion (Recipient)*				1.7258***		
				(0.2057)		
High Extraversion (Sender) × *High Extraversion (Recipient)*				0.9292		
				(0.1899)		
Low Emotional range (Sender) × *Low Emotional range (Recipient)*					1.6505***	
					(0.1314)	
Low Emotional range (Sender) × *High Emotional range (Recipient)*					1.0775	
					(0.1405)	
High Emotional range (Sender) × *Low Emotional range (Recipient)*					0.9285	
					(0.0998)	
High Emotional range (Sender) × *High Emotional range (Recipient)*					0.5493**	
					(0.1066)	
Low Openness (Sender) × *Low Openness (Recipient)*						0.9547
						(0.3474)

Table 9. (Contiuned)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Low Openness (*Sender*) ×*High Openness* (*Recipient*)						0.7840 (0.1442)
High Openness (*Sender*) ×*Low Openness* (*Recipient*)						1.7687*** (0.2819)
High Openness (*Sender*) ×*High Openness* (*Recipient*)						1.4236*** (0.1154)
Visible message = 1	1.2256** (0.0932)					
Sender/recipient controls	Yes	Yes	Yes	Yes	Yes	Yes
Message controls	Yes	Yes	Yes	Yes	Yes	Yes
Product controls	Yes	Yes	Yes	Yes	Yes	Yes
Log-likelihood	−30,325.8	−30,301.1	−30,272.9	−30,263.9	−30,199.0	−30,290.3
BIC	60,967.73	60,954.90	60,898.55	60,880.40	60,750.79	60,933.28
AIC	60,703.52	60,660.21	60,603.86	60,585.71	60,456.10	60,638.59
χ^2	4,207.497	4,256.812	4,313.163	4,331.310	4,460.924	4,278.431
p	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
N	191,355	191,355	191,355	191,355	191,355	191,355

Notes. Semiparametric survival analysis with Cox proportional hazards model on matched sample. The HRs represent the percent increase (HR > 1) or decrease (HR < 1) in postpurchase hazards associated with each attribute. The baseline case for Models 2–6 represents dyads in which the WOM message was not visible.

$^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

In addition to the aforementioned robustness checks that are reported in this section and the online appendix, we have conducted additional checks for which the results are not reported here because of space limitations. For instance, the results and findings remain qualitatively the same after removing the top 20% of users with the highest number of words posted on the social media platform. Similarly, the results remain qualitatively the same after removing 20% of users with the lowest number of words posted on the social media platform. Likewise, the findings remain the same using different samples. To address any potential remaining self-selection issues, we also further enhance the identification strategy with a Heckman selection model (Heckman 1979). The results remain qualitatively the same. Moreover, the results remain qualitatively the same employing alternative operationalizations for categorizing users to low-level and high-level groups for each personality characteristic. Other robustness checks include controlling for exogenous demand shocks based on Google Trends search volume, brand affinity, etc. To sum up, our identification strategy employs a quasi-experiment, latent variable models, and propensity score matching techniques as well as several controls for user similarity and homophily, network structure roles, user preferences, expertise, popularity, activity, strength of user relationships, message visibility and advocacy, as well as controls for products, product brands, product ads, etc., based on natural language processing, machine-learning, and deep-learning methods.

6. Discussion and Conclusions

6.1. Contributions and Theoretical Implications

Social media have transformed how consumers communicate and interact online and, consequently, how firms pursue their key marketing objectives. In particular, firms increasingly rely on leveraging electronic WOM in social media to attain their goals. At the same time, social media, thanks to their intrinsic characteristics, provide to both researchers and marketers the opportunity to directly observe the online communications and the related WOM instances (Godes and Mayzlin 2004, Trusov et al. 2009) and gain deeper insights into the users' characteristics that accentuate or attenuate the effectiveness of WOM. In this study, tapping into the idea that personality characteristics affect individual behaviors, we extend the set of users' characteristics that have been theorized as affecting the success of WOM instances (Chang 2004, Forman et al. 2008, v. Wangenheim and Bayón 2004) to include latent user characteristics, such as personality traits. In essence, this paper takes a significant step toward studying latent user characteristics that impact the effectiveness of WOM. This is the first paper to unveil that personality similarity and specific latent personality traits affect users' online purchase behavior and facilitate WOM. In particular, drawing on theories that are rooted in psychology and social sciences, we examine at a granular level how specific personality characteristics enable or constrain the effects of WOM messages in social media. Furthermore, this paper goes beyond individualistic views of factors that amplify the effectiveness of WOM and provides a more holistic view by examining the pairwise characteristics of both the source and recipient of WOM messages. To answer our research ques-

