

Does the Source Matter? How Referral Channels and Personal Communication Tools Affect Consumers' Referral Propensity

Antonia Köster
LMU Munich
koester@bwl.lmu.de

Christian Matt
University of Bern
matt@bwl.lmu.de

Thomas Hess
LMU Munich
thess@bwl.lmu.de

Abstract

Many companies are using social sharing buttons to make it easier for consumers to refer a website or app to other potential consumers. Although these buttons are ubiquitous online, it remains unclear whether consumer referral propensity (i.e. the likelihood of consumers referring other consumers) varies across the channels through which consumers arrive at the website. In particular, we test whether referral propensity is higher for consumers themselves acquired through social referrals and compare them with consumers accessing the website through other commonly used channels, such as search engines and online advertisements. In addition, we examine whether the communication tool (i.e. social networking websites or instant messaging clients) through which the referral is transmitted affects consumers' referral decisions. Our results indicate that consumers acquired through social referrals are more likely to make a referral and that the communication tools do not differ in their influence on consumers' referral propensity.

1. Introduction

Companies are increasingly relying on social referrals to generate awareness and acquire new customers for their offerings. As the effectiveness of traditional advertising decreases, companies are striving to leverage the power of interpersonal networks [9]. Several studies have shown that consumers attribute higher credibility to information received from other consumers than from advertisements [17, 43]. Many consumers use personal communication tools (PCTs), such as WhatsApp and Google+ for interpersonal communication [41]. Consequently, companies are integrating social sharing buttons in their website or app to connect them to these PCTs and thus facilitate consumers' ability to share online content [40]. Spotify and Dropbox are recent examples of companies that

have been successful in increasing their customer base through social referrals.

Consumers often do not access websites directly. They might be referred by other channels, such as search engines, other websites, or online advertisements. However, they might also be accessing the website because they received a referral from a social contact [4]. In our study, we are interested in understanding whether consumers acquired through social referrals are more likely to send a referral to others than are consumers who were acquired through organic (unpaid) search engine results or online advertisements. These referral channels represent different types: Social referrals comprise consumers acquired through interpersonal persuasion attempts, advertisements represent persuasion attempts directly from the company, and finally, referrals from search engine results contain consumers accessing the website through computational referrals. In viral marketing, consumers acquired through social referrals represent second-stage actors, whereas consumers acquired through online advertisements or search engines can be defined as first-stage actors [24].

Word of mouth (WOM) refers to the dissemination of information (e.g. opinions and recommendations) via informal face-to-face communication [1, 29]. WOM referrals, also known as social referrals, are usually unsolicited, that is, they are sent to recipients who are not actively seeking information [9]. Social referrals shared via PCTs have been researched in the electronic WOM (eWOM) literature [4, 24, 40]. Prior research has examined the impact of eWOM on firm-level outcomes, such as sales [e.g. 13, 18], and individual-level outcomes, such as consumer decision-making [e.g., 9, 40]. Furthermore, motives that lead to referral engagement [10], social referral incentive systems [40], and content characteristics [36] have been investigated.

Despite the growing prominence of social sharing buttons on websites or apps and the research call to examine consumers' referral decisions across different stages of dissemination [24], little attention has been paid to the impact of referring channels (i.e. social referrals, online advertisements, search engines) on

consumers' referral propensity. Our study seeks to fill this gap by examining the following research question: *Do consumers referred through social referrals have a higher referral propensity than consumers referred through organic search engine results or online advertisements?* In other words, will consumers acquired through social referrals from another person (i.e. second-stage actors) be more likely to click on social sharing buttons than would those coming directly from company's advertisements or from search queries (i.e. first-stage actors).

Furthermore, with the proliferation of platforms and apps for interpersonal communication, many different PCTs can be used to share referrals with friends and acquaintances. On the one hand, social networking sites (e.g. Facebook) have been widely employed to maintain and generate new relationships, and they have received considerable attention from researchers [31]. On the other hand, mobile instant messaging clients (e.g. WhatsApp) represent another type of PCT – one that has also piqued researchers' and marketers' interest [26, 38]. Although both PCTs are widely used to share referrals, it is unclear whether consumers referred by instant messaging clients have a higher referral propensity than consumers referred through social networking sites. To the best of our knowledge, there has been no relevant research that compares these two PCTs with each other in terms of how the source might influence consumers' sharing behavior. To fill this research gap, this study will also address a second research question: *Is there a difference in consumers' referral propensity acquired through different personal communication tools?* In particular, will there be a difference in the number of consumers who click on social sharing buttons on the website when they are referred through WhatsApp as opposed to Facebook?

In order to answer these research questions, we use a large data set comprising real-world online consumer behavior records. This data set was generated by one of the largest European media companies and includes information about consumer behavior on a traditional content website [35]. The website displays only proprietary, producer-generated entertainment content, which is a typical experience good [16]. Experience goods are difficult to assess prior to consumption [34]. Therefore, potential consumers are likely to resort to decision heuristics, making other consumers' information important and social referrals influential [16, 39]. In addition, entertaining content is likely to be forwarded to social contacts [36]. Hence, content websites offer a good setting to test the impact of different referral channels on consumers' referral propensity.

This study makes two important contributions to the existing literature by unveiling how referring channels

influence referral propensity. First, we compare the effectiveness of different referral channels and PCTs on consumers' referral propensity. Second, our data set incorporates consumers already acquired through social referrals, which allows us to compare the referral propensity across stages of dissemination, unlike extent research, which has focused on either first-stage [e.g., 9, 24] or second-stage actors [e.g., 4, 9].

