Social Commerce: An Empirical Examination of the Antecedents and Consequences of Commerce in Social Network Platforms

Completed Research Paper

Vilma Todri
Department of Information, Operations, and Management Sciences
Leonard N. Stern School of Business
New York University
vtodri@stern.nyu.edu

Panagiotis Adamopoulos
Department of Information, Operations, and Management Sciences
Leonard N. Stern School of Business
New York University
padamopo@stern.nyu.edu

Abstract

This paper studies a pioneering venture of integrating e-business with social network platforms and seeks to understand the antecedents and consequences of "social commerce". In particular, we conduct an econometric analysis examining how the characteristics of the users and their social networks affect their decision to participate in this novel service. Based on the empirical results, we find that the social neighbors of the users and their economic behavior, the brand loyalty of the users, and their familiarity with the platform have significant effects on the likelihood of social purchases. Additionally, we build predictive models in order to both identify the effective disseminators of information and discover their distinguishing characteristics. Finally, we both contribute to the related literature, discovering new rich findings, and provide actionable insights with major implications for brands and marketers who would like to generate direct sales on social network platforms and orchestrate word-of-mouth.

Keywords: E-Business, Social Commerce, Econometric Analysis, Predictive Modeling, Text-Mining, Social Media, Online Social Networks, Word-of-Mouth

Introduction

Businesses respond quickly to the advancements of Information Technology (IT) and the unique tools such improvements offer. Some IT developments enhance existing operations and products while others are more disruptive as they transform existing products and business models or create new ones and contribute to a paradigm shift in the industry. Social media, in particular, represent one of the most transformative impacts of IT on business (Aral et al. 2013), as they drastically revolutionize the way consumers and firms interact. As the percentage of adults using social media has surged significantly over the last years and consumers spend an increasing amount of their time online, companies invest a growing amount of their marketing budget towards online and social media advertising. Hence, companies nowadays increasingly compete in the social media space for consumers' attention and engagement with their brand. However, effective social media advertising has also recently moved beyond the one-way messaging, harnessing the social connections that exist in the platforms (Heath et al. 2013). As marketers recognize the strategic importance of social media, they invent new ways to capitalize on the unique opportunities social media generate in order to connect and engage with their customers and tap into interpersonal communication as an effective means of non-traditional advertising.

Acknowledging that consumers' choices can be vastly influenced by the online presence and behavior of their peers, the emergence of social media has enabled unprecedented opportunities for marketers to leverage consumers' social networks in order to promote and amplify the firms' marketing messages. Incentivizing referral systems, nurturing positive online word-of-mouth (WOM), and spurring the creation of online forums and communities (Dellarocas 2006) are some quintessential and effective means of utilizing social links among customers in order to achieve key marketing objectives. In the same spirit, various strategies that integrate social media and commerce have emerged during the last years. For instance, the idea of a consumer making a purchase and leading additional sales through sharing that purchase via social network platforms is known and well established; e-commerce websites, such as Amazon, have seen much success as they promoted post-purchase social shares (Owned it Ltd 2013). Moving further beyond traditional modes of commerce and driving innovation in e-business, American Express Company in partnership with the social network platform of Twitter, an online social networking and microblogging service, and various participating retail brands introduced a novel way of interconnecting commerce and social networks by unveiling the ability for its customers to buy products directly through the social platform while automatically spreading the word about these products. As retailers' attitude towards social media shifts from being considered as a marketing channel to being perceived as a novel direct sales channel, we need to better understand such a pioneering social commerce model along with the opportunities and challenges that emerge.

Social commerce, or simply "s-commerce", is a representative and promising example of leveraging the social connections between users to generate effective leads for businesses. However, the boundaries of what constitutes social commerce nowadays are at best ambiguous. The term "social commerce" was first introduced in 2005 by Yahoo! (Rubel 2006) and initially referred to a feature that allowed users to review products. During the next years, the scope of this term expanded towards various directions incorporating peer recommendations, shared shopping lists, product referrals, coupon sharing, team-buying (i.e., a group of customers gathering together in order to bargain with merchants), "deal-of-the-day" websites (i.e., e-commerce websites where a minimum number of purchases should be reached in order an offer to be activated), network-based marketing as well as integration of online merchandisers with social network platforms, and firm-controlled online communities (Wang and Zhang 2012). One of the first innovative services that leveraged social network data, the easiness of online transactions, and social connections in order to enable this new type of social commerce was the Facebook Gifts service. Recently, Amazon launched a similar social commerce venture according to which consumers can use a specific keyword when replying to a message on Twitter containing an Amazon product link in order to add the corresponding product to their electronic shopping cart (Amazon.com 2014).

Building on concepts related to the previous social commerce ventures but also exhibiting some unique features, Amex in partnership with Twitter, one of the fastest growing social networks on the Internet, introduced a novel social commerce service based on which a user can purchase a product from a participating brand by sending a short 140-character text message, called "tweet", with a designated keyword (i.e., hashtag). For instance, one of the unique features in this setting is that all the transactions are by default visible to the users of the platform turning each purchase into an advertisement to the social neighbors of the customer. Other unique characteristics of this new e-business service include that each transaction takes place within the platform itself deploying only the core features of the platform (designated hashtags) and without the need to visit the website of the retailer (e.g., Amazon Cart service), both physical and electronic products are available for purchase, the product offerings are exclusive to the platform and not available via other channels, and the products are delivered to the user that completed the transaction instead of other users (e.g., Facebook Gifts service). These features not only vastly distinguish this novel type of social commerce from other ventures of interconnecting social media and commerce but also drive innovation by transforming the way marketers can conduct e-business.

The novel s-commerce business model under study has strategically important implications for businesses with presence in social media. The 'pay by tweet' service is an exemplary model of how social media are driving innovation in electronic commerce. Enabling the customers to make a "frictionless" purchase within the social platform, replacing previous payment methods, and broadcasting the purchases, empowers the advertisers to leverage the social platforms not only as advertising channels but also as direct sales channels. Apart from generating direct revenue through these online purchases, businesses can harness consumers' social networks and create "buzz" by turning a purchase into an advertisement for the social neighbors. Last but not least, the social commerce venture is a business model that has the

potential to successfully engineer WOM by disseminating the users' advocacy into the online network. Due to the effectiveness of WOM at driving sales and generating leads, marketers have been attempting to engineer WOM, rather than expect the discussions to occur naturally (Godes et al. 2005). Hence, the emergence of new strategies for engineering WOM are of paramount importance for businesses nowadays as the ability to do so successfully implies that WOM can be, at least partially, affected by the firm.

In this paper, we study this unique e-business model, a social commerce service, and aim at elucidating the factors that drive and affect the consumers' decision to adopt this novel service and make a purchase that will automatically be disclosed to the social network. This unique characteristic of the service can essentially transform a purchase into an endorsement and thus offers the potential to engineer WOM by strategically disseminating such information. Employing both econometric and predictive models, we study how the characteristics of the user and her social network affect the decision to make a purchase and engage into WOM at the same time, in an attempt to better understand how firms can conduct business through social commerce and spur online conversations. Observing the WOM episodes, the breadth of their dissemination, and the valence of the recommendations, we are able to study the distinguishing characteristics of the disseminators that are associated with successful post-purchases (i.e., after the transaction of the disseminator has been broadcasted to the network) of their neighbors either due to awareness or influence effects. Therefore, our findings have important managerial implications for firms that would like to engage in similar e-business ventures and provide guidelines to marketers that would like to use such strategies in order to engineer WOM. Overall, we believe that this paper makes significant contributions to the IS literature and specifically the related streams of research studying social commerce, e-business, and online WOM.

Regarding the structure of this paper, the following section provides an overview of the different streams of literature that are closely related to the social commerce phenomenon under study and then develops the research hypotheses that are empirically examined. Then, the data generating process and our unique data set are thoroughly described. The next section presents the conducted empirical analysis, as well as the corresponding robust specification checks, and discusses the antecedents of participating in this novel social commerce initiative. Identifying, the potential for engineering WOM, a predictive model is built in order to both identify the effective disseminators of WOM and discover their distinguishing characteristics. Finally, the last section discusses the key managerial implications of this study and concludes with a discussion of limitations and future research directions.