Table 10. Survival Analysis (Personality Characteristics with Additional Deep-Learning Controls)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
User similarity	1.0980***	1.0990***	1.0970***	1.0949***	1.2969***	1.2949***	1.2970***	1.2803***
	(0.0044)	(0.0045)	(0.0045)	(0.0045)	(0.0218)	(0.0218)	(0.0215)	(0.0212)
Reciprocal relationship	7.1589***	7.1701***	6.7201***	6.7682***	7.0216***	6.9979***	7.0329***	6.4343***
	(0.3505)	(0.3512)	(0.3293)	(0.3316)	(0.3219)	(0.3209)	(0.3237)	(0.3041)
Number of peer-to-peer interactions	1.0008	1.0008	1.0006	1.0006	1.0007	1.0007	1.0008	1.0005
	(0.0008)	(0.0008)	(0.0009)	(0.0009)	(0.0008)	(0.0008)	(0.0008)	(0.0009)
Sentiment of message	1.5962***	1.5955***	1.5230***	1.5764***	1.6474***	1.6227***	1.5621***	1.4451***
	(0.1066)	(0.1066)	(0.1006)	(0.1051)	(0.1038)	(0.1025)	(0.0989)	(0.0948)
Personalized message	1.1075***	1.1079***	1.1071***	1.1048***	1.0546***	1.0527***	1.0473***	1.0371***
	(0.0052)	(0.0052)	(0.0052)	(0.0052)	(0.0039)	(0.0039)	(0.0052)	(0.0053)
User expertise (*Sender*)	1.2580***	1.2551***	1.2667***	1.2743***	1.2040***	1.1956***	1.2171***	1.1982***
	(0.0293)	(0.0294)	(0.0297)	(0.0299)	(0.0257)	(0.0255)	(0.0262)	(0.0266)
User leadership (*Sender*)	1.0139***	1.0139***	1.0111***	1.0109***	1.0098***	1.0100***	1.0100***	1.0046*
	(0.0017)	(0.0017)	(0.0017)	(0.0017)	(0.0018)	(0.0018)	(0.0018)	(0.0018)
Personality similarity (*Main personality type*)		0.9484						
		(0.0451)						
Personality similarity			1.4718***					
			(0.0482)					
Personality similarity (*Agreeableness*)				1.1080*	1.4614***	1.4598***	1.4430***	1.4375***
				(0.0495)	(0.1141)	(0.1141)	(0.1136)	(0.1131)
Personality similarity (*Conscientiousness*)				1.0677	0.9693	0.9742	0.9863	1.0105
				(0.0401)	(0.0474)	(0.0476)	(0.0483)	(0.0492)
Personality similarity (*Extraversion*)				1.3041***	0.9764	0.9652	0.9295	0.9028
				(0.0616)	(0.0658)	(0.0653)	(0.0636)	(0.0622)
Personality similarity (*Emotional range*)				1.0386	1.0077	1.0048	1.0155	1.0141
				(0.0417)	(0.0400)	(0.0399)	(0.0405)	(0.0407)
Personality similarity (*Openness*)				1.0487	0.9145	0.9114	0.9374	0.9237
				(0.0461)	(0.0658)	(0.0655)	(0.0669)	(0.0656)
Visible message = 1					1.3154***	1.2384**	1.1851*	1.4418***
					(0.0950)	(0.0906)	(0.0870)	(0.1079)
Visible message = 1 × *Personality similarity* (*Agreeableness*)					0.7633**	0.7737**	0.7752**	0.7661**
					(0.0684)	(0.0693)	(0.0697)	(0.0692)
Visible message = 1 × *Personality similarity* (*Conscientiousness*)					1.0869	1.0844	1.0710	1.0188
					(0.0668)	(0.0665)	(0.0661)	(0.0629)
Visible message = 1 × *Personality similarity* (*Extraversion*)					1.3469***	1.3465***	1.3917***	1.4450***
					(0.1104)	(0.1106)	(0.1152)	(0.1210)
Visible message = 1 × *Personality similarity* (*Emotional range*)					0.9980	0.9979	0.9917	1.0268
					(0.0559)	(0.0559)	(0.0558)	(0.0579)
Visible message = 1 × *Personality similarity* (*Openness*)					1.1840*	1.1937*	1.1508	1.1632
					(0.0997)	(0.1005)	(0.0965)	(0.0972)
Sender/recipient popularity controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sender/recipient latent characteristics controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Message controls	No	No	No	No	No	Yes	Yes	Yes
Product controls	No	No	No	No	No	No	Yes	Yes
Additional sender/recipient controls	No	No	No	No	No	No	No	Yes
Log-likelihood	−23,679.1	−23,678.5	−23,601.3	−23,605.3	−31,774.7	−31,761.8	−31,738.2	−31,543.9
BIC	47,489.45	47,500.13	47,345.76	47,401.48	63,829.96	63,816.37	63,805.8	63,526.98
AIC	47,380.24	47,380.98	47,226.62	47,242.62	63,595.42	63,571.63	63,530.46	63,159.87
χ^2	3,571.856	3,573.108	3,727.475	3,719.471	4,108.19	4,133.979	4,181.143	4,569.737
p	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
N	151,569	151,569	151,569	151,569	198,290	198,290	198,290	198,290