This paper is organized as follows. In the next section, we review the prior literature on social referrals. In section three, we present our hypotheses concerning the effect of referral channels and PCTs on consumers' referral propensity. The subsequent sections describe our data analysis and the results of the hypothesis testing. Finally, in section six, we discuss implications, limitations, and directions for future research.

2. Conceptual foundations

2.1. Social referrals

Social referrals transmitted via PCTs have been researched in the broader context of eWOM [4, 24, 40], which also comprises research on seller feedback and consumer reviews on online platforms, such as forums and online communities [6, 20]. The peculiarity of social referrals in an online context is that they involve direct communication between individuals with social connections to each other [28, 40]. Social referrals are therefore more personal than seller feedback or online consumer reviews that are generally posted publicly [40]. The proliferation of PCTs not only makes it possible to share information with larger audiences without temporal or geographical constraints but also simplifies the process of sharing information [4]. These advantages attract the attention of companies interested in leveraging existing consumers' social networks to acquire potential consumers [40].

2.2. Drivers of social referral behavior

Research at the intersection of IS and marketing has dealt with the outcomes of referral behavior at the firm level and with the drivers of referral behavior at the individual level [10, 14]. Self-enhancement, extreme satisfaction, and customer commitment have been identified as important motivators for consumer referrals [10]. Moreover, consumers' perceptions of information value influence the consumers' propensity to engage in referrals [24, 37]. Other researchers have claimed that individuals are concerned with whether their actions will impair or enhance their image and whether this concern will affect their decision to make referrals [49]. In addition, social norms, tie strength, and

online social referral incentive systems have been proposed as influencing factors [40]. Social benefits have also been suggested as an important motivator for consumers to share referrals [20]. The most important reason why consumers engage in referrals may be social capital, defined as “the sum of the actual and potential resources embedded within, available through, and derived from the network of relationships possessed by an individual or social unit” [33, p. 243]. It governs relations among individuals, making the maintenance and creation of it important [7].

2.3. Comparison of referral channels

Prior research has focused on comparing eWOM with online or traditional, offline advertising. For example, Lu, Ba, Huang and Feng [29] examined the effect of promotional activities (i.e. online coupons and paid search engine results) and eWOM (i.e. consumer reviews) on restaurant revenue. However, the authors did not analyze social referrals and measured the effects only on an aggregate level. In the context of social networking sites, scholars have also analyzed the importance of social referrals in a comparison with offline advertising [43]. They found that referrals have a more long-term effect on customer acquisition than traditional advertising. However, they did not compare social referrals with online advertisements or search engine results and based their findings only on aggregated user data. Furthermore, studies have examined the joint effect of several customer acquisition channels. For example, Villanueva, Yoo and Hanssens [45] compared offline and online marketing activities with WOM and eWOM (e.g. organic search engine results and newspaper articles) consumer acquisition channels based on customers’ self-reporting. They showed that customers acquired through WOM and eWOM add nearly twice as much long-term value to the company than do the customers acquired through marketing activities. Prior research has also investigated the effects of eWOM (e.g. blogs and forums) and traditional advertising on companies’ stock market performance [48]. These findings are limited, because they do not shed light on individual-level effects.

It is surprising that, despite these valuable contributions to the literature, little attention has been paid to estimating the impact of social referrals compared with other referral channels, such as online advertisements. Trusov, Bucklin and Pauwels [43] acknowledge that an analysis of individual-level data may add novel insights regarding how referring channels might yield different revenue benefits to websites. Lu, Ba, Huang and Feng [29] recommend that further research should measure consumers’ response to different information cues in order to provide a deeper

understanding of individual marketing tools. Moreover, many studies have highlighted the impact of eWOM on firm-level outcomes for products or services [e.g. 13, 18], whereas entertainments goods have received less attention [16]. We intend to address this research gap by examining the effect of social referrals compared with referrals from organic search engine results or online advertisements on consumers’ propensity to make a social referral on a content website.

3. Hypotheses development

Consumers’ inclination to forward content reflects their tendency to share information with acquaintances, colleagues, family members, and friends. Assessing this inclination has gained importance in the online world [24]. Moreover, individuals are more likely to be influenced by those with whom they share common sets of ties, given that they tend to share common understandings and interests [2]. Prior research has also shown that consumers’ fear of being negatively perceived by their social contacts for forwarding messages deters them from doing so [36, 49]. The fact that a consumer has been referred to the website by another consumer may be interpreted as a signal of higher social demand [44]. Moreover, receiving referrals from social contacts rather than from a company is believed to enhance recipients’ trust in these messages [16, 45]. One could argue that consumers acquired through social referrals have greater knowledge about the sharing button functionality and therefore will have a higher referral propensity. However, this explanation does not hold, as it is not apparent for the receiver whether the referral sender is using a sharing button or, for example, inserting a copied URL link. Therefore, we propose that consumers acquired through social referrals will be more likely to forward this information to others as social acceptance may be more likely and the perceived value of the information increases over that of both other channels owing to higher demand perception [24]. Moreover, consumers arriving at the website through search engines or advertisements are primarily driven by self-interest and might be less likely to start communicating with their social contacts by sharing a referral.