Literature Review

Online Social Networks and Social Commerce

Social media is a group of Internet-based applications that build on the foundations of Web 2.0, according to which content and applications are continuously modified by users in a participatory and collaborative fashion, and allow the creation and exchange of User Generated Content (UGC) (Kaplan and Haenlein 2010). Social networks, a prominent type of social media, enable constant connectivity among friends, acquaintances, or even strangers and, hence, have fundamentally changed the way people consume and interact (Aral et al. 2013) as well as the way users create and share content. This development, in combination with the proliferation of mobile devices, has spurred the growth of social media which have now gained great momentum; it has been reported that consumers spend more time on social networks than any other type of online activity including e-mail (Nielsen 2012). Starting to recognize the strategic opportunities that emerge from such information technology advancements and the subsequent changes in consumer behavior, several companies started using social networking sites, initially as another medium of online advertising. Some marketers, however, have already moved beyond the one-way communication and tapping into social media's central tenets, collaboration and interaction, harness the unique features of social networks by encouraging brand engagement and the creation of interactive brand communities. One of the most promising ventures trying to capitalize on these opportunities is the 'pay by tweet' service by Amex on the social network of Twitter.

Online Word-of-Mouth

Social media and social networks offer an appealing context in which we can study online word-of-mouth (WOM). Online social networks render online WOM more convenient than traditional forms of WOM (Shi

et al. 2013) because of their unprecedented scale (Dellarocas 2003) and their ability to spread information rapidly (Forman et al. 2008). Chevalier and Mayzlin (2006) were among the first to show that online WOM affects consumer purchasing behavior. Since then, WOM has attracted a great deal of attention from both academic researchers and practitioners. Online WOM is also often referred as earned media in the marketing literature; publicity gained through promotional efforts other than traditional means of advertising. Bollinger (2013) notices that earned media exposures may be difficult to track and are not in the control of the advertiser. Nevertheless, it has been shown that earned media, in combination with paid and owned media, drive consumer purchases (Bruce et al. 2012).

Capitalizing on the opportunities of online WOM, companies would like to better understand and, if possible, control this effective marketing strategy. This is especially important nowadays that traditional forms of advertising seem to be less effective (Lavinsky 2013; Wergin and Muller 2012) and earning consumers' trust is of paramount importance for a successful campaign. For instance, it has been shown that WOM from friends and family is the most influential and trustworthy source of information for consumers across the globe (Nielsen 2013). Thus, firms would like to be able to "engineer" WOM and foster conversations, rather than hope WOM to naturally occur from satisfied customers. This type of WOM is also known as "endogenous WOM" and should be distinguished from "exogenous WOM", which is not under the firm's control (Godes and Mayzlin 2004). A similar marketing practice when firms give incentives to consumers to spread information about a product via word-of-mouth is usually called viral marketing. Studying viral marketing strategies, Aral and Walker (2011) examine how firms can create word-of-mouth peer influence by designing viral features into their products ad marketing campaigns and they find that active-broadcast viral features generate an additional increase in peer influence and social contagion compared to active-personalized viral features.

Pertaining to the phenomenon under study, the social commerce initiative of Amex is also an effort to engineer WOM as a customer who publicly makes a purchase automatically endorses the service and the product. In particular, such disseminating features have the potential to impact the decisions of the users' social neighbors through a variety of influence processes, including those that raise awareness as well as those that persuade individuals to change their expectations of the utility (Aral 2011). Even though there usually exists an inherent tension between these objectives, since high persuasiveness is usually associated with less breadth of awareness (Godes and Mayzlin 2004), the unique context of the Amex social commerce service allows for both. In this setting, influencers and disseminators help the firm achieve maximal awareness as they share their purchases by default and increase awareness while, depending on the valence of the recommendation, they also have the potential to be persuasive as well. Hence, the 'pay by tweet' service can also be used as part of a strategy for engineering WOM leveraging existing social connections; it involves one-to-many communication between a customer and her social neighbors and has the potential to spread rapidly. This study contributes to this stream of research by explicitly studying the WOM episodes and their outcomes as well as evaluating at the same time the impact of the valence of the recommendations in a social commerce setting.

Hypotheses Development

Privacy Concerns and Social Sharing

A quintessential feature of the presented social commerce model is the incorporation of an audience into the purchase decision. The integration of the purchase process with social sharing renders all the transactions publicly visible and essentially the preferences of a customer are disclosed to the local social neighborhood and potentially to the rest of the network. Revealing such important information about the users might lead to critical privacy concerns, especially in the case of non-reciprocal relationships. For instance, Humphreys et al. (2010) found that users abstain from posting personally identifiable information (e.g., email addresses, home addresses and phone numbers) on Twitter and so such information is rarely included in tweets. Besides, the users might refrain from social sharing due to self-presentation concerns (i.e., the desire to control the impressions other people form of them) and the corresponding inferences their social neighborhood friends can draw based on such purchases. Even though in the real world it is feasible to adjust or change the desired impression for different groups of people based on the nature of the particular relationships, in the specific social commerce setting the same information is disclosed to all the groups. Hence, given the diversity of the groups and relationships in the social neighborhood, it is unlikely that a user would desire all her/his neighbors to form the same

impression. Therefore, taking also into consideration that lateral surveillance (i.e., the asymmetrical, non-transparent monitoring of citizens by one another) is an important component of social networks (Andrejevic 2004), the digital visibility of purchases might attenuate the aforementioned problems. Additionally, this effect might be further intensified due to the anticipatory social influence, since the individual's decisions might change simply in anticipation of the virtual presence of her peers (Rhue and Sundararajan 2013). Thus, privacy concerns and social norm effects might hinder social commerce purchases. However, online social networks on average are both vaster and have weaker ties than offline social networks (Acquisti and Gross 2006) and users, apart from close friends, also connect with acquaintances, or even strangers in the social network that might not perceive as important audience. Therefore, the users might not feel the need to conform to social expectations especially when the relationships in the social platform are not reciprocal; a follower might not be a followee. Under this scenario, privacy concerns and social norms might not affect social commerce purchases.

Based on the above discussion and the corresponding streams of research, a user might refrain from publicly propagating such a purchase and endorsement, due to self-presentation, privacy and other similar concerns. Therefore, we expect the size of the local social network to have a negative effect on the user's decision to participate in the service, make a purchase and publicly endorse a brand, if indeed the user changes her behavior in response to privacy concerns and social norms. Hence, we hypothesize the following:

> *H1: A larger local social network of followers is likely to discourage users from making a purchase and sharing a public endorsement for the brand and the corresponding product.*

User Seniority and Ease of Use

The 'pay by tweet' service that was recently initiated on Twitter's social network is a novel type of social commerce that enables consumers to make purchases on the social network using a designated hashtag. A widely accepted model that theorizes about the adoption of novel technologies, the Technology Acceptance Model (TAM), also applies to the adoption of e-commerce (Gefen and Straub 2000). According to TAM, technology adoption is affected by prior use-related beliefs and one of the most important such beliefs is the perceived ease of use; the degree to which the user believes that a particular system would be free of effort (Davis et al. 1989). In our setting, based on the previous proposition, a more senior user of the platform as well as a more active user are more familiar with the platform and its features and, hence, are more likely to have a more favorable disposition towards adopting an innovative service. Hence, we hypothesize the following:

> *H2: A more senior user of the platform is more likely to become an adopter of the social commerce service.*

> *H3: A more active user of the platform is more likely to become an adopter of the social commerce service.*

Brand Trust and Loyalty

Another related stream of research examines the effects of trust and loyalty in e-commerce. In particular, trust is a crucial aspect of many economic activities that revolve around consumers and especially those entailing economic transactions. It has been shown that online purchase intentions are the product of both consumer assessments of IT itself (e.g., ease of use) and trust in the e-vendor (Gefen et al. 2003); trust can be integrated with perceived risk given the implicit uncertainty of the e-commerce environment (Pavlou 2003). Additionally, anecdotal evidence suggests that the loss of trust towards a vendor can have detrimental economic effects. It is indicative that Target's profit significantly declined the first quarter after the company's data breach (Harris 2014). Trust is also a crucial antecedent of participation, especially in online settings, because vendors can more easily behave opportunistically (Reichheld and Schefter 2000). Trust essentially creates a "goodwill stock" that makes consumers more likely to engage in an innovative endeavor of the brand since when a situation presents uncertainty, information asymmetry, or fear of opportunism, trust plays a crucial role in decreasing the uncertainty and the lack of information. Two of the factors that enhance consumers' trust towards the vendor are increased interaction history (McKnight et al. 2000) and familiarity (Gefen 2000). Similarly, membership in brand communities

established on social media has also positive effects on brand trust and brand loyalty (Laroche et al. 2013). Thus, we hypothesize the following:

> *H4: A loyal brand follower of the service provider is more likely to participate in the social commerce service.*

Social Network Interactions and Valence of WOM

Network interactions and peer effects among individuals have been extensively studied in various settings and applications. An emerging stream of literature has paid attention to the influence of social networks on the diffusion and adoption of products or services. The traditional considerations of peer effects include homophily, a tendency of peers towards the same behavior, and peer influence and contagion, the spread of behaviors throughout the peer group (Aral et al. 2009). Nevertheless, both of these effects suggest a conformity to behaviors of social neighbors that extends to consumption choices. Hence, social network interactions facilitate to a large extent the adoption of products and services. Studying this conjecture, Hill et al. (2006) provide statistical support for the hypothesis that network linkage can directly affect product/service adoption while Bakshy et al. (2009) find that adoption rates quicken as the number of friends adopting increases. Similarly, recent studies on social media show that consumers' social activities can increase product awareness (Aral and Walker 2011), drive additional sales (Chen et al. 2011), and enhance brand loyalty (Rishika et al. 2013). In the specific phenomenon of social commerce in the context of Twitter and based on the aforementioned findings, the implicit and explicit advocacy of users or simply the visibility of their choices could lead to a conformity of the behaviors of their followers as well as a broader and/or faster diffusion through a variety of influence processes, including awareness and influence mechanisms.

Furthermore, in the context of user generated context (i.e., consumer reviews), it has been shown that the valence of recommendations, either in the form of numeric review ratings (Chevalier and Mayzlin 2006; Forman et al. 2008; Godes and Mayzlin 2004) or in the form of textual information (Archak et al. 2011), affects product sales. Given that in our setting a user may choose to personalize her advocacy message, when she publicly shares her purchase, we aim at elucidating in the particular context both the impact of the user's friends personalizing their messages and the effect of the valence of their recommendations. Hence, we hypothesize the following:

> *H5: A user with a larger number of purchasers in her social neighborhood is more likely to make a purchase through the social commerce service.*

> *H6: A user who receives a larger number of personalized messages about the social commerce service and products is more likely to make a purchase through the social commerce service.*

> *H7: A user who receives a larger number of messages with a higher average sentiment about the social commerce service and products is more likely to make a purchase through the social commerce service.*

The E-Business Model and the Data Generating Process

The data generating process of the specific social commerce transactions differs significantly from that of traditional online and offline commerce. In particular, in the social commerce setting under study, American Express broadcasting a short message in the social network announces the list of participating merchants and the products that are available for sale with the respective sale price and the designated hashtags users must use in order to make a purchase. Consumers who are interested into making a purchase must have a Twitter account and sync their American Express account with Twitter through an easy opt-in process. Hence, only American Express cardholders and Twitter users are eligible to participate in the program. Once American Express announces the products, users can purchase them by posting a short message, usually referred as tweet, in their social network profile of Twitter and including the designated hashtag. Such a tweet is publicly posted on the Twitter profile of the user and her friends (i.e., social neighbor users who follow the stream/timeline of the specific user) will automatically receive the tweet on their own newsfeed. At the same time, American Express tracks the tweets that use the designated hashtag in the social network and matches them to the desired product. Once a tweet is automatically processed by American Express, a reply is sent to the user prompting her to confirm the

purchase within fifteen minutes. After the purchase is confirmed, American Express bills the customers and ships the product for free to the billing address of the consumer's American Express credit card within 1 to 5 business days. The standard product warranties apply while products are eligible for returns within 30 days after the date of delivery. Also, consumers can purchase multiple products but they are limited to one purchase per item. During the aforementioned process, consumers receive two e-mails providing information about their orders. The first email is a purchase confirmation email that lists all the order details, such as price, tax, shipping and invoice number, and the second email is a shipment confirmation e-mail that provides the product tracking number as well as billing/shipping information.

Our unique database contains all the social commerce transactions that were generated through the aforementioned process on Twitter's social platform. Each transaction is committed from a Twitter user account and is associated with a specific product offering. We have filtered out unconfirmed and ineligible attempts to make transactions, such as the tweets that were posted after the expiration date of the product offerings. 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 tweet of the user, the tweet id, the exact date and time that it was posted, the user id, and the designated hashtag. Moreover, our database also contains social network users who were eligible to make a purchase but choose not to do so. Additionally, we have access to user specific information provided by the Twitter API, such as the user's screen name, the number of followers and the number of friends, the datetime that the user created the account, how many statuses the user has posted since the creation of the account, and the self-reported description of the user's profile, etc.

Moreover, we enhanced our database by collecting the users' social network neighbors. For each user we collect the set of followers and the set of friends (or followees as they are also called). When a user is following another user in the platform that means she subscribes to her/his stream of tweets. On Twitter, users don't have to ask for permission to follow someone; once they follow them they automatically start receiving their tweet updates in their own newsfeeds. However, the number and the set of the followers can be different from that of the friends since Twitter, as opposed to Facebook for instance, has not enforced reciprocity among follower relationships. Since the relationship is not by default reciprocal, users have to follow each other in order both to get updates of each other on their newsfeed. Hence, we are able to track how many friends of a specific user purchased any type of product in this social commerce setting and when exactly these purchases took place. To avoid introducing biases in our empirical models, we have excluded from our analysis all the user accounts who claimed product offerings but do not have any followers or were created after the launch of the service.

Additionally, we have information about all the product offerings. American Express partnered with well-known retailers and offered in total eight different products available for purchase. The products, which were offered at a reduced retail price, were available for purchase only for a specific period of time. The featured products belonged in a wide variety of categories and all of them were 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 American Express was available for purchase only through the social network of Twitter. Hence, our study does not suffer from sample selection bias issues. Sample selection bias would arise if the users could choose from which social network to make the purchase and we had analyzed only the transactions that took place on Twitter.

Table 1 summarizes the variables used in the analysis and shows the corresponding descriptive statistics computed over all the observations in our data set; in the predictive modeling and network-based targeting section additional non-numeric variables are described.

Table 1. Variables and Descriptive Statistics						
Variable	**Description**	**Mean**	**Std. Dev.**	**Median**	**Min**	**Max**
Purchase	Whether the user has made a purchase using the 'pay by tweet' service	0.027	0.163	0	0	1
MonthSeniority	The number of months since the user created the account on Twitter	27.641	15.529	27	0	83
FollowersCount	The number of followers the user has	1,435.733	82,273.402	105	1	8,057,104

FriendsCount	The number of users this user is following	859.335	4,837.867	409	0	802,984
Verified	Whether the user has a verified account	0.003	0.054	0	0	1
RankingAmexInFriends	The relative ranking of Amex in the list of friends of the user	0.017	0.107	0	0	1
FriendsPurchased	How many friends of the user have made a purchase using the social commerce service	0.598	2.850	0	0	732
PersonalizedTweets	The number of personalized tweets issued by the friends of the user until the time of the decision	0.319	0.578	0	0	2
AvgSentiment	The average sentiment in the endorsements from the friends of this user	0.080	0.209	0	-1	1
VerifiedAccounts	The number of verified users this user is following	8.803	54.043	2	0	12,336
VerifiedBrands	The number of verified brands this user is following	2.591	6.897	1	0	706
ActiveUser	The average number of tweets (including retweets) issued by the user per day	4.094	120.030	0.43	0	77,844
% Discount	The promotional discount	0.264	0.164	0.23	0	0.512
$ Price	The promotional price of the product	93.745	69.625	87.50	15	179.99

Econometric Modeling and Empirical Results

In this empirical analysis, we aim at elucidating the factors that drive and affect consumers' decision to participate in this novel social commerce service and, hence, make a purchase that will be automatically broadcasted to the social network. Understanding the impact of user characteristics and local social network on such user decisions is of paramount importance, given the unprecedented opportunities this venture provides to directly generate revenue in social media and spur online WOM (as we discuss in the next section). Additionally, this analysis will provide valuable insights on how firms can conduct business through social commerce.