Notes. Semiparametric survival analysis with Cox proportional hazards model. The presented econometric specifications include additional controls for latent user homophily and network structure roles based on deep-learning methods. The HRs represent the percent increase (HR > 1) or decrease (HR < 1) in postpurchase hazards associated with each attribute. $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

Table 11. Survival Analysis (Personality Characteristics with Additional Deep-Learning Controls)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Personality similarity (Agreeableness)	1.4430*** (0.1136)		1.3815*** (0.1069)	1.3245*** (0.0935)	1.4285*** (0.1121)	1.5230*** (0.1117)
Personality similarity (Conscientiousness)	0.9863 (0.0483)	1.1017* (0.0525)		1.0368 (0.0507)	1.1324** (0.0523)	1.0327 (0.0477)
Personality similarity (Extraversion)	0.9295 (0.0636)	1.0512 (0.0708)	0.9230 (0.0639)		0.8089** (0.0566)	0.9530 (0.0654)
Personality similarity (Emotional range)	1.0155 (0.0405)	1.0710 (0.0455)	1.0858* (0.0433)	0.9799 (0.0400)		0.9569 (0.0379)
Personality similarity (Openness)	0.9374 (0.0669)	1.0956 (0.0764)	1.0129 (0.0690)	0.8592* (0.0594)	0.9729 (0.0636)	
Visible message = 1 × *Personality similarity (Agreeableness)*	0.7752** (0.0697)		0.7681** (0.0687)	0.8593 (0.0701)	0.7475** (0.0678)	0.7331*** (0.0621)
Visible message = 1 × *Personality similarity (Conscientiousness)*	1.0710 (0.0661)	0.9840 (0.0598)		1.0251 (0.0633)	1.0132 (0.0556)	1.0287 (0.0614)
Visible message = 1 × *Personality similarity (Extraversion)*	1.3917*** (0.1152)	1.1720* (0.0941)	1.4185*** (0.1191)		1.5525*** (0.1315)	1.3800*** (0.1135)
Visible message = 1 × *Personality similarity (Emotional range)*	0.9917 (0.0558)	0.9231 (0.0541)	0.9856 (0.0514)	1.0302 (0.0588)		1.0344 (0.0572)
Visible message = 1 × *Personality similarity (Openness)*	1.1508 (0.0965)	0.9649 (0.0810)	1.0875 (0.0890)	1.1893* (0.0977)	1.1572 (0.0918)	
Low Agreeableness (Sender) × *Low Agreeableness (Recipient)*		0.6158*** (0.0491)				
Low Agreeableness (Sender) × *High Agreeableness (Recipient)*		0.5599*** (0.0556)				
High Agreeableness (Sender) × *Low Agreeableness (Recipient)*		1.3104** (0.1302)				
High Agreeableness (Sender) × *High Agreeableness (Recipient)*		1.0929 (0.1387)				
Low Conscientiousness (Sender) × *Low Conscientiousness (Recipient)*			0.6936* (0.1158)			
Low Conscientiousness (Sender) × *High Conscientiousness (Recipient)*			0.9999 (0.1031)			
High Conscientiousness (Sender) × *Low Conscientiousness (Recipient)*			1.6257*** (0.1986)			
High Conscientiousness (Sender) × *High Conscientiousness (Recipient)*			1.7309*** (0.1489)			
Low Extraversion (Sender) × *Low Extraversion (Recipient)*				1.1883* (0.0912)		
Low Extraversion (Sender) × *High Extraversion (Recipient)*				0.5756*** (0.0769)		
High Extraversion (Sender) × *Low Extraversion (Recipient)*				1.7263*** (0.2032)		
High Extraversion (Sender) × *High Extraversion (Recipient)*				0.9625 (0.1954)		
Low Emotional range (Sender) × *Low Emotional range (Recipient)*					1.6211*** (0.1252)	
Low Emotional range (Sender) × *High Emotional range (Recipient)*					1.0785 (0.1390)	
High Emotional range (Sender) × *Low Emotional range (Recipient)*					0.9211 (0.0969)	
High Emotional range (Sender) × *High Emotional range (Recipient)*					0.5442** (0.1048)	
Low Openness (Sender) × *Low Openness (Recipient)*						0.9808 (0.3567)