***H1:** Consumers acquired through social referrals will have a higher referral propensity than consumers acquired through organic search engine results or online advertisements.*

A comprehensive understanding of a PCT requires a consideration of both “its functional characteristics and the set of general symbolic meanings users attach to its nature and purpose” [41, p. 3]. For example, e-mail

tends to convey a relatively high level of formality (i.e. impersonal contact) [42]. Although the two PCTs we consider in this study are comparable to each other in that both support one-to-one and one-to-many communication, they differ regarding the levels of personal intimacy and information overload [19]. Social networking sites, such as Facebook, continuously display new content from various sources (e.g. companies, friends, interest groups), and advertisements also contribute to the amount of available information. An average Facebook user has more than 150 friends in his or her network [12, 25]. These “friends” might also include acquaintances and even strangers. Instant messaging clients, such as WhatsApp, are designed to support short dyadic message exchanges [26]. These exchanges are linked to higher feelings of intimacy and deeper conversations than those via Facebook [38]. WhatsApp is more personal and in general allows for more private, intimate conversations because it is necessary to have someone’s mobile phone number in order to communicate with them via WhatsApp [3]. These PCT characteristics are somewhat “fixed,” as both PCTs are relatively mature, and these differences allow us to compare the two PCTs.

Primarily consumers pass along information to their social contacts to build and maintain social capital [7]. Thus, consumers’ referral propensity might also be influenced by the perceived value of the information. Prior research suggests that scarcity can increase information value [24]. We argue that, because of the smaller network size and increased intimacy, a WhatsApp user perceives referrals from their social contacts as more valuable than is the case for a Facebook user. Therefore, the WhatsApp user is more likely to share the content with his or her social network as well.

H2: Consumers in the second stage will have a higher referral propensity if referred via WhatsApp (instant messaging client) than consumers referred through Facebook (social networking site).

4. Data and measures

4.1. Research context

The data set was obtained from a well-known German media company that provides professional video-on-demand content, such as short clips on their website. Primarily, the website offers videos of general interest. The website has roughly 650,000 visits per day, of which approximately 530,000 are unique consumers. Our data set includes proprietary information, supplied by the media company, associated with all daily visits. Therefore, our data does not suffer from the recall

problems associated with self-reported data [22]. The company prominently displays social sharing buttons on each page with video content to make it easier for existing consumers (i.e. first-stage and second-stage actors) to share the website with others. The company did not explicitly incentivize social referrals with rewards. Therefore, we were able to examine unsolicited and unrewarded referral decisions.

The company tracks outbound social referrals from existing consumers (i.e. first-stage actors) to identify consumers arriving at the website through these social referrals (i.e. second-stage actors). This recording allowed us to investigate peer-to-peer communication, which is typically not made publicly available and consequently difficult to study [43]. Any time a visitor accesses a URL, the website provider records the details of the session. Among other session-related details, these data points include the referral channel from which the visitor arrives, how long the visitor remains on the website, the number of page views, the device category, a visitor identification number, and whether the consumer has clicked on the sharing buttons at the website (i.e. a social click). The data collected monitors observable, session-related behavior and does not record demographic or financial data, thus protecting a visitor’s privacy [4, 5]. Our data set does not allow us to control for visitor effects because we are unable to identify them comprehensively across our data set. To account for visitor heterogeneity, we rely on the controls for device and geographical area, following prior literature [4]. Furthermore, we analyze unique visitors and do not consider multiple visits in our analysis [5].

This study aims to analyze converted visitors, and we do not strive to understand the factors that drive consumers to respond, for example, to social media advertising or social referrals [31]. Therefore, our data set includes only converted visitors who started to view a video on the website. Moreover, we limit the data to visitors from one geographical area (i.e. Germany). We do this because there might be differences in the presentation of the website’s content based on a visitor’s location. The data used in this study spans eight weeks from February 20 to April 20, 2016 and only includes the referral channels: social referral, search engine, and online advertisements. Our data set comprises 958,044 visits from consumers with a unique identification number and 21 different videos. In general, each video has been accessed by visitors referred from all three channels and each video in our data set has been shared at least twice and started more than 1,000 times.

4.2. Model specification

We develop three general categories of variables: referral propensity measures (dependent variable),

referral channel measures (independent variables), and control measures. Our dependent variable, *Referral Propensity*, is a binary choice variable that is defined on the basis of the consumer's behavior on the website; it indicates whether a consumer *i* chose to share the content via social sharing buttons on the website [24]. *Referral Propensity* takes the value of 1 if the consumer clicked on a sharing button; otherwise, it takes the value of 0. In general, we observe that consumers click on the sharing buttons before and after they start to watch a video.

Our data set also captures information about consumers' traffic source. In this study, we compare social referrals with online advertisements and search engines. The variable *Social Referral* indicates whether the consumer was acquired through a social referral from an existing consumer who has used a social sharing button (e.g. WhatsApp, Twitter, and Facebook) on the website. The variable *Search Engine* indicates whether the consumer was referred by organic search engine results (e.g. Google and Yahoo), and the variable *Advertisement* indicates whether the consumer was acquired through a Facebook advertisement. It is important to note that the company promoted Facebook ads addressing the general Facebook user and advertised their Facebook postings during the entire data collection period. As reported in Table 2, the variable *Search Engine* has a mean of 0.05, which means 5% of all consumers in our data set were referred by search engines. The *Advertisement* variable has a mean of 0.95, indicating that 95% of the consumers were referred by Facebook ads. The remaining 0.1% of the consumers were referred through social referrals from Facebook, WhatsApp, and Twitter. The referral channels were implemented in the models using dummies. The variable *Advertisement* therefore equals 1 if the consumer was acquired through an online advertisement and 0 if not and *Search Engine* equals 1 if the consumer was acquired through a search engine and 0 if not.