In particular, in the social commerce setting under study, when the product offerings are announced by American Express on Twitter, each individual faces the decision of whether to purchase an item or not. Throughout our analysis, we model this decision of a user to make a purchase or not employing discrete choice models in a hedonic-like framework. Since, there are also factors not observed by the participating firms and the researchers, such as the user's satisfaction with similar products, the agent's choice is not completely deterministic and cannot be predicted exactly. Instead, the probability of any outcome is derived in a consistent, invariant, and efficient way (Greene 2003). Based on the above discussion and since a rational decision process takes place, we use binary choice models to efficiently model the actions of the users. The model platform is an underlying random utility model or latent regression model (McFadden 1973), $y^* = x\beta + \varepsilon$, in which the continuous latent utility or measure, y^*, is observed in discrete form through a censoring mechanism:

$$buy = 1 \; if \; y^* > 0,$$

$$buy = 0 \; if \; y^* \leq 0.$$

In the specific social commerce setting, the corresponding random utility model is the following:

$$\begin{aligned} logit(buy = 1) = \; &\beta_0 + \; \beta_1 \log(FollowersCount) + \beta_2 MonthSenority + \beta_3 ActiveUser \\ &+ \; \beta_4 RankingAmexInFriends + \beta_5 FriendsPurchased + \beta_6 PersonalizedTweets \\ &+ \beta_7 AvgSentiment \end{aligned}$$

As illustrated by the above choice model, we expect that several individual-specific and social network variables might affect a user's decision to make a purchase and, as a consequence, publicly endorse the

brand and the product. In particular, the variable *followersCount* represents the size of the online social network neighborhood of the decision maker. Then, the next two variables measure the user's expertise and familiarity with the social network platform. Firstly, the *monthSeniority* variable measures the months that have elapsed since the user created her Twitter profile. Secondly, the *activeUser* variable operationalizes how actively and frequently the user is using the platform as indicated by the number of total tweets the user has posted over the number of days her account has been active. Besides, the brand community membership is also considered. Specifically, the variable *rankingAmexInFriends* provides information about whether the user is a member of the Amex's brand community on Twitter. In order to obtain a richer description of the loyalty of the user towards the brand, we capitalize on the unique feature of Twitter to rank the friends and followers based on the time that the non-reciprocal and user-initiated friendship was established online. Hence, we observe what is the relative ranking of American Express in the list of the friends of the user and we transform the binary variable into a continuous variable bounded between 0 and 1. The lower bound of 0 means that "the user does not follow Amex". As the ranking of Amex in a user's friends list increases, the ratio increases reaching the upper bound of 1 (i.e., "Amex was the first Twitter account the user followed").[1] We should note here that the newsfeed (or timeline) in the social network of Twitter is not algorithmically curated but the messages are displayed in chronological order and, hence, how highly a message is ranked in the feeds of the followers of a user is independent of the content and the sender of the message. Next, the following three variables are related to purchases made by a user's social neighbor friends. In particular, the variable *friendsPurchased* captures how many friends have made at least one purchase using the specific social commerce setting and their tweets have been projected on the user's newsfeed. In an effort to capture more refined measures of endorsement, the variable *personalizedTweets* measures the number of personalized tweets the user has received in total until the time of the decision. As previously discussed, a friend of the user could simply retweet the default tweet of Amex in order to make a purchase or may choose to personalize her/his tweet message. Finally, the variable *avgSentiment* measures on a continuous scale between -1 and 1 the average valence of the personalized endorsements from the user's friends as measured by natural language-processing (NLP) algorithms.[2] In the discussed model, the coefficients are estimated using the maximum-likelihood method. Table 2 reports the estimates of the model and the corresponding goodness-of-fit statistics.

Table 2. Results from the Discrete Binary Choice Model						
Variable	**Coefficient**	**Std. Error**	**Z**	**P>|Z|**	**[95% Conf. Interval]**	
log(FollowersCount)	-.6427388 ***	.0050328	-127.71	.000	-.6526029	-.6328747
MonthSeniority	.0290927 ***	.0005432	53.56	.000	.0280281	.0301573
ActiveUser	.0000911 *	.0000403	2.26	.024	.0000122	.00017
RankingAmexInFriends	4.094084 ***	.0314009	130.38	.000	4.032539	4.155629
FriendsPurchased	.0355443 ***	.0028636	12.41	.000	.0299318	.0411568
PersonalizedTweets	.3814432 ***	.0183014	20.84	.000	.3455731	.4173133
AvgSentiment	.257461 ***	.0460746	5.59	.000	.1671565	.3477655
Constant	-2.51326 ***	.0204786	-122.73	.000	.1671565	.3477655
LR chi2(7) = 38721.86 McFadden's Adj R^2 = 0.238 AUC = 87%						

Based on the empirical results presented in Table 2, we find statistically significant effects across all the variables included in the model. The table presents the coefficients in the standard log odds format. For a more intuitive interpretation of the results, we transform the log odds to odds ratios that represent how much the odds, rather than the log odds, of a purchase increase multiplicatively with a unit increase in the independent variable. As we observe, the user's loyalty towards the brand of American Express and the number of the personalized tweet messages received from friends in the social networks have the higher impact on increasing the odds of the user making a purchase. In particular, an increase of 0.01 units in the

---

[1] An alternative but less attractive option would be to use the ranking of the user in the Amex's list of followers. However, this variable would be problematic, as it would not take into consideration, for instance, when a user's account was created and how many friends the specific user has.

[2] Alchemy's sentiment analysis algorithms and orientation analysis mechanisms were used to extract the friends' attitude toward the service and the product. Both mechanisms yielded similar results and we choose to include only the sentiment scores to our model in order to avoid multicollinearity problems.

employed measurement of customer loyalty corresponds to a 59% increase in the likelihood of the user to purchase an item (*H4*), whereas an additional personalized tweet received in the newsfeed of the user corresponds to a 46% increase (*H6*). On the other hand, the size of the social neighborhood of a user has a significant but negative impact (*H1*), as a double fold increase in the number of followers is now likely to decrease the chances of the user making a conversion by 47.4%. Statistically significant and positive effects towards increasing the odds of a purchase have also been found for the rest of the social network peer effects variables. In particular, an additional friend who makes a purchase (before the user) increases the odds of the user making a purchase by 3.6% (*H5*). Similarly, the more intensively the friends of the user recommend the service and the product, the more likely the user is to make a purchase (i.e., 29% increase in the odds for a unit increase in the sentiment score) (*H7*). Finally, the variables that capture the user's familiarity and expertise with the platform have a more moderate, but also positive and significant, effect on increasing the odds of a purchase (*H2* and *H3*). Specifically, a user who has been familiar with the platform for one additional month increases the odds of making a purchase by 2.9% while a user who is using the platform more frequently increases the odds of making a purchase by 0.009% for a unit increase. Evaluating the performance of our model, the McFadden's Adjusted R^2 metric indicates a very good fit. Additionally, using the Receiver Operating Characteristic (ROC) curve to evaluate the performance of the model, the AUC (Area Under the ROC curve) metric shows that our model can differentiate between random true positive and true negative instances 87% of the times indicating the very good performance of the model.

Next, we model the user's decision to purchase a specific product offering, relaxing the assumption that the observations pertaining to a specific product offering are not correlated. To model the corresponding decision, we employ a generalized linear mixed model (GLMM) for binary responses according to which the log odds of outcomes are modeled as a linear combination of the predictor variables where there are both fixed and random effects. We should note here that we allow the user decision of whether to make a purchase to also depend on the attractiveness of the product offering, which is captured by the random effects we introduce in the model.