Table 11. (Continued)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Low Openness (*Sender*) × *High Openness* (*Recipient*)						0.7702 (0.1409)
High Openness (*Sender*) × *Low Openness* (*Recipient*)						1.6991*** (0.2682)
High Openness (*Sender*) × *High Openness* (*Recipient*)						1.3624*** (0.1064)
Visible message = 1	1.1851* (0.0870)					
Sender/recipient popularity controls	Yes	Yes	Yes	Yes	Yes	Yes
Sender/recipient latent characteristics controls	Yes	Yes	Yes	Yes	Yes	Yes
Message controls	Yes	Yes	Yes	Yes	Yes	Yes
Product controls	Yes	Yes	Yes	Yes	Yes	Yes
Log-likelihood	−31,738.2	−31,702.5	−31,684.7	−31,668.5	−31,599.2	−31,700.7
BIC	63,805.8	63,770.96	63,735.36	63,702.87	63,564.26	63,767.31
AIC	63,530.46	63,465.03	63,429.43	63,396.95	63,258.34	63,461.38
χ^2	4,181.143	4,252.575	4,288.173	4,320.658	4,459.267	4,256.224
p	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
N	198,290	198,290	198,290	198,290	198,290	198,290

Notes. Semiparametric survival analysis with Cox proportional hazards model. The presented econometric specifications include additional controls for latent user homophily and network structure roles based on deep-learning methods. The HRs represent the percent increase (HR > 1) or decrease (HR < 1) in postpurchase hazards associated with each attribute. The baseline case for Models 2–6 represents dyads in which the WOM message was not visible.
$^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

tions, we employ big data and machine-learning techniques to extract information from a vast amount of unstructured textual content generated by social media users and leverage a novel quasi-experimental research design combined with econometric techniques allowing us to disentangle the effects of personality on WOM from correlated user behaviors and homophily.

We find that higher levels of similarity on levels of extraversion and openness as well as lower levels of similarity on agreeableness increase the effectiveness of WOM. Besides this, we find that WOM originating from users who exhibit high levels of agreeableness, conscientiousness, and openness is more likely to be more effective, whereas for users with low levels of conscientiousness or agreeableness, the opposite effect is more likely. In addition, introvert users are more susceptible to WOM effects, in contrast to extrovert users. Finally, users with low levels of emotional range impact similar users through WOM, whereas for high levels of emotional range, increased similarity usually has the opposite effect. The examined effects are also of significant economic importance as, for instance, a WOM message from a similar user in terms of personality, rather than a dissimilar user, increases the likelihood of a purchase by 47.58%. Similarly, a WOM message from an extrovert user to an introvert peer increases the likelihood of a subsequent purchase by 71.28%. The corresponding models also exhibit very good predictive ability based on out-of-sample evaluation results.

6.2. Managerial Implications

This study provides several insights leading to actionable strategies for managers who would like to effectively utilize and engineer WOM. First, we find that the effects of WOM can be accentuated when the source of WOM messages is characterized by specific personality traits. Hence, marketers might be able to increase sales and spur buzz around their brands by taking actions to encourage social media users characterized by distinct personality traits and attributes to generate or disseminate positive WOM messages. Second, given the importance of the characteristics of the source of WOM messages and the relationship between the sender and the recipient, our analysis also provides valuable insights to firms and marketers that would be interested into associating their brands with certain characteristics and attributes as well as fostering particular perceptions that might be more appealing to specific types of personalities of users in social media platforms. Besides this, our analysis demonstrates the value of directly observing the WOM instances and extracting knowledge from analyzing granular-level data. Additionally, in this study, we demonstrate to social firms the ability to conduct such analyses leveraging machine-learning and text-mining algorithms as well as the value of unstructured user-generated content in social media.

Finally, our results have important implications for other parties of the social media ecosystem as well. In particular, the conducted analysis has significant managerial implications regarding the monetization of social media and user-generated content in such settings. Social media companies, including microblogging platforms, are increasingly moving toward a model of sponsored posts (tweets) in which advertisers can bid based on various targeting criteria (Ghose et al. 2013). In particular, the asymmetric WOM effects across different types of personalities offer actionable tactical strategies as they suggest that social media companies can charge different prices to advertisers for sponsored messages based on users' personalities. Similarly, social media platforms can use the latent personality characteristics of the social media users to curate and rank user-generated content more effectively and drive engagement in their platforms. In addition, our results show that latent characteristics of the users can also be used to better predict the diffusion of information and products in social media. Although we showed the aforementioned effects in the context of a microblogging platform, the implications are potentially broader, as social media users typically engage actively with multiple platforms (GlobalWebIndex 2015) and there are increasingly many similarities among the various social media platforms.