In addition to the referring channels, several other factors could potentially influence consumers' referral propensity. Following previous literature, we control for several variables that may affect consumer behavior. We include a binary indicator, *Mobile* (mobile = 1, desktop = 0), to differentiate mobile device (including tablets) users from desktop users, presuming that a mobile user will be less likely to engage with the website due to the limitations of the smaller interface [15]. We created *Video Starts* to account for the number of videos viewed during the visit and control for *Page Views*, since a higher number of page views might indicate a greater interest in the website [11]. A page view is recorded every time a page is viewed. We use log transforms of the number of page views because the distribution is positively skewed [46]. For hypothesis 2,

we created *PCT*, a binary indicator for the PCT (instant messaging client: WhatsApp = 1, social networking site: Facebook = 0) that the consumer used to share the content.

Table 1. Description of variables

| Variable | Variable Description |
|---------------------|---|
| Referral Propensity | A binary indicator of whether the consumer used the social sharing buttons on the website during the visit. |
| Referral Channel | Acquisition type of the consumer with the categories <i>Advertisement</i> , <i>Social Referral</i> , and <i>Search Engine</i> . In the model, the variable was dummy coded with the reference category <i>Social Referral</i> . |
| Page Views | A positive integer value indicating the total number of pages viewed. |
| Mobile | A binary indicator of whether the website was accessed via a mobile device or desktop. |
| Duration | A positive integer value indicating the number of days between the date on which the video had been posted on the website and the date of the visit. |
| Video Starts | A positive integer value indicating the number of video starts during the visit. |
| PCT | A binary indicator of whether the consumer accessed the website via WhatsApp or Facebook. |

The descriptive statistics of the variables are presented in Table 2.

Table 2. Descriptive statistics

| Variable | Mean | Std. Dev. | Min | Max |
|---------------------|--------|-----------|------|--------|
| Referral Propensity | 0.002 | 0.042 | 0.00 | 1.00 |
| Advertisement | 0.951 | 0.216 | 0.00 | 1.00 |
| Search Engine | 0.048 | 0.214 | 0.00 | 1.00 |
| Social Referral | 0.001 | 0.032 | 0.00 | 1.00 |
| Page Views | 1.850 | 7.373 | 1.00 | 915.00 |
| Mobile | 0.935 | 0.246 | 0.00 | 1.00 |
| Duration | 11.318 | 8.176 | 1.00 | 69.00 |
| Video Starts | 1.002 | 0.047 | 1.00 | 5.00 |
| PCT | 0.348 | 0.476 | 0.00 | 1.00 |

Notes: N = 958,044; PCT (WhatsApp = 1) is based on social referrals, including only WhatsApp and Facebook, N = 975.

We tested for multicollinearity among the independent variables. The maximum variance inflation factor is below 2.00, which indicates that there is no multicollinearity. To test our hypotheses, we performed logistic regression and used the software Stata/IC 12.1 for the analysis. Logistic regression was chosen because it explicitly accounts for a dichotomous dependent variable [46]. To evaluate hypotheses 1 and 2, we estimated separate models. The specification of the model for testing hypothesis 1 is:

$$\begin{aligned} \text{Logit}(\text{Referral Propensity}_i) \\ = \alpha + \beta_1 \text{Advertisement}_i + \beta_2 \text{Search Engine}_i \\ + \beta_3 \text{Mobile}_i + \beta_4 \text{Page Views}_i + \beta_5 \text{Video Starts}_i + \varepsilon_i \end{aligned}$$

The effect of *Advertisement* in comparison with *Social Referral* is reflected in β_1 , and the effect of *Search Engine* in comparison with *Social Referral* is reflected in β_2 . For example, a positive coefficient β_1 will provide evidence that consumers acquired through advertisements have a higher propensity to make a referral via social sharing buttons than do the consumers already referred by social referrals.

To test hypothesis 2, we integrated an interaction term (*Social Referral* \times *PCT*) of social referrals and the *PCT*, which was used to send the referral. For hypothesis 2, the following model was tested:

$$\begin{aligned} \text{Logit}(\text{Referral Propensity}_i) \\ = \alpha + \beta_1 \text{Advertisement}_i + \beta_2 \text{Search Engine}_i \\ + \beta_3 \text{Mobile}_i + \beta_4 \text{Page Views}_i + \beta_5 \text{Video Starts}_i \\ + \beta_6 (\text{Social Referral}_i \times \text{PCT}_i) + \varepsilon_i \end{aligned}$$

5. Empirical analysis

5.1. Referral channels

First, we measure the impact of the referral channels on referral propensity. For each variable, the logistic regression estimates the effect of the variable on the referral propensity given that all other covariates remain constant. These results of model 1 are reported in Table 3, along with model fit statistics (pseudo R^2 , log likelihood). The likelihood ratio test assesses the overall fit of the model. The analysis of each model indicates a good model fit, with a highly significant likelihood ratio ($p = 0.000$). We also report Nagelkerke's pseudo R^2 value [32]. According to Hosmer, Lemeshow and Sturdivant [21] low pseudo R^2 statistics in logistic regression are the norm. In Table 3, covariate names are shown on the left, and the first column shows the odd ratios (i.e. exponential of the estimates) of the model. In the second column, we present the results with a smaller sample size (see section 5.2).