In this specification, random effects accommodate the observed and unobserved time-invariant heterogeneity of the product offerings and allow the decisions to purchase a specific product or not to be correlated across consumers within a specific product offering. Hence, they allow us to control for the attractiveness of the product offering and evaluate the coefficients of interest conditioned on that. Ignoring necessary random effects and assuming independent observations within products could lead to inconsistent estimates and distorted (smaller than normal) standard errors (Train 2009). An alternative way to control for product offering characteristics would be to introduce product specific characteristics of the offerings in our model. However, identification of such features is not feasible in this setting since those are standardized products and, thus, there is no variation of the characteristics within a product offering across consumers. Therefore, we control for the overall effectiveness of the product with two different alternative specifications of the proposed model. Furthermore, the fixed effects parameters determine the conditional mean of the response given the random effects. Since GLMMs allow for binary outcomes; denoting as p the probability that a user will purchase a specific product offering, the link function $g(.)$, which relates the outcome to the linear predictor that we shall denote η, takes the following form:

$$g(.) = log_e\left(\frac{p}{1-p}\right).$$

Then, the linear predictor that determines the conditional mean according to the link function $g(.)$ is the combination of the fixed and random effects excluding the residuals:

$$\eta = X\beta + Z\gamma \quad \text{where} \quad \gamma \sim N(0, G),$$

where X is the matrix of fixed effects explanatory variables, β the vector of the respective coefficients, Z is a matrix which codes with binary variables to which product does the decision of each customer to convert or not relates to and γ is the vector of random effects. The outcome y is modeled as:

$$\begin{aligned} y = \eta + \varepsilon = \beta_0 + \beta_1 * log(FollowersCount) + \beta_2 * MonthSeniority + \beta_3 * ActiveUser + \beta_4 \\ * RankingAmexInFriends + \beta_5 * FriendsPurchased + \beta_6 * PersonalizedTweets + \beta_7 \\ * AvgSentiment + Z\gamma + \varepsilon. \end{aligned}$$

We should note here that we adapt the three variables of the model measuring peer effects so as to correspond to the effects of the same product offering (e.g., how many friends purchased the exact product the decision to purchase pertains to rather than how many friends in total made purchases). Since there are no closed form solutions for GLMM, we approximate the true likelihood with numerical integration. Specifically, we use the adaptive Gauss-Hermite quadrature (GHQ), which adaptively varies the step size to control for the error of approximation. The results of the model above are presented in Table 3.

Table 3. Results from Generalized Linear Mixed Model (GLMM)				
Variable	**Coefficient**	**Std. Error**	**Z**	**P>|Z|**
log(FollowersCount)	-.6349 ***	.003883	163.50	.000
MonthSeniority	.02645 ***	.0004297	61.56	.000
ActiveUser	.0000688 ***	.00001663	4.14	.000
RankingAmexInFriends	3.662 ***	.02205	166.09	.000
FriendsPurchased (Same Product)	.03457 ***	.002813	12.29	.000
PersonalizedTweets (Same Product)	.4028 ***	.04221	9.5	.000
AvgSentiment (Same Product)	.5943 ***	.01584	37.52	.000
Constant	-6.120 ***	.0204786	-122.73	.000
Random Effects : ProductId (Intercept) Std.Dev. = 1.986				
AIC = 214546.5 BIC = 214667.6 LogLik -107264.2				

As Table 3 illustrates, controlling for product offering attractiveness, our results remain qualitatively the same while all the variables are statistically significant ($p < 2e - 16$). Most of the variables have smaller coefficients since part of the variance is explained by the random effects for the product offerings. As before, we transform the presented log odds into odds ratios for a more intuitive understanding of the results and we interpret the coefficients conditional on the attractiveness of a product offering. Based on the results, the user's loyalty towards the brand of American Express (*H4*) and the number of the personalized tweet messages received from friends in the social network (*H6*) continue to have high impact on increasing the odds of the user (i.e., 38.93% increase in the likelihood of purchasing the specific product for a 0.01 unit increase and 49.6% for a unit increase, respectively) making a purchase. As expected, the size of the social neighborhood of a user has a significant but negative impact (*H1*); a double fold increase in the number of followers is now likely to decrease the chances of the user making a conversion by 47%. As presented in the previous model as well, significant and positive effects towards increasing the odds of a purchase are also found for the rest of the social network peer effects variables. In particular, an additional friend who makes a purchase (before the user) increases the odds of the user making a purchase by 2.7% (*H5*). Similarly, the more intensively the friends of the user recommend the service and the product (*H7*), the more likely the user is to make a purchase (i.e., 81% increase in the odds for a unit increase in the sentiment score). The increased effect of average sentiment and number of personalized tweets in this model could be explained by the fact that now these variables are related to the same product offering that the dependent variable refers to, rather than any product offering. Finally, variables that capture the familiarity and expertise of the user with the platform still have a moderate, but positive and significant, effect on increasing the odds of a purchase. Specifically, a user who has been familiar with the platform for one additional month increases the odds of making purchase by 2.7% (*H2*) while a user who is actively using the platform more frequently increases the odds of making a purchase by 0.007% (*H3*). Even though, the fit of the GLMM model is not directly comparable with that of the discrete choice model, because the log-likelihoods are commensurate, both of the specifications provide a very good fit. We see that in all the specifications of the model our hypotheses are supported and based on the presented results we derive conclusions in the Discussion section.

Robustness Checks

Additionally, we conducted various robustness tests including different specifications of the aforementioned models in order to examine whether the key results remain consistent and are not sensitive to model specifications. In particular, we control for observed and unobserved time-invariant product-specific effects that might be correlated with the regressors of the model and, thus, we employ a fixed effects discrete binary choice model. The results of this model are presented in Table 4. We find that the qualitative nature of our results remains unchanged while all the coefficients remain statistically

significant. Furthermore, we rerun the same models allowing for peer effects variables to include all the product offerings, independently of the product offering the dependent variable is associated with; again, the results remain qualitatively the same but the corresponding table is omitted due to space limitations. Similarly, we find that a simple binary variable of brand community participation, rather than taking into account the relative ranking of Amex in friends, also provides similar qualitative results. However, the relative ranking of Amex in friends provides a better fit of the model and thus we adopt this measure. Finally, we also allow the product to moderate the relationship between the likelihood of service adoption and the size of the social neighborhood of the user. However, this model specification does not provide statistically significant better fit to our data.

Table 4. Results from Conditional Logit (FE Logistic Regression)				
Variable	**Coefficient**	**Std. Error**	**Z**	**P>|Z|**
log(FollowersCount)	-.5652 ***	.003561	-158.707	.000
MonthSeniority	.02394 ***	.0003992	59.957	.000
ActiveUser	.00005759 ***	.00001190	4.839	.000
RankingAmexInFriends	2.812 ***	.01723	163.275	.000
FriendsPurchased (Same Product)	.01923 ***	.0007790	24.693	.000
PersonalizedTweets (Same Product)	.5337 ***	.01377	38.749	.000
AvgSentiment (Same Product)	.2396 ***	.03885	6.166	.000

Predictive Modeling and Network-based Targeting

The explanatory study that we presented in the previous sections revealed what factors influence the decision of eligible social media users to adopt this service and the relative importance of these factors. In this section, we switch from explanatory modeling to predictive modeling. In other words, the main goal now is not to explain which factors affect these customer decisions, but to examine how well we can predict whether followers of particular users will (subsequently) adopt the service and purchase specific items after they are exposed to implicit or explicit advocacy. Apart from assessing our predictive power of estimating the likelihood of specific groups of customers to convert, this method will also provide us evidence of whether we can identify potentially influential users who affect the economic behavior of their followers. Answering these questions has important managerial implications for brands and marketers since the implementation of successful marketing campaigns depends on how well these "influencers" and, in general, any effective disseminators of information can be identified and targeted (Godes and Mayzlin 2009; Hill et al. 2006). Using predictive models to accurately estimate the likelihood of conversions in the local social network of a user, estimate the user's network value and identify influential users, we can design viral marketing plans that maximize the expected monetary benefit of the firm as well as positive word-of-mouth (Domingos 2005). A significant distinguishing characteristic of the proposed approach is that we do not focus solely on identifying these influencers and disseminator in the social media platform but we put equal emphasis on identifying the differentiating characteristics and attributes of such users. The presented analysis constitutes an example of how firms could leverage the vast amount of unstructured user-generated content in order to gain competitive advantage.