6.3. Limitations

While this paper takes important steps toward studying factors that increase the effectiveness of WOM, we acknowledge that there are several limitations in our analysis mostly emerging from data availability issues. In this study, we consider only actual purchases taking place within a popular social media platform. A broader set of user behaviors, such as whether given users searched for additional information about a specific product, implying that they are sufficiently interested in making a purchase, could be examined by future research. Unfortunately, search logs and other information about similar user behaviors are not available to us. Moreover, we do not examine subsequent cascades of influence in the social network beyond the direct peers of the initial disseminators of WOM (Susarla et al. 2016). Additionally, the employed social media context allowed us to leverage rich microdata that enabled machine-learning and text-mining algorithms to extract useful knowledge regarding the users' characteristics. However, the offline world is unlikely to provide such convenient context, and hence, our analysis provides actionable insights mostly for electronic WOM, rather than offline WOM. In addition, because of data availability limitations, we do not observe private communication among the users or communication in other platforms. Similarly, we do not observe communication among the users in the offline world. Despite any limitations, our contribution may be widely relevant to managers while also seeding a number of new directions for future research. Our hope is that these limitations are viewed not as a liability but as a path toward future research that extends our research questions while strengthening the relevant theory and empirical evidence. To the extent that user personality characteristics affect WOM effectiveness, the increasing size of social media may have profound implications for the future direction of electronic and social commerce.

Endnotes

[1] Lichtenthal and Tellefsen (2001) neither disentangle the effect of personality similarity from other confounders, as they examine the composite similarity in internal characteristics (e.g., combining together personality, education, perceptions, attitudes, political views, values), nor capture personality similarity based on a personality model grounded in psychological theories (Crosby et al. 1990).

[2] We have excluded from our analysis any ineligible user as well as all user accounts that do not have any followers or were created after the service was launched. In the same fashion, we have filtered out unconfirmed and ineligible attempts to make transactions, such as messages that were posted after the expiration date of the product offerings.

[3] If a user accesses the announcement of a product offering through the profile of the social commerce provider, then the handle of the social commerce provider (e.g., "@AmericanExpress") is prepopulated in the purchasing message of that user. To make the purchase, the user can click on the response field and add the appropriate hashtag. Depending on whether the click is recorded to the left or the right of the prepopulated handle (i.e., username following the "@" symbol), the cursor will be placed at the corresponding side of the handle (see Figures A3(a) and A3(b) in the online appendix). Hence, the position of the handle (whether it will be at the beginning of the message or not) is also affected by this exogenous design feature of the social network platform. In other words, the visibility of a WOM message depends on whether another account is mentioned in the message, the position of the referred username corresponding to the mentioned account, and whether each follower of the sender of the WOM follows the mentioned account, while the visibility is also affected by the randomness in the initial position of the cursor in the WOM message and whether the username of the author of a previous message was prepopulated. Overall, the visibility of the message depends on multiple factors and not just a single factor that could be controlled by the sender.

[4] We also employed an alternative sentiment analysis approach based on assigning sentiment scores to a small number of messages and then building a regression model using machine-learning techniques. The estimated scores of the different methods are very similar and the findings remain the same.

References

Adamopoulos P, Todri V (2014) Social media analytics: The effectiveness of promotional events on brand user base in social media. *Proc. 35th Internat. Conf. Inform. Systems* (*ICIS* '14). http://aisel.aisnet.org/icis2014/proceedings/SocialMedia/8/.

Adamopoulos P, Todri V (2015a) Personality-based recommendations: Evidence from Amazon.com. *Proc. 9th ACM Conf. Recommender Systems* (*RecSys* '15) (ACM, New York), 32–33.

Adamopoulos P, Todri V (2015b) The effectiveness of marketing strategies in social media. *Proc. 21st ACM SIGKDD Internat. Conf. Knowledge Discovery Data Mining* (*KDD* '15) (ACM, New York), 1641–1650.

Adamopoulos P, Tuzhilin A (2015) The business value of recommendations: A privacy-preserving econometric analysis. *Proc. 36th Internat. Conf. Inform. Systems* (*ICIS* '15). http://aisel.aisnet.org/icis2015/proceedings/GeneralIS/12/.

AddShoppers (2013) Social commerce metrics among social network sites and email users worldwide: Q3 2013. Report, AddShoppers, Charlotte, NC.

AlchemyApi (2012) Orchestr8, LLC. Accessed October 15, 2017, http://www.alchemyapi.com/.

Allport GW (1961) *Pattern and Growth in Personality* (Holt, Reinhart & Winston, Oxford, UK).

Aral S, Walker D (2012) Identifying influential and susceptible members of social networks. *Science* 337(6092):337–341.

Arun R, Suresh V, Madhavan CV, Murthy MN (2010) On finding the natural number of topics with latent Dirichlet allocation: Some observations. Zaki MJ, Yu JX, Ravindran B, Pudi V, eds. *Advances in Knowledge Discovery and Data Mining*, Part II (Springer-Verlag, Berlin), 391–402.

AYTM (2013) March 2013 AYTM market research study. Report, AYTM Market Research, San Francisco.