Table 3. Logistic regression results

| Independent Variable | Dependent Variable: Referral Propensity | |
|--|---|--------------------|
| | Model 1 | Model 3 |
| | exp(Estimate) | exp(Estimate) |
| Intercept | 0.0004** (0.000) | 0.028** (0.000) |
| Intercept (revised) | - | 0.0005 |
| Advertisement | 0.319* (0.011) | 0.283* (0.016) |
| Search Engine | 0.229** (0.002) | 0.158** (0.001) |
| Mobile | 8.174** (0.000) | 8.767** (0.000) |
| Page Views | 1.649** (0.000) | 1.872** (0.000) |
| Video Starts | 1.431 (0.426) | 1.226 (0.702) |
| Log Likelihood | -12372.762 | -5765.03 |
| Pseudo R^2 | 0.012 | 0.032 |
| Sample Size | 958,044 | 20,835 |
| Notes: * $p < 0.05$, ** $p < 0.01$; exact p-values are reported in parentheses; Model 1: total sample, Model 3: choice-based sample. | | |

As mentioned above, our analysis of the influence of referral channels is based on dummy coding with social referrals as the reference category. The results in the first column show that visitors referred by online advertisements have a negative and significant impact on the referral propensity likelihood. The odds ratio of the variable *Advertisement* equals 0.319 and is highly significant. This coefficient indicates that consumers landing on the website via Facebook advertisements decrease the odds of a referral by an average of 68% (0.319 - 1) compared with consumers accessing the website through social referrals. We find that referrals from organic search engine results have a negative and significant impact on referral propensity. The odds ratio of the variable *Search Engine* equals 0.229 and is highly significant. This coefficient indicates that consumers landing on the website via organic search engine results decrease the odds of a referral by an average of 77% (0.229 - 1) compared with consumers accessing the website through social referrals. As shown in Table 3, the number of video starts does not have a significant influence on a consumer's referral propensity, whereas the influence of the device and number of page views is significant.

5.2. Robustness checks

Since our goal is to estimate the causal impact of the referral channels on consumers' referral propensity, we perform a series of robustness tests. We analyze the first hypothesis again and consider referral propensity to be a rare event in our data.

First, we repeat the logistic regression analysis using a model that controls for the different videos in our data set. Although we considered only videos that were shared at least two times and visited at least 1,000 times during the two-month period, the video j might have an influence on a consumer's referral propensity. By adding a dummy for each of the 21 videos (*Video*) in our model, we absorb the effects particular to each video. In addition, we included *Duration* as another control variable, indicating the number of days between the date on which the video had been posted on the website and the date of the visit [4]. Since the *Duration* variable is measured in days and includes zero, we used duration increased by one, which allows us to use the logarithm. In order to test the model, we limit the data to the first video view. Therefore, the following model does not include the number of video starts, and our dependent variable is defined as a binary indicator of whether a consumer clicked on a sharing button before or after the first video was viewed.

$$\begin{aligned} \text{Logit}(\text{Referral Propensity}_i) &= \alpha + \beta_1 \text{Advertisement}_i + \beta_2 \text{Search Engine}_i \\ &+ \beta_3 \text{Mobile}_i + \beta_4 \text{Page Views}_i + \beta_5 \text{Duration}_i \\ &+ \beta_{6j} \sum_{j=1}^{20} \text{Video}_{ij} + \varepsilon_i \end{aligned}$$

The estimates are reported in the first column of Table 4.

Table 4. Logistic regression results: first video

| Independent Variable | Dependent Variable: Referral Propensity | |
|----------------------|---|--------------------|
| | Model 2 | Model 4 |
| | exp(Estimate) | exp(Estimate) |
| Intercept | 0.0002** (0.000) | 0.014** (0.658) |
| Intercept (revised) | - | 0.0003 |
| Advertisement | 0.361* (0.025) | 0.335* (0.038) |
| Search Engine | 0.360* (0.042) | 0.301* (0.039) |
| Mobile | 7.974** (0.000) | 8.156** (0.000) |
| Page Views | 1.842** (0.000) | 1.998** (0.000) |

| | | |
|--|--------------------|--------------------|
| Duration | 1.159** (0.000) | 1.156** (0.001) |
| Video* | (0.000) | (0.000) |
| Log Likelihood | -12290.911 | -5696.130 |
| Pseudo R ² | 0.017 | 0.046 |
| Sample Size | 958,044 | 20,835 |
| Notes: *p<0.05, **p<0.01; exact p-values are reported in parentheses; Model 1: total sample, Model 4: choice-based sample; *Wald test for the hypothesis that all video coefficients are zero. | | |

Model 2 controls additionally for possible effects owing to different videos and duration. The results of model 2 show, that the variables *Search Engine* and *Advertisement* are still significantly negatively related to consumers' referral propensity compared with the *Social Referral* variable. The variable *Duration* has a positive and significant influence. Our data shows that approximately 45% of sharing consumers refer the first video they consume on the website within one week after it was available on the website, and almost all sharing consumers (94%) have used the social sharing buttons to make a social referral within 20 days after the video had been published.

Second, we noticed that the percentage of consumers using sharing buttons is low in our data set (0.18%), which is typical in the market [8]. Since the majority of consumers do not trigger the share button (N = 956,336) our dependent variable *Referral Propensity* (referral versus no referral) is imbalanced (i.e. the occurrence of social referrals is rather rare). The biases that rare events create in estimating logit models have been discussed in the literature [23]. To overcome the problem of misclassification, one should re-estimate the model while deliberately under-sampling the non-sharing consumers so that a more balanced sample of ones and zeros in the dependent variable is obtained. This sampling technique is called choice-based sampling [23]. In order to verify the robustness of the estimates of the referral channels on the referral propensity rate, we apply choice-based sampling to correct this potential bias. Because this method does not yield consistent estimates of the intercept when traditional maximum likelihood methods are used, we adjust the estimated intercepts for each alternative [30]. Our smaller data set to test hypothesis 1 comprises 1,708 consumers who decided to click on social sharing buttons and 19,127 randomly sampled consumers (1% of the sample if *Referral Propensity* = 0) who did not click on the sharing buttons during their visit. The estimates are reported in the second column of Table 3 and Table 4. We find that the logistic regression with rare event correction (i.e. reduced sample size) produces estimates

very similar to those generated with the total sample in the first columns, which further improves our confidence in the finding that social referrals have a positive effect on referral propensity compared with the other two channels. Choice-based sampling increased pseudo R² and decreased log likelihood, which indicates that the model fit has improved. To summarize, having examined the impact of referral channels via different model specifications and data sizes, we find consistent evidence that social referrals have a significant positive influence on consumers' referral propensity.