The effective use of vast amounts of unstructured user-generated content in the social media context can be efficiently achieved using relevant predictive modeling techniques from the fields of machine learning and data mining. Applying machine learning and data mining techniques, we can also easily avoid common pitfalls, such as overfitting our data and inducting models that do not generalize beyond our training data set. Besides, the use of predictive modeling techniques, instead of econometric models, allows us to build more accurate models of higher complexity using more advanced methods. In particular, for this specific data mining problem, we train Random Forests (Breiman 2001), Bagging (Breiman 1996), AdaBoost (Freund and Schapire 1996), and linear function (logistic regression) classifiers in combination with text-mining techniques. We employ a cross-validation scheme with 10 folds (Provost and Fawcett 2013), and we evaluate each model in terms of both ranking and classification tasks, based on the accuracy (i.e., percentage of correctly classified instances), Area under the ROC curve (AUC), and F1-score measures. For the predictive models, we consider an extended set of attributes, including the user related attributes discussed in the explanatory analysis as well as the inferred gender of the user, the url provided by the user in association with her/his profile on the social media platform, whether the user has altered the default theme or background of her/his profile, whether the user has

uploaded her/his own avatar as profile image of her/his profile or a default egg avatar is used instead, the number of brands and verified accounts the user follows, the implicit interests of the user, user-defined text features describing the account of the user, etc. In particular, the various text features describing the users are derived applying text-mining techniques to the user-defined description of their accounts and after some pre-processing of the words, which includes removal of stop-words and non-English words, stemming, and fuzzy matching. In addition, the user interests are captured based on whether the user is connected on the social platform with brands of specific categories (e.g., electronics, beauty and fashion, etc.). Finally, the gender of the user is inferred based on a binary classifier, with accuracy of more than 0.80, and using the name of the user, as they have defined it, and the frequency of names in the male and female population.

Using our unique database, we built different models that predict whether a user is potentially "influencer" and, thus, her/his social contacts in the platform will subsequently adopt the service. Table 5 presents the results of our experimental evaluation illustrating the predictive (out of sample) power of the employed approach. Based on the evaluation results, the inducted models perform well both in terms of classification and ranking; the difference between the AUC metric and the accuracy and F-measure is due to the low probability of converting (i.e., imbalanced classes). Estimating the lift curves of the four classifiers illustrating that the models' targeting is up to 7.75 times as good as random (and 4.37 times better when targeting about 6% of the population), in terms of the percentage of correctly classified "influencers".

Table 5. Classification and Ranking Performance of the Predictive Model			
	Accuracy	**F1-score**	**AUC**
Logistic Regression	88.12%	0.935	0.764
AdaBoost	87.10%	0.931	0.763
Bagging	88.37%	0.936	0.786
Random Forest	87.97%	0.934	0.727

Variable Importance

Enhancing the interpretability of the conducted analysis and the resulting predictive models, in order to better evaluate the proposed approach and discover richer findings, we also cast the problem as a variable importance problem. Variable importance is often defined in the literature as the effect on a measurable quantity of interest upon changing a variable of interest, holding all other factors constant. The idea of "importance" is in and of itself a vague term (Dalessandro et al. 2012). This concept was presented early on by (Achen 1982) who suggests that any inquiry into the importance of a variable should also have an explicit objective function associated with the importance. Therefore, in the presented analysis, attribute importance should be determined by the ability of a characteristic to predict the conversions of followers of particular users after being exposed to their advocacy. Thus, we measure the variable importance based on the resulting odds ratio derived from the logistic regression model. A thorough survey and analysis of methods for measuring variable importance is available in (Johnson and LeBreton 2004). The main advantage of the selected approach is that the presented results can be easily measured and interpreted. In particular, Table 6 shows the attributes with the highest and lowest odds ratios based on the results of the employed linear model and the corresponding chi-squared statistic evaluating the predictive worth of each attribute; the textual attributes correspond to word stems.

Table 6. Top and bottom attributes based on variable importance					
Variable	**Odds Ratio (Class 1)**	**Chi-squared**	**Variable**	**Odds Ratio (Class 1)**	**Chi-squared**
"strategi"	2.981515	147.1439	default_profile_image	0.428927	1231.5600
"product"	2.306805	124.3077	"mom"	0.634236	11.0996
"mba"	2.295157	61.3610	"father"	0.649815	21.2890
"founder"	2.106150	156.7443	"web"	0.654536	11.1449
"opinion"	2.039984	168.1348	"softwar"	0.654965	9.3838

Based on the results presented in Table 6, the most significant attributes for predicting whether followers of particular users will (subsequently) adopt the service and purchase specific items after they are exposed to implicit or explicit advocacy include whether the user is a strategist (i.e, "stategi"), product manager or marketer (i.e., "product"), or an entrepreneur (i.e., "founder", "co-founder"). In the most important variables are also included whether the user is highly educated (i.e., "mba") and expresses her/his personal opinions via the social profile (e.g., "opinion maker", "personal opinions", "opinionated", etc.), which might indicate that the user holds an important role or position. Another very significant attribute, which is not included in the above top 5 list, is whether the user is a blogger. This information was discovered both based on the description of user and the provided profile url domain (e.g., blogger, wordpress). Similarly, very informative were the attributes related to digital space and marketing (e.g., brand advocate / -marketing / -enthusiast / -manager, digital media / -marketing / -commerce / -innovation, etc.). On the other hand, the most significant attribute for correctly classifying a user as not an "influencer" or disseminator is the user having a default ("egg avatar") profile image and not her/his own personal avatar. Other attributes that were found to provide significant predictive evidence that a user is not influential include the information of the user being a male or female parent who has raised a child (i.e., mom, father) or having a technical job (e.g., web developers, web designers, software developers, and software engineers). We should note here that a strong indication of an "influential" or not "influential" user is not the actual corresponding characteristic of the user but that the user included the specific information in her/his limited profile description (e.g., (Provost and Fawcett 2013)).

Cost/Benefit Analysis

Furthermore, we conduct a cost/benefit analysis using the above predictive models in order to maximize the business profit, and determine how deep to target into the predicted list in order to maximize the corresponding profit, as well as assess the economic implication of the proposed approach. Assuming that the benefit of correctly targeting an "influencer" is 2 monetary units and the cost of misclassifying a "non-influencer" is 1 unit and that the total population consists of 100 users, we derive that targeting the top 5.5028% of the population maximizes our expected profit. This strategy corresponds to an expected benefit of 4.14 units, which is 7.51 units larger than randomly targeting a subset of users of same size (cost). Using our predictive framework, similar estimations can be derived based on any combination of cost/profit and population size. For instance, using the average benefit and cost from our database (i.e., $261.80 and $33.12), the expected profit for a population of the same size is $1315.69.

Discussion and Managerial Implications

In this study, we examine a novel social commerce model that interconnects social networks with e-commerce. This "social commerce" service creates unprecedented strategic opportunities for firms to both generate direct revenue through social media and successfully engineer WOM leveraging the social connections of the users (Adamopoulos and Todri 2014). In order to better understand this pioneering venture, we employ both empirical econometric models and predictive analytic approaches. In the conducted econometric analysis, we study and quantify the impact of certain factors that affect consumers' decision to participate in such an "s-commerce" model and make a purchase that will be automatically broadcasted to the social network. Our findings provide a rich understanding of how users' characteristics and the behavior of their social neighbors can affect the likelihood of a social commerce purchase. Furthermore, taking advantage of the unique features of the particular business model and our ability to observe the WOM episodes, the breadth of their dissemination and the valence of the recommendations, we are able to build predictive models that identify the effective disseminators associated with successful post-purchases as well as discover their distinguishing characteristics. The derived models are essential for orchestrating marketing strategies and constitute an example of how firms can leverage the vast amount of unstructured user-generated content in order to gain competitive advantage. In the following paragraphs, we discuss the most important findings focusing on theoretically integrating them into prior literature as well as illustrating the key managerial implications.

Leveraging the social connections of the users in social media is a central tenet of the presented e-business venture. Marketers might naturally think that targeting users with larger number of followers would be the most effective strategy, since their potential public endorsement of the brand and the product would have a larger audience. However, we find that the size of a user's social network

neighborhood negatively affects the user's likelihood of making a purchase. Hence, we find evidence that the users do take into consideration their social network and might have privacy concerns, despite some contradictory findings in the prior literature. Therefore, simple monetary incentives might not be sufficient to make a social commerce purchase attractive. In particular, marketers need to take into consideration also what type of products would be more effective in maximizing the returns from such a venture. For instance, social commerce might be more effective for socially accepted products. Alternatively, in order to increase the effectiveness of such initiatives, marketers could make prominent any socially desirable product features. Towards this idea, we tested the conjecture that the product moderates the effect of the size of the social neighborhood on a user's decision to make a purchase but the null hypothesis could not be rejected in our dataset; this is an interesting direction that can be examined in future works. As an alternative solution to the negative impact of the size of the social neighborhood on the likelihood of a purchase of the user, companies that would like to engage in social commerce ventures in the future could experiment with providing higher incentives for social sharing to the users with a larger social neighborhood. Finally, this finding highlights the need for marketers to address users' privacy concerns and consider allowing them to make a purchase through direct messages, even though such strategies could attenuate the WOM effects.