Back MD, Stopfer JM, Vazire S, Gaddis S, Schmukle SC, Egloff B, Gosling SD (2010) Facebook profiles reflect actual personality, not self-idealization. *Psych. Sci.* 21(3):372–374.

Bapna R, Umyarov A (2015) Do your online friends make you pay? A randomized field experiment on peer influence in online social networks. *Management Sci.* 61(8):1902–1920.

Barrick MR, Mount MK (1991) The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psych.* 44(1):1–26.

Barrick MR, Mount MK, Judge TA (2001) Personality and performance at the beginning of the new millennium: What do we know and where do we go next? *Internat. J. Selection Assessment* 9(1–2):9–30.

Barrick MR, Mount MK, Strauss JP (1993) Conscientiousness and performance of sales representatives: Test of the mediating effects of goal setting. *J. Appl. Psych.* 78(5):715–722.

Baumgartner H (2002) Toward a personology of the consumer. *J. Consumer Res.* 29(2):286–292.

Blei DM, Ng AY, Jordan MI (2003) Latent Dirichlet allocation. *J. Machine Learn. Res.* 3(4–5):993–1022.

Botwin MD, Buss DM, Shackelford TK (1997) Personality and mate preferences: Five factors in mate selection and marital satisfaction. *J. Personality* 65(1):107–136.

Byrne D, Griffitt W (1969) Similarity and awareness of similarity of personality characteristics as determinants of attraction. *J. Experiment. Res. Personality* 3(3):179–186.

Byrne D, London O, Reeves K (1968) The effects of physical attractiveness, sex, and attitude similarity on interpersonal attraction. *J. Personality* 36(2):259–271.

Chaiken S (1980) Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *J. Personality Soc. Psych.* 39(5):752–766.

Chaiken S (1987) The heuristic model of persuasion. Zanna MP, Olson JM, Herman CP, eds. *Soc. Influence: Ontario Sympos.*, Vol. 5 (Lawrence Erlbaum Associates, Hillsdale, NJ), 3–39.

Chaiken S, Maheswaran D (1994) Heuristic processing can bias systematic processing: Effects of source credibility, argument ambiguity, and task importance on attitude judgment. *J. Personality Soc. Psych.* 66(3):460–473.

Chang C (2004) Country of origin as a heuristic cue: The effects of message ambiguity and product involvement. *Media Psych.* 6(2): 169–192.

Chen J, Haber E, Kang R, Hsieh G, Mahmud J (2015) Making use of derived personality: The case of social media ad targeting. *Proc. 9th Internat. AAAI Conf. Web Soc. Media* (AAAI Press, Palo Alto, CA), 51–60.

Cohen GL (2003) Party over policy: The dominating impact of group influence on political beliefs. *J. Personality Soc. Psych.* 85(5): 808–822.

Costa PT, McCrae RR (1992) Four ways five factors are basic. *Personality Individual Differences* 13(6):653–665.

Crosby LA, Evans KR, Cowles D (1990) Relationship quality in services selling: An interpersonal influence perspective. *J. Marketing* 54(3):68–81.

Dellarocas C (2006) Strategic manipulation of Internet opinion forums: Implications for consumers and firms. *Management Sci.* 52(10):1577–1593.

Devaraj S, Easley RF, Crant JM (2008) How does personality matter? Relating the five-factor model to technology acceptance and use. *Inform. Systems Res.* 19(1):93–105.

Digman JM (1990) Personality structure: Emergence of the five-factor model. *Annual Rev. Psych.* 41(1):417–440.

Fast LA, Funder DC (2008) Personality as manifest in word use: Correlations with self-report, acquaintance report, and behavior. *J. Personality Soc. Psych.* 94(2):334–346.

Feldman SP, Spencer MC (1965) *The Effect of Personal Influence on the Selection of Consumer Services* (Center for Regional Studies, Lawrence, KS).

Festinger L (1954) A theory of social comparison processes. *Human Relations* 7(2):117–140.

Forman C, Ghose A, Wiesenfeld B (2008) Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets. *Inform. Systems Res.* 19(3):291–313.

Forrester (2013) North American Consumer Technographics online benchmark survey. Report, Forrester Research, Cambridge, MA.

Gardner BO, Titcomb C, Cramer RJ, Stroud CH, Bate BP (2013) Perceived personality similarity and perceptions of expert testimony. *J. Individual Differences* 34(4):185–192.

Ghose A (2017) *Tap: Unlocking the Mobile Economy* (MIT Press, Cambridge, MA).

Ghose A, Todri-Adamopoulos V (2016) Towards a digital attribution model: Measuring the impact of display advertising on online consumer behavior. *MIS Quart.* 40(4):889–910.

Ghose A, Goldfarb A, Han SP (2013) How is the mobile Internet different? Search costs and local activities. *Inform. Systems Res.* 24(3):613–631.

GlobalWebIndex (2015) GWI social report. Report, GlobalWebIndex, London.