5.3. Personal communication tools

The results for hypothesis 2 are presented in Table 5. The interaction term *Social Referral* × *PCT* shows the coefficient for social referral = 1 and PCT = 0.

Table 5. Logistic regression results: PCT

| Independent Variable | Dependent Variable: Referral Propensity | |
|--|---|--------------------|
| | Model 5 | Model 6 |
| | exp(Estimate) | exp(Estimate) |
| Intercept | 0.001** (0.000) | 0.001** (0.00) |
| Advertisement | 0.166** (0.000) | 0.184** (0.001) |
| Search Engine | 0.119** (0.000) | 0.183** (0.002) |
| Mobile | 8.124** (0.000) | 7.927** (0.000) |
| Page Views | 1.649** (0.000) | 1.841** (0.000) |
| Video Starts | 1.434 (0.423) | - |
| Duration | - | 1.161** (0.047) |
| Video* | - | (0.000) |
| Social Referral × PCT | 0.179 (0.125) | 0.168 (0.113) |
| Log Likelihood | -12371.223 | -12289.282 |
| Pseudo R ² | 0.012 | 0.017 |
| Sample Size | 958,039 | 958,039 |
| Notes: *p<0.05, **p<0.01; exact p-values are reported in parentheses; Model 5: base model, Model 6: includes video dummies and duration; *Wald test results for the hypothesis that all video coefficients are zero; Twitter has been removed (five observations). Testing both models with the choice-based sample size leads to similar results (the pseudo R ² of model 5 increases to 0.032 and the pseudo R ² of model 6 increases to 0.046). | | |

The coefficient of the interaction term is negative, as expected, but not significant. Consumers referred by Facebook friends are not significantly less likely to click on social sharing buttons than consumers acquired through WhatsApp contacts. This result holds when we control for videos and duration in model 6. Therefore, we have to reject hypothesis 2 and recommend that further research might analyze this model with a larger sample of consumers acquired through social referrals, as the direction of the influence is as expected, and the effect size is substantial in both models.

6. Discussion

This study made it possible to observe the actual referral decisions of first-stage actors (i.e. consumers acquired through advertisements or search engines) and second-stage actors (i.e. consumers acquired through social referrals). Therefore, we were able to compare consumers' referral propensity across different stages of dissemination. Referrals through first-stage actors are essential because the ability to reach second-stage actors is based on the first-stage actors' referral decisions. Although these referral decisions are important, a viral effect is only possible if second-stage actors also share the information with their social network. In general, our data set of a traditional content website shows that consumers arrive at the website less frequently via social referrals than through organic search engine results or Facebook advertisements. Although the overall potential of social referrals for customer acquisition seems limited, we provide evidence that consumers already acquired through social referrals are more likely to initiate referrals than consumers acquired through advertisements or search engines.

A consumer's decision to share the content on the website is an important measure for companies to investigate because such an act of sharing attracts potential consumers to the website. Therefore, content providers should not underestimate social referrals. For content providers, it is important to understand that consumers acquired through social referrals are more likely to start influencing their friends than if the consumer arrived at the website through other channels (i.e. online advertisements or search engines). This is especially interesting, as social referrals provide a rather cost-effective mechanism to attract consumers to the website [40]. Incorporating the analysis of actual referral behavior helps content providers assess the real value of referral channels. Although social referrals often originate from the referrer's motivation to share the information, social referral incentive systems can be designed to increase these unsolicited referrals. Content providers already embedding social sharing buttons

could facilitate referral engagement of consumers in the first and second stage by investing in a social referral incentive system or by promoting the usage of their sharing buttons [40]. For example, the content provider in our study might integrate statements such as “send to a friend” at the end of the video clip to increase referral likelihood across dissemination stages. Our results showed that there is no difference in consumers’ referral propensity acquired through WhatsApp or Facebook in the second stage. A more personal communication tool such as WhatsApp does not increase consumers’ referral propensity compared with Facebook. Content providers should therefore display both social sharing buttons on their website as the majority of consumers in our data set arrived via these two PCTs at the website.

Our results have also theoretical implications. This study extends previous research on eWOM, by showing that consumers referred through social referrals outperform consumers referred through search engines or online advertisements. Moreover, we examine consumer behavior at the individual level and not at the aggregate level [29]. Furthermore, prior studies primarily addressed the impact of eWOM on sales of tangible goods, such as books [e.g., 13, 18] and less research has focused on intangible goods, such as free online content [16]. We extend existing research by providing results for content-driven business models.

Although clickstream data represents an important source of behavioral insights, it limits our modeling effort in a number of ways. First, consumers may notice social sharing buttons but fail to click on them because of time pressure, preoccupation with the content, or the need to accomplish their navigation goals [5]. Second, the data we analyzed is based on last-click metrics, which ignores prior channel touches [27]. However, extant browsing behavior prior to the focal website visit is especially prevalent for high involvement products [22]. Third, although cookies recognize subsequent visits, we excluded repeated sessions as cookies do not allow for the possibility of exploiting similarities in click behavior across sessions for each consumer [5]. Further research could strengthen our findings by analyzing data collected at a website with registration. Fourth, future research should examine whether the same consumer clicked multiple times on the sharing buttons within a session. To describe the issue, zero-inflated binominal regression models can be specified [47]. Furthermore, as we have analyzed only three referral channels, further research might generate new insights by analyzing other referral channels as well. Moreover, the investigation of websites providing other types of content or even products could offer further insights [35]. In addition, while it is possible that our findings can be applied to websites that provide user-generated content, we have no data on such websites.