Moreover, focusing on the effects of brand trust and consumer loyalty, we find a significant and positive effect on the odds of a social commerce purchase. This finding is in agreement with prior academic research that highlights trust as a crucial factor towards the adoption of e-commerce. We contribute to this stream of research by showing that trust and loyalty to a brand are important predictors also in social commerce. Hence, the participating firms could try to directly or indirectly increase the trust of the users in an effort to further diffuse the adoption of social commerce services. Apart from the implications for network-based targeting techniques, this finding reveals how vital it is for marketers and brands to establish and promote a consumer community in social media enticing users to engage with their brand and potentially further increasing their loyalty. Fostering trust and loyalty could have a positive impact on the overall image of the brand, drive sales, and generate leads associated with key marketing objectives in the rest of the online and offline world. Finally, nowadays that social commerce offers the chance to generate direct sales, social media also provide to the firms the opportunity to better quantify the monetary value of a follower and further evaluate the corresponding marketing strategies.

Furthermore, we also find that the economic behavior of the user's social network as well as the valence of recommendations have a significant and positive effect on the odds of a social commerce purchase. In particular, the more social network friends of a user adopt the product, the more likely the user is to make a purchase. This finding agrees with prior literature examining peer effects in online and offline social networks. We contribute to this stream of research by demonstrating that peer effects, which could operate either through awareness or influence mechanisms, exist also in social commerce settings, even when WOM is firm initiated. Besides, the increased visibility in the social media allows firms to directly observe the WOM episodes and better understand how they operate and spread in the network. In this paper, we contribute to the related WOM literature by studying the impact of a personalized endorsement on the likelihood of a social commerce purchase and quantifying the effect of customer curated versus firm curated WOM messages. Such findings have important managerial implications for marketers that would like to orchestrate WOM. We also show that the personalization of a user's message has a positive and significant effect towards increasing the odds of a social purchase. Based on these findings, in the future, companies could prompt or even incentivize the users to personalize their endorsements in such e-business venture. Finally, the valence of the endorsement, as captured by the sentiment of the personalized message, has a significant and positive effect as well. Hence, in the case of predefined and structured endorsements, companies should also try to promote messages with a more positive sentiment. This observation is in accordance with prior evidence across different empirical settings.

Additionally, pertaining to the user's familiarity with the platform, we find that a user's seniority, in terms of how recently the social network profile was created, as well as the frequent and active use of the platform also have a significant and positive effect on the odds of a social commerce purchase. This finding emphasizes the need for marketers to make the process as frictionless as possible since such a strategy would also attract less experienced users. Additionally, firms should aim at educating the less familiar with the platform users by various means, such as offering guidelines, video demonstrations, and specialized online customer support. This is especially important since a potential customer who has difficulties participating in the social commerce venture could disseminate negative WOM messages to

her social network friends, as a result of her bad experience. Additionally, considering network-based targeting, this finding suggests that more experienced and advanced users should be targeted first, since they are more likely to faster adopt the service and propagate the endorsement messages in the network.

Besides, illustrating how actionable our findings are in real-world marketing problems and attempting to discover richer insights, our predictive analysis aims at identifying potentially influential users who affect the economic behavior of their followers through awareness or influence mechanisms. Our analysis demonstrates how marketers, in the era of big data, could leverage the vast amount of unstructured user-generated content in order to identify effective disseminators. The conducted analysis has significant economic benefits since such customers offer greater value to firms because they have a higher propensity to propagate product information, based on a combination of being particularly influential and having more friends (Richardson and Domingos 2002). Thus, firms that want to adopt such a social commerce business model should find these influencers and further promote their behavior. Considering that a larger social network neighborhood is likely to discourage a user from making a social commerce purchase, firms that would like such ventures to spread faster in the social network and at a larger scale need to move beyond naïve strategies and identify the key disseminators of the information. Analyzing the distinguishing characteristics of such key users, we find that users who hold or relate to key technology positions, such as marketing managers or entrepreneurs, are much more likely to be effective disseminators. Furthermore, we find that often bloggers can also play that role very effectively. Combining these findings with the explanatory analysis results, a successful network-based marketing technique for American Express with significant economic benefit, as revealed by our cost/benefit analysis, would be to target the disseminators of information and provide incentives for personalizing their messages towards their social network friends.

Finally, the social commerce initiative under study and the corresponding findings we present in this paper have important managerial implications that extend beyond social networks and the online world. In particular, the offered simplicity in completing purchases through a designated hashtag generates tremendous potentials for cross-platform marketing activities and it can fundamentally change the way consumers shop online. For instance, we could see traditional media of advertising, such as TV commercials and print ads, using hashtags to generate immediate purchases. Hence, the specific service is not just an innovative push towards social commerce but it is also a push towards mobile commerce, given the opportunity for a new independent sales channel that enables purchases without enduring the traditional checkout process or even visiting the website of the seller.

One limitation of this study is the limited number of participating brands and product offerings and the absence of variation of product characteristics within the product listings across consumers. Even though we cannot individually quantify the impact of specific product characteristics, we control for the overall effectiveness of the product using two different econometric model specifications. As part of the future work, we plan to study the effect of such social commerce initiatives on the user base of the participating firms and, in particular, the future engagement and loyalty of the corresponding users.

Conclusions

In this study, we shed light into a novel social commerce venture and seek to understand its antecedents and consequences. In particular, we study the users' decision to participate in a social commerce model that integrates e-commerce and social networks and find that the likelihood for a social purchase depends on various user characteristics, such as the size of the online social network of the user, her/his loyalty and trust towards the brand, and the familiarity of the user with the social network platform. Additionally, we find that the economic behavior of a user's immediate social network as well as the personalization and the valence of recommendations from the social neighbors of the user have also significant impact on her/his decisions to make such purchases. Furthermore, tapping into the opportunities presented in the vast amounts of unstructured user-generated data, we employ predictive modeling techniques that identify the distinguishing characteristics of the effective disseminators of information and, thus, offer the potential to vastly enhance the effectiveness of network-based targeting strategies of a firm. The findings of this study have important implications for companies that would like to adopt this groundbreaking type of e-business as well as marketers that desire to orchestrate WOM in social networks. Finally, this paper contributes to the IS literature and specifically the related streams of research studying social commerce, e-business, and online WOM.

References

Achen, C.H. 1982. *Interpreting and Using Regression*. Sage.

Acquisti, A., and Gross, R. 2006. "Imagined Communities: Awareness, Information Sharing, and Privacy on the Facebook," *Privacy enhancing technologies*: Springer, pp. 36-58.

Adamopoulos, P., and Todri, V. 2014. "Social Media Analytics: The Effectiveness of Promotional Events on Brand User Base in Social Media," *ICIS 2014*, Auckland, New Zealand: ACIS, p. 8.

Amazon.com. 2014. "#Amazoncart: Add Items to Your Cart without Leaving Twitter."   Retrieved September 2014, from http://www.amazon.com/gp/socialmedia/amazoncart

Andrejevic, M. 2004. "The Work of Watching One Another: Lateral Surveillance, Risk, and Governance," *Surveillance & Society* (2:4).

Aral, S. 2011. "Commentary-Identifying Social Influence: A Comment on Opinion Leadership and Social Contagion in New Product Diffusion," *Marketing Science* (30:2), pp. 217-223.

Aral, S., Dellarocas, C., and Godes, D. 2013. "Introduction to the Special Issue-Social Media and Business Transformation: A Framework for Research," *Information Systems Research* (24:1), pp. 3-13.

Aral, S., Muchnik, L., and Sundararajan, A. 2009. "Distinguishing Influence-Based Contagion from Homophily-Driven Diffusion in Dynamic Networks," *Proceedings of the National Academy of Sciences* (106:51), pp. 21544-21549.