Godes D, Mayzlin D (2004) Using online conversations to study word-of-mouth communication. *Marketing Sci.* 23(4):545–560.

Godes D, Mayzlin D (2009) Firm-created word-of-mouth communication: Evidence from a field test. *Marketing Sci.* 28(4):721–739.

Golbeck J, Robles C, Edmondson M, Turner K (2011) Predicting personality from Twitter. *2011 IEEE Third Internat. Conf. Privacy, Security, Risk Trust/IEEE Internat. Conf. Soc. Comput. (PASSAT/SocialCom 2011)* (IEEE Computer Society, Los Alamitos, CA), 149–156.

Goldberg LR (1981) Language and individual differences: The search for universals in personality lexicons. *Rev. Personality Soc. Psych.* 2(1):141–165.

Grover A, Leskovec J (2016) node2vec: Scalable feature learning for networks. *Proc. 22nd ACM SIGKDD Internat. Conf. Knowledge Discovery Data Mining* (ACM, New York), 855–864.

Hass RG (1981) Effects of source characteristics on cognitive responses and persuasion. Petty RE, Ostrom TM, Brock TC, eds. *Cognitive Responses in Persuasion* (Psychology Press, New York), 141–172.

Heckman JJ (1979) Sample selection bias as a specification error. *Econometrica* 47(1):153–161.

Hirsh JB, Peterson JB (2009) Personality and language use in self-narratives. *J. Res. Personality* 43(3):524–527.

Hoffman M, Bach FR, Blei DM (2010) Online learning for latent Dirichlet allocation. Lafferty JD, Williams CKI, Shawe-Taylor J, Zemel RS, Culotta A, eds. *Advances in Neural Information Processing Systems*, Vol. 23 (Curran Associates, Red Hook, NY), 856–864.

Horton RL (1979) Some relationships between personality and consumer decision-making. *J. Marketing Res.* 16(2):233–246.

Hu R, Pu P (2011) Enhancing collaborative filtering systems with personality information. *Proc. Fifth ACM Conf. Recommender Systems* (ACM, New York), 197–204.

Jensen-Campbell LA, Graziano WG (2001) Agreeableness as a moderator of interpersonal conflict. *J. Personality* 69(2):323–362.

John OP, Srivastava S (1999) The Big Five trait taxonomy: History, measurement, and theoretical perspectives. Pervin LA, John OP, eds. *Handbook of Personality: Theory and Research*, Vol. 2 (Guilford Press, New York), 102–138.

Judge TA, Bono JE (2000) Five-factor model of personality and transformational leadership. *J. Appl. Psych.* 85(5):751–765.