7. References

- [1] Arndt, J., *Word of Mouth Advertising: A Review of the Literature*, Advertising Research Foundation, New York, 1967.
- [2] Borgatti, S.P., and Everett, M.G., "Notions of Position in Social Network Analysis", *Sociological Methodology*, 22(1), 1992, pp. 1-35.
- [3] www.bristolpost.co.uk/surprise-survey-finds-average-number-contacts/story-28636810-detail/story.html#comments, accessed 25.08.2016, 2016.
- [4] Burtch, G., Ghose, A., and Wattal, S., "An Empirical Examination of Peer Referrals in Online Crowdfunding", *Proceedings of the 35th International Conference on Information Systems*, New Zealand, 2014, pp. 1-19.
- [5] Chatterjee, P., Hoffman, D.L., and Novak, T.P., "Modeling the Clickstream: Implications for Web-Based Advertising Efforts", *Marketing Science*, 22(4), 2003, pp. 520-541.
- [6] Cheung, C.M.K., and Thadani, D.R., "The Impact of Electronic Word-of-Mouth Communication: A Literature Analysis and Integrative Model", *Decision Support Systems*, 54(1), 2012, pp. 461-470.
- [7] Coleman, J.S., "Social Capital in the Creation of Human Capital", *American Journal of Sociology*, 94(1), 1988, pp. 95-120.
- [8] <http://blog.naytev.com/social-share-buttons-analysis/>, accessed 08.06.2016, 2016.
- [9] De Bruyn, A., and Lilien, G.L., "A Multi-Stage Model of Word-Of-Mouth Influence through Viral Marketing", *International Journal of Research in Marketing*, 25(3), 2008, pp. 151-163.
- [10] De Matos, C.A., and Rossi, C.A.V., "Word-Of-Mouth Communications in Marketing: A Meta-Analytic Review of the Antecedents and Moderators", *Journal of the Academy of Marketing Science*, 36(4), 2008, pp. 578-596.
- [11] Demers, E., and Lev, B., "A Rude Awakening: Internet Shakeout in 2000", *Review of Accounting Studies*, 6(2), 2001, pp. 331-359.
- [12] Ellison, N.B., Steinfield, C., and Lampe, C., "The Benefits of Facebook “Friends:” Social Capital and College Students’ Use of Online Social Network Sites", *Journal of Computer-Mediated Communication*, 12(4), 2007, pp. 1143-1168.
- [13] Forman, C., Ghose, A., and Wiesenfeld, B., "Examining The Relationship Between Reviews and Sales: The Role of Reviewer Identity Disclosure in Electronic Markets", *Information Systems Research*, 19(3), 2008, pp. 291-313.
- [14] Garnefeld, I., Eggert, A., Helm, S.V., and Tax, S.S., "Growing Existing Customers' Revenue Streams through Customer Referral Programs", *Journal of Marketing*, 77(4), 2013, pp. 17-32.
- [15] Ghose, A., Goldfarb, A., and Han, S.P., "How is the Mobile Internet Different? Search Costs and Local Activities", *Information Systems Research*, 24(3), 2012, pp. 613-631.
- [16] Godes, D., and Mayzlin, D., "Using Online Conversations to Study Word-of-Mouth Communication", *Marketing Science*, 23(4), 2004, pp. 545-560.
- [17] Goh, K.-Y., Heng, C.-S., and Lin, Z., "Social Media Brand Community and Consumer Behavior: Quantifying the Relative Impact of User- and Marketer-Generated Content", *Information Systems Research*, 24(1), 2013, pp. 88-107.