Aral, S., and Walker, D. 2011. "Creating Social Contagion through Viral Product Design: A Randomized Trial of Peer Influence in Networks," *Management Science* (57:9), pp. 1623-1639.

Archak, N., Ghose, A., and Ipeirotis, P.G. 2011. "Deriving the Pricing Power of Product Features by Mining Consumer Reviews," *Management Science* (57:8), pp. 1485-1509.

Bakshy, E., Karrer, B., and Adamic, L.A. 2009. "Social Influence and the Diffusion of User-Created Content," *Proceedings of the 10th ACM conference on Electronic commerce*: ACM, pp. 325-334.

Bollinger, B. 2013. "Measuring Asymmetric Persistence and Interaction Effects of Media Exposures across Platforms." University of British Columbia.

Breiman, L. 1996. "Bagging Predictors," *Machine learning* (24:2), pp. 123-140.

Breiman, L. 2001. "Random Forests," *Machine learning* (45:1), pp. 5-32.

Bruce, N.I., Foutz, N.Z., and Kolsarici, C. 2012. "Dynamic Effectiveness of Advertising and Word of Mouth in Sequential Distribution of New Products," *Journal of Marketing Research* (49:4), pp. 469-486.

Chen, H., De, P., and Hu, Y.J. 2011. "It-Enabled Broadcasting in Social Media: An Empirical Study of Artists' Activities and Music Sales," *Workshop on Information Systems and Economics*.

Chevalier, J.A., and Mayzlin, D. 2006. "The Effect of Word of Mouth on Sales: Online Book Reviews," *Journal of marketing research* (43:3), pp. 345-354.

Dalessandro, B., Perlich, C., Stitelman, O., and Provost, F. 2012. "Causally Motivated Attribution for Online Advertising," *Proceedings of the Sixth International Workshop on Data Mining for Online Advertising and Internet Economy*: ACM, p. 7.

Davis, F.D., Bagozzi, R.P., and Warshaw, P.R. 1989. "User Acceptance of Computer Technology: A Comparison of Two Theoretical Models," *Management science* (35:8), pp. 982-1003.

Dellarocas, C. 2003. "The Digitization of Word of Mouth: Promise and Challenges of Online Feedback Mechanisms," *Management science* (49:10), pp. 1407-1424.

Dellarocas, C. 2006. "Strategic Manipulation of Internet Opinion Forums: Implications for Consumers and Firms," *Management Science* (52:10), pp. 1577-1593.

Domingos, P. 2005. "Mining Social Networks for Viral Marketing," *IEEE Intelligent Systems* (20:1), pp. 80-82.

Forman, C., Ghose, A., and Wiesenfeld, B. 2008. "Examining the Relationship between Reviews and Sales: The Role of Reviewer Identity Disclosure in Electronic Markets," *Information Systems Research* (19:3), pp. 291-313.

Freund, Y., and Schapire, R.E. 1996. "Experiments with a New Boosting Algorithm," *ICML*, pp. 148-156.

Gefen, D. 2000. "E-Commerce: The Role of Familiarity and Trust," *Omega* (28:6), pp. 725-737.

Gefen, D., Karahanna, E., and Straub, D.W. 2003. "Trust and Tam in Online Shopping: An Integrated Model," *MIS quarterly* (27:1), pp. 51-90.

Gefen, D., and Straub, D.W. 2000. "The Relative Importance of Perceived Ease of Use in Is Adoption: A Study of E-Commerce Adoption," *J. AIS* (1), pp. 0-.

Godes, D., and Mayzlin, D. 2004. "Using Online Conversations to Study Word-of-Mouth Communication," *Marketing Science* (23:4), pp. 545-560.

Godes, D., and Mayzlin, D. 2009. "Firm-Created Word-of-Mouth Communication: Evidence from a Field Test," *Marketing Science* (28:4), pp. 721-739.

Godes, D., Mayzlin, D., Chen, Y., Das, S., Dellarocas, C., Pfeiffer, B., Libai, B., Sen, S., Shi, M., and Verlegh, P. 2005. "The Firm's Management of Social Interactions," *Marketing Letters* (16:3-4), pp. 415-428.

Greene, W. 2003. "H.(2003): Econometric Analysis," *New Jersey, ua: Prentice Hall*).

Harris, E. 2014. "Data Breach Hurts Profit at Target," in: *The New York Times*. http://www.nytimes.com/2014/02/27/business/target-reports-on-fourth-quarter-earnings.html.

Heath, D., Singh, R., Ganesh, J., and Kroll-Smith, S. 2013. "Exploring Strategic Organizational Engagement in Social Media: A Revelatory Case,").

Hill, S., Provost, F., and Volinsky, C. 2006. "Network-Based Marketing: Identifying Likely Adopters Via Consumer Networks," *Statistical Science*), pp. 256-276.

Humphreys, L., Gill, P., and Krishnamurthy, B. 2010. "How Much Is Too Much? Privacy Issues on Twitter," *Conference of International Communication Association, Singapore*.

Johnson, J.W., and LeBreton, J.M. 2004. "History and Use of Relative Importance Indices in Organizational Research," *Organizational Research Methods* (7:3), pp. 238-257.

Kaplan, A.M., and Haenlein, M. 2010. "Users of the World, Unite! The Challenges and Opportunities of Social Media," *Business horizons* (53:1), pp. 59-68.

Laroche, M., Habibi, M.R., and Richard, M.-O. 2013. "To Be or Not to Be in Social Media: How Brand Loyalty Is Affected by Social Media?," *International Journal of Information Management* (33:1), pp. 76-82.

Lavinsky, D. 2013. "Is Traditional Marketing Still Alive?," in: *Forbes.com*.

McFadden, D. 1973. "Conditional Logit Analysis of Qualitative Choice Behavior,").

McKnight, D.H., Choudhury, V., and Kacmar, C. 2000. "Trust in E-Commerce Vendors: A Two-Stage Model," *Proceedings of the twenty first international conference on Information systems*: Association for Information Systems, pp. 532-536.

Nielsen. 2012. "State of the Media: Social Media Report (Nielsen Report)," http://www.nielsen.com.

Nielsen. 2013. "Global Trust in Advertising and Brand Messages (Nielsen Report),"), September 2013.

Owned it Ltd. 2013. "Leaving Money on the Table – How Order Confirmation Pages Are Used to Increase Sales."

Pavlou, P.A. 2003. "Consumer Acceptance of Electronic Commerce: Integrating Trust and Risk with the Technology Acceptance Model," *International journal of electronic commerce* (7:3), pp. 101-134.

Provost, F., and Fawcett, T. 2013. *Data Science for Business: What You Need to Know About Data Mining and Data-Analytic Thinking*. " O'Reilly Media, Inc.".

Reichheld, F.F., and Schefter, P. 2000. "E-Loyalty," *Harvard business review* (78:4), pp. 105-113.

Rhue, L., and Sundararajan, A. 2013. "Digital Visibility in Consumer Choices," *Available at SSRN*).

Richardson, M., and Domingos, P. 2002. "Mining Knowledge-Sharing Sites for Viral Marketing," *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining*: ACM, pp. 61-70.

Rishika, R., Kumar, A., Janakiraman, R., and Bezawada, R. 2013. "The Effect of Customers' Social Media Participation on Customer Visit Frequency and Profitability: An Empirical Investigation," *Information systems research* (24:1), pp. 108-127.

Rubel, S. 2006. "Trends to Watch Part Ii: Social Commerce," *Micro Persuasion vom* (23), p. 2005.

Shi, N., Hong, Y., Wang, K., and Pavlou, P. 2013. "Social Commerce Beyond Word of Mouth: Role of Social Distance and Social Norms in Online Referral Incentive Systems,").

Train, K. 2009. *Discrete Choice Methods with Simulation*. Cambridge university press.

Wang, C., and Zhang, P. 2012. "The Evolution of Social Commerce: The People, Management, Technology, and Information Dimensions," *Communications of the Association for Information Systems* (31).

Wergin, R., and Muller, R. 2012. "A Case Study in Marketing Communications: Traditional Vs. E-Media Advertising," *International Journal of the Academic Business World* (6:1).