Kalbfleisch JD, Prentice RL (2011) *The Statistical Analysis of Failure Time Data* (John Wiley & Sons, Hoboken, NJ).
Kang Y-S, Herr PM (2006) Beauty and the beholder: Toward an integrative model of communication source effects. *J. Consumer Res.* 33(1):123–130.
Karney BR, Bradbury TN (1997) Neuroticism, marital interaction, and the trajectory of marital satisfaction. *J. Personality Soc. Psych.* 72(5):1075–1092.
Kelman HC (1961) Processes of opinion change. *Public Opinion Quart.* 25(1):57–78.
Klein KJ, Lim B-C, Saltz JL, Mayer DM (2004) How do they get there? An examination of the antecedents of centrality in team networks. *Acad. Management J.* 47(6):952–963.
Komarraju M, Karau SJ (2005) The relationship between the Big Five personality traits and academic motivation. *Personality Individual Differences* 39(3):557–567.
KPMG (2013) 2013 Retail industry outlook survey. Report, KPMG, Amstelveen, Netherlands.
Kullback S, Leibler RA (1951) On information and sufficiency. *Ann. Math. Statist.* 22(1):79–86.
Lichtenthal JD, Tellefsen T (2001) Toward a theory of business buyer-seller similarity. *J. Personal Selling Sales Management* 21(1):1–14.
Lin C-F (2002) Segmenting customer brand preference: Demographic or psychographic. *J. Product Brand Management* 11(4): 249–268.
Lin DY, Wei L-J (1989) The robust inference for the Cox proportional hazards model. *J. Amer. Statist. Assoc.* 84(408):1074–1078.
Mackie DM, Worth LT, Asuncion AG (1990) Processing of persuasive in-group messages. *J. Personality Soc. Psych.* 58(5):812–822.
Mahmud J, Zhou MX, Megiddo N, Nichols J, Drews C (2013) Recommending targeted strangers from whom to solicit information on social media. *Proc. 2013 Internat. Conf. Intelligent User Interfaces* (*IUI* '13) (ACM, New York), 37–48.
Mairesse F, Walker M (2006) Words mark the nerds: Computational models of personality recognition through language. *Proc. 28th Annual Conf. Cognitive Sci. Soc.* (Cognitive Science Society, Austin, TX), 543–548.
McCrae RR, Costa PT Jr (1987) Validation of the five-factor model of personality across instruments and observers. *J. Personality Soc. Psych.* 52(1):81–90.
McCrae RR, Costa PT (1991) Adding Liebe und Arbeit: The full five-factor model and well-being. *Personality Soc. Psych. Bull.* 17(2): 227–232.
McFadden D (1973) Conditional logit analysis of qualitative choice behavior. Zarembka P, ed. *Frontiers in Econometrics* (Academic Press, New York), 105–142.
Menon T, Blount S (2003) The messenger bias: A relational model of knowledge valuation. *Res. Organ. Behav.* 25:137–186.
Nielsen (2013) Global trust in advertising and brand messages. Report, Nielsen, New York.
Norman WT (1963) Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. *J. Abnormal Soc. Psych.* 66(6):574–583.
Ogunlade JO (1979) Personality characteristics related to susceptibility to behavioral contagion. *Soc. Behav. Personality* 7(2):205–208.
Oreg S, Sverdlik N (2014) Source personality and persuasiveness: Big Five predispositions to being persuasive and the role of message involvement. *J. Personality* 82(3):250–264.
Owoputi O, O'Connor B, Dyer C, Gimpel K, Schneider N, Smith NA (2013) Improved part-of-speech tagging for online conversational text with word clusters. *2013 Conf. North Amer. Chap. Assoc. Comput. Linguistics: Human Language Tech.* (Association for Computational Linguistics, Stroudsburg, PA), 380–390.
Parish L, Eysenck HJ, Eysenck SGB (1965) The Eysenck personality inventory. *British J. Ed. Stud.* 14(1):140.
Pennebaker JW, Chung CK, Ireland M, Gonzales A, Booth RJ (2007) The development and psychometric properties of LIWC2007. White paper, LIWC.net, Austin, TX.
Perozzi B, Al-Rfou R, Skiena S (2014) DeepWalk: Online learning of social representations. *Proc. 20th ACM SIGKDD Internat. Conf. Knowledge Discovery Data Mining* (ACM, New York), 701–710.
Pornpitakpan C (2004) The persuasiveness of source credibility: A critical review of five decades' evidence. *J. Appl. Soc. Psych.* 34(2):243–281.
Rentfrow PJ, Gosling SD (2003) The do re mi's of everyday life: The structure and personality correlates of music preferences. *J. Personality Soc. Psych.* 84(6):1236–1256.
Shaver PR, Brennan KA (1992) Attachment styles and the "Big Five" personality traits: Their connections with each other and with romantic relationship outcomes. *Personality Soc. Psych. Bull.* 18(5):536–545.
Sigurdsson JF (1991) Computer experience, attitudes toward computers and personality characteristics in psychology undergraduates. *Personality Individual Differences* 12(6):617–624.
Simpson EM, Snuggs T, Christiansen T, Simples KE (2000) Race, homophily, and purchase intentions and the black consumer. *Psych. Marketing* 17(10):877–889.
Singh R, Ho SY (2000) Attitudes and attraction: A new test of the attraction, repulsion and similarity-dissimilarity asymmetry hypotheses. *British J. Soc. Psych.* 39(2):197–211.
Susarla A, Oh J-H, Tan Y (2016) Influentials, imitables, or susceptibles? Virality and word-of-mouth conversations in online social networks. *J. Management Inform. Systems* 33(1):139–170.
Tan DTY, Singh R (1995) Attitudes and attraction: A developmental study of the similarity-attraction and dissimilarity-repulsion hypotheses. *Personality Soc. Psych. Bull.* 21(9):975–986.
Tausczik YR, Pennebaker JW (2009) The psychological meaning of words: LIWC and computerized text analysis methods. *J. Language Soc. Psych.* 29(1):24–54.
Todri V, Adamopoulos P (2014) Social commerce: An empirical examination of the antecedents and consequences of commerce in social network platforms. *Proc. 35th Internat. Conf. Inform. Systems (ICIS'14)*. http://aisel.aisnet.org/icis2014/proceedings/EBusiness/54/.
Trusov M, Bucklin RE, Pauwels K (2009) Effects of word-of-mouth versus traditional marketing: Findings from an Internet social networking site. *J. Marketing* 73(5):90–102.
Tupes EC, Christal RE (1992) Recurrent personality factors based on trait ratings. *J. Personality* 60(2):225–251.
v. Wangenheim F, Bayón T (2004) The effect of word of mouth on services switching: Measurement and moderating variables. *Eur. J. Marketing* 38(9/10):1173–1185.
Yarkoni T (2010) Personality in 100,000 words: A large-scale analysis of personality and word use among bloggers. *J. Res. Personality* 44(3):363–373.