- [18] Gu, B., Park, J., and Konana, P., "Research Note: The Impact of External Word-of-Mouth Sources on Retailer Sales of High-Involvement Products", *Information Systems Research*, 23(1), 2012, pp. 182-196.
- [19] Guo, Z., Tan, F.B., and Cheung, K., "Students' Uses and Gratifications for Using Computer-Mediated Communication Media in Learning Contexts", *Communications of the Association for Information Systems*, 27(20), 2010, pp. 339-378.
- [20] Hennig-Thurau, T., Gwinner, K.P., Walsh, G., and Gremler, D.D., "Electronic Word-of-Mouth via Consumer-Opinion Platforms: What Motivates Consumers to Articulate Themselves on the Internet?", *Journal of Interactive Marketing*, 18(1), 2004, pp. 38-52.
- [21] Hosmer, D.W., Lemeshow, S., and Sturdivant, R.X., *Applied Logistic Regression*, John Wiley & Sons, New Jersey, 2013.
- [22] Huang, P., Lurie, N.H., and Mitra, S., "Searching for Experience on the Web: An Empirical Examination of Consumer Behavior for Search and Experience Goods", *Journal of Marketing*, 73(2), 2009, pp. 55-69.
- [23] King, G., and Zeng, L., "Logistic Regression in Rare Events Data", *Political Analysis*, 9(2), 2001, pp. 137-163.
- [24] Koch, O.F., and Benlian, A., "Promotional Tactics for Online Viral Marketing Campaigns: How Scarcity and Personalization Affect Seed Stage Referrals", *Journal of Interactive Marketing*, 32(1), 2015, pp. 37-52.
- [25] Köster, A., Matt, C., and Hess, T., "The Role of Communication Types on Referral Acceptance in Social Networks", *Proceedings of the 21st Americas Conference on Information Systems*, Puerto Rico, 2015, pp. 1-11.
- [26] Li, D., Chau, P.Y., and Lou, H., "Understanding Individual Adoption of Instant Messaging: An Empirical Investigation", *Journal of the Association for Information Systems*, 6(4), 2005, pp. 102-129.
- [27] Li, H., and Kannan, P., "Attributing Conversions in a Multichannel Online Marketing Environment: An Empirical Model and a Field Experiment", *Journal of Marketing Research*, 51(1), 2014, pp. 40-56.
- [28] Lis, B., and Neßler, C., "Electronic Word of Mouth", *Business & Information Systems Engineering*, 6(1), 2013, pp. 63-65.
- [29] Lu, X., Ba, S., Huang, L., and Feng, Y., "Promotional Marketing or Word-of-Mouth? Evidence from Online Restaurant Reviews", *Information Systems Research*, 24(3), 2013, pp. 596-612.
- [30] Manski, C.F., and Lerman, S.R., "The Estimation of Choice Probabilities from Choice Based Samples", *Econometrica*, 45(8), 1977, pp. 1977-1988.
- [31] Mao, E., and Zhang, J., "What Drives Consumers to Click on Social Media Ads? The Roles of Content, Media, and Individual Factors", *Proceedings of the 48th Hawaii International Conference on System Sciences*, Hawaii, 2015, pp. 3405-3413.
- [32] Nagelkerke, N., "A Note on a General Definition of the Coefficient of Determination", *Biometrika*, 78(3), 1991, pp. 691-692.
- [33] Nahapiet, J., and Ghoshal, S., "Social Capital, Intellectual Capital, and the Organizational Advantage", *Academy of Management Review*, 23(2), 1998, pp. 242-266.
- [34] Nelson, P., "Information and Consumer Behavior", *Journal of Political Economy*, 78(2), 1970, pp. 311-329.
- [35] Oestreicher-Singer, G., and Zalmanson, L., "Content or Community? A Digital Business Strategy for Content Providers in the Social Age", *MIS Quarterly*, 37(2), 2013, pp. 591-616.
- [36] Phelps, J.E., Lewis, R., Mobilio, L., Perry, D., and Raman, N., "Viral Marketing or Electronic Word-Of-Mouth Advertising: Examining Consumer Responses and Motivations to Pass Along Email", *Journal of Advertising Research*, 44(4), 2004, pp. 333-348.
- [37] Pihlström, M., and Brush, G.J., "Comparing the Perceived Value of Information and Entertainment Mobile Services", *Psychology and Marketing*, 25(8), 2008, pp. 732-755.
- [38] Quan-Haase, A., and Young, A.L., "Uses and Gratifications of Social Media: A Comparison of Facebook and Instant Messaging", *Bulletin of Science, Technology & Society*, 30(5), 2010, pp. 350-361.
- [39] Senecal, S., and Nantel, J., "The Influence of Online Product Recommendations on Consumers' Online Choices", *Journal of Retailing*, 80(2), 2004, pp. 159-169.
- [40] Shi, N., Hong, Y., Wang, K., and Pavlou, P., "Social Commerce Beyond Word of Mouth: Role of Social Distance and Social Norms in Online Referral Incentive Systems", *Proceedings of the 34th International Conference on Information Systems*, Italy, 2013, pp. 1-19.
- [41] Tan, C.-H., Sutanto, J., Phang, C.W., and Gasimov, A., "Using Personal Communication Technologies for Commercial Communications: A Cross-Country Investigation of Email and SMS", *Information Systems Research*, 25(2), 2014, pp. 307-327.
- [42] Trevino, L.K., Webster, J., and Stein, E.W., "Making Connections: Complementary Influences on Communication Media Choices, Attitudes, and Use", *Organization Science*, 11(2), 2000, pp. 163-182.
- [43] Trusov, M., Bucklin, R.E., and Pauwels, K., "Effects of Word-of-Mouth versus Traditional Marketing: Findings from an Internet Social Networking Site", *Journal of Marketing*, 73(5), 2009, pp. 90-102.
- [44] Van Herpen, E., Pieters, R., and Zeelenberg, M., "When Demand Accelerates Demand: Trailing the Bandwagon", *Journal of Consumer Psychology*, 19(3), 2009, pp. 302-312.
- [45] Villanueva, J., Yoo, S., and Hanssens, D.M., "The Impact of Marketing-induced versus Word-of-Mouth Customer Acquisition on Customer Equity Growth", *Journal of Marketing Research*, 45(1), 2008, pp. 48-59.
- [46] Wooldridge, J., *Introductory Econometrics: A Modern Approach*, South-Western, Mason, OH, 2015.
- [47] Yang, S., and Ghose, A., "Analyzing the Relationship Between Organic and Sponsored Search Advertising: Positive, Negative, or Zero Interdependence?", *Marketing Science*, 29(4), 2010, pp. 602-623.
- [48] Yu, Y., Duan, W., and Cao, Q., "The Impact of Social and Conventional Media on Firm Equity Value: A Sentiment Analysis Approach", *Decision Support Systems*, 55(4), 2013, pp. 919-926.
- [49] Zhang, Y., Feick, L., and Mittal, V., "How Males and Females Differ in their Likelihood of Transmitting Negative Word of Mouth", *Journal of Consumer Research*, 40(6), 2014, pp. 1097-1108.